Chapter 1

Theoretical basis

This chapter is devoted to give a summary of the basic theoretical concepts used for this work. We start with shortly introducing quantum chromodynamics (QCD) as the theory of strong interactions followed by a description of its regularization by the standard Wilson lattice QCD approach, as it can be found in textbooks such as [15, 16, 17, 18, 19, 20]. Then we discuss in more detail several more recent formulations of lattice QCD and compare their properties. Aiming at computer simulations of lattice QCD we finally present methods and algorithms for this purpose.

1.1 QCD in Euclidean space-time

In this section we shortly derive the SU(3) gauge invariant QCD Lagrangian, describe its quantization by means of the path integral formalism and the transition from Minkowski to Euclidean space-time. The QCD Lagrangian is constructed such that the standard lattice regularization follows immediately.

Motivated by the demand of local gauge symmetry, we consider a continuous group of (gauge) transformations, represented by a set of SU(3) matrices $V(x)$, and a quark of flavor q, represented by a color-spinor field $\psi^{\alpha}_{c,q}(x)$ with SU(3) color index $c = 1, 2, 3$ and Dirac spinor index $\alpha = 0, 1, 2, 3$. Suppressing all these indices in the following, the fields $\psi(x)$ transform like

$$\psi(x) \quad \rightarrow \quad V(x)\psi(x)\,. \tag{1-1}$$

In order to construct a gauge invariant kinetic energy term for these quark fields, we define the gauge *covariant* derivative in direction μ by the limiting procedure:

$$D_\mu\psi(x) = \lim_{a\to 0}\frac{1}{a}[U(x, x + a\hat{\mu})\psi(x + a\hat{\mu}) - \psi(x)]\,, \tag{1-2}$$

where $\hat{\mu}$ is a unit vector in direction μ^*. For later purposes we extract the forward covariant difference operator

$$\nabla_\mu \psi(x) = \frac{1}{a}[U(x, x + a\hat{\mu})\psi(x + a\hat{\mu}) - \psi(x)]\,, \tag{1-3}$$

and define also the backward covariant difference operator

$$\nabla^*_\mu \psi(x) = \frac{1}{a}\Big[\psi(x) - U(x, x - a\hat{\mu})\psi(x - a\hat{\mu})\Big]\,. \tag{1-4}$$

In definition (1-2) we have introduced a so-called parallel transporter $U(x, y)$, a unitary SU(3) matrix, which obeys the transformation law

$$U(x, y) \quad \rightarrow \quad V(x)U(x, y)V^\dagger(y)\,, \tag{1-5}$$

and we set $U(x, x) = 1$. Since $U(x, y)$ is a SU(3) matrix and a continuous function of its arguments, for infinitesimal a it can be represented as

$$U(x, x + a\hat{\mu}) = \exp\left(-igaA^i_\mu(x + \frac{a}{2}\hat{\mu})\lambda^i + \mathcal{O}(a^3)\right)\,. \tag{1-6}$$

Here g is the bare gauge coupling, introduced for later convenience, and A^i_μ is a vector field for each generator λ^i of the transformation group. Expanding (1-6) in a and inserting it into definition (1-2) yields in the limit $a \rightarrow 0$ for the continuum covariant derivative associated with the local SU(3) gauge symmetry

$$D_\mu = \partial_\mu - igA^i_\mu\lambda^i\,, \tag{1-7}$$

which has the correct gauge transformation properties and where A^i_μ is the continuum *gauge potential*. With what we have defined so far we can already write a locally gauge invariant Lagrangian containing a kinetic energy term for the quark field $\psi(x)$ and a quark mass term. In order to complete the Lagrangian we have to find a kinetic term for the field A_μ itself. A term like this can be constructed by defining

$$\begin{aligned} U_\Box(x; \mu, \nu) =& U(x, x + a\hat{\mu})U(x + a\hat{\mu}, x + a\hat{\nu} + a\hat{\mu}) \\ & \times U(x + a\hat{\nu} + a\hat{\mu}, x + a\hat{\nu})U(x + a\hat{\nu}, x) \end{aligned} \tag{1-8}$$

as the product of the four parallel transporters at the corners of a small square in space-time, the *plaquette*. However, $U_\Box(x; \mu, \nu)$ itself is not gauge invariant, and in order to convert it to a locally gauge invariant expression we have to take the trace. Inserting (1-6) into (1-8) and using the Campbell-Baker-Hausdorff formula we achieve by expanding to order a^6

$$\mathrm{Tr}\, U_\Box(x; \mu, \nu) = \mathrm{Tr}[1] - \frac{1}{4}g^2a^4(F^i_{\mu\nu})^2 + \mathcal{O}(a^6)\,, \tag{1-9}$$

*We use the non-standard notation a (instead of for instance ϵ), because a will be identified with the lattice spacing later on.

where we introduced the field strength

$$F^i_{\mu\nu} = \partial_\mu A^i_\nu - \partial_\nu A^i_\mu + g f^{ijk} A^j_\mu A^k_\nu \,. \tag{1-10}$$

f^{ijk} are the *structure constants* of the symmetry group (see appendix A). By construction $(F^i_{\mu\nu})^2$ is locally gauge invariant. Now we can write down the QCD Lagrangian that is renormalizable, conserves parity and is invariant under time reversion [3, 4, 5]

$$\mathcal{L} = -\frac{1}{4}(F^i_{\mu\nu})^2 + \bar{\psi}(i\gamma^\mu D_\mu - m_q)\psi \,, \tag{1-11}$$

with $\bar{\psi} = \psi^\dagger\gamma_0$ being the anti-quark field. The fermionic part of Eq. (1-11) contains the covariant Dirac operator $M \equiv i\gamma^\mu D_\mu - m_q$ with quark mass m_q[†].

Eq. (1-11) exhibits the classical QCD Lagrangian. In order to quantize the theory we will use the path integral formalism introduced by Feynman [21]. The fundamental quantity in the path integral formalism is the classical action S, which is explicitly given for QCD with one quark flavor as the space-time integral over the Lagrangian density

$$S[\psi, \bar{\psi}, A] = \int d^4x \left\{ -\frac{1}{4}(F^i_{\mu\nu})^2 + \bar{\psi} M \psi \right\} . \tag{1-12}$$

The expectation value of a physical observable O is then formally given by the following functional integral:

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}A\, \mathcal{D}\psi\, \mathcal{D}\bar{\psi}\; O[\psi, \bar{\psi}, A]\; e^{iS[\psi,\bar{\psi},A]} \,, \tag{1-13}$$

with the *partition function* $\mathcal{Z}$

$$\mathcal{Z} = \int \mathcal{D}A\, \mathcal{D}\psi\, \mathcal{D}\bar{\psi}\; e^{iS[\psi,\bar{\psi},A]} \,. \tag{1-14}$$

One advantage of the path integral quantization is that it deals only with classical fields and not with operators. However, in case of fermions the (classical) fields ψ must be represented by anti-commuting variables. This is realized by the use of so called *Grassmann variables*. The integral over such Grassmann valued numbers can be defined in a sensible way and we adopt here the standard conventions (see for instance [15]).

Formally we can perform the integral over the Grassmann valued fields ψ and $\bar{\psi}$ analytically due to its Gaussian structure

$$\int \mathcal{D}\psi\, \mathcal{D}\bar{\psi}\; \exp\left\{ i \int d^4x\; \bar{\psi} M \psi \right\} \propto \det M \,, \tag{1-15}$$

[†]For simplicity we consider only one quark flavor. The generalization to N_f flavors of quarks is straightforward.

leaving us to calculate the functional determinant of the operator M[‡]. Nevertheless, we can define an effective action

$$S_{\text{eff}}[A] = \int d^4x \, \left\{ -i \log \det M - \frac{1}{4}(F^i_{\mu\nu})^2 \right\} . \tag{1-16}$$

in terms of which we can now express the expectation value of an operator O

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}A \; O[A] \; e^{iS_{\text{eff}}[A]} . \tag{1-17}$$

This is possible because due to Wick's theorem for the contraction of fields one can eliminate the time-ordered product of quark fields in the operator O by suitable factors of M^{-1} which do no longer depend on the quark fields $\bar{\psi}, \psi$.

For a numerical treatment of the functional integral in Eq. (1-17) it is useful to analytically continue the time component of the 4-vectors to purely imaginary values[§]. This rotation of the time coordinate $x^0 \to -ix^0$ leads to a Euclidean 4-vector product:

$$x^2 = (x^0)^2 - |\vec{x}|^2 \quad \to \quad -(x^0)^2 - |\vec{x}|^2 = -|x_{\text{E}}|^2 . \tag{1-18}$$

It was shown by Osterwalder and Schrader [22] that under certain conditions one can reconstruct the whole quantum field theory in Minkowski space from the Euclidean field theory. The most important condition is the so called *Osterwalder-Schrader positivity* or *reflection positivity*, which replaces Hilbert space positivity and the spectral condition of the Minkowskian formulation [22].

In Euclidean space the action is mapped to its Euclidean version

$$S \to iS \equiv -S_{\text{E}} = -\int d^4x \, \left[\frac{1}{4}(F^i_{\mu\nu})^2 + \bar{\psi}(\gamma_\mu D_\mu + m_q)\psi \right] , \tag{1-19}$$

where the Euclidean Dirac matrices γ_μ are hermitian and satisfy the anti-commutation relation

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} . \tag{1-20}$$

For the explicit representation see appendix A.1. Consequently the exponential weight, e.g. in the partition function $\mathcal{Z}$, can now be interpreted as a Boltzmann factor, as the Euclidean action is real and bounded from below (if the functional determinant $\det M$ is real)

$$\mathcal{Z}_{\text{E}} = \int \mathcal{D}\bar{\psi} \, \mathcal{D}\psi \, \mathcal{D}A \; e^{-S_{\text{E}}} . \tag{1-21}$$

Since we will work almost solely in Euclidean space-time we will skip the subscript $_{\text{E}}$ in the following and use, for convenience, t instead of x_0.

[‡]One can show that the functional determinant is equivalent to the sum of vacuum diagrams.
[§]This is often called Wick rotation.

