HISQ and Symanzik Improvement Notes

David Clarke

February 5, 2021

Chapter 1

Staggered fermions

In this section we work in a free theory to simplify the notation. In the interacting theory, each spinor will have a link attached to it originating at the same point, and the Dirac indices that we will manipulate commute with these links, so the derivation will be the same. This follows parts of Section 10.1 in Gattringer and Lang [1] and Section 4.4 in Rothe [2].

We saw before that the doubling problem could be alleviated by adding a Wilson term of the form

$$\frac{1}{a}\sum_{\mu}\mathbf{1}\big(1-\cos(p_\mu a)\big). \tag{1.1}$$

Writing this as exponentials, and doing the Fourier back transformation, we can express this as

$$a\sum_{\mu=1}^{4}\frac{1}{2a^2}\left(2\delta_{x,y}-\delta_{x,y-a\hat{\mu}}-\delta_{x,y+a\hat{\mu}}\right). \tag{1.2}$$

In this form, one clearly sees that the Wilson term vanishes in the naive continuum limit like a. Moreover it has the form of a discretized $\partial_\mu\partial_\mu$. A drawback of this term, however, is that it breaks $\mathrm{SU}_A(N_f)\times\mathrm{U}_A(1)$ explicitly because the Wilson term commutes with γ_5.

For some calculations one may desire not to break this symmetry entirely, in particular when one is investigating phenomena closely related to chiral symmetry. This motivates other possible fermion discretizations that remove doublers. *Staggered fermions* will remove some of the unphysical doublers, while keeping a remnant of this axial symmetry intact. This is accomplished

by redistributing the quark degrees of freedom over the lattice in such a way that the spacing between each degree of freedom doubles, thereby doubling the effective lattice spacing, and hence reducing the size of the Brillouin zone. New quarks called *tastes* will be constructed by linear combinations of these degrees of freedom.

We take as the starting point the naive fermion action

$$S_F = a^4 \sum_n \bar{\psi}(n) \left(\sum_{\mu=1}^{4} \gamma_\mu \frac{\psi(n+\hat{\mu}) - \psi(n-\hat{\mu})}{2a} + m\psi(n) \right). \tag{1.3}$$

In this equation and in what follows I will write the space-time coordinate as n to make the connection between the space-time coordinate and the following *staggered transformation*

$$\begin{aligned} \psi(n) &= \gamma_1^{n_1}\gamma_2^{n_2}\gamma_3^{n_3}\gamma_4^{n_4}\psi(n)' \\ \bar{\psi}(n) &= \bar{\psi}(n)'\gamma_4^{n_4}\gamma_3^{n_3}\gamma_2^{n_2}\gamma_1^{n_1} \end{aligned} \tag{1.4}$$

more clear.

The fact that these transformed fields come with power of gamma matrices, along with the fact that the gamma matrices are idempotent, will diagonalize the naive action in Dirac space. In particular one finds relationships such as

$$\bar{\psi}(n)\gamma_4\psi(n+\hat{4}) = (-1)^{n_1+n_2+n_3}\bar{\psi}(n)'\mathbf{1}\psi(n+\hat{4}), \tag{1.5}$$

which means that in the kinetic part, the gamma matrices are annihilated, and what remains are scalar *staggered phases*

$$\eta_1(n) = 1, \quad \eta_2(n) = (-1)^{n_1}, \quad \ldots, \quad \eta_4(n) = (-1)^{n_1+n_2+n_3}. \tag{1.6}$$

Substituting the staggered fields into the naive action, we thus find

$$S_F = a^4 \sum_n \bar{\psi}(n)'\mathbf{1} \left(\sum_{\mu=1}^{4} \eta_\mu(n) \frac{\psi(n+\hat{\mu})' - \psi(n-\hat{\mu})'}{2a} + m\psi(n)' \right), \tag{1.7}$$

which is now diagonal in Dirac space because the gammas are gone. The Dirac operator of eq. (1.7) represents four copies of the same equations along its diagonal, so we can throw away three copies without losing any information. We will call this last, kept copy χ; another way of looking at

this is we keep only one Dirac component. With this notation, we find a particularly simple form for the staggered action:

$$S_F^{\text{stag}} = a^4 \sum_n \bar{\chi}(n) \left(\sum_{\mu=1}^{4} \eta_\mu(n) \frac{\chi(n+\hat{\mu}) - \chi(n-\hat{\mu})}{2a} + m\chi(n) \right) \tag{1.8}$$

Equation (1.8) is convenient for simulations because it contains no Dirac spinors explicitly, but we would like a form that tells us a little more about the physics. We will now begin to rewrite this equation in such a way that we see the increase of the effective lattice spacing. The strategy will be to divide the lattice into disjoint unit hypercubes. The χ live on the corners of the hypercubes. New fields q will be linear combination of the χ, and their sites will be indexed by hypercube number.

