Global Anomalies
in
Chiral Lattice Gauge Theories

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von
Oliver Bär
aus Frankfurt am Main

Hamburg
2000

Gutachter der Dissertation:	Prof. Dr. Martin Lüscher
	Prof. Dr. Klaus Fredenhagen
Gutachter der Disputation:	Prof. Dr. Klaus Fredenhagen
	Prof. Dr. Joachim Bartels
Datum der Disputation:	4. Juli 2000
Dekan des Fachbereiches Physik und Vorsitzender des Promotionsausschusses:	Prof. Dr. Friedrich-Wilhelm Büßer

Abstract

We study global anomalies in a new approach to chiral gauge theories on the lattice, which is based on the Ginsparg-Wilson relation. In this approach, global anomalies make it impossible to define consistently a fermionic measure for the functional integral.

We show that a global anomaly occurs in an SU(2) theory if the fundamental representation is used for the fermion fields. The generalization to higher representations is also discussed. In addition we establish a close relation between global anomalies and the spectral flow of the Dirac operator and employ it in a numerical computation to prove the existence of the global SU(2) anomaly in a different way. This method is inspired by an earlier work of Witten who first discovered this type of anomalies in continuum field theory.

Zusammenfassung

Wir untersuchen globale Anomalien in einem neuen Zugang zu chiralen Eichtheorien auf dem Gitter, welcher auf der Ginsparg-Wilson Relation beruht. In diesem Zugang verhindern globale Anomalien die konsistente Definition eines fermionischen Maßes für das Funktionalintegral.

Wir zeigen, daß eine globale Anomalie in einer SU(2) Theorie auftritt, wenn für die Fermionfelder die Fundamentaldarstellung verwendet wird. Die Verallgemeinerung auf höhere Darstellungen wird ebenfalls diskutiert. Des weiteren stellen wir einen engen Zusammenhang zwischen globalen Anomalien und dem Spektralfluss des Diracoperators her und nutzen diesen, um das Vorliegen der globalen SU(2) Anomalie auf eine unterschiedliche Art zu beweisen. Diese Methode lehnt sich an eine frühere Arbeit von Witten an, der diese Art von Anomalien als erster in der Kontinuumsfeldtheorie entdeckte.

Contents

Introduction

Symmetry is one of the most important concepts in modern particle physics. In classical field theory symmetries show up as the invariance of the classical action under some symmetry transformation. A quantum field theory , however, does not need to possess all the symmetries which are present in the corresponding classical theory. If a symmetry of the classical action cannot be preserved in the process of quantization, the theory is said to have an anomaly. In the framework of Euclidean field theories, for example, anomalies can appear due to the non-invariance of the fermion measure in the functional integral.

Depending on the symmetry an anomaly has different impact on the theory. An anomaly in a global symmetry usually results in observable properties of particles. Most prominent is the decay of the neutral pion into two photons which is totally dominated by the axial anomaly. Neglecting the contribution of the anomaly the prediction for the decay rate is roughly three orders of magnitude smaller than the experimental result.

On the other hand an anomaly in a gauge symmetry is not acceptable. The violation of gauge invariance results in anomalous Ward identities and destroys the renormalizability of the theory. Consequently gauge anomalies must cancel. This is usually achieved by properly choosing the particle content of the theory. The requirement of anomaly cancellation thus restricts the number of theoretically acceptable gauge theories.

In 1982 Witten discovered a peculiar type of gauge anomaly for an SU(2) chiral gauge theory with a single doublet of Weyl fermions. Usually gauge anomalies appear already for infinitesimal gauge transformations. In that case they can be discussed in perturbation theory and the name perturbative anomaly is sometimes used for them. In contrast to this Witten's SU(2) anomaly is called a non-perturbative or global anomaly because gauge invariance is broken for particular finite gauge transformations only.

Even if Witten's argument is quite convincing it is purely formal and cannot straightforwardly translated into a regularized form of the theory. In perturbation theory, for instance, it is necessary to first fix the gauge before one can apply a regularization scheme. To discuss Witten's anomaly it is crucial to avoid gauge fixing and regularized perturbation theory is therefore not the right tool.

The method of introducing a Euclidean space time lattice as a regulator has also its drawbacks. The Nielsen-Ninomiya theorem was regarded as a proof that one has to break chiral symmetry in order to avoid unwanted fermion species or loosing the loclity properties of the theory. Therefore it seemed to be impossible to set up chiral lattice gauge theories without making any compromises with respect to gauge invariance or locality.

Only recently an exact symmetry of the classical lattice action was discovered that can be regarded as the lattice form of chiral symmetry. Based on this symmetry a general framework for chiral lattice gauge theories with manifest gauge invariance has been proposed by Lüscher. These developments give rise to the question whether global anomalies can be treated properly using this new formalism.

The answer to this question is given in this thesis. For the first time the precise meaning of a global anomaly and its presence in the SU(2) chiral gauge theory is discussed in a regularized and mathematically well-defined quantum field theory.

The organization of this thesis is as follows. For later reference and to introduce the notion of global anomalies we review Witten's SU(2) anomaly in the first chapter. In chapter 2 the concept of exact chiral symmetry on the lattice based on the Ginsparg-Wilson relation is briefly introduced. Lüscher's approach to chiral gauge theories on the lattice is presented in chapter 3. In particular the definition of a fermion functional integral is thoroughly discussed and a precise definition of global anomalies is given.

Motivated by Witten's work the gauge group SU(2) is considered in chapter 4. Indeed we find a global anomaly for the fundamental representation of this gauge group. The generalization to higher representations is given in the next chapter.

A second method to deal with global anomalies, founded on a close connection between the anomaly and the spectral flow of the Dirac operator, is used in chapter 6. We conclude with some final remarks. Several appendices contain additional material and some rather technical derivations.

Chapter 1

Witten's SU(2) anomaly

In this chapter we briefly review Witten's global SU(2) anomaly. We first present Witten's original argument that a chiral SU(2) gauge theory with a single doublet of Weyl fermions is mathematically inconsistent. A second derivation in terms of Bardeen's non-Abelian anomaly is also given.

1.1 Witten's original argument

Consider a Euclidean SU(2) gauge theory coupled to a single doublet of Dirac fermions. In terms of the gauge potential A_μ and the fermion fields $\psi, \overline{\psi}$, the action reads

$$S \;=\; S_G + S_F\,, \quad S_F \;=\; \int \mathrm{d}^4x\; \overline{\psi}(x) D(A_\mu) \psi(x)\,. \tag{1.1}$$

S_G denotes the pure gauge part of the action and $D(A_\mu)$ stands for the covariant Dirac operator

$$D(A_\mu) \;=\; \gamma_\mu \left(\partial_\mu + A_\mu\right) + m\,. \tag{1.2}$$

We use the convention that the gauge coupling appears in S_G. The partition function of the theory is expressed by the functional integral

$$\mathcal{Z} = \int \mathcal{D}[A_\mu]\mathcal{D}[\psi]\mathcal{D}[\overline{\psi}]\; \mathrm{e}^{-S_G - S_F}\,. \tag{1.3}$$

Since $\psi(x)$ and $\overline{\psi}(x)$ are anti commuting elements of a Grassmann algebra, the integration over the fermion fields formally yields

$$\int \mathcal{D}[\psi]\mathcal{D}[\overline{\psi}]\; \mathrm{e}^{-S_F} \;=\; \det D(A_\mu) \tag{1.4}$$

with the so-called *fermion determinant* on the right hand side.

Equation (1.4) should be taken with care. The determinant, i.e. the product of all eigenvalues of the Dirac operator, is actually infinite. A regularization procedure, e.g. the zeta function regularization must be employed to give (1.4) a mathematically precise meaning. We will not carry out this here and the formal character of all equations and manipulations should be kept in mind in the following.

The Dirac fields $\psi(x)$ and $\overline{\psi}(x)$ can be decomposed into a sum of left- and right-handed parts, using the chiral projectors

$$P_- = \frac{1}{2}(1-\gamma_5)\,, \quad P_+ = \frac{1}{2}(1+\gamma_5)\,. \tag{1.5}$$

Here the left-handed components satisfy the equations

$$P_-\psi_L \;=\; \psi_L\,, \quad \overline{\psi}_L P_+ = \overline{\psi}_L\,. \tag{1.6}$$

If the mass m is equal to zero the fermion action in (1.1) can be written as a sum,

$$S_F \;=\; S_{F,L} + S_{F,R}\,, \tag{1.7}$$

where $S_{F,L}$ contains left-handed components only, while the right-handed ones occur solely in $S_{F,R}$. In analogy to (1.3) the partition function of a chiral SU(2) gauge theory with a single doublet of left-handed Weyl fermion reads

$$\mathcal{Z}_L = \int \mathcal{D}[A_\mu]\mathcal{D}[\psi_L]\mathcal{D}[\overline{\psi}_L]\,\mathrm{e}^{-S_G-S_{F,L}}\,. \tag{1.8}$$

Due to the pseudo reality of the fundamental representation of SU(2) a doublet of Dirac fermions is completely equivalent to two doublets of left-handed Weyl fermions. Equation (1.4) thus implies

$$\int \mathcal{D}[\psi_L]\mathcal{D}[\overline{\psi}_L]\,\mathrm{e}^{-S_{F,L}} \;=\; \sqrt{\det D(A_\mu)}\,. \tag{1.9}$$

It is important to realize that the sign of the square root is undetermined. Fixing the sign for all gauge fields is therefore part of the complete definition of the chiral gauge theory. However, we cannot choose arbitrarily the sign by hand. We need to be careful not to spoil some fundamental property any proper gauge theory should have. In particular we want to ensure the following two properties of $\sqrt{\det D(A_\mu)}$:

a) gauge invariance,
b) smooth gauge field dependence. (1.10)

The requirement of gauge invariance is obvious. The second property, which seems to be rather technical, is necessary in deriving the field equations of the theory when one needs to differentiate $\sqrt{\det D(A_\mu)}$ with respect to the gauge field.[1] As has been shown by Witten, both conditions cannot be satisfied simultaneously. This result is known as *Witten's SU(2) anomaly.*

To sketch Witten's original proof we first have a closer look at the square root of the fermion determinant. Suppose the Dirac operator is anti-hermitian.

[1] If we relax condition b) and allow for discontinuities, the field equations will in general contain non-local terms. These terms indicate a non-local gauge theory.

Without a mass term it also anti-commutes with γ_5. In that case its eigenvalues are purely imaginary and come in complex conjugate pairs,

$$\sigma(D) = \{i\lambda_1, -i\lambda_1, i\lambda_2, -i\lambda_2, \ldots\}\,, \quad \lambda_j \geq 0\,. \tag{1.11}$$

Witten considers the one-point compactification of space-time such that the eigenvalues of D are discrete. The fermion determinant is just the product of all λ_j^2 and for the square root we find

$$\sqrt{\det D(A_\mu)} = \pm \prod_j \lambda_j(A_\mu)\,. \tag{1.12}$$

Before we can fix the sign ambiguity we need to discuss a special property of SU(2) gauge transformations. The group $\mathcal{G}$ of all gauge transformations is the set of all mappings from space-time into SU(2). Since we consider space-time to be a sphere, not all gauge transformations are homotopic to each other. In fact $\mathcal{G}$ falls into two disjoint equivalence classes, $\mathcal{G} = \{1\} \cup \{g\}$. All mappings in $\{1\}$ are homotopic to the trivial gauge transformation 1, while all elements of $\{g\}$ are not. This topological result is mathematically expressed by a non-trivial fourth homotopy group,

$$\pi_4(\mathrm{SU}(2)) = \mathbb{Z}_2\,. \tag{1.13}$$

As a consequence the gauge orbit of any gauge field A_μ falls into two disjoint parts. However, for $t \in [0,1]$ the path

$$\begin{aligned} A_\mu(t,x) &= tA_\mu(x) + (1-t)A_\mu^g(x)\,, & (1.14)\\ A_\mu^g(x) &= g(x)(A_\mu(x) + \partial_\mu)g(x)^{-1}\,, & (1.15) \end{aligned}$$

smoothly interpolates between A_μ and its gauge transform A_μ^g, even if the gauge transformation belongs to $\{g\}$. As a function of the gauge field (1.14) the eigenvalues of D depend also on t. Witten proved that along any smooth path connecting A_μ and A_μ^g, $g \in \{g\}$, an odd number of eigenvalues $i\lambda_j$ cross zero and become $-i\lambda_j$. In short:

$$\lambda_{j_k}\Big|_{t=0} = -\lambda_{j_k}\Big|_{t=1}\,, \qquad k = 1,\ldots,n\,, \qquad n \text{ odd}\,. \tag{1.16}$$

The proof is based on the Atiyah-Singer index theorem for a certain five-dimensional Dirac operator. The fifth dimension plays essentially the rôle of the path parameter t. We do not repeat the proof here and refer to [1] for details.

Let us come back to the definition of the sign in (1.12). If we want to satisfy condition b) along the path (1.14), we can choose the sign only for one particular gauge field, at $t = 0$ say. Smooth gauge field dependence fixes the sign along the entire path. Witten's result (1.16) and (1.12) then imply

$$\sqrt{\det D(A_\mu)} = -\sqrt{\det D(A_\mu^g)}\,. \tag{1.17}$$

Hence the gauge invariance condition a) is violated.

On the other hand we can define the sign to be always non-negative, setting

$$\sqrt{\det D(A_\mu)} \;=\; \prod_j \mid \lambda_j(A_\mu) \mid \; . \tag{1.18}$$

In that way gauge invariance is achieved, but we sacrificed the smoothness condition b), since the function $|\lambda_{j_k}(t)|$ is not smooth for $\lambda_{j_k} = 0$. As we claimed the conditions (1.10) are incompatible. An SU(2) chiral gauge theory with a single doublet of Weyl fermions is therefore inconsistent.

It is important to emphasize that (1.16) is valid only if the gauge transformation g belongs to the homotopically non-trivial class $\{g\}$. If g is an element of $\{1\}$, the integer k in (1.16) is even. In that case the sign in (1.17) is absent and both conditions a) and b) hold. In particular the square root of the fermion determinant can be consistently defined if we consider infinitesimal gauge transformations only. For that reason Witten's anomaly is referred to as a *non-perturbative* or *global* anomaly.

Witten generalized his result for higher fermion representations. In terms of the diagonal SU(2) generator T_3 he found the theory to be inconsistent if $2\,\mathrm{tr}\,T_3^2$ is an odd integer. The trace of T_3^2 depends only on the highest weight j and one finds

$$2\,\mathrm{tr}\,T_3^2 \;=\; \frac{2}{3} j(j+1)(2j+1). \tag{1.19}$$

Obviously the right hand side of (1.19) is even if j is an integer, so these representations are anomaly free. For half-integer values one easily proves that $2\,\mathrm{tr}\,T_3^2$ is odd if

$$j \;=\; 2n + \frac{1}{2}, \qquad n \in \mathbb{N}_0 \, . \tag{1.20}$$

Note that not all representations are anomalous, even though (1.13) is independent of the representation. A non-trivial fourth homotopy group just causes a disjoint gauge orbit. To establish Witten's anomaly we also need the result (1.16) for the spectral flow along paths connecting disjoint parts of the gauge orbit. This second part only holds for the representations (1.20).

Nevertheless the topological result (1.13) is a necessary condition for Witten's anomaly and we expect the same obstruction for other gauge groups with a non-trivial π_4. However, among all compact simple Lie groups only the symplectic Lie groups Sp(2n) have a non-trivial fourth homotopy group [2],

$$\pi_4(\mathrm{Sp(2n)}) \;=\; \mathbb{Z}_2 \, , \quad n \in \mathbb{N}. \tag{1.21}$$

The only exceptions are SU(2) and SO(n), n=3,4,5, because of the group homomorphisms [3]

$$\begin{aligned}
\mathrm{SU}(2) &= \mathrm{Sp}(2)\, , \\
\mathrm{SO}(3) &= \mathrm{Sp}(2)/\mathbb{Z}_2\, , \\
\mathrm{SO}(4) &= \mathrm{Sp}(2)/\mathbb{Z}_2 \times \mathrm{Sp}(2)/\mathbb{Z}_2\, , \\
\mathrm{SO}(5) &= \mathrm{Sp}(4)/\mathbb{Z}_2\, .
\end{aligned}$$

The result (1.13) is just a special case of (1.21).

We finally remark that Witten's anomaly can be canceled by choosing an even number n_f of Weyl doublets. In that case the sign in (1.16) is replaced by $(-1)^{n_f}$ and gauge invariance is preserved. The Standard Model, based on the gauge group SU(3)×SU(2)$_L$×U(1)$_Y$, is therefore anomaly free, since n_f is equal to 4 for each generation.

1.2 The embedding technique

Witten's global SU(2) anomaly can also be derived in terms of the perturbative non-Abelian SU(3) anomaly.[2] In the functional integral formulation of quantum field theories this anomaly stems from the non-invariance of the fermion measure under gauge transformations. If g is a gauge transformation and if we write ψ_L^g and $\overline{\psi}_L^g$ for the gauge transformed left-handed fields, this non-invariance is formally written as

$$\mathcal{D}[\psi_L^g]\mathcal{D}[\overline{\psi}_L^g] \;=\; \mathcal{D}[\psi_L]\mathcal{D}[\overline{\psi}_L]\,\exp\,\Gamma(g, A_\mu)\,, \tag{1.22}$$

where the term $\Gamma(g, A_\mu)$ is purely imaginary. If the gauge transformation g is connected to the identity, i.e. if there exists a smooth path g_t, $0 \le t \le 1$, with $g_0 = 1$ and $g_1 = g$, the derivative of Γ with respect to t is found to be

$$\frac{\mathrm{d}}{\mathrm{d}t}\Gamma(g, A) = -\frac{1}{24\pi^2}\int \mathrm{d}^4x\,\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\left(\omega\,\partial_\mu\{A_\nu\partial_\rho A_\sigma + \frac{1}{2}A_\nu A_\rho A_\sigma\}\right), \tag{1.23}$$

with $\omega = g_t\partial_t g_t^{-1}$. We have suppressed the dependency on x and t on the right hand side for notational simplicity. The integrand in (1.23) is called the *perturbative non-Abelian anomaly* [5].[3]

As has been shown by various authors [7–9] the sign in (1.17), which is responsible for Witten's SU(2) anomaly, is just a special case of (1.22) and (1.9) with $\Gamma = \pi i \pmod{2\pi i}$.

To compute this value of Γ first note that SU(2) can be embedded in SU(3) which has a trivial fourth homotopy group π_4. Therefore every gauge transformation is connected to the identity in SU(3). In particular there exists a path $\Omega(t, x)$, $0 \le t \le 1$, with boundary values

$$\Omega(0, x) \;=\; \mathbb{1}\,, \qquad \Omega(1, x) \;=\; \left(\begin{array}{c|c} g(x) & \\ \hline & 1 \end{array}\right). \tag{1.24}$$

Here $g(x)$ denotes a homotopically non-trivial gauge transformation of SU(2). Note that $g(x)$ is a two-by-two matrix for all x and therefore fits into the upper left corner of the fundamental representation of SU(3) that is used in (1.24).

[2] For a thorough introduction to anomalies see the textbook by Bertlmann [4], for example.

[3] It is also called the consistent anomaly since it satisfies the Wess-Zumino consistency condition [6].

Similarly to the gauge transformations we embed SU(2) gauge potentials A_μ, which are su(2) valued fields, into the Lie algebra su(3) according to

$$\tilde{A}_\mu(x) = \left(\begin{array}{c|c} A_\mu(x) & \\ \hline & 0 \end{array}\right) . \tag{1.25}$$

Taking such a potential we define the path

$$\tilde{A}_\mu(t,x) = \Omega(t,x)\left(\tilde{A}_\mu(x) + \partial_\mu\right)\Omega(t,x)^{-1} . \tag{1.26}$$

In contrast to (1.14) this path basically connects $A_\mu(x)$ and $A_\mu^g(x)$ by a path of gauge equivalent fields, even if g is homotopically non-trivial. This is only possible because $\Omega(t,x)$ is a true SU(3) gauge transformation for $0 < t < 1$.