1.2 Lattice regularization

In the following section we will describe how QCD as a quantum field theory can be regularized in Euclidean space-time by means of a discrete space-time lattice. As a regularization scheme it can be used equally well as e.g. dimensional regularization. But it has the crucial advantage to allow for non-perturbative computations by means of Monte Carlo methods.

To this end, we introduce a hyper-cubic space-time lattice with spacing a and extension $a\,L$ ($a\,T$) in the spatial directions (time direction). The boundary conditions can be chosen in different ways, and we will use periodic boundary conditions throughout this work. Only for the quark fields in time direction we use either periodic or anti-periodic boundary conditions. The parallel transporters $U(x,y)$ are SU(3)-valued and defined as in Eq. (1-6), where we use the notation $U_{x,\mu} \equiv U(x, x+a\hat{\mu})$. Note that $U_{x,-\mu} \equiv U(x, x-a\hat{\mu}) = U^{-1}_{x-a\hat{\mu},\mu}$. The set of all parallel transporters on the lattice $U \equiv \{U_{x,\mu}\}$ we call a *gauge field configuration*.

The finite lattice spacing provides an ultra-violet cutoff for the momenta. In the finite volume the allowed momenta are then given by

$$k = \pm\frac{2\pi n}{La}\,, \qquad n = 1, \ldots, L/2\,. \tag{1-22}$$

On lattice we are now able to specify what is meant by the functional integral over the gauge fields:

$$\int \mathcal{D}A \quad \equiv \quad \int \prod_{x,\mu} dU_{x,\mu}\,, \tag{1-23}$$

where the product is over all lattice points x and directions μ. Unlike in the continuum, the lattice gauge fields are SU(3) matrices with elements that are bounded in the range $[0,1]$. Therefore, as proposed by Wilson, we use the invariant *Haar measure* for the integration and adopt the standard definitions (cf. e.g. [18]).

The path integral of the Grassmann valued fermionic fields $\bar{\psi}$ and ψ is discretized similarly by

$$\int \mathcal{D}\bar{\psi}\mathcal{D}\psi \quad \equiv \quad \int \prod_{x} d\bar{\psi}(x) d\psi(x)\,. \tag{1-24}$$

1.2.1 Wilson lattice action

The next step is to discretize the action. Recalling the definition of the plaquette variable (1-8) the definition

$$S_{\mathrm{G}}[U] = \sum_{x} \sum_{1\le\mu<\nu} \beta \left\{ 1 - \frac{1}{3}\,\mathrm{Re}\,\mathrm{Tr}(U_{\Box}(x;\mu,\nu) \right\} \tag{1-25}$$

is locally gauge invariant, real and positive for $\beta \equiv 6/g^2 > 0$. It is obvious from Eq. (1-9) that it has the correct limit as $a \to 0$. $S_{\rm G}$ is called the Wilson plaquette gauge action [10] and in the following we will call β the bare coupling constant.

We are left to discretize the fermionic part of the action. With the definitions (1-3) and (1-4) of the forward and backward difference operators ∇_μ and ∇^*_μ, respectively, the lattice Dirac operator $D_{\rm W}$ proposed by Wilson in Ref. [10] can be read off from

$$\bar{\psi}\, D_{\rm W}[U]\, \psi = a^4 \sum_{x,\mu} \bar{\psi}(x) \frac{1}{2} \Big[\gamma_\mu (\nabla_\mu + \nabla^*_\mu) - a r \nabla^*_\mu \nabla_\mu \Big] \psi(x)\,. \tag{1-26}$$

The so-called Wilson term $-ar\nabla^*_\mu\nabla_\mu$ has been introduced in order to cure the problem that a naive discretization gives in the continuum limit rise to $2^{d=4} = 16$ fermion excitations rather than one (the so-called fermion doubling problem). With the Wilson term the extra 15 species pick up a mass proportional to $2r/a$ and decouple in the limit $a \to 0$. The Wilson parameter r can be chosen as $-1 \leq r \leq 1$. With a bare mass m_0 the Wilson lattice action then reads

$$S[\psi, \bar{\psi}, U] = S_{\rm G}[U] + \bar{\psi}(D_{\rm W}[U] - m_0)\psi\,. \tag{1-27}$$

However, while curing the fermion doubling problem, the Wilson term explicitly breaks chiral symmetry at finite values of the lattice spacing, i.e. even the massless Wilson-Dirac operator no longer anti-commutes with γ_5. Although chiral symmetry is expected to be recovered in the continuum, its breaking at finite values of a has important consequences, among others:

- It introduces discretization errors of $\mathcal{O}(a)$. With a naive discretization the lattice artifacts in the action appear only at $\mathcal{O}(a^2)$.
- The quark mass renormalizes both additively and multiplicatively. Hence, it is useful to define a subtracted bare quark mass parameter by

  $$m_q = m_0 - m_{\rm crit}\,, \tag{1-28}$$

  where $m_{\rm crit}$ is called the critical mass parameter. The value of $m_{\rm crit}$ needs to be determined such that the chiral point is properly defined, for instance by demanding that the lightest pseudo scalar state becomes massless at this point.

Both of these two consequences will play an important rôle in this work.

We note in passing that usually in the simulations the *hopping parameter representation* of the Wilson lattice action (1-27) is used. It is obtained from (1-27) by re-scaling the fermionic fields as follows:

$$\psi \;\to\; \frac{\sqrt{2\kappa}}{a^{3/2}}\psi\,, \qquad \bar{\psi} \;\to\; \frac{\sqrt{2\kappa}}{a^{3/2}}\bar{\psi}\,, \qquad \kappa = \frac{1}{2am_0 + 8r}\,, \tag{1-29}$$

with the so called *hopping parameter* κ. To m_{crit} corresponds the critical hopping parameter

$$\kappa_{\text{crit}} = \frac{1}{2am_{\text{crit}} + 8r}\,, \tag{1-30}$$

and the fermionic part of the action reads

$$\begin{aligned} S[\psi,\bar{\psi},U] = \sum_x \Bigg\{ & \bar{\psi}(x)\psi(x) - \kappa\bar{\psi}(x)\sum_{\mu=1}^{4}\Big[U_{x,\mu}(r+\gamma_\mu)\psi(x+a\hat{\mu}) \\ & + U^\dagger_{x-a\hat{\mu},\mu}(r-\gamma_\mu)\psi(x-a\hat{\mu})\Big]\Bigg\} \\ \equiv & \sum_{x,y}\bar{\psi}(x)M_{xy}\psi(y)\ . \end{aligned} \tag{1-31}$$

1.2.2 Remnant chiral symmetry on the lattice

If we consider massless continuum QCD with two quark flavors $u \equiv \psi^\alpha_{c,u}$ and $d \equiv \psi^\alpha_{c,d}$, the Lagrangian obeys isospin symmetry, the symmetry of an SU(2) unitary transformation mixing the u and d fields. But, since left- and right-handed quarks do not couple, this Lagrangian is actually symmetric under the separate unitary transformations

$$\begin{pmatrix}u\\d\end{pmatrix}_L \equiv \frac{1-\gamma_5}{2}\begin{pmatrix}u\\d\end{pmatrix} \rightarrow U_L\begin{pmatrix}u\\d\end{pmatrix}_L, \qquad \begin{pmatrix}u\\d\end{pmatrix}_R \equiv \frac{1+\gamma_5}{2}\begin{pmatrix}u\\d\end{pmatrix} \rightarrow U_R\begin{pmatrix}u\\d\end{pmatrix}_R. \tag{1-32}$$

We can separate the U(1) and the SU(2) parts of these transformations: the symmetry of the classical Lagrangian is $\mathrm{U}_V(1)\times\mathrm{U}_A(1)\times\mathrm{SU}_V(2)\times\mathrm{SU}_A(2)$. The vector part of this symmetry is a manifest symmetry of strong interactions, but the $\mathrm{U}_A(1)$ is broken by instanton contributions and the $\mathrm{SU}_A(2)$ symmetry is spontaneously broken. These two statements imply that the flavor singlet axial current

$$j^{\mu 5} = (\bar{u}\ \bar{d})\gamma^\mu\gamma^5\begin{pmatrix}u\\d\end{pmatrix} \tag{1-33}$$

is anomalous (Adler-Bell-Jackiw anomaly) and that the quark condensate acquires a non-zero expectation value in the QCD vacuum. Due to Goldstone's theorem we expect three massless particles associated with the spontaneously broken symmetries, which are in nature realized as the pion triplet.

Of course, if the quarks are not exactly massless, the isotriplet axial currents are no longer exactly conserved. Then, the quark mass terms give the pions masses of the form (see e.g. Ref. [15])

$$m_\pi^2 = \frac{M^2}{f_\pi}(m_u + m_d)\,, \tag{1-34}$$

with a mass parameter M and f_π the *pion decay constant* with dimension of a mass. The value of M has been estimated to be of the order of 400 MeV, the value of f_π is 93 MeV [23]. To give the pions the observed 140 MeV [23] one therefore needs $(m_u + m_d) \sim 10$ MeV, which is indeed a small, but important perturbation to the massless case.

Chiral symmetry plays an important rôle in continuum QCD and it is therefore desirable to preserve chiral symmetry also in lattice QCD at finite values of the lattice spacing a. But, we have seen already in the discussion around Eq. (1-26) that for the standard Wilson operator one has to live either with the doubling problem or with explicitly broken chiral symmetry.

In fact, it was proven by Nielsen and Ninomiya [24, 25, 26, 27] that for a massless lattice Dirac operator D it is not possible to achieve the following properties at the same time:

1. $D(x)$ is an essentially local operator (bounded by $e^{-\gamma|x|}$).
2. The Fourier transform of $D(x)$ fulfills $\tilde{D}(p) = i\gamma_\mu p_\mu + \mathcal{O}(ap^2)$ for $p \ll \pi/a$.
3. $\tilde{D}(p)$ is invertible for $p \neq 0$. (no massless doublers).
4. The lattice action is invariant under continuous chiral transformations, i.e. $\gamma_5 D + D\gamma_5 = 0$.