To this end, we introduce new vectors h (hypercube) and s (corner) that are related to the original site vectors by

$$n_\mu = 2h_\mu + s_\mu \quad \text{with} \quad h_\mu \in \{0, 1, ..., N_\mu/2\} \quad s_\mu \in \{0, 1\}. \tag{1.9}$$

Moving in the direction μ, h_μ counts the hypercube number, while s is a displacement vector starting at the hypercube's origin. In this indexing scheme

$$\eta_\mu(n) = \eta_\mu(2h + s) = \eta_\mu(s) \tag{1.10}$$

because contributions to powers of gamma coming from the hypercubes are always even, according to eq. (1.9).

Next we introduce Γ, which will serve as the weights in the linear combination. We define

$$\Gamma^s = \gamma_1^{s_1} \gamma_2^{s_2} \gamma_3^{s_3} \gamma_4^{s_4}. \tag{1.11}$$

These obey, respectively, orthogonality and completeness relations[1]

$$\frac{1}{4} \operatorname{tr} \left[\Gamma^{s\dagger} \Gamma^{s'} \right] = \delta_{ss'} \quad \text{and} \quad \frac{1}{4} \sum_s \Gamma^{s*}_{ba} \Gamma^s_{b'a'} = \delta_{aa'} \delta_{bb'} \tag{1.12}$$

Our new quark fields are then (see Fig. 1.1)

$$q(h)_{ab} \equiv \frac{1}{8} \sum_s \Gamma^s_{ab} \chi(2h + s) \quad \text{and} \quad \bar{q}(h)_{ab} \equiv \frac{1}{8} \sum_s \bar{\chi}(2h + s) \Gamma^{s*}_{ab}. \tag{1.13}$$

[1]The orthogonality relation can be proven using properties of the Euclidean γ matrix, in particular that $\gamma_\mu = \gamma_\mu^\dagger = \gamma_\mu^{-1}$.

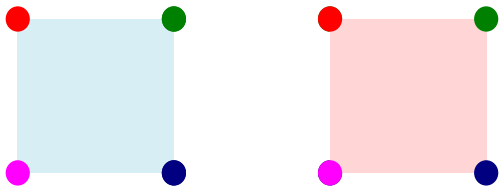

Figure 1.1: Example subdivision of one direction into disjoint hypercubes. Each hypercube is labeled by h, and the field $q(h)$ associated to the hypercube is a linear combination of the χ fields on the corners, here represented by different colors. The effective lattice spacing between a χ of one corner of the blue hypercube and the corresponding χ in the same corner of the red hypercube is $a' = 2a$.

In order to cast the staggered action in terms of $q(h)$, we need to invert eq. (1.13). This can be done using the orthogonality relation. One obtains

$$\chi(2h+s) = 2\,\mathrm{tr}\left[\Gamma^{s\dagger} q(h)\right] \quad \text{and} \quad \bar{\chi}(2h+s) = 2\,\mathrm{tr}\left[\bar{q}(h)\Gamma^{s}\right]. \tag{1.14}$$

We are now ready to carry out the substitution. The mass term follows straightforwardly from the completeness relation. The kinetic term is difficult, however, because the site offsets mix hypercubes. Therefore it is crucially important to keep track of which hypercube χ belongs to. Keeping track of the hypercube will make manifest an s dependence that makes applying the completeness relation not possible, at least superficially. Gattringer and Lang [1] suggest the following trick: we can shift the site that the first hypercube starts on by one, which gives equivalent results as the original hypercube labeling convention, and then average over these two conventions. I wasn't able to figure this out. Rothe [2] has another strategy, which I was also not able to follow through. Nevertheless, one should find