We now compute Γ by integrating (1.23) along the path (1.26). For simplicity we consider the case $\tilde{A}_\mu = 0$, i.e. the path connects the classical vacuum with the pure gauge configuration $\Omega\partial_\mu\Omega^{-1}$, which differs from $g\partial_\mu g^{-1}$ just by the dimension of the group representation. In that case we just need to replace A_μ by $\Omega\partial_\mu\Omega^{-1}$ etc. in (1.23) and find

$$\Gamma(g,0) = \frac{-1}{48\pi^2}\int_0^1 \mathrm{d}t \int \mathrm{d}^4x\, \epsilon_{\mu\nu\rho\sigma}\, \mathrm{tr}\, \left(\Omega\partial_t\Omega^{-1}\,\Omega\partial_\mu\Omega^{-1}\ldots\Omega\partial_\sigma\Omega^{-1}\right) . \tag{1.27}$$

To bring this integral into a more familiar form we define $x_4 = t$ and introduce a five dimensional tensor $\epsilon_{\mu\nu\rho\sigma\lambda}$ with $\epsilon_{01234} = 1$. This gives an extra factor 1/5 and we finally get

$$\Gamma(g,0) = \frac{-1}{240\pi^2}\int_{0\le x_4\le 1} \mathrm{d}^5x\, \epsilon_{\mu\nu\rho\sigma\lambda}\, \mathrm{tr}\, \left(\Omega\partial_\mu\Omega^{-1}\ldots\Omega\partial_\lambda\Omega^{-1}\right) . \tag{1.28}$$

This type of integrals is well known in the context of anomalies. Usually one considers paths $\Omega(t,x)$ with boundary values $\Omega(0,x) = \Omega(1,x) = \mathbb{1}$. In that case the integral is an integer multiple of $2\pi i$ [10]. In (1.28), however, the boundary values are given by (1.24). Even in this case the possible values are strongly restricted.

First note that the integral is invariant under small variations $\delta\Omega(t,x)$ as long as this variation preserves the particular form of the boundary values (1.24). Therefore, the integral depends only on the homotopy class of $g(x)$. If $g(x)$ is in the non-trivial homotopy class, $g^2(x)$ belongs to the trivial one. This already implies that the integral in (1.28) is either 0 or $i\pi$ (mod $2\pi i$).

Either by explicit computation, as performed by Witten [7], or by using the exactness of some homotopy sequence [8] one finds the value $i\pi$. Hence the phase factor in (1.22) is equal to -1 resulting again in the sign on the right hand side of (1.17).

Compare how the sign arises in both derivations. The fermion determinant is always real and passes through zero along the path (1.14) considered in Witten's original argument. Along the path (1.26) the fermion determinant picks up a phase factor that eventually sums up to -1.

Chapter 2

Exact chiral symmetry on the lattice

In the following we deal with the new developments concerning exact chiral symmetry on the lattice based on the Ginsparg-Wilson relation. We briefly repeat the Nielsen-Ninomiya theorem and discuss how the Ginsparg-Wilson relation enables us to circumvent it. Next we introduce Neuberger's Dirac operator as an explicit realization of an operator that satisfies the Ginsparg-Wilson relation. The spectral properties of this operator and an exact index theorem are also presented. Finally, the expansion of Neuberger's operator in orthogonal polynomials is discussed which we will need later on.

We start with setting our notations concerning lattice gauge field theory. For a introduction to the lattice approach to quantum field theory the reader is referred to the standard textbooks [11, 12].

2.1 Preliminary remarks

If not stated otherwise we consider a finite Euclidean lattice Γ of size L with periodic boundary conditions. The lattice spacing is denoted by a.

Lattice gauge fields assign to each lattice point x four link variables $U(x,\mu)$, where $\mu = 0,\ldots,3$ labels the lattice axes. The link variables are elements of some gauge group G, which we assume to be a compact connected Lie group.

A gauge transformation Λ is specified by $\Lambda(x) \in G$ and the link variables transform like

$$U(x,\mu) \quad\longrightarrow\quad \Lambda(x)\,U(x,\mu)\,\Lambda(x+a\hat{\mu})^{-1} \tag{2.1}$$

under Λ. Here $\hat{\mu}$ denotes the unit vector in direction μ.

Lattice Dirac fields $\psi(x)$ transform according to some unitary representation R of G,

$$\psi(x) \quad\longrightarrow\quad R[\Lambda(x)]\,\psi(x)\,. \tag{2.2}$$

Besides a Dirac index $\alpha = 1,\ldots,4$ fermion fields thus also carry a group index $a = 1,\ldots,\dim R$. For more than one flavour an additional flavour index appears too. Anti–fermion fields are denoted by $\overline{\psi}$ and transform differently under gauge transformations, namely

$$\overline{\psi}(x) \quad\longrightarrow\quad \overline{\psi}(x)\, R[\Lambda(x)]^{-1}\,. \tag{2.3}$$

The set $\mathcal{F}$ of all complex fermion fields forms a finite dimensional vector space and the canonical scalar product is defined by

$$(\psi,\phi) \quad=\quad a^4 \sum_x \sum_{\alpha,a} \psi^*_{\alpha,a}(x)\phi_{\alpha,a}(x)\,. \tag{2.4}$$

This scalar product also induces a norm for the fermion fields and subsequently an operator norm for a operators A,

$$\|A\| \quad=\quad \sup_{\psi\in\mathcal{F}} \frac{\|A\psi\|}{\|\psi\|}. \tag{2.5}$$

An operator trace is also defined via the scalar product by setting

$$\mathrm{Tr}\,A \quad=\quad \sum_j (v_j, Av_j)\,, \tag{2.6}$$

where the fields v_j are an orthonormal basis of fermion fields.

Derivatives of fields are approximated by differences on the lattice and the forward and backward difference operators are defined by

$$\partial_\mu\psi(x) \quad=\quad \frac{1}{a}\Big(\psi(x+a\hat{\mu}) - \psi(x)\Big)\,, \tag{2.7}$$

$$\partial^*_\mu\psi(x) \quad=\quad \frac{1}{a}\Big(\psi(x) - \psi(x-a\hat{\mu})\Big)\,. \tag{2.8}$$

A generalization of these operators are the gauge covariant forward and backward difference operators

$$\nabla_\mu\psi(x) \quad=\quad \frac{1}{a}\Big(R[U(x,\mu)]\psi(x+a\hat{\mu}) - \psi(x)\Big)\,, \tag{2.9}$$

$$\nabla^*_\mu\psi(x) \quad=\quad \frac{1}{a}\Big(\psi(x) - R[U(x-a\hat{\mu},\mu)]^{-1}\psi(x-a\hat{\mu})\Big)\,. \tag{2.10}$$

Under gauge transformations they transform gauge covariantly.

2.2 The Nielsen-Ninomiya theorem

Until recently it was widely believed that one cannot preserve chiral symmetry exactly on the lattice without spoiling some other fundamental principle like locality or allowing for extra physical particles, the so-called doubler states. This is basically the content of the famous Nielsen-Ninomiya theorem [13] which we briefly review.

Consider the free fermion action

$$S_F \quad=\quad a^4 \sum_x \overline{\psi}(x) D\psi(x)\,, \tag{2.11}$$

where D denotes the lattice Dirac operator. For brevity we suppress the summation over the Dirac and gauge group indices.

The action of D on fermion fields can be written as

$$D\psi(x) \;=\; a^4 \sum_x D(x,y)\psi(y)\,, \tag{2.12}$$

with $D(x,y)$ being the kernel of D. On an infinite lattice the Dirac operator should be translational invariant, $D(x,y) = D(x-y,0) \equiv D(x-y)$ and we write $\tilde{D}(p)$ for the Fourier transform of the kernel. It is desirable that D satisfies the following four conditions:

$$\begin{array}{ll} \text{a)} & D(x) \text{ is local}\,, \\ \text{b)} & \tilde{D}(p) \,=\, i\gamma_\mu p_\mu \,+\, \mathcal{O}(ap^2) \quad \text{for } p \ll \pi/a\,, \\ \text{c)} & \tilde{D}(p)^{-1} \quad \text{exists for all} \quad p \neq 0\,, \\ \text{d)} & D\gamma_5 + \gamma_5 D \,=\, 0\,. \end{array} \tag{2.13}$$

The locality property a) ensures the universality of the continuum limit. More precisely it requires the Dirac operator to satisfy the inequality

$$\|D(x)\| \;<\; Ce^{-d|x|}\,, \tag{2.14}$$

where $\|.\|$ denotes a matrix norm for the kernel $D(x)$. C and d are some positive numbers. Properties b) and c) guarantee the correct continuum limit without additional poles at non-zero momentum and d) leads to chiral invariance of the fermion action S_F. The Nielsen-Ninomiya theorem states that all four conditions cannot hold simultaneously.

An example for this result is the naive discretization of the continuum Dirac operator, where the derivative is simply replaced by a difference operator. This Dirac operator violates condition c) and describes 16 fermions instead of one in the continuum limit. On the other hand, Wilson proposed a Dirac operator that satisfies a) - c) but at the expense of chiral symmetry [14]:

$$D_w \;=\; \frac{1}{2}\sum_\mu \Big(\gamma_\mu(\partial_\mu^* + \partial_\mu) - a\partial_\mu^*\partial_\mu\Big). \tag{2.15}$$

Note that the symmetry breaking term $a\partial_\mu^*\partial_\mu/2$ is proportional to the lattice spacing. Property d) is therefore recovered in the classical continuum limit. In the presence of gauge fields the difference operators ∂_μ and ∂_μ^* are simple replaced by their covariant counterparts ∇_μ and ∇_μ^*.

2.3 Ginsparg-Wilson relation and chiral symmetry on the lattice

In 1982 Ginsparg and Wilson suggested [15] to replace condition d) by

$$D\gamma_5 + \gamma_5 D \;=\; aD\gamma_5 D\,. \tag{2.16}$$

Nowadays this condition is called *Ginsparg-Wilson relation*. The right hand side of it is proportional to a, so the naive anti-commutation law d) is obtained in the classical continuum limit.

The Ginsparg-Wilson relation did not attract that much attention in the beginning. It took more than 15 years until this relation was reconsidered, first of all in the context of perfect actions [16–18]. As has subsequently been shown in [19], the Ginsparg-Wilson relation implies an exact symmetry of the fermion action, which can be regarded as the lattice form of chiral symmetry.

To discuss this we consider infinitesimal transformations $\psi \to \psi + i\epsilon\delta\psi$ and $\overline{\psi} \to \overline{\psi} + i\epsilon\delta\overline{\psi}$ of the fermion fields. This transformation with the variations

$$\delta\psi = \gamma_5(1 - aD)\psi\,, \qquad \delta\overline{\psi} = \overline{\psi}\gamma_5\,, \tag{2.17}$$

leaves the fermion action invariant provided the Ginsparg-Wilson relation holds. The appearance of $(1-aD)$ in $\delta\psi$ originates in the right hand side of (2.16). For vanishing lattice spacing, however, the chiral variation $\delta\psi$ assumes its familiar form.

We mention two possible modifications [19] of (2.17). First of all the transformation can be cast into a symmetric form. Instead of including $(1 - aD)$ in $\delta\psi$ one can alternatively define both variations with a term $(1 - aD/2)$ without affecting the invariance of S_F. Furthermore, the variations can be generalized to flavour non-singlet chiral transformations by including the generator of the flavour symmetry group in (2.17).

It is well-known that the flavour-singlet chiral symmetry is anomalous in the presence of gauge fields. This behaviour is also found for the lattice chiral symmetry. As has been shown in [19], the axial anomaly stems from the non-invariance of the fermion measure under the transformation (2.17).

2.4 Neuberger's operator

Neuberger found an explicit realization of a Dirac operator that satisfies the Ginsparg-Wilson realization [20, 21]. In terms of the Wilson-Dirac operator (2.15) it reads

$$D = \frac{1}{a}\left(1 - A(A^\dagger A)^{-1/2}\right)\,, \qquad A = 1 - aD_w\,. \tag{2.18}$$

The validity of the Ginsparg-Wilson relation is easily checked by noting that

$$V = A(A^\dagger A)^{-1/2} \tag{2.19}$$

is unitary and in addition satisfies $V^\dagger = \gamma_5 V \gamma_5$.

It is, however, not obvious, that Neuberger's operator satisfies also the conditions a) - c) in (2.13). As has been proved in [19, 22],(2.18) has the correct continuum limit and is a local operator if $(A^\dagger A)$ has no zero mode.[1] This holds,

[1] Even if $A^\dagger A$ has an isolated zero mode the operator V can be defined and Neuberger's operator is still local [22].

for example, in the free case since then the lower bound $||A^\dagger A)|| \geq 1$ holds. In the presence of gauge fields one can strictly prove the inequality [22].

$$||A^\dagger A)|| \geq 1 - 30\epsilon\,, \tag{2.20}$$

if the gauge field satisfies

$$||1 - P(x,\mu,\nu)|| \;<\; \epsilon \tag{2.21}$$

for all plaquettes

$$P(x,\mu,\nu) \;=\; U(x,\mu)U(x+a\hat{\mu},\nu)U(x+a\hat{\nu},\mu)^{-1}U(x,\nu)^{-1}. \tag{2.22}$$

This implies that Neuberger's operator is completely acceptable for gauge fields satisfying (2.21) with $\epsilon < 1/30$. Gauge fields with this property are called *admissible* in the following. Even though one cannot prove rigorously the locality of Neuberger's operator for all gauge fields, a numerical study [22] indicates that this is the case, at least for moderate gauge couplings.

Another important property of Neuberger's operator is its differentiability with respect to the gauge field. As has also been shown in [22], this property holds for admissible gauge fields.

We make a final comment concerning the locality of (2.18). The Wilson-Dirac operator D_w contains nearest neighbor difference operators only. The kernel $D_w(x,y)$ therefore vanishes if $|x-y| > a$. Operators whose kernel vanish for $|x-y| > n\,a$ with some fixed integer n are called *ultra-local.*

On the other hand, Neuberger's operator is not ultra-local due to the inverse square root of $A^\dagger A$. As has been shown in [23–25], a Dirac operator satisfying the Ginsparg-Wilson relation cannot be ultra-local at the same time. Numerical simulations with Ginsparg-Wilson fermions are thus much more time consuming in comparison with the usage of Wilson fermions.

2.5 Spectral properties and the index theorem

Neuberger's operator is of the form $aD = 1 - V$ where V is unitary. It immediately follows that aD itself is a normal operator. Hence there exists an orthonormal basis of eigenvectors of aD for the space of all fermion fields.

We also conclude that the spectrum of aD lies on a unit circle around 1 in the complex plane and the eigenvalues are of the form

$$\lambda_j \;=\; 1 - e^{i\theta_j}. \tag{2.23}$$

On a finite lattice the spectrum is discrete and we can label the eigenvalues by an integer index j.

The unitary operator V has the additional property $V^\dagger = \gamma_5 V \gamma_5$. Hence for each eigenvector w_j with eigenvalue λ_j the vector $\tilde{w}_j = \gamma_5 w_j$ is an eigenvector of aD with eigenvalue λ_j^*. The eigenvalues thus come in complex conjugate pairs. The only possible real eigenvalues are zero and two.

The Ginsparg-Wilson relation implies that eigenvectors with real eigenvalues can be chosen to have definite chirality. Keeping this in mind and using the result $\mathrm{Tr}\,\gamma_5 = 0$ it is not difficult to prove the relation

$$-\frac{a}{2}\mathrm{Tr}\,\{\gamma_5 D\} \;=\; n_+ - n_- \,, \tag{2.24}$$

where n_+ and n_- denote the number of zero modes of D with positive and negative chirality, respectively. The integer on the right hand side is called the *index* of D. Obviously, it is invariant under continuous variations of the gauge field. The left hand side of (2.24) can alternatively be written as

$$-\frac{a}{2}\mathrm{Tr}\,\{\gamma_5 D\} \;=\; a^4 \sum_x q(x)\,, \qquad q(x) \;=\; -\frac{a}{2}\mathrm{tr}\,\{\gamma_5 D(x,x)\}\,. \tag{2.25}$$

As long as D is a local operator, the field $q(x)$ is also local. If we think of it as a topological charge density, the sum in (2.25) represents the topological charge itself and both results (2.24) and (2.25) establish an exact index theorem on the lattice [17].

For classical lattice gauge fields, originating from some smooth gauge potential $A_\mu(x)$ by the path-ordered exponential

$$U(x,\mu) = \mathcal{P}\exp\left\{a\int_0^1 \mathrm{d}t A_\mu(x+(1-t)a\hat{\mu})\right\}\,, \tag{2.26}$$

the topological charge density can be expanded in powers of the lattice spacing. Taking into account that $q(x)$ is gauge invariant and transforms as a pseudoscalar, one finds the expansion [26–28]

$$q(x) \;=\; \frac{1}{32\pi^2}\,\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,\{F_{\mu\nu}(x)F_{\rho\sigma}(x)\} \,+\, \mathcal{O}(a)\,, \tag{2.27}$$

with $F_{\mu\nu}(x)$ being the field tensor associated to the gauge potential $A_\mu(x)$. For a approaching zero we therefore recover the well-known topological charge density of continuum field theory.

2.6 Approximations of Neuberger's operator

If $A^\dagger A$ has no zero mode the square root of its inverse is well-defined. However, a power series representation of it is sometimes advantageous. In the following we first present the expansion in Legendre polynomials. Afterwards we argue that a Chebyshev approximation has its merits for approximations of Neuberger's operator.

2.6.1 Expansion in Legendre polynomials

Suppose all eigenvalues λ_i of $A^\dagger A$ are strictly positive. Then the spectrum of $A^\dagger A$ is contained in an interval $[a,b]$ with $a > 0$. Let P_i denote the projector onto the eigenspace of eigenvalue λ_i. We then find the spectral representation

$$\frac{1}{\sqrt{A^\dagger A}} \;=\; \sum_i \frac{1}{\sqrt{\lambda_i}}\,P_i \tag{2.28}$$

for the inverse square root of $A^\dagger A$. For x in $[a,b]$ the function $1/\sqrt{x}$ can be expanded in orthogonal polynomials. If we use this expansion for each $1/\sqrt{\lambda_i}$ in (2.28) we also obtain an expansion for the inverse square root of $A^\dagger A$. The use of Legendre polynomials is particularly simple [22]. For later reference we will briefly sketch the main steps and present the relevant formulae.

Any continuous function f defined on the closed interval $[-1,1]$ can be expanded in terms of Legendre polynomials $P_n(x)$ according to

$$f(x) \;=\; \sum_{n=0}^{\infty} a_n P_n(x)\,, \qquad a_n \;=\; \frac{2n+1}{2}\int_{-1}^{1} dx\, f(x) P_n(x)\,. \tag{2.29}$$

For functions f defined on more general intervals $[a,b]$ first note that the linear function

$$g(\alpha) \;=\; \frac{b+a}{2} - \frac{b-a}{2}\alpha \tag{2.30}$$

maps the interval $[-1,1]$ onto $[a,b]$. Therefore, the composite function $\tilde{f} = f\circ g$ can be expanded in terms of Legendre polynomials and for $f = \tilde{f}\circ g^{-1}$ we find

$$f(x) \;=\; \sum_{n=0}^{\infty} a_n P_n\left(g^{-1}(x)\right)\,, \qquad a_n \;=\; \frac{2n+1}{2}\int_{-1}^{1} dx\, f(g(x)) P_n(x)\,. \tag{2.31}$$

We are interested in the function $\sqrt{x}$, defined on $[a,b]$, $0<a<b$. The coefficients a_n are most easily found using the generating function of the Legendre polynomials [29],

$$\frac{1}{r\sqrt{1-2t\alpha+t^2}} \;=\; \sum_{n=0}^{\infty} \frac{t^n}{r} P_n(\alpha)\,. \tag{2.32}$$

If we adjust the parameters r and t to

$$r \;=\; \frac{1}{2}\left(\sqrt{b}+\sqrt{a}\right)\,, \qquad t \;=\; \frac{\sqrt{b}-\sqrt{a}}{\sqrt{b}+\sqrt{a}}\,, \tag{2.33}$$

the left hand side of (2.32) coincides with $1/\sqrt{g(\alpha)}$. Using the series (2.32) in in (2.31), the integrations are trivial because of the orthogonality property of the Legendre polynomials. The final result for the expansion reads

$$\frac{1}{\sqrt{x}} \;=\; \sum_{n=0}^{\infty} \frac{t^n}{r} P_n(\alpha(x))\,, \qquad \alpha(x) \;=\; \frac{b+a}{b-a} - \frac{2}{b-a}\,x\,. \tag{2.34}$$

Note that t depends only on the quotient a/b. For $a<b$ we thus find t to be smaller than 1. Consequently, since $|P_n(\alpha)| \le 1$, the expansion (2.34) is absolute convergent. Using (2.34) for $1/\sqrt{\lambda_i}$ in (2.28), we find the expansion

$$\frac{1}{\sqrt{A^\dagger A}} \;=\; \sum_{n=0}^{\infty} \frac{t^n}{r}\, P_n(\alpha(A^\dagger A)) \tag{2.35}$$

with α given in (2.34). This expansion is norm convergent since the bound $||P_n(\alpha(A^\dagger A))|| \le 1$ holds.

2.6.2 Alternation theorem and Chebyshev expansion

The expansion of $1/\sqrt{x}$ in terms of Legendre polynomials is just one example for an expansion in orthogonal polynomials. Legendre polynomials seem to be a natural choice, because the expression (2.32) for the generating function makes the computation of the expansion coefficients a_n trivial. But we can, of course, use other orthogonal polynomials $p_n(x)$, even though the integrals for the expansion coefficients cannot be solved analytically in general.

