Justus-Liebig-Universität Giessen

Institut für Theoretische Physik I

Fachbereich 07 - Mathematik und Informatik, Physik, Geographie

Master's Thesis

Instanton effects in the context of a generalised σ-model

Vorgelegt von:

Jan Fokken

August 2, 2011

Contents

I Introduction

I.1 Motivation

Over the last centuries scientists succeeded in describing the physical world with only four different foces: gravitation, electro-magnetism, the weak nuclear force and the strong nuclear force, known as Quantum Chromo Dynamics [QCD]. While the study of each of these forces yields interesting questions, this work will focus on some of the peculiar aspects of QCD.

The first property, which is much more prominent for QCD, than for the other fundamental forces, is the strong scale (or energy) dependence of its coupling strength. At high energies QCD becomes very weak (objects interact only weakly), resulting in the asymptotic freedom of elementary particles, while at low energies its coupling strength increases drastically, giving rise to the concept of 'colour confinement'. This confinement forces all low energy observables to come as 'colourless' states, such as hadrons and mesons. The curious energy dependence of QCD has a tremendous influence on theoretical treatment of this force in different energy regimes. At high energies, where the force is weak, the well established tools of perturbation theory can be used. With this advantage all related questions can be approached conceptually in a relatively simple setting. On the other hand, in the low energy region, the sharp rise of the involved strong coupling constant leads to a complete breakdown of the perturbative concepts. To arrive at (qualitatively and quantitatively) meaningful predictions for this realm, theoretical physicists have developed various concepts of effective field theories [EFT] over the last sixty years. While the conceptual progress concerning effective descriptions of fundamental theories has been very impressive in the preceding decades, there are still many open and interesting questions concerning various aspects of low energy (non-perturbative) QCD.

The second aspect that will be interesting in the context of this work, is not a unique feature of QCD, but is in principle inherent to many non-abelian theories. In the nineteen seventies it was found that non-abelian theories, like QCD, have a non-trivial vacuum configuration, $|\theta\rangle$. It turned out that it was possible to make a transition from one vacuum, $|\theta_1\rangle$, to another, $|\theta_2\rangle$, by means of certain field configurations, A_μ^{cl}. These fields were found by Belavin as explicit solutions of a pure Yang-Mills theory that minimised the Euclidean action. Later they became known as 'instanton' solutions. In the present context especially the unique interactions of instantons with fermions and scalar fields will be of interest. Ultimately these interactions allow to construct an instanton induced contribution, $\mathcal{V}_{\mathrm{inst}}$, to the effective scalar potential,

$\mathcal{V}$, in the context of a given, effective model.

This work now will investigate a possible generalisation of a very successful effective low energy model in QCD, the 'linear σ-model'. In 1960 Gell-Mann and Lévy developed this model to capture the low energy 'isospin' structure of QCD. They employed relevant low energy fields (scalar and pseudo-scalar mesons) as effective variables of their model, instead of the fundamental degrees of freedom of QCD (quarks and gluons). This effective description makes the linear σ-model much more tractable, than dealing with the full theory of QCD at low energies. Originally they used the (scalar, iso-scalar) σ-meson and (pseudo-scalar, iso-vector) $\boldsymbol{\pi}$-mesons as effective fields. These were assembled into a combined field, $\Phi = (\sigma, \boldsymbol{\pi}^{\mathrm{T}})^{\mathrm{T}}$, on which the effective Lagrange density depended, $\mathcal{L}^{\mathrm{eff}} = \mathcal{L}^{\mathrm{eff}}(\Phi)$. In this model the pions are characterised as Goldstone bosons of the effective model, while the scalar σ-meson becomes a massive, broad resonance state. Apart from these structural features another advantage of the σ-model is its large range of applicability. Essentially it can be used from the description of quarks up to the treatment of effective nucleon interactions.

The generalisation which will be pursued in this work, concerns the isospin structure of the σ-model. The effective Φ field incorporates a (scalar, iso-scalar) and a (pseudo-scalar, iso-vector) contribution. Going one step further one could generalise this structure to allow also for another combined field $\Lambda = (\widetilde{\eta}, \boldsymbol{\delta}^{\mathrm{T}})^{\mathrm{T}}$[1)], which incorporates the remaining two possibilities, namely a (pseudo-scalar, iso-scalar) and a (scalar, iso-vector) part. The model that will be developed in chapter III will depend on a combined field $\Omega = (\Phi^{\mathrm{T}}, \Lambda^{\mathrm{T}})^{\mathrm{T}}$, which then incorporates the generalisation of the isospin structure of the original σ-model.

Understandably this generalisation does not come for free. The original linear σ-model seperated the degrees of freedom into a 3-dimensional Goldstone mode and *one* massive field contribution. A simple generalisation, which just replaces Φ with the generalised Ω field would produce seven Goldstone modes and still only one massive field. As there is no experimental evidence for a physical realisation of such a model, another ingredient will be needed, if a generalisation of the σ-model in the indicated direction shall be developed. In the present work the induced scalar potential contribution from the instanton sector will be used to fill this conceptual gap. This approach is not entirely new, as it was first pursued by Saito and Shigemoto in 1979 Ref. Ref. [1]. In their paper the important mesonic mass relations, that will also be presented here have already been worked out in the context of a 'pure' instanton model[2)]. In later years it was found that these original 'pure' instantons had to be replaced by something named 'constrained' instantons in the context Saito's and Shigemoto's paper. Therefore this work will rederive the findings of Saito and Shigemoto in the 'constrained' instanton context and in addition further relations concerning fermions and gauge field couplings will be presented.

In the final model of chapter III the combination of the generalised σ-model with a Lagrange density

[1)] In the nomenclature of the PDG the $\boldsymbol{\delta}$-meson is usually called the a^0 resonance. The (pseudo-scalar, iso-scalar) part $\widetilde{\eta}$ cannot directly be identified with the physical η-meson, which will be part of the discussion in chapter III.

[2)] The exact differences of instantons or 'pure' instantons on the one hand and 'constrained' instantons on the other will be worked out explicitly in Sec. II.9.

$\mathcal{L}^{\text{eff}} = \mathcal{L}^{\text{eff}}(\Omega)$ with the instanton induced potential part, $\mathcal{V}_{\text{inst}}$, will lead to an effective model, which consists of a 3-dimensional pionic Goldstone mode and up to three different massive contributions: the σ-meson, the δ-meson and one remaining contribution of $\widetilde{\eta}$.
Naturally, before arriving at the complete, generalised Lagrange density of the linear σ-model at the end of chapter III, many aspects concerning instantons and the scalar field generalisation ($\Phi \to \Omega$) have to be worked out explicitly. As the treatment of all related subjects incorporates well known parts of modern physics as well as partly new areas, the next section will clarify the exact structure of this work, in order to offer an effective guideline to the interested reader.

I.2 Structure

As already indicated in the previous section this work will derive a possible generalisation of the 'linear σ-model' in the context of constrained instantons. A rigorous treatment of this generalisation has to rely on various concepts of modern physics. Of course all these theoretical concepts have been treated in great detail in many excellent textbooks and papers. Nevertheless, as the discussion in this work shall be as self-contained as possible, all subjects that are important to the derivation of the model in chapter III will be presented and briefly discussed beforehand in the theoretical background chapter II. It is inherent to such types of presentations, that the chapter concerning the background informations mixes parts that are original to the present work with those that have a more recitative character. Therefore this section will provide a brief outlook on the various subjects in chapter II.

- In general chapter II deals with three different aspects of modern theoretical physics. The first part, from Sec. II.1 to Sec. II.5, introduces concepts from group and representation theory that led to very successful descriptions of physical phenomena throughout the last hundred years. These sections are included for the reader's convenience as well as to clarify later notations. Thus they give a short summary of known textbook knowledge.
- The second part about effective field theories (Sec. II.6 and Sec. II.7) gives a brief summary of this large area of modern field theory. It also reports background knowledge that has only been collected here from sources that present the subject in great detail.
- Finally the last part of Sec. II.8 and Sec. II.9 introduces many ideas and results on the physics of instantons. As the model from chapter III relies heavily on the influence of instantons, this subject is presented in a more detailed fashion. While Sec. II.8 still completely relies on presentations of the subject given in literature, the situation is slightly different in Sec. II.9. Here only the first part of the section, which deals with explicit results for 'pure' instantons, is completely based on already known results. The second part, starting from Sec. II.9.8.1 mixes findings from other author's with calculations explicitly derived for the model of chapter III.

- After the loosely connected theoretical background part chapter III gives a well sorted presentation of the generalised linear σ-model. In the Sec. III.1 and Sec. III.2 the general structure of the model is discussed, including a list of preliminary assumptions, a presentation of the used effective fields and a discussion of the induced instanton potential, $\mathcal{V}_{\mathrm{inst}}$.
- The consecutive sections then derive the Lagrange densities for the included effective fields. The derivations will include a discussion of kinetic parts, mass terms, interactions and current contributions. In addition two possible schemes will be presented, that allow to identify the free parameters of the effective model with physical observables.

Therefore the reader, already familiar with the topics from chapter II might want to progress swiftly to chapter III. If one feels uneasy at some point with the used concepts or nomenclature, this should not lead to any problems, since chapter III will provide many back references to the corresponding sections in chapter II.
On the other hand, readers, unfamiliar with the concepts from chapter II are of course invited to read through the sections of this background chapter. Although most entries have a more recapitulating character, the sections always provide references to the literature, which was found to be most comprehensible to the author.
Finally, for notational conventions and abbreviations see App. A.1 and A.2.

II Theoretical background

This first chapter will set the theoretical basis for the effective model, which will be discussed in chapter III. Therefore topics in this part are chosen in order to present all important and needed concepts for the later model. Naturally this choice forestalls a closed introduction of the different subjects of discussion. Another obvious introductory remark concerns the completeness of the presented derivations in this chapter. While a full presentation and derivation of all basic concepts would surely be helpful, it would also transform this work to a full grown textbook. Instead of giving a full description, most aspects in this chapter are only briefly introduced and for further background information each section provides the important references to the literature.

II.1 Lorentz group

An important topic of this work is the Lorentz group (or symmetry) and the implications that its imposition yields for physical systems. Therefore this paragraph shall give a short introduction on the basic properties of this group. It will heavily rely on various concepts from group and representation theory. As the complete treatment of these fundamentals would blur the focus of this work, the corresponding introductory part is banned to the appendix App. B and a general treatment can found for example in Ref. [2] or [3].

Coming back to the Lorentz group its success story started with Einstein's discovery of special relativity. Since then all theories and models that do not restrict themselves to the regime of classical mechanics have to leave the speed of light c invariant. Mathematically this invariance can be expressed in terms of a group, which is named the Lorentz group. The elements of this group, $\{\Lambda^{\mu}{}_{\nu}\}$, have to leave the metric of space-time, $\eta^{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$, invariant, or in a formula:

$$(\eta^{\mu\nu})' = \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}\eta^{\alpha\beta} \tag{2.1.1}$$

This is equivalent to forcing the speed of light to be a constant. As mentioned in the introduction indices in Minkowski space are only contracted if they appear on different levels. A reason for this distinction can be found in the appendix App. A.5. The question is of course what the elements of this group are and how they can be characterised. In principle this information can completely be gained from Eq.

(2.1.1). A derivation of the various relations can be found in Ref. [4, p.55-62] or [5, p.15-19]. The most important results are summarised throughout this section.
It turns out that the Lorentz group is a continuous 6 parameter Lie group which consists of four disconnected parts: $\{\mathbb{1}, P, T, PT\}$. As it leaves the (semi-) scalar product[1)] of Minkowski space invariant, so to say the length of a 4-vector, it is the mathematical group $O^{(3,1)}$. Close to the identity its elements can be represented by

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \delta\omega^{\mu}{}_{\nu} \qquad\qquad \forall \quad \|\delta\omega\| < 1. \tag{2.1.2}$$

It can be shown that $\delta\omega_{\mu\nu} = -\delta\omega_{\nu\mu}$ is antisymmetric[2)]. Analysing the allowed structure of Lorentz transformations, $\Lambda^{\mu}{}_{\nu}$, gives a more concrete meaning to the four disconnected parts[3)]. First the determinant of all possible Lorentz transformations is $\det(\Lambda) = \pm 1$ (compare the group $O(n)$ in App. B.5.3). Apart from this another constraint turns out to be that the 00-component of Λ is either $\Lambda^{0}{}_{0} \geq 1$ or $\Lambda^{0}{}_{0} \leq -1$. These conditions lead exactly to the four disconnected parts. They are listed in Tab. 2.1.1. In the second column the $\rtimes$ symbol is the semidirect product of group theory[4)]. It allows to construct the full Lorentz group from the combination of the 'normal subgroup', $SO^{+}(3,1)$, and the 'discrete subgroup', $\{\mathbb{1}, P, T, PT\}$. Further details of the underlying mathematical aspects are not important here, so the interested reader is referred to Ref. [7, p.38-69].

Table 2.1.1: Constituent of the Lorentz group

symbol	mathematical structure	conditions	name
$\mathbb{1}$	$\mathbb{1} \rtimes SO^{+}(3,1)$	$\det(\Lambda) = +1$, $\Lambda^{0}{}_{0} \geq +1$	restricted Lorentz group
P	$P \rtimes SO^{+}(3,1)$	$\det(\Lambda) = -1$, $\Lambda^{0}{}_{0} \geq +1$	parity reversed part
T	$T \rtimes SO^{+}(3,1)$	$\det(\Lambda) = -1$, $\Lambda^{0}{}_{0} \leq -1$	time reversed part
PT	$(PT) \rtimes SO^{+}(3,1)$	$\det(\Lambda) = +1$, $\Lambda^{0}{}_{0} \leq -1$	parity-time reversed part

The decomposition from Tab. 2.1.1 into the disconnected parts allows to simply study the restricted Lorentz group and to examine the discrete transformations of parity and time reversal separately.
Compared to the full Lorentz group the $SO^{+}(3,1)$ subgroup is a connected six parameter Lie group and so, close to the identity, it can be represented via its generators, $M^{\mu\nu}$. These generators satisfy the

[1)] It is a semi-scalar product since $\|a\|^2 = a^{\mu}a_{\mu} = -a_0^2 + \mathbf{a}^2$ is not positive-definite in Minkowski space.
[2)] For this use Eq. (2.1.1) and the relation $(\Lambda^{-1})^{\mu}{}_{\nu} = \Lambda_{\nu}{}^{\mu}$
[3)] In this context disconnected means that all four parts of the Lorentz group are pairwise disconnected. Each part in itself is connected and this is why the parts are called the connected parts of the Lorentz group in mathematics.
[4)] $\rtimes$ should really be $\rtimes_{\varphi}$, where φ specifies how elements of $SO^{+}(3,1)$ are combined. It is only included here for completeness but for a correct definition of the semidirect product see Ref. [6, p.236-237].

following algebra:

$$[M^{\mu\nu}, M^{\rho\sigma}] = \mathrm{i}\hbar\left(\eta^{\mu\rho}M^{\nu\sigma} - \eta^{\nu\rho}M^{\mu\sigma}\right) - (\rho \leftrightarrow \sigma)\,. \tag{2.1.3}$$

While being true the algebra is not directly intuitive, but by using two identifications it can be split up into the well known generators of boosts and rotations. Take $J_i \equiv \frac{1}{2}\epsilon_{ijk}M^{jk}$ to be the generators of spatial rotations and $K_i \equiv M^{i0} = -M^{0i}$ to be the generators of 'Lorentz boosts' (for $\{i, j, k\} \in \{1, 2, 3\}$). These six new generators fulfil the commutation relations

$$[J_i, J_j] = \mathrm{i}\hbar\epsilon_{ijk}J_k, \tag{2.1.4}$$
$$[K_i, K_j] = -\mathrm{i}\hbar\epsilon_{ijk}J_k, \tag{2.1.5}$$
$$[J_i, K_j] = \mathrm{i}\hbar\epsilon_{ijk}K_k. \tag{2.1.6}$$

So, by using the K_i and the J_i generators, restricted Lorentz transformations can be built out of three spatial rotations and three boosts. From the generators (either K_i and J_i or $M^{\mu\nu}$) a unitary representation, $U_{\mathscr{L}}(\Lambda)$, of the Lorentz group can be built[5)], which realises the action of the Lorentz group. If one is interested in the action of the Lorentz group on another group S then an operator generating this action close to the identity can be written in the form

$$U_{\mathscr{L}}(\mathbb{1} + \delta\omega) = I + \frac{\mathrm{i}}{2\hbar}\delta\omega_{\mu\nu}M^{\mu\nu}. \tag{2.1.7}$$

Here $\mathbb{1}$ represents the unit element in Minkowski space, whereas I is the unit element in S space. While it is helpful for the physical interpretation of the Lorentz group to rewrite the original generators, $M^{\mu\nu}$, in terms of boosts and rotations, there is an even more elegant way to choose them. The following linear combination of J_j and K_j turns out to be useful:

$$N_i = \frac{1}{2}(J_i - \mathrm{i}K_i), \tag{2.1.8}$$
$$N_i^\dagger = \frac{1}{2}(J_i + \mathrm{i}K_i). \tag{2.1.9}$$

This set of generators has no easy physical interpretation but working out the algebra, by using Eq. (2.1.4)-(2.1.6), shows a crucial advantage:

$$[N_i, N_j] = \mathrm{i}\epsilon_{ijk}N_k, \tag{2.1.10}$$
$$[N_i^\dagger, N_j^\dagger] = \mathrm{i}\epsilon_{ijk}N_k^\dagger, \tag{2.1.11}$$
$$[N_i, N_j^\dagger] = 0. \tag{2.1.12}$$

[5)] In the language of App. B.5.2 this is the exponential map.

Comparing this algebra to the one for $SU(2)$ (Eq. (B.18)) shows that N_i and $N_i^\dagger$ are just the generators of two $SU(2)$ algebras and Eq. (2.1.12) implies that the algebras do not mix at all. So, this reformulation of generators allowes to identify the original, unintuitive, $M^{\mu\nu}$ algebra with either boosts and rotations or, more importantly, with two non interacting $SU(2)$ algebras. Mathematically speaking, the $SO^+(3,1)$ algebra is isomorphic to $SU(2) \otimes SU(2)$ and so this 'simpler' structure can be studied instead of the whole $SO^+(3,1)$ algebra. Any object, which transforms under $SO^+(3,1)$, can be written in terms of irreducible representations of two seperate $SU(2)$ groups.
In App. B.6.1 and B.6.2 it is shown that these representations are uniquely labelled by their highest weight, j, or equivalently by the dimension of representation, $\ell = 2j + 1$. How an object transforms under Lorentz transformations is therefore specified by two numbers, that correspond to the dimensions of the two $SU(2)$ representations. The most important objects are:

Table 2.1.2: Lorentz group indices of various sets of objects

(ℓ_1, ℓ_2)	name	total weight J_3
$(1,1)$	scalar $\widehat{=}$ singlet	0
$(2,1)$	left-handed spinor	1/2
$(1,2)$	right-handed spinor	1/2
$(2,2)$	vector	$\{0,1\}$

The last column simply gives all possible J_3 weights for the combined system ($J^3 = j_1^3 + j_2^3$, compare Eq. (B.21)). It is listed here since it is another simple connection to a well known physical quantity - the spin. With this identification scalars are spin 0 objects (bosons), spinors have spin $1/2$ and will be identified with fermions and vectors are simply identified with the $(2,2)$ objects as only these have the correct spin numbers (0 and 1). Some basic facts about spinors will be mentioned in the following section.
Taking it all together, a remarkable simplification in the study of the Lorentz group is possible. Instead of studying the whole group it is sufficient to analyse the action of the four discrete Lorentz transformations $\mathbb{1}$, P, T and PT separately and in addition to study the group $SU(2) \otimes SU(2)$ instead of $SO^+(3,1)$.

II.1.1 Left- and right-handed parts

In the last section it was shown, that the connected subgroup of the Lorentz group, $SO^+(3,1)$, falls apart into two independent $SU(2)$ subgroups. To distinguish these two representations, the first (which was called $(2,1)$ in Tab. 2.1.2) is typically named 'left-handed', or with a group symbol $SU_\mathrm{L}(2)$ and the second one (being $(1,2)$ in Tab. 2.1.2) 'right-handed', with the symbol $SU_\mathrm{R}(2)$. In this context, left- and right-handed are *only* names. From this point of view, it is not helpful to identify this handedness with some familiar connotation.

Left-handed and right-handed spinors live in different representations of $SU^{+}(3,1)$ and so they are labelled with different indices. Typically the indices of right-handed spinors get an extra 'dot', while left-handed spinors have an 'undotted' index. This dot is just part of the name, to label elements that transform under $SU_{\mathrm{L}}(2)$ *or* $SU_{\mathrm{R}}(2)$.
The action of a Lorentz transformations on a left-handed spinor shall be called $L_a^{\ b}$ throughout this work. Accordingly the change of right-handed components is done via a 'right-handed' transformation $R^{\dot{a}}_{\ \dot{b}}$[6)].
The contraction of indices in the two $SU(2)$ spaces works exactly as in Minkowski space. To get some acquaintance with this notation the transformation of a left-handed spinor, $\chi_a(x)$, and a right-handed spinor, $\xi^{\dagger\dot{a}}(x)$, is given here as

$$U_{\mathscr{L}}^{-1}(\Lambda)\,\chi_a(x)\,U_{\mathscr{L}}(\Lambda) = L_a^{\ b}\chi_b(\Lambda^{-1}x), \tag{2.1.13}$$

$$U_{\mathscr{L}}^{-1}(\Lambda)\,\xi^{\dagger\dot{a}}(x)\,U_{\mathscr{L}}(\Lambda) = R^{\dot{a}}_{\ \dot{b}}\xi^{\dagger\dot{b}}(\Lambda^{-1}x)^{7)}. \tag{2.1.14}$$

There is one other important fact about spinors, which can be seen from the definitions of the generators N_i and $N_i^\dagger$ (see Eq. (2.1.8)). Following the introduction of Sec. II.1, the N_i span the algebra of $SU_{\mathrm{L}}(2)$, while the $N_i^\dagger$ generate the algebra of $SU_{\mathrm{R}}(2)$. This means that hermitian conjugation transforms the generators N_i and $N_i^\dagger$ into each other and so the same has to be true for elements that transform according to their representations. That is $(\chi_a)^\dagger = \chi^\dagger_{\dot{a}}$ and vice versa. This is the reason why the right-handed field in Eq. (2.1.14) has been written with a dagger. As every left-handed field, ξ^a, can be turned into its right-handed partner, $\xi^{\dagger\dot{a}}$, one can express everything only in terms of left-handed fields and their hermitian conjugates.

II.1.2 Lorentz invariants and spinor notations

In Tab. 2.1.2 the most important objects and the dimensions of their $SU_{\mathrm{L}}(2)$ and $SU_{\mathrm{R}}(2)$ representations were listed. Using this table and the results on $SU(2)$ tensor products (App. B.6.3.1), one can derive various objects that are invariant under Lorentz transformations. In principle all objects that transform as the trivial representation are invariant (under group transformations) by definition. In this case, all objects that transform like a $(1,1)$ under $SU_{\mathrm{L}}(2)\otimes SU_{\mathrm{R}}(2)$ representations fall into this category. Therefore the first invariant object can be taken as an alternative definition of the scalar (see Tab. 2.1.2).
For the second invariant object a little work needs to be done. The tensor product of two Lorentz vectors gives:

$$(2,2)\otimes(2,2) = (3_s\oplus 1_a, 3_s\oplus 1_a) = (3,3)_s\oplus(1,3)_a\oplus(3,1)_a\oplus(1,1)_s \tag{2.1.15}$$

[6)]In the notation for spinor indices the same distinction between left- vs. right- and up- vs. down-indices is used as in the notation of space-time indices. In addition left- and right-handed indices are distinguished. This is not crucial throughout this work, but as it is the standard notation it is adopted here as well.
[7)]The transformation of the spinor field argument x is explained in App. A.3.

In the final result the subscript means (anti-) symmetric with respect to exchanging the elements of *both* $SU(2)$ representations. Ignoring all the higher dimensional terms the $(1,1)_s$ element shows that there is an invariant object which can be used to connect two vectors. This is again a well known object, the metric tensor, $\eta^{\mu\nu}$. It does not change under Lorentz transformations by construction (compare Eq. (2.1.1)).

So far it appears that this section only states well known relations in a fancy language, but of course there is a last one coming. The combination of a vector with a left- and a right-handed spinor yields:

$$(2,1)\otimes(1,2)\otimes(2,2) = (2,2)\otimes(2,2) = (3,3)_s \oplus (1,3)_a \oplus (3,1)_a \oplus (1,1)_s \tag{2.1.16}$$

So there is an invariant symbol combining a vector with two spinors. This means that a left- and a right-handed spinor can be combined to form an ordinary vector and there is an invariant symbol, which realises the 'translation' (mapping).

The invariant object indicated by Eq. (2.1.16) must be an object which carries a left-handed and a right-handed spinor index and, in addition, one vector index: $\sigma^\mu_{a\dot{a}}$. With this Lorentz invariant object a vector can be translated into spinors and vice versa: $A_{a\dot{a}} = \sigma^\mu_{a\dot{a}} A_\mu$. As an explicit representation $\sigma^\mu_{a\dot{a}}$ can be expressed in terms of (2×2) matrices in the following way:

$$\sigma^\mu_{a\dot{a}} = (I_{a\dot{a}}, (\boldsymbol{\sigma}_{a\dot{a}})^{\mathrm{T}})^{\mathrm{T}}, \tag{2.1.17}$$

$$\bar{\sigma}^{\dot{a}a\mu} = (I^{a\dot{a}}, (-\boldsymbol{\sigma}^{a\dot{a}})^{\mathrm{T}})^{\mathrm{T}} = \epsilon^{ab}\epsilon^{\dot{a}\dot{b}}\sigma^\mu_{b\dot{b}}. \tag{2.1.18}$$

Here $\boldsymbol{\sigma}$ are the Pauli matrices, a and $\dot{a}$ are their row and column indices and the transpositions only refers to the corresponding vectors. The derivation of this relation is a bit lengthy and so it can be reviewed in Ref. [5, p.209-217]. Eq. (2.1.18) is listed in the above equations, as it will show up in the free Lagrange density for spinors, but the derivation of both symbols is not crucial to this work. The correctness of the last equality can simply be checked numerically[8)] and gives the connection between both symbols.

The introduced invariant symbols can be used to reexpress combinations of fields in different representations (spinor vs. vector, or co- vs. contravariant vector). This will show up again in Sec. II.4. Apart from this translation into different representations the invariant symbols themselves stand in a direct relation to each other. This can be worked out explicitly using Eq. (2.1.17) and (2.1.18). The result is:

$$\frac{1}{2}\mathrm{tr}_{\mathrm{s}}\left(\bar{\sigma}^{\dot{a}a\mu}\sigma^\nu_{b\dot{b}}\right) = \frac{1}{2}\bar{\sigma}^{\dot{a}a\mu}\sigma^\nu_{a\dot{a}} = -\eta^{\mu\nu}\text{[9)]}. \tag{2.1.19}$$

The check if this equation is correct, is probably the quickest test to see, if the numerical choice of the σ-symbols is possible. As the righthand side consists only of the invariant symbol $\eta^{\mu\nu}$ the left hand side must be invariant as well. How this argument works exactly can be seen in Ref. [5, p.209-217].

8) The group theoretical argument is neat but not necessary at the moment and can be reviewed Ref. [5, p.209-217] as well.
9) tr_s means that the trace is taken to be over spinor indices only.

II.2 Quaternion groups

Originally quaternions were found by W.R. Hamilton as an extension to complex numbers. From the group theoretical point of view it turns out that the quaternion group forms a four element subgroup of $GL(2)$ (for a definition compare App. B.5.3). Historically this group is labelled as $H = \{\pm I, \pm i, j, k\}$ and the group multiplication law in the shortest way is given by $i^2 = j^2 = k^2 = ijk = -I$. While all properties of quaternions are characterised by this equation, there is a particularly useful representation of this group for the purpose of this work. This is a representation, which gives a complete basis of $\mathbb{C}^{2\times 2}$:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad i = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix} \quad , \quad j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad , \quad k = \begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}. \tag{2.2.1}$$

These elements form a 4-dimensional, irreducible basis of the quaternion group. As each element is an $SU(2)$ matrix, it acts on 2-dimensional vectors in a spin $j = 1/2$ representation. For later convenience and to distinguish them from the Lorentz spinors these spin $1/2$ objects will be called ‘Iso-spinors’ (which again is only a name!).

Objects, that are invariant under the group H, can be expressed in terms of this "vector-iso-spinor" basis (see App. B.1). For this define the quaternion symbols:

$$q^{\alpha}_{\mathrm{I}\,ab} := (I, i, j, k)^{\mathrm{T}} = (I_{ab}, (-\mathrm{i}\boldsymbol{\tau}_{ab})^{\mathrm{T}})^{\mathrm{T}}, \tag{2.2.2}$$

$$\bar{q}^{ab\,\alpha}_{\mathrm{I}} := (I, i, j, k)^{\star} = (I^{ab}, (\mathrm{i}\boldsymbol{\tau}^{ab})^{\mathrm{T}})^{\mathrm{T}} = \left(q^{\alpha\,ab}_{I}\right)^{\dagger} = \epsilon^{ac}\epsilon^{bd}q^{\mu}_{I\,cd}. \tag{2.2.3}$$

In the second line $\star$ stands for complex conjugation, $\boldsymbol{\tau}$ is a vector of Pauli matrices, all Latin subscripts are the $SU(2)$ matrix indices and the index α labels the four elements of the vector. The additional subscript I is a reminder that the object q^{α}_{I} acts in iso-spinor space. This additional distinction will be important, once instantons are examined in Sec. II.9, as there will be quaternion symbols in iso-spinor space ($\hat{=} q^{\alpha}_{\mathrm{I}}$) and in ‘euclideanised Minkowski space’ ($\hat{=} q^{\alpha}$) as well.

Both iso-spinor symbols ($q^{\alpha}_{\mathrm{I}\,ab}$ and $\bar{q}^{ab\,\alpha}_{\mathrm{I}}$) look very similar to the ones from the previous section ($\sigma^{\mu}_{a\dot{a}}$ and $\bar{\sigma}^{b\dot{b}\mu}$) and indeed they have corresponding properties. But notice that, unlike the $\sigma^{\mu}_{a\dot{a}}$ symbols, q^{α}_{ab} does not distinguish between two different $SU(2)$ representations (there are no dots). The reason for this is that, in contrast to the Lorentz group spinors, there is no additional label that introduces a distinction between different $SU(2)$ representations here. In the previous section $\sigma^{\mu}_{a\dot{a}}$ could be used to translate a combination of a left- and a right-handed spinor into an object in Minkowski space. To see the connection of the quaternion symbols one can calculate the trace over the iso-spinor components of $q^{\alpha}_{\mathrm{I}\,ab}$, in analogy to Eq. (2.1.19). Using the definitions of $q^{\alpha}_{\mathrm{I}\,ab}$ one finds the relation

$$\frac{1}{2}\mathrm{tr}_{\mathrm{I}}(\bar{q}^{cd\,\alpha}_{\mathrm{I}}\, q^{\beta}_{\mathrm{I}\,ab}) = \frac{1}{2}\bar{q}^{ab\,\alpha}_{\mathrm{I}}\, q^{\beta}_{\mathrm{I}\,ab} = \delta^{\alpha\beta}\text{[10)]}. \tag{2.2.4}$$

So, in contrast to the Lorentz group symbol, $q^{\alpha}_{\mathrm{I}\,ab}$ establishes a connection between two iso-spinors and an element in ordinary Euclidean space ($\alpha, \beta \in \{0, 1, 2, 3\}$). The analysis that the choice of $q^{\alpha}_{\mathrm{I}\,ab}$ is a correct invariant symbol can be done just as for Eq. (2.1.19).

From a group theoretical point of view, the existence of an invariant symbol, $q^{\alpha}_{\mathrm{I}\,ab}$, can be shown as well. This symbol transforms as $2 \otimes 2$ in iso-spinor space and the combination of four 2-dimensional $SU(2)$ representations yields a scalar object:

$$2^{\otimes 4} = 5_{\underline{s}} \oplus 3_a \oplus 3_{sa} \oplus 3_{as} \oplus 1_s \oplus 1_{\underline{s}} \tag{2.2.5}$$

The lines under the indices mark completely (anti-) symmetric elements. For the derivation of this equation check App. A.4 and for the general procedure see App. B.6.3.1. The existence of a singlet, $1_{\underline{s}}$, in the tensor product shows that there is a $2 \otimes 2$ object, which takes two other iso-spinors in a '2' representation into a scalar. This object is the already introduced $q^{\alpha}_{\mathrm{I}\,ab}$.

Using the invariant symbol, a scalar (or inner) product in quaternion space can be defined in analogy to the scalar product in 4 Euclidean dimensions. For $x, y \in \mathbb{R}^4$, the vectors can be translated to iso-spinor space as: $x_{ab} = x_\alpha q^{\alpha}_{\mathrm{I}\,ab}$ and $y^{ab} = y_\alpha \bar{q}^{ab\,\alpha}_{\mathrm{I}}$ and the following definition of the scalar product gives all the needed properties:

$$(y, x)_{\mathrm{H}} = \frac{1}{2}\mathrm{Tr}(y^{ab} x_{cd}) = y_\alpha x_\beta \frac{1}{2}\mathrm{Tr}(\bar{q}^{cd\,\alpha}_{\mathrm{I}}\, q^{\beta}_{\mathrm{I}\,ab}) = \sum_\alpha y_\alpha x_\alpha. \tag{2.2.6}$$

In conclusion, one can use quaternions to relate two spin $1/2$ objects to ordinary vectors in four dimensions. This connection to Euclidean space instead of Minkowski space is realised via the inclusion of an extra factor of i in the definition of the invariant symbols (compare Eq. (2.2.2)). Once the objects traditionally called iso-spinors will be included, this factor of i will play an important role (compare Sec. III.3.2).

A more thorough introduction to the connection between various groups and the quaternion group is given in the two following text books: Ref. [8, p.248-266] and Ref. [9, p.18-31].

II.3 Noether theorem

In Sec. II.1 and II.2 two symmetry groups have been introduced, which will be of utter importance in all derivations to come (in the general context of instantons as well as for the explicit construction of an effective model in chapter III). In a physical context these groups lead to certain continuous symmetries (for example rotational invariance of systems is implied by the Lorentz group). In 1918 Noether showed that continuous symmetries directly imply the conservation of an associated current and charge Ref.

[10)] tr_{I} is the trace over the matrix indices of the quaternion symbols. The I refers to isospin, as this concept will be related with these indices later.

[10]. As these currents will be discussed for the model of chapter III a short derivation of the Noether theorem shall be given here, which will be oriented along the lines of Ref. [5, p.132-139]. To approach the theorem take a Lagrange density, $\mathcal{L} = \mathcal{L}(A, \partial_\mu A)$, that depends on some field, A, and its derivative ($x \in \mathbb{M}^{(3,1)} \equiv$ Minkowski space). Take $A = A^\alpha(x)T^\alpha$ to transform under the action of a symmetry group, G, that leaves the action, $S = \int \mathrm{d}^4x\, \mathcal{L}[A(x), \partial_\mu A(x)]$, invariant. Here T^α are the generators of G (compare Sec. B.5.2). As the theorem deals with continuous symmetries it is sufficient to focus on infinitesimal transformations. All derived results can be 'lifted' to large scale transformations by repeated infinitesimal ones (compare App. B.2(ii)) in the end.
Suppose that Λ_ε characterises the group action on Minkowski space, $\mathbb{M}^{(3,1)}$, and that $U(\Lambda_\varepsilon)$ is an element of the symmetry group. The index ε denotes a set of infinitesimal parameters which is needed to specify the group transformations (for example rotations in a plane only need one parameter). The action of the group on the S space shall be characterised through the matrix $D^{\alpha\beta}(\Lambda_\varepsilon)$. The transformation is now given by

$$[A^\alpha(x)]' \equiv U^{-1}(\varepsilon)A^\alpha(x)U(\varepsilon) = D^{\alpha\beta}(\varepsilon)A^\beta(\Lambda_\varepsilon^{-1}x). \tag{2.3.1}$$

Some comments on general transformation rules can be found in App. A.3. For the theorem, only the change of $A^\alpha(x)$ under transformations will be needed: $\delta_\varepsilon A^\alpha(x) = [A^\alpha(x)]' - A^\alpha(x)$. With transformation Eq. (2.3.1) the change of the Lagrange density becomes

$$\delta_\varepsilon \mathcal{L}(x) = \mathcal{L}' - \mathcal{L} = \frac{\partial \mathcal{L}}{\partial A^\alpha(x)}\delta_\varepsilon A^\alpha(x) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu A^\alpha(x))}\delta_\varepsilon(\partial_\mu A^\alpha(x)). \tag{2.3.2}$$

The variation, δ_ε, is defined at the same point x as the derivative ∂_μ and so they commute in the second term. Together with the product rule this leads to:

$$\delta\mathcal{L}(x) = \frac{\partial \mathcal{L}}{\partial A^\alpha(x)}\delta_\varepsilon A^\alpha(x) - \partial_\mu\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu A^\alpha(x))}\right]\delta_\varepsilon A^\alpha(x) + \partial_\mu\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu A^\alpha(x))}\delta_\varepsilon A^\alpha(x)\right] \tag{2.3.3}$$

$$= \underbrace{\left(\frac{\partial \mathcal{L}}{\partial A^\alpha(x)} - \partial_\mu\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu A^\alpha(x))}\right]\right)}_{=\delta S/\delta_\varepsilon A^\alpha(x)}\delta_\varepsilon A^\alpha(x) + \partial_\mu\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu A^\alpha(x))}\delta_\varepsilon A^\alpha(x)\right]. \tag{2.3.4}$$

The indicated equivalence in the first term can be seen by a short calculation:

$$\frac{\delta S}{\delta_\varepsilon A^\alpha(x)} = \int_{\mathbb{M}} \mathrm{d}^4y \frac{\delta \mathcal{L}(y)}{\delta_\varepsilon A^\alpha(x)} \tag{2.3.5}$$

$$= \int_{\mathbb{M}} \mathrm{d}^4y \left(\frac{\partial \mathcal{L}(y)}{\partial A^\beta(y)} \frac{\delta_\varepsilon A^\beta(y)}{\delta_\varepsilon A^\alpha(x)} + \frac{\partial \mathcal{L}(y)}{\partial(\partial_\mu A^\beta(y))} \frac{\delta_\varepsilon(\partial_\mu A^\beta(y))}{\delta_\varepsilon A^\alpha(x)} \right) \tag{2.3.6}$$

$$= \int_{\mathbb{M}} \mathrm{d}^4y \left(\frac{\partial \mathcal{L}(y)}{\partial A^\beta(y)} \delta^{\alpha\beta} \delta^4(y-x) - \partial_\mu \left[\frac{\partial \mathcal{L}(y)}{\partial(\partial_\mu A^\beta(y))} \right] \delta^{\alpha\beta} \delta^4(y-x) \right) \tag{2.3.7}$$

$$= \frac{\partial \mathcal{L}(x)}{\partial A^\alpha(x)} - \partial_\mu \left[\frac{\partial \mathcal{L}(x)}{\partial(\partial_\mu A^\alpha(x))} \right]. \tag{2.3.8}$$

Line (2.3.7) is only correct if the field $A^\alpha(x)$ is sufficiently localised (in this case the boundary contribution from the partial integration does not contribute).
The second term of Eq. (2.3.4) is given the name of a current:

$$\sum_a \varepsilon_a j_a^\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu A^\alpha(x))} \delta_\varepsilon A^\alpha(x). \tag{2.3.9}$$

The sum on the left over a runs over all parameters ε. It is not obvious at the moment that the left-hand side can actually be written in such a sum. Sec. III.3.3 and III.4.1 will pick this up and answer the question for the important cases in this work. In fact it will be possible to drop ε_a completely from the equation, as a similar sum will appear on the right-hand side as well. Using Eq. (2.3.4) and (2.3.9) gives the Noether current:

$$\sum_a \varepsilon_a \left[\partial_\mu j_a^\mu(x) \right] = \delta_\varepsilon \mathcal{L}(x) - \frac{\delta S}{\delta_\varepsilon A^\alpha(x)} \delta_\varepsilon A^\alpha(x). \tag{2.3.10}$$

While this equation is not too revealing in the general case it becomes very interesting if one finds a symmetry transformation that leaves the Lagrange density unchanged. If, in addition, the A^α field satisfies the Euler-Lagrange equations (2.3.8), then the whole right-hand side of Eq. (2.3.10) vanishes leaving:

$$\partial_\mu j_a^\mu(x) = \frac{\partial}{\partial t} j_a^0(x) + \boldsymbol{\nabla} \boldsymbol{j}_a(x) = 0. \tag{2.3.11}$$

This has exactly the form of a continuity equation. So, by just assuming that the Lagrange density is invariant under a symmetry transformation of the group, G, one finds the existence of a continuity equation. (Of course, $A^\alpha(x)$ needs to fulfil the Euler-Lagrange equation as well.) If the current $j_a^\mu(x)$ is localised in space (that is it does not extend to infinity), then an integration over Minkowski space yields another neat result:

$$\frac{\partial}{\partial t}\underbrace{\int_{\mathbb{M}} \mathrm{d}^4x\, j_a^0(x)}_{=:Q_a} = \frac{\partial}{\partial t} Q_a = -\int_{\mathbb{M}} \mathrm{d}^4x \boldsymbol{\nabla}\mathbf{j}_a(x) = -\int_{\partial\mathbb{M}} \mathrm{d}^3x \mathbf{j}_a(x)\mathbf{n}(x) = 0. \tag{2.3.12}$$

For the third equality, Gauss theorem has been used. The last integral vanishes as it is performed on the boundary, where $\boldsymbol{j}_a(x)$ vanishes. Eq. (2.3.12) means that an underlying symmetry, which leaves the Lagrange density unchanged, implies the existence of a conserved quantity in time.
Even if this result comes a bit as a surprise it appears almost natural from a different perspective. Take a symmetry to 'live' in some sort of plane and a physical system that is invariant under the symmetry. The invariance implies that the physical system can only live within the symmetry plane, as well and, in addition, the physical system cannot change within the plane. Therefore any physical quantity that is related to symmetry transformations within the plane will keep its value for all time, as the system cannot leave the plane and as it is constant within it. The topic of conserved currents and charges will be picked up again for the explicit example in Sec. III.3.3 and III.4.1.

II.4 Important Lagrange densities

Using the ideas from Sec. II.1.1, one can construct all Lorentz invariant Lagrange densities by just building combinations of representations (a, b) (according to Tab. 2.1.2) that produce a singlet, $(1, 1)_s$, in the tensor product. The simplest case of a Lorentz invariant model is the Klein-Gordon Lagrange density:

$$\mathcal{L} = -(\partial_\mu\phi)^\dagger(\partial^\mu\phi) - m^2\phi^\dagger\phi. \tag{2.4.1}$$

With $\phi \in (1, 1)$ ($\hat{=}$ scalar), it is manifestly Lorentz invariant, as it contains only scalars and the invariant combination $\partial_\mu\partial^\mu$. Of course the next step is to build a Lagrange density for the next object in the Lorentz group - the spinor. But matters are significantly harder, as spinors themselves are transformed under Lorentz transformations. In addition there is no pure scalar in the decomposition into irreducible representations of the two spinor tensor products:

$$(1, 2) \otimes (1, 2) = (1, 1)_a \oplus (1, 3)_s, \tag{2.4.2}$$

$$(2, 1) \otimes (1, 2) = (2, 2). \tag{2.4.3}$$

So another way of including an object that is bilinear in spinor fields has to be developed (bilinears in the fields lead to linear equations of motion, which in turn can be used to describe free particles). These complications lead to a lengthy construction of spinor Lagrangians which mostly consists of the derivation of a Lorentz invariant, hermitian, bilinear kinetic term and one corresponding to a mass term. In full glory (following the group theoretical point of view) it is presented in Ref. [5, p.205-228] or [11, p.65-78]. There are several aspects in this derivation that will be used throughout this work but as a whole it is more a side note. One important aspect for physical systems is closely related to the constraints (Lorentz invariance, hermiticity and scalar behaviour) on a Lagrange density. They imply that any spinor Lagrangian incorporates a physical particle and its anti-particle as well. The simplest model, in which the physical particle, Ψ, and anti-particle, $\overline{\Psi}$, (defined in Eq. (2.4.7)) are not the same is built out of a combination of two left-handed spinors (χ_a and ξ_a) and their right-handed partners ($\chi^\dagger_{\dot{a}}$ and $\xi^\dagger_{\dot{a}}$). In addition to the transformation rules for spinors under the Lorentz group both, χ_a and ξ_a, are forced to transform in a particular way under an additional $U(1)$ symmetry:

$$U^{-1}(\alpha)\chi_a U(\alpha) = e^{-\mathrm{i}\alpha}\chi_a \qquad , \qquad U^{-1}(\alpha)\xi_a U(\alpha) = e^{+\mathrm{i}\alpha}\xi_a. \tag{2.4.4}$$

where $U(\alpha)$ is an element of the additional symmetry group, $U(1)$. The resulting Lagrange density for this example is discussed completely in Ref. [5, p.221-225] is:

$$\mathcal{L} = \mathrm{i}\chi^\dagger_{\dot{a}}\bar{\sigma}^{\dot{a}a\mu}\partial_\mu\chi_a + \mathrm{i}\xi^\dagger_{\dot{a}}\bar{\sigma}^{\dot{a}a\mu}\partial_\mu\xi_a - m(\xi^a\chi_a + \chi^\dagger_{\dot{a}}\xi^{\dagger\dot{a}}). \tag{2.4.5}$$

This is the Weyl formulation of the free Lagrange density of two left-handed 'Weyl spinors'. As derivations in this work mostly rely on the separation into left- and right-handed fields all spinor Lagrangians will be presented in this *Majorana-* or *Weyl-representation*. In this representation the Dirac γ-matrices take a form that makes this separation very easy. The Lagrange density Eq. (2.4.5) can be reexpressed in terms of the more familiar 'Dirac spinors' by introducing the four γ-matrices and the additional β-matrix (in Weyl-representation):

$$\gamma^\mu := \begin{pmatrix} 0 & \sigma^\mu_{a\dot{c}} \\ \bar{\sigma}^{\mu\dot{a}c} & 0 \end{pmatrix} \qquad , \qquad \beta := \begin{pmatrix} 0 & \delta^{\dot{a}}{}_{\dot{c}} \\ \delta_a{}^c & 0 \end{pmatrix}. \tag{2.4.6}$$

The particular form of the above matrices comes as a surprise, if one is used to the common 'Dirac'-representation. The very different structure of the γ-matrices is *only* related to the different choice of basis, which is used in the Weyl formulation and, again, a detailed introduction to the Weyl notation can be found for example in Ref. [5, p.205-228]. In the now introduced Weyl representation the combined spinors for physical particle and antiparticle are

$$\Psi \equiv (\chi_a, \xi^{\dagger\dot{a}})^{\mathrm{T}} \qquad \Psi^\dagger = (\chi^\dagger_{\dot{a}}, \xi^a) \qquad , \qquad \overline{\Psi} = \Psi^\dagger\beta = (\xi^a, \chi^\dagger_{\dot{a}}). \tag{2.4.7}$$

$\overline{\Psi}$ is called the Dirac conjugate of Ψ. From the definition of the Dirac spinors (Eq. (2.4.7)) one can see the advantage of this representation. The left- and right-handed components are separated in the Dirac spinors. Of course, there is a completely understandable derivation for these definitions, which is simply too long for the present purpose. Nevertheless, using the above matrices and spinors the Lagrange density of Eq. (2.4.5) can be rewritten to the standard form:

$$\mathcal{L} = \mathrm{i}\overline{\Psi}\gamma^{\mu}\partial_{\mu}\Psi - m\overline{\Psi}\Psi. \tag{2.4.8}$$

The $U(1)$ symmetry from the Weyl Lagrangian turns into the relation for the Dirac spinors:

$$U^{-1}(\alpha)\Psi U(\alpha) = e^{-\mathrm{i}\alpha}\Psi \qquad , \qquad U^{-1}(\alpha)\overline{\Psi}U(\alpha) = e^{+\mathrm{i}\alpha}\overline{\Psi}. \tag{2.4.9}$$

For the use in later sections, it is useful to introduce projection operators onto the left-handed component, P_{L}, and on the right-handed component, P_{R}, of Ψ. Eq. (2.4.7) define how a Dirac spinor can be constructed out of left- and right-handed Weyl spinors. From time to time it is useful to split Dirac spinors up into their left- and right-handed components. This can be done by introducing yet another matrix:

$$\gamma_5 \equiv \begin{pmatrix} -\delta_a^{\ c} & 0 \\ 0 & \delta^{\dot{a}}_{\ \dot{c}} \end{pmatrix}. \tag{2.4.10}$$

With this, the projection operator on the left-handed part becomes $P_{\mathrm{L}} := (\mathbb{1} - \gamma_5)/2$ and the right-handed projection is $P_{\mathrm{R}} := (\mathbb{1} + \gamma_5)/2$. Note that γ_5 anticommutes with γ^{μ} and with β! In particular this leads to the following relation for the Dirac conjugate of its left-handed component:

$$\overline{(P_{\mathrm{L}}\Psi)} = (P_{\mathrm{L}}\Psi)^{\dagger}\beta = \Psi^{\dagger}\frac{1}{2}(I - \gamma_5)\beta = \Psi^{\dagger}\beta\frac{1}{2}(I + \gamma_5) = \overline{\Psi}P_{\mathrm{R}}. \tag{2.4.11}$$

This is consistent with the change of a left-handed Weyl spinor into its right-handed partner under conjugation. By defining $\Psi_{\mathrm{L}} = P_{\mathrm{L}}\Psi$ and $\Psi_{\mathrm{R}} = P_{\mathrm{R}}\Psi$ the Lagrange density of Eq. (2.4.8) can be written in terms of the left- and right-handed components:

$$\mathcal{L} = \mathrm{i}(\overline{\Psi}_{\mathrm{L}}\gamma^{\mu}\partial_{\mu}\Psi_{\mathrm{L}} + \overline{\Psi}_{R}\gamma^{\mu}\partial_{\mu}\Psi_{\mathrm{R}}) - m(\underbrace{\overline{\Psi}_{\mathrm{R}}\Psi_{\mathrm{L}} + \overline{\Psi}_{\mathrm{L}}\Psi_{\mathrm{R}}}_{\overline{\Psi}_{\mathrm{R}}\Psi_{\mathrm{L}} + (\overline{\Psi}_{\mathrm{R}}\Psi_{\mathrm{L}})^{\dagger}}). \tag{2.4.12}$$

II.5 Chiral symmetry

In principle, the idea behind chiral symmetry is to take the additional $U(1)$ symmetry of spinor Lagrangians and enlarge it in such a way, that left- and right-handed spinors transform differently under it. That way, the single $U(1)$ symmetry is turned into two different ones $[U_{\mathrm{L}}(1)$ and $U_{\mathrm{R}}(1)]$ just like the two $SU(2)$ subgroups of the Lorentz group, and from the group theoretical point of view this enlarges $[SU_{\mathrm{L}}(2) \otimes SU_{\mathrm{R}}(2)]$ to $[U_{\mathrm{L}}(2) \otimes U_{\mathrm{R}}(2)]$. The implications of such an assumed additional symmetry are most conveniently analysed in the Weyl Lagrangian (Eq. (2.4.5)), or in the form of Lagrangian Eq. (2.4.12). Take $U_{\mathrm{ch}}(\alpha_{\ell,r})$ to be an element of $U_{\mathrm{L,R}}(1)$, so that the spinors transform under the corresponding $U_{\mathrm{L,R}}(1)$ symmetry as:

$$U_{\mathrm{ch}}^{-1}(\alpha_\ell)\psi_{\mathrm{L}}U_{\mathrm{ch}}(\alpha_\ell) = e^{-\mathrm{i}\alpha_\ell}\psi_{\mathrm{L}} \equiv L\Psi \qquad \hat{=} \qquad U_{\mathrm{ch}}^{-1}(\alpha_\ell)\chi_a U_{\mathrm{ch}}(\alpha_\ell) = e^{-\mathrm{i}\alpha_\ell}\chi_a, \tag{2.5.1}$$

$$U_{\mathrm{ch}}^{-1}(\alpha_r)\psi_{\mathrm{R}}U_{\mathrm{ch}}(\alpha_r) = e^{-\mathrm{i}\alpha_r}\psi_{\mathrm{R}} \equiv R\Psi \qquad \hat{=} \qquad U_{\mathrm{ch}}^{-1}(\alpha_r)\xi^{\dagger\dot{a}} U_{\mathrm{ch}}(\alpha_r) = e^{-\mathrm{i}\alpha_r}\xi^{\dagger\dot{a}}. \tag{2.5.2}$$

Here the linear operators $L = e^{-\mathrm{i}\alpha_\ell}P_{\mathrm{L}}$ and $R = e^{-\mathrm{i}\alpha_r}P_{\mathrm{R}}$ have been introduced just as an abbreviation for later convenience. The projection operators, P_{L} and P_{R}, correspond to the ones from previous section. From now on the chiral symmetry will be discussed using the Dirac Lagrangian Eq. (2.4.12). The transformations of the Weyl spinors are only listed to show how the formalism can be translated into the Weyl picture.

Applying the $U_{\mathrm{L,R}}(1)$ transformations Eq. (2.5.1) to the Lagrange density Eq. (2.4.12), one sees right away that the kinetic part is unchanged, as left- and right-handed fields do not mix there. But the situation is different for the mass term. Under chiral transformations it becomes

$$\begin{aligned} U_{\mathrm{ch}}^{-1}(\alpha_{\ell,r})\overline{\Psi}\Psi U_{\mathrm{ch}}(\alpha_{\ell,r}) &= \overline{\Psi}_R R^\dagger L\psi_{\mathrm{L}} + \overline{\Psi}_L L^\dagger R\psi_{\mathrm{R}} &\quad (2.5.3)\\ &= \overline{\Psi}_R e^{-\mathrm{i}(\alpha_\ell-\alpha_r)}\psi_{\mathrm{L}} + \overline{\Psi}_L e^{+\mathrm{i}(\alpha_\ell-\alpha_r)}\psi_{\mathrm{R}} &\quad (2.5.4)\\ &= \overline{\Psi} e^{-\mathrm{i}(\alpha_r-\alpha_\ell)\gamma_5}\Psi\,. &\quad (2.5.5)\end{aligned}$$

As chiral symmetry does not provide a condition that sets $\alpha_\ell = \alpha_r$, a simple mass term like in Eq. (2.4.12) violates a general chiral symmetry in such a model. The only possibility to include an invariant mass term into the model is to include another object, w, which neutralises the chiral transformations. This means that w has to transform under $U_{R,L}(1)$ as $U_{\mathrm{ch}}^{-1}(\alpha_{\ell,r})wU_{\mathrm{ch}}(\alpha_{\ell,r}) = LwR^\dagger$. With this object a combined term

$$U_{\mathrm{ch}}^{-1}(\alpha_{\ell,r})\Big[\overline{\Psi}_{\mathrm{L}}w\Psi_{\mathrm{R}} + h.c.\Big]U_{\mathrm{ch}}(\alpha_{\ell,r}) = (\overline{\Psi}_{\mathrm{L}}\underbrace{L^\dagger L}_{=\mathbb{1}}\, w\, \underbrace{R^\dagger R}_{=\mathbb{1}}\,\Psi_{\mathrm{R}} + h.c.) = (\overline{\Psi}_{\mathrm{L}}w\Psi_{\mathrm{R}} + h.c.)\,. \tag{2.5.6}$$

is invariant under $U_{\mathrm{L,R}}(1)$ transformations and therefore could qualify for a mass term in the model. A mass term has to be invariant under normal Lorentz transformations as well and so one still needs to specify how w does change under these. As w couples to a left- and to a right-handed Dirac spinor, it has to compensate their $SU_{\mathrm{L,R}}(2)$ transformations as well and so w goes into

$$U_{\mathscr{L},\mathrm{ch}}^{-1}wU_{\mathscr{L},\mathrm{ch}} = \underbrace{e^{-\mathrm{i}\alpha_\ell\tau}}_{\in SU_{\mathrm{L}}(2)}\, LwR^\dagger\, \underbrace{e^{\mathrm{i}\alpha_r\tau}}_{\in SU_{\mathrm{R}}(2)}\,. \tag{2.5.7}$$

under a chiral $U_{\mathrm{ch}}(\alpha_{\ell,r})$ *and* Lorentz $U_{\mathscr{L}}(\Lambda)$ transformation labelled by the unitary operator $U_{\mathscr{L},\mathrm{ch}} := U_{\mathscr{L},\mathrm{ch}}(\Lambda,\alpha_{\ell,r})$. Of course, in this situation the physical meaning of the newly introduced object w has to be identified. This topic will be picked up again in the construction of the model Lagrangian in Sec. III.3.2.

Typically, in chiral models the two different $U(1)$ transformations are not labelled by left and right, but by a 'vector' α_{V} and an 'axial' α_{A} component. The linear connection between these labels is:

$$\alpha_{\mathrm{V}} = \frac{1}{2}(\alpha_\ell + \alpha_r) \qquad , \qquad \alpha_{\mathrm{A}} = \frac{1}{2}(\alpha_r - \alpha_\ell)\,. \tag{2.5.8}$$

The reason for this type of labeling can be seen from Eq. (2.5.4). The vector transformation, $U_{\mathrm{V}}(1)$, contains all transformations that leave a simple mass term ($\overline{\Psi}_{\mathrm{R}}\Psi_{\mathrm{L}} + h.c.$) invariant and the axial transformations, $U_{\mathrm{A}}(1)$, correspond to the remaining possibilities.

II.6 Non-perturbative QCD

So far, the examples of Lagrange densities from the two previous sections could be applied to any kind of model, ranging from elementary particles up to the description of any kind of object that satisfies the symmetry demands. So Eq. (2.4.1) could be used to describe any scalar-like object and Eq. (2.4.12) suits to any object that transforms as a Dirac spinor.

As mentioned in the introductory part, this work will be concerned with particular questions in QCD. Due to the drastic energy dependence of the coupling constant, g_{QCD}, of strong interactions, this theory behaves very differently at different energy scales. At very high energies (well above the binding energy of nuclear matter) g_{QCD} becomes small and so the general concepts of perturbation theory are applicable in this regime. On the other hand, going to lower energies, g_{QCD} increases rapidly, making it first tedious and then impossible to use perturbative expansions of the theory. This particular behaviour poses huge problems on detailed investigations of the nature of QCD. Throughout the history of physics, most problems that people were interested in could be approached by identifying a certain 'minimum configuration' and then finding approximate solutions in terms of this minimum and small deviations from it. Thus, historically, many mathematical tools have been developed, which could be used in a perturbative context only and, in contrast, general techniques to deal with non-perturbative phenomena are still very rare. This uneven distribution is of course, also due to the fact that usually non-perturbative phenomena cannot be dealt with on a general footing, but have to be investigated in each particular aspect separately.

For questions concerning QCD physicists have come up with mainly two and a half different approaches to arrive at qualitative and/or quantitative results for the non-perturbative regime (corresponding to the low energy regime). In some sense, the 'brute force' way to low energy QCD is the so called lattice QCD. It is not an effective field theory [EFT] from the usual point of view and is briefly mentioned here for completeness. In lattice QCD the appearance of large coupling constants does not introduce any difficulties, as the direct numerical evaluation of scattering amplitudes and bound states does not rely on a particular numerical range for the coupling constants. In principle this evaluation can be achieved by discretising space-time to a lattice and then explicitly calculating the whole partition function. While this idea has the advantage to produce testable quantitative results, it has a significant drawback in the present context. Through the disrictisation of space-time one explicitly breaks the Lorentz invariance of the underlying theory and so all effects that are directly connected to the symmetry will be very hard to identify in any final result.

II.7 Effective field theories

The remaining one and a half approaches are, what is generally referred to as effective field theories. They have in common that they divide the theory of interest into different parts (energy regimes), that should be treated differently. The reason for this separation is quite intuitive: If there are effects (particles, resonances, etc.) which only occur above a given threshold energy, μ_0, then physics well below this energy should, loosely speaking, not depend on these phenomena, as the energy available is not sufficient to produce the effects. Thus one could replace the fundamental degrees of freedom, F, of the theory with effective ones, $F(\mu_0)$, that capture the most important implications of the theory up to the separation scale, μ_0. As indicated, the effective degrees of freedom loose their generality and become inherently scale dependent. The demands on such effective variables are that they have to simplify calculations (as their introduction is rather pointless otherwise) and that their scale dependence has to be unambiguous. How the scale dependence is specified turns out to be the main practical difference between the mentioned EFT approaches. It will be discussed after a rough sketch of the 'landscape' of QCD particles and resonances is drawn. Over the last sixty years this landscape has been studied quite well from the very low energy scales of hadron physics up to the realm of perturbative QCD. On the high energy side quarks and gluons have been found to be the relevant degrees of freedom. In going to lower energies, the growing coupling constant and the concept of confinement forces quarks and gluons to form compound, colourless objects (mesons and baryons). So, instead of describing the theory in terms of the fundamental quark and gluon fields (which cannot be seen at low energies anyway), the compound objects can be used as effective low energy degrees of freedom. While this idea is conceptually appealing, as it allows focussing on the relevant variables, there are problems concerning the mathematical approach. The experimentally observed spectrum of mesons, baryons and resonances thereof has been found to be spread out over all energy regimes. Therefore, it is hard to identify a suitable cutoff scale, that clearly seperates regimes of different phenomena in low energy QCD. In order to discuss the methods of EFT, suppose for now that one has identified a suitable cutoff energy, μ_0, for the questions one is interested in. In addition, assume that the energy range of interest lies in the non-perturbative part of QCD, as otherwise perturbation theory could simply be applied. The exact way how to simplify the original theory by the use of μ_0 is what seperates the remaining two approaches.

II.7.1 Wilson EFT

First take the EFT methods introduced by Wilson (Ref. [12]). They are well summarised in Ref. [5, 176-187] and here only some key ideas will be presented. Wilsons' idea was to use the additional energy scale to rewrite a given theory with the action, $S(\varphi)$, in terms of low energy fields, $\varphi_<(k) := \varphi|_{(|k|<\mu_0)}$, and the remaining high energy part, $\varphi_>(k) := \varphi|_{(|k|\geq\mu_0)}$, so that the partition function can be given as

$$Z = \int \mathscr{D}\varphi_< e^{-S^{\text{eff}}(\varphi,\mu_0)} \qquad , \qquad e^{-S^{\text{eff}}(\varphi,\mu_0)} = \int \mathscr{D}\varphi_> e^{-S(\varphi)}. \tag{2.7.1}$$

If all energies of interest are well below the cutoff ($E << \mu_0$), then all directly observable fields have to be from the low energy region. In a diagrammatic picture this means that all external lines have to be $\varphi_<$ fields, whereas any off-shell contribution in closed loops can come from a high energy field as well. Unfortunately the loop contributions cannot, in principle, be discarded here, since perturbation theory is not applicable at the energy scale, E, and therefore it is not guaranteed that the value of loop diagrams becomes small. This allows to construct new effective terms in the low energy Lagrange density with arbitrary numbers of external fields, $\varphi_<$, via loop diagrams[11)]. So by using the above separation into low and high energy fields one produces a low energy Lagrange density, whose parameters become cutoff dependent and in addition one produces an infinite series of new low energy terms:

$$\mathcal{L}^{\text{eff}} = \mathcal{L}(\varphi_<, \mu_0) + \sum_d c_d(\mu_0)\, O_d. \tag{2.7.2}$$

Here $\mathcal{L}$ indicates the Lagrange density from the original theory but augmented with the μ_0 dependence. This should be understood as all parameters becoming cutoff dependent [for example masses $m \to m(\mu_0)$ and coupling constants $\lambda \to \lambda(\mu_0)$]. The operators O_d represent terms with d external $\varphi_<$ field lines and $c_d(\mu_0)$ is the corresponding coefficient. So far the discussed consequences don't seem very appealing, since the original theory in a non-perturbative regime has been replaced by a scale dependent one, which, in addition, incorporates an infinite series of new contributions.

The power of this representation becomes visible, if only a little bit of physics is added in the description. After all, one knows that the low energy fields correspond to the relevant degrees of freedom at the energy scale $E << \mu_0$. Therefore, if one believes in the minimisation principle of the action, the physical observables that correspond to the relevant degrees of freedom must lie in a (local) minimum of the action. This observation now allows a reintroduction of perturbation theory through the back door. By simply taking the parameters of $\mathcal{L}^{\text{eff}}$: $\{m(\mu_0), \lambda(\mu_0), c_d(\mu_0), ...\}$, and tying them to physical observables, one can make a perturbative expansion of the partition function in terms of the effective parameters. Of course, in doing so all observables become scale dependent, which is somewhat bothersome at first glance. On the other hand, assuming that any theory is exact at *every* energy scale is a very tough demand.

[11)] In fact the only constraint on new terms is that they have to fulfil the symmetry requirements imposed by the original Lagrangian.