In order to circumvent this theorem it was proposed by Ginsparg and Wilson [28] to replace the property that D anti-commutes with γ_5 by the so-called Ginsparg-Wilson relation:

$$\gamma_5 D + D\gamma_5 = aD\gamma_5 RD\,, \tag{1-35}$$

where R is a local matrix that may have a nontrivial γ-matrix dependence but must have a chirally non-invariant piece.

The Ginsparg-Wilson relation implies a continuous symmetry of the fermionic lattice action, as was shown by Lüscher [29]. The infinitesimal variation corresponding to this symmetry reads

$$\delta\psi = \gamma_5(1 - \frac{1}{2}aD)\psi\,, \qquad \delta\bar{\psi} = \bar{\psi}(1 - \frac{1}{2}aD)\gamma_5\,, \tag{1-36}$$

which is a flavor singlet chiral transformation on the lattice. The flavor non-singlet transformations can be defined similarly by including a group generator in Eq. (1-36). An operator D fulfilling the Ginsparg-Wilson relation (1-35) is called a *Ginsparg-Wilson operator*.

Since the flavor singlet chiral symmetry in QCD is anomalous in the presence of gauge fields, it is interesting to see whether this anomaly is correctly reproduced

when the lattice Dirac operator fulfills the relation (1-35). In fact it was shown in Ref. [29] that the Ward identities associated with the global flavor singlet chiral transformations on the lattice have the correct anomaly. Moreover, the discussion in Ref. [29] reveals that flavor non-singlet chiral rotations are exact symmetries of the lattice theory. In addition the Atiyah-Singer index theorem [30] holds at finite values of the lattice spacing [31] with the important consequence that the difference of the number of positive chirality zero modes and negative chirality zero modes of a Ginsparg-Wilson Dirac operator in an external color gauge field is equal to the topological charge.

Furthermore, like in continuum QCD, the lattice chiral symmetry prohibits mixing between operators with different chirality. The latter proves to be rather useful in the course of renormalization, in particular in the calculation of matrix elements of the weak interactions Hamiltonian needed for instance for the extraction of B_K.

One particular solution of the Ginsparg-Wilson relation was found by Neuberger [32, 33] and is given by the so called *overlap operator*. For $R_{xy} = \delta_{xy}/\rho$ the massless operator reads

$$D_{\mathrm{ov}}^{(0)} = \frac{\rho}{a}\left[1 - A(A^\dagger A)^{-1/2}\right], \qquad A = \rho - aD_{\mathrm{W}}\,, \tag{1-37}$$

where D_{W} is the Wilson-Dirac operator (1-26) and $0 < \rho < 2$ is a real parameter. A bare mass m_{ov} can be added in the following way

$$D_{\mathrm{ov}} = \left(1 - m_{\mathrm{ov}}\frac{a}{2\rho}\right) D_{\mathrm{ov}}^{(0)} + m_{\mathrm{ov}}\,, \tag{1-38}$$

where the somewhat un-usual form avoids $\mathcal{O}(a)$ lattice artifacts [34]. The overlap operator was shown to be manifestly gauge covariant and it has no doublers. Moreover it was shown to converge to the expected classical continuum expression (up to a finite normalization constant) and the requirement of locality is fulfilled with a certain choice of the parameter ρ [35]. Note that an equivalent formulation is provided by the *domain wall* approach [36, 37] (for a review see [38]).

However, from a feasibility point of view the overlap operator is compared to the Wilson-Dirac operator computer time demanding. The reason for this is the fact that for each application of D_{ov} an evaluation of $1/\sqrt{A^\dagger A}$ is needed. This is usually done with a polynomial in the operator $A^\dagger A$ and needs therefore a certain number of applications of the latter. This number depends naturally on the parameters under investigation and might very well be of $\mathcal{O}(100)$.

It is also part of this work to compare the computational costs for two particular lattice Dirac operators – one of those being the overlap operator.

1.2.3 Symanzik effective action

We have explained in the last paragraphs that the Wilson formulation of lattice QCD breaks chiral symmetry and shows discretization errors of $\mathcal{O}(a)$, which is not a problem of principle if one is able to simulate at small enough values of the lattice spacing a. Unfortunately, nowadays lattice QCD simulations are still restricted to rather large values of a and it is therefore worthwhile to improve the scaling with the lattice spacing towards the continuum limit.

We will present here two ways to achieve on-shell $\mathcal{O}(a)$ improvement, both of them are based on the effective action introduced by Symanzik [12, 13, 14].

Close to the continuum the lattice theory as has been defined in the previous paragraphs can be described by a local effective action

$$S_{\text{eff}} = S_0 + aS_1 + a^2 S_2 + \dots \tag{1-39}$$

While the leading term, S_0, is just the continuum action[¶], the correction terms S_k are to be interpreted as operator insertions in the continuum theory and are given by

$$S_k = \int d^4x \mathcal{L}_k(x)\,. \tag{1-40}$$

The Lagrangians $\mathcal{L}_k$ have mass dimension $4+k$ and they are linear combinations of local composite fields. The list of possible fields is constraint by gauge and flavor symmetry and the exact discrete symmetries of the lattice action, including space-time lattice symmetries and charge conjugation. Moreover partial integration can be used to further reduce the number of possible terms.

Clearly the action is not the only origin for cut-off effects, also the local composite fields, from which observables of interest are built, will be a source of those. Consider a local gauge invariant composite field $\phi(x)$ on the lattice, constructed out of gluon and fermion fields. For simplicity we assume that it does not mix with other operators under renormalization. The effective field

$$\phi_{\text{eff}}(x) = \phi_0(x) + a\phi_1(x) + a^2\phi_2(x) + \dots \tag{1-41}$$

represents in the effective theory the renormalized lattice field $Z_\phi\phi(x)$ with an appropriately chosen renormalization factor Z_ϕ. The fields ϕ_k that appear in the representation (1-41) are linear combinations of local fields with appropriate dimension and symmetry properties. To leading order in the lattice spacing a connected lattice n-point function with all points $x_1, \dots, x_n$ kept at non zero distance from

[¶]The continuum theory also has to be regularized to make the expressions meaningful. One could think e.g. of a regularization with a much smaller lattice spacing $\tilde{a} \ll a$.

each other is given by

$$\begin{aligned} G_n(x_1,\ldots,x_n) = {} & \langle \phi_0(x_1)\ldots\phi_0(x_n)\rangle^{\text{cont}} \\ & - a\int d^4y\ \langle \phi_0(x_1)\ldots\phi_0(x_n)\mathcal{L}_1(y)\rangle^{\text{cont}} \\ & + a\sum_{k=1}^{n}\langle \phi_0(x_1)\ldots\phi_1(x_k)\ldots\phi_0(x_n)\rangle^{\text{cont}} + \mathcal{O}(a^2)\,. \end{aligned} \tag{1-42}$$

All the expectation values on the right hand side of Eq. (1-42) have to be evaluated in the continuum theory with action S_0, which we indicate by the superscript $^{\text{cont}}$. In the second term, which is the contribution of the $\mathcal{O}(a)$ term in the effective action, the integral over y in general diverges at the points $y = x_k$. However, each such contact term can be described as the effect of a local field placed at x_k. This local field must have the global symmetry quantum numbers of $\phi_0(x_k)\mathcal{L}(y)$ and therefore the possible fields are known from ϕ_{eff}. Hence, the contact terms lead just to a redefinition of the field $\phi_1(x)$ and the renormalization factor.

In the following we will assume that the lattice action is the Wilson action. For this case in Ref. [39] the $\mathcal{O}(a)$ effective Lagrangian $\mathcal{L}_1(x)$ was derived to be a linear combination of the following fields

$$\begin{aligned} O_1 &= \bar\psi\sigma_{\mu\nu}F_{\mu\nu}\psi\,, \\ O_2 &= \bar\psi D_\mu D_\mu\psi + \bar\psi\overleftarrow{D_\mu}\overleftarrow{D_\mu}\psi\,, \\ O_3 &= m\,\mathrm{Tr}\{F_{\mu\nu}F_{\mu\nu}\} \\ O_4 &= m\{\bar\psi\gamma_\mu D_\mu\psi - \bar\psi\overleftarrow{D_\mu}\gamma_\mu\psi\} \\ O_5 &= m^2\bar\psi\psi\,, \end{aligned} \tag{1-43}$$

where m is the quark mass and $F_{\mu\nu}$ the field-strength tensor and D_μ is here the gauge covariant partial derivative, see Ref. [39]. We do not consider the fields O_2 and O_4, because they can be eliminated by relations derived from the classical field equations. Note that in order to apply the field equations to simplify the effective Lagrangian one has to carefully treat some contact terms. However, these contact terms can be shown to amount only for a redefinition of the fields [39], see above.

Thus, aiming to improve the lattice action by adding a suitable counter-term of $\mathcal{O}(a)$ to the Wilson action, one has to add a counter-term of the form

$$a^5\sum_x\Big[c_1\hat O_1(x) + c_3\hat O_3(x) + c_5\hat O_5(x)\Big]\,, \tag{1-44}$$

where $\hat O_k$ is some lattice representation of the field O_k. Using the leftover ambiguity of $\mathcal{O}(a^2)$ we may represent the fields $\mathrm{Tr}\,F_{\mu\nu}F_{\mu\nu}$ and $\bar\psi\psi$ by the Wilson plaquette field and the local scalar density, respectively. Since these two already appear in

the Wilson action the counter-terms proportional to $\hat{O}_3$ and $\hat{O}_5$ amount to a renormalization of the bare coupling and the bare mass. Let us remark here that the latter prescription to absorb the counter-terms of O_3 and O_5 leads to the additional complication of a mass dependent renormalization scheme, which is usually not favorable. One can cope with this complication [39], but we will not go into these details since we are not going to use this improvement programme further.

Therefore, for the on-shell $\mathcal{O}(a)$ improved action one needs a counter-term

$$\delta S[U,\bar{\psi},\psi] = a^5 \sum_x c_{\rm sw}\bar{\psi}(x)\frac{i}{4}\sigma_{\mu\nu}\hat{F}_{\mu\nu}(x)\psi(x)\,, \tag{1-45}$$

where $\hat{F}_{\mu\nu}$ is a lattice representation of the field strength tensor $F_{\mu\nu}$ and $c_{\rm sw}$ is a tunable coefficient[‖]. The coefficient $c_{\rm sw}$ needs to be determined in a suitable way in order to obtain an order a improved theory on the lattice.