All expansions have in common that they result in an infinite series for the inverse square root of $A^\dagger A$. This is often good enough for theoretical studies. For instance, the discussion of the locality properties of Neuberger's operator in [22] made use of expansion (2.35).

For an implementation of Neuberger's operator on a computer, however, any infinite series must be truncated. This leads first of all to a truncation error which one must have under control. Secondly, an efficient finite approximation of Neuberger's operator depends on the type of orthogonal polynomials we use in the expansion, as we will discuss in the following.

The truncation of any series yields an approximation D_N of Neuberger's operator, where D_N and the corresponding remainder R_N are defined as

$$
\begin{aligned}
D &= D_N + R_N, && (2.36)\\
D_N &= 1 - A\sum_{n=0}^{N} a_n p_n(\alpha(A^\dagger A)), && (2.37)\\
R_N &= A\sum_{n=N+1}^{\infty} a_n p_n(\alpha(A^\dagger A)). && (2.38)
\end{aligned}
$$

So D_N contains polynomials up to degree N only. A measure of quality for the approximation is $\|R_N\|$ and we easily find the inequality

$$\|R_N\| \le \max_{x\in[a,b]} \left|1-\sqrt{x}\,s_N(x)\right| , \qquad s_N(x) = \sum_{n=0}^{N} a_n p_n(\alpha(x)) . \quad (2.39)$$

Here the norm for the remainder is completely given in terms of the finite polynomial approximation $s_N(x)$ of the function $1/\sqrt{x}$. In principle we can make $\|R_N\|$ as small as we like by choosing N sufficiently large. For *fixed* N, however, the size of $\|R_N\|$ depends considerably on the type of orthogonal polynomials we use for the expansion. It is therefore desirable to know which polynomials lead to the smallest value of $\|R_N\|$ for fixed N. To answer this question, we first state a simple form of the

Alternation theorem: Let f, ρ be continuous functions on $[a, b]$ with ρ strictly positive. $p_N(x)$ denotes the least maximum approximation of f, i.e. among all polynomials of degree less or equal N the polynomial p_N minimizes

$$\|f-\rho p_N\|_\infty = \max_{x\in[a,b]} |f(x)-\rho(x)p_N(x)| .$$

Suppose there exists a polynomial q_N of degree N such that $\rho(x)q_N(x)$ has $N+2$ points of alternation, i.e. there exist $N+2$ values x_i, $a \leq x_0 < x_1 < \cdots < x_{N+1} \leq b$, such that

$$f(x_i) - \rho(x_i)q_N(x_i) \;=\; \rho(x_{i+1})q_N(x_{i+1}) - f(x_{i+1}) \;=\; \pm\|f - \rho q_N\|_\infty \,.$$

That is the maximum discrepancy between f and ρq_N occurs with alternating sign at $N+2$ successive points. Then $p_N = q_N$, i.e. q_N is the least maximum approximation of f.

The proof of this theorem is given in appendix A. Usually the Alternation Theorem is formulated with $\rho(x) = 1$ and traces back to Chebyshev. It then states that the polynomial p_N itself has $N+2$ points of alternation.

If we set $f(x) = 1$ and $\rho(x) = \sqrt{x}$, the Alternation Theorem characterizes the polynomial approximation $p_N(x)$ of $1/\sqrt{x}$ that minimizes the right hand side of the inequality in (2.39). Expansions with different types of orthogonal polynomials can now be compared by simply checking if the function $\sqrt{x}s_N(x)$ has – at least approximately – $N+2$ points of alternation.

It turns out that an expansion in Chebyshev polynomials is almost optimal in this sense. As an illustration fig. 2.1 shows the function $\sqrt{x}s_N(x)$, $N = 16$, where $s_N(x)$ denotes the sum

$$s_N(x) \;=\; \sum_{n=0}^{N} a_n\, T_n(\alpha(x))\,, \tag{2.40}$$

with T_n being the Chebyshev polynomial of degree n. The definition of Chebyshev polynomials and all relevant formulae for the Chebyshev expansion are gathered in appendix B. The expansion coefficients a_n were numerically computed for $a = 0.2$ and $b = 49$ using MATHEMATICA. The value for b is motivated by the exact upper bound [22]

$$\|A^\dagger A\| \;\leq\; 49\,.$$

In the free case $A^\dagger A$ is also bounded from below with $1 \leq \|A^\dagger A\|$. For non-trivial gauge fields this is not the case and we have here randomly chosen $a = 0.2$.

As we see in fig. 2.1 the function $\sqrt{x}s_{16}(x)$ has to a very good approximation 18 points of alternation. The finite Chebyshev approximation (2.40) consequently almost minimizes the bound (2.39) for the remainder $\|R_N\|$.

For comparison fig. 2.2 shows the same function but this time Legendre polynomials were used in $s_{16}(x)$. The desired property of having 18 points of alternation is obviously not present. And in fact the resulting bound for $\|R_N\|$ is much weaker by a factor of almost 2 compared to the bound obtained with Chebyshev polynomials.

Even if we present here the results for $N = 16$ and $a = 0.2$ only, their qualitative features remain valid for various values N and a. We conclude that a truncated Chebyshev expansion results almost in the least maximum approximation of $1/\sqrt{x}$ and thus almost minimizes the truncation error $\|R_N\|$ of Neuberger's operator for fixed N.

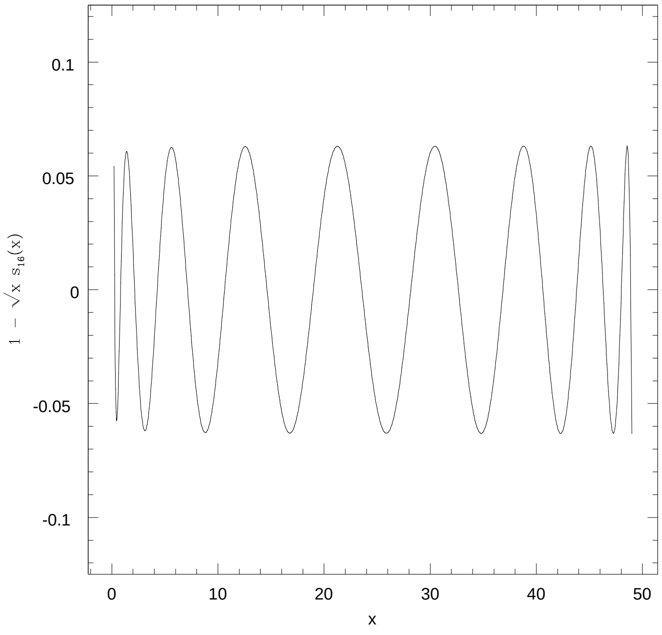

Figure 2.1: The function $1 - \sqrt{x}s_{16}(x)$, $0.2 \leq x \leq 49$, for a truncated expansion of $1/\sqrt{x}$ in Chebyshev polynomials.

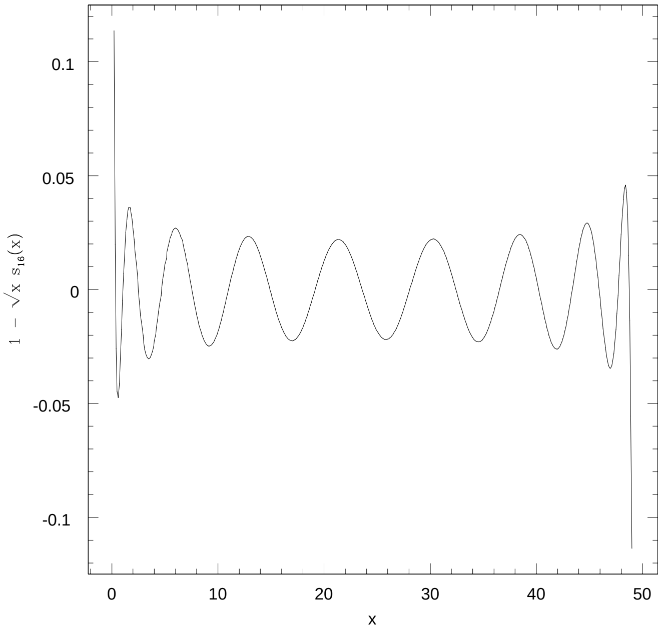

Figure 2.2: The function $1 - \sqrt{x}s_{16}(x)$, $0.2 \leq x \leq 49$, for a truncated expansion of $1/\sqrt{x}$ in Legendre polynomials.

Chapter 3

Chiral gauge theories on the lattice

Having exact chiral symmetry on the lattice we now address the problem of constructing chiral lattice gauge theories. As we will show, it is fairly easy to introduce Weyl fermions but the definition of a fermion measure for the functional integral is non-trivial. A general strategy to set up a proper fermion measure has been presented in [30] which we describe in detail. We also discuss the cancellation of perturbative gauge anomalies and explain how global anomalies emerge very naturally in this approach.

3.1 Lattice gauge fields

In order to have exact chiral symmetry we want to use Neuberger's operator for the lattice Dirac operator. As discussed in the last chapter, the locality and the differentiability of Neuberger's operator is rigorously guaranteed for the admissible gauge fields only. For that reason we will define a gauge field action that restricts the functional integral to the set of admissible fields. In that way we ensure locality and differentiability of Neuberger's operator for all physically relevant gauge fields.

For the admissible gauge fields we define the gauge action as

$$S_G = a^4 \frac{1}{g_0^2} \sum_x \sum_{\mu,\nu} \mathrm{tr}\,(1 - P(x,\mu,\nu))\; f(P), \tag{3.1}$$

$$f(P) = \left(1 - \|1 - P(x,\mu,\nu)\|^2/\epsilon^2\right)^{-1}. \tag{3.2}$$

Apart from the factor $f(P)$ this is just Wilson's plaquette action [14] with coupling constant g_0. The function $f(P)$ essentially causes the Boltzmann factor e^{-S_G} to vanish smoothly if the gauge field approaches ϵ in (2.21). Recall that ϵ must not exceed 1/30. Note furthermore that the action is gauge invariant.

The subset of admissible gauge fields and the gauge action define completely a lattice gauge theory at the classical level. To set up the corresponding quantum theory we need a functional integral for the computation of correlation functions. Let us first define the Boltzmann factor to be identically zero for all but the admissible gauge fields, i.e.

$$e^{-S_G} = 0 \qquad \text{if} \qquad \|1 - P(x,\mu,\nu)\| > \epsilon. \tag{3.3}$$

Then we can define the functional integral for arbitrary finite products $\mathcal{O}$ of link variables as usual as

$$\langle \mathcal{O} \rangle_G = \int \mathcal{D}[U]\, \mathcal{O} e^{-S_G}, \qquad \mathcal{D}[U] = \prod_x \prod_\mu \mathrm{d}U(x,\mu). \tag{3.4}$$

The measure $\mathrm{d}U$ denotes the Haar measure for the gauge group G. The integration region extends over the admissible gauge fields only because of (3.3). Note that the integrand $\mathcal{O}e^{-S_G}$ is a smooth function of the gauge fields due to the differentiability of $f(P)$.

Since the bound (2.21) constrains the gauge field fluctuations at the scale of the cutoff only, the modified action (3.1) should lie in the same universality class as the usual plaquette action, at least in the continuum limit in the weak coupling phase [30].

3.2 Fermion fields

From now on we will always assume D to be Neuberger's operator. This enables us to introduce Weyl fermions in the following way.

We already remarked that $V = 1 - aD$ is unitary with the additional property $V^\dagger = \gamma_5 V \gamma_5$. As an immediate consequence the operator $\hat{\gamma}_5 = \gamma_5 V$ satisfies

$$\hat{\gamma}_5^\dagger = \hat{\gamma}_5, \qquad \hat{\gamma}_5^2 = 1. \tag{3.5}$$

Therefore we can define the projection operators

$$\hat{P}_\pm = \frac{1}{2}\left(1 \pm \hat{\gamma}_5\right). \tag{3.6}$$

We can split every Dirac fermion field ψ in two orthogonal components $\hat{P}_-\psi$ and $\hat{P}_+\psi$ which we call the left-handed and right-handed components of ψ. If the component $\hat{P}_+\psi$ ($\hat{P}_-\psi$) vanishes the field is said to be left-handed (right-handed). The reason for this designation will become clear in a moment.

For the definition of chiral anti-fermion fields $\bar{\psi}$ we use the common projectors P_- and P_+ defined in (1.5). Any field $\overline{\psi}$ can be decomposed according to $\overline{\psi} = \bar{\psi}P_- + \overline{\psi}P_+$. The appearance of the projectors on the right hand side of the fields means that they act in spinor space from the right. Left-handed fields satisfy $\overline{\psi}P_+ = \overline{\psi}$, right-handed ones $\overline{\psi}P_- = \overline{\psi}$.

The crucial point is that left-handed and right-handed fields decouple in the fermion action [26, 31]. The properties

$$P_- D = D\hat{P}_+, \qquad P_+ D = D\hat{P}_-, \tag{3.7}$$

of the projectors lead to $S_F = S_{F,L} + S_{F,R}$, where the left- and right-handed parts of the action are defined as

$$S_{F,L} = a^4 \sum_x \bar{\psi}(x) \left[P_- D \hat{P}_+ \psi\right](x), \tag{3.8}$$

$$S_{F,R} = a^4 \sum_x \bar{\psi}(x) \left[P_+ D \hat{P}_- \psi\right](x). \tag{3.9}$$

The decoupling allows us to define a classical chiral lattice gauge theory for left-handed fields only. By imposing the constraints

$$\hat{P}_- \psi = \psi , \tag{3.10}$$
$$\overline{\psi} P_+ = \overline{\psi} , \tag{3.11}$$

on the fermion fields we eliminate the right-handed components. These constraints are local and gauge invariant because of the locality and gauge covariance of Neuberger's operator. The action $S_{F,L}$ is also local and gauge invariant. Hence the left-handed fields and the action $S_{F,L}$ together with the gauge field action of the previous section define properly a classical chiral gauge theory.

3.3 Evolution operator Q_t and the twist $\mathcal{T}$

Before turning to the definition of the *quantum* field theory by specifying a fermionic functional integral, we want to discuss a new feature inherent in our approach. The projector $\hat{P}_-$ contains the Dirac operator D. The definition of left-handed fields therefore depends on the particular background gauge field. This has some geometrical implications which will be relevant later for the discussion of global anomalies.

A path Γ in configuration space is defined by specifying the gauge field $U_t(x, \mu)$ at each value of the path parameter $t \in [0, 1]$. We require the gauge field to be a smooth function of t and we use the short-hand notation

$$P_t = \hat{P}_-|_{U=U_t} \tag{3.12}$$

to refer to $\hat{P}_-$ along the path. Equation (3.10) implies that the subspace of left-handed fermion fields depends on t and rotates according to the evolution of P_t.

We introduce the operator Q_t to be the solution of the first order linear differential equation

$$\partial_t Q_t = [\partial_t P_t, P_t] Q_t \ , \qquad Q_0 = 1\!\!1. \tag{3.13}$$

The hermiticity of P_t implies that Q_t is unitary. Furthermore it has the property

$$P_t Q_t \;=\; Q_t P_0, \tag{3.14}$$

which is easily proved by showing that $P_t Q_t$ and $Q_t P_0$ satisfy the same differential equation. Equation (3.14) tells us that if ψ_0 is a left-handed field at $t = 0$, at a different point t we have $P_t\, Q_t\, \psi_0 = Q_t\, \psi_0$. This means that $Q_t\, \psi_0$ is a left-handed field at point t. We therefore call Q_t the *evolution operator* of P_t along a path.

The definition of Q_t implicitly assumes that the gauge field $U_t(x, \mu)$ is differentiable for all t. However, it can be extended straightforwardly to piecewise smooth paths. In that case one defines Q_t^k for each individual smooth piece Γ^k of the path. The operator Q_t along the whole path is then the ordered product

$$Q_t = \prod_{k=1}^{n} Q_t^k \ . \tag{3.15}$$

This definition is consistent, that is whenever we split a path in more pieces than necessary we will obtain the same result for Q_t. The unitarity of Q_t follows from the unitarity of the individual Q_t^k.

Consider now closed loops in configuration space, that is $U_0(x,\mu) = U_1(x,\mu)$. After a whole turn the subspace of left-handed fields has rotated following the gauge field geometry along the loop. A priori there is no reason to expect the operator Q_1 to be the identity mapping. More generally, Q_1 defines a unitary mapping from the space of left-handed fermion fields at $t = 0$ onto itself. The determinant of that mapping serves us as a measure for the rotation along the loop. We therefore introduce the *twist* $\mathcal{T}$ for any closed loop Γ in configuration space, defined by

$$\mathcal{T}(\Gamma) = \det\{1 - P_0 + P_0 Q_1\} . \tag{3.16}$$

This definition contains the operator $(1 - P_0)$, representing the identity in the space of right-handed fields at $t = 0$. The unitarity of Q_t implies that $\mathcal{T}$ is a phase for all gauge groups. If Q_1 is the identity, we get trivially a twist equal to one. The twist for two closed loops with the same base point satisfies the composition law

$$\mathcal{T}(\Gamma_1 \circ \Gamma_2) = \mathcal{T}(\Gamma_1) \cdot \mathcal{T}(\Gamma_2) . \tag{3.17}$$

In the following the twist is called *non-trivial* if it is different from one. Loops with non-trivial twist will play an important rôle for the discussion of global anomalies later on.

3.4 Fermion integration measure

To set up a *quantum* field theory we need to define a functional integral in order to compute correlation functions. The definition of the fermion integration measure is a difficult task in the present approach to chiral lattice gauge theories, because the projector $\hat{P}_-$ depends on the gauge field.

3.4.1 Phase ambiguity of the fermion measure

Consider some fixed gauge field and choose an orthonormal basis of left-handed fields, i.e. a set of fields v_j, $j = 1, 2 \ldots$, that satisfy

$$\hat{P}_- v_j = v_j, \qquad (v_i, v_j) = \delta_{ij}. \tag{3.18}$$

Any left-handed field ψ can be expanded according to

$$\psi(x) \;=\; \sum_j v_j(x) c_j. \tag{3.19}$$

For classical fields the coefficients c_j are complex numbers. In a quantum field theory they are regarded as Grassmann variables such that two quantum fields ψ_1 and ψ_2 anti commute. A fermion integration measure is now given by

$$\mathcal{D}[\psi] \;=\; \prod_j \mathrm{d}c_j, \tag{3.20}$$

leading to a fermionic functional integral if we follow the usual rule

$$\int \mathrm{d}c_j \, c_k \quad = \quad \delta_{jk} \tag{3.21}$$

and its obvious generalization to multiple integrals.

Different bases for the fermion fields may lead to different fermion integration measures. To see this consider a second orthonormal basis v'_j that is related to v_j by some unitary basis transformation S according to

$$v'_j(x) \quad = \quad \sum_l v_l(x) S_{lj} \, . \tag{3.22}$$

Expanding the field ψ in this basis with coefficients c'_j, the primed fermion integration measure is related to the unprimed one by

$$\prod_j \mathrm{d}c'_j \quad = \quad \det S \prod_j \mathrm{d}c_j . \tag{3.23}$$

The two measures may differ by a phase because S is unitary. Only if S is unimodular we obtain the same measure. For that reason two bases are called equivalent if they differ by a unimodular basis transformation. Then the fermion measure depends on the equivalence class of bases only and each basis is an explicit representative of the associated measure via the rule (3.20).

The difference of two measures is of no importance if the phase ambiguity does not depend on the gauge field. In that case the phase can be taken out of the entire functional integral and cancels in normalized correlation functions. Here the situation is different. As we have seen, the subspace of left handed fermion fields depends on the gauge field and so does the corresponding fermion measure. Therefore the phase ambiguity is also gauge field dependent and the phase does not drop out in normalized correlation functions. This makes the definition of the fermion integration measure rather non-trivial.

Whether a proper fermion measure exists and how it can be consistently defined will be answered in the next subsection. But before we continue with the definition of the fermion measure for the anti-fermion fields.

The subspace of left-handed anti-fermion fields is gauge field independent. If we choose a basis $\overline{v}_k$ the expansion of anti-fermion fields and the corresponding integration measure read

$$\overline{\psi}(x) = \sum_k \overline{c}_k \overline{v}_k(x), \quad \mathcal{D}[\overline{\psi}] = \prod_k \mathrm{d}\overline{c}_k . \tag{3.24}$$

Here the ambiguity is a constant phase factor only.