If observables are not strongly cutoff dependent, it might very well be, that every known physical theory is an effective low energy approximation to some underlying, more general concept that simply has not been found, yet. So, in general a cutoff dependent theory or model is really not the surprising case, but rather the contrary. In Ref. [5, 176-187], a general investigation of cutoff dependencies is presented along with many mathematical details and, as mentioned, here only the conceptual ideas have been presented.

II.7.2 'Matching' EFT

This section is mainly inspired by Georgi's paper Ref. [13] and the reader is referred to this paper for detailed derivations and further analyses. Compared to the introduced Wilson EFT, the procedure of 'matching' EFT is in some sense even more radical than the ideas that led Wilson to his effective theory. By allowing physical observables to become energy (or scale) dependent Wilson succeeded in rewriting QCD in terms of the relevant parameters only - in a given energy regime. In this sense he found a very suitable, scale dependent representation of QCD, which still incorporates every aspect of the full theory. By altering the cutoff scale, μ_0, Wilson EFT can be adapted to any energy range and in the limit of $\mu_0 \to \infty$ the effective description recovers the full theory.

While this is perfectly fine one could ask the question if it is actually necessary to keep every detail of the complete theory. In an effective low energy description one is essentially only interested in low energy phenomena and it is not a primary concern if the description can be changed smoothly to describe any other energy regime. Following this idea leads to a drastic concept: Instead of studying the theory of QCD, one could as well study any other model, which only captures the low energy behaviour of the theory. The expressions in the effective description don't even have to match the structure of QCD. Only observables calculated from the effective description have to find their correspondence in the low energy observables of the full theory. In this sense, if a given, complete theory, $\mathcal{L}_{\mathrm{H}}(\chi, \phi)$, consists of heavy particles, χ, with $m_\chi \approx \mu_0$ and light ones, ϕ, with $m_\phi << \mu_0$, then one can 'invent' a low energy effective model, $\mathcal{L}(\phi)$, which only incorporates the light particles. Of course, in doing so one makes a mistake, as all (off-shell) high energy contributions are ignored in this description. In coming closer to the cutoff, this mistake should become more apparent, since any high energy effect becomes "less off-shell". By introducing an additional contribution, $\delta\mathcal{L}$, to the effective Lagrangian this error can be remedied. $\delta\mathcal{L}$ has to be adjusted such that the low energy description of light particles coincides with the description from the full theory at the cutoff energy. Schematically this is represented in Fig. 2.7.1.

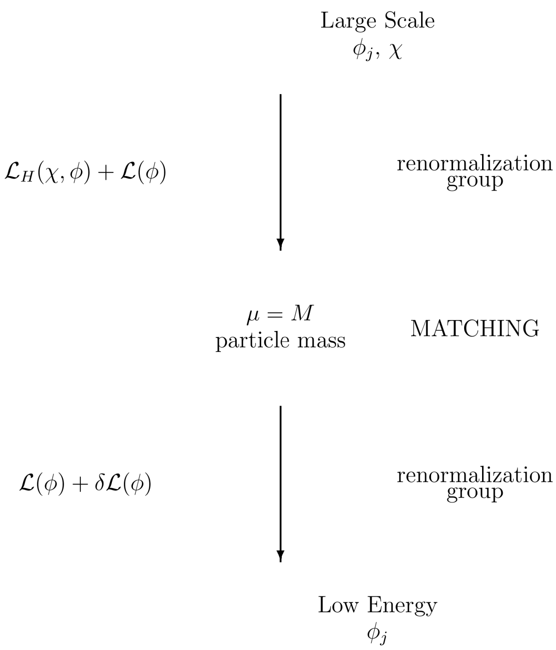


Figure 2.7.1: Matching procedure to consistently connect an effective description $\mathcal{L}^{\text{eff}}$ with a full theory $\mathcal{L}_{\text{H}}$. This figure is adopted from Ref. [13, p.21].

In essence $\delta\mathcal{L}$ corresponds to the operators O_d from the previous section. It incorporates all possibilities how high energy off-shell loops can lead to a d-point vertex in the low energy fields ($\hat{=}$ light particles). In his paper, Ref. [13, 20-27], Georgi derives the mathematical relations, that have only been captured in words here.

The renormalisation group indication in Fig. 2.7.1 highlights another necessary aspect for a rigorous treatment of matching EFT. It was already mentioned that the meson and hadron spectrum of QCD is wildly spread out over all energy ranges. Just like Wilson EFT, the matching procedure allows for a treatment of all energy regimes as well[12)], only the procedure to arrive at effective expressions is slightly different. Here one has to start with the full (known) theory at high energies. The renormalisation group can then be used to capture the energy dependence of observables as the energy scale, μ_0, is lowered. As soon as the scale arrives at the mass, M, of a physical observable (bound state or resonance) one

12) In Wilson EFT simply the cutoff, μ_0, was altered to describe different energy regimes.

changes the theory to an effective model, which does not incorporate the heavy observable anymore. The consistency of the effective model and the high energy description is guaranteed by the matching procedure at the energy $\mu_0 = M$. From there on the renormalisation group is used once again to keep track of energy dependencies of the remaining observables. As soon as the next heaviest mass is hit by the cutoff, μ_0, the procedure starts again and next effective description is used (now excluding the next heaviest particle). This goes on until one arrives at the lightest observables of the theory and thus one arrives at an effective description of QCD at every scale of interest.

II.7.2.1 Consequences

Having talked about different concepts of effective low energy descriptions and approximations to QCD the question is naturally what to make of it for the remaining parts of this work.

- The main focus of the upcoming theoretical background discussions lies on the concept of instantons and their implications. As this area of QCD is generically a non-perturbative effect within the theory, it can in principle be present at all energy scales. Therefore Sec. II.8 will give a general, non-perturbative introduction to instanton physics. Afterwards explicit results concerning instantons will be discussed in Sec. II.9. There, the descriptions will rely on expansion techniques around classical field configurations[13)]. This expansion *cannot* be seen as an application of perturbative QCD. The reason for this lies in the arguments from Sec. II.7. Since instantons are non-perturbative effects the configuration which minimises their action can lie at any energy scale, and thus as well within the range of non-perturbative QCD. In this sense, the expansions in Sec. II.9 can be understood as expansions of an effective model in the sense of the Wilson EFT.
- Later, the discussion of constrained instantons in Sec. II.9.8-II.9.10 will introduce additional constraints on the involved parameters, which ultimately lead to energy dependent instanton contributions. Thus, the constrained instanton effects become scale dependent and in principle they have to be treated in an EFT environment. Conceptually this is much easier to assess in the matching EFT picture. At a given energy scale, E, only those constrained instanton effects, that lie in the range of this scale have to be included and all contributions. While this is rather cryptic at the moment, the meaning will become clear at the end of Sec. II.9.10, when the exact energy dependencies of constrained instantons are given.
- Finally, in the model part of this work, yet another approach will be pursued. Historically, to gain a practical understanding of the possibilities of instanton physics in effective models, the simplest contribution from the instanton sector was included in conceptual studies. While the results from Sec. II.9.8-II.9.10 will suggest that this is an oversimplified approach, it will still be used in the model calculations, as further studies are very involved and thus have to be postponed. In this

[13)]Configurations that locally minimise the action.

sense the model calculations are a mere appetizer for future analyses.

II.8 Degenerate vacua and instantons

This section shall provide the quickest path to the idea of instantons in non abelian gauge theories. As always, when choosing the highway, there are countless sights along the road that one will miss for the sake of velocity. Not that this analogy is of any help in finding those sights, but the reader should be aware that the subject is much richer than the path presented in this section. More detailed introductions in the topic can be found in Ref. [5, p.576-609], [14, p.421-472], [15, p.277-289] and finally [16, p.340-344][14)] and of course, these references are the main sources for the following sections.
The key idea behind instantons is to re-examine known models in field theory and look for new, time independent solutions of the classical field equations. The reason to look for time independent configurations is that one already knows dynamical solutions, which can be gained via perturbation theory. The time independence now allows focusing on configurations that have their origin not in the dynamics of a system, but in its topology. Fortunately the topology of a system can be studied in the absence of any particle - that is to say, for a start, it is sufficient to focus on the vacuum configurations of a system.
The easiest model, one could start with, is that of a massless scalar particle, $\mathcal{L} = \partial_\mu\phi\partial^\mu\phi$. This is a rather unexciting model (lacking all dynamics) and so is its vacuum configuration, $\phi_{\text{vac}} = \text{const}$. The vacuum is unique up to a constant, which can be fixed by imposing a boundary condition at infinity: $\phi(\mathbf{x} = \infty) = 0$. Thus, it is not possible to relate something like a physical particle to the vacuum field configuration, as it is not localised and as its energy is $E_{\text{vac}} = \int_M \mathrm{d}^4x \left|\partial_\mu\phi_{\text{vac}}\right|^2 = 0$. Here M is the manifold in which the field, ϕ, lives, that is in the present case just 4-dimensional space-time.

Leaving this disappointing example, the situation becomes much richer if the scalar fields are exchanged for local gauge fields. The next natural step would be to talk about an abelian local gauge model (something like electro-magnetism), but as the non-abelian case will be of interest in later application this step will be skipped if favour of the less intuitive non-abelian version[15)].

II.8.1 Euclidean and Minkowski spaces

Before going into the indicated discussion about non-abelian local gauge models some remarks on conceptual difficulties are in order. Historically the concepts of degenerate vacuum configurations have been studied in Euclidean spaces and not in the physically interesting Minkowski space. In order to relate any insight from these studies to physical observables in Minkowskian space there are two logical approaches

[14)] While this is a ridiculously long list of citations, all sources have different approaches to the problem, which makes all of them worth reading.
[15)] There are interesting phenomena connected with the abelian case, such as magnetic monopoles, but these are not the topic in the present discussion.

possible:

- First one could derive a complete theory in Euclidean space and in the end translate it to Minkowski space via a 'Wick-rotation'. Any problem connected to this translation would of course have to be taken into account in this approach.
- The other possibility is to directly formulate the complete theory in Minkowski space, which circumvents the problem of translations between spaces. Unfortunately the very definition of path integrals is questionable in Minkowki space from a mathematical point of view. In addition - to the author's knowledge - the concepts that shall be studied in this chapter are not yet formulated from scratch in a Minkowskian topology.

The difficulty indicated above is actually not restricted to the area of degenerate vacuum configurations, but is a general problem of modern field theory. Usually the physicists approach is to take Minkowskian theories and translate space-time into the Euclidean pendant, which leads to well defined path integrals. This is unproblematic as long as scalar fields are treated. The situation becomes difficult as soon as spinor fields shall be described. In Sec. II.1 spinors have been introduced as objects that transform under one of the two $SU(2)$ subgroups of the Lorentz group, $SO^{+}(3,1) \simeq SU_{\mathrm{L}}(2) \otimes SU_{\mathrm{R}}(2)$. In the same section it was shown that left- and right-handed spinors could be related to each other via hermitian conjugation. By going from Minkowski to Euclidean space the symmetry group changes $SO^{+}(3,1) \to SO(4)$. This means that, instead of taking the relation between left- and right-handed spinors from Minkowki space one needs to use the corresponding relations from Euclidean space. The catch now is that, in Euclidean space, there are still two subgroups $SO(4) \simeq SU_{\mathrm{A}}(2) \otimes SU_{\mathrm{B}}(2)$, but they are *independent* of each other this time. Therefore the reformulation of Minkowskian theories really should include a reformulation of spinor relations.
This problem is usually dealt with by ignoring it, using the following argument: As one is not really interested in the 'in between' results in Euclidean space, one simply translates the physically meaningful Minkowski theory into Euclidean space, calculates everything and then - before interpreting it - one relates the results back to the physically useful Minkowski space theory. This procedure implicitly assumes that the change between Minkowski and Euclidean space is only a technicality to simplify calculations and that (for the physically interesting content) there is no difference between $SO^{+}(3,1)$ and $SO(4)$. While it is very well possible that this assumption is true, to the author's knowledge no-one so far managed to prove it.

While the formulation of the above problem might be seen as mathematical snootiness, it gains immediate importance, if concepts are investigated that can only be formulated in Euclidean or Minkowki space. Degenerate vacuum configurations, as will be discussed in this chapter, are connected to the topology of the underlying space and therefore it is not trivial to find equivalent concepts in spaces with different topologies, such as $SO^{+}(3,1)$ and $SO(4)$.
Now, having indicated these severe problems, the line of thought in this work will be that of faithful

ignorance: The concepts of degenerate vacua and all related subjects will be presented in Euclidean space under the assumption that the final translation back to Minkowki space is in fact possible and does not yield further problems.

II.8.2 Winding number in Yang-Mills models

For the next model - a local $SU(n)$ Yang-Mills gauge model - take t^a to be the generators of $SU(n)$ and $A_\mu(x_\mu) = t^a A^a_\mu(x_\mu)$ as the corresponding gauge field. As discussed in the previous section all coming derivations (unless explicitly indicated differently) in this chapter will be presented in Euclidean space: x_μ with $\mu \in \{1, 2, 3, 4\}$ and x_4 as the time component. The model, as before, shall only incorporate a kinetic contribution, so it is a pure Yang-Mills gauge model with the Euclidean action:

$$S_\mathrm{E}(A_\mu) = \frac{1}{2}\int_M \mathrm{d}^4 x_\mathrm{E}\, \mathrm{Tr}\left(A^{\mu\nu}A_{\mu\nu}\right). \tag{2.8.1}$$

Here, the trace runs over the generator indices of $SU(n)$, M again refers to the manifold in which the integrant lives and take $A_{\mu\nu} = A^a_{\mu\nu}t^a$ as the field tensor of the non-abelian gauge field[16]. The field strength tensor has the form:

$$\begin{aligned} A_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu - \mathrm{i}g[A_\mu, A_\nu] && (2.8.2)\\ &= \underbrace{\left(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu + g f^{abc} A^a_\mu A^b_\nu\right)}_{\equiv A^c_{\mu\nu}} t^c\,. && (2.8.3)\end{aligned}$$

For a derivation of this expression see App. A.6. For the $SU(n)$ gauge transformation,
$U(x_\mu) = e^{2\mathrm{i}g t^a \alpha^a(x_\mu)}$, a local gauge theory is invariant under the transformation:

$$t^a A^a_\mu(x_\mu) \to U(x_\mu)\, t^a A^a_\mu\, U^\dagger(x_\mu) + \frac{\mathrm{i}}{g} U(x_\mu)\partial_\mu U^\dagger(x_\mu)\,. \tag{2.8.4}$$

where the arrow gives the mapping of the field under the $SU(n)$ gauge transformation. Any field configuration can be gauged via this transformation and in particular the vacuum A^μ_vac can be gauged as well. As boundary condition for this model one can therefore not impose $A^\mu_\mathrm{vac}(x_\mu = \infty) = 0$, but only has the weaker 'pure gauge' condition holds:

$$A^\mu_\mathrm{vac}(\mathbf{x}, x_4) \xrightarrow[x_\mu\to\infty]{} \frac{\mathrm{i}}{g} U(\mathbf{x}, x_4)\partial_\mu U^\dagger(\mathbf{x}, x_4) = 2t^a \partial_\mu \alpha^a(\mathbf{x}, x_4)\,. \tag{2.8.5}$$

At this point it is convenient to assume that the gauge transformations approach a time independent limit for $x_\mu \to \infty$, which leads to a temporal gauge condition at infinity ($A^4(x_\mu = \infty) = 0$)[17]. This choice

[16] In mathematics this is often called the curvature form $F = \mathrm{d}A + A \wedge A$.
[17] This is equivalent to the earlier restriction on time independent solutions of the field equations.

leads to the temporal, pure gauge condition:

$$A^{\mu}_{\text{vac}}(\mathbf{x}, x_4) \xrightarrow[x_\mu \to \infty]{} \frac{\mathrm{i}}{g} U_{\pm}(\mathbf{x}) \partial_\mu U^{\dagger}_{\pm}(\mathbf{x}) = 2t^a \partial_\mu \alpha^a_{\pm}(\mathbf{x})\,. \tag{2.8.6}$$

where $U_+(\mathbf{x}) = U(\mathbf{x})|_{x_4=+\infty}$ and $U_-(\mathbf{x}) = U(\mathbf{x})|_{x_4=-\infty}$ and analogusly $\alpha_\pm$[18].

The question is now, if there is a unique vacuum state in this theory, so that all other vacuum states can be reached via gauge transformations from this state. In other words: can every $A^{\mu}_{\text{vac}}(U_1)$ be transformed smoothly into every other $A^{\mu}_{\text{vac}}(U_2)$? It turns out that this is not the case, but that there are different vacuum configurations and all of them are separated by finite energy barriers. To derive this, at first sight astonishing fact, some concepts from differential geometry are needed. First define the 'dual field strength tensor' as

$$\widetilde{A}^{\mu\nu} \equiv \frac{1}{2} \epsilon^{\mu\nu\sigma\tau} A_{\sigma\tau}\,. \tag{2.8.7}$$

with the Euclidean antisymmetric tensor, $\epsilon^{1234} = +1$. Note that the following equality holds: $\widetilde{A}^{\mu\nu}\widetilde{A}_{\mu\nu} = A^{\mu\nu}A_{\mu\nu}$ and with this one finds:

$$0 \leq \frac{1}{2} \mathrm{Tr}\left(\widetilde{A}_{\mu\nu} \pm A_{\mu\nu}\right)^2 = \mathrm{Tr}(A^{\mu\nu}A_{\mu\nu}) \pm \mathrm{Tr}(\widetilde{A}^{\mu\nu}A_{\mu\nu})\,. \tag{2.8.8}$$

The nonnegativity follows, since it is just the trace over a real square[19] and thus one can conclude:

$$\int_M \mathrm{d}^4x\, \mathrm{Tr}(A^{\mu\nu}A_{\mu\nu}) \geq \left| \int_M \mathrm{d}^4x\, \mathrm{Tr}(\widetilde{A}^{\mu\nu}A_{\mu\nu}) \right|\,. \tag{2.8.9}$$

Here, the integral over the manifold has just been added to make the connection to the Euclidean action and so combining Eq. (2.8.9) with (2.8.1) one gets a lower bound for the action. From this one already knows that, if the right-hand side of Eq. (2.8.9) is nonzero, then the vacuum action, corresponding to the vacuum energy, of the Yang-Mills model is nonzero, as well. For further insights on this lower bound the following identity is of help:

$$\mathrm{Tr}(\widetilde{A}^{\mu\nu}A_{\mu\nu}) = \partial_\mu \epsilon^{\mu\nu\sigma\tau} \mathrm{Tr}\left(A_\nu A_{\sigma\tau} + \frac{2}{3}\mathrm{i}g A_\nu A_\sigma A_\tau\right) \equiv \frac{1}{2}\partial_\mu J^{\mu}_{\text{CS}}\,. \tag{2.8.10}$$

Here, the Chern-Simons current, J^{μ}_{CS}, has been introduced. The derivation of the above identity (Eq. (2.8.10)) is possible by brute force 'index calculation' but it is more conveniently done in the framework of differential forms. This derivation can be reviewed in Ref. [15, p.218-235; 493]. A general, very 'applied' introduction on forms can be found in Ref. [17, p.1-11] and for a more advanced introduction Ref. [18] is helpful. Wether deriving Eq. (2.8.10) one way or another, it can be used to reexpress the

[18] Since the gauge condition only holds in the $x_\mu \to \pm\infty$ limit, U_+ does not have to equal U_-.

[19] To see that the square is real compare Eq. (2.8.3), using $f^{abc} \in \mathbb{R}$ and $A^a_\mu \in \mathbb{R}$. This is called the Bogomolny inequality.

integral over $\mathrm{Tr}(\widetilde{A}^{\mu\nu}A_{\mu\nu})$:

$$\int_M \mathrm{d}^4x\,\mathrm{Tr}(\widetilde{A}^{\mu\nu}A_{\mu\nu}) = \frac{1}{2}\int_M \mathrm{d}^4x\,\partial_\mu J^\mu_{\mathrm{CS}} = \frac{1}{2}\int_{\partial M}\mathrm{d}S^3_\mu J^\mu_{\mathrm{CS}} \tag{2.8.11}$$

$$= \int_{\partial M}\mathrm{d}S^3_\mu\,\epsilon^{\mu\nu\sigma\tau}\mathrm{Tr}\left(A_\nu A_{\sigma\tau} + \frac{2}{3}\mathrm{i}gA_\nu A_\sigma A_\tau\right). \tag{2.8.12}$$

In the first line the volume integral over the divergence of J^μ_{CS} was converted into a 3-dimensional surface integral (with the oriented measure $\mathrm{d}S^3_\mu$) by the means of Stokes theorem and the second line is just the definition of the Chern-Simons current. The clue is now that the integral is evaluated at the boundary, ∂M, of the manifold so that A_μ is in a pure gauge and therefore $A^a_{\mu\nu}(\mathbf{x} = \infty)$ vanishes (compare Eq. (2.8.6)) from Eq. (2.8.12). With this, the ‘winding number’ is defined as:

$$n_\pm := \frac{g^2}{2N\pi^2}\lim_{x_\infty\to\infty}\int_M \mathrm{d}^4x\,\mathrm{Tr}(\widetilde{A}^{\mu\nu}A_{\mu\nu})\Big(\delta^{(4)}(x_4 - t_{-\infty}) - \delta^{(4)}(x_4 + t_{+\infty})\Big) \tag{2.8.13}$$

$$= \frac{\mathrm{i}g^3}{3N\pi^2}\int_{\partial M}\mathrm{d}S^3_\mu\,\epsilon^{\mu\nu\sigma\tau}\mathrm{Tr}\,(A_\nu A_\sigma A_\tau)\Big|_{x_4=\pm\infty} \tag{2.8.14}$$

$$= \frac{1}{3N\pi^2}\int_{\partial M}\mathrm{d}S^3_\mu\,\epsilon^{\mu\nu\sigma\tau}\mathrm{Tr}\left[(U_\pm\partial_\nu U^\dagger_\pm)(U_\pm\partial_\sigma U^\dagger_\pm)(U_\pm\partial_\tau U^\dagger_\pm)\right]. \tag{2.8.15}$$

Note that the integration over x_4 in the first line is just a fake integration as the value of the Euclidean time is fixed. The prefactor $g^2/(N2\pi^2)$ is just a normalization constant, which depends on the particular choices of $A_\mu(\mathbf{x})$, the gauge transformation $U_\pm = U_\pm(\mathbf{x})$ and the dimension of the manifold. In the present case, g^2 is included as the self-interactions of the gauge fields are scaled by a coupling constant in the covariant derivative, $N = 2^{\dim(\partial M)}$ since a factor of 2 was included in the definition of $U(\mathbf{x})$[20] and the remaining $S_3 = 2\pi^2$ is the surface of the unit-sphere in 3 dimensions. The exact origin of this prefactor can be reviewed in Ref. [14, p.450-451].

Regarding Eq. (2.8.15), a lot of effort was put into rewriting the model, to get a lower bound on the action S_{E}, which looks not any better than the original action. Yet there is an important catch hidden in the winding number: it is a topological invariant! Therefore it cannot be changed by continuous gauge transformations, which in return means that any two vacuum configurations with different winding numbers correspond to inequivalent vacua.

The derivation of the invariance of Eq. (2.8.15) is a bit tricky. It is presented in a neat (but also short way) in Ref. [14, p.445-447].

[20]This factor will be dropped when explicit instanton contributions are calculated from Sec. II.9 onwards. It is included here, to see how normalisations change the result.

II.8.3 Instantons

Combining Eq. (2.8.15) with (2.8.9) and (2.8.1), the lower bound of the Euclidean action for a given winding number, n, at $x_4 = \pm\infty$ is given as $S_{\mathrm{E}} \geq N\pi^2 |n_\pm| / g^2$. If one is interested in the transition from an initial state at $x_4 = -\infty$ to a final state at $x_4 = +\infty$ then the bound is given by:

$$S_{\mathrm{E}} = \frac{1}{2}\int_M \mathrm{d}^4 x_{\mathrm{E}}\, \mathrm{Tr}\left(A^{\mu\nu} A_{\mu\nu}\right) \geq \frac{N\pi^2}{g^2} |n_+ - n_-| \,. \tag{2.8.16}$$

This follows from the derivation of Eq. (2.8.15). Solutions that minimise the transition action are those field configurations that saturate the lower limit of Eq. (2.8.9). They obey the relation:

$$\widetilde{A}_{\mu\nu} = \mathrm{sign}(n_+ - n_-) A_{\mu\nu}{}^{21)} \,. \tag{2.8.17}$$

If $n_+ - n_- = 1$ the solution is called the *instanton* and if $n_+ - n_- = -1$ it is called the *anti-instanton* (later n-instantons will refer to solutions of Eq. (2.8.17) with winding number n). If an explicit symmetry group is chosen then Eq. (2.8.17) can be used to derive the exact functional form of the instanton solution, $A^{\mathrm{cl}}(x_\mu)$. It turns out, as will be presented in Sec. II.9, that this solution is localised in space. Via Lorentz transformations it can then be boosted to any frame of reference, so that $A^{\mathrm{cl}}(x_\mu)$ has the exact properties of a particle in the Lagrange formalism.

II.8.4 θ-vacuum

While this settles the topic of the unique vacuum, it directly leads to the question, which vacuum should be included in the calculation of physically measurable quantities. As there is no apparent reason that excludes any configuration with a particular winding number, it is reasonable to assume that the physical vacuum is a superposition of all possible vacua, $|n\rangle$, with winding number, n, times a spectral function: $|vac\rangle = \sum_n f(n)|n\rangle$. To find the exact form of $f(n)$, it is useful to think about the expectation value of a localised physical observable $O(A)$, with A being a short-hand notation for all involved fields. If this observable is first measured in a large volume Ω_1 and then in an even larger volume $\Omega = \Omega_1 + \Omega_2$, where Ω_1 is just the same volume as before, then this change of the volume should not affect the outcome of the expectation value, if the volumina are much larger than the actual 'size' of the observable. Take n_j as the winding number in Ω_j. Then the winding number of Ω is simply $n = n_1 + n_2$ and with this the

[21] Inserting Eq. (2.8.17) into (2.8.16) gives an equality with fascinating mathematical properties. It connects an integral over smooth functions with the difference of two natural numbers and thus constrains the integral significantly. This equality, derived in a slightly different setting is an explicit example of the Atiyah-Singer index theorem. The derivation is shown in Ref. [14, p.362-370] and a general, mathematical introduction on the theorem is given in Ref. [19].

expectation value in the path integral formalism is:

$$\langle O\rangle_\Omega = \frac{\sum_{n_1,n_2} f(n_1+n_2)\left(\int \mathscr{D}A_{n_1}\, e^{-S_E(A,\Omega_1)}O(A)\right)\left(\int \mathscr{D}A_{n_2}\, e^{-S_E(A,\Omega_2)}\right)}{\sum_{n_1,n_2} f(n_1+n_2)\left(\int \mathscr{D}A_{n_1}\, e^{-S_E(A,\Omega_1)}\right)\left(\int \mathscr{D}A_{n_2}\, e^{-S_E(A,\Omega_2)}\right)}\,. \tag{2.8.18}$$

The index in the functional measure $\mathscr{D}A_n$ means that only those fields are included in the integral that result in the correct winding number, n. This is a rather long expression to point out a rather simple fact and it is possibly easier to visualise the situation, as in Fig 2.8.1 than to frame it in a formula.

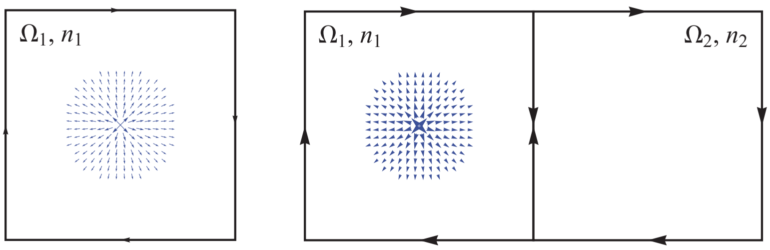


Figure 2.8.1: First, there is some physical configuration (blue vector field in volume Ω_1) shown. Then, the second figure shows the same event, only in the larger volume ($\Omega = \Omega_1 + \Omega_2$).

If $\langle O\rangle_\Omega = \langle O\rangle_{\Omega_1}$, as one needs on physical grounds, then the weight function has to be of the form $f(n_1+n_2) = f(n_1)f(n_2)$, in order to cancel out all dependencies on n_2 in Eq. (2.8.18). This means that $f(n) = e^{\mathrm{i}\theta n}$ for an arbitrary phase $\theta \in \mathbb{R}$ and so the θ-vacuum state becomes:

$$|\theta\rangle := \sum_n e^{-\mathrm{i}\theta n}|n\rangle\,. \tag{2.8.19}$$

Now the final step is at hand. So far the topological invariant, n, has been derived, the lowest possible energy difference for the change of n_- to n_+ was found and, finally, a consistent vacuum state with a new parameter, θ, for the Yang-Mills model was constructed. What remains to be done is the consistent inclusion of the new parameter into a minkowskian path integral formalism. For this suppose that a system starts out in a vacuum, $|\theta\rangle|_{x_4=-\infty}$, and ends in a state, $\langle\theta'|_{x_4=+\infty}$. Using $n \equiv n_+ - n_-$ the partition function (in the Euclidean path integral formalism) is:

$$Z_{\theta',\theta} = \langle\theta'| \int \mathscr{D}A\, e^{-S_{\mathrm{E}}}|\theta\rangle = \sum_{n_-,n_+} e^{\mathrm{i}(\theta' n_+ - \theta n_-)} \int \mathscr{D}A_{(n_+-n_-)}\langle n_+|e^{-S_{\mathrm{E}}}|n_-\rangle \tag{2.8.20}$$

$$= \sum_{n_-,n} e^{\mathrm{i}(n_-(\theta'-\theta)+n\theta')} \int \mathscr{D}A_n\, e^{-S_{\mathrm{E}}} \tag{2.8.21}$$

$$= \sum_{n} e^{\mathrm{i}n\theta'}\delta(\theta'-\theta) \int \mathscr{D}A_n\, e^{-S_{\mathrm{E}}}\,. \tag{2.8.22}$$

The δ-distribution is proportional to the n_- sum over the exponential[22] and the index in the functional measure again labels the fields that are included in the integral (for all other values of $n_+ - n_-$ the integrant vanishes). The δ-distribution in the last line shows that $Z_{\theta',\theta}$ is over-defined. This can be fixed by accepting that the partition function only depends on θ and defining:

$$Z_\theta \equiv \sum_n e^{\mathrm{i}n\theta} \int \mathscr{D}A_n\, e^{-S_{\mathrm{E}}} = \sum_n \int \mathscr{D}A_n\, \exp \int \mathrm{d}^4x_{\mathrm{E}}\, \mathrm{Tr}\left[-\frac{1}{2}F^{\mu\nu}F_{\mu\nu} + \mathrm{i}\frac{g^2\theta}{16\pi^2}\widetilde{F}^{\mu\nu}F_{\mu\nu}\right]. \tag{2.8.23}$$

For the second equality the definitions of n and $n_\pm$ were used. Now, all that is left to do is to reexpress everything in Minkowski space. This can be done by replacing $x_4 = \mathrm{i}x^0$ and $\epsilon^{1234} = -\epsilon^{1230}$, where the index 0 labels a variable in $\mathbb{M}^{(3,1)}$. Note that the term of the winding number contains an integration *and* a derivative with respect to x_4[23], so that it only catches a factor of -1 from the ε-tensor. As the other term does not have the additional derivative it gets an additional factor of i. The following partition function can then be given in Minkowski space:

$$Z_\theta = \int \mathscr{D}A\, \exp\mathrm{i} \int \mathrm{d}^4x\, \mathrm{Tr}\left[-\frac{1}{2}F^{\mu\nu}F_{\mu\nu} - \frac{g^2\theta}{16\pi^2}\widetilde{F}^{\mu\nu}F_{\mu\nu}\right]. \tag{2.8.24}$$

Here the sum over n is included in the functional measure, $\mathscr{D}A = \sum_n \mathscr{D}A_n$, as it is the only term that still depends on n. Having a closer look at the final result (Eq. (2.8.24)) one might get second thoughts about the inclusion of the new term into the Yang-Mills action. As discussed (Eq. (2.8.10)), this term is actually a 'total divergence' and usually these terms are dropped by the argument that field configurations are sufficiently localised so that they have no contribution from the boundary of space-time. But don't panic, the last pages were not just an afternoon entertainment, since, in this special case, the field configurations at the boundary do have a direct influence on the physics involved (compare Eq. (2.8.15)).

In fact, instantons have a neat connection to the idea of chiral field theory. The first is that instantons can solve the so called '$U(1)$ *problem*' by explicitly violating the axial chiral $U_A(1)$ symmetry. While this is an interesting topic, it is subject to the same fate as so many (in fact almost all) interesting things: It is not crucial for this work and thus has to be treated someplace else, for instance Ref. [14, p.243-246,450-

[22] As partition functions can be scaled by arbitrary constant, the replacement is correct in this context.
[23] To see this compare Eq. (2.8.15). n can be gained by dropping the constraints on x_4.

455]. But nevertheless in the model, introduced in chapter III, the explicit breaking of $U_A(1)$ will be used.
The second connection to chiral theories allows including a fermionic mass term in a chiral model, if instantons are included in the model as well. In Sec. II.5 it was discussed that such a term violates the chiral gauge invariance $U_{\mathrm{ch}}^{-1}\,|m|\,\overline{\Psi}e^{\mathrm{i}\phi}\Psi U_{\mathrm{ch}} = |m|\,\overline{\Psi}e^{-\mathrm{i}(\phi+2\alpha_A)\gamma_5}\Psi$. By analysing the change of a θ-term in the Lagrange density (see Eq. (2.8.24)) under chiral transformations, it turns out that $\theta \to \theta + 2\alpha_A$ changes by the same amount as the pure mass term. If one therefore takes the fermion mass to be $m_{\mathrm{F}} = |m|\,e^{\mathrm{i}(\phi-\theta)}$, then this term becomes gauge invariant. Unfortunately, there is a price to pay for this newly introduced mass term, which is known as the '*strong CP problem*'[24)] By studying effective QCD Lagrangians including the instanton effect, one finds that these models allow for CP violating terms. So far, QCD is known to preserve CP up to very high energy scales. This constraint ultimately leads to the following contidion on the vacuum angle: $|\theta| < 10^{-9}$. This is an extreme fine tuning problem and one would like to have a convenient explanation that sets $\theta = 0$. A solution to this problem can be the introduction of yet another field - the axion. Now, as this leads too far, here are two references, where the just mentioned study of chiral models and the implications are discussed in detail: Ref. [5, p.601-608] and [14, p.455-461].

II.9 Explicit results related to instantons

So far, the concept of instantons and the implications of their existence have been discussed on very general grounds to provide an insight into the richness of the topic. Unfortunately, for the later use, one has to get ones hands dirty and the general picture has to be filled with countless details. This section will provide some of the bare necessities and naturally the references to literature, where the subject is treated in full glory[25)].
First, important instanton results for a free Yang-Mills model will be presented, concerning solutions, degrees of freedom and direct changes of the path integral measure. After this, the model will be enhanced with an additional scalar Higgs field and the inevitable changes, related to the altered model, will be discussed.

II.9.1 Conventions for Euclidean space calculations

In the beginning it should be mentioned that all the calculations to come will be done for the anti-instanton ($n = -1$). There is no physical reason for this choice. It is simply the standard convention in literature to do explicit calculations for the anti-instanton and translate everything to the instanton later.

[24)] A discussion of the 'strong CP problem' from the time of the instantons' discovery can be found in Ref. [20].

[25)] This is just another alteration of yet so often repeated theme: There is not enough space (and time) to give a sufficiently complete introduction to every topic used in this work, and I apologise for the inconvenience related to any incompleteness.

These analogous calculations for the instanton are easy to reach by changing the gauge field solution from anti-instanton to the corresponding instanton. The solutions, and what changes explicitly is explained in Sec. II.9.2. The most important difference is that anti-instantons couple only to left-handed fermions and instantons only to right-handed ones. The contributions to measures and other spinor-independent quantities are the same for all (anti-) instantons, which will become apparent from the later discussions.

II.9.1.1 Gauge group conventions

For explicit calculations it is very convenient to rescale the gauge field, so that its coupling constant and the factor of $-\mathrm{i}$ is incorporated into the field (this means that coupling constants appear only in front of the kinetic term of the gauge field). In addition the generators of the involved $SU_{\mathrm{I}}(2)$ gauge group are taken to be

$$T_a = -\frac{\mathrm{i}}{2}\tau_a \qquad ; \qquad [T_a, T_b] = f_{abc}T_c \qquad \Rightarrow \qquad \mathrm{tr}_{\mathrm{I}}(T_a T_b) = -\frac{1}{2}\delta_{ab}{}^{26)}\,. \tag{2.9.1}$$

With these conventions the covariant derivative (including a gauge field, A_μ), acting on some field, X, becomes: $D_\mu X = \partial_\mu X + [A_\mu, X]$. Finally the field tensor of the gauge field is

$$A_{\mu\nu} = [D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]\,. \tag{2.9.2}$$

The rescaling of the gauge fields means that its Euclidean action incorporates a factor of the coupling constant: $S_{\mathrm{E}} = \frac{1}{g^2}\int \mathrm{d}^4 x_{\mathrm{E}} A_{\mu\nu}A^{\mu\nu}$.

II.9.1.2 Spinor conventions

In analogy to the minkowskian spinor matrices γ^μ, γ_5 and β from Sec. II.4 one can define the equivalent matrices for Euclidean space-time[27]. Instead of the invariant $\sigma^\mu_{\dot{a}b}$ symbols (Sec. II.1.2) these matrices are formed from the quaternion symbols, $q^\mu_{\dot{a}b}$ (Sec. II.2):

$$\gamma^\mu_{\mathrm{E}} := \begin{pmatrix} 0 & \mathrm{i}q^\mu_{a\dot{c}} \\ -\mathrm{i}\bar{q}^{\mu\dot{a}c} & 0 \end{pmatrix} \qquad , \qquad \beta_{\mathrm{E}} = \beta \qquad , \qquad \gamma_{\mathrm{E}\,5} = \gamma_5\,. \tag{2.9.3}$$

[26] $\mathrm{tr}_{\mathrm{I}}(\cdot)$ refers to a trace over the gauge group indices.

[27] In principle this change means that the governing symmetry group is not $SO(3,1)$ anymore but pure $SO(4)$. Therefore the underlying generators, $M^{\mu\nu}$, of Sec. II.1 are replaced by the Euclidean equivalence, $M^{\mu\nu}_{\mathrm{E}}$, and this in turn implies a change of the spinor matrices.

These are the most crucial identities for the Euclidean spinor algebra. For the explicit instanton solution the two following objects are needed as well:

$$q^{\mu\nu} := \frac{1}{2}\left(q^\mu \bar{q}^\nu - q^\nu \bar{q}^\mu\right) \qquad , \qquad \bar{q}^{\mu\nu} := \frac{1}{2}\left(\bar{q}^\mu q^\nu - \bar{q}^\nu q^\mu\right) . \tag{2.9.4}$$

These can be expressed in terms of the ‘t' Hooft-symbols’ [28]:

$$\bar{q}_{\mu\nu} = \mathrm{i}\eta_{a\mu\nu}\tau^a \qquad , \qquad q_{\mu\nu} = \mathrm{i}\bar{\eta}_{a\mu\nu}\tau^a . \tag{2.9.5}$$

Of course, there are many more details to be discussed for this algebra, but as they are merely a side note in the later topics, the interested reader is referred to Ref. [21, p.18-21;97-100] for further investigations. As indicated in Sec. II.8.1, Euclidean space consists of two linearly independent subgroups [$SO(4) \simeq SU_\mathrm{B}(2) \otimes SU_\mathrm{A}(2)$] in contrast to Minkowski space [$SO^+(3,1) \simeq SU_\mathrm{L}(2) \otimes SU_\mathrm{R}(2)$], where they are related to each other via hermitian conjugation. Therefore the left- and right-handed spinors (ψ_L, ψ_R) turn into *independent* spinors [$\psi_\mathrm{L} \hat{=} \psi_\mathrm{B} = \lambda$ and $\psi_\mathrm{R} \hat{=} \psi_\mathrm{A} = \bar{\chi}$]. This difference has to be kept in mind, when translating the spinor formalism into Euclidean space.

II.9.2 The (anti) instanton solution

For the later model, instanton fields in a $SU(2)$ gauge group will be of importance and so all solutions will be restricted to this special case. To be completely clear, the instanton solutions will be derived in Euclidean space-time [$SO(4) \simeq SU_\mathrm{B}(2) \otimes SU_\mathrm{A}(2)$], augmented with the $SU_\mathrm{I}(2)$ gauge group (corresponding to ‘isospin’). In Sec. II.8 it was shown that a n-instanton field, A_μ^cl has to solve the (anti-) selfduality equation, $\widetilde{F}_{\mu\nu} = \mathrm{sign}(n) F_{\mu\nu}$. In 1975, the BPST[29] solution for this equation was found (Ref. [22]) for the special case of $n = 1$ (normal instantons). The derivation of a more general solution is given in Ref. [21, p.17-25]. Both derivations make use of the so called ‘hedge hog’ map for the $SU_{A/B}(2)$ gauge transformations. If $x_\alpha = (x_4, x_1, x_2, x_3)^\mathrm{T}$ is taken as an Euclidean 4-vector and q_α as the quaternion symbol of Sec. II.2, then the $SU_\mathrm{B}(2)$ gauge transformation is given by $U(\hat{x}) = \sum_\alpha x_\alpha \bar{q}_\alpha / \rho$ and the $SU_\mathrm{A}(2)$ transformation has the form $U(\hat{x}) = \sum_\alpha x_\alpha q_\alpha / \rho$. With this, the topological solution for an anti-instanton ($n = -1$) was found to be:

$$A_\mu^\mathrm{reg}(x; x_0, \rho) = \frac{(x-x_0)^2}{(x-x_0)^2+\rho^2}\, U(\hat{x})\partial_\mu U^\dagger(\hat{x}) = \frac{-q_{\mu\nu}(x-x_0)^\nu}{(x-x_0)^2+\rho^2} \qquad \text{for} \quad (n=-1) , \tag{2.9.6}$$

$$A_\mu^\mathrm{sing}(x; x_0, \rho) = \frac{\rho^2}{(x-x_0)^2+\rho^2}\, U^\dagger(\hat{x})\partial_\mu U(\hat{x}) = \frac{-\rho^2\, \bar{q}_{\mu\nu}(x-x_0)^\nu}{(x-x_0)^2[(x-x_0)^2+\rho^2]} \qquad \text{for} \quad (n=-1) . \tag{2.9.7}$$

[28] $\eta_{a\mu\nu} := \epsilon_{a\mu\nu} + \delta_{a\mu}\delta_{\nu 4} - \delta_{a\nu}\delta_{4\mu}$ and $\bar{\eta}_{a\mu\nu} := \epsilon_{a\mu\nu} - \delta_{a\mu}\delta_{\nu 4} + \delta_{a\nu}\delta_{4\mu}$

[29] Belavin, Polyakov, Schwartz, Tyupkin

Here x_0 is an arbitrary parameter, marking the instanton's position and ρ is another parameter, which gives the size of the instanton solution. The second equation is called the singular configuration and can be reached from the regular solution via a simple gauge transformation $A_\mu^{\mathrm{sing}} = A_\mu^{\mathrm{reg}} - \mathrm{i}/g\, U\partial_\mu U^\dagger$. As n is negative the instanton solves the anti-selfduality equation (Eq. (2.8.17)). The gauge orientation term scales like $U^\dagger\partial^\mu U \sim (x-x_0)^\mu/(x-x_0)^2$, which means that the solution A_μ^{sing} is highly localised around x_0[30]. This makes it the preferable choice for calculations, as terms fall off rapidly for $|x| \to \infty$. For the instanton ($n = 1$) the solutions are almost the same. The only difference is that the q-tensors are exchanged ($\bar{q}_{\mu\nu} \leftrightarrow q_{\mu\nu}$), which means that the instanton solves the corresponding selfduality equation.
From Eq. (2.9.6) one could get the impression that instanton solutions depend on five variables (size and position), but this is not the whole truth. Even after a gauge is fixed, there will remain a global $SU_{\mathrm{I}}(2)$ ('isospin') gauge freedom transforming the instanton field to:

$$A_\mu(x; x_0, \rho, \boldsymbol{\theta}) = U_{\mathrm{I}}(\boldsymbol{\theta}) A_\mu(x; x_0, \rho) U_{\mathrm{I}}^\dagger(\boldsymbol{\theta})\,. \tag{2.9.8}$$

Here $U_{\mathrm{I}}(\boldsymbol{\theta}) = e^{-\mathrm{i}\theta_a\tau_a/2}$ is a constant $SU_{\mathrm{I}}(2)$ transformation with $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)^{\mathrm{T}}$. And so there is another set of three variables, that the instanton solution depends on[31]. A derivation of this additional freedom can be found in Ref. [23, p.3441-3444] and the corresponding group theoretical approach in Ref. [21, p.11-13].
The model will be presented in the presence of a so called 'background gauge' which does not violate this $SU_{\mathrm{I}}(2)$ symmetry and therefore the three additional variables of the instanton solution have to be included in all derivations. A general result shows that the ($n = 1$)-instanton solutions depend on $4N$ variables, if $A_\mu \in SU(N)$. A rather detailed derivation of the free instanton solution is given in Ref. [21, p.21-25].

II.9.3 Zero-modes

Having found the instanton solutions (Eq. (2.9.6)), the question is how instantons effect QCD calculations. Maybe the easiest way to address this question is to expand the gauge fields in a given QCD-model around the instanton solution: $A_\mu = A_\mu^{\mathrm{cl}} + a_\mu$, where A_μ^{cl} is the 'classical' instanton solution and a_μ is a perturbation of order $\mathcal{O}(\hbar)$ around it. In the path integral formalism this change of variables is not for free, as one still has to change the integration measure accordingly ($\mathscr{D}A_\mu \to \mathscr{D}a_\mu$). Some of the quantum fluctuations around instanton solutions pose difficulties on this transition, as a number of 'zeros' will show up in inconvenient places in the partition function. The problematic quantum fluctuations are related to the $4N$ degrees of freedom of A_μ^{cl} and dealing with them leads to the concept of collective

[30] For the same reason A_μ^{sing} is singular at $x = x_0$, which is not a problem since the fields only have to be square integrable in $\mathbb{R}^4$.

[31] In the following derivations the isospin symmetry will not play a major for a while and thus it will be picked up explicitly again in Sec. II.9.8.1.

coordinates.

Square integrable quantum fluctuations that preserve the (anti-) selfduality and the winding number of an n-instanton can be shown to fulfil the condition

$$\bar{q}^{\dot{a}b\,\mu} q^{\nu}_{a\dot{b}} D^{\mathrm{cl}}_{\mu} a_{\nu} = 0 \tag{2.9.9}$$

in a background gauge (for an explanation of this condition compare Ref. [21, p.25-26]). The background gauge is usually chosen in instanton calculations for convenience, as it allows for a different gauge of quantum fluctuations compared to the chosen gauge of a classical background field. The general concept of this gauge is reviewed in Ref. [5, p.478-485]. The number of solutions of Eq. (2.9.9) can be found via the Atiyah-Singer index theorem[32)], as this equation can be related to the question of counting zero-modes[33)] of the operator $\bar{q}_{\dot{a}b\,\mu} D^{cl\,\mu}$. The general result, presented in Ref. [24], shows that there are exactly $4Nn$ solutions of Eq. (2.9.9), where, as before, N is the dimension of the underlying 'isospin symmetry', $SU_{\mathrm{I}}(N)$, and n is the winding number ($n = -1$ in the present case).

Now, one is in the position to analyse the expansion of the Euclidean action in a pure Yang-Mills gauge model around $A^{\mathrm{cl}}_{\mu\nu}$:

$$\begin{aligned} S_{\mathrm{E}} &= \int \mathrm{d}^4 x_{\mathrm{E}} \left[\mathrm{tr} A^{\mu\nu} A_{\mu\nu} + \mathcal{L}_{\mathrm{gf}} + \mathcal{L}_{\mathrm{gh}} \right] && (2.9.10) \\ &= \int \mathrm{d}^4 x_{\mathrm{E}} \left[\mathrm{tr} \left(A^{\mathrm{cl}} + a \right)^{\mu\nu} \left(A^{\mathrm{cl}} + a \right)_{\mu\nu} + \mathcal{L}_{\mathrm{gf}} + \mathcal{L}_{\mathrm{gh}} \right] && (2.9.11) \\ &= \frac{8\pi^2}{g^2} |n| + \mathrm{tr} \int \mathrm{d}^4 x_{\mathrm{E}} \left[a_{\mu} M^{\mu\nu} a_{\nu} + \bar{c} M_{\mathrm{gh}} c \right] + O(\{a, c, \bar{c}\}^3) \, . && (2.9.12) \end{aligned}$$

In the last line the minimum contribution has been rewritten, as it is a constant anyway (compare Sec. II.8.3). Here a gauge fixing term and the corresponding ghost Lagrangian have been included (for a review on those two contributions see Ref. [5, p.430-434]). The concrete gauge fixing and ghost Lagrangian have to be specified according to the actual model and gauge choice. To bring the present example into a background gauge ($D^{\mathrm{cl}}_{\mu} a^{\mu} = 0$) one would need:

$$\mathcal{L}_{\mathrm{gf}} = -\frac{1}{g^2} \mathrm{tr} \left(D^{\mathrm{cl}}_{\mu} a^{\mu} \right)^2 \qquad , \qquad \mathcal{L}_{\mathrm{gh}} = -\bar{c} D^{\mathrm{cl}}_{\mu} D^{\mu} c \, . \tag{2.9.13}$$

With these contributions to Eq. (2.9.12), the operators for the quantum fluctuations can be calculated. The derivation is given in Ref. [21, p.35] leading to:

$$M^{\mu\nu} = (D^{\mathrm{cl}})^2 \delta^{\mu\nu} + 2 (A^{\mathrm{cl}})^{\mu\nu} + O(\{a\}) \qquad , \qquad M_{\mathrm{gh}} = -(D^{\mathrm{cl}})^2 + O(\{a, c\}) \, . \tag{2.9.14}$$

32) Strictly speaking this is only correct if the space, in which $\bar{q}_{\dot{a}b\,\mu} D^{cl\,\mu}$ lives is compactified. There are various ways to do this - one explicit compatification in the present case would be to fix the gauge at spacial infinity ($A_{\mu}(\infty) = \mathrm{const}$).

33) A zero-mode, $|Z_j\rangle$, is a normalisable solutions of an operator with eigenvalue, $\epsilon_j = 0$.

For any one-loop calculation the correction terms in the above equation can be dropped, so that both operators are purely built from the covariant derivative with the classical gauge field, D^{cl}. Now the introductory problem of the section is at hand: One knows from the previous discussion that D^{cl} has zero-modes and thus $M^{\mu\nu}$ and M_{gh} have zero-modes, too. By performing a path integral over all fluctuations one effectively produces - apart from finite contributions - an infinite sum over all zero-mode contributions, $\int(\mathscr{D}a_\mu)_0\, e^0 = \infty$, where the subscript 0 means that the path integral is only performed over zero-modes.
At this point it becomes clear that it is no coincidence that the instanton degrees of freedom exactly match the number of zero-modes of the quantum fluctuations. A fluctuation, which changes A_μ^{cl} in a direction $\gamma_k \in \{(x_0)_i, \rho, \theta_j\}$ does not change the total action of Eq. (2.9.12), as γ_k is an explicit degree of freedom of the instanton solution with action, $S_{\mathrm{E}} = 8\pi^2/g^2$. So the contribution from the fluctuations, $a_\mu M^{\mu\nu} a_\nu$, has to vanish, leading to the divergence in the functional integration.

II.9.4 Collective coordinates

To better deal with the zero-mode fluctuations one can rewrite their functional integration into a direct integration over the degrees of freedom γ_k. The γ_k are, what is conventionally known as 'collective coordinates'. The techniques to make the transition $(\mathscr{D}a_\mu)_0 \to \prod_k \mathrm{d}\gamma_k$ are very similar to the Faddeev-Popov concept to get rid of gauge redundencies in non-abelian gauge theories. Therefore many concepts in the following paragraph can be found in greater detail in Ref. [5, p.430-434]. The idea is to artificially include a constraint function $f_i^\mu(\gamma) \equiv f_i^\mu(\gamma; x)$[34)] for each collective coordinate into the path integral and then integrate over all of them, such that the final partition function does not depend on the arbitrary constraints, $f_i^\mu(\gamma)$, anymore. The constraint can be included via the identity:

$$1 = \int \mathrm{d}^{4N}\gamma \Delta_\gamma \prod_i^{4N} \delta\left([A_\mu - A_\mu^{\mathrm{cl}}(\gamma)] f_i^\mu(\gamma)\right) . \tag{2.9.15}$$

Here Δ_γ is a needed Jacobian (see Ref. [5, p.432]). To leading order in the fluctuation (or equivalently the coupling constant) it has the form:

$$\Delta_\gamma = \left| \det_{ij} \left(\int \mathrm{d}^4x \frac{\partial A_\mu^{\mathrm{cl}}(\gamma)}{\partial \gamma_i} f_i^\mu(\gamma) \right) \right| . \tag{2.9.16}$$

[34)] For notational convenience the explicit dependence on x will not be mentioned anymore, just as it is already convention for fields.

In Eq. (2.9.15) the δ-distribution can be replaced by a gaussian via:

$$\prod_{i=1}^{4N} \delta\left(\int \mathrm{d}^4x\, a_\mu f_i^\mu\right) = \lim_{\alpha\to 0} \frac{1}{(2\pi\alpha)^{n/2}} \exp\left[-\frac{1}{2}\int \mathrm{d}^4x\, a_\mu \left(\alpha^{-1} f_i^\mu \times f_i^\nu\right) a_\nu\right]. \tag{2.9.17}$$

The indices of the constraint functions f_i^μ has to be summed over. Using this, the zero-mode contribution of the partition function can be rewritten to:

$$Z_0 = \int (\mathscr{D}A_\mu)_0\, e^{-S_\mathrm{E}(A_\mu)} \sim e^{-S_\mathrm{E}^\mathrm{cl}} \int \mathrm{d}^{4N}\gamma\, Z_\gamma\,, \tag{2.9.18}$$

$$Z_\gamma = \Delta_\gamma \lim_{\alpha\to 0} \frac{1}{(2\pi\alpha)^{4N/2}} \int \mathscr{D}a_\mu \exp\left[-\frac{1}{2}\int \mathrm{d}^4x\, a_\mu \left(M^{\mu\nu} + \alpha^{-1} f_i^\mu \times f_i^\nu\right) a_\nu\right] \tag{2.9.19}$$

$$= \Delta_\gamma \lim_{\alpha\to 0} \frac{1}{(2\pi\alpha)^{4N/2}} \det(M^{\mu\nu} + \frac{1}{\alpha} f_i^\mu \times f_i^\nu)^{-1/2}. \tag{2.9.20}$$

Here the sum over the index, i, has been left out as well as the entire ghost field part, since its functional integral only gives the usual contribution. Notice that, without the constraint function, the partition function would diverge ($\det(M^{\mu\nu})$=0, since it contains zero-modes), as expected. With the constraint in place the situation is different. Now the modified determinant has to be evaluated. For this the reasoning from Ref. [25] has been adopted. There, the general identity

$$\det\left(M^{\mu\nu} + b_i^\mu \times b_i^\nu\right) = \det(M_{\mu\nu}) \det_{ij}\left(\delta_{ij} + b_i^\mu M_{\mu\nu}^{-1} b_i^\nu\right). \tag{2.9.21}$$

was rewritten for the case where M has zero-modes. The arbitrary vectors, b_i^μ, have been rescaled by limiting factors of α, leading to:

$$\lim_{\alpha\to 0} \det(M^{\mu\nu} + \frac{1}{\alpha} f_i^\mu \times f_i^\nu) = \det{}'(M^{\mu\nu}) \det_{ik}(f_i^\mu Z_{k\mu})[\det_{kl}(Z_k^\mu Z_{l\mu})]^{-1} \det_{lj}(Z_l^\mu f_{j\mu}). \tag{2.9.22}$$

The prime denotes that the determinant is only to be taken over non-zero eigenvalues and the Z_j^μ are the zero-modes of $M^{\mu\nu}$. The zero-modes correspond to changes of the classical solution in a direction of γ_j and thus they have the form $Z_{j\mu} \sim (\partial A_\mu^\mathrm{cl})/(\partial\gamma_j)$. Therefore the previously introduced Jacobian can be written as $\Delta_\gamma = \det_{ij}(Z_i^\mu f_{j\mu})$. Combining this with the equations Eq. (2.9.22), (2.9.18), (2.9.20) and absorbing the divergent factor, $\lim_{\alpha\to 0}\alpha^{-4N/2}$, in the normalisation of the path integral one finally arrives at a well defined expression for the partition function of quantum fluctuations around a classical instanton solution:

$$Z = \int \mathscr{D}A_\mu e^{-S_\mathrm{E}} \tag{2.9.23}$$

$$= e^{-S_\mathrm{E}^\mathrm{cl}} \int \frac{\mathrm{d}^{4N}\gamma}{(2\pi)^{4N/2}} [\det_{kl}(Z_k^\mu Z_{l\mu})]^{1/2} [\det{}'(M^{\mu\nu})]^{-1/2}. \tag{2.9.24}$$

This equation incorporates the influence of all quantum fluctuations to the instanton partition function and is the main result of this section. The fluctuations 'perpendicular' to the collective coordinates are summed over in the 'amputated determinant' and the fluctuations along those coordinates are captured in the remaining integral over $\mathrm{d}^{4N}\gamma$. In the case of an $SU_{\mathrm{I}}(2)$ gauge group the remaining integration is eight dimensional: $\mathrm{d}^{4\times 2}\gamma = \mathrm{d}^4 x_0 \, \mathrm{d}\rho \, \mathrm{d}^3\theta$. While the here presented derivation is very hasty, there is a much deeper treatment of the subject given in Ref. [26]. The derivation given there also provides a solution, if the point of expansion does not correspond to the exact classical solution A_μ^{cl}.

II.9.5 Measure of zero-modes

In the last section the partition function of an instanton has been rewritten. What is still left to be done, is to find the explicit form of the involved determinants of Eq. (2.9.24). The first of these gives a contribution from all zero-modes and can be calculated once the explicit form of all those modes is established. For this define

$$U_{ij} = \langle Z_i | Z_j \rangle = \frac{-2}{g^2} \mathrm{tr}_{\mathrm{I}} \int \mathrm{d}^4 x Z_{i\mu} Z_j^\mu \,, \tag{2.9.25}$$

so that the determinant contribution becomes $\left(\det_{ij}(U_{ij})\right)^{1/2}$. The trace runs over the $SU_{\mathrm{I}}(2)$ indeces of the gauge group. It was aready mentioned in the previous section that the zero-modes are similar to differentiations of the classical solution with respect to the collective coordinates. This would be already the end of the story if it weren't for the background gauge that still must be fulfilled. In a pure gauge field model it has the form:

$$D_\mu^{\mathrm{cl}} a^\mu = \left(\partial_\mu + A_\mu^{\mathrm{cl}}\right) a^\mu = 0 \,, \tag{2.9.26}$$

where the superscript cl refers to the classical instanton solution. Having this gauge, the zero-modes - as they are quantum fluctuations - have to respect it as well and so the eight modes are of the form:

$$Z_{i\mu} = \frac{\partial A_\mu^{\mathrm{cl}}}{\partial \gamma_i} + D_\mu^{\mathrm{cl}} \Lambda_i \,. \tag{2.9.27}$$

Here Λ_i are gauge parameters that can be used to reestablish the background gauge. Now all zero-modes of a pure $SU_{\mathrm{I}}(2)$ Yang-Mills model can be calculated. As this is a mere technical task, here simply the results are given. The complete derivation can be reviewed in Ref. [21, p.35-39].

$$Z_{\rho\,\mu} = -2\frac{\rho\bar{q}_{\mu\nu}x^\nu}{(x^2+\rho^2)^2} \qquad \text{with} \qquad \Lambda_\rho = 0\,, \tag{2.9.28}$$

$$Z_{x_0^\alpha\,\mu} = F^{\rm cl}_{\mu\alpha} \qquad \text{with} \qquad \Lambda_{x_0^\alpha} = A^{\rm cl}_\alpha\,, \tag{2.9.29}$$

$$Z_{\theta_j\,\mu} = D_\mu\left[\frac{x^2}{x^2+\rho^2}T_j\right] \qquad \text{with} \qquad \Lambda_{\theta_j} = -\frac{\rho^2}{x^2+\rho^2}T_j\,. \tag{2.9.30}$$

Now the matrix U_{ij} can be calculated[35] in terms of the instanton action $S^{\rm cl} = (8\pi^2)/g^2$:

$$U_{ij} = \begin{pmatrix} \delta_{\mu\nu}S^{\rm cl}\Big|_{4\times4} & & \\ & 2S^{\rm cl}\Big|_{1\times1} & \\ & & \frac{1}{2}g_{ab}(\boldsymbol{\theta})\rho^2 S^{\rm cl}\Big|_{3\times3} \end{pmatrix}_{8\times8}\,. \tag{2.9.31}$$

The first entry ($\delta_{\mu\nu}S^{\rm cl}$) is the 4-dimensional part from the x_0-zero mode, the second entry comes from the instanton size ρ and the last entry is the 3-dimensional contribution from the gauge group zero-modes. $g_{ab}(\boldsymbol{\theta})$ is a 3-dimensional matrix that encodes the combination of any two generators T_α, T_β of the gauge group. This matrix is independent of the group representation (apart from a normalisation) and taking its determinant gives the Haar measure of the $SU_{\rm I}(2)$ group. Given all this, the determinant over zero-modes in Eq. (2.9.24) turns out to be:

$$\sqrt{\det(U_{ij})} = \frac{1}{2}(S^{\rm cl})^4\rho^3\sqrt{\det(g_{ab}(\boldsymbol{\theta}))}\,. \tag{2.9.32}$$

If no entry in the partition function depends explicitly on the orientation in $SU_{\rm I}(2)$ space, then the integration over the collective coordinates $\boldsymbol{\theta}$ can be calculated right away:

$$\mathrm{vol}(SU_{\rm I}(2)) = \int \mathrm{d}^3\theta\,\sqrt{\det(g_{ab}(\boldsymbol{\theta}))} = 2\pi^2\,. \tag{2.9.33}$$

The last equality is actually dependent on the chosen normalisation for the generators[36] (here $T_a = -\mathrm{i}\tau_a/2$). The derivation so far appears to be easily extendable to gauge groups with arbitrary dimensions N. This is not true, as the $SU(2)$ instanton solution has to be embedded into higher dimensional gauge groups and this changes the overall factor of $\sqrt{U}$. The reason for the chosen notation is comparability

[35] To see that $\left|Z_\rho\right|^2 = 2S^{\rm cl}$ one needs the identity $\mathrm{tr}(\bar{q}_{\mu\nu}\bar{q}^{\mu\nu}) = 12$, which can be derived from Eq. (2.9.5).

[36] For an explicit calculation of various group volumina compare Ref. [21, p.100-106].

with Ref. [21]. The zero-mode contribution to the measure of the partition function in the present case is

$$\frac{\sqrt{\det(U_{ij})}}{(2\pi)^{4N/2}} = \frac{(S^{\mathrm{cl}})^4\rho^3}{2(2\pi)^4}\mathrm{vol}(SU_{\mathrm{I}}(2)) = \underbrace{\frac{2^8\pi^6}{g^8\rho^5}}_{=:m_1(\rho)}\rho^8\,. \tag{2.9.34}$$

In the definition of the collective coordinate measure, $m_1(\rho)$, a factor of ρ^5 has been included, in order to give a dimensionless integral ($\mathrm{d}^4x_0\mathrm{d}\rho\,\rho^{-5}$). The remaining factor of ρ^8 will be subject to renormalisation in Sec. II.9.7.

$m_1(\rho)$ is different by a factor of 2^2 in comparison with the literature[37)]. Usually in literature $N \geq 3$ gauge models are analysed, where the mentioned embedding of the $SU(2)$ instantons in the $SU(N)$ gauge group has to be taken into account. In Ref. [21, p.45], the difference is notable directly to be a factor of 4 in the 'group volume'.

II.9.6 Fermionic zero-modes

This section represents a slight intermission in the derivation of the bosonic partition function. In a pure gauge field model there are no fermions and thus there is no direct need to burden ones life with the complications they bring along. Of course, as the final model in this work shall describe nucleons, there is an obvious need to consider also this extended case. This section will give some basic concepts for the treatment of fermions in the later model.