Depending on the possible fields ϕ_1 it might then also be necessary to determine further improvement coefficients for the field under consideration. One example is the isovector axial current

$$A^\alpha_\mu(x) = \bar{\psi}(x)\gamma_\mu\gamma_5\frac{1}{2}\tau^\alpha\psi(x)\,,$$

that needs one improvement coefficient $c_{\rm A}$.

Finally, we want to mention that, since O_1 transforms like a mass term under chiral rotations [41] and therefore explicitly breaks chiral symmetry, it cannot appear if the operator obeys the Ginsparg-Wilson relation: the reason for this is that in the chiral limit there is exact chiral symmetry on the lattice and thus all correction terms in the action not proportional to some powers of the mass must be absent in the effective action [34]. Moreover, with a properly introduced mass term as in Eq. (1-38) terms proportional to am are forbidden due to the combined symmetry of $[m \to -m]$ and

$$\psi \to \gamma_5(1-aD)\psi\,, \qquad \bar{\psi} \to -\bar{\psi}\gamma_5\,.$$

Thus a lattice regularization with exact chiral symmetry shows no $\mathcal{O}(a)$ lattice artifacts. This, of course, provides an additional advantage of chirally symmetric lattice formulations compared to $\mathcal{O}(a)$ improvement by means of the Symanzik improvement programme: in the latter a (potentially large) number of improvement coefficients has to be tuned, while there is still no control about higher order lattice artifacts.

1.2.4 Twisted mass regularization

Having introduced in the last subsection the concept of Symanzik's effective action and one way to obtain on-shell $\mathcal{O}(a)$ improvement – that amounts in determining

[‖]The improved action was first obtained by Sheikholeslami and Wohlert [40].

improvement coefficients and adding of counter-terms to the lattice action – we will in this subsection follow a different approach, recently realized in Ref. [42], where only one parameter needs to be tuned in order to obtain $\mathcal{O}(a)$ improved results.

To this end we introduce the so called *twisted mass* regularization of lattice QCD (tmQCD) [43, 44]. The twisted mass fermion lattice action for $N_{\rm f} = 2$ flavors of mass degenerate quarks reads

$$S_{\rm tm} = a^4 \sum_x \bar{\chi}(x)\big[D_{\rm W} + m_0 + i\mu\gamma_5\tau_3\big]\chi(x) = a^4 \sum_x \bar{\chi}(x) D_{\rm tm}\chi(x)\,, \tag{1-46}$$

where μ is refered to as the *twisted mass parameter* and τ_3 is the third Pauli matrix acting in flavor space. The twisted mass Dirac operator $D_{\rm tm}$ is given by

$$D_{\rm tm} = D_{\rm W} + m_0 + i\mu\gamma_5\tau_3\,. \tag{1-47}$$

We denote the fermion fields now by χ and call them the twisted basis for a reason that will become clear later. In Ref. [43] this formulation was introduced to avoid un-physically small eigenvalues of the lattice Dirac operator, since the twisted mass serves as an infrared cutoff for the eigenvalues of the operator $D_{\rm tm}$, and it was shown that this particular regularization is equivalent to the standard Wilson regularization up to cut-off effects [43].

In fact the continuum limit of the action (1-46) reads

$$S^{\rm cont}_{\rm tm} = \int d^4x\ \bar{\chi}(x)\big[\gamma_\mu D_\mu + m_q + i\mu\gamma_5\tau_3\big]\chi(x)\,, \tag{1-48}$$

which is form-invariant under axial transformations

$$\chi\ \rightarrow\ e^{i\omega\gamma_5\tau_3/2}\chi\,, \qquad \bar{\chi}\ \rightarrow\ \bar{\chi}\ e^{i\omega\gamma_5\tau_3/2}\,. \tag{1-49}$$

The form (1-48) can be obtained from the standard continuum fermion action by rotating the fermion fields ψ with *twist angle* ω according to

$$\psi\ \rightarrow\ \chi = e^{i\omega\gamma_5\tau_3/2}\psi\,, \qquad \bar{\psi}\ \rightarrow\ \bar{\chi} = \bar{\psi}\ e^{i\omega\gamma_5\tau_3/2}\ . \tag{1-50}$$

Note that the axial transformations (1-50) are non-anomalous: they leave the fermion measure invariant since $\mathrm{Tr}\,\tau_3 = 0$. The latter rotations just transform the mass parameters according to

$$\begin{aligned} m_q\ &\rightarrow\ m_q\cos\omega + \mu\sin\omega\,, \\ \mu\ &\rightarrow\ -m_q\sin\omega + \mu\cos\omega\,, \end{aligned} \tag{1-51}$$

and the mass term in Eq. (1-48) can also be written as $m\exp(-i\omega\gamma_5\tau_3)$ with $m^2 = m_q^2 + \mu^2$. One form of particular interest – which will become clear later – is the special case with $m_q = 0$

$$S^{\rm cont}_{\rm mtm} = \int d^4x\ \bar{\chi}(x)\big[\gamma_\mu D_\mu + i\mu\gamma_5\tau_3\big]\chi(x)\,, \tag{1-52}$$

which is related to the standard action by an axial transformation with angle $\omega = \pi/2$. It is refered to as the action with maximal twist. The corresponding lattice action has the form

$$S_{\rm mtm} = a^4 \sum_x \bar{\chi}(x)\big[D_{\rm W} + m_{\rm crit} + i\mu\gamma_5\tau_3\big]\chi(x)\,, \tag{1-53}$$

since the Wilson term $-ar\nabla^*_\mu\nabla_\mu$ still breaks chiral symmetry explicitly and therefore additive renormalization of m_0 is needed. We refer to this particular lattice regularization as *maximally twisted mass lattice QCD* (mtmQCD).

Since the transition from standard to twisted form of the action corresponds to a change of fermionic variables $\psi \to \chi$, the χ basis is called *twisted basis* and the ψ basis is called *physical basis*. The lattice Wilson twisted mass action (1-46) in the twisted basis can be translated to the physical basis again by an axial transformation of the form of Eq. (1-49) and it reads

$$\begin{aligned} S^{\rm ph}_{\rm tm} = a^4 \sum_x \bar{\psi}(x) \Bigg[& \frac{1}{2}\sum_\mu \gamma_\mu(\nabla^*_\mu + \nabla_\mu) \\ & + \left(-r\frac{a}{2}\sum_\mu \nabla^*_\mu\nabla_\mu + m_{\rm crit}\right) e^{-i\omega\gamma_5\tau_3} + m_q \Bigg] \psi(x)\,. \end{aligned} \tag{1-54}$$

In this form the fermion mass term is real and, since the Wilson term is not left invariant under the axial rotation, now the Wilson term is rotated. We remark that the tmQCD lattice action can easily be translated to the hopping parameter representation by re-scaling according to Eq. (1-29), as it was done for the Wilson lattice action with Eq. (1-31).

Finally, as in the continuum the tmQCD and the standard QCD are exactly related by the transformation (1-50), they share all the symmetries. But in the twisted basis the symmetry transformations can have a different form than usual. For instance the usual parity operation

$$\mathcal{P}: \qquad \begin{cases} U_0(\vec{x},t) & \to \quad U_0(-\vec{x},t) \\ U_k(\vec{x},t) & \to \quad U_k^\dagger(\vec{x}-a\hat{k},t) \\ \chi(\vec{x},t) & \to \quad \gamma_0\chi(-\vec{x},t) \\ \bar{\chi}(\vec{x},t) & \to \quad \bar{\chi}(-\vec{x},t)\gamma_0 \end{cases} \tag{1-55}$$

is no longer a symmetry of the continuum action in the twisted basis. But if $\mathcal{P}$ is replaced by the following modified parity operation $\tilde{\mathcal{P}}$ [44], the symmetry is recovered:

$$\tilde{\mathcal{P}}: \qquad \begin{cases} \chi(\vec{x},t) & \to \ \gamma_0 \exp(i\omega\gamma_5\tau_3)\chi(-\vec{x},t) \\ \bar{\chi}(\vec{x},t) & \to \ \bar{\chi}(-\vec{x},t)\exp(i\omega\gamma_5\tau_3)\gamma_0\ , \end{cases} \tag{1-56}$$

where ω is defined by Eq. (1-51) and the gauge fields transform under $\tilde{\mathcal{P}}$ as for $\mathcal{P}$. However, for the lattice action in the twisted basis $\mathcal{P}$ is only a symmetry if it is combined with $[\mu \to -\mu]$ or a flavor exchange [44]. The former combination we will denote with $\mathcal{P}_\mu$ and the latter with $\mathcal{P}_F$. Moreover, on the lattice the isospin symmetry is explicitly broken in the twisted mass formulation even in the mass degenerate case.

$\mathcal{O}(a)$ improvement at maximal twist

The mtmQCD formulation is of particular interest, because one can show that terms proportional to $\mathcal{O}(a)$ are absent in the Symanzik expansion of certain physical observables, as was proven in Ref. [42]. In order to see this we will follow in this work a different approach than originally used in Ref. [42]. It is similar to a proof in the physical basis presented in Ref. [45] and based on the following observations:

First, it is important to notice that all the expectation values in the Symanzik expansion of an operator O on the right hand side of Eq. (1-42) are to be taken in the continuum theory with action S_0. This implies that any operator not obeying the symmetries of S_0 must have zero expectation value and therefore be absent in the expansion. The symmetry we will use is the just introduced modified parity operation $\tilde{\mathcal{P}}$.

Second, all the operators in the expansion of the operator O must share all its lattice symmetries, otherwise they must be absent. The particular symmetry of the mtmQCD lattice action is $\tilde{\mathcal{P}} \times \mathcal{D}_m \times \mathcal{D}_d$. The transformation according to $\mathcal{D}_m \times \mathcal{D}_d$ is multiplying each term with $(-1)^{d_m+d_d}$. d_m represents the mass dimension of the term and d_d its normal dimension.