Once the measures for the fermion fields are defined, fermion expectation values of any product $\mathcal{O}$ of the fermion fields are given as

$$\langle \mathcal{O} \rangle_F \quad = \quad \int \mathcal{D}[\psi] \mathcal{D}[\overline{\psi}] \mathcal{O} e^{-S_F} . \tag{3.25}$$

In particular, for the fermion determinant in the vacuum sector one finds

$$\langle 1 \rangle_F = \det M, \qquad M_{kj} = a^4 \sum_x \overline{v}_k(x) \, D v_j(x) . \tag{3.26}$$

Fully normalized correlation functions of the theory are finally defined by

$$\langle \mathcal{O} \rangle = \frac{1}{Z_0} \langle \langle \mathcal{O} \rangle_F \rangle_G \,, \qquad Z_0 = \langle \langle 1 \rangle_F \rangle_G \tag{3.27}$$

To complete the definition of the theory we need to fix the phase ambiguity of the fermion measure. However, this fixing should fulfill various conditions. For instance we insist on a smooth gauge field dependence of fermionic expectation values (3.25). In addition, the definition of the measure should neither destroy locality nor gauge invariance of the theory. In the next subsection we will formulate these and some other requirements more precisely.

3.4.2 Reconstruction Theorem

In order to discuss the gauge field dependence of the fermion measure we introduce a differential operator. Let $f(U)$ be a smooth function of the gauge field $U(x,\mu)$ and $\eta_\mu(x) = \eta_\mu^c(x) T^c$ a Lie algebra valued vector field. T^c denotes the generator of the gauge group. A differential operator δ_η is defined through

$$\delta_\eta f(U) = \frac{\partial}{\partial \epsilon} f(U_\epsilon) \bigg|_{\epsilon=0} \,, \qquad U_\epsilon(x,\mu) = U(x,\mu) + \epsilon\, a\, \eta_\mu(x) U(x,\mu) \,. \tag{3.28}$$

$\delta_\eta f(U)$ is called the *variation* of $f(U)$. Note that it is linear in η.

Now consider an admissible gauge field $U(x,\mu)$ in the vacuum sector. Since the operator $\hat{P}_-$ is differentiable with respect to the gauge field there exists a differentiable basis v_j of left-handed fermion fields. If we use this basis to define the fermion measure it will also be differentiable. Hence starting from (3.26) the variation of the effective action is well defined and given by [30]

$$\delta_\eta \ln \det M = \mathrm{Tr}\left\{ \delta_\eta D \hat{P}_- D^{-1} P_+ \right\} - i \mathfrak{L}_\eta, \tag{3.29}$$

$$\mathfrak{L}_\eta = i \sum_j (v_j, \delta_\eta v_j), \tag{3.30}$$

The trace term comes without surprise and a corresponding formal expression is also found in the continuum theory. The second term arises from the gauge field dependence of the basis vectors v_j and is called *measure term.*

Since the measure term is linear in η there exists a current $j_\mu^c(x)$ such that the measure term can alternatively be written as

$$\mathfrak{L}_\eta = a^4 \sum_x \eta_\mu^c(x) j_\mu^c(x). \tag{3.31}$$

This current is a well defined object emerging from any smooth left-handed basis. The current determines the change of phase of the fermion measure due to its gauge field dependence.

As we have seen, a fermion measure defines a current $j_\mu^c(x)$. We can ask the question, whether the opposite is also true. Does a current define in some way a fermion measure? The answer to this question has been given in [30].

Indeed, any current $j^c_\mu(x)$ defines a proper fermion measure provided it fulfills the following four conditions:

1. Smoothness condition: $j_\mu(x)$ is a smooth function of the gauge field.

2. Locality condition: $j_\mu(x)$ is a local expression in the gauge field.

3. Gauge invariance condition: The current is gauge covariant and satisfies the anomalous conservation law

$$[\nabla^*_\mu j_\mu]^c(x) = \mathcal{A}^c(x), \tag{3.32}$$

$$\mathcal{A}^c(x) = \frac{ia}{2}\mathrm{tr}\left\{\gamma_5 R(T^c) D(x,x)\right\}, \tag{3.33}$$

where $R(T^c)$ denotes the group generator for the representation R, $D(x,x)$ is the kernel of the Dirac operator and the expression $\mathcal{A}^c(x)$ represents the covariant anomaly on the lattice. The operator ∇^*_μ is the gauge covariant backward lattice derivative,

$$\nabla^*_\mu j_\mu(x) \;=\; \frac{1}{a}\left[j_\mu(x) - U(x-a\hat{\mu},\mu)^{-1} j_\mu(x-a\hat{\mu}) U(x-a\hat{\mu},\mu)\right]. \tag{3.34}$$

The formulation of the fourth condition requires one more definition. For any smooth path $U_t(x,\mu)$ we define the *Wilson line*

$$W = \exp\left\{ i\int_0^1 \mathrm{d}t\, \mathfrak{L}_\eta \right\}, \tag{3.35}$$

where the measure term $\mathfrak{L}_\eta$ is defined in (3.31) by the current and the field

$$a\eta_\mu(x) \;=\; \partial_t U_t(x,\mu) U_t(x,\mu)^{-1}. \tag{3.36}$$

The Wilson line will be the total change of phase of the reconstructed fermion measure along the path. We want to emphasize that here equation (3.31) serves as a definition of the measure term for a given current, not the other way around.

4. Integrability condition: $j_\mu(x)$ satisfies

$$W = \mathcal{T} \tag{3.37}$$

for all closed curves in configuration space. The twist $\mathcal{T}$ is defined in (3.16).

In addition we demand the current to transform as an axial vector under the lattice symmetries. This is no loss of generality because the four conditions are only sensitive to the axial vector part of the current. Once we have found a current compatible with the conditions, we can easily project out the axial vector part of it. This part will also satisfy the conditions.

Whenever a current meets all four conditions, a consistent fermion measure can be defined. Only an irrelevant constant phase ambiguity remains in the definition of the functional integral. This result goes under the name *Reconstruction Theorem* [32].

To define a fermion measure out of the current we first choose an arbitrary reference field U_{ref}. Any other field U in the same topological sector can be reached by a smooth curve U_t with $U_0 = U_{ref}$ and $U_1 = U$. If we choose an orthonormal basis ω_j of left-handed fermion fields at $t = 0$, a basis v_j at $t = 1$ is given by

$$\begin{aligned} v_1 &= Q_1\,\omega_1\,W^{-1}\,, \\ v_j &= Q_1\,\omega_j\,, \qquad\qquad j \neq 1\,. \end{aligned} \tag{3.38}$$

Recall that the Wilson line W and the measure term $\mathfrak{L}_\eta$ are defined in terms of the given current by (3.35) and (3.31) respectively. The reason why the first basis vector is multiplied by the phase W^{-1} will become clear later.

We can use basis (3.38) to define a fermion measure following the steps described in subsection 3.4.1. All we are left with is an unimportant constant phase ambiguity reflecting the ambiguity of the basis ω_j at $t = 0$.

So far we have just stated the conditions for the current and described how the current is used to define a particular basis and subsequently a fermion measure. Now we discuss the importance of the conditions for the consistency of the reconstructed measure.

If the first condition holds the basis (3.38) depends smoothly on the gauge field. Consequently, the reconstructed fermion measure and fermion expectation values $\langle \mathcal{O} \rangle_F$ also have a smooth gauge field dependence. This rather technical property is important in deriving the field equations of the theory, where one needs to differentiate $\langle \mathcal{O} \rangle_F$ with respect to the gauge field.

The second condition guarantees that the locality of the theory is preserved by the definition of the fermion measure. To see this consider the variation of the effective action (3.29). The first term on the right hand side generates the usual loop diagrams when expanded in powers of the gauge potential. The measure term acts as a counterterm. The resulting interaction vertices are local provided the current $j^c_\mu(x)$ is a local expression in the gauge field.

The discussion of the third and fourth condition needs more preparation, so we devote the next two subsections to them.

3.4.3 The gauge invariance condition

The requirement of gauge invariance of the fermion determinant leads directly to the gauge invariance condition. To work this out explicitly we consider an infinitesimal gauge transformation, generated by a Lie algebra valued lattice field $\omega(x)$. The corresponding variation of the link field $U(x, \mu)$ is obtained with

$$\eta_\mu(x) = -\nabla_\mu \omega(x)\,, \tag{3.39}$$

$$\nabla_\mu \omega(x) = \frac{1}{a}\left[U(x,\mu)\omega(x + a\hat{\mu})U(x,\mu)^{-1} - \omega(x)\right]\,. \tag{3.40}$$

The operator ∇_μ is nothing but the gauge covariant forward difference operator. Since the Dirac operator transforms covariantly, the variation of it is found to be

$$\delta_\eta D \;=\; [R(\omega), D]\,, \tag{3.41}$$

leading to [30]

$$\delta_\eta \ln\det M \;=\; ia^4 \sum_x \omega^c(x)\left\{\mathcal{A}^c(x) - [\nabla^*_\mu j_\mu]^c(x)\right\} \tag{3.42}$$

for the variation of the effective action (3.29).

Now the importance of the gauge invariance condition is obvious. The effective action is gauge invariant if and only if (3.32) holds. More generally the gauge invariance condition ensures the gauge invariance of fermionic expectation values (3.25) of any gauge invariant operator $\mathcal{O}$.

The gauge invariance condition has a simple consequence for the current. Since the variation of a gauge invariant quantity transforms covariantly, the measure term $\mathfrak{L}_\eta$ is gauge covariant according to (3.29). This implies the gauge covariance of $j_\mu(x)$ itself.

3.4.4 The integrability condition

Suppose we have chosen a reference field U_0 in the vacuum sector and a basis ω_j of left-handed fields. If we use basis (3.38) in expression (3.26) we find the formula [30]

$$\det M_1 \det M_0^\dagger \;=\; \det\left\{1 - P_+ + P_+ D_1 Q_1 D_0^\dagger\right\} W^{-1} \tag{3.43}$$

for the fermion determinant at any other gauge field $U = U_1$. The factor W^{-1} appears due to the particular definition of the first basis vector v_1 in (3.38).

The right hand side of equation (3.43) depends in general on the current $j_\mu(x)$ and on the path that is used to connect U_1 with the reference field U_0. If we consider a closed path with $U_1 = U_0$ we certainly demand the fermion determinant to be single valued, i.e. equation (3.43) should give the same value for $t = 0$ and $t = 1$.

Let us explicitly work out the consequences of this requirement. If the Dirac operator has a zero mode for $t = 0$ both sides of (3.43) are identically zero, independently of W. So let us assume the Dirac operator has no zero mode for $t = 0$. Identical values of (3.43) for $t = 0$ and $t = 1$ immediately imply

$$W \;=\; \det\left\{1 - P_+ + P_+ D_0 Q_1 D_0^{-1}\right\} \tag{3.44}$$

where we used the relation $D_1 = D_0$. According to the properties (3.7) and (3.14) the operators γ_5 and $D_0 Q_1 D_0^{-1}$ commute. Therefore the determinant of $D_0 Q_1 D_0^{-1}$ in the subspace of left-handed anti-fermion fields coincides with the right hand side of (3.44). This determinant can alternatively be written as a determinant in the subspace of left-handed fermion fields,

$$W \;=\; \det\left\{1 - P_0 + P_0 Q_1\right\}\,, \tag{3.45}$$

since the contribution of D_0 cancels. Hence the Wilson line must be equal to the twist along the loop. This is exactly the fourth condition for the current.

In conclusion the integrability condition for the current $j_\mu(x)$ ensures the single valuedness of the fermion determinant as well as of any other fermionic expectation value $\langle\mathcal{O}\rangle_F$ along closed loops in configuration space.

3.5 Anomaly cancellation

We want to show that any current which fulfills the first three conditions vanishes in the classical continuum limit. Subsequently we will find the same anomaly cancellation condition as in continuum field theory.

Let us consider classical fields (2.26). If the current j_μ^c is a local and smooth function of the gauge field, as required by the first two conditions, it can be expanded in powers of the lattice spacing,

$$j_\mu(x) \;=\; \sum_{k=0}^{\infty} a^{k-3}\mathcal{O}_\mu^k(x)\,. \tag{3.46}$$

In this asymptotic expansion the fields $\mathcal{O}_\mu^k(x)$ are local polynomials of dimension k in $A_\mu^c(x)R(T^c)$ and its derivatives. The gauge invariance condition implies that the current is gauge covariant, hence the fields $\mathcal{O}_\mu^k(x)$ are gauge covariant too.

Gauge covariant polynomials can only contain the field tensor and its derivatives and are thus at least of dimension 2. However, $F_{\mu\nu}^c(x)R(T^c)$ does not transform as a vector. Hence all terms $\mathcal{O}_\mu^k(x), k \leq 2$, must be equal to zero.

There exist three different fields of dimension 3:

$$(D_\mu F_{\rho\sigma})^c, \qquad (D_\nu F_{\mu\nu})^c, \qquad \epsilon_{\mu\nu\rho\sigma}(D_\nu F_{\rho\sigma})^c. \tag{3.47}$$

Here we have dropped the generator $R(T^c)$. The first and second do not transform as an axial vector. Although the third expression transforms correctly it vanishes identically. This is nothing but the Bianchi identity of a gauge theory. Thus we conclude that $\mathcal{O}_\mu^3(x)$ is zero. The expansion (3.46) therefore starts linearly in a and the current vanishes in the classical continuum limit.

Let us turn to the covariant anomaly $\mathcal{A}(x)$. According to (3.33) it is gauge covariant and transforms as a pseudoscalar. In the expansion

$$\mathcal{A}(x) \;=\; \sum_{k=0}^{\infty} a^{k-4}\mathcal{O}^k(x) \tag{3.48}$$

the fields $\mathcal{O}^k(x), k \leq 3$, vanish by arguments similar to the ones given above. In the end one finds the expansion starting with [30]

$$\mathcal{A}^c(x) \;=\; c_1 d_R^{abc}\epsilon_{\mu\nu\rho\sigma}F_{\mu\nu}^b(x)F_{\rho\sigma}^c(x) + \mathrm{O}(a), \tag{3.49}$$

$$d_R^{abc} \;=\; 2i\,\mathrm{tr}\left(R(T^a)\{R(T^b)R(T^c) + R(T^c)R(T^b)\}\right). \tag{3.50}$$

The coefficient $c_1 = -1/128\pi^2$ is independent of the gauge group and can be inferred from the index theorem in the U(1) case [17, 19, 32]. We see that $\mathcal{A}(x)$ converges to the covariant gauge anomaly in the continuum limit.

Since the current vanishes for $a = 0$, the gauge invariance condition (3.32) can only be satisfied if

$$d_R^{abc} \quad = \quad 0\,. \tag{3.51}$$

This is the usual anomaly cancellation condition that is well known from continuum gauge field theory. Note that one arrives at this result mainly by demanding the two fundamental principles locality and gauge invariance.

3.6 Global anomalies

It is by no means clear that for a given gauge group and a fixed lattice spacing an appropriate current $j_\mu(x)$ exists which satisfies all four conditions. So far the existence of a current has been rigorously proved only for the Abelian gauge group U(1) [32]. The proof for the non-Abelian case is still missing.

In the following we will argue that under certain circumstances a proper current satisfying all four conditions cannot exist. For instance consider a loop Γ_{class} of classical fields, defined by a smooth gauge potential $A_\mu(x,t)$ with $A_\mu(x,0) = A_\mu(x,1)$. As we have seen, any current compatible with the first three conditions vanishes in the classical continuum limit. This immediately implies

$$W = 1 + \mathcal{O}(a) \tag{3.52}$$

for the Wilson line along Γ_{class}. This can lead to a conflict with the integrability condition if the twist along this loop is non-trivial in the continuum limit. In that case, any current satisfying the first three conditions will necessarily violate the integrability condition for some small enough lattice spacing. Consequently, a proper fermion measure cannot be defined for small a.

This motivates us to the following definition. We speak of a *global anomaly* if any current satisfying the first three conditions necessarily violates the fourth one. In short:

$$\text{Global Anomaly} \quad \Longleftrightarrow \quad W \neq \mathcal{T}\,. \tag{3.53}$$

The attribute global has its origin in the formulation of the integrability condition in terms of full loops and not just in infinitesimal variations. As discussed, a global anomaly prevents the definition of a proper fermion measure and is therefore an insurmountable obstruction in formulating a chiral lattice gauge theory.

Chapter 4

Global anomaly in SU(2)

We now consider the gauge group SU(2) in the fundamental representation (highest weight $j = 1/2$). It will be shown that there exist loops in configuration space with a twist equal to -1. We find a different value for the Wilson line along these loops and thus prove the existence of a global anomaly.

4.1 Loops in configuration space

To begin with we take a homotopically non-trivial continuum gauge transformation $g(x)$, which is equal to the identity 1 for $|x| > R$ with some finite R. The particular form of $g(x)$ is not needed. Restricting space-time to the sites of an Euclidean lattice defines a lattice gauge transformation. We assume $L/2 > R + a$ for the lattice extension L, such that $g(x)$ is equal to 1 at the boundary of the finite lattice. Hence this lattice gauge transformation satisfies the periodic boundary conditions.

For $t \in [0, 1]$ we define two paths in configuration space:

$$\Gamma_1 : \qquad U_t(x, \mu) = g(x)^t g(x + a\mu)^{-t} \,, \tag{4.1}$$

$$\Gamma_2 : \qquad U_t(x, \mu) = [g(x) g(x + a\mu)^{-1}]^{1-t} \,. \tag{4.2}$$

Note that for compact Lie groups every element in the group can be written as an exponential of some Lie algebra element. In the case of SU(2) we can write every group element in terms of the Pauli matrices σ_j, $1, 2, 3$, and three real parameters ϕ_j as

$$h \;=\; \exp\, i \sum_{j=1}^{3} \phi_j \frac{\sigma_j}{2} \,, \qquad 0 \leq \left(\sum_{j=1}^{3} \phi_j^2 \right)^{1/2} < 2\pi \,. \tag{4.3}$$

The exponent on the right hand side can be multiplied by t and this defines h^t.

On the lattice all gauge transformations are continuously connected to the identity mapping. In particular, Γ_1 is a smooth curve of SU(2) gauge configurations. This path is a pure lattice artifact and has no analogue in the continuum. Γ_2 also connects the vacuum configuration and $g(x) g(x + a\mu)^{-1}$. It is just the lattice version of path (1.14) considered by Witten.

The loop $\Gamma_2 \circ \Gamma_1$ is closed. It starts and ends at the classical vacuum passing through the pure gauge configuration $g(x)g(x+a\mu)^{-1}$. Our aim is the computation of the twist along that loop. Due to the reality properties of SU(2) the twist is real and therefore equal to either +1 or −1. Since these values cannot change continuously the twist is a homotopy invariant for SU(2).

As an auxiliary tool for the calculation we define a third path on the lattice, using the path of gauge transformations $\Omega(t,x)$ defined in (1.24):

$$\Gamma_3: \quad U_t(x,\mu) = \Omega(t,x)\Omega(t,x+a\mu)^{-1}\,. \tag{4.4}$$

Γ_3 is also a path of gauge configurations, this time the link variables are in the fundamental representation of SU(3). This path connects the vacuum and the pure gauge configuration $g(x)g(x+a\mu)^{-1}$ in the $1/2 \oplus 0$ representation of SU(2). If we enlarge the first two paths also to this representation we can use the composition law (3.17),

$$\mathcal{T}(\Gamma_2 \circ \Gamma_1) \quad = \quad \mathcal{T}(\Gamma_2 \circ \Gamma_3) \cdot \mathcal{T}(-\Gamma_3 \circ \Gamma_1), \tag{4.5}$$

to compute the twist along $\Gamma_2 \circ \Gamma_1$. The additional singlet representation does not affect our final result. To see this first notice the relation

$$\mathcal{T}(\Gamma_{j_1 \oplus j_2}) \quad = \quad \mathcal{T}(\Gamma_{j_1}) \cdot \mathcal{T}(\Gamma_{j_2}) \tag{4.6}$$

that holds for a direct sum of two SU(2) representations. Here we used the short-hand notation Γ_j for a loop defined with the gauge field in the representation j. The twist is equal to 1 for the singlet representation. Therefore the left-hand side of (4.5) is also equal to the twist in the fundamental representation alone.

Two comments should be made. We want to emphasize that the path Γ_2 is well-defined in the space of admissible fields provided that the lattice spacing a is small enough. In the following we always assume this to be the case. The fields along Γ_1 and Γ_3 trivially satisfy the bound (2.21) because they are pure gauge configurations.

The paths Γ_i, $i = 1,2,3$, are smooth. However, the loops we have defined out of them are only piecewise smooth. At the contact points of the paths in the vacuum and the pure gauge configuration $g(x)g(x+a\mu)^{-1}$ the loops are not differentiable with respect to t.

4.2 Twist along closed gauge loops

A closed expression for the twist can be derived for pure gauge loops. The discussion of this subsection applies to any gauge group, only in the end we will restrict ourselves to our particular gauge loop in SU(3).

Given an initial gauge field $U_0(x,\mu)$, gauge paths in configuration space are of the form

$$U_t(x,\mu) = \Lambda_t(x)\, U_0(x,\mu)\, \Lambda_t(x+a\mu)^{-1}\,, \tag{4.7}$$

where Λ_t denotes a curve of gauge transformations with initial value $\Lambda_0 = 1$. Γ_1 and Γ_3 are special examples of (4.7) with $U_0 = 1$.