In general, if massless fermions are included in a chiral QCD-model, the influence of instantons becomes unphysical[38)]. The reason is, that the $U(1)$ anomaly of these models allows to gauge the instanton contribution in the θ-partition function (Eq. (2.8.24)) away[39)]. In the path integral formalism this can be understood as the appearence of zero-modes in the fermionic path integral. Take for example the following generic model from Ref. [5, p.601]:

$$Z = \int \mathscr{D}\{\bar\psi\psi A_\mu\}\,\exp\left[\mathrm{i}\int \mathrm{d}^4x\,\bar\psi\mathrm{i}\not{D}\psi - \frac{1}{4}A_{\mu\nu}A^{\mu\nu} - \frac{g^2\theta}{32\pi^2}\widetilde{A}_{\mu\nu}A^{\mu\nu}\right] \tag{2.9.35}$$

$$= \int \mathscr{D}A_\mu\,\det(\mathrm{i}\not{D})\,e^{\mathrm{i}S}\,e^{\mathrm{i}n\theta}\,. \tag{2.9.36}$$

The above example gives the partition function for the transition from the 0-angle to the θ-angle vacuum and indeed this expression vanishes if the Dirac operator of the massless fermion ψ has zero-modes.

The Atiyah-Singer index theorem can be used to further evaluate these zero-modes. This general treatment is presented for the present case in Ref. [21, p.28-34]. For now it is sufficient to know that fermions

37) For example: Ref. [21, p.45], [23, p.3445] (including erratum) and [27, p.8].
38) To see the full argument compare Ref. [5, p.601-602]
39) For more information on the $U(1)$ anomaly see Ref. [5, p.456-477].

in the fundamental representation have $|n|$ zero-modes in the presence of an n-instanton field (Ref. [21, p.34]). It was found that the anti-instanton ($n = -1$) leads to a zero-mode for the Euclidean Weyl-spinor, $\lambda \in SU_{\mathrm{B}}(2)$ (corresponding to the left-handed fermion ψ_{L} in Minkowski space), and the instanton ($n = 1$) gives a zero-mode for the other Weyl-spinor, $\bar{\chi} \in SU_{\mathrm{A}}(2)$ (which translates to the right-handed fermion ψ_{R} in Minkowski space). Here the $SU_j(2)$ groups correspond to the subgroups of $SO(4)$ as they have been introduced in Sec. II.9.1.2.

Anyway, having the explicit form of the instanton field, A^{cl}_μ, (Eq. (2.9.6)), the fermionic zero-modes can be calculated explicitly by solving the equation $\not{D}^{\mathrm{cl}}\,\Psi_0 = \gamma^\mu_{\mathrm{E}} D^{\mathrm{cl}}_\mu \Psi_0 = 0$, where Ψ_0 is a Euclidean Dirac-spinor. Using the definitions of the Euclidean γ-matrices Ψ can be decomposed into the 2 independent Weyl-spinors: $\lambda = \frac{1}{2}(1 - \gamma_{\mathrm{E}\,5})\Psi_0$ and $\bar{\chi} = \frac{1}{2}(1 + \gamma_{\mathrm{E}\,5})\Psi_0$ and the zero-mode equation then gives the two relations:

$$\bar{q}^\mu D^{\mathrm{cl}}_\mu \lambda = 0 \qquad ; \qquad q^\mu D^{\mathrm{cl}}_\mu \bar{\chi} = 0\,, \tag{2.9.37}$$

$$\left(\bar{q}^\mu \partial_\mu + \bar{q}^\mu A^{\mathrm{cl}}_\mu\right)\lambda = 0 \qquad ; \qquad \left(q^\mu \partial_\mu + q^\mu A^{\mathrm{cl}}_\mu\right)\bar{\chi} = 0\,. \tag{2.9.38}$$

These equations can be solved by using the explicit form of A^{cl}_μ and the identities for the combinations of q_μ, $\bar{q}_\mu$ and $\bar{q}_{\mu\nu}$ (compare Ref. [21, p.97-100]). For the anti-instanton, one finds the following normalised[40)] fermionic zero-mode:

$$\lambda^{\mathrm{sing}}(x; x_0, \rho, \mathcal{K}) = \frac{\rho}{\pi}\frac{q^\mu_{a\dot{c}}(x-x_0)_\mu}{\left[(x-x_0)^2((x-x_0)^2+\rho^2)^3\right]^{1/2}}\,\mathcal{K}\,, \tag{2.9.39}$$

$$\lambda^{\mathrm{reg}}(x; x_0, \rho, \mathcal{K}) = \frac{\rho}{\pi}\frac{\epsilon_{ac}}{\left[(x-x_0)^2+\rho^2\right]^{3/2}}\,\mathcal{K}\,. \tag{2.9.40}$$

Here $\mathcal{K}$ is a grassmann valued variable, which is the so called fermionic collective coordinate[41)]. From now on, often the explicit dependence on x_0 will be dropped, if it is unimportant for explicit derivations. The antisymmetric tensor, ϵ_{ac}, in the regular gauge zero-mode can be understood from the group theoretical point of view, as the only possible invariant symbol in the combination of two $1/2$ spinors ($2 \otimes 2 = 1_a \oplus 3_s$). Apart from this formal mathematical argument the indices of ϵ_{ac} can also be identified with the spin and 'isospin' dependence of the mode. They reflect the fact that instantons in general depend on the symmetries of the space and thus this dependence is adopted in the zero-modes as well (here the spin-space is coupled to the 'isospin'-space). A clear derivation of this interpretation for λ^{reg} is given in Ref. [28, p.283-285].

Note that the fermionic zero-mode does not have the correct units for a fermionic field in four dimen-

40) The norm is defined as: $\left[\langle\lambda|\lambda\rangle = \frac{1}{2}\mathrm{tr}_{\mathrm{I}}\int \mathrm{d}^4x\,\lambda^\dagger(x)\lambda(x)\right]$.

41) This is not a collective coordinate as the ones introduced before, as it is not a degree of freedom of the instanton solution but the name is conventional.

sions, which is normally in powers of energy: $[\psi] = 3/2$[42]. The reason for the (at first sight) wrong units is a different normalisation. The correction of all involved units is explicitly performed in Sec. II.9.9.4. λ is a zero-mode for every $\mathcal{K}$ and so the fermionic path integral can easily be split up into $\mathscr{D}\{\bar{N}N\} \to \mathscr{D}\{\bar{\mathcal{K}}\mathcal{K}\}\left(\mathscr{D}\{\bar{N}N\}\right)_{\neq 0}$, where the subscript $\neq 0$ means that the fermionic zero-modes have been extracted. Massless fermions, as already mentioned, nullify the effect of instantons, while massive fermions on the other hand force the vacuum angle, θ, to be very close to zero (compare Sec. II.8.4), so the question comes to mind what the significance of the last 15 pages could be in real physical models. Of course, there is a way out of this dilema: One could apply perturbation theory. Suppose the fermionic Lagrangian in the partition function Eq. (2.9.35) is enhanced with an additional perturbative contribution: $\mathcal{L}_N = \bar{N}(\mathrm{i}\not{D} + g_J J)N$. Then, to lowest order in perturbation theory the zero-mode contribution to the determinant in the partition function of this model becomes:

$$\det{}_0(\mathrm{i}\not{D} + g_J J) \approx \langle N_0|\mathrm{i}\not{D}|N_0\rangle + g_J\langle N_0|J|N_0\rangle = g_J\langle N_0|J|N_0\rangle\,. \tag{2.9.41}$$

Here $|N_0\rangle$ stands for a suitable zero-mode, depending on the explicit model and conventions and the subscript 0 at the determinant means that only the zero-modes are included. The above expansion is possible, as the Dirac operator, $\not{D}$, can be expanded in eigenfunctions. This perturbative concept for the treatment of fermionic zero-modes will be used in the later model, as well.

II.9.7 Non-zero quantum fluctuations

Coming back to the calculation of the partition function of Eq. (2.9.24), the last contribution, still waiting for evaluation is the amputated determinant from the non-zero quantum fluctuatioins $(\det'(M^{\mu\nu}))^{-1/2}$. While one could start to find the eigenfunctions, and values of $M^{\mu\nu}$ for an explicit evaluation of the determinant, it is far more convenient to understand what this path integral over vacuum fluctuations actually corresponds to: In the so far developed formalism a pure Yang-Mills model has been expanded around a classical field configuration, A_μ^{cl}. In the language of Feynman perturbation theory this means that external lines in diagrams correspond to classical fields, while internal loops are produced by the quantum fluctuations a_μ. Thus evaluating the determinant of $M^{\mu\nu}$ can be translated to calculating the one-loop quantum correction to a background field in a Yang-Mills model. In principle this is already it and one is almost done. There is just one conceptual difficulty, which should be mentioned: As the instanton field A_μ^{cl} depends crucially on the topology of the space one is working in, it is not straight forward to use the standard tool of dimensional regularisation in the context of instantons. A way around this problem is to use the Pauli-Villars regularisation scheme instead. Ultimately the results of different regularisation procedures for renormalisable models are equivalent and do only alter by an overall constant. Therefore

[42] This notation basically countes the powers of energy as a unit-dimension of objects and will be explained completely in Sec. III.4.2. The full machinery for analysis in this notation is presented in Ref. [5, p.90-91].

the final result derived via one scheme can easily be translated into another one. In his original work 't Hooft presented a complete derivation of both schemes Ref. [23] and also gave the overall factor by which they differ.
Before coming to the β-function, and by this the actual results for the quantum fluctuations, it is useful to generalise the free Yang-Mills Lagrangian from Eq. (2.9.24). It is possible to give the one-loop β-function for a model including n_f Dirac fermions in some group representation, R_{DF}, and n_s complex scalars in a group representation, R_CS. Such a model has the Lagrange density:

$$\mathcal{L}_\mathrm{gen} = \mathcal{L}^\mathrm{f}_\mathrm{YM} + \mathcal{L}^\mathrm{f}_N + \mathcal{L}^\mathrm{f}_\Omega + g_\Omega \bar{N}\Omega N + \mathcal{L}_\mathrm{gf} + \mathcal{L}_\mathrm{gh} \,. \tag{2.9.42}$$

where the 'free' Lagrangians in order of appearance are: Yang-Mills, nucleon, N, scalar field, Ω, gauge fixing and ghost field. The nucleon and scalar field are coupled to the gauge field via the covariant derivative and the Ω-nucleon coupling is given explicitly. *Only* the free Lagrangians for Ω and N may contain a mass term, as this will not change the final result. This model can be expanded around the instanton field, A^cl_μ, with all other fields only contributing as quantum fluctuations:

$$Z_\mathrm{gen} = \int \mathscr{D}\{\bar{N}N\Omega A_\mu \bar{c}c\}\, e^{-\int \mathrm{d}^4x_\mathrm{E}\mathcal{L}_\mathrm{gen}} \tag{2.9.43}$$

$$= \int \mathscr{D}\{\bar{n}na_\mu\}_0 \mathscr{D}\{\bar{n}n\omega a_\mu \bar{c}c\}_{\neq 0}\, e^{-\int \mathrm{d}^4x_\mathrm{E}\mathcal{L}_\mathrm{gen}} \,. \tag{2.9.44}$$

The second line already shows the expansion around the instanton solution in the integration measure (n, ω and a_μ stand for the fluctuation fields). As earlier the subscript $\neq 0$ means an integration over non-zero modes and 0 stands for the integration only over zero-modes. Note that only the fermionic fields and the gauge field have zero-modes. Now it is time to recollect the results from the previous sections, namely equation: Eq. (2.9.24), (2.9.34) and (2.9.41). Using this and performing the standard gaussian path integrals over bosonic and fermionic functionals gives:

$$Z_\mathrm{gen} = \int \mathrm{d}^4x_0 \mathrm{d}\rho\, m_1(\rho)\, \underbrace{e^{-S^\mathrm{cl}_\mathrm{E}}\rho^8 \frac{\det'(M_N)\det(M_c)}{\left[\det'(M^{\mu\nu})\det(M_\Omega)\right]^{1/2}}}_{=:m_2(\rho,\mu_0)} \times \det_0(\mathrm{i}\not{D} + g_\Omega\Omega) \,. \tag{2.9.45}$$

Here M_j stands for the operator that is sandwiched between the two quantum fields, j, before the functional integration is performed. For the moment ignore the last zero-mode determinant, $\det_0(\mathrm{i}\not{D} + g_\Omega\Omega)$, as it is actually zero in this example ($\Omega = 0$ at the expansion point from the earlier assumption). Later this will not be the case but there is still some work to be done, in order to promote Ω to non-zero expectation values. The determinant fractional is actually the expression that just corresponds to the one-loop quantum corrections to the classical solution, $e^{-S^\mathrm{cl}_\mathrm{E}}$. This can now be calculated, either following 't Hoofts' derivation[43] (Ref. [23]), or calculating the one-loop β-function of this generic model (Eq. (2.9.42)), like it is presented in the background gauge formalism in Ref. [5, p.439-447;478-485]. The result for the renormalised correction to the measure at the renormalisation scale μ_0 is:

$$m_2(\rho,\mu_0) = \exp\left[-\frac{8\pi^2}{(g_{\mathrm{R}}^{\mathrm{D}}(\mu_0))^2} + \log(\rho\mu_0)\widetilde{\beta}(N,\Omega) + A - n_{\mathrm{f}}B - \sum_I n_{\mathrm{s}}(I)A(I)\right] \tag{2.9.46}$$

$$\widetilde{\beta}(N,\Omega) = \left[\frac{22}{3} - \frac{2}{3}n_{\mathrm{f}} - \sum_I \frac{1}{6}n_{\mathrm{s}}(I)C(I)\right]^{44)} \tag{2.9.47}$$

The above equation is given in 't Hoofts' notation (Ref. [23, p.3448]). $g_{\mathrm{R}}^{\mathrm{D}}(\mu_0)$ is the dimensionally renormalised coupling constant, $n_{\mathrm{s}}(I)$ is the number of complex scalars in the isospin representation I, n_{f} is the number of $I = 1/2$ fermions and A, $A(I)$, B and $C(I)$ are numerical constants, whose values are given in Tab. 2.9.1. They arise from the regularisation procedure. The summation runs over all appearing isospin representations of the complex scalar field Ω.

Table 2.9.1: Numerical constants of the β-function for the present case of instanton calculations in a 'background gauge' without 'higgs-like' interactions. For a derivation, including the analytical dependencies compare Ref. [23, p.3445-3446].

I	$C(I)$	$A(I)$	A	B
0	0	0	6.998435	0.49412
1/2	1	0.239246		
1	4	0.816799		
3/2	10	1.786912		

At first sight one might feel uneasy about the transition of the instanton radius from ρ^8 to $\rho^{\widetilde{\beta}}$ in the correction to the measure. The simplest argument for this is a dimensional reasoning. The renormalisation scale μ_0 has the dimension of energy. So, in order to make $m_2(\rho,\mu_0)$ dimensionless, one needs another dimensionfull quantity. The only options in the present model are the instanton position, x_0, and its size, ρ. As the model is translationally invariant it does not make sense to use x_0 in the measure and so the only meaning- and dimensionfull quantity that could multiply μ_0 is ρ[45]. The model is explicitly renormalised in ρ. This ends the discussion of free instantons and one has the final result of the renormalised partition

43) This derivation also gives explicitly the non-vanishing constant factors for the readers' convenience.

44) The analog β-function in Ref. [5, p.484] has the form $\widetilde{\beta}(N,\Omega) \propto \left[\frac{11T(A)}{3} - \frac{4T(R_{DF})}{3}n_{\mathrm{f}} - \frac{T(R_{\mathrm{CS}})}{3}n_{\mathrm{s}}\right]$, where $T(j)$ is the representation index Ref. [5, p.422].

45) Carefully following the renormalisation procedure produces the correct factor of $\rho^{\widetilde{\beta}}$ as well.

function for the generic model:

$$Z_{\text{gen}} = \int \mathrm{d}^4 x_0 \mathrm{d}\rho\, \mathfrak{m}_1(\rho)\mathfrak{m}_2(\rho, \mu_0) \times \det{}_0(\mathrm{i}\not{D} + g_\Omega \Omega)\,. \tag{2.9.48}$$

II.9.8 Constrained instantons

The previous sections gave a broad introduction to instantons starting from some general concepts over instanton solutions and collective coordinates to the correction to the path integral measure for a generic QCD model. Although the concepts are partly very technical, one has fine, analytic and (for the peresent purpose) at first sight apparently useless solutions[46)]. The problems with the so far presented techniques are related to the purpose of the later model in chapter III and the role of the $SU_{\mathrm{I}}(2)$ instantons in it. Later, an effective model in terms of nucleons and mesons shall be built and therefore the degrees of freedom should to be identified with fermionic- or suitable bosonic-fields. The (instanton) gauge field now is a vector field with isospin $I = 1$, which would make it a suitable candidate for the ϱ-meson, for example. But the ϱ-meson, like all effective low energy variables with vectorial properties, is quite massive ($m_\varrho \approx 770\,\mathrm{MeV}$) and so the instanton would have to be massive for this identification. A consistent way to generate masses for gauge bosons is the Higgs-mechanism (Sec. III.5.2) but unfortunately the whole concept of instantons breaks down, if a Higgs-Lagrangian is included in the generic model (Eq. (2.9.42)). The reason for this direct breakdown lies in the intimate relation between instantons and topology. The instanton solution is a minimum of the action, which is stabelised by the topology of the system. By adding a Higgs-field with a non-zero expectation value, effectively the action can be lowered without bounds and thus the instanton solution vanishes. This statement can be verified by taking a Yang-Mills-Higgs [YMH] model (with H as Higgs field):

$$S_{\mathrm{E}} = \frac{1}{g^2}\int \mathrm{d}^4 x\, \mathrm{tr}_{\mathrm{I}} \left\{ \frac{1}{2} A_{\mu\nu}(x) A^{\mu\nu}(x) + \underbrace{\frac{g^2}{\lambda^2}}_{=:\tilde{g}^2} \left[\left| D_\mu H(x) \right|^2 + \frac{1}{4}\left(|H(x)|^2 - \langle H \rangle^2 \right)^2 \right] \right\}. \tag{2.9.49}$$

and rescaling it according to $A_\mu(x) \to a A_\mu(ax)$ and $H(x) \to H(ax)$. This scaling preserves the correct behaviour of both fields at infinity. If $S_{\mathrm{E}}(a)$ is evaluated at the free instanton solution, it turns out that taking $a > 1$ leads to a smaller value for the action and with this S_{E} can be lowered without bounds. The scaling argument is known as 'Derrick's theorem' (Ref. [29]). It can be reviewed for the above example in Ref. [30, p.439-440].

The transformation $A_\mu(x) \to a A_\mu(ax)$ is nothing but a replacement of the instanton size with $\rho \to \rho/a$ (compare Eq. (2.9.6)) and so the rescaling can be understood as diminishing the instanton's size -

46) Now this is a drastic comment on 20 pages, which gives a strong hint that the author is still interested in saving the instanton concept for the later model.

ultimately to zero (this gives the whole contribution the appearance of a δ-distribution). For an $SU(2)$ instanton this is the only direction, in which the action can be lowered by a suitable rescaling. That is to say, if it wasn't for the size, ρ, degree of freedom, the instanton would still minimise the action of the model. This is the crucial observation that led to the concept of 'constrained instantons'. In principle the above argument rules out ordinary instantons completely. So, if one would like to keep something comparable to the instanton mechanism, *the chosen model has to be changed* via the inclusion of a constraint that forbids the action to be lowered indefinitely in the ρ-direction.
The first to give a qualitative discussion of this very difficult area of instanton physics was Affleck in 1980 (Ref. [30]). It was already known that the free instanton solution exists, if the vacuum expectation value [VEV] of the Higgs field is identically zero (compare Sec. II.9.7). From this he developed a formalism that allowed preseving the free solution, at least in a small region, if the VEV was sufficiently small. For this he introduced the following two constraints:

$$\mathcal{L}_A^{\text{con}} = \sigma_1 \int \mathrm{d}^4x \left[\mathrm{tr}A^3 - c_1\rho^{-2}\right] , \tag{2.9.50}$$

$$\mathcal{L}_H^{\text{con}} = \sigma_2 \int \mathrm{d}^4x \left[(\mathrm{tr}(H^\dagger H) - \langle H\rangle^2)^2 - c_2\rho^{-2}\right] . \tag{2.9.51}$$

for the gauge field, A_μ, and the Higgs field, H. The parameters σ_j are functions, that have to be adjusted order by order in perturbation theory, to meet the boundary conditions for the fields and the parameters c_j can be fixed such that the constrained instanton solutions still coincides with the free solution in a small volume around the instanton position, x_0. Outside this region the constrained instantons decay exponentially. The neatness of Affleck's concept is, that it gives an expansion of the constrained instanton solution, which leads to only one new term in the partition function - a gaussian cutoff. All other results concerning the original instanton calculations are untouched to first order in Affleck's expansion. On the other hand it seems rather hard to give an analytic expression for the constrained instanton field in this procedure.

Later, a slightly different point of view was proposed by Wang (Ref. [25]). He did not fix a constraint in order to find a perturbative solution for the field equations, but in some sense took the opposite direction. He argued that the constrained instanton in a YMH model corresponds to a 'valley direction'[47], parametrised by the instanton's size, ρ, which has to be compatible with all constraint-independent results of the model. To find the exact valley direction, he employed the ansatz for the constrained instanton, $A_\mu^{\text{con}}(x)$[48], (in singular gauge) and the corresponding Higgs field, $H^{\text{con}}(x)$ (as Lagrangian he used Eq. (2.9.49), as before):

$$A_\mu^{\text{con}}(x) = B(|x|)\frac{-\bar{q}_{\mu\nu}x^\nu}{x^2}{}^{49)} \qquad , \qquad H^{\text{con}}(x) = (1 - h(|x|))\langle H\rangle\,. \tag{2.9.52}$$

[47] A given function varies slowly in the 'valley direction' compared to other possible directions.
[48] For concistency with earlier results one has to replace $x = x' - x_0$.

By expanding $B(|x|)$ and $h(|x|)$ with respect to the dimensionless parameters $(\rho\,|x|^{-1}, \langle H\rangle\,|x|)$ and forcing them to fulfil the limiting (constraint independent) cases of

$$B(|x|) = \frac{\rho^2}{x^2+\rho^2} \quad , \quad h(|x|) = 1 - \sqrt{\frac{x^2}{x^2+\rho^2}} \quad \text{for} \quad |x| << \rho\,, \tag{2.9.53}$$

$$B(|x|) = m_{\mathrm{A}}^2 x^2 K_2(m_{\mathrm{A}}\,|x|) \quad , \quad h(|x|) = \frac{m_{\mathrm{H}}\,|x|}{2} K_1(m_{\mathrm{H}}\,|x|) \quad \text{for} \quad |x| >> \rho\,. \tag{2.9.54}$$

he ultimately gained an analytic expression for the "best" valley direction:

$$A_\mu^{\mathrm{con}}(x) = \frac{-\rho^2 m_{\mathrm{A}}^2 K_2(m_{\mathrm{A}}\,|x|)}{2+\rho^2 m_{\mathrm{A}}^2 K_2(m_{\mathrm{A}}\,|x|)}\,\frac{\bar{q}_{\mu\nu}x^\nu}{x^2}\,, \tag{2.9.55}$$

$$H^{\mathrm{con}}(x) = \left(1+\rho^2 m_{\mathrm{H}}\frac{K_1(m_{\mathrm{H}}\,|x|)}{|x|}\right)^{-1/2}\langle H\rangle\,. \tag{2.9.56}$$

Here $K_j(x)$ is the modified Bessel function of the second kind (and order j) and the masses are identified as coefficients of the quadric terms in the Lagrangian: $m_{\mathrm{A}} = (\widetilde{g}/2)^{1/2}\langle H\rangle$ and $m_{\mathrm{H}} = \langle H\rangle$. Note that $H^{\mathrm{con}}(x)$ is induced by the constrained instanton field. So the classical field configuration, around which the model will be expanded has a non trivial space dependence. The expansion does not depend on the last dimensionless parameter $\rho\langle H\rangle$ and one has to assume independently that $\rho\langle H\rangle$ is small in order for the expansion to work. While this requirement will not be investigated any further, one can take it inversely as an additional constraint on the instanton's size: If the constrained instanton formalism is correct for a given $\langle H\rangle$, then the instantons must have small radius such that $\rho\langle H\rangle < \varepsilon << 1$, for a given ε.

There is a complication related to the dynamically generated instanton mass, m_{A}. This mass, much like a mass term for fermions (compare Sec. II.4), couples the instanton in the $SU_{\mathrm{A}}(2)$ representation to the anti-instanton in the $SU_{\mathrm{B}}(2)$ representation. If complicated vacuum configurations for the scalar Higgs-field are considered, then different combinations of instanton and anti-instanton parts can even acquire different effective masses. This topic will be investigated in detail in Sec. III.5.3.

Wang then used the concepts from Ref. [26] to write down a partition function for quasi-zero modes in the presence of a general constraint[50)] and used the explicit form of the solutions $(A_\mu^{\mathrm{con}}, H^{\mathrm{con}})$ to determine the constraint that would be needed to construct them. The advantage of this procedure over the somewhat more direct method by Affleck is that it allows a qualitative investigation of the new pre-exponential factors in the partition function. Although it is possible, a detailed analysis of this pre-exponential contribution will be postponed. In fact an analytic analysis of the upcoming measure corrections is probably not possible in the non-perturbative regime and the reason to give the remaining

49) The ansatz for A_μ^{con} is actually slightly different, compared to Ref. [25, p.7], but this way it is consistent with the introduced conventions.

50) This technique is similar to the concept introduced in Sec. II.9.4 only that now the constraint function does not vanish completely anymore in the final expression, as one is not expanding around a real extremum of the functional anymore.

results of this section is to provide a starting point for a detailed nummerical investigation of the important contributions.

The constrained instanton also delivers a new contribution to the exponential part of the path integral measure, which can be found by inserting the classical solutions Eq. (2.9.55) and (2.9.56) into (2.9.49)[51)]:

$$m_{\text{higgs}}(\rho, \lambda) = e^{-S_{\text{E}}(A_\mu^{\text{con}}, H^{\text{con}})} \tag{2.9.57}$$

$$= \exp\left[-2\pi^2 \lambda^{-2}(\rho\langle H\rangle)^2 - O\left(\lambda^{-2}(\rho\langle H\rangle)^4 \log(\rho\langle H\rangle)\right)\right] . \tag{2.9.58}$$

This relation indicates that the integration over all sizes of the constrained instanton is damped by an exponential factor for large instantons. This damping resolves a general divergence problem in the ρ-integral of the free instanton solution.

Finally it should be mentioned, that, to first order, the measure from the kinetic part of constrained instantons coincides with the original measure for instantons

$$-S_{\text{E}}(A^{\text{con}}) = -\int d^4x_{\text{E}} \mathcal{L}^{\text{kin}}(A^{\text{con}}) = -\int d^4x_{\text{E}} (A^{\text{con}})^{\mu\nu} A^{\text{con}}_{\mu\nu} = \frac{8\pi^2}{g^2}|n| + O((\rho^4\langle H\rangle)^4) . \tag{2.9.59}$$

and therefore the only direct change in the partition function is the inclusion of the discussed new Higgs measure, m_{higgs}.

II.9.8.1 Approximate fermionic zero-modes

In Sec. II.9.6 the fermionic zero-modes in the presence of a free instanton have been discussed and presented. Knowing that the free instantons have to be replaced with the constrained solution from Eq. (2.9.55), one has to rederive the corresponding fermionic zero-mode in the constrained instanton context. For the derivation the constrained instanton is assumed to be localised at the origin, $x_0 = 0$. Any other position can be gained, by a simple shift (compare Sec. II.9.6). Unlike before, the derivation is much harder, as the equations to be solved are now coupled:

$$0 = -\mathrm{i}\bar{q}^\mu \left[\partial_\mu + A_\mu^{\text{con}}\right]\psi_{\text{B}} + \widetilde{g}_{\text{H}} H^{\text{con}} \psi_{\text{A}} , \tag{2.9.60}$$

$$0 = \mathrm{i} q^\mu \partial_\mu \psi_{\text{A}} + \widetilde{g}_{\text{H}} H^{\text{con}} \psi_{\text{B}} . \tag{2.9.61}$$

The factor of $\widetilde{g}_{\text{H}} = g_{\text{H}}\lambda^{-1/2}$ is the 'modified' coupling of the Higgs field to the fermion fields. In it the quadric Higgs coupling, λ, shows up, as the Higgs Lagrangian has been represented in a rescaled fashion (compare Eq. (2.9.49), just like the modified gauge coupling, $\widetilde{g}$, to the Higgs field). Before proceding any further it is useful to add in another ingredient, which will be important in the later model. It was already mentioned that the instanton concept, as presented here, relies on a Euclidean space-time, $SO(4)$, and an

[51)]The explicit derivation of these factors is given in Ref. [21, p.62-65].

additional $SU_{\rm I}(2)$ symmetry, which was taken to be the 'isospin' symmetry group earlier. As some parts of the solutions to the equations Eq. (2.9.60) and (2.9.61) will depend on particular isospin components, it is necessary to give the explicit iso-spinor representations of all involved fields. The Higgs-field will be replaced by a (2×2) iso-spinor field, $H\hat{=}\Omega_\alpha q^{\alpha\, ab}$, the fermionic zero-modes will have two iso-spinor components and the gauge field will remain as it was - the $SU_{\rm I}(2)$ gauge field in Euclidean space-time.
As the VEV of interest, $\langle\Omega\rangle = \mathrm{diag}(\omega_1, \omega_2)$, will have different diagonal entries and is therefore allowed to explicitly break the isospin symmetry. This leads to different fermionic zero-modes in the iso-spinor components, as will be shown in a moment. In addition the different VEV components imply gauge field components with different masses for the long distance regime, $\rho >> |x|$ (compare Eq. (2.9.54)), but the exponential decay of the gauge field solution [$\sim K_2(m_{\rm A}\,|x|)$] will help to circumvent dealing with this complication.
There is not much hope to solve the coupled differential equations Eq. (2.9.60) and (2.9.61) analyticaly, as the fields $A_\mu^{\rm con}$ and $H^{\rm con}$ have a nontrivial space dependence and so, in order to make any progress one can at least give approximate solutions for important limits. The detailed derivation of these limits was done by Espinosa in 1989 (Ref. [31]) and by Kastening two years later (Ref. [32]). In their analysis they assumed to have a left-handed 2-component spinor field, $\psi_{\rm B}$, and a right-handed 1-component field, $\psi_{\rm A}$. The situation in this context is slightly different, as both left- and right-handed fields are assumed to be 2-component spinors. Nevertheless most steps of the derivation from Espinosa can be used in this derivation as well, since the diagonal choice of the Higgs field VEV will decouple the different iso-spinor components from one another. Take as a suitable iso-spinor approach:

$$\langle H\rangle = \begin{pmatrix} h_{11} & 0 \\ 0 & h_{22}\end{pmatrix} \quad , \quad \psi_{\rm B} = \begin{pmatrix} \varphi^1\, q^\alpha x_\alpha \\ \varphi^2\, q^\beta x_\beta \end{pmatrix} \quad , \quad \psi_{\rm A} = \begin{pmatrix} \omega^1_{(x^2)}\, I \\ \omega^2_{(x^2)}\, I \end{pmatrix}, \tag{2.9.62}$$

where $h_j := h_{jj}$ are the just numbers, $\varphi^j := \varphi^j(x^2)$ and $\omega^j := \omega^j(x^2)$ are scalar functions and q^α are the quaternion symbols[52]. The unit matrix I in iso-spinor space was included for completeness in the $\psi_{\rm A}$ field. The gauge field, $A_\mu^{\rm con}$, will have a more complicated structure. In the short distance ($|x| << \rho$) the isospin symmetry will remain approximately valid and thus $A_\mu^{\rm con} = (A_\mu^{\rm con})^\alpha \tau^\alpha/2$ will have the appearance of an ordinary $SU_{\rm I}(2)$ gauge field. In the other limit ($|x| >> \rho$) the VEV of the scalar field will lead to a symmetry breaking, which induces different the gauge field components.
Now the first approximate solution of the coupled differential equations can be found in the small distance limit ($|x| << \rho$). Combining Eq. (2.9.53) and (2.9.52) gives the approximate Higgs and gauge field in

[52] These quaternion symbols have their origin in the euclidianisation of Minkowski space and are unrelated to the iso-spinor space.

this limit:

$$A_\mu^{\mathrm{con}\,\alpha}(x) \approx \frac{-\rho^2}{x^2+\rho^2}\frac{\bar{q}_{\mu\nu}x^\nu}{x^2}\frac{\tau^\alpha}{2} \quad , \quad H^{\mathrm{con}}(x) \approx \left(\frac{x^2}{x^2+\rho^2}\right)^{1/2} \underbrace{\begin{pmatrix} h_{11} & 0 \\ 0 & h_{22} \end{pmatrix}}_{=\langle H\rangle} \qquad \text{for} \quad |x| << \rho\,. \tag{2.9.63}$$

In the expression for A_μ^{con} the $SU_{\mathrm{I}}(2)$ generator $\tau^\alpha/2$ is given explicitly to show the exact isospin structure. In earlier derivations this was not necessary as this symmetry was explicitly integrated out leading to a multiplicative factor in the functional deteminant of $\sqrt{\det(U_{ij})}$ in Eq. (2.9.31). The isospin matrix structure of the Higgs field is generated by $\langle H\rangle$. Comparing the approximate solution for the gauge field, A_μ^{con}, with the free instanton solution (Eq. (2.9.7)) one can see that the constrained instanton indeed gives the free instanton solution in a small area around the instanton location[53]. The expression for the Higgs field shows that its influence is deminuished by factors of $|x|/\rho$ [54]. Therefore the Higgs field will be set to zero in Eq. (2.9.60), which then turns into the differential equation for the fermionic zero-modes of normal instantons (Eq. (2.9.38)). This gives the approximate fermionic zero-mode for ψ_{B} in the small distance limit (in singular gauge):

$$0 = -\mathrm{i}\bar{q}^\mu\left[\partial_\mu + A_\mu^{\mathrm{con}}\right]\psi_{\mathrm{B}} + O\left(\frac{x}{\rho}\right)\psi_{\mathrm{A}}\,, \tag{2.9.65}$$

$$\Rightarrow \psi_{\mathrm{B}}^j(x) \approx \frac{\rho}{\pi}\frac{q^\mu x_\mu}{\left[x^2(x^2+\rho^2)^3\right]^{1/2}} \qquad \text{for} \qquad |x| << \rho\,. \tag{2.9.66}$$

Here the ρ-dependence has been chosen to normalise the mode ($\langle\psi_{\mathrm{B}}^j|\psi_{\mathrm{B}}^j\rangle = 1$), according to the earlier results (compare Sec. II.9.6). In powers of energy the mode's dimension is $[\psi_{\mathrm{B}}^j] = 2$. The concept of 'energy power counting' will be discussed in details in Sec. III.4.2 and can be reviewed in the literature Ref. [5, 90-92]. Here the energy dimension is given, as it will be needed for a dimensional analysis in Sec. II.9.9.3. In the approximate solution the superscript j labels the iso-spinor comonent. In this limit both components coincide, which is expected from the earlier derivation for normal instantons. Using

[53] As this was one of the original puproses of the constrained instantons this finding is actually mandatory.

[54] To see this use the Taylor expansion of the square root:

$$\left(\frac{x^2}{x^2+\rho^2}\right)^{1/2} = \frac{|x|}{\rho}\left[1 - \frac{x^2}{2\rho^2} + O\left(\frac{x^4}{\rho^4}\right)\right] \tag{2.9.64}$$

the approximate result for ψ_{B} and H^{con} in Eq. (2.9.61) leads to the approximation for ψ_{A}:

$$0 = \mathrm{i}q^{\mu}\partial_{\mu}\psi_{\mathrm{A}} + \widetilde{g}_{\mathrm{H}}H^{\mathrm{con}}\psi_{\mathrm{B}}\,, \tag{2.9.67}$$

$$0 = q^{\mu}\left(\mathrm{i}\partial_{\mu}\psi^{j}_{\mathrm{A}} + \widetilde{g}_{\mathrm{H}}h_{j}\frac{\rho}{\pi}\frac{x_{\mu}}{(x^{2}+\rho^{2})^{2}}\right), \tag{2.9.68}$$

$$\Rightarrow \psi^{j}_{\mathrm{A}} \approx -\mathrm{i}\frac{(\rho M_{j})}{2\pi}\frac{1}{x^{2}+\rho^{2}}\,, \tag{2.9.69}$$

$$\approx -\mathrm{i}\frac{(\rho M_{j})}{2\pi}\frac{1}{\rho^{2}}\left[1-\frac{x^{2}}{\rho^{2}}+O\left(\frac{x^{4}}{\rho^{4}}\right)\right] \qquad \text{for} \qquad |x| << \rho\,. \tag{2.9.70}$$

Here the VEV of the Higgs field has been replaced with the fermionic mass: $M_{j} = \widetilde{g}_{\mathrm{H}}h_{j}$. So the solutions for ψ_{B} and ψ_{A} are approximately the same as the earlier derived free fermionic zero-mode for normal instantons in the singular gauge (although earlier one had $\psi_{\mathrm{A}} = 0$). One sees that for $|x| << \rho$ the solution of Eq. (2.9.60) will be dominated by A^{con}_{μ} and the influence of H^{con} is neglectible. This argument does not hold for the second equation (Eq. (2.9.61)), as A^{con}_{μ} does not contribute here. This gives the slightly different behaviour of ψ_{A}. Another observation concerns the VEV of the Higgs field. If $\langle H\rangle \to 0$ one recovers the exact zero-modes from the free instanton sector, as it should be. In fact this could already be deduced from the defining equations (Eq. (2.9.60) and (2.9.61)), so here it only gives a small consistency check.

For the other limiting case - far away from the instanton - one can also give an analytic expression for the approximate zero-modes, but this time a little more computation is necessary. First, the limiting Higgs and gauge fields are:

$$A^{\mathrm{con}}_{j\mu}(x) \approx -\frac{\rho^{2}m^{j\,2}_{\mathrm{A}}}{2}\frac{K_{2}(m^{j}_{\mathrm{A}}|x|)}{x^{2}}\bar{q}_{\mu\nu}x^{\nu} \quad , \quad H^{\mathrm{con}}_{ij}(x) \approx \left(1-\frac{\rho^{2}m^{ij}_{\mathrm{H}}}{2}\frac{K_{1}(m^{ij}_{\mathrm{H}}|x|)}{|x|}\right)h^{ij} \quad \text{for} \quad |x| >> \rho\,. \tag{2.9.71}$$

Both fields are given here in components for notational reasons and, as earlier, $m^{ij}_{\mathrm{H}} = h^{ij}$. The instanton mass, m^{j}_{A}, and the general instanton/anti-instanton structure turns out to be a bit more difficult in the low energy region. For the present derivation the exact instanton masses are not important, as long as they are non-zero. Therefore details on the low energy instanton structure and masses will be dealt with in the explicit model calculations (Sec. III.5.3) but for now it is sufficient to assume $m^{j}_{\mathrm{A}} \neq 0$.

To derive the corresponding fermionic zero-modes note that the Bessel function, for large arguments has the limes: $K_{\nu}(x) \to_{x\to\infty} \left(\frac{\pi}{2x}\right)^{1/2} e^{-x}$. Using this, the limit ($|x| >> \rho$) for the gauge and Higgs fields simplifies to: $A^{\mathrm{con}}_{j\mu}(x) \to 0$ and $H^{\mathrm{con}}_{ij}(x) \to h^{ij}$. So one finds that far away from the instanton the solution to the differential equations (Eq. (2.9.60) and (2.9.61)) is governed by the VEV of the Higgs field, while the gauge field contribution vanishes[55)]. Using this simplification the coupled differential equations turn into:

$$0 = -\mathrm{i}\bar{q}^{\mu}\partial_{\mu}\psi_{\mathrm{B}} + \widetilde{g}_{\mathrm{H}}\langle H\rangle\psi_{\mathrm{A}}\,, \tag{2.9.72}$$
$$0 = \mathrm{i}q^{\mu}\partial_{\mu}\,\psi_{\mathrm{A}} + \widetilde{g}_{\mathrm{H}}\langle H\rangle\psi_{\mathrm{B}}\,. \tag{2.9.73}$$

The derivative in the first equation (Eq. (2.9.72)) can be evaluated with the 'Ansatz' of Eq. (2.9.62). In spinor component notation one finds:

$$0 = -\mathrm{i}\bar{q}^{\mu}q^{\alpha}\Big[\delta_{\mu\alpha}\varphi_j(x^2) + x_{\alpha}(\partial_{\mu}\varphi^j(x^2))\Big] + gh^j\omega^j(x^2)I\,, \tag{2.9.74}$$
$$0 = -\mathrm{i}\Big[4\varphi^j I + 2\bar{q}^{\mu}q^{\alpha}x_{\mu}x_{\alpha}\frac{\partial\varphi^j}{\partial(x^2)}\Big] + gh^j\omega^j I\,, \tag{2.9.75}$$
$$\omega^j = \frac{\mathrm{i}}{gh^j}\Big[4\varphi^j + 2x^2(\varphi^j)'\Big]\,, \tag{2.9.76}$$
$$\Rightarrow (\omega^j)' = \frac{2\mathrm{i}}{gh^j}\Big[3(\varphi^j)' + x^2(\varphi^j)''\Big]\,. \tag{2.9.77}$$

To get the correct spinor structure remember that $h^j = h^{jj}$ has a matrix character. In the second and third line the identity for quaternion symbols has been used: $\sum_{\mu\nu}\bar{q}^{\mu}q^{\nu} = \delta_{\mu\nu}I$. The prime in the third and forth line denotes the derivative $a' = (\partial a)/(\partial(x^2))$. This result can now be combined with Eq. (2.9.73), leading to the differential equation:

$$0 = \mathrm{i}q^{\mu}\partial_{\mu}\omega^j + \widetilde{g}_{\mathrm{H}}h^j\varphi^j q^{\mu}x_{\mu}\,, \tag{2.9.78}$$
$$0 = q^{\mu}x_{\mu}\Big[2\mathrm{i}(\omega^j)' + \widetilde{g}_{\mathrm{H}}h^j\varphi^j\Big]\,, \tag{2.9.79}$$
$$\Rightarrow 0 = x^2(\varphi^j)'' + 3(\varphi^j)' - \frac{(gh^j)^2}{4}\varphi^j\,. \tag{2.9.80}$$

In the second line the derivative of $\omega^j(x^2)$ was taken, just as the φ^j derivative before and, to arrive at the last line, Eq. (2.9.77) was inserted. At this point the work by Espinosa Ref. [31, 328] can be adopted. In addition, his solution for the remaining ω^j field also applies, since all involved differential equations are exactly the same, regardless of the different iso-spinor structure in Espinosa's derivation. Including the correct boundary conditions, he found the solutions to these equations to be:

$$\varphi^j = \frac{\rho M_j^2}{2\pi}\frac{K_2(M_j\,|x|)}{x^2} \quad\Rightarrow\quad \psi^j_{\mathrm{B}}(x) \approx \frac{\rho M_j^2}{2\pi}\frac{K_2(M_j\,|x|)}{x^2}q_{\mu}x^{\mu} \quad \text{for} \quad |x| >> \rho\,, \tag{2.9.81}$$
$$\omega^j = \frac{-\mathrm{i}\rho M_j^2}{2\pi}\frac{K_1(M_j\,|x|)}{|x|} \quad\Rightarrow\quad \psi^j_{\mathrm{A}}(x) \approx -\frac{\mathrm{i}\rho M_j^2}{2\pi}\frac{K_1(M_j\,|x|)}{|x|}\,I \quad \text{for} \quad |x| >> \rho\,. \tag{2.9.82}$$

[55)] The vanishing of the gauge field is in fact necessary, since it has to approach a 'pure gauge' transformation at infinity (compare Eq. (2.8.6)).

Here the normalisation leads again to the same fermion field units, $[\psi^j_{\mathrm{A/B}}] = 2$ (in powers of energy). This time the fermionic masses, $M_j = \widetilde{g}_{\mathrm{H}} h_j$, have been introduced. The dependence on the different components of the scalar VEV of the fermionic zero-modes is explicitly visible in this limit. The appearance of the fermion mass couples the two subgroups $SU_{\mathrm{A}}(2)$ and $SU_{\mathrm{B}}(2)$ of Euclidean space-time implicitly together (compare Eq. (2.4.12) for the Minkowski version of this effect). Therefore it is more convenient to write the large distance zero-modes in a doubled dimensional 'Dirac-spinor'-like structure, which incorporates the contribution from the instanton and the anti-instanton[56]. This combination gives a mode, which will correspond to a left-handed field, if translated back into Minkowski space (indicated already by the subscript):

$$\psi^j_{\mathrm{L}}(x) \sim -\mathrm{i}\frac{\rho M_j^2}{2\pi}\left[\frac{K_2(M_j\,|x|)}{x^2}(\gamma_{\mathrm{E}\mu}x^\mu) + \frac{K_1(M_j\,|x|)}{|x|}\mathbb{1}\right] \qquad \text{for} \quad |x| >> \rho\,. \tag{2.9.83}$$

Here γ^μ_{E}[57] and $\mathbb{1}$ only act in Euclidean space-time and the superscript j labels the component in iso-spinor space (the different iso-spinor components only differ by their masses M_j). In later calculations the approximate Fourier transforms of these expressions will be useful. These can be gained by the means of the following integral identities:

$$\int \mathrm{d}^4x\, e^{\mathrm{i}px} f(x^2) = \frac{4\pi^2}{p}\int_0^\infty \mathrm{d}r\, J_1(pr) r^2 f(r^2)\,, \tag{2.9.84}$$

$$\int \mathrm{d}^4x\, e^{\mathrm{i}px} x_\mu f(x^2) = -\frac{4\pi^2\mathrm{i}p_\mu}{p^2}\int_0^\infty \mathrm{d}r\, J_2(pr) r^3 f(r^2)\,, \tag{2.9.85}$$

$$\int_R^\infty \mathrm{d}r\, rJ_n(pr)K_n(Mr) = \int_0^\infty \mathrm{d}r\, rJ_n(pr)K_n(Mr) - \int_0^R \mathrm{d}r\, rJ_n(pr)K_n(Mr) \tag{2.9.86}$$

$$= \frac{p^n}{M^n(p^2+M^2)} - \int_0^R \mathrm{d}r\, rJ_n(pr)K_n(Mr)\,. \tag{2.9.87}$$

$J_n(x)$ are the Bessel functions of first kind and $K_n(x)$, as before, are modified Bessel functions of second kind (both of order n). As the functions $\psi_{\mathrm{A}}(x)$ and $\psi_{\mathrm{B}}(x)$ are not known over the entire x-domain, it is not possible to compute the exact Fourier transforms, but only estimates for the low- and high momentum region can be given. The regions can be identified by analysing the shape of the space dependent functions.

[56] For the instanton basically the zero-modes are interchanged ($\psi^j_{\mathrm{B}} \leftrightarrow -\psi^j_A$) and q_μ turns into $\bar{q}_\mu$.
[57] For a definition of γ^μ_{E} compare 2.9.3.

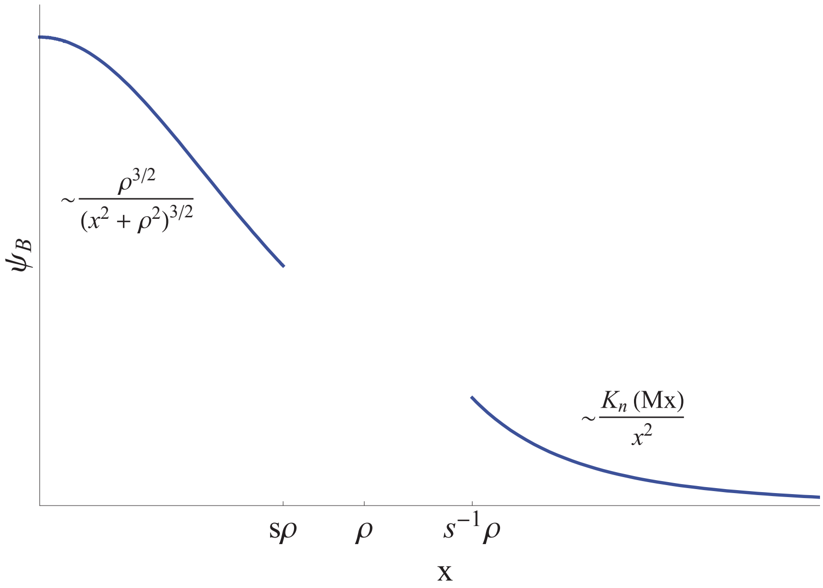


Figure 2.9.1: Schematic $|x|$ dependence of the approximate fermionic zero-mode ψ_{B}.

For low momenta (corresponding to low frequencies in the Fourier transformation) the small values of $\psi_{\mathrm{B}}(x)$ are supressed and the large x-approximation $\psi_{\mathrm{B}} \sim K_2(M_{\mathrm{N}}\,|x|)$ dominates the transformation. The contrary is true for high momenta. There the 'sampling rate' is high enough to capture the contributions from the (relatively small) $|x| \leq s\rho << \rho$ region. In the integral this region (if not supressed) gives the main contribution to the total result, as the function is peaked around $x = 0$.
The only scale to distinguish both regions is ρ and so the Fourier transform will be split up according to the inverse of this scale, ρ^{-1}[58]. A similar argument holds for ψ_{A}. For the low momentum region the

[58]The inverse has to be taken for dimensional reasons.

integral identities Eq. (2.9.84)-(2.9.87) give (in index notation):

$$\hat{\psi}^j_{\mathrm{B}}(p) = -2\pi\rho\left[\frac{\mathrm{i}q_\mu p^\mu}{p^2+M_j^2} - \frac{M_j^2}{p^2}\int_0^{s^{-1}\rho}\mathrm{d}r\, rJ_n(pr)K_2(M_jr)\right] \approx -2\pi\rho\frac{\mathrm{i}q_\mu p^\mu}{p^2+M_j^2} \qquad \text{for} \quad |p| << \rho^{-1}\,, \tag{2.9.88}$$

$$\hat{\psi}^j_{\mathrm{A}}(p) = -2\pi\rho\left[\frac{M_j}{p^2+M_j^2} - \frac{M_j}{|p|}\int_0^{s^{-1}\rho}\mathrm{d}r\, rJ_n(pr)K_1(M_jr)\right] \approx -2\pi\rho\frac{M_j}{p^2+M_j^2} \qquad \text{for} \quad |p| << \rho^{-1}\,. \tag{2.9.89}$$

And the corresponding terms for the high momentum part Eq. (2.9.66) and (2.9.70) are (although not nicely analytic):

$$\hat{\psi}_{\mathrm{B}}(p) = 4\pi\rho\frac{\mathrm{i}q_\mu p^\mu}{p^2}\int_0^\infty \mathrm{d}rJ_2(pr)\frac{r^3\theta(s\rho-r)}{[r^2(r^2+\rho^2)^3]^{1/2}} \qquad \text{for} \quad |p| >> \rho^{-1}\,, \tag{2.9.90}$$

$$\hat{\psi}^j_{\mathrm{A}}(p) = -2\pi\mathrm{i}\rho\frac{M_j}{p}\int_0^\infty \mathrm{d}rJ_1(pr)\frac{r^2\theta(s\rho-r)}{r^2+\rho^2} \qquad \text{for} \quad |p| >> \rho^{-1}\,. \tag{2.9.91}$$

Note that the energy dimensions of the Fourier transformed fields fit nicely to the original ones. In momentum-space the zero-mode energy dimension is $[\hat{\psi}^j_{\mathrm{A/B}}] = -2$, compared to the earlier $[\psi^j_{\mathrm{A/B}}] = 2$. In the above equation the usual Heaviside function, $\theta(x)$, was used. Of course, the integrands are only correct, if the Heaviside function is included. Analogously the approximation in Eq. (2.9.88) and (2.9.89) are only useful, if the contribution from the neglected integrals is small. Knowing this, one can approximate low energy phenomena by using the simplified low energy zero-modes and simultaneously introducing an effective cutoff, $\Lambda_\rho \sim s^{-1}\rho$. This cutoff will essentially label the minimal distance down to which the low energy approximation is useful and from where on intermediate approximations are needed. Going to extremely high energies the free instanton zero-modes can be used to approximate physical contributions. For this regime the cutoff will label the inverse ($\Lambda_\rho \sim s\rho$), which is a largest distance, up to which the approximation makes sense.

As the final step, the low energy zero-modes can be retranslated to Minkowski space, just giving the free fermion propagator:

$$\hat{\psi}^j_{\mathrm{L}}(p) \approx 2\pi\rho\frac{(\not p + M_j)}{p^2 - M_j^2}P_{\mathrm{L}}{}^{59)}\,. \tag{2.9.92}$$

If the derivation is done for the instanton instead, then one arrives at the right-handed fermion propagator. The approximate low energy momentum space solutions (Minkowski space) for the gauge field and the

[59] Here the West coast metric was chosen, as the original derivation relied on this convention.

shifted Higgs field ($\widetilde{H} = \langle H \rangle - H$)[60] can be calculated in a similar fashion, giving:

$$\rho^{-1}\hat{A}^{\rm con}_\mu(p) \approx \frac{4\pi^2}{p^2 - m_{\rm A}^2}\sigma_{\mu\nu}(\rho p^\nu) \qquad ; \qquad \rho^{-1}\hat{\widetilde{H}}(p) \approx -\frac{2\pi}{p^2 - m_{\rm H}^2}(\rho m_{\rm H}) \qquad \text{for} \quad |p| << \rho^{-1} \,. \tag{2.9.93}$$

In the equation for $\hat{A}^{\rm con}_\mu$ the object $\bar{\sigma}_{\mu\nu} \sim [\gamma_\mu, \gamma_\nu]$ is the Minkowski space equivalent to the $\bar{q}_{\mu\nu}$ from Euclidean space. Both equations have been rescaled by a factor of ρ^{-1}, so that their energy dimension is symmetryic[61] under Fourier transformation in this representation.
The above results are actually quite pleasant. They mean that all fields behave as free particles in the low energy regime and do not see the constrained instanton effects. If one goes to higher energies the situation is more complicated and the equations Eq. (2.9.90) and (2.9.91) have to be solved there, explicitly encoding the influence of the instantons.

II.9.9 Zero-mode contributions

The last contribution to the partition function, that needs to be calculated is the determinant over fermionic pseudo zero-modes $\det_0(i\not{D} + g_\Omega \Omega)$. Here $\Omega = \Omega_\alpha q_{\rm I}^\alpha$ is a matrix in iso-spinor space (and a scalar otherwise). The subscript I at the quaternion symbol shall only label that this quaternion symbol acts in iso-spinor space. The reason for this extra labeling is that there are also the quaternion symbols of euclideanised space-time. These appear in the definition of the gauge fields and pseudo zero-modes from the previous section (e.g. Eq. (2.9.81) or (2.9.83)). It is important to keep in mind that the fermionic pseudo zero-modes carry a vector index in iso-spinor space and represent a matrix in $SO(4)$ space-time. The problem, indicated in the previous section, is that these pseudo zero-modes are only known analytically in two extreme regimes. So there is no hope to produce an exact analytic expression for the determinant over these field solutions. This section will give an estimate of the determinant under the assumption that the contribution from the non analytic regime is neglectible. Whether or not this simplification is justified is a topic to further numerical investigations.
In Sec. II.9.7 the partition function of a generic model was determined and given for the case of normal instantons in Eq. (2.9.48). In the consecutive section the constrained instanton was discussed and this led to a very complicated correction to the partition function (Eq. (2.9.57)). Taking this together the partition function yields another correction to the measure:

$$Z_{\rm gen} = \int d^4x_0 d\rho \; m_1(\rho) m_2(\rho, \mu_0) m_{\rm higgs}(\rho, \lambda) \times \det_0(i\not{D} + g_\Omega \Omega) \,. \tag{2.9.94}$$

Here it is important that the determinant over zero-modes contains the classical field configurations $A^{\rm con}_\mu$, $\Omega^{\rm con}$ from the previous sectioin as well as a fluctuation around these, indicated by a_μ and ω.

[60]This shifted field just represents the dynamical contribution of the Higgs field.
[61]Symmetric in this situation means that if the original field has dimension $[F] = 2$, then the transformed field has $[\hat{F}] = -2$.

If the operator $O = (\mathrm{i}\not{D} + g_\Omega \Omega)$ is diagonalisable, its determinant is given by the product of all eigenvalues. For the present discussion it will be assumed that this decomposition is possible[62]. The 'sub-determinant' $\det_0(\cdot)$ only incorporates the zero eigenvalues of O and therefore it vanishes if the operator has exact zero-modes: $O|N_j\rangle = 0|N_j\rangle$. Now, as the scalar field has a non vanishing expectation value, $\langle\Omega\rangle \neq 0$, the situation is more complicated and only pseudo zero-modes exist, as discussed in the previous section. If $|N_j\rangle$ is such a pseudo zero-mode where j incorporates all indices and variables that the mode depends on, then one needs to calculate the eigenvalues of O acting on it:

$$\det_0(O) = \prod_j \langle N_j| O |N_j\rangle = \det_{\mathrm{rem}} \int \mathrm{d}^4 x \left[\langle N((x-x_0),\rho)| O(x) |N((x-x_0),\rho)\rangle\right] . \tag{2.9.95}$$

Here the constrained instanton is not at the origin anymore, but at position, x_0 (leading to a space dependence of $(x - x_0)$ for the pseudo zero-modes). The reason for the explicit inclusion of x_0 is that the operator, $O = O(x)$, has a different space dependence. The $\det_{\mathrm{rem}}$ refers to any remaining symmetry space of the operator, O (this will later be the isospin space and so it will be called $\det_{SU(2)}$ from now on). In the preceding section it was found that the pseudo zero-modes show a very different behaviour in the two regions: $(x - x_0) << \rho$ and $(x - x_0) >> \rho$. To better deal with this complication the integral in Eq. (2.9.95) can be rewritten by substituting $\tilde{x} = (x - x_0)$:

$$\det_0(O) = \det_{SU(2)} \int \mathrm{d}^4\tilde{x} \left[\langle N(\tilde{x},\rho)| O(\tilde{x}+x_0) |N(\tilde{x},\rho)\rangle\right] , \tag{2.9.96}$$

$$= \det_{SU(2)} \left\{ \int_0^{s\rho} \mathrm{d}^4\tilde{x} \left[\langle N(\tilde{x},\rho)| O(\tilde{x}+x_0) |N(\tilde{x},\rho)\rangle\right] \right. \tag{2.9.97}$$

$$+ \int_{s\rho}^{s^{-1}\rho} \mathrm{d}^4\tilde{x} \left[\langle N(\tilde{x},\rho)| O(\tilde{x}+x_0) |N(\tilde{x},\rho)\rangle\right] \tag{2.9.98}$$

$$\left. + \int_{s^{-1}\rho}^{\infty} \mathrm{d}^4\tilde{x} \left[\langle N(\tilde{x},\rho)| O(\tilde{x}+x_0) |N(\tilde{x},\rho)\rangle\right] \right\} . \tag{2.9.99}$$

The integral boundaries are meant to be boundaries of the radial part, if the integration is performed in spherical coordinates. For a suitable scaling parameter $s \in [0, 1]$ the above equation seperates the integral into a short range part (Eq. (2.9.97)) and a long range contribution (Eq. (2.9.99)). The integral in between from $s\rho$ to $s^{-1}\rho$ is not accessible to analytic considerations, as the pseudo zero-modes are not known in this regime. In order to use the approximate zero-modes from the previous section (Eq. (2.9.66), (2.9.70) and Eq. (2.9.81), (2.9.82)) one needs $s << 1$. This unfortunately means that the non-analytic contribution (Eq. (2.9.98)) becomes larger. This work will not deal with an exact analysis of this issue, but it will simply be assumed that a suitable value of s can be found, such that the approximate zero-modes from the previous section are still usable in the analytic integrals and at the same time the contribution from Eq. (2.9.98) is neglectible (compared to the other integrals). Nevertheless it should be

[62]For all configurations of the scalar field Ω, which are interesting in physical situations, this operator will be diagonalisable.

mentioned that the separation of these integrals is neither trivial nor unimportant and that it should be subject to future investigations. Using the simplification the zero-mode determinant becomes:

$$\det{}_0(O) \approx \det_{SU(2)} \left\{ \int_0^{s\rho} \mathrm{d}^4\tilde{x}\, [\langle N(\tilde{x},\rho)|\, O(\tilde{x}+x_0)\, |N(\tilde{x},\rho)\rangle] \right. \tag{2.9.100}$$

$$\left. + \int_{s^{-1}\rho}^{\infty} \mathrm{d}^4\tilde{x}\, [\langle N(\tilde{x},\rho)|\, O(\tilde{x}+x_0)\, |N(\tilde{x},\rho)\rangle] \right\}^{63)} . \tag{2.9.101}$$

II.9.9.1 High energy contribution

The first line (Eq. (2.9.100)) can now be expressed using the short range approximation from the previous section (Eq. (2.9.66)), which is exactly the same as exanding the determinant in terms of the free fermionic zero-modes from Sec. II.9.6 and treating the scalar field, $\Omega(\tilde{x}+x_0)$, as a small perturbation. Taking $|{}^{\mathrm{h}}N_{\tilde{x}}\rangle \equiv |{}^{\mathrm{h}}N(\tilde{x},\rho)\rangle$ as the small distance (or equivalently high energy) approximation of the full pseudo zero-mode, one has

$$\bar{q}_\mu D^\mu_{(\tilde{x}+x_0)}\, |{}^{\mathrm{h}}N_{\tilde{x}}\rangle = \underbrace{\bar{q}_\mu (D^{\mathrm{con}})^\mu_{(\tilde{x}+x_0)}\, |{}^{\mathrm{h}}N_{\tilde{x}}\rangle}_{=0} + \bar{q}_\mu a^\mu_{(\tilde{x}+x_0)}\, |{}^{\mathrm{h}}N_{\tilde{x}}\rangle^{64)} . \tag{2.9.103}$$

Here the gauge field (and with it the covariant derivative, $D_\mu = \partial_\mu + A_\mu$) has been written as the constrained instanton contribution and a fluctuation around it: $A_\mu = [(A^{\mathrm{con}})^c_\mu + a^c_\mu](-\mathrm{i}\tau^c_{\mathrm{I}}/2)^{65)}$. In many applications (and in the later model as well) one can argue, that the fluctuation contribution is neglectible, but as this argument depends on the expansion point of a given model, it will be kept in the following analysis, to give a more general picture. With this the integral can be simplified to:

$$\text{Ⓐ} := \int_0^{s\rho} \mathrm{d}^4\tilde{x}\, [\langle N_{\tilde{x}}|\, O(\tilde{x}+x_0)\, |N_{\tilde{x}}\rangle] = \int_0^{s\rho} \mathrm{d}^4\tilde{x} \Big[\langle {}^{\mathrm{h}}N_{\tilde{x}}|\, \{\mathrm{i}\bar{q}_\mu D^\mu_{(\tilde{x}+x_0)} + g_\Omega \Omega(\tilde{x}+x_0)\}\, |{}^{\mathrm{h}}N_{\tilde{x}}\rangle\Big] \tag{2.9.104}$$

$$= \int_0^{s\rho} \mathrm{d}^4\tilde{x}\, \langle {}^{\mathrm{h}}N_{\tilde{x}}| \Big[\mathrm{i}\bar{q}_\mu a^\mu_{(\tilde{x}+x_0)} + g_\Omega \Omega(\tilde{x}+x_0)\Big] |{}^{\mathrm{h}}N_{\tilde{x}}\rangle . \tag{2.9.105}$$

At this point it becomes crucial to keep track of the different spaces the various quaternion symbols live in. The high energy pseudo zero-mode, $|{}^{\mathrm{h}}N_{\tilde{x}}\rangle$, is identical for all iso-spinor components and therefore it

63) The ignored contribution is bound by the expression:

$$\int_{s\rho}^{s^{-1}\rho} \mathrm{d}^4\tilde{x} \det_{SU(2)} \Big[O(\tilde{x},x_0)\Big] \leq \Big(\rho s^{-1} - \rho s\Big) \max_{\tilde{x}} \Big[\det_{SU(2)} (O(\tilde{x},x_0))\Big] \tag{2.9.102}$$

64) The different arguments of the zero-mode and the derivative operator corresponds to the situation of Sec. II.9.6, only with a constant shift of x_0.