$$S_F^{\mathrm{stag}} = a'^4 \sum_h \left(m\,\mathrm{tr}\left[\bar{q}q\right] + \mathrm{tr}\left[\bar{q}\gamma_\mu \partial_\mu q\right] - \frac{a'}{2}\,\mathrm{tr}\left[\bar{q}\gamma_5 \Box_\mu q \gamma_\mu \gamma_5\right] \right) \tag{1.15}$$

where we have an implied summation over μ, we have written $q = q(h)$, introduced the effective spacing $a' \equiv 2a$, and the discretized derivatives are

$$\partial_\mu f(h) \equiv \frac{f(h+\hat{\mu}) - f(h-\hat{\mu})}{2a'} \tag{1.16}$$

and

$$\Box_\mu f(h) \equiv \frac{f(h+\hat{\mu}) - 2f(h) + f(h-\hat{\mu})}{a'^2}. \tag{1.17}$$

Our final step in the exploration of the staggered action will be to identify the unphysical fermions. These *tastes* are hidden in the q. This should not be surprising because the q_{ab} have 16 components, but Dirac spinors should have 4. This tells us each q corresponds to 4 spinors, and we will identify from the a and b a taste index and Dirac index. By comparing e.g. the $\mathrm{tr}[\bar{q}q] = \bar{q}_{ab}q_{ba}$ term with what we expect from the physical, continuum theory, it makes sense to identify the Dirac index with b. Thus our taste spinors are

$$\psi^t(h)_\alpha \equiv q(h)_{\alpha t} \quad \text{and} \quad \bar{\psi}^t(h)_\alpha \equiv \bar{q}(h)_{t\alpha}. \tag{1.18}$$

The staggered fermion action then becomes

$$S_F^{\text{stag}} = a'^4 \sum_h \left(m\bar{\psi}^t\psi^t + \bar{\psi}^t\gamma_\mu\partial_\mu\psi^t - \frac{a'}{2}\bar{\psi}^t\gamma_5(\tau_5\tau_\mu)_{tt'}\Box_\mu\psi^{t'} \right), \tag{1.19}$$

where we have introduced new matrices

$$\tau_\mu \equiv \gamma_\mu^T. \tag{1.20}$$

With the staggered action in the form of eq. (1.19) we can begin to discuss some physics. The last term is called the *taste-breaking* term. It is similar to the Wilson term, in the sense that it also represents a second derivative. However unlike the Wilson term, it allows for interactions between fermions of different taste, i.e. it allows for *taste mixing*. If not for the taste-breaking term, fermions of different taste would be mass-degenerate, which one sees clearly from the first term. In the naive continuum limit, this taste breaking terms vanishes like a.

The Wilson term broke axial symmetry completely, but in the taste-breaking term, the remnant $\mathrm{U}_A(1) \times \mathrm{U}_A(1)$ remains. In particular this term is invariant under transformations

$$\psi' = e^{i\omega}\psi, \qquad \bar{\psi}' = \bar{\psi}e^{-i\omega} \tag{1.21}$$

and

$$\psi' = e^{i\omega\gamma_5\otimes\tau_5}\psi, \qquad \bar{\psi}' = \bar{\psi}e^{i\omega\gamma_5\otimes\tau_5}. \tag{1.22}$$

This latter symmetry follows from the fact that γ_5 commutes through the taste-breaking term, while τ_5 will pick up a minus sign. One can identify

this symmetry with a subgroup of the axial taste symmetry group $\mathrm{SU}_A(N_t)$, where N_t is the number of tastes.

At finite lattice spacing, taste mixing lifts the taste mass degeneracy. One way to get some feeling for taste-breaking effects, then, is to look at the mass spectrum of staggered fermions. It has been found that these taste breaking effects can be reduced by improved gauge actions or smearing [3; 4], and in particular smearing seems to drive masses to 4-fold degeneracy. An intuition for why smearing might help is as follows: In the interacting theory, each χ is attached to links according to its site, and since tastes are linear combinations of these, it follows that different tastes touch different links. So the more "distance" in $\mathrm{SU}(N)$ space between the links, i.e. the more the links fluctuate, the greater the taste-breaking effects will be. Since smearing algorithms tend to drive links to more typical values given their neighbors, they reduce these fluctuations, and hence the taste-breaking.