The projector to the subspace of left-handed fermion fields along gauge paths is thus given by

$$P_t = \Lambda_t P_0 \Lambda_t^{-1} \,. \tag{4.8}$$

For gauge paths the differential equation (3.13) for the evolution operator Q_t can be solved exactly. In terms of $X_t = \Lambda_t^{-1}\partial_t \Lambda_t$ the solution reads

$$Q_t = \Lambda_t \, \mathcal{P} \exp\left(- \int_0^t \mathrm{d}t' \, Y_{t'} \right) \,, \tag{4.9}$$

where we defined $Y_t = \{P_0 X_t P_0 + (1 - P_0) X_t (1 - P_0)\}$. Evidently this operator is block diagonal in the space of chiral fields, $Y_t = Y_{t,L} + Y_{t,R}$. For closed gauge loops ($\Lambda_1 = 1$) the operator Q_1 is just the path ordered exponential of the integral over Y_t and as usual one finds

$$\det Q_1 \; = \; \exp\left(- \int_0^1 \mathrm{d}t \; \mathrm{Tr}\, Y_t \right) \,. \tag{4.10}$$

The twist is just the determinant of Q_1 in the subspace of left-handed fermion fields at $t = 0$ and thus given as

$$\mathcal{T} \; = \; \exp\left(- \int_0^1 \mathrm{d}t \; \mathrm{Tr}\, P_0 Y_t \right) \,. \tag{4.11}$$

The projector restricts the trace to the space of left-handed fields only. With definition (3.33) for the covariant anomaly equation (4.11) can be rewritten as

$$\mathcal{T} = \exp\left(i a^4 \sum_x \mathcal{A}^c_{t=0}(x) \int_0^1 \mathrm{d}t \; X^c_t(x) \right) \,. \tag{4.12}$$

That is, along closed gauge loops the twist is proportional to the exponential of the covariant anomaly for the starting configuration at $t = 0$.

Our derivation of this result assumes the gauge loop to be differentiable for all t. This is not the case for $-\Gamma_3 \circ \Gamma_1$ we are interested in. However, eq. (4.9) can be applied independently to Γ_1 and Γ_3. Taking the product $Q(\Gamma_3)^{-1} \cdot Q(\Gamma_1)$ the factor Λ cancels and we finally arrive again at (4.12).

The loop $-\Gamma_3 \circ \Gamma_1$ starts in the vacuum configuration. In that case the covariant anomaly (3.33) for $t = 0$ is given as

$$\mathcal{A}^c_{t=0}(x) = \frac{ia}{2} \mathrm{tr}\,\{T^c\}\, \mathrm{tr}\,\{\gamma_5 D_{t=0}(x,x)\} \,, \tag{4.13}$$

with $D_{t=0}$ being the free Dirac operator. The group generators T^c of the fundamental representation of SU(3) are traceless and the anomaly thus vanishes. An immediate consequence is the result

$$\mathcal{T}(-\Gamma_3 \circ \Gamma_1) = 1 \tag{4.14}$$

for the twist along our closed gauge loop.

4.3 The twist along $\Gamma_2 \circ \Gamma_3$

We are now left with the calculation of the twist for the loop $\Gamma_2 \circ \Gamma_3$. Instead of solving the differential equation for Q_t we will first derive a general integral representation for the twist which we will subsequently use for the computation of $\mathcal{T}(\Gamma_2 \circ \Gamma_3)$.

Consider two loops Γ and $\tilde{\Gamma}$, defined by two gauge fields U_t and $\tilde{U}_t$. They are called to be *homotopic* if there exists a gauge field $U_{t,s}$, depending continuously on two parameters t and s, such that

$$\begin{array}{llll} 1) & U_{t,0} = U_t\,, \quad U_{t,1} = \tilde{U}_t\,, & 0 \leq t \leq 1\,, \\ 2) & U_{0,s} = U_{1,s} = U_{ref}\,, & 0 \leq s \leq 1\,. \end{array} \tag{4.15}$$

According to 1) both loops can be deformed into each other with s playing the rôle of the deformation parameter. The second property tells us that t parameterizes closed loops with constant base point U_{ref} for all values of s.

In the following we will simply write $\Gamma(s)$ instead of $\Gamma(U_{t,s})$ and call it the *homotopy* between Γ and $\tilde{\Gamma}$. The homotopy is said to be smooth if the dependency on s is smooth. Note that two loops need not to be homotopic in the space of admissible gauge fields. The bound (2.21) is responsible for a highly non-trivial topology of that space.

Suppose we are given a smooth homotopy $\Gamma(s)$. If the gauge field depends on two parameters, the same is true for the projector $\hat{P}_- = P_{t,s}$ and the evolution operator $Q = Q_{t,s}$. Differentiating the twist for the loop $\Gamma(s)$ we find

$$\partial_s \ln \mathcal{T}(\Gamma(s)) = \mathrm{Tr}\,(P_{0,s}\, Q_{1,s}^{-1}\, \partial_s Q_{1,s})\ . \tag{4.16}$$

Since $P_{0,s} = P_{1,s}$ the projector $P_{0,s}$ commutes with $Q_{1,s}$.

For the derivative of $Q_{t,s}$ with respect to s we can find an integral representation. Let us make the ansatz $\partial_s Q_{t,s} = Q_{t,s} R_{t,s}$ with some operator $R_{t,s}$. If we differentiate this equation with respect to t and the differential equation (3.13) with respect to s, we find

$$\partial_s Q_{t,s} = Q_{t,s} \int_0^t \mathrm{d}r\; Q_{r,s}^{-1} \partial_s X_{r,s}\, Q_{r,s}\ . \tag{4.17}$$

Here we defined $X_{r,s} = [\partial_r P_{r,s}, P_{r,s}]$. Setting t equal to 1 equation (4.16) can alternatively be written as

$$\partial_s \ln \mathcal{T}(\Gamma(s)) = \int_0^1 \mathrm{d}t\; \mathrm{Tr}\, P_{t,s}\, [\partial_t P_{t,s}, \partial_s P_{t,s}]\ . \tag{4.18}$$

If we integrate this result over s we finally find

$$\mathcal{T}(\Gamma(1)) = \mathcal{T}(\Gamma(0))\, \exp \int_0^1 \int_0^1 \mathrm{d}t\, \mathrm{d}s\; \mathrm{Tr}\, P_{t,s}\, [\partial_t P_{t,s}, \partial_s P_{t,s}]\ . \tag{4.19}$$

Keep in mind that this integral representation is valid provided the loop $\Gamma(1)$ can be smoothly deformed into $\Gamma(0)$. This equation essentially expresses the twist along $\Gamma(1)$ as an integral over a surface in configuration space defined by

$U_{t,s}$. Notice that the right-hand side is independent of the particular parameterization of that surface.

The integral in (4.19) looks not very accessible. However, the integrand can be expanded in powers of the lattice spacing. This is sufficient to compute the twist for small a.

To work this out explicitly we consider a homotopy of loops with classical fields. According to (2.26) the link field $U_{t,s}$ is given as the path-ordered exponential of a gauge potential $A_\mu(t,s,x)$, depending also on the parameters t and s. The expansion of the trace in (4.19) was already discussed in ref. [30]. Taking the locality and the symmetry properties of the integrand into account the leading term is found to be

$$\mathrm{Tr}\, P_{t,s}[\partial_t P_{t,s}, \partial_s P_{t,s}] = \frac{-i}{32\pi^2} \int \mathrm{d}^4x\, d_R^{abc} \epsilon_{\mu\nu\rho\sigma} \partial_t A_\mu^a(t,s,x) \partial_s A_\nu^b(t,s,x) F_{\rho\sigma}^c(t,s,x)\,. \tag{4.20}$$

Here $F_{\rho\sigma}^c(t,s,x)$ denotes the field tensor associated with the gauge potential. Using expression (4.20) together with (4.19) we are in principle able to compute the twist in the classical continuum limit.

Let us now turn to the loop $\Gamma_2 \circ \Gamma_3$ we are interested in. This loop is defined from an underlying loop of classical fields $A_\mu(t,x)$. Surely there exists a homotopy of classical fields $A_\mu(t,s,x)$ such that this loop can be contracted to the constant loop with $A_\mu = 0$. For small lattice spacings the corresponding lattice field $U_{t,s}(x,\mu)$ will satisfy the bound (2.21) for all t and s. Hence $\Gamma_2 \circ \Gamma_3$ is homotopic to the constant loop with $U_{ref} = 1$ in the subset of admissible fields and we can apply (4.19) and (4.20).

We consider a homotopy $\Gamma(s)$ with $\Gamma(1) = \Gamma_2 \circ \Gamma_3$. The path $\Gamma(0)$ is the constant loop such that the twist $\mathcal{T}(\Gamma(0))$ in (4.19) is equal to 1. All we need to apply (4.20) is a parameterization of the surface enclosed by $\Gamma_2 \circ \Gamma_3$ in the classical continuum limit. An example is given by

$$A_\mu(t,s,x) = (1-t)\Omega(s,x)\partial_\mu\Omega(s,x)^{-1}\,. \tag{4.21}$$

For $t = 0$ we recover the correct continuum limit of Γ_3, defined in (4.4). On the other hand, setting $s = 1$, equation (4.21) coincides with (1.14) for $A_\mu = 0$, the continuum limit of Γ_2. The parameterization (4.21) therefore has the correct boundary.

Note that the parameters t and s do not coincide with the ones given in (4.15). We instead make use of the parameterization invariance of the integral in (4.19) and define the surface in a particular simple way.

The surface parameterization (4.21) leads directly to the following relations, needed in (4.20):

$$\begin{aligned}
\partial_s A_\nu(t,s,x) &= (1-t)\left\{[\Omega\partial_\nu\Omega^{-1}, \Omega\partial_s\Omega^{-1}] + \partial_\nu(\Omega\partial_s\Omega^{-1})\right\}\,, \\
\partial_t A_\mu(t,s,x) &= -\Omega\partial_\mu\Omega^{-1}\,, \\
F_{\rho\sigma}(t,s,x) &= (t^2-t)\,[\Omega\partial_\rho\Omega^{-1}, \Omega\partial_\sigma\Omega^{-1}]\,.
\end{aligned} \tag{4.22}$$

For brevity we have suppressed the dependency on s and x of Ω.

The integration over t in (4.19) is easily performed and gives a factor 1/12. Making use of the anti-symmetry property of the epsilon tensor we finally find

$$\ln \mathcal{T}(\Gamma_2 \circ \Gamma_3) = \frac{-1}{48\pi^2} \int_0^1 \mathrm{d}s \int \mathrm{d}^4x \, \epsilon_{\mu\nu\rho\sigma} \, \mathrm{tr}\left(\Omega\partial_s\Omega^{-1}\Omega\partial_\mu\Omega^{-1}\cdots\Omega\partial_\sigma\Omega^{-1}\right) \quad (4.23)$$

for the leading term in the expansion in a. Note that we multiplied back the derivative of the potential and the field tensor into the trace over the group generators. Recall that they generate the fundamental representation of SU(3).

The right hand side of (4.23) is exactly the integral in equation (1.27) which we have encountered in section 1.2 in the context of Witten's anomaly. As discussed the value of this integral is $i\pi$ (mod $2\pi i$). Hence we obtain

$$\mathcal{T}(\Gamma_2 \circ \Gamma_3) = -1 \quad (4.24)$$

for the twist in the classical continuum limit.

4.4 Global anomaly in SU(2)

Taking into account the result (4.14) the composition law (4.5) reads

$$\mathcal{T}(\Gamma_2 \circ \Gamma_1) = \mathcal{T}(\Gamma_2 \circ \Gamma_3). \quad (4.25)$$

As we have already mentioned, the left hand side of this equation can only take the values ± 1. Since (4.25) is valid for all lattice spacings, it implies that (4.24) is not only the result in the continuum limit, but it also gives the twist for sufficiently small but finite lattice spacings a. The absence of terms proportional to a in (4.24) has its reason in our special loop which starts in the vacuum configuration. So our final result reads

$$\mathcal{T}(\Gamma_2 \circ \Gamma_1) = -1. \quad (4.26)$$

We want to emphasize that the introduction of Γ_3 is merely a technical tool. If we were able to compute directly the evolution operator Q_t along $\Gamma_2 \circ \Gamma_1$, we would have found the same twist using directly its definition (3.16).

We made use of the fact that $\Gamma_2 \circ \Gamma_1$ starts in the classical vacuum configuration. Our result (4.26), however, is not restricted to such loops. As we have previously remarked, the twist is a homotopy invariant in SU(2). Having found a loop with a twist equal to -1, all loops in the same homotopy class have the same twist.

To establish a global anomaly we have to show $W \neq \mathcal{T}$ along $\Gamma_2 \circ \Gamma_1$. The semi-classical argument in section 3.6 cannot directly be applied to this loop because Γ_1 has no continuum limit. However, the current is of order a along Γ_2. Since the path Γ_1 is a curve of gauge configurations and because the current is gauge-covariant we conclude that the current is of order a along the entire loop. Hence for the Wilson line along $\Gamma_2 \circ \Gamma_1$ we find

$$W(\Gamma_2 \circ \Gamma_1) \,=\, 1 + \mathcal{O}(a). \quad (4.27)$$

According to our discussion in section 3, the loop $\Gamma_2 \circ \Gamma_1$ thus violates the integrability condition for small enough lattice spacings and an SU(2) gauge theory exhibits a global anomaly.

The introduction of an SU(3) path Γ_3 for the computation of the twist has an obvious similarity with the calculation of Witten's anomaly in section 1.2. In both cases we are led to the five dimensional integral (1.27). This strongly indicates that the global obstruction we established here is the lattice analogue of Witten's SU(2) anomaly.

Chapter 5

Global anomaly in higher SU(2) representations

5.1 Preliminaries

So far we considered the fundamental representation of SU(2). For computational reasons we introduced an auxiliary path Γ_3 with the gauge field in the fundamental representation of SU(3) and embedded the links of the other paths, Γ_1, Γ_2, also in this representation.

In the following we compute the twist for higher representations of SU(2) along the same lines. However, the embedding turns out to be less trivial. Basically, for any given SU(2) representation j we cannot find an appropriate SU(3) representation R that will decompose into j and a finite number of singlets. More generally, we always end up with a decomposition

$$R \;\to\; \sum_j c(R,j)\, j\,, \tag{5.1}$$

where the coefficients $c(R,j)$ give the multiplicity of the representation j. To turn (5.1) the other way around, only particular sums of SU(2) representations can be embedded in an SU(3) representation R, such that a path of gauge transformations similar to (1.24) exists.

This in mind we start with the definition of Γ_3. In analogy to (4.4) we define $\Gamma_3(R)$ with a curve $\Omega_R(t,x)$, this time in the representation R of SU(3). The boundary value $\Omega_R(1,x)$ is still in SU(2), but now in the representation given by the right hand side of (5.1). Hence we define the paths Γ_1, Γ_2, for this direct sum of SU(2) representations.

Using (4.6) several times we find the generalization of the composition law (4.5) to be

$$\prod_j \mathcal{T}(j)^{c(R,j)} \;=\; \mathcal{T}(R)\,. \tag{5.2}$$

Notice that $\mathcal{T}(R)$ denotes the twist along $\Gamma_2 \circ \Gamma_3$ and R indicates the SU(3) representation used for Γ_3. On the other hand, $\mathcal{T}(j)$ stands for the twist along the pure SU(2) loop $\Gamma_2 \circ \Gamma_1$. In (5.2) we already used the fact that the twist along the pure gauge loop $-\Gamma_3 \circ \Gamma_1$ is 1 for all representations R, since the simple replacement of T^c by $R(T^c)$ in (4.13) still results in (4.14). Equation (5.2) is our master formula that will enable us to compute recursively $\mathcal{T}(j)$ for all SU(2) representations.

5.2 The twist $\mathcal{T}(R)$

Let us discuss the twist $\mathcal{T}(R)$ along $\Gamma_2 \circ \Gamma_3$ for arbitrary R. The integral in (4.23) contains the product of five generators of SU(3). Expressing twice the product of two of them as a commutator, we may write our result (4.23) as

$$\ln \mathcal{T} \quad = \quad d^{abc} I^{abc}, \tag{5.3}$$

where d^{abc} is the d-symbol of SU(3). The remaining part I^{abc} is independent of the group representation and the explicit expression of it is not needed in the following. For some other representation R the result (4.23) can be similarly written as

$$\ln \mathcal{T}(R) \quad = \quad d_R^{abc} I^{abc}. \tag{5.4}$$

It differs from (5.3) only in the d-symbol that is now defined with the generators $R(T^a)$ of the representation R. One can show that d_R^{abc} can always be expressed in terms of d^{abc} [33]. Introducing the *anomaly coefficient* $A(R)$ we write

$$d_R^{abc} \quad = \quad A(R)\, d^{abc}, \tag{5.5}$$

leading to the relation

$$\mathcal{T}(R) \quad = \quad \mathcal{T}^{A(R)} \tag{5.6}$$

for the twist along $\Gamma_2 \circ \Gamma_3$ for higher representations. Once the twist $\mathcal{T}$ for the fundamental representation is known, its value for higher representations is completely determined by the anomaly coefficient. The computation of $\mathcal{T}(R)$ is therefore reduced to the purely group theoretic task of calculating $A(R)$.

By definition $A(R)$ is equal to one for the fundamental representation. Furthermore one can straightforwardly establish the relations

$$\begin{aligned} A(R_1 \oplus R_2) &= A(R_1) + A(R_2), \\ A(R_1 \otimes R_2) &= A(R_1) d_2 + A(R_2) d_1, \end{aligned} \tag{5.7}$$

where d_i denotes the dimension of the representation R_i.

Let us consider a particular example. The fundamental representation of SU(3) is three dimensional and in the following denoted by 3. For the n-fold tensor product $3 \otimes \ldots \otimes 3 = 3 \otimes^n$ the anomaly coefficient and the twist are easily computed to be

$$A(3\otimes^n) \, = \, n 3^{n-1}, \qquad \mathcal{T}(3\otimes^n) \, = \, (-1)^n. \tag{5.8}$$

We will use this result in the next section to classify the anomalous SU(2) representations.

$$
\begin{aligned}
(\frac{1}{2}\oplus 0)\otimes^2 &= 2\cdot 0 \oplus 2\cdot\frac{1}{2}\oplus 1\cdot 1\\
(\frac{1}{2}\oplus 0)\otimes^3 &= 4\cdot 0 \oplus 5\cdot\frac{1}{2}\oplus 3\cdot 1\oplus 1\cdot\frac{3}{2}\\
(\frac{1}{2}\oplus 0)\otimes^4 &= 9\cdot 0 \oplus 12\cdot\frac{1}{2}\oplus 9\cdot 1\oplus 4\cdot\frac{3}{2}\oplus 1\cdot 2\\
(\frac{1}{2}\oplus 0)\otimes^5 &= 21\cdot 0 \oplus 30\cdot\frac{1}{2}\oplus 25\cdot 1\oplus 14\cdot\frac{3}{2}\oplus 5\cdot 2\oplus 1\cdot\frac{5}{2}
\end{aligned}
$$

Table 5.1: The decomposition of $(\frac{1}{2}\oplus 0)\otimes^n$ up to $n=5$.

5.3 Global anomaly in higher SU(2) representations

As already mentioned, the fundamental representation 3 of SU(3) decomposes into a doublet and a singlet of SU(2). Hence, for the n-fold tensor product $3\otimes^n$ the decomposition (5.1) reads

$$3\otimes^n \;\to\; (\frac{1}{2}\oplus 0)\otimes^n \;=\; \sum_j c(n,j)\, j\,. \tag{5.9}$$

Up to n equals 5 we have collected the explicit decompositions in table 5.1. Consider first $n=2$. Using (5.2) and (5.8) we find

$$\mathcal{T}(1) \;=\; 1\,, \tag{5.10}$$

because the anomalous doublet representation with $\mathcal{T}(\frac{1}{2})=-1$ appears twice in the decomposition. On the other hand, $c(3,\frac{1}{2})$ equals 5 and we find

$$\mathcal{T}(\frac{3}{2}) \;=\; 1\,. \tag{5.11}$$

Here the minus sign on the right hand side of (5.8) is compensated by an additional sign due to an odd number of doublet representations. Compare this with $n=5$. Here $c(5,\frac{1}{2})$ is even and, because one can check before $\mathcal{T}(2)=1$, we get

$$\mathcal{T}(\frac{5}{2}) \;=\; -1\,. \tag{5.12}$$

These explicit examples should be enough to illustrate the idea of our method. We are now in the position to prove

Lemma I: $\mathcal{T}(n)=1$ for $n\in\mathbb{N}_0$.