65) The strange appearance of the $SU_{\mathrm{I}}(2)$ generator is due to the mathematical convention for gauge field generators (see Sec. II.9.1.1).

commutes with the iso-spinor matrix, q_{I}^{α}, from the scalar field, Ω. But as the zero-mode has a matrix structure in Euclidean space-time $|^{\mathrm{h}}N_{\tilde{x}}\rangle \sim q^{\nu}x_{\nu}$, it does not commute with the gauge field, $\bar{q}_{\mu}a^{\mu}$. Taking the explicit form of the fermionic zero-modes in singular gauge from Eq. (2.9.39) the integral can be rewritten to:

$$Ⓐ = \int_0^{s\rho} \mathrm{d}^4\tilde{x} \left[g_{\Omega} \left| \frac{\rho}{\pi\left[\tilde{x}^2+\rho^2\right]^{3/2}} \right|^2 \Omega_{\alpha}(\tilde{x}+x_0)\, q_{\mathrm{I}}^{\alpha} + \langle {}^{\mathrm{h}}N_{\tilde{x}}|\, \mathrm{i}\bar{q}_{\mu}a^{\mu}_{(\tilde{x}+x_0)}|{}^{\mathrm{h}}N_{\tilde{x}}\rangle \right] \tag{2.9.106}$$

$$= \int_0^{s\rho} \mathrm{d}^4\tilde{x} \left[g_{\Omega} \left|{}^{\mathrm{h}}N_{\tilde{x}}\right|^2 \Omega_{\alpha}(\tilde{x}+x_0)\, q^{\alpha} + \langle {}^{\mathrm{h}}N_{\tilde{x}}|\, \mathrm{i}\bar{q}_{\mu}a^{\mu}_{(\tilde{x}+x_0)}|{}^{\mathrm{h}}N_{\tilde{x}}\rangle \right]. \tag{2.9.107}$$

Here it is also possible to rewrite the gauge field part further, but as the later model will focus mainly on the scalar sector, the gauge field contribution will mainly be listed throughout this part to allow futher investigations in the future. To arrive at a consistent expression (for the scalar part) later it is useful to already introduce a slightly different notation. First define

$${}^{\mathrm{h}}G^{ij}(x,s,\rho) := \left|{}^{\mathrm{h}}N_x\right|^2 \theta(s\rho - |x|)\,, \tag{2.9.108}$$

$$|{}^{\mathrm{h}}N_{\tilde{x},s}\rangle := |{}^{\mathrm{h}}N_{\tilde{x}}\rangle\theta(s\rho - |x|)\,. \tag{2.9.109}$$

where $\theta(x)$ is the Heaviside function. The second line is meant as a replacement of the old approximate zero-mode with the one that is enhanced with the Heaviside function (this part is only included, to clear up the notation concerning the gauge field fluctuations). The iso-spinor indices, i, j, are not yet important, as $\left|{}^{\mathrm{h}}N_x\right|^2$ is the same for all entries of the iso-spinor matrix, $q_{\mathrm{I}}^{\alpha\, ij}$. Only in the later discussed low energy regime they will have a distinct meaning. Now the scalar field can be splitt up into two parts, just as the gauge field before. It was mentioned that the scalar field consists of a VEV and a fluctuation $\Omega = \Omega^{\mathrm{con}} + \omega$ and so Eq. (2.9.107) turns into:

$$Ⓐ = g_{\Omega}\, q_{\mathrm{I}}^{\alpha\, ij} \int_0^{\infty} \mathrm{d}^4\tilde{x}\, {}^{\mathrm{h}}G^{ij}(\tilde{x},s,\rho) \left[\Omega^{\mathrm{con}}_{\alpha}(\tilde{x}+x_0) + \omega_{\alpha}(\tilde{x}+x_0)\right]$$
$$+\mathrm{i} \int_0^{\infty} \mathrm{d}^4\tilde{x}\, \langle {}^{\mathrm{h}}N_{\tilde{x},s}|\, \bar{q}_{\mu}a^{\mu}_{(\tilde{x}+x_0)}|{}^{\mathrm{h}}N_{\tilde{x},s}\rangle\,. \tag{2.9.110}$$

The indices of ${}^{\mathrm{h}}G^{ij}$ are *not* to be traced out with $q_{\mathrm{I}}^{\alpha\, ij}$. They only mean that the (ij) entry of $q_{\mathrm{I}}^{\alpha\, ij}$ is multiplied by the corresponding term, ${}^{\mathrm{h}}G^{ij}$. This may seem as a weird notation at the moment, but it will help to identify the meaning of x_0 after the low energy contribution is included in the next section.

II.9.9.2 Low energy contribution

Now the second determinant contribution (line Eq. (2.9.101)) has to be brought to a more explicit form. In this long distance (or low energy) regime the approximate pseudo zero-mode, $|{}^{\mathrm{l}}N_{\tilde{x},j}\rangle \equiv |{}^{\mathrm{l}}N(\tilde{x},\rho,M_j)\rangle$, is given by Eq. (2.9.83) or equivalently Eq. (2.9.92) and solves Eq. (2.9.60) and (2.9.61). This means that the classical gauge field, A^{con}, and the scalar Higgs field, $H^{\mathrm{con}} \hat{=} \Omega^{\mathrm{con}}$, solve the zero-mode equation:

$$\left[\mathrm{i}\bar{q}_\mu (D^{\mathrm{con}})^\mu_{(\tilde{x}+x_0)} + g_\Omega \Omega^{\mathrm{con}}(\tilde{x}+x_0)\right] |^{\mathrm{l}}N_{\tilde{x},j}\rangle = 0\,. \tag{2.9.111}$$

Compared to the previously discussed high energy contribution, the low energy zero-modes do depend on the fermion mass, M_j, as well. The reason for this, as mentioned in Sec. II.9.8.1, is that the scalar field approaches its VEV and thus generates a fermionic mass term in the equations of motion.
Using the separation into a classical part and the fluctuations around it [$\omega = \omega_\alpha q_{\mathrm{I}}^\alpha$ and $a_\mu = a_\mu^\beta(-\mathrm{i}\tau_{\mathrm{I}}^\beta/2)$], as mentioned in the previous section, Eq. (2.9.111) can be used to evaluate the action of O on the zero-mode:

$$\begin{aligned} O(\tilde{x}+x_0)\,|^{\mathrm{l}}N_{\tilde{x},j}\rangle &= \left[\mathrm{i}\bar{q}_\mu (D^{\mathrm{con}})^\mu + g_\Omega \Omega^{\mathrm{con}}\right]_{(\tilde{x}+x_0)} |^{\mathrm{l}}N_{\tilde{x},j}\rangle + \left[\mathrm{i}\bar{q}_\mu a^\mu + g_\Omega \omega\right]_{(\tilde{x}+x_0)} |^{\mathrm{l}}N_{\tilde{x},j}\rangle && (2.9.112)\\ &= \left[\mathrm{i}\bar{q}_\mu a^\mu + g_\Omega \omega\right]_{(\tilde{x}+x_0)} |^{\mathrm{l}}N_{\tilde{x},j}\rangle\,. && (2.9.113) \end{aligned}$$

As before the fluctuation of the gauge field, $\bar{q}_\mu a^\mu$, is listed here for reasons of completeness and its influence has to be determined in later works.
While the analogous expression for the high energy could directly be used to evaluate the integral in the zero-mode determinant (Eq. (2.9.100)) the situation is now a bit more complicated. Earlier, due to the high energies, one could treat the whole influence of the scalar field as a perturbative contribution and use standard expansion techniques for eigenvalue equations. In the low energy regime the eigenfunctions are significantly influenced by the scalar VEV, $\langle\Omega\rangle$, which leads to different fermionic masses for different iso-spinor components. Therefore in Sec. II.9.8.1 the Dirac-like low energy zero-mode was given as (Eq. (2.9.83)):

$$|^{\mathrm{l}}N_{\tilde{x},j}\rangle \sim -\mathrm{i}\frac{\rho M_j^2}{2\pi}\left[\frac{K_2(M_j|x|)}{x^2}(\gamma_{\mathrm{E}\mu}x^\mu) + \frac{K_1(M_j|x|)}{|x|}\mathbb{1}\right]. \tag{2.9.114}$$

where the index j labels the iso-vector components of the mode. The exact structure in the Dirac-like space will only be evaluated after Fourier transforming the above expression into momentum space, where the approximate zero-mode has a much nicer appearance (Eq. (2.9.92)). Now, the following expression can be calculated:

$$\begin{aligned} \langle^{\mathrm{l}}N_{\tilde{x},i}|O(\tilde{x}+x_0)\,|^{\mathrm{l}}N_{\tilde{x},j}\rangle\Big|_{a_\mu=0} &= g_\Omega\langle^{\mathrm{l}}N_{\tilde{x},i}|\,\omega(\tilde{x}+x_0)\,|^{\mathrm{l}}N_{\tilde{x},j}\rangle && (2.9.115)\\ &= \langle^{\mathrm{l}}N_{\tilde{x},i}|^{\mathrm{l}}N_{\tilde{x},j}\rangle\,\omega^\alpha(\tilde{x}+x_0)\,q_{\mathrm{I}}^{\alpha\,ij}\,. && (2.9.116) \end{aligned}$$

Just as before the indices, i, j, are *not* traced out, but it means that the (ij)-element of $q_{\mathrm{I}}^{\alpha\,ij}$ gets the new factor of $\langle^{\mathrm{l}}N|^{\mathrm{l}}N\rangle_{ij}$. Here, the reason for this strange notation can be given. In principle one needs to find the eigenvalues of $\omega^\alpha q_{\rm I}^\alpha$ in iso-spinor space. As not all Pauli matrices are diagonalisable at the same time one needs to specify a basis. The above notation now allows to attribute the correct zero-modes to the corresponding elements of ω and simply taking the normal determinant of a (2×2)-matrix later. The fluctuation gauge field was set to zero, to eleminate it from the equation. It will be added back into the final expression later. Now the low energy part of the weight function can be defined as:

$$ {}^{\rm l}G^{ij}(x,s,\rho) := \langle {}^{\rm l}N_{x,i}|{}^{\rm l}N_{x,j}\rangle\theta(|x|-s^{-1}\rho)\,, \tag{2.9.117} $$

$$ |{}^{\rm l}N_{x,i,s}\rangle := |{}^{\rm l}N_{x,i}\rangle\theta(|x|-s^{-1}\rho)\,. \tag{2.9.118} $$

and, as promised, this part gives different contributions to the different elements of the iso-spinor matrix. In addition one finds that the classical solution, $\Omega^{\rm con}$, does not get a contribution from the low energy regime. The low energy contribution to the scalar field determinant (Eq. (2.9.101)) can be given as:

$$ \text{Ⓑ} := \int_{s^{-1}\rho}^{\infty} {\rm d}^4\tilde{x}\,[\langle N(\tilde{x},\rho)|\,O(\tilde{x}+x_0)\,|N(\tilde{x},\rho)\rangle] \tag{2.9.119} $$

$$ = \int_0^\infty {\rm d}^4\tilde{x}\,\Big[g_\Omega\big(q_{\rm I}^{\alpha\,ij}\big)\,{}^{\rm l}G^{ij}(\tilde{x},s,\rho)\omega_\alpha(\tilde{x}+x_0) + {\rm i}\langle {}^{\rm l}N_{\tilde{x},i,s}|\,\bar{q}_\mu a^\mu_{(\tilde{x}+x_0)}\,|{}^{\rm l}N_{\tilde{x},j,s}\rangle\Big] \tag{2.9.120} $$

$$ = g_\Omega q_{\rm I}^{\alpha\,ij}\int_0^\infty {\rm d}^4\tilde{x}\,{}^{\rm l}G^{ij}(\tilde{x},s,\rho)\underbrace{\Big[\Omega_\alpha^{\rm con}(\tilde{x}+x_0)\theta(s\rho-|x|)+\omega_\alpha(\tilde{x}+x_0)\Big]}_{=:\tilde{\Omega}_\alpha(\tilde{x}+x_0)} $$

$$ + \,{\rm i}\int_0^\infty {\rm d}^4\tilde{x}\,\langle {}^{\rm l}N_{\tilde{x},i,s}|\,\bar{q}_\mu a^\mu_{(\tilde{x}+x_0)}\,|{}^{\rm l}N_{\tilde{x},j,s}\rangle\,. \tag{2.9.121} $$

For notational reasons in the third line the zero term, $\Omega^{\rm con}\theta(s\rho-|x|)$, was added. This term becomes zero with the Heaviside function from the low energy weight function Eq. (2.9.117). The newly introduced scalar field, $\tilde{\Omega}$, can also be used in the earlier discussed high energy part (Eq. (2.9.110)). In this regime the additional Heaviside function at $\Omega^{\rm con}$ is the same as in the high energy weight and therefore doesn't change the integral anymore[66)] (Eq. (2.9.108)).

II.9.9.3 Full determinant

Now the results from the two previous subsections can be combined. Taking the Eq. (2.9.110) and (2.9.121) in terms of the redefined field, $\tilde{\Omega}$, the approximate zero-mode determinant of Eq. (2.9.100)

[66)] Upon Fourier transforms into momentum space this additional Heaviside function generates some problems and so one should switch back to the normal Ω field beforehand.

turns into:

$$\det{}_0\big(O\big) \approx \det_{SU(2)}\big\{Ⓐ + Ⓑ\big\} \tag{2.9.122}$$

$$\approx \det_{SU(2)}\Bigg\{g_\Omega\,\bar{q}^{\alpha\,ij}\int_0^\infty \mathrm{d}^4\tilde{x}\,\underbrace{\Big[{}^{\mathrm{h}}G^{ij}(\tilde{x},s,\rho) + {}^{\mathrm{l}}G^{ij}(\tilde{x},s,\rho)\Big]}_{=:G^{ij}(\tilde{x},s,\rho)}\tilde{\Omega}_\alpha(\tilde{x}+x_0) \tag{2.9.123}$$

$$+\,\mathrm{i}\int_0^\infty \mathrm{d}^4\tilde{x}\Big[\underbrace{\langle{}^{\mathrm{h}}N_{\tilde{x},s}|\,\bar{q}_\mu a^\mu_{(\tilde{x}+x_0)}|{}^{\mathrm{h}}N_{\tilde{x},s}\rangle}_{=:\,{}^{\mathrm{h}}A^c_{(\tilde{x}+x_0)}(-\mathrm{i}\tau^c_I/2)} + \underbrace{\langle{}^{\mathrm{l}}N_{\tilde{x},i,s}|\,\bar{q}_\mu a^\mu_{(\tilde{x}+x_0)}\,|{}^{\mathrm{l}}N_{\tilde{x},j,s}\rangle}_{=:\,{}^{\mathrm{l}}A^c_{(\tilde{x}+x_0)}(-\mathrm{i}\tau^c_I/2)}\Big]\Bigg\} \tag{2.9.124}$$

$$\approx \det_{SU(2)}\Bigg\{g_\Omega\,\bar{q}^{\alpha\,ij}\int_0^\infty \mathrm{d}^4\tilde{x}\,G^{ij}(\tilde{x},s,\rho)\tilde{\Omega}_\alpha(\tilde{x}+x_0) \tag{2.9.125}$$

$$+\,\frac{\tau^c_I}{2}\int_0^\infty \mathrm{d}^4\tilde{x}\Big[\,{}^{\mathrm{h}}A^c_{(\tilde{x}+x_0)} + {}^{\mathrm{l}}A^c_{(\tilde{x}+x_0)}\Big]\Bigg\}^{67)}\,. \tag{2.9.126}$$

The above equation is not exact, as the middle part of the integration, $[s\rho, s^{-1}\rho]$, is explicitly left out. The indices, i, j, are not traced out and the lines concerning the gauge fields, a_μ, will not be analysed further, just as before. The rewriting of the gauge fields in iso-spinor components, ${}^{\mathrm{h/l}}A^c_{(\tilde{x}+x_0)}$, was done as a reminder, that the determinant over isospin space elements also includes elements of the gauge field fluctuation. So, in examinations beyond this work the induced coupling of gauge and scalar fields via the fermionic pseudo zero-modes could give further insights in the underlying structure of instantons in a Higgs field environment.

The scalar part of the equation can be brought into a nicer form by using the definition of the convolution $\big(a(x_0) * b(x_0) := \int \mathrm{d}x a(x) b(x - x_0)\big)$. Again, this rewriting is only approximately correct, as a convolution needs an integral over the whole domain. So, the smaller the contribution from the omitted part is, the smaller is the error by expressing the above expression in a convolution (indicated by the symbol $*$). One finds:

$$\det{}_0\big(\mathrm{i}\not{D} + g_\Omega\Omega\big)_{a_\mu=0} \approx \det_{SU(2)}\Big\{g_\Omega\, q_{\mathrm{I}}^{\alpha\,ij} G^{ij}(-x_0,s,\rho) * \tilde{\Omega}_\alpha(-x_0)\Big\} \tag{2.9.127}$$

$$\approx g_\Omega^2 \det_{SU(2)}\begin{Bmatrix} G^{11} * [\tilde{\Omega}_0 + \mathrm{i}\tilde{\Omega}_3] & G^{12} * [\tilde{\Omega}_2 + \mathrm{i}\tilde{\Omega}_1] \\ -G^{21} * [\tilde{\Omega}_2 - \mathrm{i}\tilde{\Omega}_1] & G^{22} * [\tilde{\Omega}_0 - \mathrm{i}\tilde{\Omega}_3]\end{Bmatrix}_{(-x_0)} \tag{2.9.128}$$

$$\approx g_\Omega^2\Big\{(G^{11} * \tilde{\Omega}_0)(G^{22} * \tilde{\Omega}_0) + (G^{11} * \tilde{\Omega}_3)(G^{22} * \tilde{\Omega}_3)$$

$$+(G^{12} * \tilde{\Omega}_1)(G^{21} * \tilde{\Omega}_1) + (G^{12} * \tilde{\Omega}_2)(G^{21} * \tilde{\Omega}_2)\Big\}_{(-x_0)}\,. \tag{2.9.129}$$

67) The missing coupling constant g_{A} in front of the gauge field contribution is due to the chosen normalisation for gauge fields. It is simply absorbed in the field. So in a 'symmetric' notation this factor can just be put in front of the gauge field integral.

In the second and third line the arguments of the involved functions have been dropped for apparent notational reasons and only the final subscript of x_0 indicates the point of evaluation. Using the definition of the weight function (Eq. (2.9.124)) one can use that $G^{12} = G^{21}$. To evaluate the four products of convolutions in Eq. (2.9.129) the following abbreviation will be employed: $G^a(x) := G^{ij}(x, s, \rho)$, where the superscript $a \in \{1, 2, 3, 4\}$ labels the elements of the (2×2)-matrix above. In addition, one needs to include the integral over the collective coordinate x_0. Combining all this, one can work through some tedious algebra [remember the definition of $\tilde{\Omega}$ (Eq. (2.9.121)), of ${}^{\mathrm{h}}G^a(x)$ (Eq. (2.9.108)) and of ${}^{\mathrm{l}}G^a(x)$ (Eq. (2.9.117))] to arrive at:

$$
\begin{aligned}
© : &= \int \mathrm{d}^4 x_0 \, (G^a * \tilde{\Omega}^\alpha)_{(-x_0)} (G^b * \tilde{\Omega}^\beta)_{(-x_0)} && (2.9.130)\\
&= \int \mathrm{d}^4 x_0 \mathrm{d}^4 y \mathrm{d}^4 z \Big\{ G^a(y) \big[\Omega^\alpha_0(y + x_0)\theta(s\rho - |y|) + \omega^\alpha(y + x_0) \big] \\
&\qquad\qquad \times G^b(z) \big[\Omega^\beta_0(z + x_0)\theta(s\rho - |z|) + \omega^\beta(z + x_0) \big] \Big\} && (2.9.131)\\
&= \int \mathrm{d}^4 x_0 \Big\{ \Big([{}^{\mathrm{h}}G^a * \Omega^\alpha_0] + [({}^{\mathrm{l}}G^a + {}^{\mathrm{h}}G^a) * \omega^\alpha] \Big) \Big([{}^{\mathrm{h}}G^b * \Omega^\beta_0] + [({}^{\mathrm{l}}G^b + {}^{\mathrm{h}}G^b) * \omega^\beta] \Big) \Big\}_{(-x_0)} && (2.9.132)\\
&= \int \mathrm{d}^4 x_0 \Big\{ \Big([{}^{\mathrm{h}}G^a * \Omega^\alpha] + [{}^{\mathrm{l}}G^a * \omega^\alpha] \Big) \Big([{}^{\mathrm{h}}G^b * \Omega^\beta] + [{}^{\mathrm{l}}G^b * \omega^\beta] \Big) \Big\}_{(-x_0)} && (2.9.133)\\
&= \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \Big\{ \Big({}^{\mathrm{h}}\hat{G}^a \hat{\Omega}^\alpha + {}^{\mathrm{l}}\hat{G}^a \hat{\omega}^\alpha \Big) * \Big({}^{\mathrm{h}}\hat{G}^b \hat{\Omega}^\beta + {}^{\mathrm{l}}\hat{G}^b \hat{\omega}^\beta \Big) \Big\}_k . && (2.9.134)
\end{aligned}
$$

In line Eq. (2.9.132) the Heaviside function was absorbed in the weight function and the definitions of the partial weight functions were used. Line Eq. (2.9.133) is simply a rearrangement, using the definition of $\Omega^\alpha = \Omega^\alpha_0 + \omega^\alpha$ and the final expression (Eq. (2.9.134)) can be reached by performing a Fourier transform and using that convolutions turn into products in the process. Before proceeding, the role of x_0 should be discussed briefly. Line Eq. (2.9.130) shows that the determinant only depends on the position of the instanton, x_0, and this translates nicely into the momentum representation (transition from line Eq. (2.9.133) to (2.9.134)). So the zero-mode contribution in total gives a local contribution to the partition function. Later this can be exploited to form a contribution to the effective potential in Sec. III.2. Now © can be further transformed to:

$$
\begin{aligned}
© = &\int \frac{\mathrm{d}^4 k \mathrm{d}^4 u}{(2\pi)^8} \Big[{}^{\mathrm{h}}\hat{G}^a_{(k)} \, {}^{\mathrm{h}}\hat{G}^b_{(k-u)} \, \hat{\Omega}^\alpha_{(k)} \hat{\Omega}^\beta_{(k-u)} \Big] && (2.9.135)\\
+ &\int \frac{\mathrm{d}^4 k \mathrm{d}^4 u}{(2\pi)^8} \Big[\Big({}^{\mathrm{h}}\hat{G}^a_{(k)} \, {}^{\mathrm{l}}\hat{G}^b_{(k-u)} + {}^{\mathrm{l}}\hat{G}^a_{(k-u)} \, {}^{\mathrm{h}}\hat{G}^b_{(k)} \Big) \hat{\Omega}^\alpha_{(k)} \hat{\omega}^\beta_{(k-u)} \Big] && (2.9.136)\\
+ &\int \frac{\mathrm{d}^4 k \mathrm{d}^4 u}{(2\pi)^8} \Big[{}^{\mathrm{l}}\hat{G}^a_{(k)} \, {}^{\mathrm{l}}\hat{G}^b_{(k-u)} \, \hat{\omega}^\alpha_{(k)} \hat{\omega}^\beta_{(k-u)} \Big] . && (2.9.137)
\end{aligned}
$$

Here it was only used that the convolution is symmetric ($a * b = b * a$). The first term (the high energy contribution) is in fact independent of the super-scripts a and b, since ${}^{\mathrm{h}}G \equiv {}^{\mathrm{h}}G^{a} \sim \left|{}^{\mathrm{h}}N\right|^2$ is identical for all a. For the low energy part (Eq. (2.9.137)) and the cross term (Eq. (2.9.136)), mixing low and high energy contributions, this is not true, as the approximate zero-modes are not the same, in general[68)]. Combining Ⓒ with the determinant (Eq. (2.9.129)) gives the final result for the fermionic zero-mode determinant in the scalar sector:

$$\text{Ⓓ} := \mathcal{F}\left\{\int \mathrm{d}^4x_0 \det\nolimits_0\left(\mathrm{i}\not{D} + g_\Omega\Omega\right)_{a_\mu=0}\right\} \tag{2.9.138}$$

$$\approx \int \frac{\mathrm{d}^4k\mathrm{d}^4u}{(2\pi)^8}\left\{{}^{\mathrm{h}}G_{(k)}\,{}^{\mathrm{h}}G_{(k-u)}\sum_\alpha \hat{\Omega}^{\alpha}_{(k)}\hat{\Omega}^{\alpha}_{(k-u)}\right. \tag{2.9.139}$$

$$+\ {}^{\mathrm{h}}\hat{G}_{(k)}\left[\left({}^{\mathrm{l}}\hat{G}^{11}_{(k-u)} + {}^{\mathrm{l}}\hat{G}^{22}_{(k-u)}\right)\left(\hat{\Omega}^{0}_{(k)}\hat{\omega}^{0}_{(k-u)} + \hat{\Omega}^{3}_{(k)}\hat{\omega}^{3}_{(k-u)}\right) + 2\,{}^{\mathrm{l}}\hat{G}^{12}_{(k-u)}\left(\hat{\Omega}^{1}_{(k)}\hat{\omega}^{1}_{(k-u)} + \hat{\Omega}^{2}_{(k)}\hat{\omega}^{2}_{(k-u)}\right)\right] \tag{2.9.140}$$

$$\left.+\left[{}^{\mathrm{l}}\hat{G}^{11}_{(k)}\,{}^{\mathrm{l}}\hat{G}^{22}_{(k-u)}\left(\hat{\omega}^{0}_{(k)}\hat{\omega}^{0}_{(k-u)} + \hat{\omega}^{3}_{(k)}\hat{\omega}^{3}_{(k-u)}\right) + {}^{\mathrm{l}}\hat{G}^{12}_{(k)}\,{}^{\mathrm{l}}\hat{G}^{12}_{(k-u)}\left(\hat{\omega}^{1}_{(k)}\hat{\omega}^{1}_{(k-u)} + \hat{\omega}^{2}_{(k)}\hat{\omega}^{2}_{(k-u)}\right)\right]\right\}. \tag{2.9.141}$$

This lengthy expression represents the main result of this section. It encodes the contribution to the partition function from approximate fermionic zero-modes interacting with a scalar field, Ω. Line Eq. (2.9.139) gives the part of the determinant, which depends on high energies, the last line (Eq. (2.9.141)) gives the low energy contribution and the line in between (Eq. (2.9.140)) gives an effective interaction of the low and high energy regime. This part is naturally generated, as the determinant gives products of two convolutions (Eq. (2.9.129)). The low energy contribution only modifies the fluctuation ω without changing the VEV part and so in this sector only the fluctuations acquire the quadratic correction (Eq. (2.9.141)). With the concepts of effective field theories from Sec. II.6 in mind Ⓓ nicely seperates the constrained instanton contributions into terms at different energy scales. To arrive at an explicit equation the G-functions have to be inserted, which have been defined as:

$${}^{\mathrm{h}}G_{(x)} = {}^{\mathrm{h}}G^{a}_{(x)} = \left|{}^{\mathrm{h}}N_{(x)}\right|^2\theta(s\rho - |x|) = \left|\frac{\rho}{\pi\left[x^2+\rho^2\right]^{3/2}}\right|^2\theta(s\rho - |x|)\,, \tag{2.9.142}$$

$${}^{\mathrm{l}}G^{ij}_{(k)} = \mathcal{F}\left[\langle{}^{\mathrm{l}}N_{x,i}|{}^{\mathrm{l}}N_{x,j}\rangle\theta(|x| - s^{-1}\rho)\right] = \left(\frac{2\pi\rho}{\not{k}+M_i}\frac{2\pi\rho}{\not{k}+M_j}\right) * \mathcal{F}\left[\theta(|x| - s^{-1}\rho)\right]_{(k)}. \tag{2.9.143}$$

This ends the general derivation of the fermionic zero-mode contribution and only as a reminder, the fluctuation contribution from the gauge field is missing in the above equation, as it has been set to zero. Before turning to the complete, instanton induced, partition function one can make the connection to the work by Saito and Shigemoto (Ref. [1]). If one focusses only on the high energy contribution to the

[68)] If the VEV of Ω is simply given by a constant $\sigma_0 I$, then all low energy approximate zero-modes are in fact the same, as the masses of the iso-spinor M_j are the same.

zero-mode determinant, one is left with the terms that they proposed in their work in 1979:

$$\mathbb{D}\Big|_{\text{high}} \sim \int \frac{\mathrm{d}^4k\mathrm{d}^4u}{(2\pi)^8}\, {}^{\mathrm{h}}\hat{G}_{(k)}\, {}^{\mathrm{h}}\hat{G}_{(k-u)} \sum_{\alpha} \hat{\Omega}^{\alpha}_{(k)}\hat{\Omega}^{\alpha}_{(k-u)} \tag{2.9.144}$$

The only difference, compared to their results, are the Fourier transformed Heaviside functions (compare Eq. (2.9.142)). But as these are directly induced by the constrained instantons and their different zero-modes in different energy regimes, this is an expected deviation. In their derivation, they split the weight function up into a constant part and a momentum dependent remainder. The constant term they used to explicitly construct an effective quadratic contribution to the scalar Lagrangian, $\mathcal{L}_{\text{inst}} = \frac{a^2}{2}(\Omega^\alpha)^2$.

When their paper was published the implications of constrained instantons and the importance of their employment instead of normal instantons was not yet known. The above derivation shows that the basic idea of their paper is still applicable but it turns into a high energy effect, if one believes in the existence of constrained instantons.

Due to the limited time resources the later model will only make use of the constant high energy contribution (Eq. (2.9.144)), as it was already used by Saito and Shigemoto. Nevertheless future works should of course, include investigations of the momentum dependent parts and of the effects from the low energy contributions, which are explicitly excluded when focussing only on Eq. (2.9.144).

II.9.9.4 Dimensional analysis

As a final step in the calculation of approximate zero-mode determinants a short dimensional analysis is in order. When the whole business of constrained instantons was started the original question was to calculate the generic partition function

$$Z_{\text{gen}} = \int \mathrm{d}^4x_0\mathrm{d}\rho\; m_1(\rho)m_2(\rho,\mu_0)\; m_{\text{higgs}}(\rho,\lambda) \times \det{}_0(\mathrm{i}\not{D} + g_\Omega\Omega)\,. \tag{2.9.145}$$

Using this equation it is relatively straight forward to check the dimensionful quantities. The analysis will be done in terms of powers of energy and further details can be reviewed in Sec. III.4.2 and in Ref. [5, 90-92]. The 'energy-power' of a quantity will be labelled by $[\cdot]$, so a mass for example has $[M] = 1$. Using this, the dimension of the partition function is $[Z_{\text{gen}}] = 0$, as it is just the total number of configurations of the system. For the right-hand side one already knows that the second measure correction and the 'Higgs correction' have zero energy dimension as well.: $[m_2(\rho,\mu_0)] = [m_{\text{higgs}}(\rho,\lambda)] = 0$. (compare Sec. II.9.7 and Sec. II.9.8). Thus one knows:

$$\left[\int \mathrm{d}^4x_0\mathrm{d}\rho\; m_1(\rho)\; \det{}_0(\mathrm{i}\not{D} + g_\Omega\Omega)\right] = 0\,. \tag{2.9.146}$$

Further, using the definition of $m_1(\rho)$ (Eq. (2.9.34)) one finds $\left[\int \mathrm{d}^4x_0 \mathrm{d}\rho\, m_1(\rho)\right] = -4 - 1 + 5 = 0$. So, all parameters have been normalised such that the determinant must have zero energy dimension as well. In terms of the original fields (not Fourier transformed) the determinant is a sum of various terms of the following form (compare Eq. (2.9.129)):

$$\int \mathrm{d}^4x \mathrm{d}^4y G^a(x - x_0) G^b(y - x_0) \tilde{\Omega}^\alpha(x) \tilde{\Omega}^\alpha(y) . \tag{2.9.147}$$

where $G^a = G^{ij} = \langle N_i | N_j \rangle$ and $\tilde{\Omega}^\alpha$ have the same definitions as in the previous section. The energy dimension of a scalar field is known[69)] to be $[\tilde{\Omega}] = 1$ and this now allows to determine the needed dimension of G^a and with this the normalisation of the fermionic pseudo zero-modes:

$$0 = \left[\int \mathrm{d}^4x \mathrm{d}^4y G^a(x - x_0) G^b(y - x_0) \tilde{\Omega}^\alpha(x) \tilde{\Omega}^\alpha(y)\right] \tag{2.9.148}$$

$$= \left[\int \mathrm{d}^4x \mathrm{d}^4y\right] + 2\,[G^a] + 2\left[\tilde{\Omega}^\alpha\right] = -8 + 2\,[G^a] + 2\,, \tag{2.9.149}$$

$$\Rightarrow [G^a] = \left[\langle N_i | N_j \rangle\right] = 3\,, \tag{2.9.150}$$

$$\Rightarrow \left[| N_j \rangle\right] = 3/2\,. \tag{2.9.151}$$

So the zero-modes need to have an energy dimension of 3/2, which does not agree with the so far used normalisation. In Sec. II.9.8.1 the dimension of the zero-modes was found to be $[|^{\mathrm{h}}N\rangle] = [\psi^j_{\mathrm{B}}] = 2$ and thus the modes have to be rescaled by an additional factor of $\rho^{1/2}$, to maintain a dimensionless determinant. So, in a final step the zero-modes have to be replaced by:

$$|^{\mathrm{h/l}}N\rangle \to |^{\mathrm{h/l}}\tilde{N}\rangle = \rho^{1/2} |^{\mathrm{h/l}}N\rangle\,. \tag{2.9.152}$$

This analysis is correct for all contributions to the zero-mode determinant. The rescaled pseudo zero-modes can be used throughout the whole derivation of instantons, as no result depends on their normalisation. The dimensional analysis does also apply to the Fourier transformed expression, if the integration over the instanton position is included with the correct measure contribution, $\int \mathrm{d}^4x_0\, \rho^{-4}$. In fact the final result of the determinant in momentum space from the previous section (Eq. (2.9.141)) does only change by a factor of ρ^{-2} since the new normalisation generates a factor of ρ^2 which is absorbed by the normalised x_0 integration.

II.9.10 The full partition function

After a long derivation of all contributions to the partition function that are generated through the inclusion of instantons, one is now in the position to write down the complete partition function. In the

[69)] This will be explained in detail in Sec. III.4.2.

derivation of explicit results (Sec. II.9-II.9.9.3) a partition function of the form

$$Z_{\text{gen}} = \int \mathrm{d}^4 x_0 \mathrm{d}\rho \; m_1(\rho) m_2(\rho,\mu_0) \; m_{\text{higgs}}(\rho,\lambda) \times \det{}_0(\mathrm{i}\not{D} + g_\Omega \Omega) \tag{2.9.153}$$

was used. This function was motivated by the expansion of a gauge field model around the classical anti-instanton ($n = -1$) solution with a nonzero VEV of the scalar field. Before this, in the introductory part, a slightly different point of view was used. There the starting point was a pure gauge field model which has been enriched with a sensible vacuum state (Sec. II.8.4) leading to a partition function of the form:

$$Z_\theta = \sum_n \int \mathscr{D}A_n \, \exp \int \mathrm{d}^4 x_{\mathrm{E}} \, \mathrm{Tr}\left[-\frac{1}{2} F^{\mu\nu} F_{\mu\nu} + \mathrm{i} \frac{g_{\mathrm{A}}^2 \theta}{16\pi^2} \widetilde{F}^{\mu\nu} F_{\mu\nu} \right] . \tag{2.9.154}$$

To bring both ideas together the derived Z_{gen} for one ($n = -1$) anti-instanton has to be generalised to arbitrary winding numbers, n (which then includes instanton terms with positive n as well), and in addition the earlier discussed vacuum state has to be included. The inclusion of the vacuum is very straight forward, by just adding in a phase for the corresponding vacuum angle, $e^{\mathrm{i}n\theta}$, as it was done in Sec. II.8.4. Before discussing the treatment of different winding numbers it is useful to reorganise the generic partition function slightly, so that different contributions can be separated nicely:

$$\begin{aligned} Z_{\text{gen}} &= \int \mathrm{d}^4 x_0 \mathrm{d}\rho \; m_1(\rho) \; m_{\text{higgs}}(\rho,\lambda) \det{}_0(\mathrm{i}\not{D} + g_\Omega \Omega) \; m_2(\rho,\mu_0) & (2.9.155) \\ &= \left[\int \mathrm{d}^4 x_0 \mathrm{d}\rho \; m_1(\rho) \; m_{\text{higgs}}(\rho,\lambda) \det{}_0(\mathrm{i}\not{D} + g_\Omega \Omega) \, e^{-8\pi^2/g_{\mathrm{A}}^2} \rho^8 \right] \int \mathscr{D}\{\overline{N} N a_\mu \Omega \overline{c} c\}_{\neq 0} e^{-S_{\mathrm{E}}} . & (2.9.156) \end{aligned}$$

Here the measure contribution, m_2, has been written out explicitly, as its treatment varies strongly in different applications. So far all results of the constrained instantons have been presented from a perturbative approach, as mathematical tool are very developed in this field. In this context the measure contribution, m_2, represented a renormalised correction due to quantum fluctuations of order $O(\hbar)$. If one leaves the perturbative regime this is no longer true. In hadron physics one works with the so called chiral condensate and the fluctuations around this VEV can be taken to be observable resonances. Therefore, in this field, the assumption that fluctuations around classical solutions are small is not justified for all fields[70]. Of course, there are many mathematical difficulties in this low energy regime. One needs a different renormalisation procedure and the effective cutoff contribution from the Higgs field m_{higgs} turns out to be problematic. In fact the Higgs field measure leads to an upper bound on the size of constrained instantons, as $m_{\text{higgs}} \to 0$, if $\langle H \rangle \rho \geq 1$. To see this compare the definition of m_{higgs} (Eq. (2.9.58)).

[70] Later only the scalar field will be assumed to have 'large' fluctuations around the VEV ($\Omega = \langle \Omega \rangle + \omega$), while the gauge field will still only have quantum fluctuations around the instanton configuration ($A_\mu = A_\mu^{\text{con}} + O(\hbar)$). The reason for this unequal treatment is, that the mathematical tools in the derivation of instantons become invalid, if fluctuations around A_μ^{con} become 'larger' than $O(\hbar)$.

But as the instanton formalism yields many promising features, and as the QCD vacuum 'demands' its inclusion on a fundamental level, it is tempting to 'guess' an effective low energy Lagrangian (or partition function) and see what kind of physics can be modeled with it. The reason behind these scentences of motivation is, that the later model will not bother with a mathematically sound derivation of correct low energy measures from the Higgs field, $\mathcal{m}_{\text{higgs}}(\langle H\rangle, \rho)$, and the renormalisation procedure, $\mathcal{m}_2(\mu_0, \rho)$, but simply assume that both can be adjusted to reproduce any connected observable. If explicit equations were worked out from this demand, one would arrive at constraints on the maximal allowed size of instantons in this model and at an explicit renormalisation scale, μ_0. While the perturbative regime will mostly be left behind from now on, the earlier presented results can easily adopted into the general equations to come and their treatment is rather straight forward.

Now back to the generalisation to arbitrary winding number instantons. This will be done by following the 'historical' approach. The inclusion of other winding numbers becomes very simple, if one employs two assumptions. First, suppose that the distribution of instantons throughout space corresponds to a dilute gas and secondly take all higher winding number configurations as multiple times ($n = \pm 1$) winding number configurations. Looking at Eq. (2.9.156), one sees, that the instanton part has been nicely separated from the fluctuation contribution. If one now goes to a ($n = n_-$) configuration, this means that the part in square brakets appears n_- times. In addition one needs to include a symmetry factor of $1/n_-!$ as the order of the n_- instanton terms does not matter. If one now includes the vacuum angle $e^{\mathrm{i}\theta n}$, and allows n to be positive as well as negative, the generic partition function becomes:

$$Z_{\text{gen}}(n_+, n_-) = \frac{1}{|n_-|!}\left[\int \mathrm{d}^4x_0 \mathrm{d}\rho\, \mathcal{M}\, \det{}_0(\mathrm{i}\not{D} + g_\Omega \Omega)\, e^{-\mathrm{i}\theta}\right]^{|n_-|} \times \frac{1}{n_+!}\left[\int \mathrm{d}^4x_0 \mathrm{d}\rho\, \mathcal{M}\, \det{}_0(\mathrm{i}\not{D} + g_\Omega \Omega)\, e^{\mathrm{i}\theta}\right]^{n_+} \\ \times \int \mathscr{D}\{\overline{N}N a_\mu \Omega \bar{c} c\}_{\neq 0} e^{-S_{\mathrm{E}}}\,. \tag{2.9.157}$$

Here $\mathcal{M} = \mathcal{m}_1\, \mathcal{m}_{\text{higgs}} e^{-8\pi^2/g_{\mathrm{A}}^2} \rho^8$ incorporates all measure contributions. Naturally, if n becomes positive, one has to switch from the anti-instanton to the corresponding instanton equations. It was implicitly assumed that the remaining fluctuation contribution is independent of the winding number. The terms with positive and negative winding number have been separated, as the instanton configurations with negative winding number couple to $\psi_{\mathrm{B}} \in SU_{\mathrm{B}}(2)$ fermionic modes and the ones with positive winding number couple to fields from the group $SU_{\mathrm{A}}(2)$. In Minkowski space this translates to anti-instantons only coupling to left-handed fields and instantons only coupling to right-handed ones. Therefore the fermionic zero-modes are different in both contributions and thus the determinants are, as well.

If one now sums over all winding numbers the partition function just gets two new exponential factors:

$$Z_{\text{gen}} = \sum_{n_+,n_-} Z_{\text{gen}}(n_+, n_-) \tag{2.9.158}$$

$$= \exp\left\{\int \mathrm{d}^4x_0 \mathrm{d}\rho\, \mathcal{M}\left[e^{-\mathrm{i}\theta}\mathrm{DET}\Big|_{n_-} + e^{\mathrm{i}\theta}\mathrm{DET}\Big|_{n_+}\right]\right\}\int \mathscr{D}\{\overline{N}Na_\mu\Omega\bar{c}c\}_{\neq 0}e^{-S_\mathrm{E}} \tag{2.9.159}$$

$$= \exp\left\{\int \mathrm{d}\rho\, \mathcal{M}\left[e^{-\mathrm{i}\theta}\mathbb{D}|_{n_-} + e^{\mathrm{i}\theta}\mathbb{D}|_{n_+}\right]\right\}\int \mathscr{D}\{\overline{N}Na_\mu\Omega\bar{c}c\}_{\neq 0}e^{-S_\mathrm{E}}\,. \tag{2.9.160}$$

In the first expression $\mathrm{DET} := \det_0(\mathrm{i}\not{D} + g_\Omega\Omega)$ is just an abbreviation and in the third line the definition from Eq. (2.9.141) was used. The subscripts $n_\pm$ are a reminder, which zero-modes have to be chosen (instanton or anti-instanton). If one now rewrites S_E into an effective action plus a quantum perturbation around it $S_\mathrm{E} = S_\mathrm{E}^{\text{eff}} + \delta S_\mathrm{E}$, then one finds that the instanton contributions give a contribution to the effective action:

$$Z_{\text{gen}} = \exp\left\{-S_\mathrm{E}^{\text{eff}} + \int \mathrm{d}\rho\, \mathcal{M}\left[e^{-\mathrm{i}\theta}\mathbb{D}|_{n_-} + e^{\mathrm{i}\theta}\mathbb{D}|_{n_+}\right]\right\}\int \mathscr{D}\{\overline{N}Na_\mu\Omega\bar{c}c\}_{\neq 0}e^{-\delta S_\mathrm{E}}\,. \tag{2.9.161}$$

This generic partition function will be the starting point for the model in the next chapter. As for the model most aspects of this equation will be dropped right away (to arrive at tractable equations), they should be mentioned at least once to give possible starting points for future explorations.

- The measure contribution, $\mathcal{M}$, incoporates a Higgs contribution from the interaction of Higgs and instanton field (see Sec. II.9.8), the classical instanton action, $e^{-8\pi^2/g_\mathrm{A}^2}$, and a renormalisation, that has to be specified according to the fluctuation action, δS_E.
- The factor $\mathbb{D}$ is given in Eq. (2.9.141). This term explicitly gives the zero-mode contribution for the anti-instanton in the case of vanishing gauge field fluctuations ($a_\mu = 0$). If these fluctuations shall be included as well then one has to revert to Eq. (2.9.126) and work out the additional terms. In the other direction Eq. (2.9.144) gives an estimate of the determinant under the assumption, that low energy phenomena and gauge field contributions can be neglected. In principle the ideas from Sec. II.6 suggest that the low energy part of $\mathbb{D}$ should be included in an effective model at very low energies, while the high energy contribution and the mixed terms should be included, if the cutoff scale is shifted to higher energies.
- Then, there is the path integral over the fluctuation contributions, δS_E. This part of the partition function has to be adjusted according to the specific needs of the model one is interested in. In effective field theories this adjustment will be hard in general, as, on the one hand, one needs to find a separation scale, that cuts off phenomena that are regarded as unimportant and on the other hand all remaining processes have to be important and make physical sense.
- Finally, if one is interested in the earlier discussed perturbative regime of instanton physics, then

one can come back to Eq. (2.9.155) and follow the steps with the original measure correction m_2 in place. In the final expression this will mean that $S_{\mathrm{E}} = 0$, as its contribution has already been absorbed in m_2, and the factor is changed to $\mathcal{M} = m_1\, m_2\, m_{\mathrm{higgs}}$.

III The model

Finally, after a long theoretical journey, enough bits and pieces have been accumulated to build a model that simulates the interaction of fermionic and bosonic degrees of freedom based on certain symmetry assumptions, the concept of instantons and other appealing prerequisites. In principle nuclear matter is almost completely characterised by the nature of the 'strong force', that is Quantum Chromo Dynamics. In Sec. II.6 it was discussed that the high energy behaviour of QCD is accessible via perturbation theory, while the low energy phenomena have to be handled using some kind of effective model. The exact point, from which on perturbative methods are a legitimite tool is not of importance for the present model. As it shall give an effective description of systems in, or near the ground state it is certainly in the non-perturbative regime of QCD. The model as it will be built in this chapter is originally designed to describe nuclear matter and therefore all examples will be discussed concerning this application but it should be mentioned that, in principle, it is as well suited for the description of other fermionic and bosonic systems in a non-perturbative context. Further 'initial assumptions' are discussed below.

III.1 Defining assumptions

The following 'initial assumptions' are chosen for mainly two reasons. The first category incorporates assumtpions that are widely to believed true among physicists (like for example Lorentz invariance) while the second category assembles all assumptions that are chosen for convenience reasons, as calculations would become significantly more involved if they weren't included.

EFT: In Sec. II.6 the concepts of effective field theories have been briefly introduced. The reason for this has of course been that the model in the following sections will be built as an effective approximation to QCD. While the general concepts for a vigoreous design of an effective model have been discussed in the introductory section (Sec. II.6), the needed steps will not be performed here. Naturally this lacking detail should be investigated in later studies.

Chiral symmetry: In Sec. II.1 the structure of the Lorentz group, $SO^{+}(3,1) \simeq SU_{\mathrm{L}}(2) \otimes SU_{\mathrm{R}}(2)$, was presented and in Sec. II.5, as a possible extension of this symmetry, the chiral symmetry was presented. This symmetry enlarged the Lorentz group to: $U_{\mathrm{L}}(2) \otimes U_{\mathrm{R}}(2)$. Later, in the introductory part to instantons in Sec. II.8.4 it was found that the concept of instantons explicitly violates the axial $U_{\mathrm{A}}(1)$ part of a general chiral symmetry. Therefore, in the model the underlying group structure of space-time will be assumed to have the broken chiral symmetry:

$$U_{\mathrm{V}}(2) \otimes SU_{\mathrm{A}}(2) \simeq U_{\mathrm{V}}(1) \otimes SU_{\mathrm{V}}(2) \otimes SU_{\mathrm{A}}(2) \simeq U_{\mathrm{V}}(1) \otimes SU_{\mathrm{L}}(2) \otimes SU_{\mathrm{R}}(2) \tag{3.1.1}$$

Transformations in the combined Lorentz and partial chiral symmetry space will be named $U_{\mathscr{L},ch}$, in agreement with Sec. II.5.

Isospin symmetry: In addition to the symmetry of space-time an additional *local* isospin symmetry, $SU_{\mathrm{I}}(2)$, will be assumed to hold for the model. This assumption fits exactly to all derivation parts about instantons and so it simply establishes one of the prerequisits for an inclusion of instantons. As the symmetry is local, an additional gauge field, $A_\mu \in SU_{\mathrm{I}}(2)$, will be needed, which will give the instanton contributions.

Spontaneous symmetry breaking: The mass generation of all constituents of the model will be done via a spontaneous symmetry breaking in the scalar sector. Some cornerstones for this have already been carved in Sec. II.9.8 for the instanton sector. The implications for fermionic, scalar and dynamical gauge field parts will be presented as the model is developed.

Degrees of freedom: In principle the relevant degrees of freedom for this model in a generic model setting are: a number of n_{f} fermions, N, and n_{s} scalars, Ω, that transform under isospin transformations. In addition a gauge field, A_μ, is needed to account for the local character of the isospin symmetry. As a concrete application to a physically relevant case the model will be presented as an effective nucleon model.

Fermion structure: The fermion iso-spinor in the model will be taken to be $N = (p, n)^{\mathrm{T}}$, where the constituents are proton and neutron Dirac-spinors. Thus N will live in isospin space and p and n will transform under combined Lorentz and chiral transfromations, $U_{\mathscr{L},\mathrm{ch}}$.

Scalar field structure: The scalar degrees of freedom will be taken to be σ, $\boldsymbol{\pi}$, η and $\boldsymbol{\delta}$ ($\triangleq a^0(980)$ in particle data group notation)[1)]. In the model it will be assumed that they obey an overall rotational symmetry. This means they can be assembled in one vectorial structure of the form $\Omega^\alpha = (\sigma, \boldsymbol{\pi}^{\mathrm{T}}, \eta, \boldsymbol{\delta}^{\mathrm{T}})^{\mathrm{T}}$ and one has an eight dimensional rotational symmetry, $\Omega \in SO(8)$. Later, in Sec. III.2.1, this structure will be reassembled in a more convenient biquaternion representation. With this the scalar field, Ω, will transform under isospin and transformations and $U_{\mathscr{L},\mathrm{ch}}$.
Apart from this it will be assumed that Ω can be separated into a VEV contribution, $\langle\Omega\rangle$, and a part that captures the the dynamical (physical) mesonic contributions.

Gauge field structure: Finally the effective gauge field, A_μ, will be expanded around the classical field configuration - that is the constrained instanton solution. So the gauge field can be written

[1)] $\sigma(x), \eta(x) \in \mathbb{R}$ and $\boldsymbol{\pi}(x), \boldsymbol{\delta}(x) \in \mathbb{R}^3$.

as $A_\mu = A_\mu^{\text{con}} + a_\mu$, where A_μ^{con} represents the constrained instanton solution, as introduced in Sec. II.9.8 (equation Eq. (2.9.55)) and a_μ is a fluctuation of order $O(\hbar)$. Writing the gauge field in this form allows to use the main results of (constrained) instanton calculations from chapter II.
'Large' fluctuations (which are larger than the quantum fluctuations of order $O(\hbar)$) around A_μ^{con} will not be included, as fluctuations of this magnitude would lead to a complete breakdown of the instanton calculations. For example the calculations in Sec. II.9.3-II.9.5 explicitly relied on the expansion $A_\mu = A_\mu^{\text{con}} + a_\mu$ and thus would not be usable in the context of 'large' fluctuations.

Euclidean/Minkowski space: In Sec. II.8.1 the conceptual problems with the transition between euclideanised Minkowski space and ordinary Minkowski space have been discussed (Usually this transition is called a 'Wick-rotation'). As path integrals are only well defined in the mathematical sense in Euclidean space-time and as the concept of instantons is only rigorously accessible in Euclidean space, it is sensible to start out with a model in eudclideanised Minkowski space. Later, in order to relate any finding from the Euclidean model to physical observables, one needs to make the transition to Minkowski space, even if this is not well defined from a mathematical perspective. Therefore, the general structure of the model (Sec. III.2) will be presented in Euclidean space-time, but from there on all Lagrange densities of interest will be taken to be in Minkowski space. Sec. III.5.2 will explicitly discuss the related problems for the gauge field Lagrangian, as the conceptual problems are most apparent there.

III.2 General model structure

Combining the assumptions, discussed in the previous section allows writing down the partition function in Euclidean space of the model. For this suppose that the contributing fields, F, can be splitt up into a classical part and a quantum fluctuation of order $O(\hbar)$ around it (indicated by δF):

$$N^{\text{full}} = N + \delta N\,, \tag{3.2.1}$$

$$\Omega_\alpha^{\text{full}} = \Omega_\alpha + \delta\Omega_\alpha = \langle\Omega_\alpha\rangle + \omega_\alpha + \delta\Omega_\alpha\,, \tag{3.2.2}$$

$$A_\mu^{\text{full}} = A_\mu^{\text{con}} + \delta a_\mu\,. \tag{3.2.3}$$

The Euclidean action, S_{E}, is then minimised by the fields without the fluctuations, δF. It can then be written into an effective part and a perturbation around it:

$$\begin{aligned} S_{\text{E}} &= S_{\text{E}}^{\text{eff}}(N, \overline{N}, A, \Omega, \overline{c}, c) + \delta S_{\text{E}} && (3.2.4)\\ &= \int \mathrm{d}^4x_{\text{E}} \left[\mathcal{L}_N^{\text{eff}} + \mathcal{L}_\Omega^{\text{eff}} + \mathcal{L}_A^{\text{eff}} + \mathcal{L}_{\text{gfh}}^{\text{eff}}\right] + \delta S_{\text{E}} && (3.2.5)\end{aligned}$$

Here the effective nucleon Lagrangian, $\mathcal{L}_N^{\text{eff}} = \mathcal{L}_N^{\text{eff}}(\overline{N}, N, \Omega, A)$, incorporates all nucleon contributions - that is to say the free nucleon Lagrangian and the coupling terms to scalar and gauge fields. In analogy $\mathcal{L}_\Omega^{\text{eff}}$ contains all scalar terms and their coupling to the gauge field and finally the effective gauge field Lagrangian, $\mathcal{L}_A^{\text{eff}}$, only consists of the kinetic term from the gauge field. The part $\mathcal{L}_{\text{gfh}}^{\text{eff}}$ is the combined gauge fixing and ghost Lagrangian, which will depend on the gauge, scalar and ghost field later.
Apart from the ordinary terms in the particular contributions, $\mathcal{L}_{\text{K}}^{\text{eff}}$, which will be discussed in the sections to come, there are in principle high energy contributions, $\delta\mathcal{L}_{\text{K}} \sim \sum_d c_d(\mu_0) O_d$, in each effective Lagrangian. They are part of the EFT formalism, how it was discussed in Sec. II.6. Here these terms are only listed as a reminder, but from now on any such contribution will be ignored, as the exact EFT formalism is not worked out in the present model (corresponding to setting $c_d(\mu_0) = 0$ and ignoring EFT renormalization corrections).
Now the action of the effective nucleon model can be inserted into the earlier derived expression of the partition function in Sec. II.9.10. In the form of Eq. (2.9.161) this expression already contains the full dependence on instanton induced effects:

$$Z_{\text{gen}} = \exp\left\{-S_{\text{E}}^{\text{eff}}\Big|_{A_\mu^{\text{con}}=0} + \int \mathrm{d}\rho\, \mathcal{M}\left[e^{-\mathrm{i}\theta} \text{Ⓓ}|_{n_-} + e^{\mathrm{i}\theta} \text{Ⓓ}|_{n_+}\right]\right\} \int \mathcal{D}\{\overline{N} N a_\mu \Omega \bar{c} c\}_{\neq 0} e^{-\delta S_{\text{E}}} \,. \tag{3.2.6}$$

Here Ⓓ was given in Eq. (2.9.141) and the definition of the measure is $\mathcal{M} = m_1\, m_{\text{higgs}} e^{-8\pi^2/g_{\text{A}}^2} \rho^8$. Notice that above the instanton contribution, A_μ^{con}, is set to zero in the effective action, $S_{\text{E}}^{\text{eff}}$. The reason for this is that the kinetic term of a (constrained) instanton can be calculated explicitly, leading only to an effective contribution to the measure:

$$m_{\text{higgs}} e^{-8\pi^2/g_{\text{A}}^2} = \exp\left\{-\int \mathrm{d}^4 x_{\text{E}} \left[\mathcal{L}_\Omega^{\text{eff}}(A_\mu^{\text{con}}, \langle\Omega\rangle) + \mathcal{L}_A^{\text{eff}}(A_\mu^{\text{con}})\right]\right\} \,. \tag{3.2.7}$$

For the measure contribution from the Higgs sector compare Eq. (2.9.58) and for the effect of constrained instantons (instead of ordinary ones) compare Eq. (2.9.59). The reason why the instanton action occurs as a measure in the effective action was discussed in Sec. II.9.10. For the purpose of the model presentation in the remaining sections of this work the influence of quantum fluctuations will be left out and thus the remaining path integral over the fluctuations will be set to one.
The last thing to specify is, what part of the fermionic zero-mode contribution will be included in the model. In the end of Sec. II.9.9.3 it was already indicated, that only the high energy contribution will be

(partly) included in the model. In Eq. (2.9.144) this high energy part of Ⓓ was found to be:

$$\int \mathrm{d}\rho\, \rho^2\, \mathcal{M} e^{-\mathrm{i}\theta} Ⓓ|_{n_-} = e^{-\mathrm{i}\theta} \int \mathrm{d}\rho\, \rho^2\, \mathcal{M} \int \frac{\mathrm{d}^4k \mathrm{d}^4u}{(2\pi)^8}\, {}^{\mathrm{h}}\hat{G}_{(k)}\, {}^{\mathrm{h}}\hat{G}_{(k-u)} \sum_\alpha \hat{\Omega}^\alpha_{(k)} \hat{\Omega}^\alpha_{(k-u)} \tag{3.2.8}$$

$$= e^{-\mathrm{i}\theta} \int_0^{s\rho} \mathrm{d}^4x \int_0^{s\rho} \mathrm{d}^4y \underbrace{\int \mathrm{d}\rho\, \rho^2\, \mathcal{M} \int \mathrm{d}^4x_0\, |{}^{\mathrm{h}}N_{(x-x_0)}|^2 |{}^{\mathrm{h}}N_{(y-x_0)}|^2}_{=:\mathcal{G}(x,y)} \sum_\alpha \Omega^\alpha_{(x)} \Omega^\alpha_{(y)}\,. \tag{3.2.9}$$

The additional factor of ρ^2 is included to produce the correct normalisation of the zero-modes (compare Sec. II.9.9.4). In the second line only the definition (Eq. (2.9.142)) has been inserted to get to the earlier used spacial representation. In order to find the part of the introduced function, $\mathcal{G}(x, y)$, which alters an ordinary effective potential assume that it can be separated into a local and a non-local contribution:

$$\mathcal{G}(x, y) = \frac{a^2}{4} \delta^{(4)}(x - y) + \mathcal{G}_2(x, y)\,. \tag{3.2.10}$$

Here the constant factor of $a^2/4$ is only chosen for later convenience reasons. The remaining function, $\mathcal{G}_2(x, y)$, incorporates all non-local contributions. Inserting this in the above expression gives:

$$\int \mathrm{d}\rho\, \rho^2\, \mathcal{M} e^{-\mathrm{i}\theta} Ⓓ|_{n_-} = e^{-\mathrm{i}\theta} \int_0^{s\rho} \mathrm{d}^4x \int_0^{s\rho} \mathrm{d}^4y \left[\frac{a^2}{4}\delta^{(4)}(x-y) + \mathcal{G}_2(x,y)\right] \sum_\alpha \Omega^\alpha_{(x)} \Omega^\alpha_{(y)} \tag{3.2.11}$$

$$= \int \mathrm{d}^4x\, e^{-\mathrm{i}\theta} \theta(s\rho - x) \sum_\alpha \frac{a^2}{4} (\Omega^\alpha_{(x)})^2 + \big(\text{non-local}\big)\,. \tag{3.2.12}$$

Now, dropping the non-local contributions and repeating the procedure for the n_+ expression, all terms of Eq. (3.2.6) can be inserted. For n_+ note that the only difference is that the approximate zero-modes are the ones of instantons, instead of anti-instantons. This will only change Ω to its hermitian conjugate $\Omega^\dagger$, as will become clear at the end of Sec. III.3.2.

$$Z_{\text{model}} = \exp\left\{ - \int \mathrm{d}^4x_{\mathrm{E}} \left[\mathcal{L}^{\text{eff}}_N + \mathcal{L}^{\text{eff}}_\Omega + \mathcal{L}^{\text{eff}}_{\text{gfh}} + \underbrace{\frac{a^2}{4}\Big(e^{-\mathrm{i}\theta}\Omega^2_\alpha + e^{+\mathrm{i}\theta}\Omega^{\dagger\,2}_\alpha\Big)\theta(s\rho - x)}_{=:\mathcal{L}_{\text{inst}}} \right] \right\}. \tag{3.2.13}$$

As a final change the Heaviside function in $\mathcal{L}_{\text{inst}}$ will be dropped, in order to simplify calculations. Effectively this sets the status of the model back to a model that uses ordinary instead of contrained instantons. Thus once the calculations for this 'over-simplfied' version of an effective instanton model are completed, one should really investigate the differences that occur if one employs constrained instantons instead of the 'older version'. In the remaining sections the different parts $\mathcal{L}^{\text{eff}}_{\text{K}}$ of the effective partition function Z_{model} will be worked out explicitly.

III.2.1 Scalar field structure generalisation of the σ-model

Before coming to the explicit effective Lagrange densities, it is sensible to translate the vectorial structure of the scalar field, $\Omega^\alpha = (\sigma, \boldsymbol{\pi}^{\mathrm{T}}, \eta, \boldsymbol{\delta}^{\mathrm{T}})^{\mathrm{T}}$, from Sec. III.1 into a corresponding structure in iso-spinor space. The reason for this is that a local isospin symmetry was among the defining assumptions. Therefore one needs to work out how all constituents transform under this symmetry and it turns out that working in iso-spinor space is very convenient for most questions concerning the effective Lagrange densities[2)].

In Sec. II.2 it was discussed at length that a $(2 \otimes 2)$ invariant symbol in iso-spinor space could be used to produce a mapping, $q^\alpha_{\mathrm{I}\,ab}$, from a matrix valued element in iso-spinor space to a 4-dimensional vector in Euclidean space. So, as a start, focus on the first four components of Ω. If one takes these to be $\Phi^\alpha := (\sigma, \boldsymbol{\pi}^{\mathrm{T}})^{\mathrm{T}}$, then this Euclidean vector is translated to iso-spinor space via: $\Phi_{ab} = \Phi^\alpha q^\alpha_{\mathrm{I}\,ab}$.

In 1960 Gell-Mann and Lévy presented the key ideas to a model which became known as the '(linear) σ-model' Ref. [33, p.717-719]. For the σ-model they assumed that the effective Lagrange density is invariant under an additional symmetry, which rotates four meson fields into each other. They combined the three pseudo-scalar pions, $\boldsymbol{\pi}$, with an additional scalar particle, σ. In other words they used the 4-dimensional Φ^α field and assumed that the Lagrange density would be invariant under $SO(4)$ transformations acting on Φ^α.

In order to generalise the scalar structure of the 4-dimensional σ-model to the present (8-dimensional) case it is useful to have another look at the definition of the scalar product in iso-spinor space, given in Eq. (2.2.6):

$$(A, B)_{\mathrm{H}} = \frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[(A^\alpha q^\alpha_{\mathrm{I}\,ab})^\dagger B^\beta q^\beta_{\mathrm{I}\,cd}\right] = (A^\alpha)^\dagger B^\beta \left(\frac{1}{2}\bar{q}^{ab\,\alpha}_{\mathrm{I}} q^\beta_{\mathrm{I}\,ab}\right) = \sum_{\alpha=1}^{4} (A^\alpha)^\star B^\alpha\,. \tag{3.2.14}$$

The last equality in the above equation shows that, if the vectors are complex ($A^\alpha, B^\alpha \in \mathbb{C}^4$), instead of real, then $(A, B)_{\mathrm{H}}$ simply gives the regular scalar product in a complex, 4-dimensional vector space. As a $2n$-dimensional real vector space can be identified with a n-dimensional complex one, this observation allows to rewrite the 8-dimensional scalar field into a 4-dimensional complex one: $\Omega \in \mathbb{R}^8 \leftrightarrow \Omega \in \mathbb{C}^4$. Thus the 4-dimensional scalar field structure of the σ-model can 'natrually' be generalised to an 8-dimensional structure by an identification of the form $\Phi = (\sigma, \boldsymbol{\pi}^{\mathrm{T}})^{\mathrm{T}}$ and $\Lambda = (\eta, \boldsymbol{\delta}^{\mathrm{T}})^{\mathrm{T}}$. This gives for the

[2)] In fact, the only the gauge fixing and thus the ghost Lagrangian are easier to work out in a different representation (compare Sec. III.5.5).

complete scalar field:

$$\Omega_\alpha = \Phi_\alpha + \mathrm{i}\Lambda_\alpha = \begin{pmatrix}\sigma \\ \boldsymbol{\pi}\end{pmatrix} + \mathrm{i}\begin{pmatrix}\eta \\ \boldsymbol{\delta}\end{pmatrix}, \tag{3.2.15}$$

$$\Omega_{ab} = \Omega^\alpha q^\alpha_{\mathrm{I}\,ab}\,, \tag{3.2.16}$$

$$\Rightarrow \|\Omega\|^2 = (\Omega,\Omega)_{\mathrm{H}} = \frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[(\Omega^{ab})^\dagger \Omega_{cd}\right] = \sum_{\alpha=1}^{4}(\Omega^\alpha)^\star\Omega^\alpha \tag{3.2.17}$$

$$= |\Phi|^2 + |\Lambda|^2 = \sigma^2 + \boldsymbol{\pi}^2 + \eta^2 + \boldsymbol{\delta}^2\,. \tag{3.2.18}$$

From the calculation of the $\|\Omega\|^2$ one finds that the scalar product, defined in Eq. (3.2.14), preserves the assumed $SO(8)$ symmetry of Ω. Thus it is possible to reexpress Ω in a terms of the iso-spinor representation. The only difference, compared to the ordinary σ-model is, that now the iso-spinor representation $\Omega_{ab} = \Omega^\alpha q^\alpha_{\mathrm{I}\,ab} = (\Phi_\alpha + \mathrm{i}\Lambda_\alpha) q^\alpha_{\mathrm{I}\,ab}$ is complex, instead of real.
How the scalar field transforms under the imposed symmetries (isospin, chiral and Lorentz) of the model will be discussed in the context of its coupling to nucleons (Sec. III.3.2).