And last, the fields contributing to $\mathcal{L}_1$ are restricted to those that obey the symmetries of the lattice action, as for instance $\mathcal{P}_\mu$ and $\mathcal{P}_F$.

By using these arguments we will now show that in the expansion of an operator O even under $\tilde{\mathcal{P}}$, which means O goes exactly into itself under a $\tilde{\mathcal{P}}$ transformation, terms linear in a are absent. To this end we first have to accumulate the action counter-terms needed for $\mathcal{L}_1$: in addition to the fields listed in Eq. (1-43) one finds among others the following:

$$\begin{aligned} O_6 &= \mu^2 \bar{\chi}\chi\,, \\ O_7 &= \mu\,\mathrm{Tr}\{F_{\mu\nu}F_{\mu\nu}\}\,, \\ O_8 &= im\mu\bar{\chi}\gamma_5\tau_3\chi\,, \\ O_9 &= m\mu\bar{\chi}\chi\,, \\ O_{10} &= \mu\,\mathrm{Tr}\{F_{\mu\nu}\tilde{F}_{\mu\nu}\}\,. \end{aligned} \tag{1-57}$$

We listed some more fields than are actually important in order to show how they can be removed by the help of the symmetries $\mathcal{P}_\mu$ and $\mathcal{P}_F$. For instance O_{10} is parity

odd, but not affected by a flavor exchange and is therefore absent in $\mathcal{L}_1(x)$ due to $\mathcal{P}_F$. O_9 and O_7 are also absent, because they are odd under $\mathcal{P}_\mu$. At maximal twist m is identical zero and hence we are left with O_6 in addition to O_1 (with ψ replaced by χ), where we again used the classical equations of motion to reduce the number of possible fields.

Consider now an operator O even under $\tilde{\mathcal{P}}$. In the second term of its Symanzik expansion (1-42), containing an insertion of $\mathcal{L}_1$, we have to evaluate the expectation value of the $\tilde{\mathcal{P}}$ even operator O and the fields O_i listed above: $\langle OO_i\rangle^{\text{cont}}$. O_1 and O_6 are odd under $\tilde{\mathcal{P}}$, O is even and thus these terms are absent in the expansion.

Since this means that lattice artifacts originating from the lattice action are absent we can conclude already that quantities extracted from the transfer matrix, like for instance hadron masses, are not affected by $\mathcal{O}(a)$ cut-off effects at maximal twist.

In order to show that also the third term on the right hand side of Eq. (1-42) is absent for $\tilde{\mathcal{P}}$ even operators we use the aforementioned symmetry $\tilde{\mathcal{P}} \times \mathcal{D}_m \times \mathcal{D}_d$. The operator O is even under $\tilde{\mathcal{P}}$ and must therefore be even under $\mathcal{D}_m \times \mathcal{D}_d$. As the third term in Eq. (1-42) is multiplied with one power of a it must be odd under $\mathcal{D}_m \times \mathcal{D}_d$ and therefore odd under $\tilde{\mathcal{P}}$. Since O is even under $\tilde{\mathcal{P}}$ also this term must be absent in the expansion.

This is a remarkable result, because with tmQCD at maximal twist it is possible to avoid un-physically small eigenvalues and to obtain $\mathcal{O}(a)$ improved expectation values needing no improvement coefficients**. Quantities even in $\tilde{\mathcal{P}}$ are for instance hadron masses and on-shell matrix elements at zero three-momentum. More examples can be found in Ref. [42].

Overlap versus mtmQCD

We will close this section with a short comparison of the overlap lattice approach to the twisted mass lattice formulation at maximal twist. Both formulations allow to extract $\mathcal{O}(a)$ improved results and the two corresponding lattice Dirac operators are protected against un-physically small eigenvalues [43]. Of course, one has to keep in mind that for the overlap formulation $\mathcal{O}(a)$ improvement comes automatic while for mtmQCD the tuning of the twist angle is required. The first difference of principle is the fact that the overlap operator exhibits exact chiral symmetry on the lattice, while for mtmQCD chiral properties are only improved, since the symmetry is still explicitly broken, even if the effects become visible only at $\mathcal{O}(a^2)$. But the price for exact chiral symmetry on the lattice is that the cost for one application of

**One can show [42] that $\mathcal{O}(a)$ improvement for parity even operators can be achieved also for the Wilson formulation, when one averages over independent simulations with positive and negative value of the Wilson parameter r or with positive and negative value of m_q.

the overlap Dirac operator is significantly larger than the cost for the application of the twisted mass operator.

We note in passing that also with mtmQCD it is possible to avoid unwanted mixing of operators with opposite chirality, although chiral symmetry is broken. For details we refer to Ref. [46].

1.3 Observables

In the preceding sections we have discussed several formulations of lattice QCD and in particular, how one can obtain $\mathcal{O}(a)$ improvement. So far we did not yet explain how physical quantities like masses and decay constants can be extracted in lattice simulations, which is the content of the following section. Since we will work with three different formulations, the pure Wilson, the tmQCD and the overlap formulation, we have to consider in this section all of the three formulations. The composite fields and correlation functions for the Wilson formulation can equivalently be used in the tmQCD formulation, they only differ by the fact that for tmQCD we work in the twisted basis, where all the fermion fields are rotated. There are, however, some specialties in the case of maximal twist. We will discuss the three formulations in separate subsections.

Before coming to the two point functions needed in this work it is useful to discuss the relation between Euclidean and Minkowski quantities. Consider the following Euclidean time ordered two point function

$$\langle 0|T[O_1(x)O_2(0)]|0\rangle\,, \tag{1-58}$$

with operators O_1 and O_2, representing the amplitude for the creation of a state with quantum numbers of operator O_2 at space-time point 0, its propagation to space-time point $x = (\vec{x}, t)$ and finally its annihilation by O_1. By integrating over space-like coordinates we project to zero momentum states. If we then insert a complete set of energy eigenstates we obtain

$$\langle 0|\int d^3x\; O_1(x)O_2(0)|0\rangle = \sum_n \frac{\langle 0|O_1|n\rangle\langle n|O_2|0\rangle}{2E_n} e^{-E_n t}\,. \tag{1-59}$$

If there is a stable single-particle state $|n\rangle$ with the corresponding quantum numbers, then its Energy E_n is equal to the mass M_n of the particle, since we have projected to zero momentum. In general there might be several such states, but asymptotically for large enough values of t the correlation function will be dominated by the state with the lowest mass, e.g. M_1. Thus one gets

$$\langle 0|\int d^3x\; O_1(x)O_2(0)|0\rangle \xrightarrow{t\,\to\,\infty} \frac{\langle 0|O_1|1\rangle\langle 1|O_2|0\rangle}{2M_1} e^{-M_1 t}\,, \tag{1-60}$$

and the mass of the lightest particle can be extracted from the exponential fall-off of the correlation function at large Euclidean times without analytical continuation to Minkowski space. This simple connection also holds for matrix elements. These arguments, however, break down as soon as the sum over states is not just over stable single-particle states. For instance the ρ meson mass can no longer be extracted by simply measuring the ground state energy (cf. [47, 48]).

1.3.1 Wilson formulation

Hadron and quark masses are – in the simplest cases – extracted from two point correlation functions of suitable composite fields. The most common bilinears are the scalar and the pseudo scalar densities, which read in the physical basis ψ:

$$S^0(x) = \bar{\psi}(x)\psi(x)\,, \qquad P^\alpha(x) = \bar{\psi}(x)\gamma_5\frac{\tau_\alpha}{2}\psi(x)\,, \tag{1-61}$$

and the vector and the axial currents

$$\begin{aligned} A^\alpha_\mu(x) &= \bar{\psi}(x)\gamma_\mu\gamma_5\frac{\tau_\alpha}{2}\psi(x)\,, \\ V^\alpha_\mu(x) &= \bar{\psi}(x)\gamma_\mu\frac{\tau_\alpha}{2}\psi(x)\,. \end{aligned} \tag{1-62}$$

Here τ_α, $\alpha = 1, 2, 3$ are the usual Pauli matrices acting in isospin space. Moreover the local bilinear with tensor structure is of interest

$$T^\alpha_k(x) = \bar{\psi}(x)\gamma_0\gamma_k\frac{\tau_\alpha}{2}\psi(x)\,. \tag{1-63}$$

All local interpolating field operators for mesons in a Wilson like theory can be found in table 1.1. In practice a general two point correlator with three-momentum $\mathbf{p}$ can be re-written in terms of the quark propagators as follows

$$\begin{aligned} C_{AB}(\mathbf{p}, t) &= \sum_{\mathbf{x}} e^{-ipx}\langle\bar{\psi}_2(x)\Gamma_B\psi_1(x)\bar{\psi}_1(0)\Gamma_A\psi_2(0)\rangle \\ &= -\sum_{\mathbf{x}} e^{-ipx}\langle\mathrm{Tr}\,(S_2(0,x)\Gamma_B S_1(x,0)\Gamma_A)\rangle\,. \end{aligned} \tag{1-64}$$

Here $\Gamma_{A,B}$ represent the Dirac structure, the quark propagators are denoted with $S_{1,2}$ and the trace is taken over spin and color indices. The sum over the space-like points $\mathbf{x}$ on a given time slice projects to zero momentum states. In case the quark propagators fulfill the hermiticity property $S(0,x) = \gamma_5 S(x,0)^\dagger\gamma_5$ the computation of both $S(0,x)$ and $S(x,0)$ can be avoided in favor of computing only one of the two.