Proof: The statement is true for $n=0,1$. Now we assume $\mathcal{T}(k)=1$ for $k\le n$. Consider

$$(\frac{1}{2}\oplus 0)\otimes^{2n+2} \;=\; (1\oplus 2\cdot\frac{1}{2}\oplus 2\cdot 0)\otimes^{n+1} \;=\; 1\,\otimes^{n+1} \;+\; \sum_j \tilde{c}(2n+2,j)j\,. \tag{5.13}$$

Because both the doublet and the singlet representation appear twice in the decomposition of $(1/2\oplus 0)\otimes^2$, the coefficients $\tilde{c}(2n+2,j)$ are even for all values of j. The tensor product of the triplet representation can be decomposed further and we write

$$1 \otimes^{n+1} \;=\; (n+1) \oplus \sum_{k=0}^{n} a_k \, k \,, \tag{5.14}$$

where we introduced the not necessarily even integers a_k. In that decomposition only integer representations occur. Taking into account (5.8) equation (5.2) reads

$$\mathcal{T}(n+1) \prod_k \mathcal{T}(k)^{a_k} \prod_j \mathcal{T}(j)^{\tilde{c}(2n+2,j)} \;=\; 1 \,. \tag{5.15}$$

According to our assumption $\mathcal{T}(k) = 1$, this implies $\mathcal{T}(n+1) = 1$. □

The general case for half-integer representations is slightly more involved. For that reason we first compute explicitly $\mathcal{T}(\frac{7}{2})$ and $\mathcal{T}(\frac{9}{2})$ without having the decomposition as in table 5.1. This example already exhibits all features of the proof for arbitrary j.

For any representation j we define $\Sigma_a(j)$ to be the number of anomalous representations in the decomposition of $(\frac{1}{2} \oplus 0) \otimes^{2j}$ with highest weight less than j. For example,

$$\Sigma_a(\frac{5}{2}) \;=\; c(5, \frac{1}{2}) \,. \tag{5.16}$$

Similarly, we define $\Sigma_f(j)$ to be the number of anomaly free *half-integer* representations in the same decomposition. For the 5/2 representation this quantity is given by the multiplicity of the 3/2 representation,

$$\Sigma_f(\frac{5}{2}) \;=\; c(5, \frac{3}{2}) \,. \tag{5.17}$$

Equation (5.2) and (5.8) connect the twist $\mathcal{T}$ and Σ_a by the relation

$$\mathcal{T}(j) \cdot (-1)^{\Sigma_a(j)} \;=\; -1, \tag{5.18}$$

that is valid for half integer values of j.

Now we consider the tensor product $(\frac{1}{2} \oplus 0) \otimes^7$ that can be decomposed as

$$(1 \oplus 2 \cdot \frac{1}{2} \oplus 2 \cdot 0) \otimes \left((\frac{1}{2} \oplus 0) \otimes^5 \right) \;=\; 1 \otimes \left((\frac{1}{2} \oplus 0) \otimes^5 \right) \oplus \sum_j \tilde{c}(5,j) j \,, \tag{5.19}$$

where the coefficients $\tilde{c}$ stem from the tensor product with the doublet and singlet representation. The tensor product of the triplet representation can be further decomposed with the result

$$1 \otimes \left((\frac{1}{2} \oplus 0) \otimes^5 \right) \;=\; \sum_{k=1}^{4} b(7, \frac{2k-1}{2}) \frac{2k-1}{2} \oplus \text{integer representations} \,, \tag{5.20}$$

where the coefficients $b(n+2, j)$ are given in terms of $c(n, j)$ as

$$\begin{array}{lclclcl} b(7,\frac{1}{2}) & = & c(5,\frac{1}{2}) & + & c(5,\frac{3}{2}) & & , \\ b(7,\frac{3}{2}) & = & c(5,\frac{1}{2}) & + & c(5,\frac{3}{2}) & + & c(5,\frac{5}{2}) \,, \\ b(7,\frac{5}{2}) & = & & + & c(5,\frac{3}{2}) & + & c(5,\frac{5}{2}) \,, \\ b(7,\frac{7}{2}) & = & & & & & c(5,\frac{5}{2}) \,. \end{array} \tag{5.21}$$

Let us denote the contribution of (5.20) to $\Sigma_a(\frac{7}{2})$ by $\tilde{\Sigma}_a(\frac{7}{2})$. Because the 1/2 and the 5/2 representation are anomalous, it is given by the sum of the first and third row in (5.21). Taking into account $c(5,\frac{5}{2})=1$ we obtain

$$\tilde{\Sigma}_a(\frac{7}{2}) \quad = \quad 1+\Sigma_a(\frac{5}{2})+2\Sigma_f(\frac{5}{2})\,. \tag{5.22}$$

It follows that $\tilde{\Sigma}_a(\frac{7}{2})$ is odd because $\Sigma_a(\frac{5}{2})$ is even.

The contribution to $\Sigma_a(\frac{7}{2})$ of the second term on the right hand side of (5.19) amounts in adding various coefficients $\tilde{c}$ to $\tilde{\Sigma}_a(\frac{7}{2})$. Because all $\tilde{c}$ are even, $\Sigma_a(\frac{7}{2})$ is odd and thus

$$\mathcal{T}(\frac{7}{2}) \;=\; 1\,. \tag{5.23}$$

Having established this result we can proceed along the same lines for the $j=9/2$ representation. Starting from the tensor product $(\frac{1}{2}\oplus 0)\otimes^9$ we find equations corresponding to (5.19), (5.20) and (5.21), leading to

$$\tilde{\Sigma}_a(\frac{9}{2}) \quad = \quad 1+\Sigma_a(\frac{7}{2})+2\Sigma_f(\frac{7}{2})\,. \tag{5.24}$$

Following the argumentation given above we find $\Sigma_a(\frac{9}{2})$ to be even because $\Sigma_a(\frac{7}{2})$ is odd. This implies

$$\mathcal{T}(\frac{9}{2}) \;=\; -1\,. \tag{5.25}$$

It is not difficult to generalize this example and prove the general result

Lemma II: $\mathcal{T}(2n+\frac{1}{2})=-1$ and $\mathcal{T}(2n+\frac{3}{2})=1$ for $n\in\mathbb{N}_0$.

The proof goes along the lines we have described above. The key observation is the generalization of equation (5.22) and (5.24),

$$\tilde{\Sigma}_a(n+\frac{3}{2}) \;=\; 1+\Sigma_a(n+\frac{1}{2})+2\Sigma_f(n+\frac{1}{2})\,. \tag{5.26}$$

This formula implies that $\Sigma_a(2n+\frac{1}{2})$ is even and $\Sigma_a(2n+\frac{3}{2})$ is odd for all integers n. Together with (5.18) this gives the desired result. The complete proof can be found in appendix C.

Our argument that lead to (3.52) for the Wilson line W does not depend on the representation for the gauge group, so for all j we find

$$W(j) \quad = \quad 1+\mathcal{O}(a)\,. \tag{5.27}$$

Hence we find the representations with

$$j \;=\; 2n+\frac{1}{2}\,, \qquad n\in\mathbb{N}_0, \tag{5.28}$$

to be anomalous.

As we have shown in section 1.1 these are exactly the representations that satisfy Witten's anomaly condition, namely that $2\,\mathrm{tr}\,T_3^2$ is an odd integer. Hence we find a global anomaly for the same representations as Witten did in the continuum. Notice that this coincidence is highly non-trivial. Witten's condition emerges from the investigation of the zero modes of a five dimensional Dirac operator in a certain instanton field. This differs completely from the embedding technique we used here.

Chapter 6

Spectral flow of Neuberger's operator

In the previous two sections we proved the existence of a global SU(2) anomaly by introducing an auxiliary path Γ_3 in SU(3). We were led to a five-dimensional integral we were familiar with from the discussion of Witten's anomaly in section 1.2. This already indicates that the global anomaly we have discovered is nothing but Witten's SU(2) anomaly. There is, however, one piece missing, namely the connection between the anomaly and the spectral flow of the Dirac operator, which we discuss in this chapter.

6.1 A spectral flow theorem

Consider a smooth closed path U_t in field space and assume D_0 to have no zero modes. The function

$$f(t) = \det\left(1 - P_+ + P_+ D_t Q_t D_0^\dagger\right) \tag{6.1}$$

depends smoothly on t and is real for the gauge group $SU(2)$. In addition it satisfies

$$f(0) > 0\,, \qquad f(1) = \mathcal{T} f(0). \tag{6.2}$$

This implies that f passes through zero an odd number of times for $0 \le t \le 1$ if and only if $\mathcal{T} = -1$. Next one can show

$$f^2(t) = \det D_t \det D_0^\dagger. \tag{6.3}$$

Since Neuberger's operator satisfies $D^\dagger = \gamma_5 D \gamma_5$, the eigenvalues λ_i of D come in complex conjugate pairs and we find

$$\det D_t = \prod_i \lambda_i(t)\lambda_i^*(t)\,. \tag{6.4}$$

The eigenvalues depend smoothly on t for smooth paths. Equation (6.3) implies that if $f(t)$ passes through zero at some point t_0 the same happens for an odd number of eigenvalues $\lambda_i(t)$. This is most easily proved by expanding both $f(t)$ and $\lambda_i(t)$ around t_0 and using these expansions in (6.3). We conclude that an odd number of eigenvalues λ_i of D cross zero along a loop if and only if $\mathcal{T} = -1$.[1] In short we may write

$$\lambda_{j_k}\Big|_{t=0} = \lambda^*_{j_k}\Big|_{t=1}, \qquad k = 1, \ldots, n, \qquad n \text{ odd}. \tag{6.5}$$

This result is completely analogous to (1.16). Thus we find the same behaviour of the spectral flow on the lattice that Witten proved in the continuum using the Atiyah–Singer index theorem.

As has been shown in chapter 4 the twist along $\Gamma_2 \circ \Gamma_1$ is equal to -1. Because Γ_1 is a curve of gauge transformations the function f remains constant along that piece of the loop. The crossing of the eigenvalues occurs along Γ_2, the lattice analogue of Witten's path (1.14).

6.2 Numerical computation of the spectral flow

On a finite Euclidean lattice we are able to compute numerically the spectral flow along a loop U_t. Doing so we should find the result (6.5) for the eigenvalues along loops with $\mathcal{T} = -1$.

In the following we present numerical results for the spectral flow of Neuberger's operator along a loop homotopic to $\Gamma_2 \circ \Gamma_1$. As expected we find an odd number of eigenvalues that cross zero. Independently of chapter 4 this result shows $\mathcal{T}(\Gamma_2 \circ \Gamma_1) = -1$ by a completely different method.

6.2.1 A topologically non-trivial gauge transformation

An explicit example of the non-trivial element of $\pi_4(\mathrm{SU}(2))$ is constructed in ref. [34]. The crucial point of the construction is the observation that SU(3) is a non-trivial SU(2) bundle over S^5. Starting from local trivialisations of that bundle on the upper and lower half-sphere of S^5, the transition function on the equator is a homotopically non-trivial mapping from S^4 into SU(2). In terms of Pauli matrices σ_j, $j = 1, 2, 3$, the final result in [34] is written as

$$g(n) \;=\; \frac{\xi_0}{1+n_4^2}\sigma_0 + i\frac{\xi_j}{1+n_4^2}\sigma_j, \tag{6.6}$$

where σ_0 denotes the two dimensional identity matrix and n is an element of $S^4 = \{n \in \mathbb{R}^5, n^2 = 1\}$. The functions ξ_μ are given by

$$\begin{aligned} \xi_0(n) &= 2n_4, \\ \xi_1(n) &= 2(n_0 n_2 + n_1 n_3), \\ \xi_2(n) &= 2(n_0 n_3 - n_1 n_2), \\ \xi_3(n) &= n_0^2 + n_1^2 - n_2^2 - n_3^2. \end{aligned} \tag{6.7}$$

[1] At the same time, of course, the complex conjugate eigenvalues λ_i^* also cross zero. To be precise one should speak of an odd number of *pairs* (λ_i, λ_i^*) of eigenvalues that cross zero. Moreover, the crossing can happen an odd number of times. To avoid awkward and lengthy formulations we will not be linguistically pedantic in describing this phenomenon and use the simple phrase that an odd number of eigenvalues cross zero along a loop.

These functions are well known in the theory of fiber bundles. According to Hopf the manifold S^3 is a U(1) bundle over S^2. The projection map π from S^3 to S^2, sometimes called *Hopf map*, is exactly given by the functions $\xi_i, i = 1, 2, 3$. This can be seen by using the relation

$$\sum_{i=1}^{3} \xi_i^2 = \left(\sum_{i=0}^{3} n_i^2 \right)^2 , \tag{6.8}$$

which follows immediately from (6.7). Hence $\xi = (\xi_1, \xi_2, \xi_3)$ is an element of S^2 if $\tilde{n} = (n_0, n_1, n_2, n_3)$ is an element of S^3. Using (6.8) one can also easily check that $g(n)$ in (6.6) is indeed an element of SU(2).

In Witten's original argument S^4 is a compactification of $\mathbb{R}^4$. To introduce an Euclidean space time lattice we need to go back to $\mathbb{R}^4$. This is done by regarding the points of S^4 as the stereographic coordinates of $\mathbb{R}^4$.

In general, stereographic coordinates are introduced as the dimensionless quantities

$$n_\mu = \frac{2x_\mu R}{x^2 + R^2}, \tag{6.9}$$

$$n_4 = \frac{x^2 - R^2}{x^2 + R^2}, \tag{6.10}$$

where x denotes the Euclidean distance of the space–time point x_μ. The north pole $n_4 = 1$ corresponds to spatial infinity. R is an arbitrary reference scale one has to introduce. It is the radius of the sphere in the usual visualization of stereographic coordinates in one or two dimensions.

The inverse of (6.9) and (6.10) for $n_4 \neq 1$ is given by

$$q_\mu \equiv \frac{x_\mu}{R} = \frac{n_\mu}{1 - n_4}.$$

Given the stereographic coordinates n_μ one can only recover the scaled coordinates q_μ.

Coming back to the map (6.6) we replace n_μ by the formulae (6.9), (6.10) and find the corresponding map from $\mathbb{R}^4$ into SU(2) as

$$g(q) = \frac{q^4 - 1}{q^4 + 1}\sigma_0 + i\frac{2\xi_j}{q^4 + 1}\sigma_j, \tag{6.11}$$

where the abbreviation q^2 is used for x^2/R^2. The functions ξ_j are the same as in (6.7) but with n_μ replaced by q_μ.

In the limit $q^2 \to \infty$, $g(q)$ becomes the identity. Because of the fourth power of q in (6.11) this limit is practically reached for rather moderate values q. At the origin g is identical to $-\sigma_0$. The non–trivial part of the gauge transformation g is thus concentrated within a finite sphere of some radius R'.

The mapping g has various symmetry properties. For instance, it is invariant under the transformation $q_\mu \to -q_\mu$. In addition, there exists a transformation T that transforms g into its inverse. This transformation is explicitly given by

$$T: \quad \begin{aligned} q_0 &\to -q_3, \\ q_1 &\to -q_2, \\ q_2 &\to q_1, \\ q_3 &\to q_0. \end{aligned} \tag{6.12}$$

This unitary transformation satisfies $T^2 = -1$ and we find

$$g(Tq) \;=\; g^{-1}(q) \tag{6.13}$$

for the gauge transformation (6.11).

6.2.2 Definition of the path

Restricting the coordinates x_μ in (6.11) to the points of a lattice defines a lattice approximation of the gauge transformation g. The lattice spacing a must be chosen carefully so that this approximation exhibits the non-trivial topology of the continuum gauge transformation. Hence we require the scaled lattice spacing

$$\hat{a} \;=\; \frac{a}{R} \tag{6.14}$$

to be much smaller than 1. When we use a finite lattice we simply 'cut' $g(x)$ for x_μ larger than the lattice size L. Proceeding that way we do not loose the topology of g provided L is large enough. So the ratio L/R should be at least of order 1.

We construct a path similar to (4.2) that connects a constant gauge field V and its gauge transform V^g. The reason why we use a constant gauge field instead of the classical vacuum will become clear later.

We define the constant gauge field

$$V(x,\mu) \;=\; \exp\left(i\theta(\mu)\sigma_3\right) \tag{6.15}$$

with some angles θ we will specify later. We still write $V(x,\mu)$ even though the field is independent of x. In contrast the gauge transformed field $V^g(x,\mu)$ depends on x. Using V and V^g we define the path

$$\begin{aligned} \tilde{\Gamma}_2: \qquad U_t(x,\mu) &= N(t)^{-1}\left\{ tV(x,\mu) + (1-t)V^g(x,\mu) \right\}, \\ N(t) &= \det\left\{ tV(x,\mu) + (1-t)V^g(x,\mu) \right\}. \end{aligned} \tag{6.16}$$

As long as $V \neq V^g$ the link $U_t(x,\mu)$ is unequal to zero for all values of t. The normalization factor $N(t)$ ensures that $U_t(x,\mu)$ is an element of SU(2). Notice that U_t is an analytic function of t.

The path (6.16) has a nice symmetry property. The normalization factor satisfies $N(t) = N(1-t)$. Taking this into account one finds for the link variables the relation

$$\begin{aligned} U_t(x,\mu) &= g(x)\tilde{U}_{1-t}(x,\mu)g(x+a\mu)^{-1}, \\ \tilde{U}_t(x,\mu) &= N(t)^{-1}\left\{tV(x,\mu) + (1-t)V^{g^{-1}}(x,\mu)\right\}, \end{aligned} \tag{6.17}$$

i.e. U_t and $\tilde{U}_{1-t}$ are gauge equivalent. The gauge field $\tilde{U}_t$ differs from U_t by using the inverse gauge transformation g^{-1} for the gauge transform of V.

Consider an arbitrary lattice Dirac operator D. The spectra of $D(U_t)$ and $D(\tilde{U}_{1-t})$ are identical. On the other hand the spectra of $D(\tilde{U}_{1-t})$ and $D(U_{1-t})$ are the same by means of (6.13). Therefore we find the symmetry relation

$$\sigma\left(D(U_t)\right) \;=\; \sigma\left(D(U_{1-t})\right) \tag{6.18}$$

for the spectrum of $D(U_t)$ as a function of t.

The path $\tilde{\Gamma}_2$ can be closed by a second path $\tilde{\Gamma}_1$, defined as

$$\tilde{\Gamma}_1 : \qquad U_t(x,\mu) \;=\; g(x)^t V(x,\mu) g(x+a\mu)^{-t}\,. \tag{6.19}$$

The loop $\tilde{\Gamma}_2 \circ \tilde{\Gamma}_1$ lies in the same homotopy class as $\Gamma_2 \circ \Gamma_1$, defined in section 4.1. The twist along $\tilde{\Gamma}_2 \circ \tilde{\Gamma}_1$ is therefore equal to -1.

6.2.3 Spectral flow of Neuberger's operator

Our aim is to detect an odd number of eigenvalues that cross zero as a function of the path parameter t. In the case of Neuberger's operator this means a crossing of the imaginary part of the eigenvalues.

We proceed in two steps. First of all we compute the lowest eigenvalues of $D^\dagger D$ and obtain the absolute value of the eigenvalues of D.[2] In a second step we calculate the imaginary part only for the eigenvalues that vanish for some value t_0 in order to establish the crossing.

We proceed that way because the Conjugate Gradient algorithm [35, 36] we want to employ in the numerical computation can only compute the lowest eigenvalues of hermitian operators. As mentioned in section 2.5 Neuberger's operator is normal but not hermitian. For that reason we compute the eigenvalues of $D^\dagger D$, which is obviously hermitian, and calculate separately the imaginary part of the eigenvalues. Let us first discuss how this second step can be done before we present the numerical results.

As discussed in section 2.5 there exists an orthonormal basis of eigenvectors of D because D is normal. If w_j is an eigenvector with eigenvalue λ_j, $\tilde{w}_j = \gamma_5 w_j$ is an eigenvector with eigenvalue λ_j^*. For all j the corresponding two dimensional subspace $U_j = \mathrm{span}(\omega_j, \tilde{\omega}_j)$ is invariant under the action of D. If the gauge group is SU(2) in the fundamental representation, one can find a chiral basis for all U_j, which has the additional property

$$\omega_{+,j}^* = K\omega_{+,j}\,, \qquad \omega_{-,j}^* = K\omega_{-,j}\,, \tag{6.20}$$

with K being an unitary operator.[3] Vectors which have the property (6.20) are called *real* in the following. The reality of the basis vectors fixes their phase up to a sign. The existence of a real basis is intimately related to the pseudo reality of the fundamental representation of SU(2) and to the property

$$D^* \;=\; KDK^{-1} \tag{6.21}$$

[2] From now on we set $a = 1$ for notational simplicity.

[3] The explicit form of K depends on the representation for the γ matrices used for the Dirac operator. In appendix D the operator K is explicitly given for the Weyl representation.

of Neuberger's operator for this representation. Here the complex conjugate operator is defined in terms of the complex conjugate kernel.