III.3 The Nucleon Lagrangian

With the initial conditions being set in Sec. III.1, it is time to lay the cornerstone of the upcoming model. At the end of the day one would like to describe fermions interacting with mesons and so the first part of the full action to be worked out will be the effective nucleon Lagrangian, $\mathcal{L}_N^{\mathrm{eff}}$. As mentioned, the fermions are combined in a nucleon iso-spinor of the form $N = (p, n)^{\mathrm{T}}$, where the proton and neutron fields are Dirac-spinors with the detailed structure:

$$p = ({}^{\mathrm{p}}\chi_a, {}^{\mathrm{p}}\xi^\dagger_{\dot{a}})^{\mathrm{T}} = p_{\mathrm{L}} + p_{\mathrm{R}}\,, \tag{3.3.1}$$

$$n = ({}^{\mathrm{n}}\chi_a, {}^{\mathrm{n}}\xi^\dagger_{\dot{a}})^{\mathrm{T}} = n_{\mathrm{L}} + n_{\mathrm{R}}\,. \tag{3.3.2}$$

The separation into left- and right-handed parts of p and n directly gives a corresponding decomposition of the nucleon spinor into N_{L} and N_{R} [3)]. To construct a Lagrange density, which is invariant under the Lorentz, chiral and isospin group one needs to find terms that are invariant under all these symmetries. Apart from the symmetry requirements the terms that will be included have to be identifiable with important physical quantities, such as kinetic, mass and interaction terms of physical fields.

[3)] Formally the decomposition into left- and right-handed fields is not entirely correct as long as one works in euclideanised Minkowski space, since there the $SU(2)$ subgroups are independent of each other (compare the discussion in Sec. II.8.1).

III.3.1 Kinetic contribution

For a start the kinetic contribution of nucleons with the mentioned symmetry properties shall be included. Using the work in Sec. II.4 and II.5 the simplest realisation is:

$$\mathcal{L}_N^{\mathrm{kin}} = \sum_{\substack{iso-\\spin}} \overline{N}(\mathrm{i}\gamma^\mu\partial_\mu)N = \mathrm{i}(\overline{p}\gamma^\mu\partial_\mu p + \overline{n}\gamma^\mu\partial_\mu n) \tag{3.3.3}$$

$$= \mathrm{i}\sum_{j\in\{\mathrm{L,R}\}}(\overline{p}_j\gamma^\mu\partial_\mu p_j + \overline{n}_j\gamma^\mu\partial_\mu n_j)\,. \tag{3.3.4}$$

The invariance under the Lorentz and Chiral group has been discussed in Sec. II.4 and II.5. For the invariance under the isospin group one has to know that the operator $\mathrm{i}\gamma^\mu\partial_\mu$ does not transform under $SU_{\mathrm{I}}(2)$ and $\overline{N}N$ is just the scalar combination of two vectors in iso-spinor space. Therefore the kinetic term is invariant under $SU_{\mathrm{I}}(2)$ as well. If one goes to the local isospin symmetry, this means that ordinary derivative has to be replaced by a covariant one ($\partial_\mu \to D_\mu = \partial_\mu - \mathrm{i}g_{\mathrm{A}}A_\mu$). The details of this transition will be discussed in context of the gauge field Lagrangian in Sec. III.5.

III.3.2 Nucleon masses and interactions

Having the kinetic term of the nucleon Lagrangian, the question is what other gauge invariant terms could be included. The simplest idea is to take the already derived vectorial structure $\overline{N}\gamma^\mu N$ and couple it to another Minkowski space vector. This term would be invariant following the same lines of thought as for Eq. (3.3.3). While being conveniently effortless, the coupling to Minkowsky space vectors means that one needs this vectorial structure in the first place. In Sec. III.5 this idea leads to the coupling of the gauge field but for now scalar couplings shall be discussed, as these ultimately lead to possible mass terms.

Sec. II.5 already raised the subject of a mass term in a chirally symmetric model. There it was found, that an ordinary fermionic mass term is prohibited in chirally symmetric systems. Instead one needed an additional field, that could compensate the chiral symmetry transformations ($\overline{N}_{\mathrm{L}}wN_{\mathrm{R}} + h.c.$)[4]. In the present case Ω qualifies for the role of the unspecified w field. So, using the notation from Eq. (2.5.7), if Ω transforms as

$$U^{-1}_{\mathscr{L},\mathrm{ch}}\Omega^\dagger U_{\mathscr{L},\mathrm{ch}} = e^{-\mathrm{i}\varepsilon_\ell^j T^j} L\Omega^\dagger R^\dagger e^{\mathrm{i}\varepsilon_r^j T^j}\,. \tag{3.3.5}$$

then the term ($\overline{N}_{\mathrm{L}}\Omega^\dagger N_{\mathrm{R}} + h.c.$)[5] becomes manifestly invariant under chiral and Lorentz symmetries. Here $T^j = \tau^j/2$ are the generators of the left- and right-handed chiral/Lorentz transformations. For the

[4]By including the conjugate field the expression becomes directly hermitian.
[5]The choice of $\Omega^\dagger$ instead of Ω is arbitrary. It is only used here, as it matches earlier conventions.

isospin symmetry the group theoretical point of view allows a quick analysis. In terms of representations $\overline{N}_{\mathrm{L}}\Omega^{\dagger}N_{\mathrm{R}}$ corresponds to the product of four fundamental $SU(2)$ representations and in App. A.4 it is shown, that this tensor product of four '2'-representations contains the trivial representation ($2^{\otimes 4} = 1_{\underline{s}} \oplus ...$). Therefore, out of $\overline{N}_{\mathrm{L}}\Omega^{\dagger}N_{\mathrm{R}}$ an object can be constructed, which is invariant under isospin transformations.

Having the needed symmetry properties, one can work out the explicit couplings between the constituent fields of $\Omega = \Phi + \mathrm{i}\Lambda$ and the nucleons. The choice that Ω lives in a Euclidean space (positive metric $\delta^{\mu\nu} = \mathrm{diag}(1,1,1,1)$), rather than in a Minkowsky-like space ($\eta^{\mu\nu} = \mathrm{diag}(-1,1,1,1)$) is directly related to the fact that some parts of Ω show pseudo-scalar and other parts scalar behavior in $\mathbb{M}^{(3,1)}$. To see this, take the scalar field in its iso-spinor representation $\Omega_{ab} = \Omega^{\alpha} q^{\alpha}_{\mathrm{I}\,ab} = \Omega^0 I_{ab} - \mathrm{i}\boldsymbol{\Omega}\boldsymbol{\tau}_{ab}$[6)] and combine it with the left- and right-handed nucleon spinors. The iso-spinor indices are not strictly necessary at the moment, but are added here, as a reminder of the full structure contained in the scalar field, Ω. Before writing down the explicit fermion-scalar coupling, it is helpful to clarify on what spaces all involved operators act. There are the projection operators, P_{L} and P_{R}, which act only on spinors in $\mathbb{M}^{(3,1)}$ (here the Dirac-spinors, p and n) and there are the isospin operators I_{ab} and $\boldsymbol{\tau}_{ab}$. These act only on 2-dimensional iso-spinors, N. As all operators act on different spaces, they commute and knowing this, the expansion comes down to:

$$g_{\Omega}(\overline{N}_{\mathrm{L}}\Omega^{\dagger}N_{\mathrm{R}} + h.c.) = g_{\Omega}\sum_{\alpha}\left[\overline{N}^a_{\mathrm{L}}\,\Omega^{\alpha}\bar{q}^{\alpha}_{\mathrm{I}\,ab}\,N^b_R + \left(\overline{N}^a_{\mathrm{L}}\,\Omega^{\alpha}q^{\alpha}_{ab}\,N^b_{\mathrm{R}}\right)^{\dagger}\right] \tag{3.3.6}$$

$$= g_{\Omega}\overline{N}^a P_{\mathrm{R}}\left[\Omega^{\star\,0}I_{ab} + \mathrm{i}\boldsymbol{\Omega}^{\star}\boldsymbol{\tau}_{ab}\right]P_{\mathrm{R}}N^b + g_{\Omega}\overline{N}^a P_{\mathrm{L}}\left[\Omega^0 I_{ab} - \mathrm{i}\boldsymbol{\Omega}\boldsymbol{\tau}_{ab}\right]P_{\mathrm{L}}N^b \tag{3.3.7}$$

$$= g_{\Omega}\overline{N}^a\left[\Omega^{\star\,0}I_{ab} + \mathrm{i}\boldsymbol{\Omega}^{\star}\boldsymbol{\tau}_{ab}\right]\frac{1}{2}(\mathbb{1}+\gamma_5)N^b + g_{\Omega}\overline{N}^a\left[\Omega^0 I_{ab} - \mathrm{i}\boldsymbol{\Omega}\boldsymbol{\tau}_{ab}\right]\frac{1}{2}(\mathbb{1}-\gamma_5)N^b \tag{3.3.8}$$

$$= g_{\Omega}\overline{N}^a\Big\{I_{ab}\mathbb{1}\,\Re[\Omega^0] + \mathrm{i}\boldsymbol{\tau}_{ab}\gamma_5\,\Re[\boldsymbol{\Omega}] - \mathrm{i}I_{ab}\gamma_5\,\Im[\Omega^0] + \boldsymbol{\tau}_{ab}\mathbb{1}\,\Im[\boldsymbol{\Omega}]\Big\}N^b \tag{3.3.9}$$

$$= g_{\Omega}\overline{N}\Big\{(I\otimes\mathbb{1})\Re[\Omega^0] - \mathrm{i}(\boldsymbol{\tau}\otimes\gamma_5)\Re[\boldsymbol{\Omega}] - \mathrm{i}(I\otimes\gamma_5)\Im[\Omega^0] + (\boldsymbol{\tau}\otimes\mathbb{1})\Im[\boldsymbol{\Omega}]\Big\}N\,. \tag{3.3.10}$$

In the third line the commutativity of Minkowski- with $SU_{\mathrm{I}}(2)$-operators was used and in addition the basic properties of the projection operators, $P_{\mathrm{L,R}}$. After that, in the fourth line the definition of real and imaginary part were used ($\Re[Z] = (Z + Z^{\star})/2$ and $\Im[Z] = (Z - Z^{\star})/2$). The last line represents the mathematical notation for operators that act on different spaces. For matrices a and b the tensor product $a \otimes b$ can be understood as if every component of a is (scalar-) multiplied by the matrix b and so line 3.3.10 is the mathematically unique notation of line 3.3.9. Typically in physics, one does not bother with the notation of line 3.3.9 but writes for example in the linear σ-model $\overline{N}\left[\sigma + \mathrm{i}\boldsymbol{\tau}\gamma_5\boldsymbol{\pi}\right]N$ and the reader is

6) Here $\boldsymbol{\Omega} = (\Omega^1, \Omega^2, \Omega^3)^{\mathrm{T}}$ was used as a shorthand notation.

obliged to translate this colloquial expression into the meaning of equation 3.3.10.
Through the transformations of different spaces and the production of invariants maybe the focus of this section has fallen slightly out of sight. To remedy this, the derived interaction in terms of the 'fundamental' meson fields shall be given (in colloquial notation):

$$\mathcal{L}^{\text{eff}}_{N\Omega} = g_\Omega(\overline{N}_\text{L}\Omega^\dagger N_\text{R} + \overline{N}_\text{R}\Omega N_\text{L}) = g_\Omega\overline{N}\left[(\sigma - \mathrm{i}\gamma_5\eta) + \boldsymbol{\tau}(\boldsymbol{\delta} + \mathrm{i}\gamma_5\boldsymbol{\pi})\right]N \tag{3.3.11}$$

$$= g_\Omega\overline{N}\left[(\sigma + \boldsymbol{\tau\delta}) - \mathrm{i}\gamma_5(\eta - \boldsymbol{\tau\pi})\right]N\,. \tag{3.3.12}$$

This expression gives the complete coupling of nucleons to all mesons $\{\sigma, \eta, \boldsymbol{\pi}, \boldsymbol{\delta}\}$ under the assumption of an overall rotational symmetry in the space of the mesons. The coupling term is hermitian and invariant under chiral, Lorentz and isospin transformations. So, by now a model can be constructed that contains a kinetic term for nucleons and a coupling to the meson field, Ω.
In addition a mass term for nucleons is almost at hand. The coupling of the σ-meson and nucleon has the correct structure to qualify for a mass term. If one assumes that the σ-meson in the model contains of a static part, v, and a variation (in other words $\sigma(x) = v + \widetilde{\sigma}(x)$), then the σ-nucleon coupling becomes:

$$g_\Omega\overline{N}(v + \widetilde{\sigma})N = g_\Omega v\overline{N}N + g_\Omega\overline{N}\widetilde{\sigma}N\,. \tag{3.3.13}$$

The dependence on space-time of N and $\widetilde{\sigma}$ has been suppressed to simplify the notation. Only the Ω-nucleon coupling, g_Ω, and the static v are constants. So, by this expansion the nucleon acquires the mass, $M_\text{N} = g_\Omega v$. If also the third component of the $\boldsymbol{\delta}$-meson has a non-vanishing VEV ($\delta_3(x) = v_\delta + \widetilde{\delta}_3(x)$) in a physical realisation, then one is in the peculiar situation, that the iso-spinor, N, has different masses in the first and second component, since δ_3 couples to the third Pauli matrix in iso-spinor space. How the expansion of the σ- and potentially δ_3-field comes about will be discussed, once the scalar Lagrangian will be investigated closely in Sec. III.4.
Gernerally the connection between the nucleon mass and the the σ-nucleon coupling leads to the Goldenberger-Treimann relation. This is nicely presented in Ref. [34, p.126-128] or Ref. [5, p.516-523]. Finally, combining the kinetic and interaction term, the full effective Lagrangian becomes:

$$\mathcal{L}^{\text{eff}}_{N} = \overline{N}\mathrm{i}\gamma_\mu D^\mu N + g_\Omega[\overline{N}_\text{R}\Omega N_\text{L} + \overline{N}_\text{L}\Omega^\dagger N_\text{R}] \tag{3.3.14}$$

$$= \overline{N}[\mathrm{i}\gamma_\mu\partial^\mu + g_\text{A}\gamma_\mu A^\mu]N + g_\Omega[\overline{N}_\text{R}\Omega N_\text{L} + \overline{N}_\text{L}\Omega^\dagger N_\text{R}]\,. \tag{3.3.15}$$

Now one is in the position, to clarify the postponed issue concerning Eq. (3.2.13). There it was stated, that the fermionic zero-mode contribution leads to the term $\mathcal{L}_\text{inst} \propto \left(e^{-\mathrm{i}\theta}\Omega_\alpha^2 + e^{+\mathrm{i}\theta}\Omega_\alpha^{\dagger\,2}\right)$ and it wasn't further explained, why the instanton contribution comes with the conjugate scalar field and the anti-instanton field with the ordinary field, Ω. Having the explicit nucleon Lagrangian (equation Eq. (3.3.14)), the answer to this question is at hand. For this remember how the zero-mode determinant was calculated. In Sec. II.9.8.1 the approximate fermionic zero-modes for anti-instantons were derived. As the anti-instantons couple to fermions of the $SU_\mathrm{B}(2)$ subgroup, or equivalently to left-handed fermions in Minkowski space, this means that the zero-modes of anti-instantons are left-handed. Therefore, with an eye on Eq. (3.3.14), one finds, that anti-instantons lead to the ordinary Ω field in $\mathcal{L}_\mathrm{inst}$. Correspondingly instantons have right-handed zero-modes and thus their contribution leads to the conjugate field, $\Omega^\dagger$, in $\mathcal{L}_\mathrm{inst}$.

III.3.3 Nucleon currents

This subsection will only give the first contribution to the complete currents of the model. While other parts of the model will be derived the corresponding current contributions will be given along the way.
In the introductory (Sec. III.1) it was stated that the model should be invariant under the Lorentz group, $SU_\mathrm{L}(2) \otimes SU_\mathrm{R}(2)$. In addition the whole model contains a $U_\mathrm{V}(1)$ symmetry, which treats left- and right-handed parts equally. This symmetry will be called $U_\mathrm{V}(1)$, where the V stands for vectorial. The invariance of the fermionic Lagrangian under this symmetry can be seen directly from Eq. (3.3.4)[7)]. In the following derivations this symmetry will mostly be left out, since its inclusion is conceptually much easier than the $SU(2)$ symmetries - the $U_\mathrm{V}(1)$ transformation simply multiplies the same complex phase to every fermionic spinor.
Noether's theorem implies, that any field which transforms non-trivially under the symmetry and fulfils the Euler-Lagrange equation leads to a conserved current (for a derivation of the theorem compare the original publication Ref. [10] or in Sec. II.3). The conserved currents connected to the model Lagrangian can be deduced by using the general relation Eq. (2.3.9):

$$\sum_a \varepsilon_a j_a^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu A^\alpha)} \delta_\varepsilon A^\alpha \,, \tag{3.3.16}$$

where A^α refers to any involved field and ε represents a small variation in direction a in the symmetry space. So far these fields are only the left- and right-handed nucleon fields. The only part missing to determine the conserved current is the variation of the nucleon fields under the Lorentz group. In Sec. II.1.1 the transformations of left-handed spinors have been presented to be $N'_\mathrm{L} = U_\mathscr{L}^{-1} N_\mathrm{L} U_\mathscr{L} = L N_\mathrm{L}$, where $U_\mathscr{L}$ was a general Lorentz transformation and $L = L_{(\alpha_\ell)} = e^{-\mathrm{i}\alpha_\ell \boldsymbol{\tau}/2}$ an element of the $SU_\mathrm{L}(2)$ group[8)]. Analogously the right-handed transformed field transformed into $N'_\mathrm{R} = R N_\mathrm{R}$. Using the

[7)] In complete chiral models there is in fact another $U_\mathrm{A}(1)$ symmetry, which treats left- and right-handed parts differently, but this symmetry will be explicitly broken once instantons are included in the picture (compare Sec. II.8.4 and later Sec. III.5.2).

[8)] The actual equations in Sec. II.1.1 looked slightly different, but the connection can be made by noting that the χ_a fields are left-handed and the $\xi^{\dot{a}}$ right-handed.

transformations for the nucleon fields the variations under the Lorentz group can be determined:

$$\delta_{\varepsilon_\ell} N_{\mathrm{L}} = L_{(\varepsilon_\ell)} N_{\mathrm{L}} - N_{\mathrm{L}} = -\frac{\mathrm{i}}{2}\varepsilon_\ell^a \tau^a N_{\mathrm{L}}\,, \tag{3.3.17}$$

$$\delta_{\varepsilon_r} N_{\mathrm{R}} = R_{(\varepsilon_r)} N_{\mathrm{R}} - N_{\mathrm{R}} = -\frac{\mathrm{i}}{2}\varepsilon_r^a \tau^a N_{\mathrm{R}}\,. \tag{3.3.18}$$

Here the exponential has been expanded in a power series and truncated after the first term as ε can be assumed to be small[9)]. Now, using the kinetic part of the nucleon Lagrange density (Eq. (3.3.3)), the components of the left- and right-handed nucleon currents are at hand:

$$j^a_{\mathrm{L}\,\mu} = \overline{N}_{\mathrm{L}} \mathrm{i}\gamma_\mu \left(-\frac{\mathrm{i}}{2}\tau^a N_{\mathrm{L}}\right) = \frac{1}{2}\overline{N}_{\mathrm{L}}\gamma_\mu \tau^a N_{\mathrm{L}}\,, \tag{3.3.19}$$

$$j^a_{\mathrm{R}\,\mu} = \overline{N}_{\mathrm{R}} \mathrm{i}\gamma_\mu \left(-\frac{\mathrm{i}}{2}\tau^a N_{\mathrm{R}}\right) = \frac{1}{2}\overline{N}_{\mathrm{R}}\gamma_\mu \tau^a N_{\mathrm{R}}\,. \tag{3.3.20}$$

Here the factors of $\varepsilon^a_{\ell/r}$ have been left out. They are arbitrary and appear on both sides. Therefore, in order to preserve the equality the prefactors of each ε^a have to match. This is exactly guaranteed by the above equations.
The left- and right-handed currents can be reexpressed into a vectorial and an axial part, as these expressions can nicely be identified with physical observables. To do this, the $U(1)$ connection between of vectorial/axial vectors and left/right elements can be employed (Eq. (2.5.8)), giving the two linear combinations:

$$\boldsymbol{\varepsilon}_\ell = \boldsymbol{\varepsilon}_{\mathrm{V}} - \boldsymbol{\varepsilon}_{\mathrm{A}} \qquad , \qquad \boldsymbol{\varepsilon}_r = \boldsymbol{\varepsilon}_{\mathrm{V}} + \boldsymbol{\varepsilon}_{\mathrm{A}}\,. \tag{3.3.21}$$

From this the vector and axial currents become:

$$\varepsilon^a_\ell j^a_{\mathrm{L}\,\mu} + \varepsilon^a_r j^a_{\mathrm{R}\,\mu} = \frac{1}{2}\Big(\overbrace{(\varepsilon^a_{\mathrm{V}} - \varepsilon^a_{\mathrm{A}})}^{\varepsilon^a_\ell}\overline{N}_{\mathrm{L}}\gamma_\mu\tau^a N_{\mathrm{L}} + \overbrace{(\varepsilon^a_{\mathrm{V}} + \varepsilon^a_{\mathrm{A}})}^{\varepsilon^a_r}\overline{N}_{\mathrm{R}}\gamma_\mu\tau^a N_{\mathrm{R}}\Big) \tag{3.3.22}$$

$$= \frac{1}{2}\left(\varepsilon^a_{\mathrm{V}}\left(-\overline{N}_{\mathrm{L}}\gamma_\mu\tau^a N_{\mathrm{L}} + \overline{N}_{\mathrm{R}}\gamma_\mu\tau^a N_{\mathrm{R}}\right) + \varepsilon^a_{\mathrm{A}}\left(\overline{N}_{\mathrm{L}}\gamma_\mu\tau^a N_{\mathrm{L}} - \overline{N}_{\mathrm{R}}\gamma_\mu\tau^a N_{\mathrm{R}}\right)\right) \tag{3.3.23}$$

$$= \varepsilon^a_{\mathrm{V}}\underbrace{\overline{N}\gamma_\mu\frac{\tau^a}{2}N}_{=:j^a_{\mathrm{V}\,\mu}} + \varepsilon^a_{\mathrm{A}}\underbrace{\overline{N}\gamma_\mu\gamma_5\frac{\tau^a}{2}N}_{=:j^a_{\mathrm{A}\,\mu}}\,. \tag{3.3.24}$$

In the last line the left- and right-handed components have been reassembled into the complete Dirac-spinor, $N = N_{\mathrm{L}} + N_{\mathrm{R}}$. In the axial current the $\gamma_5 = \mathrm{diag}(-I, I)$ matrix corrects for the extra minus sign in the left-handed contribution (compare line Eq. (3.3.23) and for the definition of γ_5 in Eq. (2.4.10)).
Before ending this section two direct examples of the conserved currents shall be given, as they will be

[9)] A general reason for using only the truncated exponentials in the context of group transformations can be found in Sec. B.5.1.

important once the scalar currents are introduced in Sec. III.4.1 and III.4.1.1. As both, the vectorial and the axial nucleon currents are conserved, one could for example focus on the third vectorial component:

$$(j_{\mathrm{V}}^3)_\mu = \overline{N}\gamma_\mu \frac{\tau^3}{2} N = \frac{1}{2} N^\dagger \gamma_0 \gamma_\mu \tau^3 N = \frac{1}{2}(p^\dagger \gamma_0 \gamma_\mu p - n^\dagger \gamma_0 \gamma_\mu n)\,. \tag{3.3.25}$$

Of course, giving the explicit forms of the remaining two vectorial components, $j_{\mathrm{V}\,\mu}^{\pm}$, and the complete axial current, $\boldsymbol{j}_{\mathrm{A}\,\mu}$, is also possible. They are left out here, as these currents will not be conserved, once the scalar current contribution is calculated. The other conserved current of interest is connected to the $U_{\mathrm{V}}(1)$ symmetry and was not explicitly derived previously. But as it only multiplies a complex phase to the nucleon spinors, $N' = e^{-\mathrm{i}\varphi}N$, its current, j_{B}^μ, follows almost directly:

$$(j_{\mathrm{B}})_\mu = \overline{N}\gamma_\mu\, I\, N = (p^\dagger \gamma_0 \gamma_\mu p + n^\dagger \gamma_0 \gamma_\mu n)\,. \tag{3.3.26}$$

Now following the lines of thought from Sec. II.3 both currents can be included in seperate continuity equations (Eq. (2.3.12)), giving the third component of the conserved vectorial charge and the $U_{\mathrm{V}}(1)$ charge:

$$0 = \partial_t \int \mathrm{d}^4x (j_{\mathrm{V}}^3)_0 = \partial_t \frac{1}{2} \int \mathrm{d}^4x (p^\dagger p - n^\dagger n) \equiv \partial_t Q_{\mathrm{V}}^3\,, \tag{3.3.27}$$

$$0 = \partial_t \int \mathrm{d}^4x (j_{\mathrm{B}})_0 = \partial_t \int \mathrm{d}^4x (p^\dagger p + n^\dagger n) \equiv \partial_t Q_{\mathrm{B}}\,. \tag{3.3.28}$$

Essentially the first equality means that the difference of protons and neutrons in the complete volume is preserved and the second equality enforces a conservation of the corresponding sum (therefore it is called baryon number, Q_{B}). Together, one finds that the number of protons and the number of neutrons is conserved separately in this model.
The charged vectorial 'charges', $Q_{\mathrm{V}}^{\pm}$, and the axial 'charges', Q_{A}^{a}, are not given, as the corresponding currents are not conserved in the complete model, and thus their charge conservation does not hold either. For the moment this ends the discussion of conserved currents. The topic will be picked up again once the scalar Lagrangian is developed.

III.4 The scalar Lagrangian

In Sec. III.3.2 the coupling between all scalar mesons and nucleons has been established. This section will focus on the details of the purely scalar contributions, $\mathcal{L}_{\Omega}^{\mathrm{eff}}$. The concepts that are used actually show a great similarity to the ideas from the normal 'linear σ-model'. The derivation of this model is presented in a very educative fashion in Ref. [34, p.111-128] and most techniques applied in this section are very similar. In general the model shall be expanded around a vacuum configuration and thus the first goal is to give a coherent representation of the model in a mean field context. For a start take the following Lagrange density:

$$\mathcal{L}_\Omega = \frac{1}{2}||\partial_\mu\Omega||^2 - \widetilde{\mathcal{V}}\left(||\Omega||^2\right) = \frac{1}{2}\mathrm{tr}_\mathrm{I}\left[(\partial_\mu\Omega)^\dagger(\partial^\mu\Omega)\right] - \left(\frac{\mu^2}{2}||\Omega||^2 + \frac{\lambda^2}{4}||\Omega||^4\right). \tag{3.4.1}$$

Here, just as before, the norm is defined via the quaternion definition (2.2.6), giving: $||\Omega||^2 = \mathrm{tr}_\mathrm{I}\left[\Omega^\dagger\Omega\right]/2$ and the scalar field is given as in Sec. III.2.1:

$$\Omega = (\Phi^\alpha + \mathrm{i}\Lambda^\alpha)q_\mathrm{I}^\alpha \quad , \quad q_\mathrm{I}^\alpha = (I, -\mathrm{i}\boldsymbol{\tau}^\mathrm{T})^\mathrm{T}\,, \tag{3.4.2}$$

$$\Phi^\alpha = (\sigma, \boldsymbol{\pi}^\mathrm{T})^\mathrm{T} \quad , \quad \Lambda^\alpha = (\eta, \boldsymbol{\delta}^\mathrm{T})^\mathrm{T}\,, \tag{3.4.3}$$

$$\Rightarrow ||\Omega||^2 = \sigma^2 + \boldsymbol{\pi}^2 + \eta^2 + \boldsymbol{\delta}^2\,. \tag{3.4.4}$$

Up to now the given Lagrange density has a global $SO(8)$ symmetry. This is why a factor of $1/2$ was included in the kinetic term[10)]. The included potential has the 'Mexican hat' form and is usually used for spontanious symmetry breaking[11)]. In order to produce a well defined model one needs $\lambda^2 > 0$. If $\mu^2 > 0$ as well, then one has a unique vacuum state and nothing interesting happens. However if $\mu^2 < 0$ is chosen, then the vacuum state does not preserve all the symmetries of the potential and one gets the typical Mexican hat shape. In addition this choice for μ means that the original scalar fields are massless, as the quadratic term in the Lagrangian comes with the wrong sign. Before performing the usual expansion around the minimum some modifications of the above potential shall be introduced.

As the given Lagrange density depends only on the modulus of the scalar field, Ω, it follows that this model is invariant under rotations in the 4-dimensional complex vector-space, or equivalently in the corresponding 8-dimensional real vector-space[12)]. This nice symmetry is disturbed in the effective model due to two reasons, which are connected to the interactions with nucleons. The first disturbance has its origin in the interaction of the scalar fields with the QCD instanton background, that was introduced in Sec. II.8 and II.9. In Sec. III.2 it was qualitatively discussed that the zero-mode contribution from the instanton sector gives rise to the additional contribution in the scalar potential, $\mathcal{L}_\mathrm{inst}$ (compare Eq. (3.2.13)). As the explicit structure of $\Omega = \Phi + \mathrm{i}\Lambda$ was discussed in Sec. III.2.1 this can now be used, to

10)If one works with complex fields this factor usually shows up in the field definition, but as it is a mere convention in all following derivations, it is given explicitly (as it is standard while dealing with real fields).

11)The spontanious symmetry breaking will be worked out here in a special case but for a general introduction see Ref. [34, p.119-128], Ref. [5, p.188-202;538-541].

12)n-dimensional complex fields can be exchanged for 2n-dimensional real fields, just as in the typical analysis context.

write down the explicit contribution of the instanton sector:

$$\mathcal{L}_{\text{inst}} = \frac{a^2}{4}\Big(e^{-\mathrm{i}\theta}\Omega_\alpha^2 + e^{+\mathrm{i}\theta}\Omega_\alpha^{\star\,2}\Big), \tag{3.4.5}$$

$$= \frac{a^2}{4}\Big(e^{-\mathrm{i}\theta}(\Phi_\alpha + \mathrm{i}\Lambda_\alpha)^2 + e^{+\mathrm{i}\theta}(\Phi_\alpha - \mathrm{i}\Lambda_\alpha)^2\Big), \tag{3.4.6}$$

$$= \frac{a^2}{2}\Big(\cos(\theta)(\Phi_\alpha^2 - \Lambda_\alpha^2) - 2\sin(\theta)\Phi_\alpha\Lambda_\alpha\Big). \tag{3.4.7}$$

Here θ is the vacuum angle from Sec. II.8.4. As mentioned earlier, there is experimental evidence that this angle is very close to zero, as any other value would lead to CP violating terms in strong interactions. Therefore, it will be assumed that it actually is zero from now on ($\theta \equiv 0$). With this the correction arising from instantons to the scalar potential becomes:

$$\mathcal{V}_{\text{inst}} = \frac{a^2}{2}\left(\|\Lambda\|^2 - \|\Phi\|^2\right). \tag{3.4.8}$$

where the exact form of a can be found by solving Eq. (3.2.9) exactly. As $\mathcal{V}_{\text{inst}}$ treats the Φ and the Λ part of the scalar field, Ω, differently, it explicitly breaks the overall $SO(8)$ rotational symmetry down to 2 decoupled $SO(4)$ rotational symmetries (one for the $\|\Phi\|$ part and the other for the $\|\Lambda\|$ part).
The other symmetry spoiling contribution arises from the nucleon background in a vacuum configuration. If a vacuum configuration which contains a certain density of protons, n_p, and neutrons, n_n, shall be described, then this background acts as a source term in the equations of motion for all neutral scalar mesons, namely the σ- and δ_3-meson. All other mesons will not be changed, if the vacuum configuration of the system is parity even, charge neutral and not direction dependent. In a potential such a source can be represented by introducing a linear term in the field with a suitable arbitrary, but fixed parameter. This means the nucleon background gives rise to the symmetry breaking contribution:

$$\mathcal{V}_{\text{nucl}} = -\alpha\sigma - \beta\delta_3\,, \tag{3.4.9}$$

$$\alpha \sim \langle \overline{N} I N\rangle = \langle(\overline{p}p + \overline{n}n)\rangle\,, \tag{3.4.10}$$

$$\beta \sim \langle \overline{N}\tau_3 N\rangle = \langle(\overline{p}p - \overline{n}n)\rangle\,. \tag{3.4.11}$$

Here the parameters α and β correspond to the source influences. They will be adjusted later in this section, by forcing the model to realise the effective physical nucleon masses and certain meson masses. With this the variations of the original potential are complete and one now has:

$$\mathcal{V} = \widetilde{\mathcal{V}} + \mathcal{V}_{\text{inst}} + \mathcal{V}_{\text{nucl}} \tag{3.4.12}$$

$$= \frac{\mu^2 - a^2}{2}\|\Phi\|^2 + \frac{\mu^2 + a^2}{2}\|\Lambda\|^2 + \frac{\lambda^2}{4}\left(\|\Phi\|^2 + \|\Lambda\|^2\right)^2 - \alpha\sigma - \beta\delta_3\,. \tag{3.4.13}$$

To shorten the notation it is useful to replace $\mu_\pm^2 = \mu^2 \pm a^2$. In Fig. 3.4.1 an example of the effective scalar potential in the (σ, δ_3)-plane is visualised for extreme parameter values. Originally (setting $a = \alpha = \beta = 0$) one has the usual 'Mexican hat'-potential (App. A.1a), which is then deformed into the present shape. To see the influence of the parameters a, α and β, there are further examples in App. A.9.

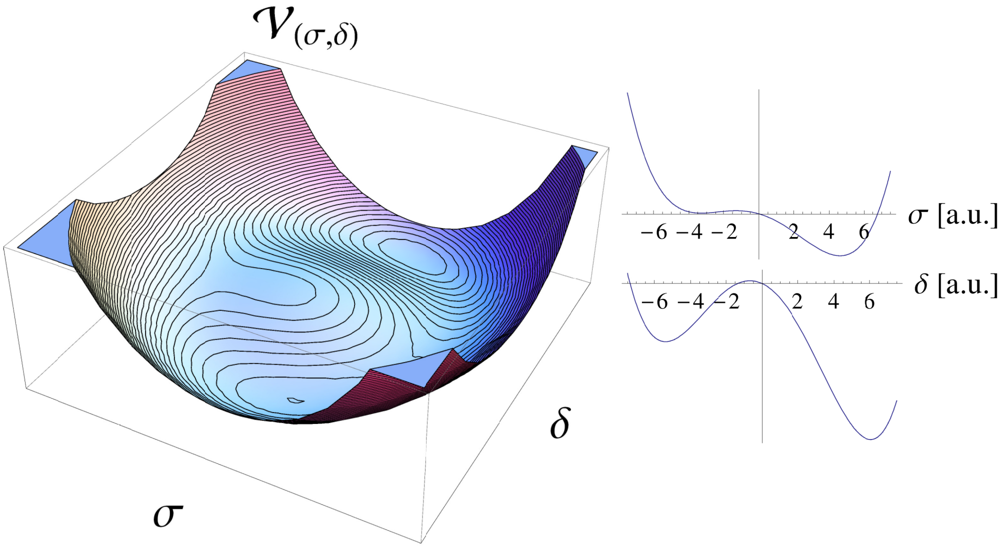


Figure 3.4.1: Example potential for extreme parameter values. The units are arbitrary and are just included for reasons of comparability.

Having the potential, now the expansion around the minimum configuration has to be obtained. Due to the nucleon source terms the vacuum expectation value (VEV) has only non-vanishing contributions in the σ- and δ_3-direction and so the derivation will be focused on these directions. The conditions on the minimum are most readily obtained in spherical coordinates in the (σ, δ_3)-plane:

$$\sigma = r \cdot \cos(\varphi) \quad , \quad \delta_3 = r \sin(\varphi) \, , \tag{3.4.14}$$

$$r^2 = \sigma^2 + \delta_3^2 \quad , \quad \varphi = \arctan\left(\frac{\delta_3}{\sigma}\right) . \tag{3.4.15}$$

Using this, the potential reads:

$$\mathcal{V} = \frac{\mu_-^2}{2}\left[r^2\cos^2(\varphi) + \boldsymbol{\pi}^2\right] + \frac{\mu_+^2}{2}\left[r^2\sin^2(\varphi) + \delta_1^2 + \delta_2^2 + \eta^2\right] + \frac{\lambda^2}{4}\Big[r^2 + \overbrace{\boldsymbol{\pi}^2 + \delta_1^2 + \delta_2^2 + \eta^2}^{=:F^2}\Big]^2$$
$$- \alpha r\cos(\varphi) - \beta r\sin(\varphi)\,. \tag{3.4.16}$$

To find the minimum take $\pi_j = \eta = \delta_1 = \delta_2 = 0$ and search for the points, where the derivative (actually the 2-dimensional gradient) vanishes:

$$\boldsymbol{\nabla}_2 = \left(\frac{\partial}{\partial r}, \frac{1}{r}\frac{\partial}{\partial\varphi}\right)^{\mathrm{T}}\,, \tag{3.4.17}$$

$$\Rightarrow \frac{\partial V}{\partial r} = \left[\mu_-^2\cos^2(\varphi) + \mu_+^2\sin^2(\varphi)\right]r + \lambda^2\|\Omega\|^2 r - \alpha\cos(\varphi) - \beta\sin(\varphi)\,, \tag{3.4.18}$$

$$\Rightarrow \frac{1}{r}\frac{\partial V}{\partial\varphi} = \left[\mu_+^2 - \mu_-^2\right]r\cos(\varphi)\sin(\varphi) + \alpha\sin(\varphi) - \beta\cos(\varphi)\,. \tag{3.4.19}$$

Suppose that the minimum lies at the coordinates, (r_0, ϕ_0). At the moment it is not helpful to give a more explicit form of the minimum, as all other parameters of the model aren't fixed anyway. For a suitable choice of parameters (in fact any choice that is of interest later) one can ensure that the extremum is an actual minimum. Then, in the minimum, one has $\|\Omega\|_{min}^2 = r_0^2$ and with this, the above equations translate to the following two conditions:

$$0 = \left[\mu_-^2\cos^2(\varphi_0) + \mu_+^2\sin^2(\varphi_0)\right]r_0 + \lambda^2 r_0^3 - \alpha\cos(\varphi_0) - \beta\sin(\varphi_0)\,, \tag{3.4.20}$$

$$0 = \left[\mu_+^2 - \mu_-^2\right]r_0\cos(\varphi_0)\sin(\varphi_0) + \alpha\sin(\varphi_0) - \beta\cos(\varphi_0)\,. \tag{3.4.21}$$

Going back to cartesian coordinates

$$R_\sigma \equiv r_0\cos(\varphi_0) \qquad , \qquad R_\delta \equiv r_0\sin(\varphi_0)\,. \tag{3.4.22}$$

the minimum conditions can be rewritten, by multiplying with r_0, to a form which will be convenient later:

$$0 = \mu_-^2 R_\sigma^2 + \mu_+^2 R_\delta^2 + \lambda^2 r_0^4 - \alpha R_\sigma - \beta R_\delta\,, \tag{3.4.23}$$

$$0 = \left[\mu_+^2 - \mu_-^2\right]R_\sigma R_\delta + \alpha R_\delta - \beta R_\sigma\,. \tag{3.4.24}$$

Now the potential can be expanded around the minimum, $\boldsymbol{R} = (R_\sigma, R_\delta)^{\mathrm{T}}$, by defining the variations, $\boldsymbol{\rho} = (\rho_\sigma, \rho_\delta)^{\mathrm{T}}$, and plugging the expansion $(\boldsymbol{R} + \boldsymbol{\rho})$ back into the potential (Eq. (3.4.13)). The fluctuations in the other directions are included as well and one gets:

$$\begin{aligned}\mathcal{V} &= \frac{\mu_-^2}{2}\Big[(R_\sigma+\rho_\sigma)^2+\boldsymbol{\pi}^2\Big]+\frac{\mu_+^2}{2}\Big[(R_\delta+\rho_\delta)^2+\delta_1^2+\delta_2^2+\eta^2\Big]\\ &+\frac{\lambda^2}{4}\Big[(R_\sigma+\rho_\sigma)^2+(R_\delta+\rho_\delta)^2+F^2\Big]^2-\alpha(R_\sigma+\rho_\sigma)-\beta(R_\delta+\rho_\delta) \qquad (3.4.25)\\ &= \frac{\mu_-^2}{2}\Big[R_\sigma^2+2R_\sigma\rho_\sigma+\rho_\sigma^2+\boldsymbol{\pi}^2\Big]+\frac{\mu_+^2}{2}\Big[R_\delta^2+2R_\delta\rho_\delta+\rho_\delta^2+\delta_1^2+\delta_2^2+\eta^2\Big]\\ &+\frac{\lambda^2}{4}\Big[r_0^2+\boldsymbol{\rho}^2+2(R_\sigma\rho_\sigma+R_\delta\rho_\delta)+F^2\Big]^2-\alpha(R_\sigma+\rho_\sigma)-\beta(R_\delta+\rho_\delta)\,. \qquad (3.4.26)\end{aligned}$$

In the last line it was used that $\boldsymbol{R}^2=r_0^2$. The quadric term can now be simplified alone to:

$$\begin{aligned}\frac{\lambda^2}{4}\Big[...\Big]^2 &= \frac{\lambda^2}{4}\Big\{r_0^4+\boldsymbol{\rho}^4+4(R_\sigma^2\rho_\sigma^2+R_\delta^2\rho_\delta^2+2R_\sigma R_\delta\rho_\sigma\rho_\delta)+F^4\\ &+2r_0^2\Big[\boldsymbol{\rho}^2+F^2+2(R_\sigma\rho_\sigma+R_\delta\rho_\delta)\Big]+4(R_\sigma\rho_\sigma+R_\delta\rho_\delta)\Big[\boldsymbol{\rho}^2+F^2\Big]+2\boldsymbol{\rho}^2F^2\Big\} \qquad (3.4.27)\\ &= \frac{\lambda^2 r_0^4}{4}+\frac{\lambda^2}{4}\Big[\boldsymbol{\rho}^2+F^2\Big]^2+\lambda^2(R_\sigma\rho_\sigma+R_\delta\rho_\delta)\Big[\boldsymbol{\rho}^2+F^2\Big]\\ &+\lambda^2(R_\sigma^2\rho_\sigma^2+R_\delta^2\rho_\delta^2+2R_\sigma R_\delta\rho_\sigma\rho_\delta)+\frac{\lambda^2}{2}r_0^2\Big[\boldsymbol{\rho}^2+F^2\Big]\\ &+\lambda^2 r_0^2(R_\sigma\rho_\sigma+R_\delta\rho_\delta)\,. \qquad (3.4.28)\end{aligned}$$

In this rather ugly looking expression the real trick of spontanious symmetry breaking is hidden[13)]. Through the quadric interaction it is possible to couple the varying fields quadratically to the VEV of the model. Thus the special shape of the potential allows an originally massless field to acquire a mass in the transition to a stable vacuum configuration. Now the equations Eq. (3.4.26) and (3.4.28) can be combined and ordered in powers of the various fields by using $\boldsymbol{\rho}^2=\rho_\sigma^2+\rho_\delta^2$.

$$\begin{aligned}\mathcal{V} &= \frac{\lambda^2}{4}\Big[\rho_\sigma^2+\rho_\delta^2+F^2\Big]^2+\lambda^2(R_\sigma\rho_\sigma+R_\delta\rho_\delta)\Big[\rho_\sigma^2+\rho_\delta^2+F^2\Big]\\ &+\frac{1}{2}\Big[\mu_-^2+\lambda^2r_0^2\Big]\boldsymbol{\pi}^2+\frac{1}{2}\Big[\mu_+^2+\lambda^2r_0^2\Big](\delta_1^2+\delta_2^2+\eta^2)\\ &+\frac{1}{2}\Big[\mu_-^2+\lambda^2r_0^2+2\lambda^2R_\sigma^2\Big]\rho_\sigma^2+\frac{1}{2}\Big[\mu_+^2+\lambda^2r_0^2+2\lambda^2R_\delta^2\Big]\rho_\delta^2\\ &+2\lambda^2R_\sigma R_\delta\rho_\sigma\rho_\delta\\ &+\Big[\mu_-^2R_\sigma+\lambda^2r_0^2R_\sigma-\alpha\Big]\rho_\sigma+\Big[\mu_+^2R_\delta+\lambda^2r_0^2R_\delta-\beta\Big]\rho_\delta\,. \qquad (3.4.29)\end{aligned}$$

Before proceding any further the minimum conditions (Eq. (3.4.20) and (3.4.21)) can be used to get rid of the last line. This line has to vanish as the expansion is performed around the minimum of the

13)The reason for the ugliness actually come from the introduced symmetry breaking contributions. In the pure Mexican hat potential the situation is much cleaner.

potential. Around an extremum the linear variation in the expansion variables, ρ_σ and ρ_δ, vanishes and so this is a nice check for the previous calculations. By multiplying Eq. (3.4.23) with R_δ and adding Eq. (3.4.24) times R_σ one obtains:

$$0 = (3.4.23) \cdot R_\delta + (3.4.24) \cdot R_\sigma = \mu_-^2 R_\delta (R_\sigma^2 - R_\sigma^2) + \mu_+^2 R_\delta (R_\delta^2 + R_\sigma^2) + \lambda^2 r_0^4 R_\delta - \beta(R_s^2 + R_\sigma^2)\,, \quad (3.4.30)$$

$$0 = r_0^2 \Big[\mu_+^2 R_\delta + \lambda^2 r_0^2 R_\delta - \beta\Big]\,. \quad (3.4.31)$$

And thus the coefficient in front of ρ_δ vanishes as required. Similarly the coefficient for ρ_σ vanishes by calculating:

$$(3.4.23) \cdot R_\sigma - (3.4.24) \cdot R_\delta = r_0^2 \Big[\mu_-^2 R_\sigma + \lambda^2 r_0^2 R_\sigma - \alpha\Big] = 0\,. \quad (3.4.32)$$

Now all the work is completed and some useful replacements are at hand. The parameters in front of the quadratic terms are identified with the corresponding mass terms, ρ_j is identified with the actual physical meson, j, and the tilde means that the corresponding field is the physical field this time:

$$\widetilde{\sigma} \equiv \rho_\sigma \quad , \quad \widetilde{\delta_3} \equiv \rho_\delta\,, \quad (3.4.33)$$

$$m_\sigma^2 := [\mu_-^2 + \lambda^2(r_0^2 + 2R_\sigma^2)] \quad , \quad m_\delta^2 := [\mu_+^2 + \lambda^2(r_0^2 + 2R_\delta^2)]\,, \quad (3.4.34)$$

$$m_\pi^2 := [\mu_-^2 + \lambda^2 r_0^2] \quad , \quad m_{\delta\eta}^2 := [\mu_+^2 + \lambda^2 r_0^2]\,, \quad (3.4.35)$$

$$\alpha = m_\pi^2 R_\sigma \quad , \quad \beta = m_{\delta\eta}^2 R_\delta\,. \quad (3.4.36)$$

The last line was just added for later convenience. It is just a rewriting of Eq. (3.4.32) and (3.4.31) and relates the symmetry breaking factors with the pion and the pseudo η mass. In the original σ-model R_σ would correspond to the pion decay constant f_π. More on this connection can be found for example in Ref. [34, p.126-128].

For calculations it may be useful to reexpress the cartesian VEV's back into spherical coordinates: $R_\sigma = r_0 \cos(\varphi_0)$ and $R_\delta = r_0 \sin(\varphi_0)$. Note that there is already a flaw apparent in the definitions of the mass terms. In Eq. (3.4.35) the masses of the η, δ_1 and δ_2 mesons are all forced to the same value, which is simply incorrect in the vacuum. This problem is lessened if one trades the η- for the η'-meson, but the difference is still large enough to pose difficulties in later parameter fixings (compare Sec. III.4.3.1). In addition, to make matters worse, the mass of δ_3 differs from the other δ-meson masses, which directly spurns all concepts of representation theory. So in total this seems like a bad idea, if it wasn't for two different ways out: First the violation of the symmetries in the potential was motivated from background vacuum configurations in the nucleon sector. Now if, as stated in the beginning of the section, the vacuum is charge neutral and parity-even, then all charged and parity-odd meson contributions will vanish anyway. So to say the evident flaws of the model are neatly swept below the rug in the vacuum configuration. Problems that are related to leaving this vacuum configuration are very intersting, but unfortunately beyond the scope of this work.

The second way to approach the problem is to abandon the strict enforcement of the η-meson mass. As the final model shall give an effective description of nucleons rather than meson it is somewhat more useful to fix all parameters of the model such that the effective nucleon masses in nuclear matter are reproduced. In Sec. III.4.3 all relevant relations for this are derived and discussed. It will turn out that it is sufficient to change the mass, $m_{\delta\eta}$, such that it is close to m_δ in order to meet the requirements of physical nucleon masses.

For now the potential can be rewritten in terms of the physical meson fields and the so far undetermined parameters:

$$\begin{aligned}\mathcal{V} =& \frac{1}{2}\Big[m_\sigma^2\widetilde{\sigma}^2 + m_\delta^2\widetilde{\delta}_3^2 + m_\pi^2\boldsymbol{\pi}^2 + m_{\delta\eta}^2(\delta_1^2 + \delta_2^2 + \eta^2)\Big] \\ &+ \lambda^2(R_\sigma\widetilde{\sigma} + R_\delta\widetilde{\delta}_3)\|\widetilde{\Omega}\|^2 + \frac{\lambda^2}{4}\|\widetilde{\Omega}\|^4 + 2\lambda^2 R_\sigma R_\delta\widetilde{\sigma}\widetilde{\delta}_3\,. \end{aligned} \tag{3.4.37}$$

While the correct notion for the physical fields is given above (with the tilde), this notation is dropped after this section again, as it does not reveal further insights. In comparison with the perturbed linear σ-model the last term in Eq. (3.4.37) is conceptionally new. This term comes from the double expansion in σ and δ_3 direction and gives a direct coupling between the fields with non vanishing VEV. Apart from this it should be mentioned that the limit of the normal perturbed σ-model (without the contribution from Λ) is completely regained if the VEV lies exactly in the σ-direction, corresponding to $\varphi_0 \to 0$ (and of course taking $\Lambda \to 0$).

Having a second look at Eq. (3.4.37), the last term is somewhat disturbing, as it mixes σ and δ_3 linearly. This leads to an effective coupling in the quadratic terms for the scalar fields, as will be shown in a moment. The mixing indicates, that one might not have chosen the most convenient set of variables and that it is potentially possible to define a linear combination of σ and δ_3, which then decouples the quadratic mass terms of the fields. To find this linear combination, one has to insert the definitions from the equations Eq. (3.4.34) and (3.4.35) into the potential (Eq. (3.4.37)), leading to:

$$\begin{aligned}\mathcal{V} =& \frac{1}{2}\Big[m_\pi^2(\sigma^2 + \boldsymbol{\pi}^2) + m_{\delta\eta}^2(\delta_1^2 + \delta_2^2 + \delta_3^2 + \eta^2) + 2\lambda^2(R_\sigma\sigma + R_\delta\delta_3)^2\Big] \\ &+ \lambda^2(R_\sigma\sigma + R_\delta\delta_3)\|\Omega\|^2 + \frac{\lambda^2}{4}\|\Omega\|^4\,. \end{aligned} \tag{3.4.38}$$

If one now defines the combined field, $\Gamma := [\sigma + \tan(\varphi_0)\delta_3]$, where $\tan(\varphi_0) = R_\delta/R_\sigma$ and a corresponding mass term, $m_\Gamma^2 := 2(\lambda R_\sigma)^2$, then the potential turns into:

$$\mathcal{V} = \frac{1}{2}\Big[m_\pi^2\|\Phi\|^2 + m_{\delta\eta}^2\|\Lambda\|^2 + m_\Gamma^2\Gamma^2\Big] + \frac{m_\Gamma^2}{2R_\sigma}\,\Gamma\,\|\Omega\|^2 + \frac{1}{4}\frac{m_\Gamma^2}{2R_\sigma^2}\|\Omega\|^4\,. \tag{3.4.39}$$

Note that the combined Γ field incorporates one free parameter φ_0 and with this, the total number of paramters ist still five, $\{m_\pi, m_{\delta\eta}, m_\Gamma, R_\sigma, \varphi_0\}$, as in the original representation. The above representation was introduced for completeness, but the following sections will mainly make use of the earlier found part (Eq. (3.4.37)).

III.4.1 Conserved scalar currents

As the potential part of the scalar Lagrangian has been presented in the previous section one could directly work out explicit (more detailed) relations for the model parameters. But just as for the fermionic sector the conserved currents should be included before going into detailed parameter fixing schemes.
For further introductory comments on the conserved currents and for nomencalture conventions compare Sec. III.3.3 or the general introduction in Sec. II.3. To establish the scalar contribution to the current, the change of the scalar field under small symmetry transformations is needed. In Sec. III.3.2 it was found that the scalar field has to change under left- and right-handed transformations as

$$\Omega' = U^{-1}_{\mathscr{L},\mathrm{ch}} \Omega U_{\mathscr{L},\mathrm{ch}} = L \Omega R^\dagger \,. \tag{3.4.40}$$

The additional $U_\mathrm{V}(1)$ symmetry, which was briefly introduced in Sec. III.3.3 will be included here in the same ignorant fashion as before. Eq. (3.4.40) immediately shows that Ω is invariant under a $U_\mathrm{V}(1)$ transformation, as it treats left- and right-handed parts equally and as the scalar phase commutes with the internal iso-spinor structure of Ω. Therefore the scalar Lagrangian does not contribute to the $U_\mathrm{V}(1)$ current ($j_\mathrm{B} = 0$). The fact that $j_\mathrm{B} = 0$ for scalar fields (e.g. mesons) nicely fits to the interpretation of j_B as the baryon current.

For the $SU(2)$ part (as earlier) the expansion of the symmetry group elements yields the change $\delta\Omega := \delta_{\varepsilon_{\ell/r}}\Omega$ under the group action. The only difference this time is that Ω changes under simultaneous action of left- and right-handed parts. To find the explicit change of all constituent fields (σ, $\boldsymbol{\pi}$, η, $\boldsymbol{\delta}$), it is useful to use the index notation $\Omega = \Omega^\alpha q_\mathrm{I}^\alpha$, where $q_\mathrm{I}^\alpha = (I, -\mathrm{i}\boldsymbol{\tau}^\mathrm{T})^\mathrm{T}$ is the quaternion symbol of the isospin group and $\alpha \in \{0, 1, 2, 3\}$.

$$\begin{aligned}
\delta\Omega &= L_{(\boldsymbol{\varepsilon}_\ell)} \Omega R^\dagger_{(\boldsymbol{\varepsilon}_r)} - \Omega && (3.4.41)\\
&= -\frac{\mathrm{i}}{2} \varepsilon_\ell^a \Omega^\alpha \tau^a q_\mathrm{I}^\alpha + \frac{\mathrm{i}}{2} \varepsilon_r^a \Omega^\alpha q_\mathrm{I}^\alpha \tau^a + O(\varepsilon_\ell^a \varepsilon_r^b) && (3.4.42)\\
&= -\frac{\mathrm{i}}{2} (\varepsilon_\mathrm{V}^a - \varepsilon_\mathrm{A}^a) \Omega^\alpha \tau^a q_\mathrm{I}^\alpha + \frac{\mathrm{i}}{2} (\varepsilon_\mathrm{V}^a + \varepsilon_\mathrm{A}^a) \Omega^\alpha q_\mathrm{I}^\alpha \tau^a + O(\varepsilon_\ell^a \varepsilon_r^b) && (3.4.43)\\
&= \varepsilon_\mathrm{A}^a \Omega^\alpha \frac{\mathrm{i}}{2} \{\tau^a, q_\mathrm{I}^\alpha\} - \varepsilon_\mathrm{V}^a \Omega^\alpha \frac{\mathrm{i}}{2} [\tau^a, q_\mathrm{I}^\alpha] + O(\varepsilon_\ell^a \varepsilon_r^b) && (3.4.44)\\
&= \underbrace{\varepsilon_\mathrm{A}^j \Omega^j}_{(\delta\Omega)^0} \underbrace{- \mathrm{i} \left(\epsilon^{ajk} \varepsilon_\mathrm{V}^a \Omega^j + \Omega^0 \varepsilon_\mathrm{A}^k \right)}_{(\delta\Omega)^k} \tau^k \,. && (3.4.45)
\end{aligned}$$

In the third line $\varepsilon_{\ell/r}$ have been replaced with the vectorial and axial variations from Sec. III.3.3. In the last line the Latin superscripts only run over the 'spacial' components $\{1,2,3\}$. Also the following relations were employed $\{-\mathrm{i}\tau^a, I\} = -2\mathrm{i}\tau^a$; $-\mathrm{i}\{\tau^a,\tau^b\} = -2\mathrm{i}\delta^{ab}$ and $-\mathrm{i}[\tau^a,\tau^b] = 2\epsilon^{abc}\tau^c$.
Comparing line Eq. (3.4.45) with the definition of $\Omega^\alpha q_\mathrm{I}^\alpha$ shows that it is possible to identify a vector $\delta\Omega = (\delta\Omega)^\alpha q_\mathrm{I}^\alpha$ with the same structure as the original Ω field. So the changed scalar field becomes:

$$(\Omega^\alpha)' = (\Omega + \delta\Omega)^\alpha \hat{=} \begin{pmatrix}\Omega^0\\ \boldsymbol{\Omega}\end{pmatrix} + \begin{pmatrix}\boldsymbol{\varepsilon}_A\boldsymbol{\Omega}\\ -\Omega^0\boldsymbol{\varepsilon}_A\end{pmatrix} + \begin{pmatrix}0\\ \boldsymbol{\varepsilon}_V\times\boldsymbol{\Omega}\end{pmatrix}. \tag{3.4.46}$$

Here the vectorial notation for Ω was used. The 'spacial' components are written as $\boldsymbol{\Omega} = (\Omega^1,\Omega^2,\Omega^3)^\mathrm{T}$ and '$\times$' labels the usual vector product: $(\boldsymbol{a}\times\boldsymbol{b})^\gamma := \epsilon^{\alpha\beta\gamma}a^\alpha b^\beta$. What is worth noting about this transformation behaviour is, that the six arbitrary parameters $\boldsymbol{\varepsilon}_A$ and $\boldsymbol{\varepsilon}_V$ are real and thus the real and imaginary parts of $\Omega = (\Phi + \mathrm{i}\Lambda)$[14] don't get mixed under the action of the left- and right-handed transformations. Therefore the constituent fields (Φ and Λ) transform in the same way as the whole field (Ω). Now the original σ-model would be regained by setting $\Lambda = 0$. Thus Eq. (3.4.46) reveals that the extended model transforms just as its ancestor and even more importantly its imaginary part Λ tranforms as the original σ-model, as well. In other words the extension from the quaternion description of the σ-model ($\Phi = \Phi^\alpha q_\mathrm{I}^\alpha$) to the biquaternion description with the field $\Omega = (\Phi + \mathrm{i}\Lambda)^\alpha q_\mathrm{I}^\alpha$ gives essentially a double copy of the original model. The only difference is that the roles of scalars and pseudo-scalars is exchanged in the new Λ part, which is essential for the structural similarity of the extended model and the original $\sigma-$model.
With this knowledge the needed variations of the constituent fields can be gained from Eq. (3.4.46).

$$\sigma' = \sigma + \boldsymbol{\varepsilon}_\mathrm{A}\boldsymbol{\pi} \quad , \quad \eta' = \eta + \boldsymbol{\varepsilon}_\mathrm{A}\boldsymbol{\delta}\,, \tag{3.4.47}$$

$$\pi' = \pi - \boldsymbol{\varepsilon}_\mathrm{A}\sigma + \boldsymbol{\varepsilon}_\mathrm{V}\times\boldsymbol{\pi} \quad , \quad \delta' = \delta - \boldsymbol{\varepsilon}_\mathrm{A}\eta + \boldsymbol{\varepsilon}_\mathrm{V}\times\boldsymbol{\delta}\,. \tag{3.4.48}$$

[14] A detailed definition of the components of Ω was given in the previous section (3.4.2-3.4.4).

Now, using the kinetic part of the scalar Lagrange density from the previous section (Eq. (3.4.1)) the vectorial and axial scalar currents can be constructed in analogy to the construction in Sec. III.3.3:

$$\varepsilon^a_{\mathrm{A}} j^{a\,\mu}_{\mathrm{A}} = \varepsilon^a_{\mathrm{A}} \left\{ \{\partial^\mu \sigma\}\pi^a - \{\partial^\mu \pi^a\}\sigma + \{\partial^\mu \eta\}\delta^a - \{\partial^\mu \delta^a\}\eta \right\}, \tag{3.4.49}$$

$$\Rightarrow \boldsymbol{j}^\mu_{\mathrm{A}} = \Re\left\{ \{\partial^\mu \Omega_0\}\boldsymbol{\Omega}^\star - \Omega_0\{\partial^\mu \boldsymbol{\Omega}^\star\} \right\}, \tag{3.4.50}$$

$$\varepsilon^a_{\mathrm{V}} j^{a\,\mu}_{\mathrm{V}} = \left\{ \{\partial^\mu \pi^a\}(\epsilon^{abc}\varepsilon^b_{\mathrm{V}}\pi^c) + \{\partial^\mu \delta^a\}(\epsilon^{abc}\varepsilon^b_{\mathrm{V}}\delta^c) \right\} \tag{3.4.51}$$

$$= \varepsilon^a_{\mathrm{V}} \left\{ \epsilon^{abc}\left(\pi^b\{\partial_\mu \pi^c\} + \delta^b\{\partial_\mu \delta^c\}\right) \right\}, \tag{3.4.52}$$

$$\Rightarrow \boldsymbol{j}^\mu_{\mathrm{V}} = \Re\left\{ \boldsymbol{\Omega} \times \{\partial^\mu \boldsymbol{\Omega}^\star\} \right\}. \tag{3.4.53}$$

In line Eq. (3.4.52) the cyclicality of ϵ^{abc} was used. The last line for vectorial or axial current is just convenient, short notations for the corresponding first line, which will simplify the comparission with the results from Sec. III.5.4. $\star$ stands for the complex conjugation and $\Re(\cdot)$ gives the real part of the argument. The equality can be verified by using the component definitions for Ω from the previous section.

The equations Eq. (3.4.49) and (3.4.51) reveal another detail of the so far derived model. There is no term in the vectorial or axial currents that connects the $\Phi = (\sigma, \boldsymbol{\pi}^{\mathrm{T}})$ part with the $\Lambda = (\eta, \boldsymbol{\delta}^{\mathrm{T}})^{\mathrm{T}}$ part of the model. Therefore, if current conservation holds, it must be satisfied independently for the the original σ-model (Φ) and for the new part (Λ). Also this was not mentioned so far, this feature is important for the model, since the effective instanton interaction in the potential (Eq. (3.4.8)) does not preserve the overall $SO(8)$ symmetry of the original Higgs potential (Eq. (3.4.1)), but breaks the model into a $\|\Phi\|^2$ and a $\|\Lambda\|^2$ dependent part. At this point one can be relieved, as the underlying symmetries are compatible with this explicit symmetry breaking $[SO(8) \to SO(4) \otimes SO(4)]$.

III.4.1.1 Violation of conservation laws

So far only the nice features of preserved symmetries and related conserved currents in the scalar sector have been discussed, but as often there is another side to the presented model. Throughout the presentation of the scalar potential in Sec. III.4 there was also a contribution presented, which explicitly violated the underlying symmetries. This was done via a presumed nuclear background contribution in Eq. (3.4.9).