As explained at the beginning of this section the meson masses can then be extracted from the exponential fall-off of suitable correlation functions at large Euclidean times. In terms of the local interpolating fields introduced above for instance

State	$I^G(J^{PC})$	Operator
Scalar	$1^-(0^{++})$	$\bar{u}(x)d(x)$
	$1^-(0^{++})$	$\bar{u}(x)\gamma_0 d(x)$
Pseudo-scalar	$1^-(0^{-+})$	$\bar{u}(x)\gamma_5 d(x)$
	$1^-(0^{-+})$	$\bar{u}(x)\gamma_5\gamma_0 d(x)$
Vector	$1^+(1^{--})$	$\bar{u}(x)\gamma_i d(x)$
	$1^+(1^{--})$	$\bar{u}(x)\gamma_0\gamma_i d(x)$
Axial	$1^+(1^{++})$	$\bar{u}(x)\gamma_5\gamma_i d(x)$
Tensor	$1^+(1^{+-})$	$\bar{u}(x)\gamma_i\gamma_j d(x)$

Table 1.1: The local interpolating field operators for light mesons in a Wilson like theory. The quark fields corresponding to the two quark flavors *up* and *down* are labeled with u and d.

the pseudo scalar mass m_{PS} and the vector meson mass m_{V} can be extracted from the following two point correlation functions:

$$C^{\alpha}_{PP}(t) = a^3 \sum_{\mathbf{x}} \langle P^{\alpha}(x) P^{\alpha}(0) \rangle \qquad \alpha = 1, 2\,, \tag{1-65}$$

$$C^{\alpha}_{VV}(t) = \frac{a^3}{3} \sum_{k=1}^{3} \sum_{\mathbf{x}} \langle V^{\alpha}_k(x) V^{\alpha}_k(0) \rangle \qquad \alpha = 1, 2\,, \tag{1-66}$$

$$C^{\alpha}_{TT}(t) = \frac{a^3}{3} \sum_{k=1}^{3} \sum_{\mathbf{x}} \langle T^{\alpha}_k(x) T^{\alpha}_k(0) \rangle \qquad \alpha = 1, 2\,. \tag{1-67}$$

In a similar way the correlation functions $C^{\alpha}_{AA}(t)$, $C^{\alpha}_{AP}(t)$ and others can be defined. Then, due to periodic boundary conditions, for example the correlation function C^{α}_{PP} is expected to behave for large Euclidean times t like $f(t) = A\cosh(m_{\mathrm{PS}}(t - T/2))$, where T is the lattice time extent, m_{PS} is the corresponding mass and A the amplitude. By fitting the functional form $f(t)$ to the data in a certain window $t_{\mathrm{min}} \leq t \leq t_{\mathrm{max}}$, where all the excited states have vanished, the values for the masses are actually extracted. Moreover, the amplitude A gives an estimate for the matrix element $\langle 0|P^{\alpha}(0)|\pi\rangle$, where we denote with $|\pi\rangle$ the pseudo scalar state with the lightest mass.

Besides the meson masses, the quark mass and the pseudo scalar decay constant

$$f_{\mathrm{PS}} \equiv m^{-1}_{\mathrm{PS}} \langle 0|A^{\alpha}_0(0)|\pi\rangle \tag{1-68}$$

are quantities of interest. For calculating f_{PS} the following two methods are possible. The first one is to obtain the amplitude $\langle 0|A^{\alpha}_0(0)|\pi\rangle$ from the asymptotic behavior of the correlation function $C^{\alpha}_{AA}(t)$ while the pseudo scalar mass is extracted from $C^{\alpha}_{PP}(t)$. The second method [49] is to obtain the ratio (we skip the flavor index α

in the following)

$$r_{AP} = \frac{\langle 0|A_0(0)|\pi\rangle}{\langle 0|P(0)|\pi\rangle} \tag{1-69}$$

from the asymptotic behavior of

$$\frac{C_{AP}(t)}{C_{PP}(t)} = r_{AP}\, \tanh[m_{\mathrm{PS}}(T/2-t)]\,, \tag{1-70}$$

where again m_{PS} is extracted from $C_{PP}(t)$. The value for f_{PS} can then be obtained from

$$f_{\mathrm{PS}} = m_{\mathrm{PS}}^{-1}\, r_{AP}\, \langle 0|P(0)|\pi\rangle\,. \tag{1-71}$$

Clearly, both methods agree asymptotically for large Euclidean times. With one of these definitions for f_{PS} also the bare current quark mass can be defined by the PCAC relation

$$m^{\mathrm{PCAC}} = \frac{f_{\mathrm{PS}}}{2\langle 0|P(0)|\pi\rangle}\, m_{\mathrm{PS}}^2\,. \tag{1-72}$$

Note that this quantity – in contrast to the masses extracted from the exponential fall-off – requires multiplicative renormalization, which is also the case for f_{PS} determined in the way explained above. In the pure Wilson case m^{PCAC} serves as an estimate for the bare quark mass.

1.3.2 tmQCD formulation

Introducing a twisted mass term in the Wilson-Dirac operator is equivalent to transforming the fermion fields $\psi \to \chi$ according to Eq. (1-50). Therefore also the composite fields have to be transformed. The result for the axial and vector currents is the following:

$$A_\mu'^{\alpha} \equiv \bar\chi\gamma_\mu\gamma_5\frac{\tau_\alpha}{2}\chi = \begin{cases} A_\mu^\alpha\, \cos(\omega) + \epsilon^{3\alpha\beta}\, V_\mu^\beta\, \sin(\omega) & (\alpha = 1,2)\,, \\ A_\mu^3 & (\alpha = 3)\,, \end{cases} \tag{1-73}$$

$$V_\mu'^{\alpha} \equiv \bar\chi\gamma_\mu\frac{\tau_\alpha}{2}\chi = \begin{cases} V_\mu^\alpha\, \cos(\omega) + \epsilon^{3\alpha\beta} A_\mu^\beta\, \sin(\omega) & (\alpha = 1,2)\,, \\ V_\mu^3 & (\alpha = 3)\,, \end{cases} \tag{1-74}$$

whereas the rotated scalar and the pseudo scalar densities are given by:

$$P'^{\alpha} \equiv \bar\chi\gamma_5\frac{\tau_\alpha}{2}\chi = \begin{cases} P^\alpha & (\alpha = 1,2)\,, \\ P^3\, \cos(\omega) + i\tfrac{1}{2}S^0 \sin(\omega) & (\alpha = 3)\,, \end{cases} \tag{1-75}$$

$$S'^{0} \equiv \bar\chi\chi = S^0\, \cos(\omega) + 2iP^3\, \sin(\omega)\,. \tag{1-76}$$

In the special case of $\omega = \pi/2$ the vector and the axial currents with $\alpha = 1,2$ transform into each other, while the pseudo scalar densities ($\alpha = 1,2$) do not rotate

with ω. Therefore the charged pseudo scalar ("pion") mass can be extracted for all values of ω from the pseudo scalar density also in the twisted basis. In the following we will skip the $'$ on the quantities in the twisted basis again, since we will always work in the twisted basis if tmQCD is concerned.

The PCAC and PCVC relations assume in the twisted basis the following form [44]:

$$\begin{aligned} \partial_\mu A_\mu^\alpha &= \quad 2m_0 P^\alpha + i\mu\delta^{3\alpha} S^0 \,, \\ \partial_\mu V_\mu^\alpha &= -2\mu\epsilon^{3\alpha\beta} P^\beta \,. \end{aligned} \tag{1-77}$$

While in the pure Wilson-Dirac operator there is only one mass term, which is aligned to the Wilson term (proportional to the unit matrix in flavor space), in the twisted mass operator (1-54) there is an additional mass term in the three-direction of flavor space, which is dis-aligned to the Wilson term. Therefore, at general values of the twist angle the quark mass estimate has to contain both of them, μ and m_q

$$m = \sqrt{(Z_{m_q} m_q)^2 + (Z_\mu \mu)^2} \,. \tag{1-78}$$

Of course, the twisted mass lattice action can be studied in the full parameter space (μ, m_q), but automatic $\mathcal{O}(a)$ improvement is only realized at full twist corresponding to $m_q = 0$. A sensible definition for an estimate of m_q is given by

$$m_\chi^{\mathrm{PCAC}} = \frac{\langle \nabla_\mu^* \bar{\chi}(x)\gamma_\mu\gamma_5 \frac{\tau^\pm}{2}\chi(x)\bar{\chi}(y)\frac{\tau^\pm}{2}\gamma_5\chi(y)\rangle}{2\langle \bar{\chi}(x)\frac{\tau^\pm}{2}\gamma_5\chi(x)\bar{\chi}(y)\frac{\tau^\pm}{2}\gamma_5\chi(y)\rangle} \,, \tag{1-79}$$

where we introduced $\tau^\pm = \tau_1 \pm i\tau_2$. Comparing to Eqs. (1-73,1-74,1-75) it is clear that m_χ^{PCAC} is identical to m^{PCAC} for $\omega = 0$. Moreover, at the special value of $\omega = \pi/2$, m_χ^{PCAC} is zero due to the exact vector symmetry in the lattice theory, and the quark mass is purely given by the twisted mass parameter μ, as it should be. Thus, m_χ^{PCAC} is a quantity that can be used to determine the parameters at which $\omega = \pi/2$.

In the twisted basis we can also define a quantity f_χ^{PS}, given by Eq. (1-68), but now in the twisted basis. We denote with the subscript χ that this is a quantity extracted in the twisted basis, and in fact it does not correspond to the physical pseudo scalar decay constant. A further estimate estimate for m_q in the twisted basis is then given by

$$m_\chi^{\mathrm{PCAC}} = \frac{f_\chi^{\mathrm{PS}}}{2\langle 0|P(0)|\pi\rangle} m_{\mathrm{PS}}^2 \,, \tag{1-80}$$

representing again the quark mass term aligned to the Wilson term.

Pseudo scalar decay constant at maximal twist

Of course, given the relations of the composite fields (1-73) connecting the twisted basis χ to the basis ψ, the value for f_{PS} can be extracted from f_{χ}^{PS}. But for the special choice $\omega = \pi/2$ this does not work (see above) and there is another possibility to extract a value for f_{PS} even without the need of a renormalization factor (cf. [50, 51, 52]).

Of particular interest in this context is the PCVC relation, which takes in the twisted basis the following form (1-77):

$$\partial_\mu \tilde{V}_\mu^\alpha \;=\; -2\mu\epsilon^{3\alpha\beta}P^\beta \;, \tag{1-81}$$

which holds exactly when the point split vector current $\tilde{V}$ as defined in Ref. [44] is used. From this follows that the vector current is protected against renormalization – in analogy to Ginsparg-Wilson fermions, which means the renormalization factor $Z_{\tilde{V}} = 1$. Therefore, Eq. (1-81) implies that $Z_P = Z_\mu^{-1}$, where Z_P is the renormalization factor of the pseudo scalar density and Z_μ the one for the twisted mass parameter μ.