If we map U_j onto $\mathbb{C}^2$ by the assignment

$$\begin{aligned} \omega_{+,j} &\longrightarrow \binom{1}{0}, \\ \omega_{-,j} &\longrightarrow \binom{0}{1}, \end{aligned} \tag{6.22}$$

we define a two dimensional matrix representation D_j of Neuberger's operator on U_j. Since $V = 1 - D$ is an unitary operator that satisfies $V^\dagger = \gamma_5 V \gamma_5$, the matrix D_j representing D can be parameterized by a single parameter θ_j:

$$D_j \longrightarrow \begin{pmatrix} 1 - \cos\theta_j & \sin\theta_j \\ -\sin\theta_j & 1 - \cos\theta_j \end{pmatrix} . \tag{6.23}$$

The parameter θ_j is defined by the off diagonal element D_j^{12}, i.e.

$$D_j^{12} = (\omega_{+,j}|D|\omega_{-,j}) = \sin\theta_j \,. \tag{6.24}$$

Note that D_j^{12} is real because of (6.20) and (6.21). However, the sign of it is not fixed because of the sign ambiguity of the basis vectors $\omega_{+,j}$ and $\omega_{-,j}$.

The eigenvalues of (6.23) are easily found as

$$\lambda_\pm = 1 - e^{\pm i\theta_j} . \tag{6.25}$$

We already know this result but now the relation between the eigenvalues and the off diagonal matrix element D_j^{12} is transparent and reads

$$D_j^{12} = \pm \mathrm{Im}\,\lambda_j \,. \tag{6.26}$$

This is our desired result. Whenever we have the basis vectors $\omega_{+,j}$ and $\omega_{-,j}$ at hand, we are able to compute the imaginary part of λ_j up to a sign by calculating the scalar product in (6.24). Since equation (6.26) remains valid for all values of t along the path $\tilde{\Gamma}_2 \circ \tilde{\Gamma}_1$, a crossing of the eigenvalue $\lambda_j(t)$ is accompanied by a change of sign in the matrix element $D_j^{12}(t)$.

The Conjugate Gradient (CG) algorithm is a method to compute numerically the lowest lying eigenvalues of hermitian operators. In addition to the eigenvalues it also gives the corresponding eigenvectors. Moreover, it is able to obtain the vectors $\omega_{+,j}$ and $\omega_{-,j}$, which are eigenvectors of $D^\dagger D$. Note that they are not eigenvectors of D itself. Hence the CG algorithm enables us to compute the imaginary part of the eigenvalues of Neuberger's operator.

To fix the sign ambiguity in (6.26) we finally remark that the basis vectors $\omega_{+,j}$ and $\omega_{-,j}$ depend smoothly on t along the path $\tilde{\Gamma}_2 \circ \tilde{\Gamma}_1$. Once we have arbitrarily fixed the sign of the basis vectors for some value of t, continuity fixes the sign along the entire path.

6.2.4 Numerical results

In the following we present the numerical results for the eigenvalues of $D^\dagger D$ and for the matrix element $D^{12}(t)$. But before we make some remarks concerning the technical details.

According to our discussion in section 2.6 we used a truncated Chebyshev expansion for the inverse square root of $A^\dagger A$ in Neuberger's operator. For numerical stability we have used Clenshaw's recurrence formula [37] for the finite sum of Chebyshev polynomials in D_N.

In all computations we have chosen $N = 50$ resulting in the bound

$$||R_{50}|| \leq 5 \cdot 10^{-5} \tag{6.27}$$

along the path (6.16). The use of the truncated operator D_{50} instead of D in the numerical computation results in truncation errors for both the eigenvalues of $D^\dagger D$ and for D^{12}.

Another source of uncertainty is the numerical error. As already mentioned, we used the CG algorithm for the computation of the lowest lying eigenvalues of $D_{50}^\dagger D_{50}$. This algorithm is described in detail in [35,36]. Here we just remark that the main idea of the algorithm is the minimization of the Ritz functional. This method yields reliably the eigenvalues with controlled numerical error. In addition it gives approximations of the corresponding eigenvectors ω_+, ω_-.

Both the truncation and the numerical error can be treated analytically and strict error bounds can be derived for the eigenvalues as well as for D^{12}. The derivation of the error bounds is given in detail in appendix C. These bounds are taken into account in the numerical results presented below.

The angles θ in the constant gauge field (6.15) are set to the values

$$\begin{aligned} \theta(0) &= 0.6\,\pi\,, \\ \theta(1) &= 0.55\,\pi\,, \\ \theta(2) &= 0.5\,\pi\,, \\ \theta(3) &= 0.45\,\pi\,. \end{aligned} \tag{6.28}$$

We have chosen slightly different values for each direction to remove as much degeneracy of the spectrum as possible. The spectrum of Neuberger's operator for a constant background gauge field can be computed analytically, but the analytic expression is rather lengthy. Nevertheless, for the lowest eigenvalue one can find the approximation

$$|\lambda_0|^2 \approx \sum_\mu \sin^2 \frac{\theta(\mu)}{L/a}\,. \tag{6.29}$$

This value is well above zero for large enough angles θ. This is exactly the reason why we have used a constant gauge field for the starting point of our path. If we had used the classical vacuum configuration the lowest eigenvalue would have been very close to zero along the whole path. Detecting a crossing in that case requires very small truncation and numerical errors and is therefore fairly time consuming if it is numerically possible at all.

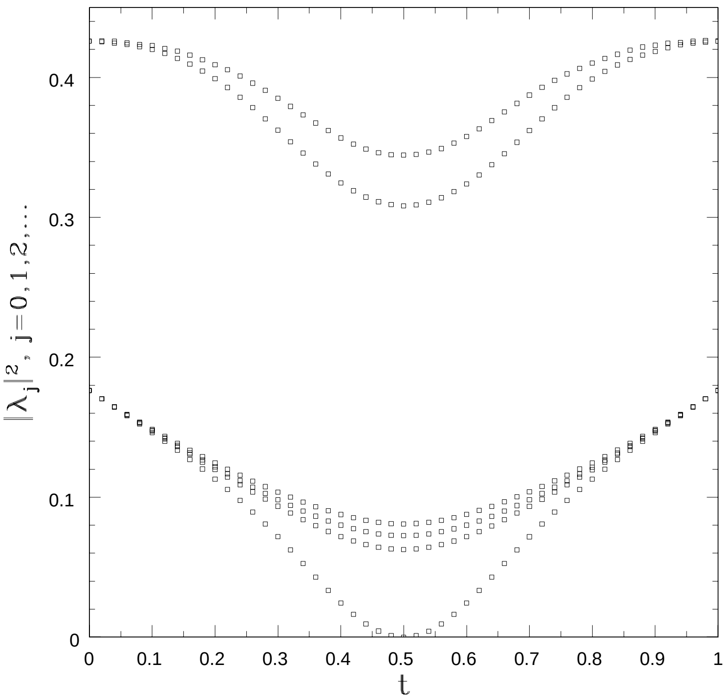

Figure 6.1: The data for the lowest six eigenvalues of $D^\dagger D$ on a 8^4 lattice. The scaled lattice spacing $\hat{a}$ is set to 0.275. The total error is smaller than the size of the data points. Only the lowest eigenvalue becomes zero for $t = 0.5$.

Let us now turn to the numerical results. Figure 6.1 shows the first 6 eigenvalues of $D^\dagger D$ along $\tilde{\Gamma}_2$ for an 8^4 lattice with scaled lattice spacing $\hat{a}$ equal to 0.275. Each eigenvalue is twofold degenerate. The total error is smaller than the size of the symbols for the data points. The values at $t = 0$ and $t = 1$ agree well with (6.29) for $L/a = 8$ and the symmetry property (6.18) of the spectrum is clearly visible.

The absolute value for all but one eigenvalue is unequal to zero. Only the lowest eigenvalue becomes zero once along the path. If there is a crossing at all only the lowest eigenvalue crosses zero at $t = 0.5$. In order to decide if the imaginary part of λ_0 really changes sign we computed D_0^{12}. Figure 6.2 shows the numerical result for the imaginary part of λ_0 and we indeed find a crossing.

We also made computations on a 12^4 lattice and for various scaled lattice spacings $\hat{a}$. The qualitative result was always the same, only the lowest eigenvalue crossed zero. But for $\hat{a} > 0.5$ no crossing took place. In that case the lattice is not fine enough and the non-trivial topology of the gauge transformation g is lost.

One important issue deserves a final comment. As already mentioned, the differentiability of Neuberger's operator is guaranteed rigorously for smooth gauge fields only. As discussed in section 2.4 this property holds if $A^\dagger A$ has no

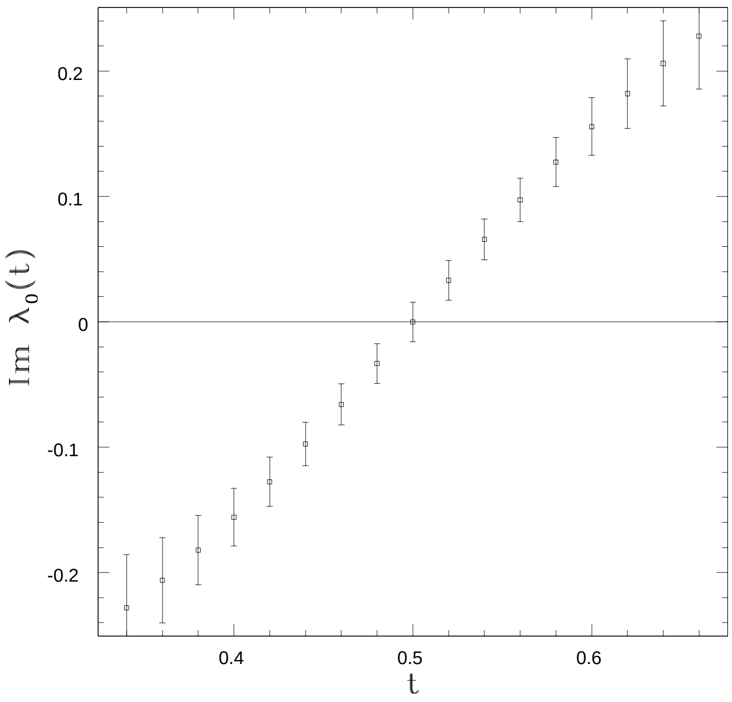

Figure 6.2: The imaginary part of λ_0 on a 8^4 lattice and for $\hat{a} = 0.275$. The error bars incorporate both the truncation and the numerical error.

zero mode. Numerically we found

$$0.4 \quad < \|A^\dagger A\| \tag{6.30}$$

along our particular path $\tilde{\Gamma}_2$ and the differentiability is therefore given. However, the bound (2.21) was not fulfilled. We found numerically that this bound is violated unless we use at least a 24^4 lattice and scaled lattice spacings smaller than 0.1. This could not have been done on the computers available to us. However, if the crossing is present on small lattices it is also present for bigger ones with smaller lattice spacings.

6.2.5 Final remarks

We have numerically shown that only the lowest eigenvalue of Neuberger's operator crosses zero along the path $\tilde{\Gamma}_2$. This result remains unchanged along the loop $\tilde{\Gamma}_2 \circ \tilde{\Gamma}_1$, because all configurations along $\tilde{\Gamma}_1$ are gauge equivalent.

The numerical result for the spectral flow together with section 6.1 serves as a simple 'numerical proof' of

$$\mathcal{T}(\tilde{\Gamma}_2 \circ \tilde{\Gamma}_1) \quad = \quad -1. \tag{6.31}$$

We, of course, expected this result because $\tilde{\Gamma}_2 \circ \tilde{\Gamma}_1$ is homotopic to the loop $\Gamma_2 \circ \Gamma_1$ which we considered in chapter 4. Turning the argument the other way around we have shown $\mathcal{T}(\Gamma_2 \circ \Gamma_1) = -1$ by a completely different and independent method.

In the analytic treatment in chapter 4 we were not precise concerning the lattice size L and the lattice spacing a. It was enough to say that L should be sufficiently large and a sufficiently small in order to establish $\mathcal{T}(\Gamma_2 \circ \Gamma_1) = -1$. As an interesting by-product the numerical result of this section tells us that loops with a non-trivial twist already exist on small lattices with 8^4 lattice points.

Chapter 7

Final remarks

In this thesis global anomalies in chiral lattice gauge theories are discussed. The presence of these anomalies makes the definition of a fermion measure for the functional integral impossible unless one violates some other fundamental property like gauge invariance or locality.

The key quantity in our treatment is the twist along loops in configuration space. Global anomalies are likely to be present whenever loops with a non-trivial twist exist. An explicit example for such a loop is given for the case of an SU(2) gauge theory in the fundamental representation.

The twist along this loop was computed twice by two completely independent methods. In the first method SU(2) was regarded as a subgroup of SU(3) and the twist has been computed completely analytically. Secondly the twist has been inferred from a numerical computation of the spectral flow of Neuberger's operator along the path.

Our investigation of global anomalies on the lattice was inspired by Witten's SU(2) anomaly in continuum field theory. And in fact, as has become clear throughout our derivation, the global SU(2) anomaly on the lattice is basically the same as Witten's obstruction. However, the treatment in this thesis is not at all a dull carbon copy of some continuum discussion.

Even though Witten's original argument is very smart, it is purely formal and maybe not very convincing from a mathematical point of view. Perhaps most importantly, the results presented here were obtained in a fully regularized and mathematically well-defined gauge theory. In particular, the global anomaly is already present for finite lattice spacings and not just in the continuum limit.

Aside from the fundamental representation of SU(2) we found the theory to be anomalous for representations with highest weight $j = 2n + 1/2$. It is important to realize that this does not automatically imply that all the other representations are anomaly-free. To establish this one has to show that the twist is equal to 1 along *all* loops in configuration space. This has not been done yet and is left for future research.

There is another open problem that is interesting to address. Throughout this thesis the vacuum sector of the theory has been considered only. The twist, our main tool, is also defined for the topologically non-trivial sectors and it should be possible to investigate the existence of global anomalies in these sectors too. This might result in the exclusion of further fermion representations.

Appendix A

Proof of the Alternation Theorem

Alternation theorem: Let f, ρ be continuous functions on $[a, b]$ with ρ strictly positive. $p_N(x)$ denotes the least maximum approximation of f, i.e. among all polynomials of degree less or equal N the polynomial p_N minimizes

$$\|f - \rho p_N\|_\infty \;=\; \max_{x \in [a,b]} |f(x) - \rho(x) p_N(x)| \,.$$

Suppose there exists a polynomial q_N of degree N such that $\rho(x) q_N(x)$ has $N + 2$ points of alternation, i.e. there exist $N + 2$ values x_i, $a \leq x_0 < x_1 < \cdots < x_{N+1} \leq b$, such that

$$f(x_i) - \rho(x_i) q_N(x_i) \;=\; \rho(x_{i+1}) q_N(x_{i+1}) - f(x_{i+1}) \;=\; \pm \|f - \rho q_N\|_\infty \,.$$

That is the maximum discrepancy between f and ρq_N occurs with alternating sign at $N + 2$ successive points. Then $p_N = q_N$, i.e. q_N is the least maximum approximation of f.

Proof: We prove the theorem by contradiction. We assume $p_N \neq q_N$. Since the least maximum approximation is unique[1] we find

$$\|f - \rho p_N\|_\infty \;<\; \|f - \rho q_N\|_\infty \,. \tag{A.1}$$

We define the two error functions

$$\begin{aligned} e(x) &= f(x) - \rho(x) p_N(x), \\ \tilde{e}(x) &= f(x) - \rho(x) q_N(x), \end{aligned}$$

and the difference $d(x) = \tilde{e}(x) - e(x)$. At the points of alternation we find

$$\operatorname{sgn} d(x_i) \;=\; -\operatorname{sgn} d(x_{i+1}) \,. \tag{A.2}$$

Hence $d(x)$ has at least $N+1$ zeros. The same is true for the quotient $d(x)/\rho(x)$ because $\rho(x)$ does not change sign. Since this quotient is a polynomial of degree N it must vanish identically. This implies

$$p_N(x) \;=\; q_N(x)$$

and contradicts (A.1). □

[1] Proofs for the existence and uniqueness of $p_N(x)$ can be found in [38].

Appendix B

Chebyshev expansion

The literature about Chebyshev polynomials and expansions in terms of them is vast. As a reference for the following we just quote [39–42], where more details can be found.

Chebyshev polynomials are defined by

$$T_n(x) \;=\; \cos n\theta\,, \qquad x \;=\; \cos\theta\,. \tag{B.1}$$

Here x is restricted to $[-1,1]$. Using trigonometric identities one can rewrite $\cos n\theta$ as a polynomial of degree n in $\cos\theta$. This leads to a polynomial of degree n in x. Explicitly, the first six Chebyshev polynomials read

$$\begin{array}{rclrcl} T_0(x) &=& 1 & T_3(x) &=& 4x^3-3x, \\ T_1(x) &=& x, & T_4(x) &=& 8x^4-8x^2+1, \\ T_2(x) &=& 2x^2-1, & T_5(x) &=& 16x^5-20x^3+5x. \end{array} \tag{B.2}$$

The Chebyshev polynomials satisfy the bound

$$|T_n(x)| \;\leq\; 1 \tag{B.3}$$

for x in $[-1,1]$. There exist $n+1$ points x_i where $T_n(x_i)$ attains one of its extremal values ± 1.

In principle, every trigonometric identity has its analogue in terms of Chebyshev polynomials. For example the identity $\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta$ leads to the recurrence relation

$$T_{n+1}(x) \;=\; 2x\,T_n(x) - T_{n-1}(x)\,. \tag{B.4}$$

Note that the coefficients in front of T_n and T_{n-1} are independent of n. The recurrence relation can be used to show that $T_n(x)$ is either even or odd. Furthermore one can prove that the leading coefficient in T_n, $n \geq 1$, is given by 2^{n-1}.

The Chebyshev polynomials satisfy the relation

$$\int_{-1}^{+1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}\,dx \;=\; \int_0^{\pi} \cos n\theta\,\cos m\theta\,d\theta \;=\; \begin{cases} 0 & m \neq n \\ \frac{\pi}{2} & m = n \neq 0 \\ \pi & m = n = 0 \end{cases} \tag{B.5}$$

The Chebyshev polynomials thus form an orthogonal set over the interval $[-1,1]$ with weight function $\omega(x) = (1-x^2)^{-1/2}$. The corresponding expansion of real valued functions f reads

$$
\begin{aligned}
f(x) &= \sum_{n=0}^{\infty}{}' a_n T_n(x)\,, &\text{(B.6)}\\
a_n &= \frac{2}{\pi}\int_{-1}^{+1} \omega(x) f(x) T_n(x)\,dx &\text{(B.7)}\\
&= \frac{2}{\pi}\int_0^{\pi} f(\cos\theta)\cos n\theta\,d\theta\,. &\text{(B.8)}
\end{aligned}
$$

The prime at the summation sign indicates that the first term in the sum is taken with a factor $1/2$. This results from the different values of the integral (B.5) for $m = n = 0$ and $m = n \neq 0$. For sufficiently well behaved functions f the Chebyshev expansion (B.6) will converge uniformly to f. For example this is true if f is continuously differentiable with bounded derivative.

To conclude we summarize the formulae for functions f defined on arbitrary intervals $[a,b]$. The function

$$
g(\alpha) = \frac{b+a}{2} - \frac{b-a}{2}\,\alpha \qquad \text{(B.9)}
$$

maps the interval $[-1,1]$ onto $[a,b]$ with $g(1) = a$ and $g(-1) = b$. The composition $\tilde{f} = f \circ g$ is defined on $[-1,1]$ and can be expanded in Chebyshev polynomials. The expansion coefficients read

$$
a_n = \frac{2}{\pi}\int_0^{\pi} f\left(\frac{b+a}{2} - \frac{b-a}{2}\cos\theta\right)\cos n\theta\,d\theta\,. \qquad \text{(B.10)}
$$

if we use the integral representation (B.8). Since the function g is invertible we find the expansion

$$
f(x) = \sum_{n=0}^{\infty}{}' a_n T_n\left(\frac{b+a}{b-a} - \frac{2}{b-a}\,x\right) \qquad \text{(B.11)}
$$

for $f = \tilde{f} \circ g^{-1}$.