We would also like to suppress the effects of unphysical tastes. Absent taste-breaking, the Dirac operator corresponding to the staggered action would be block diagonal in taste space, i.e. we would have for one physical flavor

$$D^{\mathrm{stag}} = \begin{pmatrix} D & & & \\ & D & & \\ & & D & \\ & & & D \end{pmatrix}. \tag{1.23}$$

One commonly used strategy to remove the effects of taste-breaking is therefore *rooting*. The idea is that $\det D^{\mathrm{stag}}$ is the contribution to the probability distribution from four mass-degenerate flavors. To isolate one of the flavors, it is sufficient to use $\det D = (\det D^{\mathrm{stag}})^{1/4}$. For two degenerate light flavors m_l and one heavier flavor m_s, i.e. for $N_f = 2+1$ fermions, one then samples with probability distribution

$$\mathrm{dP} = \mathrm{d}U e^{-S_G} (\det D_l^{\mathrm{stag}})^{1/2} (\det D_s^{\mathrm{stag}})^{1/4}. \tag{1.24}$$

In practice, taste-breaking is present at each lattice spacing, and therefore it is not clear whether there is some leftover effect of rooting in the continuum limit. In spite of this danger, people who employ staggered actions often tend to use rooting anyway. One way to make this step more justified is to reduce taste-breaking effects. Improved actions such as the HISQ action, to be discussed in the next section, have greatly reduced taste-breaking and, reassuringly, results from HISQ actions seem to agree with experiment.

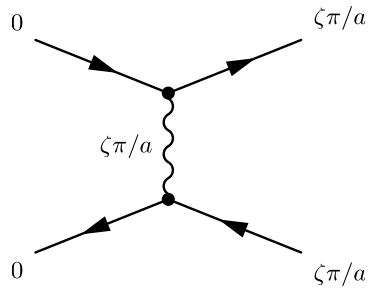

Figure 1.2: A low-energy quark absorbing a gluon with momentum $\zeta\pi/a$ can become a quark of another taste. Here ζ is a vector labelling some unphysical corner of the Brillouin zone. Image taken from Ref. [5].

1.1 Highly improved staggered fermions

Highly improved staggered quarks (HISQ) were introduced in 2007 by Follana *et al.* [5]. This paper summarizes also some of the history behind staggered quarks, and is written in a generally friendly way; therefore I encourage the reader to have a look at it.

Taste breaking can be thought of through *taste exchange*, where one quark changes its taste by exchanging a virtual gluon with momentum $p = \pi/a$; a quark with low enough momentum can thereby be pushed into another corner of the Brillouin zone. See Fig. 1.2 for a tree-level diagram. One goal of HISQ actions, then, is to suppress these processes, which is done by smearing. Another way to look at taste violations is that they are a lattice artifact, and so like any other lattice artifact, you would like to remove its leading order contributions to improve the approach to the continuum limit. Along this vein, another improvement made by HISQ is to reduce lattice artifacts from the derivative discretization.

The basic strategy of HISQ can be summarized as:

1. Improve finite difference derivatives by

$$\partial_\mu \to \partial_\mu - \frac{a^2}{6}(1+\epsilon)\partial_\mu^3,$$

   where ϵ depends on charm physics. Without considering charm physics, this is called a *Naik term*.

2. Find a smear $U_\mu \to \mathcal{F}_\mu U_\mu$ that vanishes for links carrying momentum π/a. We start with one called *Fat7*.