Fixing the flavor index to $\alpha = 1$ we now again start with the standard definition for the pseudo scalar decay constant in the physical basis Eq. (1-68)

$$\langle 0|A_0^1(0)|\pi\rangle = f_{\mathrm{PS}} m_{\mathrm{PS}} \,. \tag{1-82}$$

In the twisted basis at maximal twist the rôle of the axial and vector current is just interchanged, and therefore we can write in the twisted basis

$$\partial_\mu \langle 0|V_\mu^2(0)|\pi\rangle = f_{\mathrm{PS}} m_{\mathrm{PS}}^2 \,. \tag{1-83}$$

Using then the vector Ward identity (1-81), we can finally relate the divergence of the vector current to the pseudo scalar density and obtain

$$f_{\mathrm{PS}}\, m_{\mathrm{PS}}^2 = \partial_\mu \langle 0|V_\mu^2(0)|\pi\rangle = 2\,\mu\,\langle 0|P^1(0)|\pi\rangle \,. \tag{1-84}$$

Thus, by fitting the pseudo scalar correlation function for large time separations, we can obtain m_{PS} and the amplitude $|\langle 0|P^1(0)|\pi\rangle|^2/m_{\mathrm{PS}}$ from which we then compute the desired matrix element $|\langle 0|P^1(0)|\pi\rangle|$. Hence, we have all necessary ingredients to determine f_{PS} from Eq. (1-84), without the need of any renormalization factor.

1.3.3 Overlap formulation

As we have discussed before the overlap formulation obeys exact chiral symmetry on the lattice and the theory is $\mathcal{O}(a)$ improved, as long as cut-off effects originating from the action are considered. We have seen that besides the action also the operators

are possible origins of cut-off effects. For the overlap formulation $\mathcal{O}(a)$ improved bilinears with a Dirac structure Γ can be constructed as follows ($\bar{a} = a/\rho$):

$$O_\Gamma = \bar{\psi}_q \Gamma \left(1 - \frac{\bar{a} D_{\text{ov}}^{(0)}}{2}\right) \psi_{q'} = \frac{1}{1 - \frac{\bar{a} m_{q'}}{2}} \left(\psi_q \Gamma \psi_{q'}\right) , \tag{1-85}$$

where ψ_q and $\psi_{q'}$ represent two different quark flavors q and q' with masses m_q and $m_{q'}$, respectively. The pseudo scalar meson mass is then extracted from the exponential fall-off at large Euclidean times of

$$C_{PP}^{\text{ov}}(t) = a^3 \sum_{\mathbf{x}} \langle P^\dagger(x) P(0) \rangle , \tag{1-86}$$

where the pseudo scalar density now is given by

$$P(x) = \bar{\psi}_q(x) \gamma_5 \left[\left(1 - \frac{\bar{a}}{2} D_{\text{ov}}^{(0)}\right) \psi_{q'}\right](x) . \tag{1-87}$$

In order to remove contributions of topological zero modes (which are finite volume effects, cf. [53, 54, 55, 56, 57, 58]) the same quantity can be extracted from

$$C_{PP-SS}^{\text{ov}}(t) = a^3 \sum_{\mathbf{x}} \langle P^\dagger(x) P(0) - S^\dagger(x) S(0) \rangle , \tag{1-88}$$

where S can be defined via Eq. (1-85) as follows

$$S(x) = \bar{\psi}_q(x) \left[\left(1 - \frac{\bar{a}}{2} D_{\text{ov}}^{(0)}\right) \psi_{q'}\right](x) . \tag{1-89}$$

While C_{PP-SS}^{ov} has the advantage that contributions from topological zero modes are canceled, it has the drawback that the scalar meson appears as an excited state and can affect the extraction of the ground state mass for large quark masses. The vector meson mass m_{V} is obtained with the overlap operator from

$$C_{VV}^{\text{ov}} = \frac{a^3}{3} \sum_{k=1}^{3} \sum_{\mathbf{x}} \langle V_k^\dagger(x) V_k(0) \rangle , \tag{1-90}$$

where the vector current is defined to be

$$V_\mu(x) = \bar{\psi}_q(x) \gamma_\mu \left[\left(1 - \frac{\bar{a}}{2} D_{\text{ov}}^{(0)}\right) \psi_{q'}\right](x) . \tag{1-91}$$

For later purposes we define the axial vector current in the overlap formulation

$$A_\mu(x) = \bar{\psi}_q(x) \gamma_\mu \gamma_5 \left[\left(1 - \frac{\bar{a}}{2} D_{\text{ov}}^{(0)}\right) \psi_{q'}\right](x) . \tag{1-92}$$

The extraction of the bare quark mass and the pseudo scalar decay constant from the Ward identity in case of overlap fermions is identical to the one described

here for Wilson fermions. Only the bilinears have to be replaced by the suitable counterparts as given in Eqs. (1-87) and (1-92).

We remark here that with overlap fermions the value for $f_{\rm PS}$ can also be obtained in the same way as for twisted mass fermions by using the PCAC relation. In fact $f_{\rm PS}$ can be computed from

$$f_{\rm PS}^{\rm ov} = \frac{2m_{\rm ov}}{m_{\rm PS}^2}|\langle 0|P(0)|\pi\rangle|\,, \tag{1-93}$$

where as in the tmQCD case no renormalization constant is needed. The quark mass $m_{\rm ov}$ is the quark mass parameter in the overlap operator (1-38).

1.3.4 Setting the scale

In lattice calculations, we need to fix one dimension-full quantity in order to set the overall scale and to translate lattice units into physical units. This can be done by using a hadronic scale r_0, which is introduced by the force $F(r)$ between static quarks at intermediate distance r [59].

The hadronic length scale r_0 defined by the implicit equation

$$r^2F(r)|_{r=r(c)} = c\,, \qquad r_0 = r(1.65) \tag{1-94}$$

has turned out to be a good choice to set the scale in lattice QCD calculations: it can be computed on the lattice with high precision, both statistically and systematically, and r_0 is known to have a value of about 0.5 fm in QCD.

The force can be computed from the static quark potential that in turn can be determined on the lattice from Wilson loops, which are $r \times t$ loops of gauge links. Wilson loops are defined – similar to the plaquette variable – as the traces of products of parallel transports around a closed loop. This is, as discussed before, a gauge invariant object due to cyclic invariance in the trace. They represent a static quark/anti-quark pair separated with distance r in space and propagating in time the distance t. In order to improve the overlap with the ground state usually several levels of APE smearing [60] are applied to the space like gauge links.

Following the variational approach of Ref. [61] one gets a correlation matrix $W_{ij}(r,t)$ with i,j representing the smearing levels applied at the two space like gauge link products. The correlation matrix can be used to solve the generalized eigenvalue problem

$$W_{ij}(r,t)v_j(r) = \lambda(r,t_0,t)W_{ij}(r.t_0)v_j(r)\,. \tag{1-95}$$

The eigenvector of the largest eigenvalue can be used to project W_{ij} on the ground state and the ground state energy can then be obtained form the exponential fall-off

of the latter at large enough values of t, leading to an estimate of the potential $V(r)$. Equivalently, the generalized eigenvalue can be used as an estimate for the ground state energy.

From the potential $V(r)$ the force can be computed by numerical differentiation, which in general reads

$$F(r') = \left[V(r') - V(r'-a)\right]/a\,, \tag{1-96}$$

at a certain distance r'. We decided to use a distance $r' = r_{\mathrm{I}}$ for which the lattice force does not deviate from the tree-level continuum value [59]

$$F(r_{\mathrm{I}}) = \frac{4}{3}\frac{g^2}{4\pi r_{\mathrm{I}}^2} + \mathcal{O}(g^4)\,. \tag{1-97}$$

The value of r_0/a can then be determined by interpolating the force to the value of r/a where Eq. (1-94) is fulfilled. A second possibility is to fit

$$V(r) = V_0 + \sigma r + \alpha\frac{1}{r}$$

to the data and determine the force from the best fit function. σ is the so called string tension and α parameterizes the Coulomb like part proportional to the gauge coupling. The final results, however, should agree within the errors for both methods.

The time like gauge links represent a static propagator of the static quarks that can be derived from an effective static action. Here the naive choice is the Eichten-Hill action [62], but one can also use static actions that show an improved signal to noise ratio [63, 64, 65, 66]. The improved actions formally correspond to actions with APE or HYP [67] smeared time like gauge links or a differently discretized time derivative. The APE and in particular the HYP action lead to improved statistical and systematical precision and therefore to a better estimate for r_0 in lattice units [63, 68, 69].

Our experience is that using a HYP static action is superior to using a APE static action, which is then superior to the Eichten-Hill static action. This concerns in particular the signal to noise ratio at large distances, where it is only with a improved static action possible to determine a value for the force. This becomes important close to the continuum and if one wants to use the potential to investigate for instance string breaking [69, 70]. We remark that changing the static action is equivalent to a different choice of the operator. Therefore, the different lattice estimates for r_0/a deviate from each-other by lattice artifacts, while the continuum extrapolated value should of course agree.

1.4 Lattice simulations

In this section we will shortly discuss some general principles for constructing algorithms used in lattice simulations.

In order to compute the partition function (1-21) or an expectation value of an operator O in principle one has to solve a high dimensional integral over the gauge fields and the Grassmann valued fermionic fields. However, most of the contributions to the integral have only low weight and therefore stochastic integration with *importance sampling* is an effective method to evaluate such integrals.

Stochastic integration with importance sampling preferentially chooses such configurations that have a strong weight. At the same time it is assured that the sample average estimates the ensemble average. This means that the sample is representative for the ensemble. In particular, such a stochastic integration can be performed by means of Markov chains:

Markov chain

Consider a stochastic process in which a finite set of configurations $U_{\tau_1}, U_{\tau_2}, \ldots$ is generated sequentially according to some transition probability $P_{ij} \equiv P(U_i \to U_j)$. The state of the system at any given simulation time τ_i will be a multi-dimensional random variable, whose distribution depends only on the preceding state, if P_{ij} depends only on the state U_i. A set of configurations generated in this way is called a *Markov chain*.