Appendix C

Proof of lemma 2

II. $\mathcal{T}(2n+\frac{1}{2}) = -1$ and $\mathcal{T}(2n+\frac{3}{2}) = 1$ for $n \in \mathbb{N}_0$.

Proof: The statement is true for $n = 0$. To prove it for arbitrary integers we first define

$$\Sigma_a(2n+\frac{1}{2}) = \sum_{k=0}^{n-1} c(4n+1, \frac{4k+1}{2}), \qquad n \geq 1\,, \tag{C.1}$$

$$\Sigma_a(2n+\frac{3}{2}) = \sum_{k=0}^{n} c(4n+3, \frac{4k+1}{2}), \qquad n \geq 0\,. \tag{C.2}$$

$\Sigma_a(j)$ gives the total number of anomalous representations in the decomposition (5.1) of $(\frac{1}{2} \oplus 0)\otimes^{2j}$ with highest weight less than j. In addition we define

$$\Sigma_f(2n+\frac{1}{2}) = \sum_{k=1}^{n} c(4n+1, \frac{4k-1}{2}), \qquad n \geq 1\,, \tag{C.3}$$

$$\Sigma_f(2n+\frac{3}{2}) = \sum_{k=1}^{n} c(4n+3, \frac{4k-1}{2}), \qquad n \geq 1\,. \tag{C.4}$$

$\Sigma_f(j)$ gives the number of anomaly free *half integer* representations in the decomposition of $(\frac{1}{2} \oplus 0)\otimes^{2j}$, again with highest weight less than j. Formulae (5.2) and (5.8) imply

$$\mathcal{T}(j) \cdot (-1)^{\Sigma_a(j)} = -1 \tag{C.5}$$

for half-integer j. In order to prove the statement all we need to show is that $\Sigma_a(2n+\frac{1}{2})$ is even and $\Sigma_a(2n+\frac{3}{2})$ is odd.

Using table (5.1) one can easily check that $\Sigma_a(5/2)$ is even and $\Sigma_a(3/2)$ odd. Now let us assume that $\Sigma_a(2n+\frac{1}{2})$ is even for some fixed integer $n \geq 1$. We consider the tensor product

$$(\frac{1}{2} \oplus 0)\otimes^{4n+3} = 1 \otimes \left((\frac{1}{2} \oplus 0)\otimes^{4n+1} \right) \oplus \sum_j \tilde{c}(4n+3, j)j\,, \tag{C.6}$$

where all the coefficients $\tilde{c}$ are even integers. The tensor product with the triplet representation can be decomposed into irreducible representations with the result

$$1\otimes\left((\frac{1}{2}\oplus 0)\otimes^{4n+1}\right) = \sum_{k=1}^{2n+2} b(4n+3,\frac{2k-1}{2})\frac{2k-1}{2} \oplus \text{integer representations}. \tag{C.7}$$

The coefficients b are given in terms of the coefficients c as

$$\begin{array}{lclclcl}
b(4n+3,\frac{1}{2}) & = & c(4n+1,\frac{1}{2}) & + & c(4n+1,\frac{3}{2}), & & \\
\vdots & & \vdots & & \vdots & & \\
b(4n+3,\frac{2k+1}{2}) & = & c(4n+1,\frac{2k-1}{2}) & + & c(4n+1,\frac{2k+1}{2}) & + & c(4n+1,\frac{2k+3}{2}), \\
\vdots & & \vdots & & \vdots & & \vdots \\
b(4n+3,\frac{4n+1}{2}) & = & & & c(4n+1,\frac{4n-3}{2}) & + & c(4n+1,\frac{4n-1}{2}), \\
b(4n+3,\frac{4n+3}{2}) & = & & & & & c(4n+1,\frac{4n-1}{2}).
\end{array} \tag{C.8}$$

We denote the contribution of (C.7) to $\Sigma_a(2n+\frac{3}{2})$ by $\tilde{\Sigma}_a(2n+\frac{3}{2})$. Using $c(m,\frac{m}{2})=1$ we find

$$\tilde{\Sigma}_a(2n+\frac{3}{2}) \quad = \quad 1+\Sigma_a(2n+\frac{1}{2})+2\Sigma_f(2n+\frac{1}{2}). \tag{C.9}$$

We assumed $\Sigma_a(2n+\frac{1}{2})$ to be even, hence $\tilde{\Sigma}_a(2n+\frac{3}{2})$ is odd. Because the coefficients $\tilde{c}$ in (C.6) are all even, $\Sigma_a(2n+\frac{3}{2})$ is also odd, i.e.

$$\Sigma_a(2n+\frac{1}{2}) \text{ is even} \quad\Longrightarrow\quad \Sigma_a(2n+\frac{3}{2}) \text{ is odd}. \tag{C.10}$$

Having established this result, we consider now the tensor product (C.6) with exponent $4n+5$. Equations (C.6), (C.7) and (C.8) need just be modified by the replacement $n\to n+1/2$. Similar to (C.9) we find

$$\tilde{\Sigma}_a(2n+\frac{5}{2}) \quad = \quad 1+\Sigma_a(2n+\frac{3}{2})+2\,\Sigma_f(2n+\frac{3}{2}). \tag{C.11}$$

We therefore conclude:

$$\Sigma_a(2n+\frac{3}{2}) \text{ is odd} \quad\Longrightarrow\quad \Sigma_a(2(n+1)+\frac{1}{2}) \text{ is even}. \tag{C.12}$$

This completes the proof. □

Appendix D

The operator K in the Weyl representation

Here we want to present the explicit expression of the operator K appearing in (6.20) for the Weyl representation.

The Weyl representation for the Dirac matrices is given by

$$\begin{aligned}\gamma_0 &= \begin{pmatrix} 0 & -\sigma_0 \\ -\sigma_0 & 0 \end{pmatrix}, \\ \gamma_k &= \begin{pmatrix} 0 & -i\sigma_k \\ i\sigma_k & 0 \end{pmatrix}, \qquad k = 1, 2, 3\,.\end{aligned} \tag{D.1}$$

The σ_k denote the Pauli matrices and σ_0 denotes the unity matrix. The γ matrices have the properties

$$\begin{aligned}\gamma_\mu^\dagger &= \gamma_\mu, \\ \{\gamma_\mu, \gamma_\nu\} &= 2\delta_{\mu\nu}\,.\end{aligned} \tag{D.2}$$

In the Weyl representation $\gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3$ is diagonal and reads

$$\gamma_5 = \begin{pmatrix} \sigma_0 & 0 \\ 0 & -\sigma_0 \end{pmatrix}. \tag{D.3}$$

This matrix has the properties

$$\begin{aligned}\gamma_5^\dagger &= \gamma_5, \\ \{\gamma_\mu, \gamma_5\} &= 0\,.\end{aligned} \tag{D.4}$$

Let us define the matrix

$$C = \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix}, \tag{D.5}$$

with $\epsilon = i\sigma_2$. The matrix C has the properties

$$C^\dagger = C^T = -C = -C^* = C^{-1}\,. \tag{D.6}$$

With the help of C complex conjugation of the γ matrices can be written as

$$\gamma_\mu^* \;=\; C\gamma_\mu C^{-1}. \tag{D.7}$$

Note that C is not the matrix $\mathcal{C}$ that one uses commonly in connection with charge conjugation. Using this matrix instead of C in (D.7) one ends up with $-\gamma_\mu^*$ on the left hand side. Expressed in terms of the matrices (D.1) we find $\mathcal{C} = i\gamma_0\gamma_2$ in comparison with $C = -\gamma_1\gamma_3$.

The fundamental representation of SU(2) is pseudo real [3] and complex conjugation of any SU(2) matrix U is given by

$$U^* \;=\; \epsilon\, U\, \epsilon^{-1}\,. \tag{D.8}$$

The matrix ϵ is explicitly given by

$$\epsilon \;=\; \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{D.9}$$

and one easily checks the properties

$$\epsilon^\dagger \;=\; \epsilon^T \;=\; -\epsilon \;=\; -\epsilon^* \;=\; \epsilon^{-1}\,. \tag{D.10}$$

Now we define K as the tensor product of ϵ and the matrix C in (D.5). Explicitly we find

$$K = C\epsilon \;=\; \begin{pmatrix} \begin{matrix} 0 & \epsilon \\ -\epsilon & 0 \end{matrix} & 0 \\ 0 & \begin{matrix} 0 & \epsilon \\ -\epsilon & 0 \end{matrix} \end{pmatrix}. \tag{D.11}$$

Either from this explicit representation or from the properties of ϵ and C itself one can establish the identities

$$K^\dagger \;=\; K^T \;=\; K \;=\; K^* \;=\; K^{-1}\,. \tag{D.12}$$

Having defined K one can show that the usual Wilson–Dirac operator D_w (2.15) satisfies $D_w^* = K D_w K^{-1}$. This also holds for Neuberger's operator because it is essentially a power series of D_w and $D_w^\dagger$.

Appendix E

Error bounds

The expansion of $(A^\dagger A)^{-1/2}$ into orthogonal polynomials results in an infinite series and must be truncated in any practical computation. The eigenvalues of the truncated Neuberger operator will differ from the eigenvalues of Neuberger's operator itself we are interested in. To make this quantitatively more precise we split D into a sum of two terms (see section 2.6),

$$D = D_N + R_N \,. \tag{E.1}$$

In D_N all powers $(A^\dagger A)^n$ with $n > N$ are neglected and appear in the remainder R_N only. Note that the truncated operator D_N is no longer normal. Now we will make use of the following lemma that is proved in [43], page 109.

Lemma A: Let A, B be hermitian n×n matrices. The eigenvalues $\lambda_j(A)$ and $\lambda_j(B)$ are labeled according to

$$\begin{array}{ccccccc} \lambda_1(A) & \geq & \lambda_2(A) & \geq & \ldots & \geq & \lambda_n(A)\,, \\ \lambda_1(B) & \geq & \lambda_2(B) & \geq & \ldots & \geq & \lambda_n(B)\,. \end{array}$$

Then the inequality

$$|\lambda_j(A) - \lambda_j(B)| \leq \|A - B\|$$

holds where $\|\cdot\|$ denotes the spectral norm.

This lemma is easily applied to the case $A = D^\dagger D$ and $B = D_N^\dagger D_N$. Replacing D_N by $D - R_N$ and using $\|D\| \leq 2$ we find

$$\Big|\, |\lambda_j|^2 - |\lambda_j^N|^2 \,\Big| \leq \omega_t \,, \tag{E.2}$$

where we introduced the truncation error ω_t, given by

$$\omega_t = 4\|R_N\| + \|R_N\|^2 \,. \tag{E.3}$$

Let us now turn to the numerical error. For the numerical computation of the eigenvalues we used the Conjugate Gradient (CG) algorithm [35, 36]. This algorithm computes iteratively the lowest lying eigenvalues of a hermitian matrix

A. It generates a sequence of normalized vectors v_i and real numbers $\mu_i = \mu(v_i)$ such that the norm of the gradient g_i, defined by

$$g_i \quad = \quad (A - \mu_i) v_i , \tag{E.4}$$

decreases in each step. The algorithm stops if the stopping criterion

$$\|g_i\| \quad < \quad \omega_n \tag{E.5}$$

is satisfied for some fixed number ω_n . At this final stage of the algorithm a lemma states that an eigenvalue λ of A exists that satisfies

$$|\lambda - \mu_i| \quad < \quad \omega_n \, . \tag{E.6}$$

μ_i is therefore an eigenvalue approximation and ω_n is the numerical error. In our case A is given by $D_N^\dagger D_N$. Using the triangle inequality we easily find the bound

$$\Big| \, |\lambda_j|^2 - \mu_j \, \Big| \quad < \quad \omega_t + \omega_n \, . \tag{E.7}$$

The total error is just the sum of the truncation error and the numerical error.

Let us now turn to the matrix element (6.24). Using the triangle and the Cauchy – Schwartz inequality one establishes the relation

$$\Big| \, |\langle\omega_+|D|\omega_-\rangle| - |\langle v_+|D_N|v_-\rangle| \, \Big| \quad \leq \quad \|R_N\| + 2\|\omega_+ - v_+\| + 2\|\omega_- - v_-\| \tag{E.8}$$

for four normalized vectors $\omega_\pm, v_\pm$. If ω_+, ω_- are the vectors in (6.24) this inequality serves as a starting point for an error bound of Im λ.[1]

As already mentioned, the CG algorithm generates intermediately a sequence of vectors v_i. The eigenvectors of $D^\dagger D$ can be chosen real and chiral. Hence, the CG algorithm still converges for a sequence of real and chiral vectors. Assume v_+ to be the last vector in a sequence of vectors with positive chirality. Then the stopping criterion (E.5) is fulfilled and according to (E.4) v_+ satisfies the relation

$$\|(D^\dagger D - \mu) v_+\| \quad < \quad \omega_t + \omega_n \, . \tag{E.9}$$

We conclude that if the total error is small the vector v_+ is a good approximation of the eigenvector ω_+ of $D^\dagger D$. Here we assume the eigenvalue $|\lambda|^2$ of $D^\dagger D$ to be only double degenerate. Hence there exists only one real eigenvector ω_+ with positive chirality. Now we are in the position to employ the following lemma.

Lemma B: Let A be hermitian and λ_i the eigenvalues of it. The minimal distance δ_i between λ_i and all the other eigenvalues $\lambda_j \neq \lambda_i$ is defined by

$$\delta_i \quad = \quad \min_{\lambda_j \neq \lambda_i} | \, \lambda_j - \lambda_i \, | \quad > \quad 0 \, . \tag{E.10}$$

[1] We suppress the label j for notational simplicity.

P_{λ_i} denotes the projector onto the eigenspace of λ_i. If v is a normalized vector and $\mu \in \mathbb{R}$ such that $\|(A-\mu)v\| < \delta_i/2$, then $P_{\lambda_i}v$ is an eigenvector of A and v is an approximation of it that satisfies the error bound

$$\|v - P_{\lambda_i}v\| \quad \leq \quad \frac{2}{\delta_i}\,\|(A-\mu)v\|\,. \tag{E.11}$$

The proof of Lemma B can be found in [43], page 230. We can apply this lemma to the case $A = D^\dagger D$ and its lowest eigenvalue $|\lambda_0|^2$. If $\omega_t + \omega_n < \delta_0/2$ we have

$$\|v_+ - \langle\omega_{+,0}|v_+\rangle\omega_{+,0}\| \quad \leq \quad \frac{2}{\delta_0}(\omega_t + \omega_n)\,. \tag{E.12}$$

To use (E.8) we rather need a bound for $\|v_+ - \omega_{+,0}\|$. Using simple trigonometry we find

$$\|v_+ - \omega_{+,0}\| \quad \leq \quad \frac{\|v_+ - \langle\omega_{+,0}|v_+\rangle\omega_{+,0}\|}{\sqrt{1 - \|v_+ - \langle\omega_{+,0}|v_+\rangle\omega_{+,0}\|^2}}\,. \tag{E.13}$$

provided that $\langle\omega_{+,0}|v_+\rangle$ is smaller than 1. Both relations (E.12) and (E.13) lead to the desired bound for $\|\omega_{+,0} - v_+\|$. Of course we can obtain the same relations for the vector with negative chirality. Coming back to (E.8) we finally find an error bound for the imaginary part of λ_0. It reads

$$\Big|\ |\mathrm{Im}\,\lambda_0| - |\langle v_+|D_N|v_-\rangle|\ \Big| \quad \leq \quad \|R_N\| + 4\frac{\kappa}{\sqrt{1-\kappa^2}}\,, \tag{E.14}$$

where we use the abbreviations

$$\kappa \quad = \quad \frac{2}{\delta_0}(\omega_n + \omega_t)\,, \tag{E.15}$$

$$\delta_0 \quad = \quad \min_{j\neq 0}\ \Big|\ |\lambda_0|^2 - |\lambda_j|^2\ \Big|\ . \tag{E.16}$$

Recall that v_+, v_- are the eigenvector approximations of $\omega_{+,0}$, $\omega_{-,0}$ given by the CG algorithm.

Bibliography

[1] E. Witten, *Phys. Lett.* **B 117** (1982) 324.

[2] M. Nakahara, *Geometry, Topology and Physics*, Graduate Students Series In Physics, Institute of Physics Publishing, 1990.

[3] L. O'Raiffeartaigh, *Group structure of gauge theories*, Cambridge monographs on mathematical physics, Cambridge University Press, 1986.

[4] R. A. Bertlmann, *Anomalies in Quantum Field Theory*, Oxford University Press, 1996.

[5] W. A. Bardeen, *Phys. Rev.* **184** (1969) 1848.

[6] W. A. Bardeen and B. Zumino, *Nucl. Phys.* **B 244** (1984) 421.

[7] E. Witten, *Nucl. Phys.* **B223** (1983) 422.

[8] S. Elitzur and V. P. Nair, *Nucl. Phys.* **B 243** (1984) 205.

[9] F. R. Klinkhamer, *Phys. Lett.* **B 256** (1991) 41.

[10] B. Zumino, in: Relativity, groups and topology II (Les Houches, 1983), eds. B. DeWitt and R. Stora (North Holland, Amsterdam, 1984).

[11] I. Montvay and G. Münster, *Quantum Fields on a Lattice*, Cambridge Monographs in Mathematical Physics, Cambridge University Press, 1994.

[12] H. J. Rothe, *Lattice Gauge Theories – An Introduction*, World Scientific, 1996.

[13] H. B. Nielsen and M. Ninomiya, *Phys. Lett.* **B 105** (1981) 219, *Nucl. Phys.* **B 185** (1981) 20 [E: **B 195** (1982) 541], *ibid.* **B 193** (1981) 173.

[14] K. G. Wilson, *Phys. Rev.* **D 10** (1974) 2445.

[15] P. H. Ginsparg and K. G. Wilson, *Phys. Rev.* **D 25** (1982) 2649.

[16] P. Hasenfratz, *Nucl. Phys.* **B** (Proc. Suppl.) **63A-C** (1998) 53.

[17] P. Hasenfratz, V. Laliena and F. Niedermayer, *Phys. Lett.* **B 427** (1998) 125.

[18] P. Hasenfratz, *Nucl. Phys.* **B 525** (1998) 401.

[19] M. Lüscher, *Phys. Lett.* **B 428** (1998) 342.

[20] H. Neuberger, *Phys. Lett.* **B 417** (1998) 141.

[21] H. Neuberger, *Phys. Lett.* **B 427** (1998) 353.

[22] P. Hernández, K. Jansen and M. Lüscher, *Nucl. Phys.* **B 552** (1999) 363.

[23] I. Horvath, *Phys. Rev.* **D 60** (1999) 034510.

[24] I. Horvath, *Phys. Rev. Lett.* **81** (1998) 4063.

[25] W. Bietenholz, hep-lat/9901005.

[26] F. Niedermayer, *Nucl. Phys.* **B** (Proc. Suppl.) **73** (1999) 105.

[27] M. Lüscher, *Nucl. Phys.* **B 538** (1999) 515.

[28] Y. Kikukawa and A. Yamada, *Phys. Lett.* **B 448** (1999) 265.

[29] I. Bronstein, K. Semendjajew, G. Musiol and H. Mühlig, *Taschenbuch der Mathematik*, Verlag Harri Deutsch, 1993.

[30] M. Lüscher, *Nucl. Phys.* **B 568** (2000) 162.

[31] R. Narayanan, *Phys. Rev.* **D 58** (1998) 097501.

[32] M. Lüscher, *Nucl. Phys.* **B 549** (1999) 295.

[33] J. Banks and H. Georgi, *Phys. Rev.* **D 14** (1976) 1159.

[34] M. Lüscher, An explicit example of a generator of $\pi_4(SU(2))$, private notes (unpublished).

[35] B. Bunk, K. Jansen, M. Lüscher and H. Simma, Conjugate gradient algorithm to compute the low-lying eigenvalues of the Dirac operator in lattice QCD, ALPHA collaboration internal report (unpublished), 1994.

[36] T. Kalkreuther and H. Simma, *Comp. Phys Comm.* **93** (1996) 33.

[37] W. H. Press, S. A. Teukolsy, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, 1992.

[38] H. Heuser, *Funktionalanalysis*, B. G. Teubner Stuttgart, 1992.

[39] L. Fox and I. B. Parker, *Chebyshev Polynomials in Numerical Analysis*, Oxford Mathematical Handbooks, Oxford University Press, 1968.

[40] J. F. Hart, *Computer Approximations*, The SIAM Series in Applied Mathematics, John Wiley & Sons, Inc., 1968.

[41] D. C. Handscomb, *Methods of Numerical Approximation*, Pergamon Press, 1966.

[42] F. John, *Lectures on Advanced Numerical Analysis*, Notes on Mathematics and Its Applications, Gordon and Breach, 1967.

[43] F. Stummel and K. Hainer, *Praktische Mathematik*, Teubner Studienbücher, B. G. Teubner, 1982.

Acknowledgments

In the first place I would like to thank my advisor, Prof. Dr. Martin Lüscher, for the suggestion of the topic and for numerous helpful discussions. Without him this thesis would not have been successfully completed.
I also take the opportunity to thank Isabel Campos who partly collaborated with me on the subject.
I am grateful to Arnd Brandenburg for carefully proofreading the first draft of the manuscript.
I would like to thank Daniel Lenz for clarifying discussions on exact sequences of homotopy groups.
It is a pleasure to thank my friend Robert Kirchner. In my future places I will certainly miss the stimulating and creative discussions with him.
I would like to thank all my colleagues at DESY, in particular Ulrich Nierste and York Schröder, for the pleasant social environment.
Finally I would like to thank Margarete Mühlleitner for her constant encouragement and support at all stages of the thesis.
Financial support by Deutsches Elektronen-Synchrotron (DESY) is gratefully acknowledged.