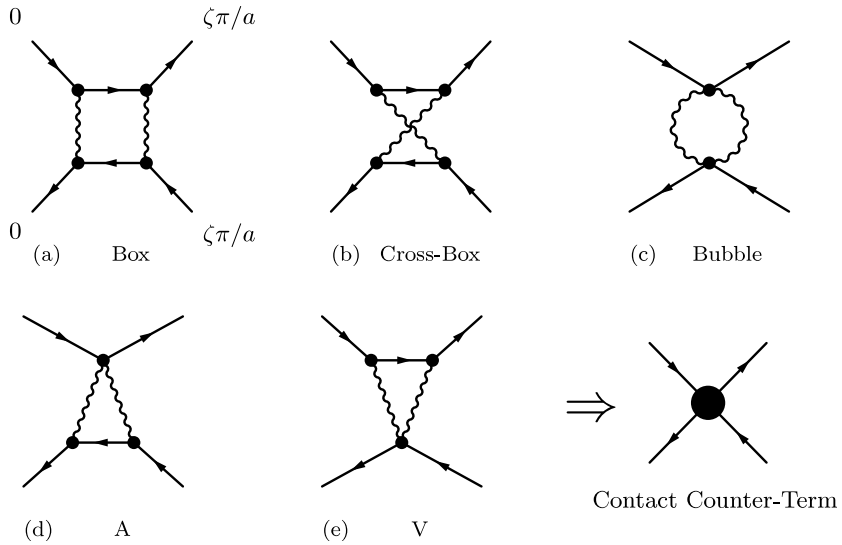

Figure 1.3: Some example one-loop processes contributing to taste exchange. Image taken from Ref. [5].

3. Multiple smearing reduces mass splittings more, so do that.

4. Remove new $\mathcal{O}(a^2)$ errors introduced by Fat7.

5. Smearing enhances some one-loop taste exchange processes, which can be suppressed by reunitarizing.

Overall the HISQ smear is then

$$\mathcal{F}^{\text{HISQ}} \equiv \mathcal{F}_{\text{corr}}^{\text{Fat7}} \mathcal{U} \mathcal{F}^{\text{Fat7}}, \tag{1.25}$$

where $\mathcal{F}_{\text{corr}}^{\text{Fat7}}$ has been corrected for the errors referenced in step 4, and $\mathcal{U}$ is the reunitarization operator needed for step 5. The HISQ action uses link variables smeared by eq. (1.25), and its kinetic term uses improved Naik discretization.

To judge how well the HISQ action does, the authors of Ref. [5] calculated several one-loop taste exchange processes (that fall into several broad categories indicated in Fig. 1.3) and determined their coefficients, d. Table 1.1 compares the suppression of taste exchange to other smearing programs. Before HISQ, ASQTAD was a popular dynamical fermion action, and one sees that, measured in this way, taste effects in the HISQ action are reduced by an order of magnitude.

		Unimproved		Improved
	ASQTAD	HISQ	HYP	ASQTAD
Avg d,$\tilde{d}$	0.23	0.02	0.02	0.13

Table 1.1: Average coefficients for the taste processes indicated in Fig. 1.3 for a few types of lattice actions. Columns indicate whether the gluons are improved. Excerpt from Table II of Ref. [5].

References

[1] C. Gattringer and C. B. Lang. *Quantum Chromodynamics on the Lattice*. Springer, Berlin, 2010. ISBN 978-3-642-01849-7.

[2] H. J. Rothe. *Lattice Gauge Theories*. World Scientific, Singapore, 2005. ISBN 981-256-168-4.

[3] S. Dürr, C. Hoelbling, and U. Wenger. Staggered eigenvalue mimicry. *Phys. Rev. D*, 70:094502, 2004. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.70.094502. URL https://link.aps.org/doi/10.1103/PhysRevD.70.094502.

[4] E. Follana, A. Hart, and C. T. H. Davies. Index Theorem and Universality Properties of the Low-Lying Eigenvalues of Improved Staggered Quarks. *Phys. Rev. Lett.*, 93:241601, 2004. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.93.241601. URL https://link.aps.org/doi/10.1103/PhysRevLett.93.241601.

[5] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong. Highly improved staggered quarks on the lattice with applications to charm physics. *Phys. Rev. D*, 75(5):054502, 2007. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.75.054502. URL https://link.aps.org/doi/10.1103/PhysRevD.75.054502.

Chapter 2

Improved actions and interpolators

In LFT the correlation length ξ is related to the mass m of the lightest particle by

$$\xi = \frac{1}{ma}. \tag{2.1}$$

This can be seen from looking at the spectral decomposition of the connected correlation function

$$\langle X(x)X(y)\rangle_c = \sum_{n>0} |\langle 0|\, X\, |n\rangle|^2\, e^{-(E_n - E_0)\Delta t}, \tag{2.2}$$

which has been evaluated at $\vec{x} = \vec{y} = 0$ for notational simplicity and we define Δt as the temporal distance between x and y. For large Δt we find schematically

$$\langle X(x)X(y)\rangle_c \sim e^{(E_1 - E_0)\Delta t} = e^{-m\Delta t} = e^{-\Delta t/\xi}. \tag{2.3}$$

Equation (2.2) measures how correlated the field X at x is with the field at y. In the continuum limit, ξ will diverge since m is constant, i.e. the system experiences a second-order phase transition. Correspondingly the correlation function increases to maximum correlation. The divergence of ξ implies a divergence in the integrated autocorrelation time

$$\tau_{\mathrm{int},X} \sim \xi_X^z \tag{2.4}$$

with *dynamical critical exponent* z that depends on the algorithm details. For $z \neq 0$, eq. (2.4) implies a deterioration in the efficacy of the updating process toward the continuum limit; this is *critical slowing down*.