Using the results from the previous section one can now determine how this symmetry breaking affects the conserved vectorial and axial currents. As discussed in Sec. II.3 both currents would be conserved if the symmetry breaking terms would not be included in the model, leading to two continuity equations:

$$\partial_\mu \boldsymbol{j}_{\mathrm{V}}^\mu = \delta\mathcal{L} = 0 \qquad , \qquad \partial_\mu \boldsymbol{j}_{\mathrm{A}}^\mu = \delta\mathcal{L} = 0\,. \tag{3.4.54}$$

Here the notation from Sec. II.3 was adopted. If one includes the symmetry breaking factors, then the righthand side of both equations changes, as the Lagrangian will not be invariant under symmetry transformations anymore. By assembling all symmetry preserving parts in a Lagrangian, $\mathcal{L}_0$, and the symmetry breaking contributions in $\mathcal{L}' = -\alpha\sigma - \beta\delta_3$[15], a general continuity equation becomes:

$$\partial_\mu \boldsymbol{j}^\mu = \delta\mathcal{L} = \delta\mathcal{L}_0 + \delta\mathcal{L}' = \delta\mathcal{L}'\,. \tag{3.4.55}$$

For the present case it is most convenient to analyse the symmetry breaking of the σ- and δ_3-meson separately. Starting out with the σ part, one finds the known result for the σ-model, that the axial current is broken by the VEV:

$$\boldsymbol{\varepsilon}_{\mathrm{A}}\left(\partial_\mu \boldsymbol{j}_{\mathrm{A}}^\mu\right)\Big|_{\delta_3=0} = -\alpha(\delta_{\boldsymbol{\varepsilon}_{\mathrm{A}}}\sigma) = \boldsymbol{\varepsilon}_{\mathrm{A}}(-\alpha\boldsymbol{\pi})\,, \tag{3.4.56}$$

$$\boldsymbol{\varepsilon}_{\mathrm{V}}\left(\partial_\mu \boldsymbol{j}_{\mathrm{V}}^\mu\right)\Big|_{\delta_3=0} = 0\,. \tag{3.4.57}$$

In the normal σ-model this equation is used to derive the Goldenberger-Treimann relation, which connects the pion decay constant f_π with the nucleon mass. A derivation for this context can be found in Ref. [34, p.106-110;p.126-128]. As the vectorial variation of σ vanishes (Eq. (3.4.47)), the vectorial current is still preserved in this model.

Now the symmetry breaking δ_3 term can be analysed similarly. The main difference is, that δ_3 transforms under axial **and** vectorial transformations (Eq. (3.4.48)), leading to

$$\boldsymbol{\varepsilon}_{\mathrm{A}}\left(\partial_\mu \boldsymbol{j}_{\mathrm{A}}^\mu\right)\Big|_{\sigma=0} = -\beta(\delta_{\boldsymbol{\varepsilon}_{\mathrm{A}}}\delta_3) = \beta\varepsilon_{\mathrm{A}}^3\eta\,, \tag{3.4.58}$$

$$\boldsymbol{\varepsilon}_{\mathrm{V}}(\partial_\mu \boldsymbol{j}_{\mathrm{V}}^\mu)\Big|_{\sigma=0} = -\beta(\delta_{\boldsymbol{\varepsilon}_{\mathrm{V}}}\delta_3) = -\beta(\varepsilon_{\mathrm{V}}^1\delta_2 - \varepsilon_{\mathrm{V}}^2\delta_1)\,. \tag{3.4.59}$$

Combining these two results with the earlier terms (Eq. (3.4.56) and (3.4.57)) the complete breaking of the conserved currents turns out to be:

$$\partial_\mu \boldsymbol{j}_{\mathrm{A}}^\mu = -\alpha\boldsymbol{\pi} + \beta\eta\,\hat{\boldsymbol{e}}_3\,, \tag{3.4.60}$$

$$\partial_\mu j_{\mathrm{V}}^{1\,\mu} = -\beta\delta_2\,, \tag{3.4.61}$$

$$\partial_\mu j_{\mathrm{V}}^{2\,\mu} = \beta\delta_1\,. \tag{3.4.62}$$

[15] This approach and notation was mainly adopted from Ref. [34, p.121-128].

So in the full model the three axial symmetries are broken by the σ VEV, R_σ, and (the third component) by the δ_3 VEV, R_δ, as well. In contrast to the ordinary σ-model the first two vectorial symmetries are also broken by the VEV of δ_3. The only remaining symmetry is the third vectorial component.
Remembering the final results from Sec. III.3.3, it is very helpful that the third component of the vectorial current, $j_{\mathrm{V}}^{3\,\mu}$, is still conserved. As the conservation of this component, combined with the conserved $U_{\mathrm{V}}(1)$ current, j_{B}^{μ}, lead to the conservation of proton and neutron numbers, it would be very disturbing for an effective nucleon model, if those symmetries would break down. Phrased differently, the scalar model as it was presented so far, preserves the two crucial symmetries in the context of an effective nucleon model. The combined scalar and nucleon Lagrangian preserves independently the number of protons and neutrons.

Finally, it its useful to rewrite the non-conserved currents (Eq. (3.4.60)-(3.4.62)) in terms of the physical (ladder operator) fields instead of the 'euclidean' representations. For this use the identification of the generators $\tau_\pm = (\tau^1 \pm \mathrm{i}\tau^2)/2$ and $\tau_0 = \tau^3$. The same relation holds then for the fields and so for example the divergence of $j_{\mathrm{V}\,+}^{\mu}$ becomes:

$$2\partial_\mu j_{\mathrm{V}\,+}^{\mu} = \partial_\mu (j_{\mathrm{V}\,1}^{\mu} + \mathrm{i} j_{\mathrm{V}\,2}^{\mu}) = -\beta(\delta_2 - \mathrm{i}\delta_1) = \beta\mathrm{i}(\delta_1 + \mathrm{i}\delta_2) = 2\mathrm{i}\beta\,\delta_+ \,. \tag{3.4.63}$$

For this rewriting the equations Eq. (3.4.61) and (3.4.61) have been used. Analogously the remaining current components can be calculated. In the ladder operator representation all non-conserved current components then become:

$$\partial_\mu j_{\mathrm{A}\,+}^{\mu} = -m_\pi^2 R_\sigma\, \pi_+ \quad , \quad \partial_\mu j_{\mathrm{V}\,+}^{\mu} = +\mathrm{i}\, m_{\delta\eta}^2 R_\delta\, \delta_+ \,, \tag{3.4.64}$$
$$\partial_\mu j_{\mathrm{A}\,-}^{\mu} = -m_\pi^2 R_\sigma\, \pi_- \quad , \quad \partial_\mu j_{\mathrm{V}\,-}^{\mu} = -\mathrm{i}\, m_{\delta\eta}^2 R_\delta\, \delta_- \,, \tag{3.4.65}$$
$$\partial_\mu j_{\mathrm{A}\,0}^{\mu} = -m_\pi^2 R_\sigma\, \pi_0 - m_{\delta\eta}^2 R_\delta\, \delta_0 \,. \tag{3.4.66}$$

Here the symmetry breaking factors α and β were replaced using Eq. (3.4.36). In doing so, one finds that the breakdown of the axial current conservation is related to the pion mass and the partial breakdown of the vectorial conservation law is connected to the pseudo-scalar meson with mass, $m_{\delta\eta}$.
This representation allows in addition to compare the very similar structure of the charged components of the non-conserved vectorial and axial currents (Eq. (3.4.64) and (3.4.65)). The only conceptual difference seems to be the factor $\pm\mathrm{i}$. This factor can be understood by rewriting the vectorial (Eq. (3.4.53)) and axial (Eq. (3.4.50)) currents from the previous section in the ladder operator basis. By expanding the equations Eq. (3.4.49) and (3.4.52) in terms of the $\{\pm, 0\}$-basis and collecting all survivng terms the

currents turn into:

$$j^{\mu}_{\mathrm{A}\,\pm} = -\Re\left\{\Omega_0 \overleftrightarrow{\partial^{\mu}} \Omega^{\star}_{\pm}\right\} \quad , \quad j^{\mu}_{\mathrm{V}\,\pm} = \pm\mathrm{i}\,\Re\left\{\Omega_3 \overleftrightarrow{\partial^{\mu}} \Omega^{\star}_{\pm}\right\}, \tag{3.4.67}$$

$$j^{\mu}_{\mathrm{A}\,3} = -\Re\left\{\Omega_0 \overleftrightarrow{\partial^{\mu}} \Omega^{\star}_{3}\right\} \quad , \quad j^{\mu}_{\mathrm{V}\,3} = 2\mathrm{i}\,\Re\left\{\Omega_+ \overleftrightarrow{\partial^{\mu}} \Omega^{\star}_{-}\right\}. \tag{3.4.68}$$

The symbol $\left[a \overleftrightarrow{\partial^{\mu}} b := a(\partial^{\mu} b) - (\partial^{\mu} a) b\right]$ was only introduced for notational reasons. Now, comparing the prefactors of the charged currents (vectorial and axial) with the prefactors in Eq. (3.4.64) and (3.4.65), one observes, that indeed, the breaking factors for vectorial and axial components simply give contributions to the non-conserved currents at the corresponding terms.

In the normal σ-model the non-conserved axial symmetry allows pions to decay via the processes of the type:

$$\langle \mathrm{vac}|\partial^{\mu}(j^{i}_{\mathrm{A}})_{\mu}|\pi^{j}\rangle \propto -\delta^{ij} f_{\pi} m_{\pi}^{2} e^{-\mathrm{i}kx}\,. \tag{3.4.69}$$

A full discussion and derivation of this aspect can be found in Ref. [34, p.106-110]. In the nomenclature from Sec. III.4 the pion decay constant is $f_{\pi} = R_{\sigma}$. Enlarging this picture to the present model means that here not only the pions are allowed to decay, but also the charged $\delta_{\pm}$-mesons, leading to the non-conserved vectorial current components. Note that a detailed derivation of these processes still needs to be done, but the general trend can already be observed.

III.4.2 Parameter conventions and dimensions

In Sec. II.9.8 a slightly different choice of parameter conventions for the Mexican hat potential was chosen for reasons of comparability. This small paragraph shall give the connection the two conventions and, in addition, gives the corresponding dimensional analysis for the used parameters.

Using the Lagrangian from Eq. (2.9.49) the connection can easily be made to the Lagrangian from the previous section (Eq. (3.4.1)) by the means of the following replacement:

$$H \to \lambda\Omega \quad , \quad \langle H\rangle \to \upsilon\lambda = \sqrt{\frac{-\mu^2}{\lambda^2}}\lambda\,. \tag{3.4.70}$$

Here υ means the VEV of the model, which is for the 'free' Mexican hat potential given above. For the final potential from Eq. (3.4.13) one gets $\upsilon = r_0$, where r_0 is the minimum of the potential, defined in the previous section.

In addition one needs to make the analytic continuation from Euclidean to Minkowski space, as Eq. (2.9.49) is given in Euclidean and Eq. (3.4.1) in Minkowski space. Depending on the conventions this directly reproduces the sign difference between both equations:

$$-S_{\mathrm{E}} = -\left(\int \mathrm{d}^d x \mathcal{H}\right)_{\mathrm{E}} \leftrightarrow \mathrm{i}\left(\int \mathrm{d}^d x \mathcal{L}\right)_{\mathrm{M}} = \mathrm{i}S \,. \tag{3.4.71}$$

Here, as earlier, the index E refers to Euclidean and M means Minkowski space. The analytic continuation is done (using the 'east coast metric') by replacing the time component with $x_0 = \mathrm{i}\tau$. This leads to the above equation (compare Ref. [5, p.176-177]).

The dimensional analysis of the introduced parameters will be done in the 'natural units' of energy. In these units one only needs to count the powers of energy, giving for fundamental quantities:

$$\text{energy:} \quad [m] = +1 \quad , \quad \text{length:} \quad [\ell] = -1\,, \tag{3.4.72}$$

$$\text{time:} \quad [t] = -1 \quad , \quad \text{derivatives:} \quad [\partial_\mu] = +1\,. \tag{3.4.73}$$

With this one knows that the Lagrange density has the units of inverse volume, $[\mathcal{L}] = d = 4$, and thus one can use $[(\partial\Omega)^2] = [\mu^2\Omega^2] = [\lambda^2\Omega^4] = [\alpha\sigma] = d$ to find the dimensions of the involved parameters. This leads to

$$[\Omega] = 1 \quad , \quad [\mu] = [a] = 1\,, \tag{3.4.74}$$

$$[\lambda] = 0 \quad , \quad [\alpha] = [\beta] = 3\,. \tag{3.4.75}$$

In the above list the dimensions of a and β follow by analogy (compare Eq. (3.4.13)). The dimensions of the two missing parameters can be found directly as well from Eq (3.4.15), giving: $[r_0] = 1$ and $[\phi_0] = 0$. For more information on dimensional analysis see also Ref. [5, p.90-91].

III.4.3 Parameter fixing

In the preceding three sections all important parameters for the scalar sector of the model have been introduced. In this context it only remains to fix the new model parameters by the means of some physical observables. It was briefly mentioned that one could either choose to tie them to the vacuum masses of the involved mesons or enforce the in medium proton and neutron masses plus some of the four possible meson masses. In the following both possibilities will be discussed.

III.4.3.1 Pure scalar model

Starting out with a pure scalar meson model, one would fix the model parameters via the meson masses $\{m_\sigma, m_\pi, m_\delta, m_\eta\}$. The equations Eq. (3.4.34), (3.4.35), (3.4.32) and (3.4.31) can be used to find the effective values. As there are four conditions (the physical meson masses), but a total of five parameters (e.g.: μ, a, λ, r_0 and φ_0)[16], one parameter will stay free. In this calculation the 4-point parameter λ will be chosen to be free. In principle this parameter can be fixed by adjusting it to a measured 4-point coupling strength of the involved mesons[17]. To arrive at relations that only depend on the meson masses the six equations Eq. (3.4.34), (3.4.35) and (3.4.36) have to be combined. In the following all relevant parameters will be rewritten, also only four are needed. For this, the pion mass and the η-δ mass can be substituted in all other equations, leading to:

$$m_\pi^2 = (\mu_-^2 + \lambda^2 r_0^2) \quad , \quad m_{\delta\eta}^2 = (\mu_+^2 + \lambda^2 r_0^2)\,, \tag{3.4.76}$$

$$m_\sigma^2 = m_\pi^2 + 2\lambda^2 R_\sigma^2 \quad , \quad m_\delta^2 = m_{\delta\eta}^2 + 2\lambda^2 R_\delta^2\,, \tag{3.4.77}$$

$$\alpha = m_\pi^2 R_\sigma \quad , \quad \beta = m_{\delta\eta}^2 R_\delta\,. \tag{3.4.78}$$

The second line can be used to find $R_\sigma = r_0\cos(\varphi_0)$ and $R_\delta = r_0\sin(\varphi_0)$. From these r_0 and φ_0 can be constructed as well:

$$R_\sigma^2 = \frac{m_\sigma^2 - m_\pi^2}{2\lambda^2} \quad , \quad R_\delta^2 = \frac{m_\delta^2 - m_{\delta\eta}^2}{2\lambda^2}\,, \tag{3.4.79}$$

$$r_0^2 = R_\sigma^2 + R_\delta^2 = \frac{1}{2\lambda^2}\Big[(m_\sigma^2 + m_\delta^2) - (m_\pi^2 + m_{\delta\eta}^2)\Big]\,, \tag{3.4.80}$$

$$\tan^2(\varphi_0) = \frac{R_\delta^2}{R_\sigma^2} = \frac{m_\delta^2 - m_{\delta\eta}^2}{m_\sigma^2 - m_\pi^2}\,. \tag{3.4.81}$$

With the equations for R_σ, R_δ and r_0 now the remaining parameters can be determined. Eq. (3.4.78) turns into

$$\alpha^2 = \frac{1}{2\lambda^2} m_\pi^4 (m_\sigma^2 - m_\pi^2) \quad , \quad \beta^2 = \frac{1}{2\lambda^2} m_{\delta\eta}^4 (m_\delta^2 - m_{\delta\eta}^2) \tag{3.4.82}$$

and finally μ and a the (anti-)symmetric combination of the equations Eq. (3.4.76) can be used including the earlier definition $\mu_\pm^2 = (\mu^2 \pm a^2)$:

$$\mu^2 = \frac{1}{2}(m_{\delta\eta}^2 + m_\pi^2) - \lambda^2 r_0^2 = \Big[(m_{\delta\eta}^2 + m_\pi^2) - \frac{1}{2}(m_\sigma^2 + m_\delta^2)\Big]\,, \tag{3.4.83}$$

$$a^2 = \frac{1}{2}(m_{\delta\eta}^2 + m_\pi^2)\,. \tag{3.4.84}$$

In Tab. 3.4.1 the effective couplings for the present case are given, as calculated from these conditions. λ is taken to be the remaining free parameter.

[16] Equivalently one could choose the original parameters of Eq. (3.4.13), or another suitable set.
[17] The 4-point term in the Lagrangian is $\lambda^2|\widetilde{\Omega}|^4/4$.

Table 3.4.1: Effective coupling constants for the scalar sector of the model. (The units for λ and φ_0 differ from the over all indication.)

conditions [MeV]	$m_\eta = 548$	$m_\pi = 140$	$m_\sigma = 571$	$m_\delta = 962$	
parameters I [MeV]	$\mu = \mathrm{i}\,553$	$a = 375$	$\alpha^{1/3} = 197\,\lambda^{-1/3}$	$\beta^{1/3} = 552\,\lambda^{-1/3}$	λ
parameters II [MeV]	$\mu_+ = \mathrm{i}\,407$	$\mu_- = \mathrm{i}\,668$	$r_0 = 682\,\lambda^{-1}$	$\varphi_0 = 0.96\,\mathrm{rad}$	λ

Some words are in order, concerning the calculated parameters from Tab. 3.4.1.

- The factor of i in the effective quadratic coupling, μ_j, might be surprising at first sight, but it is actually mandatory, in order for the potential Eq. (3.4.13) to have a non zero minimum.
- Also the parameter a is in a convenient numerical range. This parameter shall be interpreted as an effective instanton contribution. As it is related to a particular contribution in the partition function, it has to be real. Right now there are no further constraints on the parameter range and so the above result has to be used to fix the remaining parameters in instanton calculation (compare Eq. (3.2.9)).
- The radial position of the minimum in the effective potential, r_0, still depends on the quadric potential parameter, λ. As the actual minimum position is not connected to a direct observable or physical constraint, there is not much to do about this at the moment. For the same reason the angular position of the minimum, φ_0, is left out of the focus as well.
- The λ dependence of the symmetry violating terms on the other hand is very useful at the moment. Right now λ is only a free parameter and can be adjusted in a suitable fashion. Therefore it can be used to define the degree of symmetry breaking, which takes place in the model. For this note, that higher values of λ pronounce the overall $||\Omega||^4$ symmetry, while scaling down the values of α and β.
- Finally the symmetry breaking factors, α and β, do introduces some problems with the so far introduced interpretations. As one finds that the symmetry breaking factor of the δ-meson is larger than that of the σ-meson ($\beta > \alpha$), there is a conceptual problem with the argument that the breaking terms have their origin in the background nucleon condensate. Such a relation would enforce the symmetry breaking factors to be proportional to the scalar nucleon densities: $\alpha \propto \langle \overline{N}\, I\, N\rangle = [\langle \overline{p}p\rangle + \langle \overline{n}n\rangle]$ and $\beta \propto \langle \overline{N}\, \tau_3\, N\rangle = [\langle \overline{p}p\rangle - \langle \overline{n}n\rangle]$. As the scalar densities independently fulfil $\langle \overline{N}_j N_j\rangle \geq 0$, the condensate interpretation is only possible if $\alpha \geq \beta$. As seen in Tab. 3.4.1 this is not the case.
  Therefore either the interpretation of the symmetry breaking factors has to be altered, or (at least) one of the meson masses has to be changed in order to fulfil $\alpha \geq \beta$. Eq. (3.4.82) provides the information, how the masses can be changed to achive this goal. In principle it is sufficient to change the value of the, so far, unphysical mass combination, $m_{\delta\eta}$, for the first two δ-meson components

and the η-meson. One possible way to adjust the symmetry breaking factors conveniently will be presented in the next section.

III.4.3.2 Nucleon induced scalar model

In the previous section it was found that the interpretation of the symmetry breaking parameters α and β as nucleon condensates is incompatible with the found parameter values, if the physical meson masses are used to adjust them. In this section a slightly different approach shall be presented, which will remedy the previous problem at the cost of a changed $m_{\delta\eta}$ mass. Therefore, in contrast to the previous parameter fixing, this time the physical 'observables' $\{m_\sigma,\ m_\pi,\ m_\delta,\ M_\mathrm{p},\ M_\mathrm{n}\}$ will be used. Here $M_{\mathrm{p/n}}$ refers to the proton or neutron mass (more generally the masses of the two isospin components of the fermion spinor). Using this parameter set means, that all values which do not depend on $m_{\delta\eta}$ will keep the earlier derived dependencies.
Compared to the earlier parameter fixing one now has an additional parameter and thus might be tempted to fix all five model parameters (compare Eq. (3.4.13)). But this is not possible[18)] as the fermionic mass term comes with an additional parameter - the scalar coupling constant, g_Ω - which has to be determined. When the coupling between scalar mesons and fermions was discussed, it was already indicated in Eq. (3.3.13) how the fermions can obtain a mass through the VEV of the scalar field. In Sec. III.4 this VEV was found to be $\langle\Omega\rangle = R_\sigma I + R_\delta\tau_3$ (compare for example Eq. (3.4.22)). Remember that Eq. (3.4.79) showed that R_σ is independent of $m_{\delta\eta}$, while $R_\delta = R_\delta(m_{\delta\eta})$. This equation can also be used to calculate the fermion masses, depending on the condensates R_σ and R_δ:

$$M^{ij}\overline{N}^i N^j = g_\Omega \overline{N}\langle\Omega\rangle N = g_\Omega \overline{N}\,\overbrace{[R_\sigma I + R_\delta\tau_3]}^{M^{ij}}\,N\,, \tag{3.4.85}$$

$$\Rightarrow M_\mathrm{p} = g_\Omega(R_\sigma + R_\delta)\,, \tag{3.4.86}$$

$$\Rightarrow M_\mathrm{n} = g_\Omega(R_\sigma - R_\delta)\,. \tag{3.4.87}$$

To get the proton and neutron masses the explicit iso-spinor structure of N was used. In order to fix the remaining parameters of the model it is useful to define two combined quantities from the seperate proton and neutron masses. In an ideal nucleon model the proton and neutron masses are exactly degenerate and thus one would only have a common, single valued nucleon mass. In contrast, if the densities of protons and neutrons are different, this leads to different effective masses for both iso-spinor components. Both

[18)] In fact it would be strange if the introduction of an additional, purely fermionic constraint would fix the ambiguities of a purely scalar model.

effects can be captured in the quantities:

$$M := \frac{1}{2}(M_\mathrm{p} + M_\mathrm{n}) = g_\Omega R_\sigma \,, \tag{3.4.88}$$

$$\epsilon_\mathrm{M} := \frac{M_\mathrm{n}}{M_\mathrm{p}} = \frac{R_\sigma - R_\delta}{R_\sigma + R_\delta} \,. \tag{3.4.89}$$

M represents the isospin symmetry and can be used to adjust the scalar coupling constant g_Ω, while ϵ_M is directly related to the effective mass difference of protons and neutrons and thus to the isospin symmetry breaking. If $\epsilon_\mathrm{M} = 1$, then $M_\mathrm{p} = M_\mathrm{n}$ and from the definition one sees immediately that $R_\delta = 0$. This is the natural case of symmetric nuclear matter. If the ratio of neutrons to protons changes, then ϵ_M will change. In the extreme of vanishing neutron mass one has $\epsilon_\mathrm{M} = 0$. In terms of formulas the definition of ϵ_M can be used to replace the $m_{\delta\eta}$-dependence in the equations for R_δ, β, φ_0 and $m_{\delta\eta}$ itself. Starting out with the vacuum angle, from Eq. (3.4.81) one finds:

$$\epsilon_\mathrm{M} = \frac{1 - R_\delta R_\sigma^{-1}}{1 + R_\delta R_\sigma^{-1}} = \frac{1 - \tan(\varphi_0)}{1 + \tan(\varphi_0)} \,, \tag{3.4.90}$$

$$\Rightarrow \tan(\varphi_0) = \frac{1 - \epsilon_\mathrm{M}}{1 + \epsilon_\mathrm{M}} \,. \tag{3.4.91}$$

Note that the second line fits nicely in the general interpretation of φ_0, as $\varphi_0 = 0$ means that the VEV lies completely in the σ-direction (compare Eq. (3.4.79)). Now the relation for $\tan(\varphi_0)$ can be employed to solve for the remaining parameters:

$$R_\delta^2 = R_\sigma^2 \tan^2(\varphi_0) = R_\sigma^2 \left(\frac{1 - \epsilon_\mathrm{M}}{1 + \epsilon_\mathrm{M}}\right)^2 \,, \tag{3.4.92}$$

$$m_{\delta\eta}^2 = m_\delta^2 - 2\lambda^2 R_\delta^2 = m_\delta^2 - (m_\sigma^2 - m_\pi^2)\left(\frac{1 - \epsilon_\mathrm{M}}{1 + \epsilon_\mathrm{M}}\right)^2 \,, \tag{3.4.93}$$

$$\beta^2 = m_{\delta\eta}^4 R_\delta^2 = \left(m_\delta^2 - (m_\sigma^2 - m_\pi^2)\left(\frac{1 - \epsilon_\mathrm{M}}{1 + \epsilon_\mathrm{M}}\right)^2\right)^2 R_\sigma^2 \left(\frac{1 - \epsilon_\mathrm{M}}{1 + \epsilon_\mathrm{M}}\right)^2 \,. \tag{3.4.94}$$

To arrive at these relations the equations Eq. (3.4.79) and (3.4.82) were used. Naturally the parameters r_0, μ and a have to be changed as well, since they depend on $m_{\delta\eta}^2$ but this replacement is trivial using Eq. (3.4.93). This equation also delivers a very convenient feature of this type of parameter fixing. Beforehand, it was discussed that the difference of the masses m_δ and $m_{\delta\eta}$ are very unnatural from a group theoretical point of view. Especially the different masses for the components of the δ-meson gave a strange picture of this triplet. Now, in the parameter fixing of Eq. (3.4.93) one sees, that the problem with different masses vanishes for $\epsilon_\mathrm{M} = 1$ (in symmetric matter) and only increases with the asymmetry of the nuclear matter sector. Thus, for applications that only deal with small proton-neutron asymmetries the mass difference of $\delta_{1,2}$ compared to δ_3 can be treated as a small perturbation. This would effectively state that the original symmetry, leading to the η singlet and the $\boldsymbol{\delta}$ triplet, is only approximately realised, with some correction effects of order ϵ_{M}. This idea of approximate symmetries is a well established concept in various fields of modern theoretical physics and this would be another explicit example in a long list of applications.

Another advantage of the nucleon induced parameter fixing is related to the values of β in Eq. (3.4.94). Now, the limit of equal proton and neutron mass does not yield a problem concerning the value of α and β anymore. As ϵ_{M} goes to zero, β vanishes as well. This gives the correct behaviour, if the origin of β shall be related to the difference of proton and neutron densities. So, in this description, the original argument for the symmetry breaking factors α and β is applicable.

While this gives indeed a very neat interpretation of the original parameters in the scalar potential, there arises also an upper limit for the symmetry breaking. To see this compare Eq. (3.4.94) with the defining equation for the other breaking factor, $\alpha^2 = m_\pi^4 R_\sigma^2$. If one insists on the constraint $\alpha \geq \beta$, both equations can be combined and lead to a limit for ϵ_{M}:

$$\alpha^2 \geq \beta^2 \qquad \Rightarrow \qquad m_\pi^4 \geq \left(m_\delta^2 - (m_\sigma^2 - m_\pi^2) \left(\frac{1-\epsilon_{\mathrm{M}}}{1+\epsilon_{\mathrm{M}}} \right)^2 \right)^2 \left(\frac{1-\epsilon_{\mathrm{M}}}{1+\epsilon_{\mathrm{M}}} \right)^2 . \tag{3.4.95}$$

Solving this equation for ϵ_{M} gives the desired limit, up to which the above equations are compatible with the discussed interpretation of α and β.
The reason that not all symmetry breaking values of ϵ_{M} are applicable in this description is that the parameters are not only tied to fermionic observables, but also to the meson masses m_δ, m_σ and m_π. As these masses are kept fixed at their in-vacuum values the above interpretation should break down at some point. Of course, the presented argument gives room for more elaborate schemes to fix the free parameters of the so far presented model.

III.4.3.3 Conclusion

Conclusively one can state that there are various possible applications for the so far developed model. If one allows for example the in-medium masses of all involved mesons to change and takes as absolute input parameters only the proton and neutron masses, then various self-consistency methods could be used to determine the parameters of the model (and with these the effective meson masses as well).

Also the pure meson model, for which a possible parameter fixing was presented in Sec. III.4.3.1 yields interesting applications, as it allows studying various interacting scalar mesons in a group theoretical environement with incuded effective instanton interactions. Especially the different constrained instanton effects from Sec. II.9.9 and II.9.10, which have been ignored so far, could lead to interesting relations. In this context the clear advantage of the scalar model is, that it allows these kind of investigations in an otherwise relatively simple environment of the scalar Lagrangian from Sec. III.4.

III.5 The gauge field Lagrangian

Having introduced the scalar sector of the model, one still needs to develope the concrete connection to the instanton part - that is to say the connection to the gauge field sector has to be implemented.
Gauge fields are commonly introduced in order to maintain a certain invariance in a given model. Suppose a Lagrange density [e.g. $\mathcal{L} = \mathcal{L}(\|\Omega\|^2)$] is invariant under a symmetry transformation, $U(\boldsymbol{\theta})$, with a set of real parameters, $\{\theta_j\}$. If the parameters, θ_j, are constant in space-time, then the symmetry is said to be a global symmetry. In this case the dynamics of the model are directly invariant under the symmetry transformation, as the constituents of the Lagrange density $\|\Omega\|^2 = \frac{1}{2}\mathrm{tr}_{\mathrm{I}}[\Omega^\dagger\Omega]$ and $\|\partial\Omega\|^2 = \frac{1}{2}\mathrm{tr}_{\mathrm{I}}[(\partial_\mu\Omega)^\dagger(\partial^\mu\Omega)]$ are invariant themselves. This situation was treated in the previous sections (Sec. III.3 - III.4.1.1).
In contrast, if the parameters change in space-time [$\theta_j = \theta_j(x)$], the model has a 'local symmetry' and is not directly gauge invariant with respect to the given symmetry transformation, $U(\boldsymbol{\theta}(x))$. The reason is that the kinetic part now produces a gauge dependent derivative contribution [$\partial_\mu(U\Omega) = (\partial_\mu U)\Omega + U\partial_\mu\Omega$]. In order to restore the desired gauge invariance one needs to include explicitly a gauge field that absorbs the additional derivative contribution, $(\partial_\mu U)\Omega$, under gauge transformations. The inclusion of such a gauge field is a standard procedure in modern physics and can be reviewed nicely in Ref. [34, p.135-146] and more formal for non-abelian symmetries in Ref. [5, p.416-434]. In App. A.7 a natural reason for the existence of gauge fields is briefly discussed. Gauge field models are invariant under local symmetry transformations, if the ordinary derivative is replaced by the gauge covariant derivative, $D_\mu = \partial_\mu + \mathrm{i}g_{\mathrm{A}}A_\mu$, which was already extensively used in chapter Sec. II.8. The instanton sector, which originates from considering local gauge symmetries, can be included in the so far derived model if a covariant derivative is identified, which fits to the assumed local symmetries (for the explicit symmetries compare Sec. III.1).

Although this is a nice way to establish the needed connection, there is one catch to the procedure. In Sec. III.4 there were two terms included, which explicitly violated certain symmetries. Therefore, stricly speaking, if the terms are included, then the global symmetries do not exist, which directly implies that their local equivalences do not exist either. The general approach to circumvent this issue is to assume an underlying symmetry, which is broken by explicit physical realisations. Thus in this final part the influence of the symmetry breaking terms, $\mathcal{V}_{\mathrm{nucl}}$ (Eq. (3.4.9)), will be explicitly ignored to derive the 'underlying' symmetry relations.

III.5.1 Flavour or colour gauge fields

As previously mentioned the present model shall be an effective low energy approximation to QCD. Ideally an effective model only depends on observables, that are accessible at the defining low energy scale and high energy contributions should only give corrections as the cutoff is approached. Any explicit dependence on high energy parameters corresponds to an 'imperfection' of the model. These parameters are not predictable in a low energy approximation and thus have to be adjusted using external inputs (such as experiments or high energy models). Problems connected with the exact high energy parameter fixing have to be adressed in a Wilson or 'matching' EFT approach (for the concept compare Sec. II.6). In 1979 Saito and Shigemoto built an effective Lagrangian for (pseudo-)scalar mesons by using a massless $SU(2)$ colour gauge field as instanton field in Ref. [1]. Conceptually this idea is appealing as the absence of mass in the instanton sector significantly simplifies all related calculations (compare Sec. II.9.8). On the other hand the experimental evidence stands in sharp contrast to this point of view. The only 'observed' massless, non-abelian gauge fields so far correspond to the gluons (the $SU(3)$ colour gauge fields). Explicit gluon induced terms in a Lagrange density are high energy contributions, as confinement sets a lower energy bound on their influence. By using the colour gauge field they effectively introduced (high energy) gluonic contributions at all energy scales. So Saito's and Shigemoto's idea gives a nice conceptual argument for the different masses of $\sigma-$, $\pi-$, $\delta-$ and $\eta-$meson (in the context of the linear $\sigma-$model), but its explicit realisation does not seem to appear in nature. In addition there is another reason against this particular realisation, which was already discussed in the context of constrained instantons in Sec. II.9.8. As soon as the gauge field is coupled to a scalar field with non vanishing VEV, instantons have to be replaced by their constrained relatives, if one is interested in (partly) preserving the instanton features. Using constrained instantons with a non-trivial VEV in the scalar sector, one obtains an effective gauge field mass in the process, which will become apparent in the following sections.

Due to these reasons a slightly different approach is pursued in this work. If the effective instanton gauge field originates from a flavour instead of the colour symmetry, then explicit gluonic degrees of freedom can be left out of the picture completely. In this picture the effective mass of the gauge fields turns into an advantage, as the flavour gauge fields ultimately lead to vector meson contributions such as the ϱ-meson. As these are indeed very heavy, the effective gauge field mass, due to spontaneous symmetry breaking in the scalar sector, does not lead to any conceptual complications this time. While this is a very neat feature, explicit vector meson calculation for the model will not be discussed in detail in this work, as the correct implementation of the pure constrained instanton mechanism is already involved enough. Any analysis on the inclusion of explicit vector meson contributions thus has to be postponed to future works. For the present model the non-abelian gauge field lives in a $SU_{\mathrm{I}}(2)$ iso-spinor space instead of the $SU(2)$ colour space from Saito and Shigemoto.

III.5.2 Gauge field couplings

There is no kinetic term for constrained instantons, as was discussed in Sec. III.2 and so one can directly come to the couplings between the gauge field with fermion or scalar fields. The different aspects of these couplings will be worked out in the following sections (until Sec. III.5.5) for the situation, where the gauge field consists of the constrained instanton and a quantum fluctuation around it $A_\mu = A_\mu^{\mathrm{con}} + a_\mu$.

But before going into the details of new terms in the Lagrange density another leeway has to be taken care of.
The concept of instantons emerged from the euclidianisation of Minkowski space combined with the demand of a *local* Lorentz symmetry. The locality of the symmetry group gave rise to a gauge field, A_μ, and the eucidianisation allowed a calculation of the instanton field as classical configuration A_μ^{clas}[19] of the gauge field (compare the first part of chapter Sec. II.8 and II.1 for properties of the restricted Lorentz group). But so far the actual coupling of the gauge field, A_μ, to the various fields of the model have not been discussed.
Recall that 4-dimensional Euclidean space can be represented via two independent $SU(2)$ groups ($SO(4) \simeq SU_{\mathrm{A}}(2) \otimes SU_{\mathrm{B}}(2)$). Therefore, in the iso-spinor representation, the gauge field, $A^\mu = A^\mu_A \oplus A^\mu_B$, consists of 2 sub-fields - one for each $SU(2)$ subgroup. As the gauge fields shall be connected to physical observables it is most convenient to switch back to Minkowski space at this point. Again one has to emphasise that this alteration of the underlying symmetry group is non-trivial. Some issues concerning this topic have already been discussed in Sec. II.8.1. Nevertheless (ignoring all possible problems), in Minkowski space the two $SU(2)$ subgroups are ($SO^+(3,1) \simeq SU_{\mathrm{L}}(2) \otimes SU_{\mathrm{R}}(2)$). In the spirit of chiral approaches (as briefly introduced in Sec. II.5) the symmetry group is promoted to $U_{\mathrm{L}}(2) \otimes U_{\mathrm{R}}(2)$[20]. This would lead to a total of 8 symmetry generators (3 for each $SU(2)$ and 1 for each $U(1)$ group). In the introductory part to the model (Sec. III.1) and at the end of Sec. II.8.4 it was mentioned that instanton interactions explicitly violate the abelian axial symmetry, $U_{\mathrm{A}}(1)$. So, before running through the whole machinery this symmetry is directly excluded again, leading to the subgroup:

$$U_{\mathrm{L}}(2) \otimes U_{\mathrm{R}}(2) \to SU_{\mathrm{L}}(2) \otimes SU_{\mathrm{R}}(2) \otimes U_{\mathrm{V}}(1)\,, \tag{3.5.1}$$

where the subscript V means a vectorial symmetry, $U_{\mathrm{V}}(1)$[21]. This symmetry group will be assumed in the following derivations. Using the Minkowski space iso-spinor representation, now the covariant derivatives can be derived. The derivation is related to a derivation concerning the electroweak interactions for hadrons in Ref. [5, p.562-571]. Take the local symmetry transformations to be:

$$L \equiv L_{(\boldsymbol{\varepsilon}_\ell(x))} := \exp\Big[-\mathrm{i}\varepsilon^j_{\ell\,(x)} T^j\Big] \qquad , \qquad R \equiv R_{(\boldsymbol{\varepsilon}_r(x))} := \exp\Big[-\mathrm{i}\varepsilon^j_{r\,(x)} T^j\Big]\,. \tag{3.5.2}$$

Note that the definition for left- and right-handed symmetry transformations has been slightly modified compared to Sec. II.5, as this notation is a bit more convenient. The $\boldsymbol{\varepsilon}_{\ell,r}(x) \in R^3$ are the six gauge parameters and the $SU(2)$ generators, $T^j = \tau^j/2$, have the same normalisation as the instanton generators

[19] The superscriptclas refers to ordinary instantons, whereas the label con indicates constrained instantons.
[20] Compared to $SU_{\mathrm{L}}(2) \otimes SU_{\mathrm{R}}(2)$ this symmetry group incorporates additional $U(1)$ symmetries: $U_{\mathrm{L}}(2) \otimes U_{\mathrm{R}}(2) \simeq U_{\mathrm{L}}(1) \otimes U_{\mathrm{R}}(1) \otimes SU_{\mathrm{L}}(2) \otimes SU_{\mathrm{R}}(2)$.
[21] In vectorial symmetries the left- and right-handed transformations are equal (e.g.: $U(\alpha_{\mathrm{V}}) = U((\alpha_\ell + \alpha_r)/2) = e^{-\mathrm{i}(\alpha_\ell+\alpha_r)}$). See also Sec. II.5.

from Sec. II.9. The lacking factor of $(-\mathrm{i})$ is just a different convention[22]. One already knows how the scalar and fermion fields transform under the action of the $SU_{\mathrm{L,R}}(2)$ groups:

$$N_{\mathrm{L}} = (p_{\mathrm{L}}, n_{\mathrm{L}})^{\mathrm{T}} \quad , \quad (N_{\mathrm{L}})' = LN_{\mathrm{L}} \,, \tag{3.5.3}$$
$$N_{\mathrm{R}} = (p_{\mathrm{R}}, n_{\mathrm{R}})^{\mathrm{T}} \quad , \quad (N_{\mathrm{R}})' = RN_{\mathrm{R}} \,, \tag{3.5.4}$$
$$\Omega = \Omega^{\alpha} q_{\mathrm{I}}^{\alpha} \quad , \quad (\Omega)' = L\Omega R^{\dagger} \,. \tag{3.5.5}$$

Here the primes denote the action of the symmetry group: $(Z)' = U^{-1}ZU$, where U is the unitary operator realising the group action according to Eq. (3.5.1). This notation is adopted throughout the whole section. If one now defines the gauge fields to have the following transformation law under $SU_{\mathrm{L}}(2) \otimes SU_{\mathrm{R}}(2) \otimes U_{\mathrm{V}}(1)$:

$$\ell_{\mu} = \ell_{\mu}^{a} T^{a} + b_{\mu} I \quad , \quad (\ell_{\mu})' = L\ell_{\mu}L^{\dagger} + \mathrm{i}L\partial_{\mu}L^{\dagger} \,, \tag{3.5.6}$$
$$r_{\mu} = r_{\mu}^{b} T^{b} + b_{\mu} I \quad , \quad (r_{\mu})' = Rr_{\mu}R^{\dagger} + \mathrm{i}R\partial_{\mu}R^{\dagger} \,, \tag{3.5.7}$$

then the covariant derivatives can be constructed such that the Lagrangian becomes gauge invariant under Eq. (3.5.2). Here b_{μ} is the gauge field corresponding to the $U_{\mathrm{V}}(1)$ group, while ℓ_{μ}^{a} and r_{μ}^{a} are the components of the left- and right-handed $SU(2)$ symmetries. Inserting all transformation rules (Eq. (3.5.3)-(3.5.7)) one finds that

$$D_{\mu}N_{\mathrm{L}} = (\partial_{\mu} - \mathrm{i}g_{\mathrm{A}}\ell_{\mu})N_{\mathrm{L}} \quad , \quad (D_{\mu}N_{\mathrm{L}})' = L(D_{\mu}N_{\mathrm{L}}) \,, \tag{3.5.8}$$
$$D_{\mu}N_{\mathrm{R}} = (\partial_{\mu} - \mathrm{i}g_{\mathrm{A}}r_{\mu})N_{\mathrm{R}} \quad , \quad (D_{\mu}N_{\mathrm{R}})' = R(D_{\mu}N_{\mathrm{R}}) \,, \tag{3.5.9}$$
$$D_{\mu}\Omega = \partial_{\mu}\Omega - \mathrm{i}g_{\mathrm{A}}\ell_{\mu}\Omega + \mathrm{i}g_{\mathrm{A}}\Omega r_{\mu} \quad , \quad (D_{\mu}\Omega)' = L(D_{\mu}\Omega)R^{\dagger} \,, \tag{3.5.10}$$
$$(D_{\mu}\Omega)^{\dagger} = \partial_{\mu}\Omega^{\dagger} + \mathrm{i}g_{\mathrm{A}}\Omega^{\dagger}\ell_{\mu} - \mathrm{i}g_{\mathrm{A}}r_{\mu}\Omega^{\dagger} \quad , \quad [(D_{\mu}\Omega)^{\dagger}]' = R(D_{\mu}\Omega)^{\dagger}L^{\dagger} \tag{3.5.11}$$

lead to gauge invariant kinetic terms for the fermionic and scalar fields. Note that these are of the form

$$\mathcal{L}_{N}^{\mathrm{kin}} = (\overline{N}_{\mathrm{L}}\not{D}N_{\mathrm{L}} + \overline{N}_{\mathrm{R}}\not{D}N_{\mathrm{R}}) = \overline{N}[\not{D}P_{\mathrm{L}} + \not{D}P_{\mathrm{R}}]N \,, \tag{3.5.12}$$
$$\mathcal{L}_{\Omega}^{\mathrm{kin}} = \frac{1}{4}\mathrm{tr}_{\mathrm{I}}[(D_{\mu}\Omega)^{\dagger}(D^{\mu}\Omega)]^{23)} \,. \tag{3.5.13}$$

[22] The local gauge field formalism has lots physics related roots, while the instanton formalism was developed in Euclidean space-time, which is historically closer to the mathematitians approach to field theory. The mathematical convention is to absorb all prefactors into the gauge fields, whereas phycisists tend to factor out a coupling constant and a factor of $-\mathrm{i}$.

Looking at line Eq. (3.5.10) also reveals that the $b_\mu I$ part of the gauge fields explicitly drops out, since I commutes with Ω.

$$D_\mu\Omega = \partial_\mu\Omega - \mathrm{i}g_\mathrm{A}\ell^a_\mu T^a\Omega + \mathrm{i}g_\mathrm{A}\Omega r^a_\mu T^a - \mathrm{i}g_\mathrm{A}b_\mu(I\Omega - \Omega I) = \partial_\mu\Omega - \mathrm{i}g_\mathrm{A}\ell^a_\mu T^a\Omega + \mathrm{i}g_\mathrm{A}\Omega r^a_\mu T^a\,. \tag{3.5.14}$$

As gauge fields in a Higgs formalism only acquire masses via the kinetic term of the Higgs field (here Ω), one can directly conclude that b_μ remains massless in this picture. Using Eq. (3.5.8) to (3.5.11) all interactions between the gauge fields and the other constituents of the model can be worked out explicitly. Just to keep the connections in mind, remember that the action minimising configuration of (left-) right-handed gauge fields corresponds to the (anti-) instanton field. So all couplings for ℓ_μ and r_μ essentially describe instanton and fluctuation couplings.

III.5.3 Gauge field masses

While the coupling to fermions will be important if conserved currents are of interest, the coupling between the scalar field, Ω, and the gauge fields, r_μ and ℓ_μ, leads to the effective gauge field masses. To find these masses the covariant kinetic term of the scalar fields has to be expanded:

$$\begin{aligned}
\mathcal{L}^\mathrm{kin}_\Omega &= \frac{1}{4}\mathrm{tr_I}\Big[(D_\mu\Omega)^\dagger(D^\mu\Omega)\Big] = \frac{1}{4}\mathrm{tr_I}\Big[(\partial_\mu\Omega^\dagger + \mathrm{i}g_\mathrm{A}\Omega^\dagger\ell_\mu - \mathrm{i}g_\mathrm{A}r_\mu\Omega^\dagger)\,(\partial^\mu\Omega - \mathrm{i}g_\mathrm{A}\ell^\mu\Omega + \mathrm{i}g_\mathrm{A}\Omega r^\mu)\Big] && (3.5.15)\\
&= \frac{1}{4}\mathrm{tr_I}\Big[(\partial_\mu\Omega)^\dagger(\partial^\mu\Omega)\Big] && (3.5.16)\\
&+ \frac{g_\mathrm{A}^2}{4}\Big(\mathrm{tr_I}\Big[\Omega\Omega^\dagger\ell_\mu\ell^\mu\Big] + \mathrm{tr_I}\Big[\Omega^\dagger\Omega\, r_\mu r^\mu\Big] - 2\mathrm{tr_I}\Big[\Omega^\dagger\ell_\mu\Omega r^\mu\Big]\Big) && (3.5.17)\\
&- \frac{\mathrm{i}g_\mathrm{A}}{4}\Big(\mathrm{tr_I}\Big[\ell_\mu\Big[\Omega(\partial^\mu\Omega^\dagger) - (\partial^\mu\Omega)\Omega^\dagger\Big]\Big] + \mathrm{tr_I}\Big[r_\mu\Big[\Omega^\dagger(\partial^\mu\Omega) - (\partial^\mu\Omega^\dagger)\Omega\Big]\Big]\Big)\,. && (3.5.18)
\end{aligned}$$

To arrive at the final result the cyclicality of the trace was used and the fact that $a_\mu b^\mu = a^\mu b_\mu$. In this expansion line Eq. (3.5.16) just gives the kintetic energy of the scalar fields, the second line (Eq. (3.5.17)) gives the quadratic interactions between gauge and scalar fields (and through the VEV the gauge field masses as well) and the last line (Eq. (3.5.18)) leads to the current contributions from the scalar fields. As the title of this section indicates line Eq. (3.5.17) will be in the focus for the remaining paragraph. The currents of the present model will be discussed in Sec. III.5.4.

As discussed in the previous section, the contribution from the $U_\mathrm{V}(1)$ gauge field, b_μ, vanishes from the covariant derivative, $D_\mu\Omega$, and therefore in this section the field will be set to zero ($b_\mu = 0$). To analyse the quadratic interaction between gauge and scalar fields some traces in iso-spinor space need to be calculated. The derivations can be reviewed in App. A.8 and here only the results will be employed.

[23] Using the cyclic permutivity of the trace gives the invariance of the term. The factor of $1/4$ combines the $1/2$ from a kinetic term with the $1/2$ from the isospin trace.

The different terms in line Eq. (3.5.17) turn out to be:

$$\frac{1}{4}\mathrm{tr}_{\mathrm{I}}\Big[\Omega\Omega^{\dagger}\ell_{\mu}\ell^{\mu}\Big]_{b_{\mu}=0} = \frac{1}{8}\Omega^{\alpha}(\Omega^{\star})^{\beta}\,\ell^{a}_{\mu}\ell^{b\,\mu}\,\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\Big[q^{\alpha}\bar{q}^{\beta}\tau^{a}\tau^{b}\Big] \tag{3.5.19}$$

$$= \frac{1}{8}\Omega^{\alpha}(\Omega^{\star})^{\beta}\,\ell^{a}_{\mu}\ell^{b\,\mu}\left(\delta_{\alpha\beta}\delta_{ab}\Big|^{3}_{\alpha,\beta=0} + \Big[-\delta_{\alpha 0}\epsilon_{\beta ab} + \delta_{\beta 0}\epsilon_{\alpha ab} - \delta_{\alpha a}\delta_{\beta b} + \delta_{\alpha b}\delta_{\beta a}\Big]^{3}_{\alpha,\beta=1}\right)$$

$$= \frac{1}{8}\Omega^{\alpha}(\Omega^{\star})^{\alpha}\,\ell^{b}_{\mu}\ell^{b\,\mu} = \frac{1}{8}\|\Omega\|^{2}\boldsymbol{\ell}_{\mu}\boldsymbol{\ell}^{\mu}\,. \tag{3.5.20}$$

Here $\star$ denotes the complex conjugation (remember that $\Omega^{\alpha} \in \mathbb{C}^{1}$) and the vectorial gauge field, $\boldsymbol{\ell}_{\mu} = (\ell^{1}_{\mu}, \ell^{2}_{\mu}, \ell^{3}_{\mu})^{\mathrm{T}}$, contains just the $SU_{\mathrm{L}}(2)$ contribution. The additional factor of $1/4$ in the second equality comes from the generators of ($\ell_{\mu} = \ell^{a}_{\mu}\tau^{a}/2 + b_{\mu}I$). To arrive at the third line it was used, that $\ell^{a}_{\mu}\ell^{b\,\mu}$ is symmetric in the iso-spinor indices, a and b, while the part in square brackets in the second line is antisymmetric in a and b. Therefore this contribution vanishes and only the product of 'norms' survives. The quadratic term in r_{μ} can be calculated completely analogously, since the only change takes place in the antisymmetric part, which vanishes anyway. One finds:

$$\frac{1}{4}\mathrm{tr}_{\mathrm{I}}\Big[\Omega^{\dagger}\Omega\, r_{\mu}r^{\mu}\Big]_{b_{\mu}=0} = \frac{1}{8}\|\Omega\|^{2}\,\boldsymbol{r}_{\mu}\boldsymbol{r}^{\mu}\,. \tag{3.5.21}$$

The final contribution to the quadratic interaction comes from the remaining mixed term. Due to this mixing and the slightly different structure, this contribution is a bit more tedious.

$$-\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\Big[\Omega^{\dagger}\ell_{\mu}\Omega r^{\mu}\Big]_{b_{\mu}=0} = -\frac{1}{4}(\Omega^{\star})^{\alpha}\Omega^{\beta}\,\ell^{a}_{\mu}\,r^{b\,\mu}\,\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\Big[\bar{q}^{\alpha}\tau^{a}q^{\beta}\tau^{b}\Big] \tag{3.5.22}$$

$$= -\frac{1}{4}(\Omega^{\star})^{\alpha}\Omega^{\beta}\,\ell^{a}_{\mu}\,r^{b\,\mu}\Bigg(\Big(\delta_{\alpha 0}\delta_{\beta 0} - \delta_{\alpha\beta}\Big|^{3}_{\alpha,\beta=1}\Big)\delta_{ab}$$

$$+\Big[\delta_{\alpha 0}\epsilon_{a\beta b} + \delta_{\beta 0}\epsilon_{a\alpha b} + \delta_{\alpha a}\delta_{\beta b} + \delta_{\alpha b}\delta_{\beta a}\Big]^{3}_{\alpha,\beta=1}\Bigg) \tag{3.5.23}$$

$$= -\frac{1}{4}\Big[(|\Omega_{0}|^{2} - |\Omega_{j}|^{2})\ell^{k}_{\mu}r^{k\,\mu} + 2\Re\Big[\Omega^{\star}_{0}\Omega_{j}\epsilon_{jab} + \Omega^{\star}_{a}\Omega_{b}\Big]\ell^{a}_{\mu}r^{b\,\mu}\Big]. \tag{3.5.24}$$

In the last line the indices of Ω have been changed from upper to lower ones only for notational reasons and all Latin indices run from 1 to 3. $2\Re[Z] = (Z+Z^{\star})$ is the real part of the argument. One can get from the second to the third line by noticing that this time the bracket is symmetric in α and β[24]. Combining this with the fact that α and β connect a complex scalar field, Ω_{j}, with its complex conjugate, $\Omega^{\star}_{j}$, leads to the above identification with the real part.

Now the complete quadratic interaction between the gauge fields and the scalar field, Ω, can be written

[24] The reason for the symmetry in indices rather than the previously encountered antisymmetry is the different structure of Pauli matrices in the trace.

down explicitly. Combining Eq. (3.5.20) with (3.5.21) and (3.5.24) leads to:

$$\mathcal{L}^{\text{quad}} = \frac{g_{\text{A}}^2}{4}\Big(\text{tr}_{\text{I}}\big[\Omega\Omega^{\dagger}\ell_\mu\ell^\mu\big] + \text{tr}_{\text{I}}\big[\Omega^{\dagger}\Omega\, r_\mu r^\mu\big] - 2\text{tr}_{\text{I}}\big[\Omega^{\dagger}\ell_\mu\Omega r^\mu\big]\Big) \tag{3.5.25}$$

$$= \frac{g_{\text{A}}^2}{4}\text{tr}_{\text{I}}\Big[\big(\ell_\mu\Omega - \Omega r_\mu\big)^{\dagger}\big(\ell^\mu\Omega - \Omega r^\mu\big)\Big] \tag{3.5.26}$$

$$= \frac{g_{\text{A}}^2}{2}|\Omega_0|^2\Big|\frac{1}{2}(\ell_\mu^k - r_\mu^k)\Big|^2 + \frac{g_{\text{A}}^2}{2}|\Omega_j|^2\Big|\frac{1}{2}(\ell_\mu^k + r_\mu^k)\Big|^2 - \frac{g_{\text{A}}^2}{2}\Re\big[\Omega_a^\star\Omega_b + \Omega_0^\star\Omega_j\epsilon_{jab}\big]\ell_\mu^a r^{b\mu}\,. \tag{3.5.27}$$

The second line is just inserted here as it is a neat abbreviation of the first line, which can be gained by using the cyclicality of the trace. In addition it may help to understand the structural complications that will show up in the final mass terms (Eq. (3.5.29)). In line Eq. (3.5.27) the absolute value of the gauge field is just an abbreviation standing for: $|(\ell_\mu^k - r_\mu^k)/2|^2 = (\ell_\mu^k - r_\mu^k)(\ell^{k\mu} - r^{k\mu})/4$. Again all indices are to be summed over. Using this result (Eq. (3.5.27)) it is possible to work out all allowed quadratic interactions between scalar fields and the gauge fields[25)]. While the interactions are interesting in any later calculation, for now it is only important, that no prohibited interactions, such as charge creating terms, occur. As any explicit derivation of the vanishing problematic terms would directly result in calculating all of them, the reader is left with the unpleasant comment: "Rest assured, they all vanish." Of course, everyone is invited to check these calculations.

In contrast to this slightly ignorant comment, there is still the gauge field mass, which is crucial in the presented model. The gauge field mass is considered as an intersting topic, as this mass led to the complications of constrained instantons in the earlier presentations (Sec. II.9.8) and so the outcome for this mass in the present model is of interest. It can be calculated by analysing Eq. (3.5.27) in the case where Ω going to its VEV. As in Sec. III.4 the vacuum is assumed to be charge neutral and parity even. This means that the only possible contributions come form the σ- and the δ_3-fields: $\langle\Omega^\alpha q^\alpha\rangle = R_\sigma I + R_\delta\tau_3$. In complex vector notation one has: $\langle\Omega_0\rangle = R_\sigma$, $\langle\Omega_3\rangle = \text{i}R_\delta$ and $\langle\Omega_1\rangle = \langle\Omega_2\rangle = 0$. Using this, Eq. (3.5.27) gives the following masses for combined gauge fields:

$$\mathcal{L}^{\text{quad}}(\langle\Omega\rangle) = \frac{g_{\text{A}}^2}{8}\Big(R_\sigma^2|\ell_\mu^k - r_\mu^k|^2 + R_\delta^2|\ell_\mu^k + r_\mu^k|^2 - 4R_\delta^2\ell_\mu^3 r^{3\mu}\Big) \tag{3.5.28}$$

$$= \frac{g_{\text{A}}^2}{2}(R_\sigma^2 + R_\delta^2)\Big|\frac{\ell_\mu^3 - r_\mu^3}{2}\Big|^2 + \frac{g_{\text{A}}^2}{2}R_\sigma^2\Big(\Big|\frac{\ell_\mu^1 - r_\mu^1}{2}\Big|^2 + \Big|\frac{\ell_\mu^2 - r_\mu^2}{2}\Big|^2\Big)$$

$$+ \frac{g_{\text{A}}^2}{2}R_\delta^2\Big(\Big|\frac{\ell_\mu^1 + r_\mu^1}{2}\Big|^2 + \Big|\frac{\ell_\mu^2 + r_\mu^2}{2}\Big|^2\Big) \tag{3.5.29}$$

$$= \frac{1}{2}(g_{\text{A}}R_\sigma)^2\,|\boldsymbol{A}^\mu|^2 + \frac{1}{2}(g_{\text{A}}R_\delta)^2\Big(|A_3^\mu|^2 + |V_1^\mu|^2 + |V_2^\mu|^2\Big)\,. \tag{3.5.30}$$

[25)]For this the ladder operator representation $A_+ \sim (A_1 + \text{i}A_2)$; $A_- \sim (A_1 - \text{i}A_2)$ is useful.

Note that only one term survives, which mixes left- and right-handed fields from Eq. (3.5.27). This is due to the particular choice of the VEV, that leads to $\left(2\Re[\langle\Omega\rangle_a^\star\langle\Omega\rangle_b + \langle\Omega\rangle_0^\star\langle\Omega\rangle_j\epsilon_{jab}]_{(a\neq b)} = 0\right)$. In the last line (Eq. (3.5.30)) the (axial-) vector basis has been introduced, $A_\mu^j := (\ell_\mu^j - r_\mu^j)/2$ and $V_\mu^j := (\ell_\mu^j + r_\mu^j)/2$. The mixing of scalar and gauge fields gives rise to the masses of ℓ_μ, r_μ and some linear combinations of them. As the instantons parts can be constructed from the general left- and right-handed gauge fields, the above equations determine their mass spectrum and effective low energy structure.

One sees that the strict separation into instanton and anti-instanton (which would be ℓ_μ^{con} and r_μ^{con} here) breaks down and both parts become coupled through the scalar field VEV. This does not lead to new complications, as it is just the already expected low energy behaviour of constrained instantons - compare Sec. II.9.8.1 and Eq. (2.9.71). While the constraints that produced the constrained instanton solution guaranteed the existence of instanton-like structures for small distances ($|x| << \rho$), they implicitly forced these structures to break down at large distances ($|x| >> \rho$). Ultimately this meant that the gauge fields behaved as free, massive particles in the low energy regime (compare Eq. (2.9.71)). So, just as for fermions (Sec. II.5), the left- and right-handed gauge field parts become coupled through the mass term.

To get some acquaintance with the mass terms first focus on the case of vanishing δ-meson VEV ($R_\delta \to 0$). This scenario establishes the connection of the present model to its ancestor - the linear σ-model. In Eq. (3.5.30) only the first term survives, which means that only the axial gauge field acquires a mass of $M^{\text{ax}} = R_\sigma$. In the language of group theory the axial generators, T_{ax}^{ij}, are broken by the VEV $[T_{\text{ax}}^{ij}\langle\Omega\rangle_j \neq 0]$ and the vector generators, T_{vec}^{ij}, remain unbroken $[T_{\text{vec}}^{ij}\langle\Omega\rangle_j = 0]$[26)]. So by expanding the model around its VEV the axial symmetry is lost, while the vectorial part remains a symmetry of the model. The general structure of spontaneously broken gauge symmetries is very interesting itself, but unfortunately more details on this analysis have to be discussed in one of the introductory textbooks of field theory (e.g. Ref. [5, p.526-542]). For now it is sufficient to observe that the six gauge fields are grouped into two triplets - the massive axial fields and the massless vectorial fields.

The situation significantly changes if one allows for a nonzero VEV of the δ-meson ($R_\delta \neq 0$). In this case the well ordered structure of the σ-model gets somewhat messed up. Now the nice triplet separation vanishes. In the axial sector the third component couples to both VEV contributions, leading to a mass of $M_3^{\text{ax}} = g_{\text{A}}\sqrt{R_\sigma^2 + R_\delta^2}$, while the other two axial components only couple to the σ-meson VEV, which gives them only a mass of $M_{1,2}^{\text{ax}} = g_{\text{A}}R_\sigma$. In the vectorial sector the triplet breaks down as well. Here the generators of V_1^μ and V_2^μ are broken by the δ-meson VEV and thus get a mass of $M_{1,2}^{\text{vec}} = g_{\text{A}}R_\delta$. The third component of the vectorial gauge field remains the only massless field in this model. So the V_3-direction

[26)]The VEV breakinging of the generators for the linear σ-model is nicely accessible in a 4 dimensional $SO(4)$ representation. Here the VEV is $\langle\Omega\rangle = (\sigma_0, 0, 0, 0)^{\text{T}}$ and the generators are given in the following matrix:

$$T^4 = \begin{pmatrix} 0 & a_1 & a_2 & a_3 \\ -a_1 & 0 & v_3 & v_1 \\ -a_2 & -v_3 & 0 & v_2 \\ -a_3 & -v_1 & -v_2 & 0 \end{pmatrix} \tag{3.5.31}$$

corresponds to the symmetry, which remains intact in the transition to the VEV of the model[27)].

III.5.4 Currents and gauge fields

In the previous section the quadratic interaction between gauge and scalar fields was discussed. Looking at Eq. (3.5.15), one finds that the last contribution (Eq. (3.5.18)) still has to be discussed. This part couples the scalar currents to the gauge fields, ℓ_μ and r_μ. With the two relation Eq. (A.12) and (A.13) the terms can be evaluated in a similar way as the quadratic contributions in the previous section. Just as before, b_μ does not contribute in the final result, as all terms are still part of the scalar kinetic term (compare the discussion in Sec. III.5.2).

$$\mathcal{L}^{\text{cur}}_{\Omega,\ell} := -\frac{\mathrm{i}g_\mathrm{A}}{4}\mathrm{tr}_\mathrm{I}\Big[\ell_\mu\Big(\Omega\{\partial^\mu\Omega^\dagger\} - \{\partial^\mu\Omega\}\Omega^\dagger\Big)\Big]_{b_\mu=0} \tag{3.5.32}$$

$$= -\frac{\mathrm{i}g_\mathrm{A}}{4}\ell^a_\mu\Big\{\Omega^\alpha\{\partial^\mu(\Omega^\star)^\beta\} - \{\partial^\mu\Omega^\alpha\}(\Omega^\star)^\beta\Big\}\frac{1}{2}\mathrm{tr}_\mathrm{I}\Big[\tau^a q^\alpha\bar{q}^\beta\Big] \tag{3.5.33}$$

$$= -\frac{\mathrm{i}g_\mathrm{A}}{4}\boldsymbol{\ell}_\mu\,\mathrm{i}\Big\{\Big(\Omega_0\{\partial^\mu\boldsymbol{\Omega}^\star\} - \boldsymbol{\Omega}\{\partial^\mu\Omega_0^\star\}\Big) - \Big(\{\partial^\mu\Omega_0\}\boldsymbol{\Omega}^\star - \{\partial^\mu\boldsymbol{\Omega}\}\Omega_0^\star\Big)$$
$$+\boldsymbol{\Omega}\times\{\partial^\mu\boldsymbol{\Omega}^\star\} - \{\partial^\mu\boldsymbol{\Omega}\}\times\boldsymbol{\Omega}^\star\Big\} \tag{3.5.34}$$

$$= \frac{g_\mathrm{A}}{2}\boldsymbol{\ell}_\mu\,\Re\Big\{-\Big[\{\partial^\mu\Omega_0\}\boldsymbol{\Omega}^\star - \Omega_0\{\partial^\mu\boldsymbol{\Omega}^\star\}\Big] + \boldsymbol{\Omega}\times\{\partial^\mu\boldsymbol{\Omega}^\star\}\Big\}. \tag{3.5.35}$$

From the third line on a vector notation was used to prevent a cluttered index notation. As in Sec. III.4.1, bold symbols correspond to a '3-vector' (e.g.: $\boldsymbol{\Omega} = (\Omega_1, \Omega_2, \Omega_3)^\mathrm{T}$) and the symbol $\times$ stands for the cross product[28)]. To arrive at line Eq. (3.5.35) it was used, that the components of Ω are scalars and thus commute. Analogously the right-handed current from Eq. (3.5.15) can be evaluated. One finds a very similar result:

$$\mathcal{L}^{\text{cur}}_{\Omega,r} := -\frac{\mathrm{i}g_\mathrm{A}}{4}\mathrm{tr}_\mathrm{I}\Big[r_\mu\Big(\Omega^\dagger\{\partial^\mu\Omega\} - \{\partial^\mu\Omega^\dagger\}\Omega\Big)\Big]_{b_\mu=0} \tag{3.5.36}$$

$$= \frac{g_\mathrm{A}}{2}\boldsymbol{r}_\mu\,\Re\Big\{\{\partial^\mu\Omega_0\}\boldsymbol{\Omega}^\star - \Omega_0\{\partial^\mu\boldsymbol{\Omega}^\star\} + \boldsymbol{\Omega}\times\{\partial^\mu\boldsymbol{\Omega}^\star\}\Big\}. \tag{3.5.37}$$

[27)] The present jargon is heavily based on the standard treatment of spontaneous symmetry breaking in the electroweak Higgs model, which can be reviewed in Ref. [34, p.147-155] or [5, p.526-542].