For an observable O we can define a simulation time average over a given set of configurations $\{U_{\tau_i}\}$ generated in a Markov chain by

$$\langle O \rangle_{\mathrm{N}} = \frac{1}{N} \sum_{i=1}^{N} O(U_{\tau_i}) \,. \tag{1-98}$$

We want to set up the transition probability in such a way that $\langle O \rangle_{\mathrm{N}}$ is in the limit $N \to \infty$ equal to the ensemble average corresponding to a given Boltzmann distribution e^{-S}. In order to achieve this it is sufficient that the transition probability fulfills as a sufficient condition the so called *detailed balance* condition

$$e^{-S(U)} P(U \to U') = e^{-S(U')} P(U' \to U) \,. \tag{1-99}$$

There are many algorithms known that correctly implement the condition of detailed balance. One of these is the *Metropolis algorithm* which is given by the following two steps:

1. Chose an arbitrary test configuration U'.

2. Accept configuration U' as successor of configuration U with probability

$$P(U \to U') \;=\; \begin{cases} \frac{e^{-S(U')}}{e^{-S(U)}} & \text{if } e^{-S(U')} < e^{-S(U)} \\ 1 & \text{else}\,. \end{cases} \tag{1-100}$$

This step is called a Metropolis accept/reject step.

The Metropolis algorithm is in principle applicable to a any system and it can be very efficient, if one can efficiently produce test configurations in such a way that the acceptance rate is high. One only has to take into account that the generation of test configurations must be *ergodic*. This means nothing but that the probability to generate configuration U' as a next test configuration must not be zero for any possible configuration.

1.4.1 Quenched approximation

As explained before, the dependence of expectation values on the fermionic fields ψ can be removed by the help of Wick's theorem. An expectation value of an operator O then reads

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}U \; O[U] \; e^{-S_{\mathrm{G}}[U] - \log\det M[U]} \;. \tag{1-101}$$

This means that only the integral over the color gauge configurations has to be performed and we need an algorithm to generate color gauge configurations U with the desired distribution. However, for instance for each accept/reject step the determinant of M needs to be computed, which is an highly non-local object.

Therefore, solely due to limited computer resources, computations in lattice QCD were often performed in a crude approximation: it consists of neglecting the fermion contribution to the path integral, i.e. setting $\det M = \textit{constant}$. This approximation corresponds to neglecting vacuum polarization effects of quark loops. As a consequence for instance the string between a quark and an anti-quark does not break at any distance.

Even though one could expect this approximation to be bad, since quenched lattice QCD is confining, asymptotically free and shows spontaneous chiral symmetry breaking, it is reasonable to use it as a model of QCD. As a side effect one can also extract physical results and compare it to experiment. And in fact the quenched approximation seems to work surprisingly well: the deviations from experimental measurements are only of the order of 10%, even though the systematic errors are hard to estimate.

From a practical point of view setting $\det M = \textit{constant}$ corresponds to simulating a pure Yang-Mills gauge theory [71] with for instance the Wilson gauge action (1-25) on the lattice. For pure gauge theories there are efficient Metropolis

Monte-Carlo algorithms available, like the *heat-bath* and the *overrelaxation* algorithm. While the latter of the two being not ergodic, their combination gives rise to an efficient algorithm to generate gauge configurations with the correct distribution. Details on this algorithm can be found in text books, see for instance Refs. [18, 19].

Finally we remark that, since r_0 as defined in section 1.3.4 is a quantity depending only on the gauge fields, the scale in quenched simulations does not depend on the fermionic mass, but only on the coupling β in the here discussed quenched approximation. It was computed for the Wilson plaquette gauge action in a range of coupling constants between $\beta = 5.7$ and $\beta = 6.75$ in Ref. [72] (see also [73]). Therefore, in the quenched simulations for this work we did not compute r_0 but rather used the values from Ref. [72].

1.4.2 Dynamical simulations

Of course, the aim of lattice QCD computations is to take the fermion determinant into account[††]. The reason for full QCD simulations to be much more expensive than the one in the quenched approximation is – as mentioned above – that the computation of the determinant is rather demanding.

Nevertheless, there are two widely used algorithms, which include the determinant in the generation of gauge configurations: on the one hand Multi-Boson like algorithms [74] and on the other hand Hybrid Monte Carlo (HMC) like algorithms [75]. The latter is discussed in detail in chapter 3, while the former is not used for this work. However, a Two-Step Multi-Boson (TSMB) algorithm [76] was used in the collaboration in addition to the HMC for checks and for production runs. Naturally, we cross-checked the results and found full agreement within errors.

Details for Multi-Boson like algorithms can be found in the mentioned references.

1.4.3 Error estimates

In lattice simulations with Monte-Carlo methods there exist several sources of errors. Apart from errors due to discretization and finite volume effects, the most important contribution – and usually also the largest – is the statistical error.

The latter arises from the fact that due to finite computer and human resources it is not possible to average over an infinitely large sample. Nevertheless, it is possible to estimate the error one makes by approximating the infinite large sample with a finite one. It is the content of this subsection to discuss the effects of taking only a finite sub-sample.

[††]Only because for a long time most of the simulations have been performed in the quenched approximation it became common to call full QCD simulations *dynamical simulations*.

Statistical error

When averaging over a sample of M configurations in order to measure the (primary) observable O, the root-mean-square deviation[‡‡] can be computed by

$$\sigma^2_{O,M} = \frac{1}{M-1}\left(\frac{1}{M}\sum_{i=1}^{M} O_i^2 - \left[\frac{1}{M}\sum_{i=1}^{M} O_i\right]^2\right) = \frac{1}{M-1}\left(\langle O^2\rangle - \langle O\rangle^2\right) , \quad (1\text{-}102)$$

where O_i represents the i-th measurement of the observable O. For this estimate it is necessary that the single measurements are not correlated, which is in general not true. Therefore we call it the *naive* error estimate. In order to account for the correlation one can use the *integrated autocorrelation time* τ_{int}. It can be defined as follows

$$\tau_{\text{int}} = \frac{1}{2}\sum_{\tilde{t}=-\infty}^{\infty} \frac{\Gamma_O(|\tilde{t}|)}{\Gamma_O(0)} , \quad (1\text{-}103)$$

where we label the "Monte Carlo time" with $\tilde{t}$ and we introduced the autocorrelation function Γ_O for an operator O

$$\Gamma_O(|i-j|) = \langle (O_i - \langle O\rangle)(O_j - \langle O\rangle)\rangle .$$

$\Gamma_O(|i-j|)$ depends only on the distance $\tilde{t} = |i-j|$ between measurements and decays typically exponentially fast with the so called autocorrelation time τ_{c}

$$\Gamma_O(\tilde{t}) \propto \exp(-\tilde{t}/\tau_{\text{c}}) .$$

Typically τ_{c} and τ_{int} are found to be of the same order. There are two possibilities to incorporate the integrated autocorrelation time in the estimate of the statistical error:

On the one hand one can leave out during the course of production of configurations τ_{int} many configurations until one is used for the measurements. Then the sample of configurations is uncorrelated and the naive error (1-102) can serve as a good approximation for the real error. This approach has the disadvantage that the value of τ_{int} is in general not known before the measurements were performed and one might perform measurements which are then not used for the final result.

On the other hand one can use the value of τ_{int} for the error computation taking all the measurements into account: it is possible to show that the statistical error of correlated measurements can be computed from the naive error and the integrated autocorrelation time in the following way [77]

$$\sigma^2_{O,M} = 2\tau_{\text{int}}\left(\sigma^{\text{naiv}}_{O,M}\right)^2 . \quad (1\text{-}104)$$

[‡‡]The denominator arises from the fact that the exact mean value $\overline{O}$ was replaced by its estimate over the sample, because the exact mean value is not known.

This means de facto a reduction of the number of independent measurements to $M/(2\tau_{\text{int}})$.

Of course, when the integrated autocorrelation time shall be used to estimate the error, a reliable estimate of τ_{int} itself and its error is needed. First of all a reliable determination of τ_{int} is only possible, if the sample is large enough. Moreover the estimation of the statistical error on τ_{int} is a delicate procedure. It is discussed in detail in Ref. [78] and we will solely perform our error analysis along the lines of this reference (cf. also [77]).

Another method to estimate the real error is to pre-average the measurements on blocks of the total sample:

$$O_{l,B} = \frac{1}{B} \sum_{i=1+(l-1)B}^{lB} O_i, \qquad l = 1, ..., M_B = \frac{M}{B}. \tag{1-105}$$

If the blocks B become large enough – in units of τ order of τ_{int} – then the block averages are uncorrelated and the error can be estimated by

$$\sigma^2 = \frac{1}{M_B(M_B-1)} \sum_l \left(O_{l,B} - \frac{1}{M_B} \sum_{l'} O_{l',B} \right)^2 . \tag{1-106}$$

The value of the latter will increase with increasing B and will, if the block-averages become uncorrelated, reach a plateau. Of course the plateau will be only reached, if the sample is large enough. This method is called *binning*.

A further method, which is similar to binning, is the so called *Jackknife binning*. Instead of using the blocks itself to pre-average the measurements, the blocks complementary to the binning blocks are used. Therefore the blocks are significantly larger:

$$O_{l,\bar{B}} = \frac{1}{M-B} \left(\sum_{i=1}^{(l-1)B} O_i + \sum_{i=lB+1}^{M} O_i \right) .$$

One can show that the error is now estimated by

$$\sigma^2 = \frac{M_B - 1}{M_B} \sum_l \left(O_{l,\bar{B}} - \frac{1}{M_B} \sum_{l'} O_{l',\bar{B}} \right)^2 .$$

Especially for derived quantities the Jackknife binning is a widely used method (cf. [18]).

In this work we solely rely on the error analysis with help of the integrated autocorrelation time. The other methods are only used for checks. The statistical error on the integrated autocorrelation time as well as the error for derived quantities is determined along the lines of [78] (cf. [77]).