Due to critical slowing down, configuration generation becomes increasingly expensive as one lowers the lattice spacing. On the other hand, one wishes to minimize discretization error. One way to proceed is as follows: There are multiple discretizations of a lattice action that will lead to the same continuum limit action, so one can design a lattice action where one suppresses the discretization effects by hand. Such actions are examples of *improved actions*. They are in general slower to generate configurations at a given a than unimproved actions, but have the advantage of being in some sense "closer to the continuum limit" than the unimproved actions at a given N_τ.

One knows that the Wilson action has $\mathcal{O}(a^2)$ discretization errors. If we add to the Wilson action some $\mathcal{O}(a^2)$ terms out, we will have eliminated all discretization error up to $\mathcal{O}(a^3)$. This is called an $\mathcal{O}(a^3)$ *improved action* or an $\mathcal{O}(a^4)$ *action*. Ultimately, our goal is to calculate expectation values, so we will generally also want to improve our lattice interpolators.

In what follows, we will look at some improvement strategies, guided by Chapter 10 of [1] and Chapter 9 of [2], where there is also more discussion to be found. Improvement schemes are rather technically involved, so I will not be able to verify all details carefully.

2.1 Symanzik improvement

The Symanzik improvement scheme [3; 4; 5; 6] is a systematic way to improve lattices actions. Gattringer and Lang summarize the strategy as carrying out the following steps:

1. Start with a discretized version of your quantity Q.

2. Identify correction terms using the continuum theory, keeping in mind what it allowed given the symmetries of Q, and organize them according to their mass dimension.

3. Add discretized versions of the correction terms, multiplied with suitable coefficients, to the discretized version of Q so that lattice artifacts vanish up to the desired order.

There is often more than one way to pick correction terms.

We begin by improving the gauge action. In the continuum theory, the gauge action is proportional to the dimension-four operator

$$O_4 = \operatorname{tr} F_{\mu\nu} F_{\mu\nu}. \tag{2.5}$$

Our first correction term is at $\mathcal{O}(a^2)$, and there are no dimension-five operators anyway, so we write the three dimension-six operators[1]

$$\begin{aligned} O_{6a} &= \sum_{\mu\nu} \operatorname{tr} D_\mu F_{\mu\nu} D_\mu F_{\mu\nu}, \\ O_{6b} &= \sum_{\mu\nu\rho} \operatorname{tr} D_\mu F_{\nu\rho} D_\mu F_{\nu\rho}, \\ O_{6c} &= \sum_{\mu\nu\rho} \operatorname{tr} D_\mu F_{\mu\rho} D_\nu F_{\nu\rho}. \end{aligned} \tag{2.6}$$

If we expand the plaquette in the lattice spacing, all of the above operators will appear:

$$\operatorname{tr} U^{\square} = N_c + \frac{1}{2} a^4 O_4 + a^6 \sum_i r_i O_{6i}, \tag{2.7}$$

where $i \in \{\text{a, b, c}\}$.

To improve our gauge action thus requires at least three dimension-six lattice operators. The simplest are the rectangle "rt", the parallelogram "pg", and the chair "ch", which are depicted in Fig. 2.1. Their multiplicities are listed in Table 2.1. An improvement on the Wilson action could then be written as

$$S_{\text{LW}} = \frac{\beta}{3} \sum_i c_i \operatorname{Re} \operatorname{tr} \left(\mathbf{1} - U_i \right), \tag{2.8}$$

where $i \in \{\text{pl, rt, pg, ch}\}$, with "pl" standing for plaquette, and U_i is the corresponding lattice operator. In their original paper, Lüscher and Weisz chose the normalization

$$c_{\text{pl}} + 8c_{\text{rt}} + 8c_{\text{pg}} + 16c_{\text{ch}} = 1. \tag{2.9}$$

[1] You can convince yourself this is all there is by keeping track of the dimensions of the comprising operators, for example D_μ has mass dimension one and $F_{\mu\nu}$ therefore has mass dimension two, and by demanding gauge and Lorentz invariance. In the end you need two D_μ's and two $F_{\mu\nu}$'s, and there are three different ways to contract their Lorentz indices.