[28)] $(\boldsymbol{a}\times\boldsymbol{b})_i := \epsilon_{ijk}a_j b_k$

Compared to the left-handed current the only difference is the sign of the 'bracket part'. Now Eq. (3.5.35) and (3.5.37) can be combined to find the total scalar current:

$$\mathcal{L}_{\Omega}^{\text{cur}} := \mathcal{L}_{\Omega,l}^{\text{cur}} + \mathcal{L}_{\Omega,r}^{\text{cur}} \tag{3.5.38}$$

$$= g_{\text{A}} \underbrace{\frac{1}{2}(\boldsymbol{r}_\mu - \boldsymbol{\ell}_\mu)}_{=\boldsymbol{A}_\mu} \Re\Big\{\{\partial^\mu \Omega_0\}\boldsymbol{\Omega}^\star - \Omega_0\{\partial^\mu \boldsymbol{\Omega}^\star\}\Big\} + g_{\text{A}} \underbrace{\frac{1}{2}(\boldsymbol{\ell}_\mu + \boldsymbol{r}_\mu)}_{=\boldsymbol{V}_\mu} \Re\Big\{\boldsymbol{\Omega} \times \{\partial^\mu \boldsymbol{\Omega}^\star\}\Big\}. \tag{3.5.39}$$

In Sec. III.4.1 the conserved scalar currents (Noether currents) where derived in the case of a global $SU_{\text{L}}(2) \otimes SU_{\text{R}}(2) \otimes U_{\text{V}}(1)$ symmetry. Now, comparing Eq. (3.4.50) and (3.4.53) with the result above (Eq. (3.5.39)), shows that the local gauge model in this section has a nice connection to the model with a global symmetry discussed in Sec. III.4.1. This time the symmetry transformations are local ($\varepsilon_{\ell/r} = \varepsilon_{\ell/r}(x)$), leading to the left- and right-handed gauge fields and ultimately to their vectorial and axial equivalences. The above equation (Eq. (3.5.39)) now shows that these gauge fields couple to currents of the same form as the earlier found Noether currents of the global symmetry case. From Sec. III.4.1.1 one knows that the currents connected to global symmetries where directly connected to the symmetry breaking terms in the scalar Lagrangian (Eq. (3.4.9)). Therefore this connection gives a nice starting point to examine the influence of the symmetry breaking terms on a gauge field model with local symmetries.

Before ending the discussion of the gauge field coupling terms it should be mentioned that a derivation of conserved currents in this (local gauge) model is not as straight forward as it was before in the global symmetry case. The reason for this is that one has to choose a particular gauge in order to remove gauge redundancies from the model. This gauge fixing contribution complicates things slightly and makes it more difficult to identify the complete conserved currents. The next section will give the most important relations to fix a gauge for the derived model.

III.5.5 Gauge fixing and ghost fields

In preceding sections from Sec. III.5 up to this point, the global gauge model from Sec. III.3 to III.4.1.1 was step by step promoted to its *local* counter part. This was done in the left- and right-handed spinor basis, as the scalar and nucleon sector of the model was explicitly constructed in this representation. The last step that remains to be done is to fix a gauge for the non-abelian gauge fields, $\boldsymbol{A}_\mu$ and $\boldsymbol{V}_\mu$. Compared to the abelian case there are some complications, when chosing a gauge for a non-abelian gauge model. The reason for this is directly connected to the non commutiativity of the gauge field generators in such a model. They lead to additional contributions, that will spoil the advantages of a naively chosen gauge. The procedure, that leads to a suitable, convenient gauge in non-abelian models is nicely introduced in Ref. [5, p.430-434]. For the present discussion only the main general results will be adopted.

- If a particular gauge fixing term $\mathcal{L}_{\rm gf}$ is included in the Lagrangian, then this leads to another additional term in the Lagrange densitiy, which is usually called the ghost Lagrangian $\mathcal{L}_{\rm gh}$.
- Take a gauge fixing term $\mathcal{L}_{\rm gf} = -K^k K^k/(2\xi)$, with some gauge fixing function $K^k = K^k(x)$. Then the corresponding ghost Lagrangian has the form:

$$\mathcal{L}_{\rm gh} = -(c^\dagger)^k \frac{\delta K^k}{\delta \vartheta^l} c^l \,. \tag{3.5.40}$$

- Here $(c^\dagger)^k$ and c^l are the ghost fields. They are complex grassmann valued fields. They are no physical particles, but in the formal derivation of the gauged Lagrangian they show up (and affect observables, like scattering amplitudes of physical particles) and thus they have to be included.
- The transformation of a gauge field, A^a_μ, under a symmetry transformation with the parameters $\vartheta^c = \vartheta^c(x)$ can be written as:

$$U^{-1}(\vartheta^c) A^a_\mu U(\vartheta^c) = A^a_\mu - (D_\mu)^{ac}\vartheta^c \,. \tag{3.5.41}$$

With these tools the gauge of the earlier presented model can be fixed and the corresponding ghost Lagrangian can be found. For this procedure only the interactions of the gauge fields and the scalar sector will be of interest.
Earlier the spinor basis was chosen to derive all terms in the Lagrange density. For the present task it is more convenient to rewrite the spinor representation into a higher dimensional real representation of the same symmetry group. The non-abelian gauge symmetries of interest have their generators in iso-spinor space ($\ell_\mu = \ell^a_\mu \tau^a/2$ and r_μ analogously). From the construction of the field $\Omega = (\Phi + {\rm i}\Lambda)^\alpha q^\alpha_{\rm I}$ and the complete scalar Lagrangian one knows that the model is invariant under a 4-dimensional rotational symmetry acting on the real or on the imaginary part of Ω (compare the first part of Sec. III.4). This symmetry corresponds to a translation of the iso-spinor representation into a real 4-dimensional one ($SO(4) \simeq SU(2) \otimes SU(2)$ compare Sec. II.2).
But, since $\Omega^\alpha = (\Phi^\alpha + {\rm i}\Lambda^\alpha)$ is a complex prefactor to the quaternion symbol $q^\alpha_{\rm I}$, the translation into a 4-dimensional space is not sufficient to generate a real representation of the symmetry group in the present case. For this one has to enlarge the 4-dimensional representation to an 8-dimensional one. In the first four elements of the space lives the ordinary σ-model (Φ-field) and in the second set lives the newly introduced Λ-field: $\Omega_{\rm v} = (\Phi^{\rm T}, \Lambda^{\rm T})^{\rm T}$. Unfortunately in the 8-dimensional representation one does not know the correct generators of the symmetry group, $SU_{\rm L}(2) \otimes SU_{\rm R}(2)$. To find them the results from Sec. III.5.3 and III.5.4 come in quite handy. Eq. (3.5.27) and (3.5.39) show, that there are no interactions between the Φ- and the Λ-part of the model. On the other hand, looking at the appearing terms in these equations reveals that they just correspond to the ordinary linear σ-model interactions. As one already knows that this model has a $SO(4)$ symmetry, one can now construct the 8-dimensional generators. Take

all generators of 4-dimensional rotations in the vectorial/axial-representation as:

$$\widetilde{G}^4 = \mathrm{i}T_{SO(4)} = \begin{pmatrix} 0 & a_1 & a_2 & a_3 \\ -a_1 & 0 & v_3 & v_1 \\ -a_2 & -v_3 & 0 & v_2 \\ -a_3 & -v_1 & -v_2 & 0 \end{pmatrix}^{29)} \qquad a_i = v_i = 1\,. \tag{3.5.42}$$

From this define a single generator as $\widetilde{G}^4_j$, with $j \in \{1, 2, ..., 6\}$. This shall be understood as $\widetilde{G}^4_1$, $\widetilde{G}^4_2$, $\widetilde{G}^4_3$ being the axial generators and $\widetilde{G}^4_4$, $\widetilde{G}^4_5$, $\widetilde{G}^4_6$ the vectorial generators. The 8-dimensional generators are then:

$$G_k = \begin{pmatrix} \widetilde{G}^4_k & 0 \\ 0 & \widetilde{G}^4_k \end{pmatrix} = \widetilde{G}^4_k \oplus \widetilde{G}^4_k \qquad ; \qquad k \in \{1, 2, ..., 6\}\,. \tag{3.5.43}$$

Each entry here represents a (4×4)-matrix. The off-diagonal zeros reflect the fact that there are no interactions between the Φ and the Λ part of the model. With the 8-dimensional generators of the symmetry group the six gauge field, ℓ^k_μ and r^l_μ, in this representation are: $A_\mu = -\mathrm{i}A^k_\mu G_k$. The factor of $-\mathrm{i}$ is included, to compensate the additional factor of $+\mathrm{i}$ in the generators (Eq. (3.5.42)). And now the covariant derivative becomes:

$$D^{ab}_\mu = (\partial_\mu - \mathrm{i}g_\mathrm{A} A_\mu)^{ab} = \delta^{ab}_8 \partial_\mu - g_\mathrm{A} A^k_\mu (G_k)^{ab} \tag{3.5.44}$$

$$= \delta^{ab}_8 \partial_\mu - g_\mathrm{A}[(A^\mathrm{con})_\mu + a_\mu]_k (G_k)^{ab} = \bar{D}^{ab}_\mu - g_\mathrm{A} a_{k\mu} (G_k)^{ab}\,. \tag{3.5.45}$$

As usually the unit matrix in front of the derivative will be left out from now on. In the second line the gauge field has been splitt up into an instanton contribution, $(A^\mathrm{con})_{k\mu}$, and a fluctuation, $a_{k\mu}$. For the final equality the abbreviation $\bar{D}_\mu := \partial_\mu - g_\mathrm{A}(A^\mathrm{con})_\mu$ was introduced. Using such a separation allows to express the model in an extended background gauge. The pure background gauge for the instanton formalism (without a scalar VEV) was introduced in Sec. II.9.3. The results from that section will also be used, but in the present case the presence of Ω makes the situation a bit more complicated.

Note that all terms in the covariant derivative are real in this representation. This is a general property of the real $SO(N)$ representation and the reason, why it is convenient to work out the explicit gauge fixing terms in this picture. The scalar field in the vectorial representation is: $|\Omega\rangle := \Omega_\mathrm{v} = (\sigma, \boldsymbol{\pi}^\mathrm{T}, \eta, \boldsymbol{\delta}^\mathrm{T})^\mathrm{T}$. For the calculations to come it is useful to split the field up into a VEV contribution, $|v\rangle := (\sigma_0 \hat{\boldsymbol{e}}_1 + \delta_0 \hat{\boldsymbol{e}}_8)$, and a space-time dependent part, $|\omega\rangle$, which incorporates the physical meson fields. The covariant derivative

[29] The factor of i was included to make the matrices purely real (the generators of rotations are purely imaginary).

acting on the scalar field then becomes

$$(D^\mu)^{ab}|\Omega^b\rangle = D^\mu|\Omega\rangle = \Big[\partial^\mu - g_\mathrm{A}A_k^\mu G_k\Big]\Big(|v\rangle + |\omega\rangle\Big) \tag{3.5.46}$$

$$= \bar{D}^\mu|\omega\rangle - g_\mathrm{A}(A_k^\mathrm{con})^\mu G_k|v\rangle - g_\mathrm{A}a_k^\mu\Big(G_k|v\rangle + G_k|\omega\rangle\Big) \tag{3.5.47}$$

$$= |\bar{D}^\mu\omega\rangle - g_\mathrm{A}(A_k^\mathrm{con})^\mu|F_k\rangle - g_\mathrm{A}a_k^\mu(|F_k\rangle + |\omega\rangle)\,. \tag{3.5.48}$$

In the last line the definitions $|F_k\rangle := G_k|v\rangle$ and $|\bar{D}^\mu\omega\rangle := \bar{D}^\mu|\omega\rangle$ have been used. As this expression is completely real, the kinetic energy is just given by the scalar product of this expression with itself:

$$\mathcal{L}_\Omega^\mathrm{kin} = \frac{1}{2}\Big(D_\mu^{ab}\Omega_\mathrm{v}^b\Big)\Big((D^\mu)^{ac}\Omega_\mathrm{v}^c\Big) = \frac{1}{2}\langle\Omega|\overleftarrow{D}_\mu^\mathrm{T} D^\mu|\Omega\rangle \tag{3.5.49}$$

$$= \frac{1}{2}\Big(\langle\bar{D}_\mu\Omega| - g_\mathrm{A}\langle\Omega|a_\mu\Big)\Big(|\bar{D}^\mu\Omega\rangle - g_\mathrm{A}a^\mu|\Omega\rangle\Big) \tag{3.5.50}$$

$$= \frac{1}{2}\langle\bar{D}_\mu\Omega|\bar{D}^\mu\Omega\rangle + g_\mathrm{A}a_k^\mu\langle\Omega|G_k|\bar{D}_\mu\Omega\rangle + \frac{g_\mathrm{A}^2}{2}a_{k\mu}a_l^\mu\langle\Omega|G_kG_l|\Omega\rangle \tag{3.5.51}$$

$$= \frac{1}{2}\langle\bar{D}_\mu\omega|\bar{D}^\mu\omega\rangle \tag{3.5.52}$$

$$+ \frac{g_\mathrm{A}^2}{2}a_{k\mu}a_l^\mu\Big(\langle F_k|F_l\rangle + 2\langle F_k|G_l|\omega\rangle + \langle\omega|G_kG_l|\omega\rangle\Big) - g_\mathrm{A}a_k^\mu\langle\omega|G_k|\bar{D}_\mu\omega\rangle \tag{3.5.53}$$

$$+ \underbrace{\frac{1}{2}g_\mathrm{A}^2\langle F_k|F_l\rangle\,(A_k^\mathrm{con})^\mu A_{l\mu}^\mathrm{con}}_{\to m_\mathrm{higgs}} - g_\mathrm{A}\underbrace{\Big((A_k^\mathrm{con})^\mu + a_k^\mu\Big)}_{=A_k^\mu}\langle F_k|\bar{D}_\mu\omega\rangle\,. \tag{3.5.54}$$

To arrive at the last equality only fundamental algebra was used[30)], and everything was rearranged such that the different constituents of scalar and gauge field are well separated. The above equation is in fact already well known from Eq. (3.5.15) and does only come in a little disguise here. The advantage of this representation is the separation into distinct contributions. Line Eq. (3.5.52) gives the kinetic term of the physical scalar mesons, ω, and, via the covariant derivative, $\bar{D}_\mu$, their interactions with instantons, A_μ^con. In the next line (Eq. (3.5.53)) the interactions with the gauge field fluctuations, a_μ, are treated and the final line incorporates the somewhat special contributions. By integrating over the first term of line Eq. (3.5.54), this gives the Higgs measure contribution, m_higgs, discussed in Sec. II.9.8. The final term in line Eq. (3.5.54) is slightly odd, as it couples the gauge field via the VEV to the derivative of the scalar field. Fortunately this contribution will not spread confusion for too long, since it can be gauged away, using the so called R_ξ gauge. Before doing this, the 'mass matrix', $M_\mathrm{A}^{kl} = g_\mathrm{A}^2\langle F_k|F_l\rangle$, can be introduced, which generates the masses of instantons and the fluctuations a_μ. Also the earlier discussed coupling of the gauge field to currents (Eq. (3.5.39)) can be identified: $\mathcal{L}_\Omega^\mathrm{cur} \sim -g_\mathrm{A}A_k^\mu\langle\omega|G_k|\partial_\mu\omega\rangle$.

[30)] For this note that

$$\langle\bar{D}_\mu\Omega|\bar{D}^\mu\Omega\rangle = \langle\bar{D}_\mu\omega|\bar{D}^\mu\omega\rangle - 2g_\mathrm{A}A_{k\mu}^\mathrm{con}\langle F_k|\bar{D}^\mu\omega\rangle + g_\mathrm{A}^2(A_k^\mathrm{con})^\mu A_{l\mu}^\mathrm{con}\langle F_k|F_l\rangle \tag{3.5.55}$$

Now, as the new aspects of this representation are discussed, the R_ξ gauge can be introduced. For this take the gauge fixing functional, $K_k = \left(\bar{D}_\mu A_k^\mu - \xi g_{\rm A}\langle\omega|F_k\rangle\right)$, to construct the gauge fixing Lagrangian. Again $\bar{D}_\mu$ is the covariant derivative only containing the instanton gauge field. This gauge corresponds to the so called background gauge, which was already discussed (in the absence of a scalar VEV in Sec. II.9.3). As gauge fixing Lagrangian one obtains

$$\mathcal{L}_{\rm gf} = \frac{1}{2\xi}K_k K_k = \frac{1}{2\xi}\Big(\bar{D}_\mu a_k^\mu - \xi g_{\rm A}\langle\omega|F_k\rangle\Big)\Big(\bar{D}_\mu A_k^\mu - \xi g_{\rm A}\langle\omega|F_k\rangle\Big) \tag{3.5.56}$$

$$= \frac{1}{2\xi}(\bar{D}_\mu A_k^\mu)(\bar{D}_\nu A_k^\nu) + \frac{\xi g_{\rm A}^2}{2}\langle\omega|F_k\rangle^2 \underbrace{-g_{\rm A}\bar{D}_\mu A_k^\mu\langle\omega|F_k\rangle}_{=g_{\rm A}A_k^\mu\langle\bar{D}_\mu\omega|F_k\rangle}\,. \tag{3.5.57}$$

The last term now nicely compensates the last contribution in line Eq. (3.5.54). The parameter ξ is an arbitrary gauge parameter, which means on the one hand that it can be adjusted to simplify calculations and on the other hand no physical quantity can depend on it, as it is arbitrary. To arrive at the gauge from Sec. II.9.3 one needs $\xi = -1/2$.

Finally, as a last step, the ghost Lagrangian can be calculated. For this, one needs the transformation of the gauge fixing term, K_k, under the symmetry group. From Eq. (3.5.41) one knows how the gauge field transforms under symmetry transformations and one only needs find the transformation of the scalar field. In Eq. (3.5.43) the 8-dimensional generators of the symmetry group where given and so $|\omega\rangle$ transforms under an infintessimal transformation as:

$$|\omega\rangle \to |\omega\rangle - g_{\rm A}\vartheta_k G_k|\Omega\rangle\,. \tag{3.5.58}$$

With this, the derivative of the gauge fixing term, K_a, with respect to ϑ_b becomes

$$\frac{\delta K_a}{\delta\vartheta_b} = -\bar{D}_\mu(D^\mu)^{ab} + g_{\rm A}^2\xi\langle F_a|G_b|v\rangle + g_{\rm A}^2\xi\langle F_a|G_b|\omega\rangle \tag{3.5.59}$$

$$= -\bar{D}_\mu(D^\mu)^{ab} + \xi(M_A^2)^{ab} + \xi g_{\rm A}^2\langle F_a|G_b|\omega\rangle\,. \tag{3.5.60}$$

and thus the ghost Lagrangian is at hand using Eq. (3.5.40):

$$\mathcal{L}_{\rm gh} = -(\bar{D}_\mu c^{\dagger\,a})(D^\mu)^{ab}c^b - \xi c^{\dagger\,a}(M_A^2)^{ab}c^b - \xi g_{\rm A}^2 c^{\dagger\,a}\langle F_a|G_b|\omega\rangle c^b\,. \tag{3.5.61}$$

This closes the calculation of the $\mathcal{L}_{\rm gf}$ and $\mathcal{L}_{\rm gh}$. If the fluctuation field a_μ only has the size of quantum fluctuations (that is to say there is no dynamical vector meson contribution), then

III.6 The complete model

As a closing step the different parts of the effective model, which have been developed throughout the previous sections, are reassembled to give the complete Lagrange density. As all details, concerning any contribution have been discussed in detail already, this section will only give a very brief summary on the different parts. The purpose of this 'rediscussion' of the complete Lagrange density is to emphasise the main aspects of the introduced model without cluttering everything with too many details.
As a starting point of the model the Lagrange density from Eq. (3.2.13) was used:

$$\mathcal{L}_{\text{model}} = \mathcal{L}_N + \mathcal{L}_\Omega + \mathcal{L}_{\text{gfh}} + \frac{a^2}{4}\Big(e^{-\mathrm{i}\theta}\Omega_\alpha^2 + e^{+\mathrm{i}\theta}\Omega_\alpha^{\dagger\,2}\Big)\theta(s\rho - x)\,, \tag{3.6.1}$$

where a was an effective parameter, due to induced instanton zero-mode interactions. Of course, although not explicitly indicated, all terms were effective contributions. For calculatory reasons the Heaviside function was dropped and from experimental observations (there are no observed parity violating contributions to strong interactions) the vacuum angle was set to $\theta = 0$. With these simplifications the instanton related term in the Lagrange density turns into an effective contribution to the scalar potential and thus the effective Lagrange density becomes:

$$\mathcal{L}_{\text{model}} = \mathcal{L}_N + \mathcal{L}_\Omega + \mathcal{L}_{\text{gfh}} \tag{3.6.2}$$

$$= \mathcal{L}_N^{\text{f}} + \mathcal{L}_\Omega^{\text{f}} + \mathcal{L}_{\text{gfh}} + \mathcal{L}_{NA} + \mathcal{L}_{N\Omega} + \mathcal{L}_{\Omega A} + \mathcal{L}_{\Omega\Omega}\,. \tag{3.6.3}$$

In the first line the instanton part is absorbed in $\mathcal{L}_\Omega$ and the second line is only a rearrangement. Here the superscript f stands for the 'free' Lagrange densities, while the parts with two subscript fields are the interaction parts between 'nucleons', N, scalars, Ω, and the (instanton) gauge field, A_μ. By scavenging through the previous sections all different parts in Eq. (3.6.3) can be identified.
The nucleon Lagrangian was given in Eq. (3.3.14) and with this one finds:

$$\mathcal{L}_N^{\text{f}} = \overline{N}[\mathrm{i}\gamma_\mu\partial^\mu + M_{\text{N}}]N\,, \tag{3.6.4}$$

where $M_{\text{N}} = g_\Omega\,\text{diag}(R_\sigma + R_\delta, R_\sigma - R_\delta)$ is the nucleon's mass matrix with the VEV contributions from the scalar sector, as it was introduced in Sec. III.4. This mass matrix leads to a generalisation of the Goldenberger-Treimann relation for the case of non-vanishing $\delta-$meson VEV.
The free scalar Lagrangian can be obtained from the final effective potential (expanded around the minimum) of Eq. (3.4.37):

$$\begin{aligned}\mathcal{L}_\Omega^{\text{f}} = &\frac{1}{2}\Big[\partial_\mu\sigma\partial^\mu\sigma + \partial_\mu\boldsymbol{\pi}\partial^\mu\boldsymbol{\pi} + \partial_\mu\eta\partial^\mu\eta + \partial_\mu\boldsymbol{\delta}\partial^\mu\boldsymbol{\delta}\Big]\\ &-\frac{1}{2}\Big[m_\sigma^2\sigma^2 + m_\delta^2\delta_3^2 + m_\pi^2\boldsymbol{\pi}^2 + m_{\delta\eta}^2(\delta_1^2 + \delta_2^2 + \eta^2)\Big]\,.\end{aligned} \tag{3.6.5}$$

The intimate relations between the different parameters of the free nucleon and scalar Lagrangian can be reviewed in Sec. III.4 and in more detail, for explicit parameter fixing schemes in Sec. III.4.3. As earlier, one has to keep in mind, that the mesons, indicated above, could be identified with various physically observable states. So far the lowest lying meson states have been used, but in principle one is free to choose other states, as long as the particular mass splitting is maintained ($m_\sigma > m_\pi$ and $m_\delta > m_\eta$). Especially the here labelled $\eta-$meson cannot be identified naively with its physical counterpart. The problems, connected to this identification were discussed in Sec. III.4.3.1.
The interactions between fermionic and scalar fields were chosen in a very simple fashion in Eq. (3.3.11):

$$\mathcal{L}_{N\Omega} = g_\Omega(\overline{N}_\mathrm{L}\Omega^\dagger N_\mathrm{R} + \overline{N}_\mathrm{R}\Omega N_\mathrm{L}) = g_\Omega\overline{N}\left[(\sigma - \mathrm{i}\gamma_5\eta) + \boldsymbol{\tau}(\boldsymbol{\delta} + \mathrm{i}\gamma_5\boldsymbol{\pi})\right]N \tag{3.6.6}$$

and there is not much more to say about these. Coming to self-interactions of the scalar field, there is again not too much to say about them. In principle they are a direct consequence of the employed 'Higgs-mechanism' in the scalar sector and their coefficients were found in Eq. (3.4.37).

$$\begin{aligned}\mathcal{L}_{\Omega\Omega} =&\lambda^2(R_\sigma\sigma + R_\delta\delta_3)(\sigma^2 + \boldsymbol{\pi}^2 + \eta^2 + \boldsymbol{\delta}^2) + \frac{\lambda^2}{4}(\sigma^2 + \boldsymbol{\pi}^2 + \eta^2 + \boldsymbol{\delta}^2)^2\\ &+ 2\lambda^2 R_\sigma R_\delta \sigma\delta_3\,.\end{aligned} \tag{3.6.7}$$

While the first line, with the cubic and quadric interactions, is just a standard form for 'Higgs-like' potentials, the second line is slightly strange. It gives a quadratic interaction between the $\sigma-$ and the δ_3-meson. The reason for its appearance lies in the structure of the VEV configuration, which explicitly connects both mesons. At the end of Sec. III.4 this topic was discussed and an alternative representation of the scalar potential was presented (Eq. (3.4.39)).
The coupling of the scalar and gauge fields consists of two components. While the current coupling is most conveniently presented together with the fermionic current coupling, there is also a quadric interaction between gauge and scalar fields, arising from the covariant kinetic term for scalar fields. All possible interactions where given in the most compact form in Eq. (3.5.27)

$$\mathcal{L}^\mathrm{quad}_{\Omega A} = \frac{g_\mathrm{A}^2}{2}|\Omega_0|^2\left|\frac{1}{2}(\ell^k_\mu - r^k_\mu)\right|^2 + \frac{g_\mathrm{A}^2}{2}|\Omega_j|^2\left|\frac{1}{2}(\ell^k_\mu + r^k_\mu)\right|^2 - \frac{g_\mathrm{A}^2}{2}\Re\left[\Omega^\star_a\Omega_b + \Omega^\star_0\Omega_j\epsilon_{jab}\right]\ell^a_\mu r^{b\mu}\,, \tag{3.6.8}$$

where $\boldsymbol{A}_\mu = (\boldsymbol{\ell}_\mu - \boldsymbol{r}_\mu)/2$ is the axial gauge field and $\boldsymbol{V}_\mu = (\boldsymbol{\ell}_\mu + \boldsymbol{r}_\mu)/2$ the vectorial one. Apart from interactive contributions this term, in combination with the scalar VEV led to the effective gauge field masses, which were summarised in Eq. (3.5.30) and turned out to be $M_3^\mathrm{ax} = g_\mathrm{A}\sqrt{R_\sigma^2 + R_\delta^2}$, $M^\mathrm{ax}_{1,2} = g_\mathrm{A}R_\sigma$, $M^\mathrm{vec}_{1,2} = g_\mathrm{A}R_\delta$ and $M^\mathrm{vec}_3 = 0$. The label ax and vec distinguishes axial and vectorial parts of the complete gauge field.
Coming to the current interactions, in the sections Sec. III.3.3 and III.4.1 the nucleon and scalar currents in the case of a global symmetry where presented. Later, in Sec. III.5.4 it was explicitly shown for the

scalar part, that the gauge fields couple to the found Noether currents from the global symmetry case (Eq. (3.5.39)). While not derived explicitly, the definitions of the covariant derivatives (Eq. (3.5.8) and (3.5.9)) can be used to find the analogous result for the coupling between gauge fermionic fields. Thus the gauge field coupling terms can be split up into an axial, a vectorial and a $U_\mathrm{V}(1)$ contribution.

$$\mathcal{L}_{\Omega A} + \mathcal{L}_{NA} = \mathcal{L}^{\mathrm{ax}} + \mathcal{L}^{\mathrm{vec}} + \mathcal{L}^{\mathrm{B}}\,, \tag{3.6.9}$$

$$\mathcal{L}^{\mathrm{ax}} = g_\mathrm{A} \boldsymbol{A}_\mu \left[\overline{N} \gamma^\mu \gamma_5 \frac{\boldsymbol{\tau}}{2} N + \Re\left\{ \{\partial^\mu \Omega_0\} \boldsymbol{\Omega}^\star - \Omega_0 \{\partial^\mu \boldsymbol{\Omega}^\star\} \right\} \right], \tag{3.6.10}$$

$$\mathcal{L}^{\mathrm{vec}} = g_\mathrm{A} \boldsymbol{V}_\mu \left[\overline{N} \gamma^\mu \frac{\boldsymbol{\tau}}{2} N + \Re\left\{ \boldsymbol{\Omega} \times \{\partial^\mu \boldsymbol{\Omega}^\star\} \right\} \right], \tag{3.6.11}$$

$$\mathcal{L}^{\mathrm{B}} = g_\mathrm{A} b_\mu \overline{N} \gamma^\mu N\,. \tag{3.6.12}$$

Here the nucleon currents have been adopted from Eq. (3.3.24) and (3.3.26) and the scalar currents come from Eq. (3.5.39). $\boldsymbol{A}_\mu$ stands for the axial gauge field, $\boldsymbol{V}_\mu$ for the vectorial one and b_μ is the $U_\mathrm{V}(1)$ gauge field. In Sec. III.4.1.1 it was discussed that the VEV of the scalar field leads to an explicit violation of conserved currents and it was found that only the third vectorial and the baryon number currents remained conserved in the presence of $\langle \Omega \rangle = (\sigma_0 I + \delta_0 \tau_3)$. Thus the conserved charges in the model are:

$$Q_\mathrm{V}^3 = (j_\mathrm{V}^3)^0 = (p^\dagger p - n^\dagger n) + 2\mathrm{i}(\pi_+ \overleftrightarrow{\partial_t} \pi_- + \delta_+ \overleftrightarrow{\partial_t} \delta_-)\,, \tag{3.6.13}$$

$$Q_\mathrm{B} = (j_\mathrm{B})^0 = (p^\dagger p + n^\dagger n)\,. \tag{3.6.14}$$

The first equation is obtained by combining Eq. (3.3.27) with (3.4.68) and the second equality was given in Eq. (3.3.28). The effects of breaking the current conservation in the remaining components of the vectorial and axial currents where discussed in Sec. III.4.1.1 and summarised in Eq. (3.4.64)-(3.4.66).
Now the remaining contribution to the effective Lagrangian (Eq. (3.6.3)) is the gauge fixing and ghost term. These contributions have been discussed in Sec. III.5.5. The results for an extended background field gauge were found to be:

$$\mathcal{L}_{\mathrm{gf}}(\xi) = \frac{1}{2\xi} (D_\mu^{\mathrm{con}} A_k^\mu)(D_\nu^{\mathrm{con}} A_k^\nu) + \frac{\xi g_\mathrm{A}^2}{2} \langle \omega | G_k | v \rangle^2 - g_\mathrm{A} \bar{D}_\mu A_k^\mu \langle \omega | G_k | v \rangle\,, \tag{3.6.15}$$

$$\mathcal{L}_{\mathrm{gh}}(\xi) = -(D_\mu^{\mathrm{con}} c^{\dagger\, a})((D^{\mathrm{con}})^\mu)^{ab} c^b - \xi c^{\dagger\, a} (M_A^2)^{ab} c^b - \xi g_\mathrm{A}^2 c^{\dagger\, a} \langle v | G_a G_b | \omega \rangle c^b\,. \tag{3.6.16}$$

Here the notation has not been translated back to the earlier used conventions. Some important connections are the VEV, $|v\rangle = \langle \Omega \rangle$, the physical scalar fields, $|\omega\rangle$, the generators of the gauge transformations (in a real representation), G_k, the mass matrix, $(M_\mathrm{A}^2)_{ab} = g_\mathrm{A}^2 \langle v | G_a G_b | v \rangle$, and the covariant derivative containing only the instanton field, D_μ^{con}. For further clarifications one should have another look into the derivations of Sec. III.5.5.
With this, the main aspects of the model, derived in this chapter, are summarised and of course further details concerning interpretations and derivations can be found in the corresponding sections.

IV Conclusive remarks

This work has picked up on an old topic of non-perturbative QCD. In the late nineteen seventies the isospin structure of the linear σ-model was generalised in the context of instantons for the first time by Saito and Shigemoto (Ref. [1]). Their generalised model included four degrees of freedom: the pionic Goldstone mode, $\boldsymbol{\pi}$, (pseudo-scalar, iso-vector), the σ-meson (scalar, iso-scalar) with mass, m_σ, a (pseudo-scalar, iso-scalar) contribution, $\widetilde{\eta}$, and the (scalar, iso-vector) part, which was named $\boldsymbol{\delta}$-meson in the present context. In their derivation the $\widetilde{\eta}$- and the $\boldsymbol{\delta}$-field acquired the same mass via an induced potential contribution from the instanton sector. At the time, the generalisation was done with 'pure' instantons.

In the present work the old derivation has been redone by the means of the later found constrained instanton fields. In addition the formalism was enlarged to accomodate fermionic contributions as well and a detailed analysis of the involved interactions and current contributions was presented.

IV.1 Summary

In chapter II the theoretical concepts have been presented that are needed for the generalisation of the linear σ-model in the context of instanton physics. In the general introduction of Sec. II.8 instantons were introduced as explicit gauge field configurations that minimised the Euclidean action, S_E. Later, in Sec. II.9.8, it was found that the coupling of instanton fields to scalars with non-vanishing VEV opened the possibility to lower S_E without bounds and therefore this kind of coupling led to a breakdown of the concept of instantons. This observation is the key argument against Saito's and Shigemoto's generalisation of the linear σ-model. Since the linear σ-model explicitly includes a VEV contribution in the scalar sector, its coupling to instanton configurations inevitably destabelises the instanton and forces it to vanish.

In order to save the general ideas, concerning Saito's and Shigemoto's extended σ-model, in Sec. (II.9.8) suitable constraints were added to the original model that prevented the destabelisation of the instanton solution. This gave rise to the concept of constrained instantons, A_μ^con. These field configurations are constructed such that the original instanton field is maintained in a small neighbourhood around the position, x_0, of A_μ^con and then falls off exponentially at large distances. This behaviour allows to give analytic expressions for constrained instantons in the 'small' and 'large' distance limit.

In the remaining sections of chapter II (from Sec. II.9.8 onwards) the contribution to the partition function, which arises from constrained instantons was calculated. For this derivation the scalar field was explicitly assumed to have the iso-spinor representation $\Omega = \Omega^\alpha q_\mathrm{I}^\alpha = (\Phi^\alpha + \mathrm{i}\Lambda^\alpha) q_\mathrm{I}^\alpha$, with the quaternion symbol, q_I^α, as it was defined in Sec. II.2. The scalar field components, Ω^α, Φ^α and Λ^α were given in Eq. (3.2.15). In addition, the VEV configuration of the scalar field was needed to be of the form $\langle\Omega\rangle = \Omega^0 I - \mathrm{i}\Omega^3\tau^3 = \mathrm{diag}[(\Omega^0 - \mathrm{i}\Omega^3), (\Omega^0 + \mathrm{i}\Omega^3)]$. Under these assumptions the complete partition function of a generic model, including constrained instanton effects in the context of the scalar field, Ω, was derived in Eq. (2.9.161).

In chapter III the findings from chapter II were combined to rederive the generalised σ-model in the context of constrained instantons, including an exemplary fermionic contribution, which was realised in terms of the nucleon iso-spinors, N. In order to arrive at computationally tractable expressions some simplifying assumptions concerning the constrained instanton contribution, $\mathcal{L}_\mathrm{inst}$, had to be introduced. Effectively these assumptions supressed all contributions that weren't compatible with the interpretation of $\mathcal{L}_\mathrm{inst}$ as a dynamical, local contribution to the scalar potential of the form: $\mathcal{L}_\mathrm{inst} = \mathcal{V}_\mathrm{inst}(\Phi, \Lambda)$.
In Sec. III.3 the nucleon Lagrangian was discussed. Important results of this discussion were the dynamically generated proton and neutron masses, $M_\mathrm{p} = g_\Omega(R_\sigma + R_\delta)$ and $M_\mathrm{n} = g_\Omega(R_\sigma - R_\delta)$, where R_k was the VEV contribution from the scalar field, k. Additionally the fermion-scalar interactions, given in Eq. (3.3.11) and the fermionic contributions to the axial and vector currents (Eq. (3.3.24)) were derived. While the fermionic current contributions exactly matched the findings from the ordinary σ-model, the dynamical nucleon masses experienced an alteration compared to the σ-model. This alteration arose naturally as a consequence of the new contribution to the scalar VEV from the (scalar, iso-vector) field, $\boldsymbol{\delta}$.
Subsequently, in Sec. III.4, the scalar sector of the Ω-field was investigated. As the induced instanton contribution, $\mathcal{V}_\mathrm{inst}(\Phi, \Lambda)$, directly gave a new term in the scalar potential, the effects of instantons on the scalar sector were very prominent. The scalar potential incorporated a 'Higgs-like' potential, as it is standard for the ordinary σ-model, the instanton part, $\mathcal{V}_\mathrm{inst}(\Phi, \Lambda)$, and a linear perturbation from nucleonic source terms, $\mathcal{V}_\mathrm{nucl}(\sigma, \delta_3)$. Expanding this potential around its minimum configuration generated the final effective potential of Eq. (3.4.37). This expansion dynamically generated four different masses, m_σ, m_π, m_δ, $m_{\delta\eta}$, and various cubic and quadric interaction terms. Later, in Sec. III.4.3, two different approaches were proposed to identify the effective parameters of the derived model with physical observables. The first was focussed on a pure scalar model, in which parameters were tied to the in-vacuum meson masses. In contrast to this 'scheme', the second approach was designed produces an effective, low energy nucleon model, where the parameters were coupled to the effective proton and neutron masses and three additional meson masses.
While the evaluation of the scalar currents led to similar structures as already known from the ordinary σ-model (Eq. (3.4.50) and (3.4.53)), the parts that produced violations of conserved currents introduced new terms in the generalised model. In the original σ-model current violating terms only appear in the

axial sector, but the situation is different in the generalised model from chapter III. The reason for this has been found to be connected to the newly introduced linear potential term, $\mathcal{V}_{\text{nucl}} \sim \beta\delta_3$, which coupled the third component of the $\boldsymbol{\delta}$-meson (scalar, iso-vector) to an external source. This term produced additional current violating contributions in the first two vectorial components of the current, so that, in the complete generalised model, only the third vectorial component of the current, $(j_{\mathrm{V}}^3)_\mu$, remained conserved. All equations indicating current conservation violations are listed in Eq. (3.4.64)-(3.4.66).
To complete the discussion of the generalised σ-model the gauge field sector was investigated. As instantons do not have a kinetic contribution, this analysis was limited to the derivation of the effective gauge field masses and the interactions between instantons and fermionic or scalar fields. In the presented model interactions, as well as the mass terms, had their origin in the covariant derivative, which couples the gauge fields (and with these the instantons) to the remaining fields of the model. The introduction of the correct covariant derivatives, which rendered the complete model locally gauge invariant, was presented in Sec. III.5.2. Using these relations, the covariant kinetic term of the scalar field produced the instanton-Ω coupling (compare Eq. (3.5.27)), and via the VEV of the scalar field, $\langle\Omega\rangle$, the dynamical gauge field masses were generated. Only the third direction of the vectorial gauge field in iso-spinor space remained a symmetry, if the model was expanded around the scalar VEV and therefore only this component remained massless. The remaining five gauge field components acquired the masses $M_3^{\mathrm{ax}} = g_{\mathrm{A}}\sqrt{R_\sigma^2 + R_\delta^2}$, $M_{1,2}^{\mathrm{ax}} = g_{\mathrm{A}} R_\sigma$ and $M_{1,2}^{\mathrm{vec}} = g_{\mathrm{A}} R_\delta$.
Finally the three-body interactions between gauge fields and fermions or scalars were discussed in Sec. III.5.4. In the model they were realised via current couplings which had the same structure as in the original σ-model. Most conveniently these interactions were given in Eq. (3.6.9)-(3.6.12). In the closing part of the gauge field discussion (Sec. III.5.5) also the gauge fixing Lagrangian and the corresponding ghost terms were derived for an extended background gauge formalism.
With this final contribution the conceptual derivation of the generalised σ-model was ended. Many important aspects concerning the generalisation procedure, the applicability and general implications of the resulting model have been presented and discussed. Of course, as has been indicated from time to time throughout this work, there are still open questions concerning the generalised σ-model and various details still have to be worked out explicitly. In a nutshell, the proposed generalisation of the linear σ-model in the context of constrained instantons incorporates interesting possibilities to study a generalised isospin structure in the low energy regime of QCD. Hopefully, in the future the presented formalism can be extended to a rigorous effective field theory description including constrained instantons, nucleons and scalar fields.

IV.2 Outlook

Throughout the derivations and presentations of chapter II and III some interesting parts had to be left out of the focus to make room for more pressing topics. All these parts present natural starting points for future investigations. While any left out derivation was directly mentioned in the corresponding section, this outlook will give a brief summary of the most important aspects that should be studied in future investigations.

The derivation of the approximate fermionic zero-mode determinant of Sec. II.9.9 led to the very involved analytic expression of Eq. (2.9.139). For the explicit calculations of chapter III only the high energy contribution of this expression was used (Eq. (2.9.144)). In addition this high energy part was assumed to give contributions at *all* energies and not only at the correct high energy scale. Therefore, an important subject in further analyses is to improve the oversimplified expression for the fermionic zero-mode determinant, $\det_0\left(i\not{D} + g_\Omega \Omega\right)$. If a more accurate expression for this determinant can be derived, this will directly affect the effective instanton contribution, $\mathcal{L}_{\text{inst}}$, in the effective model (compare Eq. (3.2.6)).

Another interesting topic for the future is the numerical analysis of the approximate fermionic zero-modes. In Sec. II.9.8.1 analytic approximations for these modes have been given for the limiting cases of extremely high, or low energies. The analysis on these modes can be approached, by using the expressions for the constrained instanton and Higgs field (Eq. (2.9.55) and (2.9.56)) to solve the set of partial differential equations (Eq. (2.9.60) and (2.9.61)) for the fermionic modes numerically. A detailed knowledge of the fermionic pseudo zero-modes would also allow to produce more accurate expressions for the fermionic zero-mode determinant.

A final aspect for future investigations in the instanton sector is concerned with the correct pre-exponential factor of the effective instanton contribution. In Eq. (3.2.6) this factor is summarised in the measure, $\mathcal{M}$. If one obtains a quantitative prediction on this factor, then it will be possible to determine the absolute magnitude of instanton effects more acurately, as these are closely related to the parameter a in the instanton induced potential, $\mathcal{V}_{\text{inst}} \sim a^2(||\Phi||^2 - ||\Lambda||^2)$. To make progess concerning this question, one needs to derive the constrained instanton parts in the context of a more thorough effective field theory approach. The conceptual introduction on this approach have been introduced in Sec. II.7.

Concerning developments of the effective model from chapter III future works should derive a rigorous inclusion of dynamical vector mesons (not only instanton-like contributions). In addition the exact meaning of the (pseudo-scalar, iso-scalar) contribution, $\widetilde{\eta}$, in a physical context needs to be investigated. Finally the derived model should of course be tested and tuned in physical situations. As the generic range of this model is rather large, starting from an effective quark model up to the description of nucleons, there are many possible scenarios, to which the model could be applied.

A Appendix

A.1 Abbreviations

Sec.	-	section
Tab.	-	tabular
Eq.	-	equation
Fig.	-	figure
VEV	-	vacuum expectation value
YMH	-	Yang-Mills Higgs (model)
QCD	-	Quantum Chromo Dynamics
System A is invariant under group G	-	A is invariant under the action of the group G.
Symmetry space	-	The space that a symmetry group acts on.

A.2 Notation

$\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$	-	metric of Minkowski space ($\eta_{\mu\nu} = (\eta_{\mu\nu})^{-1} = \eta^{\mu\nu}$)
$\{a\}$	-	set of all a (where a can contain any number of elements)
$\{a_j\}_{j=1}^{N}$	-	set of all a_j (where j runs from 1 to N)
$g(\{\alpha\}) \equiv g(\alpha_1, \alpha_2, \alpha_3, ...)$	-	a function of a set, is to be understood as a function of all elements of the set. Sometimes it is simply referred to as $g(\alpha)$ instead of $g(\{\alpha\})$.
$a \otimes b = (a, b)$	-	$\otimes$ means that a and b form an ordered pair (a, b)
$\mathbb{M}^{(3,1)}$	-	Minkowski space (Minkowski metric: $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$)
$x^\mu \in \mathbb{M}^{(3,1)}$	-	vector in Minkowski space *or* element of an vector in Minkowski space (in index contractions)
$\mathbf{x} \in \mathbb{R}^3$	-	vector in 3-dimensional Euclidean space
$f(x) \in O(g(x))$	$\Leftrightarrow$	$0 \leq \limsup_x \left\lvert\frac{f(x)}{g(x)}\right\rvert < \infty$ (definition)
$O(\alpha_j\alpha_k)$	-	term of order $\alpha_j\alpha_k$ for arbitrary j and k
$O(\{a_1, a_2, ...\}^j)$	-	term of order j in any combination of the elements of the set $\{\cdot\}$
'diagonalised'	-	to diagonalise a matrix via a similarity transformation ($A \rightarrow S^{-1}\Lambda S = \mathrm{diag}(\lambda_1, ...\lambda_n)$, where $\det(S) \neq 0$ and λ_j are eigenvalues of A)
Lie group	-	exponential representation of Lie group ($D(\alpha) = e^{\mathrm{i}\alpha_j X_j}$)
$\mathbb{1}$ and I	-	if various spaces are treated in the same context then typically for transformations the identity elements are distinguished. Usually $\mathbb{1}$ is the identity in Minkowski space and I is the unit element in the remaining space.
$P_\mathrm{L} = (I - \gamma_5)/2$	-	projection on the left-handed components of a Dirac-spinor. Note that $P_\mathrm{L}^2 = P_\mathrm{L}$.
$P_\mathrm{R} = (I + \gamma_5)/2$	-	projection on the right-handed components of a Dirac-spinor. Note that $P_\mathrm{R}^2 = P_\mathrm{R}$.
$U_\mathrm{ch}(\Lambda)$	-	unitary operator, which realises chiral gauge transformations.
$U_\mathscr{L}(\Lambda)$	-	unitary operator, which realises Lorentz transformations.
$U_{\mathscr{L},\mathrm{ch}}(\Lambda)$	-	unitary operator, which realises Lorentz and chiral transformations.
$L = e^{-\mathrm{i}\alpha_L} P_\mathrm{L}$	-	linear operator, which realises the chiral transformation on the left-handed parts.
$R = e^{-\mathrm{i}\alpha_R} P_\mathrm{R}$	-	linear operator, which realises the chiral transformation on the right-handed parts.
$\hat{a}$	-	unit vector in a-direction

A.3 Reminder on transformations

As confusion tends to spread as soon as functions with arguments in the transformation space are investigated here is a little reminder on the subject. For $x \in \mathbb{M}^{(3,1)}$ let $f(x)$ be a scalar function of a space-time argument living in some space S. Now as usual let G be a group and Λ be the action of this group on $\mathbb{M}^{(3,1)}$. Take T_Λ to be an operator acting on $f(x)$ that realises the group transformation in the S space. The question is how $f(x)$ transforms under group transformations. One would like to have that the transformed function of the transformed argument gives the same value as the original function of the original argument, or short:

$$U^{-1}(\Lambda)f(x)U(\Lambda) = T_\Lambda f(\Lambda x) \stackrel{!}{=} f(x) \tag{A.1}$$

Here the first term is just the generic expression for the complete action of a group transformation on a given argument. Using the similarity transformation $x \to x' = \Lambda^{-1}x$ the standard result is at hand:

$$T_\Lambda f(x) = f(\Lambda^{-1}x) \tag{A.2}$$

So the transformation of a function is realised by the inverse transformation of its argument.

In a similar manner functions in spaces that transform non trivially under group transformations can be analysed as well. For this suppose that $f(x) = f^\alpha(x)$ now lives not in the space of scalar functions but some other space W whose elements transform under the group action according to some representation $D^{\alpha\beta}(\Lambda)$. In this case the complete action of the group becomes:

$$U^{-1}(\Lambda)f^\alpha(x)U(\Lambda) = D^{\alpha\beta}(\Lambda)f^\beta(\Lambda^{-1}x) \tag{A.3}$$

Note that Eq. (A.3) holds as well if $f^\alpha(x)$ transforms as an element of $\mathbb{M}^{(3,1)}$ [e.g.: $f^\mu(x) = \partial^\mu\phi(x)$].

A.4 $SU(2)$ Tensor products

Sec. B.6.3.1 gave an explicit example of the state construction via a highest weight decomposition. As the results of several of these decompositions are needed for the construction of the presented model here is a set of all used decompositions. All equations correspond to representations of the $SU(2)$ group and can be constructed in analogy to Sec. B.6.3.1.

$$2 \otimes 2 = 3_s \oplus 1_a \tag{A.1}$$
$$3 \otimes 2 = 4 \oplus 2 \tag{A.2}$$
$$3 \otimes 3 = 5_s \oplus 3_a \oplus 1_s \tag{A.3}$$
$$4 \otimes 2 = 5 \oplus 3 \tag{A.4}$$

The index in the direct sum only applies to tensor products of identical representations, as only in this case it makes sense to speak about (anti-) symmetry under exchange of the constituent representations.

Using equation A.1 to A.4 the combined results can be derived:

$$2^{\otimes 3} \equiv 2 \otimes 2 \otimes 2 = (3_{\underline{s}} \oplus 1_{\underline{a}}) \otimes 2 = 4_{\underline{s}} \oplus 2_a \oplus 2_a \tag{A.5}$$

$$2^{\otimes 4} \equiv 2 \otimes 2 \otimes 2 \otimes 2 = (3_{\underline{s}} \oplus 1_{\underline{a}}) \otimes (3_{\underline{s}} \oplus 1_{\underline{a}}) \tag{A.6}$$

$$= 5_{\underline{s}} \oplus 3_a \oplus 3_{sa} \oplus 3_{as} \oplus 1_s \oplus 1_{aa} \tag{A.7}$$

Here a new label has been introduced: the 'sub-bar' in the indices. This additional information specifies if the representation is *completely* (anti-) symmetric. This new distinction becomes necessary since there are several 'layers' of exchange symmetry one is dealing with in the case of $2^{\otimes 3}$ and $2^{\otimes 4}$. As mentioned in Sec. B.6.3.1 only the highest weight irreducible representation is guaranteed to be completely symmetric. Typically in tensor products of multiple constituents the other irreducible representations are then partially (anti-) symmetric. This is denoted with the ordinary subscript (without a 'sub-bar'). If there are several indices this means that the states are symmetric under certain permutations and antisymmetric under certain other ones.
Examining equation A.7 then reveals the delema that there is no completely symmetric singulett in the tensor decomposition and thus it seems that there could not be a term combining 4 spin 1/2 objects in a Lagrange density. Fortunately this statement is wrong. By working out all basis states of the symmetric $5_{\underline{s}}$ and the three partially antisymmetric 3_a (these are 14 basis states) one findes that it is possible to choose the two singuletts such that one of them is completely symmetric and the other one is partially symmetric. So in this case (after a lot of calculations) one has:

$$2^{\otimes 4} = 5_{\underline{s}} \oplus 3_a \oplus 3_{sa} \oplus 3_{as} \oplus 1_s \oplus 1_{\underline{s}} \tag{A.8}$$

Working out the symmetry properties of states in higher tensor products already shows the limitations of the highest weight scheme. There are more sophisticated methods to adress the problems of decompositions but their introduction leads even further away than the already time consuming introduction of the highest weight scheme. The interested readet can learn more about the methods of representation theory in Ref. [2] or Ref. [3].

A.5 Inner products in Minkowski space

The reason for the distinction of upper and lower indices lies within the mathematical structure of forms in Minkowsky space. Suppose a system lives within a differentiable manifold. Then to every point in this manifold a tangent space can be associated by taking all possible 'directional derivatives' at this point. These derivatives form a vector space and $\mathbf{e}_\mu$ (with lower indices) denotes an unit vector in it.
The cotangent space is the dual space to this tangent space. So if the tangent space is a vector space V over a field F, then the dual space V^* is the set of linear maps from V to F ($\phi : V \to F$). It can now be

guessed that $\mathbf{e}^{\mu}$ is a unit vector in the dual space - or cotangent space.
Now combining an element of the tangent space with one of the cotangent space gives exactly what one is interested in - a scalar product on manifolds: $\langle\cdot,\cdot\rangle : V^{*} \times V \rightarrow F$. With the Minkowski metric the scalar product is:

$$\langle a, b\rangle = \sum_{\mu\nu} \eta_{\mu\nu} a^{\mu} b^{\nu} \tag{A.1}$$

This inner product is bilinear, symmetric and nondegenerate.

A.6 Non abilean field tensors

There are various approaches to construct the field or curvature tensor for non-abilean gauge fields. One idea for the construction of this tensor is to take a vector field $A_{\mu}(x)$ at a given point x_0 and to see how its value changes, if one takes it around an infinitesimal placket back to the starting point. This can be done with the parallel transport operator:

$$\mathscr{P}_c(A) = e^{-\int_c \mathrm{d}^d x A(x)} \tag{A.1}$$

$\mathscr{P}$ transports the field $A(x)$ along the path c in d dimensions. If you choose an infinitesimal square of size ε^2 in the directions μ and ν as the path c then the integral can be approximated to third order with:

$$\int_{x_0}^{x_0+\varepsilon} \mathrm{d}x A(x) \cdot \hat{\mu} = \varepsilon A_{\mu}(x + \frac{\varepsilon}{2}\hat{\mu}) + O(\varepsilon^3) \tag{A.2}$$

Here a hat on the direction means an unit vector in this direction. Using this, the parallel transport along the placket can be expressed via the expansion of the exponential

$$\mathscr{P}_\square(A) = \mathscr{P}_{x;\,x+\varepsilon\hat\mu}\mathscr{P}_{x+\varepsilon\hat\mu;\,x+\varepsilon\hat\mu+\varepsilon\hat\nu}\mathscr{P}_{x+\varepsilon\hat\mu+\varepsilon\hat\nu;\,x+\varepsilon\hat\nu}\mathscr{P}_{x+\varepsilon\hat\nu;\,x} \tag{A.3}$$

$$= \left[1 + \varepsilon A_\mu(x + \frac{\varepsilon}{2}\hat\mu) + \frac{\varepsilon^2}{2}A_\mu^2(x + \frac{\varepsilon}{2}\hat\mu) + O(\varepsilon^3)\right] \tag{A.4}$$

$$\cdot\left[1 + \varepsilon A_\nu(x + \varepsilon\hat\mu + \frac{\varepsilon}{2}\hat\nu) + \frac{\varepsilon^2}{2}A_\nu^2(x + \varepsilon\hat\mu + \frac{\varepsilon}{2}\hat\nu) + O(\varepsilon^3)\right] \tag{A.5}$$

$$\cdot\left[1 - \varepsilon A_\mu(x + \varepsilon\hat\nu + \frac{\varepsilon}{2}\hat\mu) + \frac{\varepsilon^2}{2}A_\mu^2(x + \varepsilon\hat\nu + \frac{\varepsilon}{2}\hat\mu) + O(\varepsilon^3)\right] \tag{A.6}$$

$$\cdot\left[1 - \varepsilon A_\nu(x + \frac{\varepsilon}{2}\hat\nu) + \frac{\varepsilon^2}{2}A_\nu^2(x + \frac{\varepsilon}{2}\hat\nu) + O(\varepsilon^3)\right] \tag{A.7}$$

$$= 1 + \varepsilon^2\left[\left.\frac{A_\nu(\tilde x + \varepsilon\hat\mu) - A_\nu(\tilde x)}{\varepsilon}\right|_{\tilde x = x+\hat\nu\varepsilon/2} - \left.\frac{A_\mu(\tilde x + \varepsilon\hat\nu) - A_\mu(\tilde x)}{\varepsilon}\right|_{\tilde x = x+\hat\mu\varepsilon/2}\right] \tag{A.8}$$

$$+ \varepsilon^2\left[A_\mu A_\nu - A_\mu A_\nu + A_\mu A_\nu - A_\nu A_\mu\right] \tag{A.9}$$

$$+ \frac{\varepsilon^2}{2}\left[\left(A_\mu - A_\mu\right)^2 + (A_\nu - A_\nu)^2\right] + O(\varepsilon^3) \tag{A.10}$$

$$\lim_{\varepsilon\to 0}\mathscr{P}_\square(A) = 1 + \varepsilon^2\left(\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]\right) \tag{A.11}$$

For the third equality only the arguments for the derivative part have been given explicitly in order to keep the notation as clear as possible. The arguments of the various A_μ are different, as can be seen from the first equality, but they become equal in the limit $\varepsilon \to 0$. So equation A.11 describes how a vector A_μ changes along an infinitesimal square and thus it encodes the curvature of the underlying manifold and $R_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$ is the corresponding curvature tensor. In physics A_μ typically is rescaled by a coupling constant, so that $A_\mu \to \widetilde{A}_\mu = -\mathrm{i}gA_\mu$ and the curvature tensor translates to the field strength tensor as:

$$F_{\mu\nu} = \frac{\mathrm{i}}{g}\left(\partial_\mu\widetilde{A}_\nu - \partial_\nu\widetilde{A}_\mu - [\widetilde{A}_\mu, \widetilde{A}_\nu]\right) = \partial_\mu A_\nu - \partial_\nu A_\mu + \mathrm{i}g[A_\mu, A_\nu] \tag{A.12}$$

An alternative and very elegant way to derive this object makes use of differential forms. So, to follow this derivation, one needs a a small introduction to the mathematical concepts of these forms. A short, and for this purpose sufficient one can be found in Ref. [15, p.217-230]. Without going into the details one big advantage of forms should be mentioned (so to say as an advertisement) and that is: They allow to study the structure of orientable manifolds in a basis-independent framework. Along these lines of derivation (again, for details compare Ref. [15, p.217-230]) the field strenght tensor turns out to be:

$$D_\mu = \partial_\mu - \mathrm{i}gA_\mu \tag{A.13}$$

$$F_{\mu\nu} = \frac{\mathrm{i}}{g}[D_\mu, D_\nu] \tag{A.14}$$

This definition of $F_{\mu\nu}$ is in agreement with the previous definition A.12.

A.7 Derivatives in curved spaces

In modern physics one often deals with local gauge field models. In these models gauge fields have to be included in order to preseve a desired symmetry. From the mathematical point of view there is a very natural explanation, why these fields (in mathematical contexts they are usually called connections) emerge, if one examins models with local symmetries. A local symmetry simply means that the underlying space(-time), in which the model lives, is not flat anymore, but has a non-trivial topology[1]. Compared to a flat (or globally symmetric) space the definition of derivatives has to be changed in this space in order to give a meaningful quantity. In general derivatives shall give the change of a function, if its argument undergoes a small change[2]. While in flat space the derivative is easily accessible as the quotient $\Delta f/\Delta x$ the situation is more complicated if the function lives in a topologically non-trivial space. In the non-trivial space one can take the coordinates $\xi = \xi(x)$ as functions of a flat space x. If one compares the value of the function at a given point ξ_0 with its value at another point ξ_1, then the difference $[f(\xi_0) - f(\xi_1)]$ depends on the change of the 'space function' $\xi(x)$ and on the actual properties of the function $f(\cdot)$. As derivatives shall give general properties of functions regardless of the space they live in, one needs to correct for the additional, space-related change in the difference. In mathematics this is done by redifining ordinary derivatives: $D_\mu^{ab} := \delta^{ab}\partial_\mu + \Gamma_\mu^{ab}$, where Γ_μ^{ab} is called the connection symbol. This object is exactly constructed such that the undesired changes in Δf (related to the coordinates $\xi(x)$) are canceled. So, by using the new derivative D_μ^{ab} instead of the normal one, all contributions vanish that emerge from the non-trivial topology of the underlying space. In particle physics the connection symbol is usually called 'gauge field', but it serves the very same purpose. So gauge fields are actually a direct consequence of a non-trivial underlying space.

A.8 Iso-spinor traces

This section only provides some calculations of iso-spinor traces that are used throughout the work. In all calculations of this section the following conventions are used:

- The sum convention is always used.
- $\star$ means complex conjugation, while $\dagger$ means hermitian conjugation as usual.
- τ^j is the j^{th} Pauli matrix in iso-spinor space.
- I is a unit matrix in iso-spinor space.

[1] Colloquial speaking the space is 'hilly'. Take for example the surface of the earth (including local hills) as input space.
[2] In flat space one has: $\partial_\alpha f(x) := \lim_{\epsilon\to 0}[f(x+\epsilon\hat{e}_\alpha) - f(x)]/[x+\epsilon\hat{e}_\alpha]$

- $q^\alpha = (I, -\mathrm{i}\boldsymbol{\tau}^\mathrm{T})^\mathrm{T}$ is the component α of the quaternion symbol (introduced in Sec. II.2).
- $\bar{q}^\beta = (I, \mathrm{i}\boldsymbol{\tau}^\mathrm{T})^\mathrm{T}$ is the component β of the conjugate quaternion symbol (introduced in Sec. II.2).
- Quaternion symbols (including the conjugates) always carry greek indices. The index of the quaternion symbol is adopted for its Pauli matrix components as well. For example:

$$\sum_{\alpha=0}^{3} q^\alpha = I - \mathrm{i}\sum_{\alpha=1}^{3} \tau^\alpha \tag{A.1}$$

  So if α runs from 1 to 3 or rather from 0 to 3 depends on the symbol it is connected to. Also counterintuitive this abbreviation will help to keep notation clear.
- Latin indices belong to 'pure' Pauli matrices.
- The change from super- to subscripts is only introduced here to clear up notation.