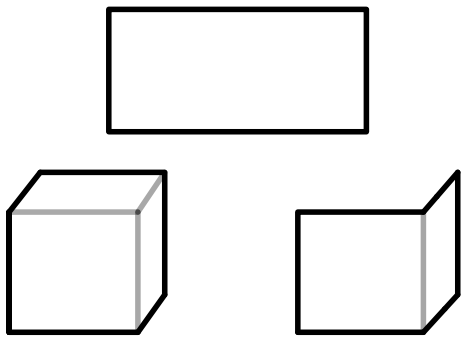

Figure 2.1: Lattice dimension-six operators. The top is the rectangle, the bottom-left is the parallelogram, and the bottom-right is the chair. Light grey lines are to guide the eye.

Operator type	Number elements per site
Rectangle	12
Parallelogram	16
Chair	48

Table 2.1: Dimension-six operator multiplicities. Loops that differ by orientation only are considered equal.

To fix the coefficients, one chooses an improvement condition, for example that the zero-temperature static potential gives

$$V(r) \approx \frac{1}{r} \tag{2.10}$$

up to $\mathcal{O}(a^4, g^2a^2)$. This yields

$$c_{\rm pg} = 0 \quad \text{and} \quad c_{\rm rt} + c_{\rm ch} = -\frac{1}{12}. \tag{2.11}$$

Furthermore since there are the most chair operators, it is convenient to pick $c_{\rm ch} = 0$. According to the normalization condition eq. (2.9), one thus finds

$$c_{\rm pl} = \frac{5}{3}, \quad c_{\rm rt} = -\frac{1}{12}, \quad c_{\rm pg} = c_{\rm ch} = 0, \tag{2.12}$$

which is called the *Lüscher-Weisz* action [6].

Next we turn to an improvement of the Wilson fermionic action. Now the leading correction is $\mathcal{O}(a)$, and owing to $\bar{\psi}\psi$ being dimension-three, there is now the possibility for dimension-five operators. Demanding that our operators satisfy the symmetries of the Wilson action, e.g. charge conjugation $\mathcal{C}$ and parity $\mathcal{P}$, we end up in the continuum with

$$\begin{aligned}
O_{5a} &= \bar{\psi}\sigma_{\mu\nu}F_{\mu\nu}\psi, \\
O_{5b} &= \bar{\psi}\left(\overrightarrow{D}_\mu\overrightarrow{D}_\mu + \overleftarrow{D}_\mu\overleftarrow{D}_\mu\right)\psi, \\
O_{5c} &= m\,\mathrm{tr}\,F_{\mu\nu}F_{\mu\nu}, \\
O_{5d} &= m\bar{\psi}\left(\overrightarrow{\not{D}} - \overleftarrow{\not{D}}\right)\psi, \\
O_{5e} &= m^2\bar{\psi}\psi,
\end{aligned} \tag{2.13}$$

where

$$\sigma_{\mu\nu} \equiv \frac{1}{2i}[\gamma_\mu, \gamma_\nu]. \tag{2.14}$$

This list of operators can be reduced by using the field equation

$$(\not{D} + m)\psi = 0, \tag{2.15}$$

from which it follows

$$O_{5a} - O_{5b} + 2O_{5e} = 0 \quad \text{and} \quad O_{5d} + 2O_{5e} = 0, \tag{2.16}$$

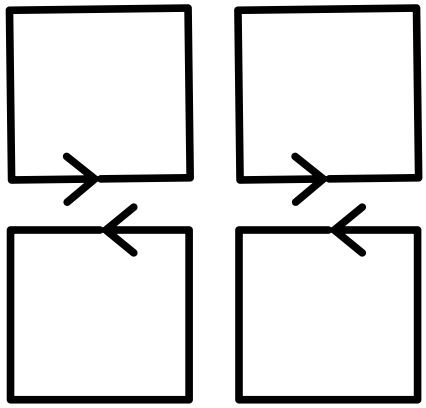

Figure 2.2: The "clover" sum $Q_{\mu\nu}(x)$ used in the discretization of the field strength $\hat{F}_{\mu\nu}$ in eq. (2.18). The point x is in the middle of the squares.

i.e. O_{5b} and O_{5d} are linearly dependent on the other operators, and can therefore be eliminated. Of the remaining operators, O_{5c} and O_{5e} are, up to the factor m, present in the original action, so they can be absorbed in the original action by redefining the bare parameters m and g.