In the calculations the following relations come in quite handy:

$$\tau^a\tau^b = \frac{1}{2}\{\tau^a, \tau^b\} + \frac{1}{2}[\tau^a, \tau^b] = \delta_{ab}I + \mathrm{i}\epsilon_{abc}\tau^c \tag{A.2}$$

$$\mathrm{tr_I}[I] = 2 \tag{A.3}$$

$$\mathrm{tr_I}[\tau^a] = 0 \tag{A.4}$$

$$\mathrm{tr_I}[\tau^a\tau^b] = 2\delta_{ab} \tag{A.5}$$

$$\mathrm{tr_I}[\tau^a\tau^b\tau^c] = \mathrm{i}\epsilon^{abx}\mathrm{tr_I}[\tau^x\tau^d] = 2\mathrm{i}\epsilon^{abc} \tag{A.6}$$

$$\mathrm{tr_I}[\tau^a\tau^b\tau^c\tau^d] = \delta_{ab}\mathrm{tr_I}[\tau^c\tau^d] + \mathrm{i}\epsilon^{abx}\mathrm{tr_I}[\tau^x\tau^c\tau^d] \tag{A.7}$$

$$= 2(\delta_{ab}\delta_{cd} - \epsilon_{abx}\epsilon_{xcd}) = 2(\delta_{ab}\delta_{cd} - \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{cb}) \tag{A.8}$$

The later relations can be deduced by using Eq. (A.2) through Eq. (A.4). This should be enough and here are the promised traces:

$$\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[q^{\alpha}\bar{q}^{\beta}\right] = \frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[(\delta_{\alpha 0}I - \mathrm{i}\tau^{\alpha})(\delta_{\beta 0}I + \mathrm{i}\tau^{\beta})\right] = \delta_{\alpha\beta} \tag{A.9}$$

$$\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[\tau^{a}q^{\alpha}\bar{q}^{\beta}\right] = \delta_{\alpha 0}\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[\tau^{a}(\delta_{\beta 0}I + \mathrm{i}\tau^{\beta})\right] - \mathrm{i}\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[\tau^{a}\tau^{\alpha}(\delta_{\beta 0}I + \mathrm{i}\tau^{\beta})\right] \tag{A.10}$$

$$= \mathrm{i}\delta_{\alpha 0}\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[\tau^{a}\tau^{\beta}\right] - \mathrm{i}\delta_{\beta 0}\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[\tau^{a}\tau^{\alpha}\right] + \frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[\tau^{a}\tau^{\alpha}\tau^{\beta}\right] \tag{A.11}$$

$$= \mathrm{i}(\delta_{\alpha 0}\delta_{a\beta} - \delta_{\beta 0}\delta_{a\alpha}) + \mathrm{i}\epsilon_{a\alpha\beta} \tag{A.12}$$

$$\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[\tau^{a}\bar{q}^{\alpha}q^{\beta}\right] = -\mathrm{i}(\delta_{\alpha 0}\delta_{a\beta} - \delta_{\beta 0}\delta_{a\alpha}) + \mathrm{i}\epsilon_{a\alpha\beta} \tag{A.13}$$

$$\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[q^{\alpha}\bar{q}^{\beta}\tau^{a}\tau^{b}\right] = \frac{1}{2}\left(\delta_{\alpha 0}\mathrm{tr}_{\mathrm{I}}\left[\bar{q}^{\beta}\tau^{a}\tau^{b}\right] - \mathrm{i}\mathrm{tr}_{\mathrm{I}}\left[\tau^{\alpha}\bar{q}^{\beta}\tau^{a}\tau^{b}\right]\right) \tag{A.14}$$

$$= \frac{1}{2}\left(\delta_{\alpha 0}\delta_{\beta 0}\mathrm{tr}_{\mathrm{I}}\left[\tau^{a}\tau^{b}\right] + \mathrm{i}\delta_{\alpha 0}\mathrm{tr}_{\mathrm{I}}\left[\tau^{\beta}\tau^{a}\tau^{b}\right] - \mathrm{i}\delta_{\beta 0}\mathrm{tr}_{\mathrm{I}}\left[\tau^{\alpha}\tau^{a}\tau^{b}\right] + \mathrm{tr}_{\mathrm{I}}\left[\tau^{\alpha}\tau^{\beta}\tau^{a}\tau^{b}\right]\right) \tag{A.15}$$

$$= \delta_{\alpha 0}\delta_{\beta 0}\delta_{ab} + \Big[-\delta_{\alpha 0}\epsilon_{\beta ab} + \delta_{\beta 0}\epsilon_{\alpha ab} + \delta_{\alpha\beta}\delta_{ab} - \delta_{\alpha a}\delta_{\beta b} + \delta_{\alpha b}\delta_{\beta a}\Big]_{\alpha,\beta=1}^{3} \tag{A.16}$$

$$= \delta_{\alpha\beta}\delta_{ab}\Big|_{\alpha,\beta=0}^{3} + \Big[-\delta_{\alpha 0}\epsilon_{\beta ab} + \delta_{\beta 0}\epsilon_{\alpha ab} - \delta_{\alpha a}\delta_{\beta b} + \delta_{\alpha b}\delta_{\beta a}\Big]_{\alpha,\beta=1}^{3} \tag{A.17}$$

$$\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[\bar{q}^{\alpha}q^{\beta}\tau^{a}\tau^{b}\right] = \delta_{\alpha\beta}\delta_{ab}\Big|_{\alpha,\beta=0}^{3} + \Big[\delta_{\alpha 0}\epsilon_{\beta ab} - \delta_{\beta 0}\epsilon_{\alpha ab} - \delta_{\alpha a}\delta_{\beta b} + \delta_{\alpha b}\delta_{\beta a}\Big]_{\alpha,\beta=1}^{3} \tag{A.18}$$

$$\frac{1}{2}\mathrm{tr}_{\mathrm{I}}\left[\bar{q}^{\alpha}\tau^{a}q^{\beta}\tau^{b}\right] = \frac{1}{2}\left(\delta_{\alpha 0}\mathrm{tr}_{\mathrm{I}}[\tau^{a}q^{\beta}\tau^{b}\right] + \mathrm{i}\mathrm{tr}_{\mathrm{I}}\left[\tau^{\alpha}\tau^{a}q^{\beta}\tau^{b}\right]\right) \tag{A.19}$$

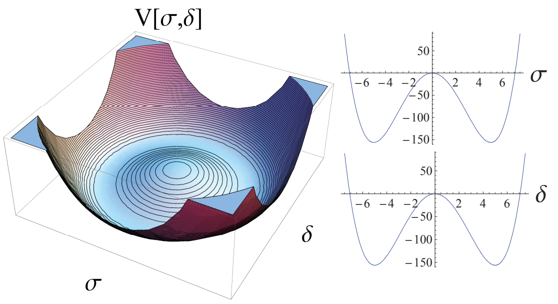
$$= \frac{1}{2}\left(\delta_{\alpha 0}\delta_{\beta 0}\mathrm{tr}_{\mathrm{I}}\left[\tau^{a}\tau^{b}\right] - \mathrm{i}\delta_{\alpha 0}\mathrm{tr}_{\mathrm{I}}\left[\tau^{a}\tau^{\beta}\tau^{b}\right] + \mathrm{i}\delta_{\beta 0}\mathrm{tr}_{\mathrm{I}}\left[\tau^{\alpha}\tau^{a}\tau^{b}\right] + \mathrm{tr}_{\mathrm{I}}\left[\tau^{\alpha}\tau^{a}\tau^{\beta}\tau^{b}\right]\right) \tag{A.20}$$

$$= \delta_{\alpha 0}\delta_{\beta 0}\delta_{ab} + \Big[\delta_{\alpha 0}\epsilon_{a\beta b} - \delta_{\beta 0}\epsilon_{\alpha ab} + \delta_{\alpha a}\delta_{\beta b} - \delta_{\alpha\beta}\delta_{ab} + \delta_{\alpha b}\delta_{\beta a}\Big]_{\alpha,\beta=1}^{3} \tag{A.21}$$

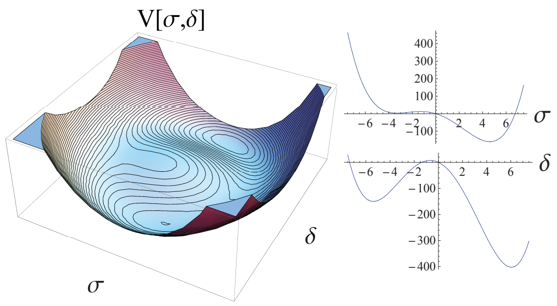
$$= \left(\delta_{\alpha 0}\delta_{\beta 0} - \delta_{\alpha\beta}\Big|_{\alpha,\beta=1}^{3}\right)\delta_{ab} + \Big[\delta_{\alpha 0}\epsilon_{a\beta b} + \delta_{\beta 0}\epsilon_{a\alpha b} + \delta_{\alpha a}\delta_{\beta b} + \delta_{\alpha b}\delta_{\beta a}\Big]_{\alpha,\beta=1}^{3} \tag{A.22}$$

A.9 Visualisations of the scalar potential

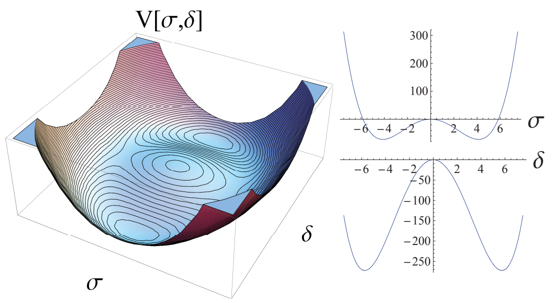
Here are some visualisations of the scalar potential for various paramters, that help to identify the role of each parameter in this potential model. All values are given in arbitrary units.



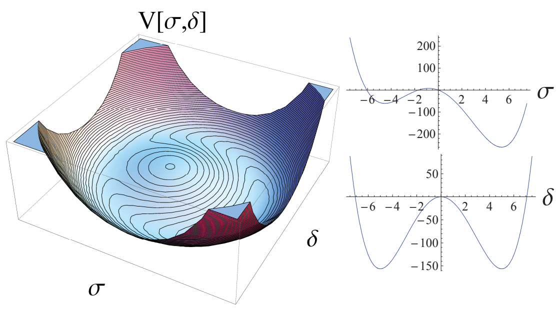
(a) Free 'Higgs potential' with a constant μ and λ. ($a = 0$, $\alpha = 0, \beta = 0$)



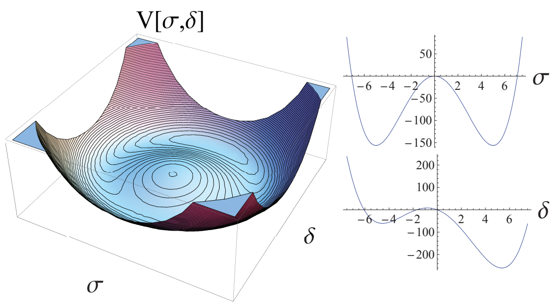
(b) Complete scalar potential as in section Sec. III.4. ($a \neq 0$, $\alpha \neq 0, \beta \neq 0$)



(c) Only instanton effects ($a \neq 0$, $\alpha = 0, \beta = 0$)



(d) Only the symmetry breaking in the σ-direction ($a = 0$, $\alpha \neq 0, \beta = 0$)



(e) Only the symmetry breaking in the δ-direction ($a = 0$, $\alpha = 0, \beta \neq 0$)

B Appendix (Group Theory)

Talking about group theory a physicist, unfamiliar with the subject, might ask: "Is this actually of any importance in physics?" And throughout the last century a very decisive "YES!" has been found to be the answer to this question.

What promotes group theory to such an useful tool in modern physics are (somewhat naturally) the things that can be described with it. A closer look at the mathematical structure of groups (see B.1) reveals that no viewer thing than symmetries can be characterised via groups. Now this should wake up every last physicist. Symmetries have been exploited from the very beginning of physics research in order to solve otherwise unsolvable problems. The idea behind the use of symmetries has always been that symmetries in general put tremendous constraints on systems. If one finds a method to include these constraints directly in the mathematical description of the system then this should simplify the problem. A formalism derived like this would allow its constituents only to evolve into configurations that agree with the symmetry of the system. In doing so the formalism itself would eliminate the huge number of possible system configurations that do not agree with the given symmetry.

Even though symmetries have often been used in physics, mostly during the last century a mathematical formalism has been developed and applied that allows to analyse general implications of symmetries on physical systems.

While the subject of group theory itself is very interesting, the focus of this work is not a full presentation of this field of research. Of course, the interested reader is invited to follow more complete introductions to group theory. Two possible sources are Ref. [2] and Ref. [3]. Both are used intensively in this chapter.

B.1 Definitions and basic ideas

Definition B.1. (Group)

A group $(G, .)$ is an object that consists of a set of elements G and an operation (group law) $O(x, y) \equiv x.y$ that specifies how group elements x and y are combined. The following relations have to be satisfied:

(i) $h = g_i.g_j \in G \quad \forall g_i,\ g_j \in G$ (G is closed under $O(\cdot,\cdot)$)

(ii) $g_i.(g_j.g_k) = (g_i.g_j).g_k \quad \forall g_i,\ g_j, g_k \in G$ (associativity)

(iii) G contains an identity element e such that $e.g = g.e = g$ for every element $g \in G$.

(iv) For every group element $g \in G$ there is an inverse g^{-1} within the group such that

$$g.g^{-1} = g^{-1}.g = e \qquad .$$

Note that the group concept can be applied to finite sets of elements (finite groups) and infinite ones as well (infinite groups). Both types of groups can be found in physics, or more exactly in the description of symmetries. Finite groups usually are used for objects with discrete symmetries (such as crystals) while infinite groups find their applications in the context of continuous symmetries (rotations, spatial translations, internal symmetries in particle physics). The amount of examples already indicates what kind of symmetries will be most important throughout this work. The reason why groups can be used to describe symmetries lies within the structure of symmetry operations. Without giving a proof the following example shall clarify how symmetry transformations obey all requirements from definition B.1:

For a square there are several symmetry operations (rotations) that leave its shape invariant. Mathematical speaking these transformations form a group which is called the dihedral group of order 8 ($\widehat{=}D_4$). (The order of the group G is the number of elements of G.) Fig. B.1 scetches all possible rotation axis for this system.

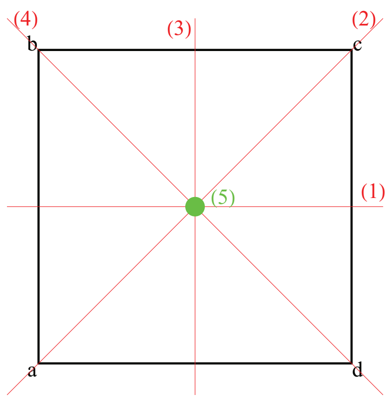


Figure B.1: Symmetry axis of a square ($D_4 \equiv$ dihedral group of order 8).

There are four rotations by 180° (red lines) and four rotation by 90°, 180°, 270° and 360° $= 0° \equiv e$ (green dot). Using this little picture one can easily work out that performing any two rotations (symmetry operations) is equivalent to just one other rotation.
For example rotating first around axis (1) and then around (3) is equivalent to a rotation of 180ř around the green dot. By calculating all possible combinations of two rotations (creating the multiplication table of the D_4-group) one can verify that this system (symmetry transformations on a square) fulfils all parts of definition B.1. This verification is tedious and not very enlightening so it is left out here. The key point is that it is inherent to any symmetric system that its symmetry transformations fulfil definition B.1. From here on 'groups' and 'symmetry groups' will be used synonymously.

Secondly representations shall be introduced. This concept allows to adjust the ideas of group theory to typical situations in theoretical physics. Usually, in any quantum mechanics related field, physicists are working with states and operators in some Hilbert space. So it would be neat to produce a coherent picture that allows to talk about symmetries of Hilbert space systems using the results of group theory. This is what representations allow to do:

Definition B.2. (Representation)

(a) A linear representation associates a set of linear operators D with the elements g of a group G, thereby giving a map $D(g)$ from the group elements to the linear operators. This map has to fulfil the two relations:

  (i) $D(e) = \mathbb{1}$[1)]

  (ii) $D(g_i)D(g_j) = D(g_i{\cdot}g_j)$

  In (i) the $\mathbb{1}$ is the identity in the space on which the linear operators act and (ii) basically means that the multiplication law for group elements translates into an ordinary multiplication in this "linear operator space".

(b) The dimension of a representation is the dimension of the space V it acts on.

Mathematically speaking, this means a representation D is a map from the group to the general linear group on a given vector space $GL(V)$ (even shorter: D is a group homomorphism $D : G \to GL(V)$).

Note that the second condition in B.2 implies that the inverse of a representation element is given by $D^{-1}(g) = D(g^{-1})$.
From the definition two very distinct types of representations can be constructed - the *faithful representations*, which associate a distinct linear operator to each element of the group and the unfaithful ones, where this is not the case. The simplest example for an unfaithful representation is called the 'trivial representation'. It simply sets every element of the group to the identity and thereby one clearly looses all informations on substructures of the group.

[1)]In the mathematical framework this postulate is obsolete but it is included here for for clarity.

As an example for a faithful representations one could think about rotations in a two dimensional space $\mathbb{R}^2$ (rotations around the third axis). Every rotation in two dimensions can be written as:

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{B.1}$$

It is easy to check that $R(\theta)$ satisfies B.2(i) with $R(0) = \mathbb{1}$ and B.2(ii) as $R(\alpha)R(\beta) = R(\alpha+\beta)$ is correct. The rotation group is infinite and this property appears in its representation $R(\theta)$ as well ($\theta \in \mathbb{R}$). Examples for finite groups are not any harder to do, but as they are not too relevant in this context and therefore most examples will be on infinite groups.
It is important to understand that the *dimension of a representation* is by no means absolute. The just given example could be represented in one dimension as well. In $\mathbb{C}^1$ rotations can be expressed simply via the unitary operator $R(\theta) = e^{i\theta}$. Of course, higher dimensional representations of a group are possible as well as long as they satisfy B.2(i) and B.2(ii).

Throughout this work representations are limited to linear operators, which yields enormous simplifications. One can bring them into any desired form via simliarity transformations without changing any of the multiplication rules of the underlying group structure. In other words if $D(g)$ is a representation of the group G then $D'(g) := S^{-1}D(g)S$, with the similarity transformation S, is a representation of the same group[2)]. If the elements of S form a group themselves, then then the similarity transformation is said to be the action of S on G in the representation $D(g)$. In this example this seems to be just an uptight name for an ordinary matrix multiplication, but as the expression is used very regularly and as the concept is important in a slightly different context here is a formal definition.

Definition B.3. (Group action) Take G as a group and X as an arbitrary set. The (left) group action is defined as the map $l(g,x) \equiv g.x$ with the following properties:

(i) $(gh).x = g.(h.x) \qquad \forall\, g,h \in G$ and $x \in X$

(ii) $e.x = x$

Analogously the (right) group action $r(x,g) \equiv x.g$ can be defined by a composition from the righthand side with adapted conditions on the multiplication. It turns out that a left group action can be written in terms of a right group action with the inverse group element: $l(g,x) = r(x,g^{-1})$.

As there are no restrictions on the set X in definition B.3, one can study the action of groups on any kind of set. In particular this allows to study the action of a group G (or its representation $D(g)$) on itself. This action is given by:

$$D'(g) := D(h^{-1})D(g)D(h) \qquad \forall g,h \in G \tag{B.2}$$

[2)]Mathematically S is an automorphism on the group structure.

and will play an important role for the characterisation of representations in the next section.

B.2 Irreducible representations and direct sums

The freedom to change a representation of a group via a similarity transformation leads to the question if there is a favourable basis in which a representation has an easy structure. The answer to that question depends on the particular group one is interested in but still there are some general remarks at this point. It is possible that a group G has an invariant subspace G_s. This means that the action of the group on itself leaves elements of G_s in this subspace. For $s \in G_s$ an $g \in G$ this means for any explicit representation of the group $D(g^{-1})D(s)D(g) = D'(s) \in G_s$ (compare the previous section).
As an example think about 3-dimensional rotations in a 4-dimensional space $\mathbb{R}^4$. Any combination of two rotations in the 3-dimensional space will still be in this subspace of $\mathbb{R}^4$. Therefore a representation of this symmetry could be written in a form:

$$D(g_j) = \begin{pmatrix} 1 & 0 \\ 0 & {}^{3}\!R(g_j) \end{pmatrix} \tag{B.1}$$

where ${}^{3}\!R(g_j)$ would be some representation of 3-dimensional rotations. In general, if a representation has an **invariant subspace** it is said to be **reducible**. If this is not the case it is called **irreducible**. Finally it is called **completely reducible** if it can be written in a complete block diagonal form where every block is an irreducible representation:

$$D(g) = \begin{pmatrix} D_1(g) & 0 & \cdots \\ 0 & D_2(g) & \cdots \\ \vdots & 0 & \ddots \end{pmatrix} \tag{B.2}$$

Such a block diagonal representation is referred to as a *direct sum* of irreducible representations:

$$D(g) = D_1(g) \oplus D_2(g) \oplus ... \tag{B.3}$$

If a representation can be decomposed into a direct sum in the above fashion then one has found the simplest possible building blocks in this representation (up to isomorphisms). This concept will find many applications in later chapters Sec. II.1.2. The dimension of a direct sum representation is easy to see from equation B.2:

$$\dim(D(g)) = \dim(D_1(g)) + \dim(D_2(g)) + ... \tag{B.4}$$

B.3 Tensor products

Apart from slicing representation into irreducible building blocks, one could try to furnish something like a product of representations. Of course, the question then is what this product should describe. Before adressing it the mathematical concept of a tensor product shall be introduced:
Suppose there are two systems A and B that live in some configuration spaces Γ_A with a complete basis $\{a\}$ and Γ_B[3)] with a complete basis $\{b\}$. To build a combined system, one needs a space where the combination of A and B lives in. One space, built of Γ_A and Γ_B, is gained by simply arranging the basis elements of the two subspaces as *ordered pairs* (a, b). The pairs of all basis elements form a basis in the tensor space[4)] $A \otimes B$:

$$(a, b) \equiv a \otimes b \tag{B.1}$$

where $\otimes$ is just a symbol saying: build a pair, where the order matters. The tensor product satisfies rules that look like distributive and associative laws but obviously it does not satisfy something like a commutative law ($a \otimes b \neq b \otimes a$). The *dimension* of the tensor product is

$$\dim(A \otimes B) = \dim(A) \cdot \dim(B) \tag{B.2}$$

It should be mentioned that the above situation is very common in physics. $A \otimes B$ simply describes the configuration space of two non-interacting subsystems (particles). Each of the subsystems lives in a space with a particular basis ($\{a\}$ or $\{b\}$). A vector in tensor space in physics is typically denoted by:

$$|ab\rangle = |a\rangle|b\rangle = |a\rangle \otimes |b\rangle := a \otimes b = (a, b) \tag{B.3}$$

Now, (as the playground is defined) the initial question about the physical purpose of tensor products can be adressed. For this take system A to be invariant under the action of a symmetry group G_A (for brevity call this: A is invariant under G_A) and B invariant under the group G_B. Working in a particular basis $\{a\}$ and $\{b\}$ gives a fixed representation of the groups: $D(g_A)$ and $E(g_B)$. Here $D(g_A)$ only acts on elements of Γ_A and $E(g_B)$ likewise only acts on elements of Γ_B. Then a *tensor representation* can be defined to be:

$$T_{D\otimes E}(g_A, g_B) := D(g_A) \otimes E(g_B) \tag{B.4}$$

It describes the combined action of both groups G_A and G_B on the combined states (tensor states). With this, a symmetry transformation in the tensor space from a given initial state $|my\rangle$ to some final state $\langle lx|$

[3)] Γ_A and Γ_B are vector spaces.
[4)] Normally this space is called tensor product

would look like:

$$\langle lx|\Big[T_{D\otimes E}(g_A, g_B)\Big]|my\rangle = \langle l|\langle x|\Big[D(g_A)\otimes E(g_B)\Big]|m\rangle|y\rangle \tag{B.5}$$

$$= \langle l|\langle x|\Big[D(g_A)|m\rangle \otimes E(g_B)|y\rangle\Big] \tag{B.6}$$

$$= \langle l|D(g_A)|m\rangle \otimes \langle x|E(g_B)|y\rangle \tag{B.7}$$

$$= \langle l|D(g_A)|m\rangle \, \langle x|E(g_B)|y\rangle \tag{B.8}$$

$$= \Big[T_{D\otimes E}(g_A, g_B)\Big]_{lm\,xy} \tag{B.9}$$

It still needs to be shown that $T_{D\otimes E}$ forms a representation but this is very easy as $T_{D\otimes E}$ is constructed such that the representations of the subspaces ($D\&E$) are only combined with the fitting vectors of their 'own' subspaces ($\Gamma_D\&\Gamma_E$). So as D and E form representations $F_{D\otimes E}$ forms one as well.
Now having such a tensor representation, it is not at all clear if it is irreducible (and usually it is not). So finding the direct sum decomposition of tensor products will be a very important task in most applications that deal with combined systems (see Sec. II.1.2 and Sec. II.4).

B.4 Symmetries, states and operators

So far the concept of groups, symmetries and their representations has been introduced without the direct need of physical observations. Nevertheless at this point it is already possible to state a subtle but important implication for physical systems:

Theorem B.1. *Let H be a hermitian operator and $D(g)$ a representation of a symmetry group G. If the commutator $[H, D(g)]$ vanishes for all elements of $D(g)$, then H can be written in terms of irreducible representations of G and so the eigenstates of H transform according to irreducible representations of G.*

Without giving a proof the theorem is kind of natural in physical situations. If the hermitian operator H describes an observable, then a symmetry of the system can be understood as a transformation that does not change the eigenvalues and -vectors of the observables. This is exactly the case if $[H, D(g)] = 0$ for all elements of the representation $D(g)$[5)]. Now $D(g)$ can be expressed in terms of irreducible representations and therefore this has to be true for H as well because otherwise $[H, D(g)] = 0$ would not hold for the whole representation.
If the operator H can be written in terms of irreducible representations of G this means $H = \sum_g H_g D^i(g)$ for an irreducible representation $D^i(g)$ and scalar functions H_g. This on the other hand means that the eigenstates of H transform according to the transformations $D^i(g)$.
For an *abelian* group the elements $D^i(g)$ and $D^j(g)$ commute and so H can be diagonalised completely.

5) $D^{-1}(g)HD(g) = D^{-1}(g)D(g)H = H$

However the situation changes if the group is *non-abelian*. In this case there are at least some elements of the representation that cannot be diagonalised simultaneously. Therefore the hermitian operator H cannot be diagonalised completely and so not all its eigenvalues can be measured simultaneously. This is an important consequence that explains for example why the quantum mechanical angular momentum can only be measured in one direction. In Sec. B.6 an example of this situation will be picked up.

Notice that the theorem has an almost philosophical implication if one starts to think about it inversively: If a system in nature possesses a certain symmetry then this alone restricts what kind of observable quantities that can live within the system!
By remembering the introductory comments on the history of symmetries in physics on the other hand, this statement is not very astonishing. As symmetries in general put huge constraints on systems they better affect possible observables. At the core this is what is exploited, whenever symmetries are used to simplify problems.
With the so far introduced concepts on group theory one is already in a position to derive the very important Noether theorem, which connects continuous global symmetries in physical systems to conserved currents and charges. A introduction to this theorem is given in Sec. II.3.

B.5 Lie Groups

Coming back to the introductory part on group theory the next topic of importance are Lie groups. These groups are infinite (continuous) groups with certain properties. They are named after Sophus Lie, who found the Lie algebras associated with continuous groups.
Non mathematical speaking: Lie groups are groups that depend smoothly on a set of continuous parameters α Ref. [2, p.43]. With smooth dependence is meant that the group operation and the inversion shall be smooth maps. Now smooth itself means that whenever two elements of the group are close toghether in 'group space' then the parameters that describe them $\{\alpha_j\}$ are close together as well. (This closeness leads to differentiability for Lie groups.)
Another way to define linear Lie groups is to see them as subgroups of $Gl(V)$ that are (closed) $\mathbb{C}^\infty$ manifolds as well Ref. [3, p.172-173].

B.5.1 Generators

Apart from these, at first glance, unfamiliar definitions Lie groups can be parameterised via 'generators' which gives them a very practical appearance. The key idea behind generators is that the closeness of group elements can be exploited in order to express them all in terms of their 'distance' to the identity e, at least in a small neighborhood. The identity is promoted to be the reference element as it appears in

every group.
So using a set of N real parameters $\{\alpha_j\}_{j=1}^N$ that characterises the group elements $g(\{\alpha\}) \equiv g(\alpha)$, the closeness of group elements allows to use a parameterisation such that

$$g(\alpha)|_{\{\alpha\}=0} = e \tag{B.1}$$

For a representation $D[g(\{\alpha\})] \equiv D(\alpha)$ of the group this means that $D(\alpha)|_{\{\alpha\}=0} = \mathbb{1}$. Using the differentiability of Lie groups (and therefore of their representations) $D(\alpha)$ can be expressed via a Taylor series in a close neighborhood of the identity:

$$D(\delta\alpha) = \mathbb{1} + \mathrm{i}(\delta\alpha_j)X_j + \mathcal{O}(\delta\alpha_j\delta\alpha_k) \tag{B.2}$$

Here $\delta\alpha$ has to be within the radius of convergence of the Taylor series and X_j is just the missing part of the expansion, meaning:

$$X_j = -\mathrm{i}\,\frac{\partial}{\partial\alpha_j}D(\alpha)\bigg|_{\{\alpha\}=0} \tag{B.3}$$

These X_j are called *generators*. As representations $D(\alpha)$ need to be linear operators, the X_j have to be linear operators as well. For practical purposes they can simply be thought of as matricies. The inclusion of i in Eq. (B.2) is not necessary but it makes the generators hermitian if the whole representation is unitary.
Equation Eq. (B.2) defines how the representation of a Lie group looks very close to the identity. So this can be used to describe infinitesimal group transformations in a particular direction (namely the α_jX_j-dircetion). If the group operation in $(G, .)$ is given by a simple addition of group elements (this depends on the parameterisation $\{\alpha\}$) then this behaviour translates to the nice multiplication law for representations: $D(\delta\alpha)\cdot D(\delta\alpha) = D(\delta\alpha+\delta\alpha)$ (compare B.2(ii)). Using this, one can get from the infinitesimal transformation to large scale transformations by combining the multiplication law with Eq. (B.2):

$$D(\alpha) = \lim_{k\to\infty}\left(1 + \mathrm{i}\frac{\alpha_j}{k}X_j\right)^k = e^{\mathrm{i}\alpha_jX_j} \tag{B.4}$$

In the limit $(1 + \mathrm{i}\alpha_jX_j/k)$ becomes an element of the representation as α_j/k becomes small for constant α_j. Raising any element of $D(\alpha)$ to some power still is in the representation[6)] and so the exponential is part of $D(\alpha)$ as well. Usually this parameterisation of Lie groups is called the *exponential representation*. Close to the identity it has the nice feature that the group is completely specified by the behaviour of its generators as Eq. (B.2) approximates all elements. So instead of studying the group elements themselves one can study its generators which is extremely helpful as they form a vector space (unlike the group

[6)] Any power of $D(g)$ is in the representation, since a group is closed under its multiplication law (see B.1)

elements).
Most of physical applications (and this work as well) only make use of the exponential representation of Lie groups. So from now on talking about Lie groups always is equivalent to talking about group representations of the form $D(\alpha) = e^{\mathrm{i}\alpha_j X_j}$ with a *fixed* and *finite* set of N generators $\{X_j\}_{j=1}^N$ and a corresponding set of free parameters $\{\alpha_j\}_{j=1}^N$. This means that in all later applications the structure of the generators $\{X\}$ will characterise the group completely.

B.5.2 Lie algebras

Now Lie groups have been introduced and their general concept has already been omitted in favour of the idea of generators. For a certain parameterisation of Lie groups it was shown that these generators could be used to specify the group structure completely. This section now analyses the structure, which is produced by any exponential representation of Lie groups.
In order to do so, the combination of different representation elements has to be examined. From the construction (see previous section) one knows already how two transformations in the same direction can be combined:

$$D(\alpha^1)D(\alpha^2) = e^{\mathrm{i}\alpha_j^1 X_j} e^{\mathrm{i}\alpha_j^2 X_j} = e^{\mathrm{i}(\alpha_j^1+\alpha_j^2)X_j} = D(\alpha^1 + \alpha^2) \tag{B.5}$$

Here α^1 and α^2 are two parameter-sets in the same representation-direction, but with possibly different magnitudes. So for transformations in the same direction one sees from Eq. (B.5) that their parameters simply add for the combined transformation. Unfortunately for two transformations in different directions (α and β) this is not true. The only thing one knows from the definition of groups is that the combination of any two representation elements must be within the whole representation as well (B.1(i)):

$$D(\widetilde{\gamma}) = D(\alpha)D(\beta) \tag{B.6}$$

$$e^{\mathrm{i}\widetilde{\gamma}_c X_c} = e^{\mathrm{i}\alpha_a X_a} e^{\mathrm{i}\beta_b X_b} \tag{B.7}$$

$$\Rightarrow \quad \mathrm{i}\widetilde{\gamma}_c X_c = \log\left(e^{\mathrm{i}\alpha_a X_a} e^{\mathrm{i}\beta_b X_b}\right) \tag{B.8}$$

This equation can be used to derive the combination rules for two different generators X_a and X_b. If the generators commute with each other, then the solution is trivial, giving just the combination rule Eq. (B.5) for all generators. But in general this is not true, since the generators can be viewed as matricies, which do not always commute. Nevertheless there is a solution to equation B.8, which is called the Baker-Campbell-Hausdorff formula. It has the form:

$$\mathrm{i}\widetilde{\gamma}_c X_c = \log\left(e^{\mathrm{i}\alpha_a X_a} e^{\mathrm{i}\beta_b X_b}\right) = \mathrm{i}(\alpha_a X_a + \beta_b X_b) - \frac{1}{2}[\alpha_a X_a, \beta_b X_b] + ... \tag{B.9}$$

The derivation of this formula is simple but mostly technical. Essentially a double Taylor expansion in $\{\alpha\}$ and $\{\beta\}$ is used. Crucial is only that the righthand side can be expressed completely in powers of commutators of $\alpha_a X_a$ and $\beta_b X_b$. So if all commutation relations for the generators are known then in principle all terms of the expansion could be calculated. In essence this is what makes generators of Lie groups so powerful. The relatively easy commutation relations of generators determine the complete group structure.
Equation Eq. (B.9) can be brought to a more standard form, if all higher order contributions are ignored (this corresponds to the infinitesimal case):

$$\mathrm{i}\widetilde{\gamma}_c X_c = \mathrm{i}(\alpha_a X_a + \beta_b X_b) - \frac{1}{2}[\alpha_a X_a, \beta_b X_b] \tag{B.10}$$

$$\alpha_a\beta_b[X_a, X_b] = \mathrm{i}\underbrace{2\,(\alpha_c + \beta_c - \widetilde{\gamma}_c)}_{\equiv\gamma_c} X_c \tag{B.11}$$

$$\alpha_a\beta_b[X_a, X_b] = \mathrm{i}\gamma_c X_c \tag{B.12}$$

$$[X_a, X_b] = \mathrm{i} f_{abc} X_c \tag{B.13}$$

In the second line the different summation indicies have been renamed and in the last line the renaming $\gamma_c = \alpha_a\beta_b f_{abc}$ was used. Equation B.13 defines an antisymmetric, bilinear operation law $[\cdot,\cdot]$. The Lie algebra (associated with a Lie group G) is a vector space $\mathfrak{g}$ over a field F with this operation law. In addition the Lie algebra has to fulfil the Jacobi identity[7)].
The f_{abc} are called the structure constants of the group. They are specific for each Lie group. The generators $\{X\}$ can have different forms, depending on the dimension of the space they live in.
From equation B.13 two properties of the structure constants can be read off directly:

- As the commutator is antisymmetric it follows that $f_{abc} = -f_{bac}$ is antisymmetric as well.
- If the representation $U(\alpha) = e^{\mathrm{i}\alpha_a X_a}$ is unitary, then the structure constants are real.
  For this one needs to remember that the generators are hermitian for a unitary representation. Knowing this the result can simply be calculated:

$$[X_a, X_b]^\dagger = (\mathrm{i} f_{abc} X_c)^\dagger \tag{B.14}$$

$$[X_b, X_a] = -\mathrm{i} f^*_{abc} X_c \tag{B.15}$$

$$\rightarrow [X_a, X_b] = \mathrm{i} f^*_{abc} X_c \tag{B.16}$$

  Equation B.16 in combination with equation B.13 lead to the desired result $f^*_{abc} = f_{abc}$.

[7)]Jacobi identity: $[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$

B.5.3 Important Lie groups

So far Lie groups have been introduced on a general footing without focussing on their connection to symmetries. This section will make up for the leeway and some of the most important Lie groups in physics shall be introduced briefly. But before doing so it is useful to provide a little list of important expressions in the characterisation of groups.

Notation - The mathematical notation for Lie groups consists of a name, the dimension of the group generators in the fundamental representation ($\hat{=}$ loosely speaking the smallest dimensional faithful representation) and the field over which the group is defined. So $A_n(\mathbb{C})$ is the group A, with dimension n over the complex numbers. In physics the field typically is $\mathbb{C}$ and so the notation is often shortend to $A_n(\mathbb{C}) = A(n)$. This convention is adopted throughout this work as well.

Subgroups - are groups within larger groups. A subgroup itself must fulfil all the group axioms (B.1) and it is embedded in the structure of a larger group.

Connected - is a mathematical concept that does not only apply to groups. In the context of group theory a group is connected if there exists a similarity transformation S for every element that connects it to the identity ($S m_G S^{-1} = \mathbb{1}$). A simple example where this fails are matrices with negative determinant. Consequently a group that is not connected is called disconnected[8)].

Traceless - The trace is a tool that allows to characterise different Lie algebras. Traces for Lie algebras are equivalent to something called 'characters' for groups. Without going into the details of characters it should be mentioned, that they do not change under basis transformations of the group and so they qualify for a tool of characterisation.
Another feature, that is important for physical applications is that traceless Lie algebras do not change the volume and orientations in infinitesimal transformations. The associated groups to these Lie algebras are called 'special'. They have the constraint that every group element needs to have unit determinant. These groups leave the volume and orientations in general unchanged.

Using these features one can now give promised list of important groups. For this take $M_n(K)$ denote the square-matrices of size $n \times n$ with entries from the field K.

$GL(n) = GL_n(\mathbb{C})$ is the group of invertible, complex matrices (under matrix multiplication). It is called general linear group and its definition is $GL_n(K) := \{A \in M_n(K) | A \text{ is invertible}\}$. As there are no other constraints this is indeed a very general structure. All other examples will be subgroups of the $GL(n)$.

$SL(n) = SL_n(\mathbb{C})$ is the special linear group. It is the group of all invertible matrices whose determinant is one - $SL_n(K) := \{A \in M_n(K) | \det(A) = 1\}$. This means that these are all linear transformations that

8) The groups of interest here are the so called topological groups. For such groups topological properties have simple analogies in the space of the group.

leave volume and orientations fixed. Although being special, is still a bit too general to describe symmetries in physics.

$O(n) = O_n(\mathbb{R})$ is the first group with an important interpretation in physical applications. It is called the orthogonal group with the definition $O_n(K) := \{A \in M_n(K) | AA^{\mathrm{T}} = A^{\mathrm{T}}A = \mathbb{1}\}$. The orthogonality condition leads to the condition for the determinants $\det(A) = \pm 1$ and so this group corresponds to symmetry transformations that leave the length of a real vector invariant. This means that $O(n)$ leaves volumina unchanged but it can reverse the orientations of vectors. A relevant example for this group in physics is the Lorentz group ($O^{(3,1)}$), which will be discussed in Sec. II.1.

$SO(n) = SO_n(\mathbb{R})$, just like $SL(n)$, has the extra condition that its constituent matrices have unit determinant. It is called the special orthogonal group $SO_n(K) := \{A \in O(n) | \det(A) = 1\}$. In contrast to $O(n)$ this group does not change the orientations of vectors. For the Lie algebra the conditions of the group lead to *traceless*, *real* and orthogonal matrices. This group finds a direct application in physics as well - it represents ordinary rotations in n dimensions.

$U(n) = U_n(\mathbb{C})$ is the analogue of $O(n)$ in the complex case and so its definition is $U_n(\mathbb{C}) := \{A \in GL_n(\mathbb{C}) | A^\dagger A = AA^\dagger = \mathbb{1}\}$. As $O(n)$ it contains all transformations that leave volumina and the origin unchanged, but this time in a complex vector space and so the condition on the determinant is $|\det(A)| = 1$.

$SU(n) = SU_n(\mathbb{C})$ is very similar to $SO(n)$. It is the special unitary group with the definition $SU_n(\mathbb{C}) := \{A \in U_n(\mathbb{C}) | \det(A) = 1\}$. So it is the analogon of $SO(n)$ for complex vector spaces.
For the algebra things look very similar as well. It consists of *traceless* matrices that are *complex* this time. This group as well is used in a lot of situations in physics. It will play a key role for the symmetries regarded in this work.

B.5.4 Examples: Lie algebras

Throughout the last sections a lot of material has been introduced without any explicit examples. While the whole theory of Lie algebras can be introduced without working in a explicit basis, it is far more practical for the present purpose to give explicit matrix representations for the generators of the important Lie groups.

U(1) - This is the first and easiest Lie group to be presented. It only has a single one dimensional generator $X_1 = 1$ and so the group elements are $U(\alpha) = e^{\mathrm{i}\alpha}$. It follows directly that $U(1)$ is abelian ($[U(\alpha), U(\beta)] = 0$).

SU(2) - has the first non-trivial algebra. There are three linear independend complex traceless matrices in two dimensions, the Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \quad , \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{B.17}$$

Including a factor of $1/2$ these are the generators $X_a = \sigma_a/2$ of $SU(2)$. The commutation relation resulting from these matrices turn out to be

$$[X_a, X_b] = \frac{1}{4}[\sigma_a, \sigma_b] = \mathrm{i}\epsilon_{abc}\frac{1}{2}\sigma_c = \mathrm{i}\epsilon_{abc}X_c \tag{B.18}$$

The group elements are then analogous to the $U(1)$ case: $U(\alpha) = e^{\mathrm{i}\frac{\alpha_a}{2}\sigma_a}$. Here the notation is slightly confusingly, as $U(1)$ refers to a group, while $U(\alpha)$ means elements of a particular group.

SO(3) - As this group corresponds to rotations in three dimensions it should better have three generators, resulting in three free parameters[9)]. Indeed there are three linearly independend, orthogonal and traceless matrices in three dimensions:

$$J_1 = \mathrm{i}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad J_2 = \mathrm{i}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad , \quad J_3 = \mathrm{i}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \tag{B.19}$$

And again, by including a normalisation factor of $1/2$ one gets the generators $X_a = J_a/2$ for $SO(3)$. The algebra of $SO(3)$ now produces something remarkable:

$$[X_a, X_b] = \frac{1}{4}[J_a, J_b] = \mathrm{i}\epsilon_{abc}\frac{1}{2}J_c = \mathrm{i}\epsilon_{abc}X_c \tag{B.20}$$

Comparing Eq. (B.18) with Eq. (B.20) one could think that $SU(2)$ and $SO(3)$ are the same groups. But this is not true! The algebras are the same but these do only describe the groups *locally*. On a global scale both groups are different. This is not important in later derivations, but as the problem appears in this context it should at least be mentioned that there is a difference between Lie groups and algebras.

These examples are only the most simple ones but higher dimensional generators for $SU(n)$ and $SO(n)$ can be built by just mimicking the matrices from the two examples. $SU(3)$ for example consists of eight generators. There are six generators of the type σ_1 and σ_2 and two diagonal traceless generators.

The given examples are chosen partly to show how well known symmetries translate into the formalism of Lie algebras. But this formalism is capable of going far beyond a description of 'classical' symmetries. The structure of the exponential map $D(\{\alpha\}) = e^{\mathrm{i}\alpha_a X_a}$ allows for many different symmetries by simply changing the set of generators $\{X_a\}$. In addition one could mess with the set of parameters. They could

[9)] They can be chosen to be the Euler angles be made space-time dependend for example ($\{\alpha\} \rightarrow \{\alpha(x)\}$). The variety of possible symmetries that can be described via Lie groups is another feature that makes them important in modern physics.

B.6 Observables in $SU(2)$

Of course, the $SU(2)$ group is a Lie group and following this terminology this section should rather be included in the part on Lie groups. To make matters worse the main concept this section will deal with (the 'highest weight decomposition') is not unique to $SU(2)$ but can be used in the characterisation of $GL(n)$ groups in general. The proof that it can be used for groups of $GL(n)$ is given in Ref. [35, p.p.126-130]. As always, whenever logical structure is bluntly put aside in this work, it is due to the limited time. While a deeper introduction to Lie groups would generalise the findings of this section to other groups this will not be done here, as the techniques for $SU(2)$ are sufficient for later derivations. Just keep in mind that this is not the end of the game.

Before going into details of construction, suppose that a system is invariant under $SU(2)$ transformations. As $SU(2)$ is non-abelian, not all generators ($\hat{=} J_i$) can be diagonalised simultaneously. For physical systems only the eigenvalues of the simultaneously diagonalised operators are measurable at the same time. So only the diagonalised generators of $SU(2)$ can be related to observable quantities. The idea of the 'highest weight decomposition' is to furnish a basis for exactly those states, which are physically measurable.

For the 2-dimensional fundamental representation of $SU(2)$ this is a rather trivial decomposition. In this representation one can see right away, that only one generator can be diagonalised at a time (as they are hermitian and non commutative - compare Eq. (B.17), Eq. (B.18)). This generator can be taken to be $J_3 = \mathrm{diag}(1/2, -1/2)$. The eigenstates of J_3 then build a *complete* basis of physically observable states in the space, $SU(2)$ acts on. Typically these states are labelled with the maximal eigenvalue j of the diagonalised generator and with the eigenvalue m, that the particular state corresponds to. The m value of a state is also called '*weight*'. So in the present case one has the basis set $B = \{|j, m\rangle\} = \{|1/2, 1/2\rangle, |1/2, -1/2\rangle\}$, with the eigenvalue (weight) equation and orthonormality relation:

$$J_3 |j, m\rangle = m |j, m\rangle \tag{B.1}$$

$$\langle j_1, m_1 | j_2, m_2\rangle = \delta_{j_1 j_2} \delta_{m_1 m_2} \tag{B.2}$$

The $\delta_{j_1 j_2}$ part of this relation will be important in the next section, once tensor products are discussed. It means that representations with different maximal weights can be chosen to be orthogonal. For the present $j = 1/2$ case the above equations mean that any operator, which can be written in terms of the irreducible 2-dimensional representation of $SU(2)$, can at best be associated with one observable quantity, which can take on the values $1/2$ or $-1/2$ (up to a normalisation). The states of physical systems can be

expressed in terms of the basis B.
Although equation B.1 was introduced from the fundamental representation of $SU(2)$, it holds for representations with any dimension. For these other examples the possible values of j and m change but they are still sufficient to label and distinguish all states, which can correspond to physical systems. Ultimately the reason for this is that any representation of $SU(2)$ has to fulfil the algebra equation ($[J_i, J_j] = \mathrm{i}\epsilon_{ijk}J_k$), but a more explicit construction of general $SU(2)$ states $|j,m\rangle$ will be given in the following two sections.

B.6.1 Raising & Lowering operators

So far J_3 has been used to give a unique label to different states of physical observables in a $SU(2)$-invariant system. The underlying symmetry allows to transform a state $|j,m_1\rangle$ into another one $|j,m_2\rangle$. This transformation can be expressed in terms of the missing two generators of $SU(2)$ by introducing raising & lowering operators $J^{\pm} = (J_1 \pm \mathrm{i}J_2)/\sqrt{2}$. Using the definition of the algebra, one finds the following commutation relations:

$$[J_3, J^{\pm}] = \pm J^{\pm} \tag{B.3}$$

$$[J^+, J^-] = J_3 \tag{B.4}$$

With these definitions one can work out how a basis state changes under the action of $J^{\pm}$:

$$J_3(J^{\pm}|j,m\rangle) = J^{\pm}J_3|j,m\rangle \pm J^{\pm}|j,m\rangle = (m \pm 1)(J^{\pm}|j,m\rangle) \tag{B.5}$$

and this explains the names for $J^{\pm}$. They raise, or lower the m value of a state $|j,m\rangle$ by 1 and thereby (up to a normalisation factor) $J^{\pm}$ can be used to transform $|j,m_1\rangle$ into $|j,m_2\rangle$.
To work out the normalisation of each state one needs to assume that the representation is finite dimensional. In this case there is a highest weight j and a lowest weight $j-\ell$, for a particular ℓ and one has the conditions:

$$\text{finite highest weight:} \qquad J^+|j,j\rangle = 0 \tag{B.6}$$

$$\text{finite lowest weight:} \qquad J^-|j,j-\ell\rangle = 0 \tag{B.7}$$

From these two requirements the normalisation of each state can be constructed recursively using equation B.1 and B.2. This derivation is presented nicely in Ref. [2, p.56-63]. It leads to the relations:

$$J_3|j,m\rangle = m|j,m\rangle \tag{B.8}$$

$$J^+|j,m\rangle = N_{jm+1}|j,m+1\rangle \tag{B.9}$$

$$J^-|j,m\rangle = N_{jm}|j,m-1\rangle \tag{B.10}$$

$$N_{jm} = \sqrt{(j+m)(j-m+1)/2} \tag{B.11}$$

The number of possible states, constructed by the three equations Eq. (B.8)-(B.11), is $\ell = 2j+1$ (since $N_{j\ell} = 0$). This is the reason why the labels j and m are sufficient to distinguish the elements of a complete basis B for all matrix representations of $SU(2)$. A $2j+1$ dimensional matrix can have up to $2j+1$ non-trivial eigenvectors. As Eq. (B.8) generates exactly $2j+1$ linearly independent eigenvectors of the matrix J_3, these vectors form a complete basis for the matrix representation[10)].
This, by the way, gives another neat relation between the highest weight of a representation and its dimension: $\dim(rep) = 2j+1$.

B.6.2 Highest weight decomposition

With the construction rules Eq. (B.8)-(B.11) from the previous section the highest weight decomposition now is a mere recipe. Nevertheless the scheme is useful as it is can be applied not only to irreducible representations (as in the previous section), but also to tensor product representations (following section). For tensor products the neat side effect of the highest weight decomposition is that it decomposes a reducible representation into a direct sum of irreducible ones.
For the generic scheme (according to Ref. [2, p.62]) suppose that J_3 is the generator to be diagonalised and states are labelled by $|j,m;\alpha\rangle$. Here j is the highest weight of the representation, m is the weight of the state and α is a label for any other measurable observable, which is independend of the symmetry transfromation.

1 - Diagonalise J_3

2 - Take the state with the highest J_3 weight ($|j,j;\alpha\rangle$)

3 - For each such highest weight state build all related states by applying the lowering operator J^- as often as possible. This constructs the irreducible j representation (also called spin j representation).

4 - Set aside the states of the constructed irreducible representations. (The remaining states are orthogonal to the constructed ones.)

[10)]Mathematically this can be seen as J_3 is hermitian, which leads to diagonalisability and so the eigenvectors of J_3 can form an orthonormal basis

5 - Find the highest J_3 weight of the remaining states ($|j, \widetilde{j}; \alpha\rangle$) and procede with step three.

Following this scheme until all existing states are used generates a complete orthonormal basis of the Hilbert space corresponding to the problem.

$$\langle j', m'; \alpha' | j, m; \alpha\rangle = \delta_{j' j}\delta_{m' m}\delta_{\alpha' \alpha} \tag{B.12}$$

If the starting representation is irreducible, then the highest weight is unique and so the scheme ends rather quickly. For a reducible representation the highest weights of each irreducible representations is still unique (see Sec. B.6.3). Therefore the scheme simply runs through several times, each time giving the states of different representations.
Only if there is a non-trivial other label α for each state, the highest weight might not be unique. In this case one simply gets copies of the irreducible j representations for each value of α. The proper Lorentz group can be viewed as an example for this case. Each basis state of this group transforms under two *independent* $SU(2)$ algebras. More details on the Lorentz group can be found in Sec. II.1. In general there are more sophisticated ways to determine the decomposition in irreducible representations of tensor products, but for the present purpose this scheme will do.

B.6.3 Tensor products for Lie groups

In Sec. B.3 the general idea of tensor products has been introduced. For Lie groups the proximity of group elements to the identity can be used to simplify the action of tensor operators on tensor states significantly. In other words, expressing group elements via infinitesimal transformations gives simple relations for the generators of the tensor algebra.
In analogy to the situation of Sec. B.3 take two systems 1 and 2. 1 is in some representation with basis $\{m\}$ and it is invariant under the representation $D_1(g)$ of some Lie group G with $g \in G$. At the same time 2 is in a representation with basis $\{y\}$ and is invariant under the representation $D_2(g)$. The transformation of a tensor state is then (just as in Sec. B.3):

$$D(g)|my\rangle = D_{1\otimes 2}(g)|m\rangle|y\rangle = (D_1(g)|m\rangle) \otimes (D_2(g)|y\rangle) \tag{B.13}$$

As D_1 and D_2 are Lie groups in the exponential representations, they can be rewritten in a power series and for elements close to the identity one gets:

$$
\begin{aligned}
D(g)|mx\rangle &= D_1(g)|m\rangle \otimes D_2(g)|y\rangle & \text{(B.14)}\\
&= \left(I + \mathrm{i}\alpha_a X_a^1 + O(\alpha^2)\right)|m\rangle \otimes \left(I + \mathrm{i}\beta_b X_b^2 + O(\beta^2)\right)|y\rangle & \text{(B.15)}\\
&= \left(\delta_{lm} + \mathrm{i}\alpha_a [X_a^1]_{lm} + O(\alpha^2)\right)|m\rangle \otimes \left(\delta_{xy} + \mathrm{i}\beta_b [X_b^2]_{xy} + O(\beta^2)\right)|y\rangle & \text{(B.16)}\\
&= \Big(\delta_{lm}\delta_{xy} + \mathrm{i}\underbrace{\left(\alpha_a [X_a^1]_{lm}\delta_{xy} + \beta_b \delta_{lm}[X_b^2]_{xy}\right)}_{\equiv \gamma_c [X_c^{1\otimes 2}]_{lm\,xy}} + O(\alpha_a\beta_b)\Big)|ym\rangle & \text{(B.17)}
\end{aligned}
$$

And thus one finds the generators of the tensor representation:

$$
\begin{aligned}
[X_c^{1\otimes 2}]_{lm\,xy} &= [X_c^1]_{lm}\delta_{xy} + \delta_{lm}[X_c^2]_{xy} & \text{(B.18)}\\
X_c^{1\otimes 2} &= X_c^1 \otimes I + I \otimes X_c^2 & \text{(B.19)}
\end{aligned}
$$

This is a very helpful result as it directly tells that a generator in the tensor representation acts on states by simply acting consecutively on each constituent representation:

$$
X_a^{1\otimes 2}\left(|m\rangle|x\rangle\right) = \left(X_a^1|m\rangle\right)|x\rangle + |m\rangle\left(X_a^2|x\rangle\right) \tag{B.20}
$$

B.6.3.1 Example

As mentioned in Sec. B.3 typically tensor product representations are reducible. The transformation rules for tensor generators combined with the highest weight decomposition now allows to decompose tensor products of various $SU(2)$ representations. Since these decompositions will be important in later derivations one explicit example will be given:

Regard the tensor product space of two independent spin 1 systems ($j = 1$). So each subsystem lives in a three dimensional $SU(2)$ invariant space and from Sec. B.3 one knows that the tensor space has dimension $9 = 3 \cdot 3$ (the basis of the tensor space are simply all possible combinations of two spin 1 states $\{|1, m_1\rangle|1, m_2\rangle\}_{m_1,m_2}$).

It is conventional to label generators of different representations with their dimension (e.g. J_3^3 is the diagonal generator of $SU(2)$ in the three dimensional representation). The $J_3^{3\otimes 3}$ generator in tensor space has the nice feature that its weights are simply the sum of the weights in the subsystems (compare Eq. (B.20)):

$$
\begin{aligned}
J_3^{3\otimes 3}|j_1, m_1\rangle|j_2, m_2\rangle &= \left(J_3^3|j_1, m_1\rangle\right)|j_2, m_2\rangle + |j_1, m_1\rangle\left(J_3^3|j_2, m_2\rangle\right) & \text{(B.21)}\\
&= (m_1 + m_2)|j_1, m_1\rangle|j_2, m_2\rangle & \text{(B.22)}
\end{aligned}
$$

The highest weight state in each subsystem is $|j = 1, m = 1\rangle$. Using the preceding equation (Eq. (B.22)) one sees that the highest weight in the tensor representation comes from the highest weight states of both subsystems.

$$|2,2\rangle = |1,1\rangle|1,1\rangle \tag{B.23}$$

This finding allows now to decompose the 9 dimensional reducible tensor representation into lower dimensional irreducible ones. The first irreducible representation is found by applying the (normalised) lowering operator to both sides of Eq. (B.23) recursively and by the use of the equations Eq. (B.8)-(B.11) one gets:

$$|2,2\rangle = |1,1\rangle|1,1\rangle \tag{B.24}$$

$$|2,1\rangle = \frac{1}{N_{22}}J^-|2,2\rangle = \frac{N_{11}}{N_{22}}\Big(|1,0\rangle|1,1\rangle + |1,1\rangle|1,0\rangle\Big) \tag{B.25}$$

$$= \frac{1}{\sqrt{2}}\Big(|1,0\rangle|1,1\rangle + |1,1\rangle|1,0\rangle\Big) \tag{B.26}$$

$$|2,0\rangle = \frac{1}{\sqrt{3\cdot 2}}J^-|2,1\rangle = \frac{1}{\sqrt{6}}\Big(|1,-1\rangle|1,1\rangle + |1,1\rangle|1,-1\rangle + 2|1,0\rangle|1,0\rangle\Big) \tag{B.27}$$

$$|2,-1\rangle = \frac{1}{\sqrt{3\cdot 6}}J^-|2,0\rangle = \frac{1}{\sqrt{18}}\Big(3|1,-1\rangle|1,0\rangle + 3|1,0\rangle|1,-1\rangle\Big) \tag{B.28}$$

$$|2,-2\rangle = \frac{3}{\sqrt{2\cdot 18}}J^-|2,-1\rangle = |1,-1\rangle|1,-1\rangle \tag{B.29}$$

The apperance of the normalisation constants is only presented explicitly in the first step, but it works similarly for all other lines (simply by following the rules of equations Eq. (B.8)-(B.11)). As this representation consists of 5 states there are still 4 missing for a complete basis of the 9 dimensional tensor space.

Ignoring the states with weight 2 the highest weight of the remaining missing states is 1, since the constituent weights are natural numbers and so they have to be lowered at least by 1. Different irreducible representations shall be orthogonal to each other. In order to guarantee this, the spin 1 tensor states can be chosen to be antisymmetric. After antisymmetrising the highest weight state, the rest of the spin 1 representation can be found by applying the lowering operator once again.

$$|1,1\rangle = \frac{1}{\sqrt{2}}\Big(|1,0\rangle|1,1\rangle - |1,1\rangle|1,0\rangle\Big) \tag{B.30}$$

$$|1,0\rangle = \frac{1}{\sqrt{2}}\Big(|1,-1\rangle|1,1\rangle - |1,1\rangle|1,-1\rangle\Big) \tag{B.31}$$

$$|1,-1\rangle = \frac{1}{\sqrt{2}}\Big(|1,-1\rangle|1,0\rangle - |1,0\rangle|1,-1\rangle\Big) \tag{B.32}$$

One can check that the antisymmetric construction ensures, that all scalar products of the spin 2 with the spin 1 representation give zero (remember the normalisation B.2). This sets the counter of states up to 8, leaving one missing state. The maximal remaining weight is 0 and so the final state is:

$$|0,0\rangle = \frac{1}{\sqrt{6}}\Big(|1,-1\rangle|1,1\rangle + |1,1\rangle|1,-1\rangle - 2|1,0\rangle|1,0\rangle\Big) \tag{B.33}$$

Notice that the spin 2 and the spin 0 representations are symmetric under the exchange of subsystems, while the spin 1 representation is antisymmetric. There is a general pattern behind this. In a combined system, that is built out of two subsystem in equivalent spin j representations, the state with the (global) highest weight is always symmetric. Since $J^{\pm}$ does not change the symmetry properties of a state, the irreducible representation belonging to this weight is always symmetric. The next representation (here spin 1) has to be chosen antisymmetric in order to fulfil the orthonormality relation (B.2). In the previous case the last remaining state ($|0,0\rangle$) is symmetric. But in a general setting (e.g. if there are more than three irreducible representations) the symmetry properties of all remaining representations aren't so obvious. Nevertheless a lengthy analysis yields the very convenient result that symmetric and antisymmetric irreducible representations alternate in the highest weight scheme if the original subsystems are in equivalent (spin j) $SU(2)$ representations.

The lengthy procedure of decomposition can be encoded in nice little equations which will only capture the most important facts. For this an irreducible $SU(2)$ representation is pictured only by its dimension and a subscript is introduced that denotes if the representation is symmetric (s) or antisymmetric (a). Using ths notation the above example of the tensor product decomposition reads:

$$3 \otimes 3 = 5_s \oplus 3_a \oplus 1_s \tag{B.34}$$

While this notation at the moment appears only as a fancy way of writing, it turns out to be very useful for the mathematical construction of physical systems. Suppose one would like to construct a Hamiltonian density out of a set of constituent elements that transform according to some symmetry group. The Hamiltonian density itself must be invariant under the action of any symmetry transformation as it is a scalar (it simply corresponds to the total energy density). So *all* possible terms that can appear in the Hamiltonian density must have scalar properties under symmetry transformations. A scalar, in the language of Eq. (B.34) is an object in the symmetric singlet (1_s) irreducible representation (there is only one state in the representation and it is symmetric, so the transformation cannot change anything about it). And so, by studying the tensor products of all constituents of the physical system, one can find all terms that can be included in the Hamiltonian density (all terms that can be decomposed into a 1_s representation + stuff).
In appendix App. A.4 all $SU(2)$ tensor products that are needed in this work are given.

Bibliography

[1] K. S. Takesi Saito, Progress of Theoretical Physics **61**, 608 (1979).

[2] H. Georgi, *Lie Algebras in Particle Physics* (Westview Press, 1999).

[3] S. Sternberg, *Group Theory and Physics* (Cambridge University Press, 1995).

[4] S. Weinberg, *The Quantum Theory of Fields I* (Cambridge University Press, 2005).

[5] M. Srednicki, *Quantum Field Theory* (Cambridge University Press, 2007).

[6] V. S. Varadarajan, *Geometry of Quantum Theory* (Springer-Verlag, 2007).

[7] M. Artin, *Algebra* (Prentice Hall, 1991).

[8] J. Gallier, *Geometric Methods and Applications: For Computer Science and Engineering* (Springer-Verlag, 1994).

[9] P. R. Girard, *Quaternions, Clifford Algebras and Relativistic Physics* (Birkenhäuser Verlag AG, 2007).

[10] E. Noether, Göttinger Nachrichten , 235 (1918).

[11] V. A. Miransky, *Dynamical Symmetry Breaking in Quantum Field Theories* (World Scientific Publishing Company Inc., 1994).

[12] K. G. Wilson, Rev. Mod. Phys. **55**, 583 (1983).

[13] H. Georgi, Annu. Rev. Nucl. Part. Sci. **43**, 209 (1993).

[14] S. Weinberg, *The Quantum Theory of Fields II* (Cambridge University Press, 2005).

[15] A. Zee, *Quantum Field Theory in a Nutshell* (Princeton University Press, 2003).

[16] T. Schäfer and E. V. Shuryak, Rev. Mod. Phys. **70**, 323 (1998).

[17] S. H. Weintraub, *Differential forms: a complement to vector calculus* (Academic Press, 1996).

[18] J. Lee, *Introduction to Smooth Manifolds* (Springer-Verlag, 2003).

[19] J. Roe, *Elliptic operators, topology and asymptotic methods* (Chapman & Hall/CRC, 1998).

[20] S. Weinberg, Phys. Rev. Lett. **40**, 223 (1978).

[21] S. Vandoren and P. van Nieuwenhuizen, (2008), 0802.1862.

[22] A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Y. S. Tyupkin, Physics Letters B **59**, 85 (1975).

[23] G. 't Hooft, Phys. Rev. D **14**, 3432 (1976).

[24] L. S. Brown, R. D. Carlitz, and C. Lee, Phys. Rev. D **16**, 417 (1977).

[25] M.-y. Wang, (1994), hep-ph/9502333.

[26] L. G. Yaffe, Nuclear Physics B **151**, 247 (1979).

[27] D. E. Morrissey, T. M. P. Tait, and C. E. M. Wagner, Phys. Rev. **D72**, 095003 (2005), hep-ph/0508123.

[28] A. I. Vaĭnshteĭn, V. A. Novikov, M. A. Shifman, and V. I. Zakharov, Sov. Phys. Usp. **25**, 201 (1982).

[29] G. Derrick, Mathematical Physics **5**, 1252 (1964).

[30] I. Affleck, Nuclear Physics B **191**, 429 (1981).

[31] O. Espinosa, Nuclear Physics B **343**, 310 (1990).

[32] B. Kastening, Physics Letters B **291**, 92 (1992).

[33] M. Gell-Mann and M. Levy, Il Nuovo Cimento (1955-1965) **16**, 705 (1960), 10.1007/BF02859738.

[34] U. Mosel, *Fields, Symmetries, and Quarks* (Springer-Verlag, 2004).

[35] D. P. Zelobenko, *Compact Lie Groups and Their Representations* ((Russian) Moscow, Nauka; Second edition, MCCME publishers, 1970).

Danksagung

Im Zusammenhang mit dieser Arbeit möchte ich besonders meinem Betreuer, Prof. Dr. Lenske für die Möglichkeit danken eine Masterarbeit mit diesem Thema in seiner Arbeitsgruppe anzufertigen. Die ungezwungene, stets hoch motivierte Atmosphäre in Besprechungen und Diskussionen hat zum erheblichen Teil zum Gelingen dieser Arbeit beigetragen. Besonders für den Freiraum, im Bezug auf das Ausarbeiten eigener Akzente bin ich Herrn Prof. Dr. Lenske sehr zu Dank verpflichtet. Außerdem wäre die thematische Abrundung vieler Aspekte dieser Arbeit nicht ohne die vielen themenübergreifenden Anmerkungen meines Betreuers möglich gewesen.
Des Weiteren gilt mein Dank Ivan Lappo-Danilevski und meinem Bruder, Eike Fokken. Beide haben mir die Erschließung der unzähligen Themenkomplexe durch spannende Diskussionen und bohrende Fragen stark erleichtert. Außerdem haben beide die Bürde auf sich genommen Teile meiner Ausarbeitungen Korrektur zu lesen, was ich ihnen nicht hoch genug anrechnen kann.

Erklärung zur Urheberschaft

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig verfasst habe und dabei keine anderen als die angegebenen Quellen und Hilfsmittel zum Einsatz kamen.

Giessen, den 24. Mai 2011 Jan Fokken