Taken altogether, all that remains for improvement is the operator O_{5a}, so that the $\mathcal{O}(a)$ improved Wilson fermion action becomes

$$S_{\text{clover}} = S_{\text{Wilson}} + \frac{c_{\text{SW}}}{2}\, a^5 \sum_{x;\, \mu<\nu} \bar{\psi}(x)\sigma_{\mu\nu}\hat{F}_{\mu\nu}\psi(x), \tag{2.17}$$

with the real *Sheikholeslami-Wohlert* coefficient c_{SW} [7] and $\hat{F}_{\mu\nu}$ being the following discretization of the field strength tensor:

$$\hat{F}_{\mu\nu}(x) \equiv -\frac{i}{8a^2}\left(Q_{\mu\nu}(x) - Q_{\nu\mu}(x)\right) \tag{2.18}$$

where $Q_{\mu\nu}(x)$ is the plaquette sum (see Fig. 2.2)

$$Q_{\mu\nu}(x) \equiv U_{\mu\nu}(x) + U_{\nu,-\mu}(x) + U_{-\mu,-\nu}(x) + U_{-\nu,\mu}(x). \tag{2.19}$$

Due to the shape of $Q_{\mu\nu}$, the latter term in eq. (2.17) is called the *clover term*, and the action is said to be *clover-improved*. This discretization of $F_{\mu\nu}$ is

chosen because it is more symmetric than the simple plaquette discretization. The coefficient c_{SW} can be calculated perturbatively; one finds [7]

$$c_{\mathrm{SW}} = 1 + 0.2659g^2 + \mathcal{O}\big(g^4\big). \tag{2.20}$$

To round out our discussion of Symanzik improvement, we now turn to the improvement of some interpolators. The hitherto discussed action improvements suffice to improve on-shell quantities, e.g. hadron masses, at $\mathcal{O}(a)$. For these on-shell quantities, only eigenstates of the Hamiltonian contribute. However it is not enough for off-shell quantities such as correlators, which have off-diagonal contributions, and their interpolators need to be improved. We will discuss now two examples, the isovector axial current A^a_μ and the pseudoscalar density P^a,

$$\begin{aligned} A^a_\mu &= \frac{1}{2}\bar{\psi}\gamma_\mu\gamma_5\sigma^a\psi, \\ P^a &= \frac{1}{2}\bar{\psi}\gamma_5\sigma^a\psi. \end{aligned} \tag{2.21}$$

The dimension-four operators needed to improve A^a_μ are

$$\begin{aligned} O^a_{4a,\mu} &= \frac{1}{2}\bar{\psi}\gamma_5\sigma_{\mu\nu}\left(\overrightarrow{D}_\nu - \overleftarrow{D}_\nu\right)\sigma^a\psi, \\ O^a_{4b,\mu} &= \frac{1}{2}\partial_\mu\left(\bar{\psi}\gamma_5\sigma^a\psi\right) \\ O^a_{4c,\mu} &= \frac{m}{2}\bar{\psi}\gamma_\mu\gamma_5\sigma^a\psi. \end{aligned} \tag{2.22}$$

Again, as before, we can apply the Dirac equation to show $O_{4a,\mu}$ is linearly dependent on the other two, and again $O_{4c,\mu}$ is equal to the original current up to the factor m, so it can be absorbed in a redefinition. The improved current is therefore

$$A^a_{\mathrm{imp},\mu} = A^a_\mu + c_A a\hat{\partial}_\mu P^a, \tag{2.23}$$

where $c_A \in \mathbb{R}$ and $\hat{\partial}$ is the symmetric difference discretization of the derivative. From perturbation theory one finds

$$c_A = -0.00756g^2 + \mathcal{O}\big(g^4\big). \tag{2.24}$$

With this discussion we see that the operator $O_{4c,\mu}$ can be absorbed into a redefinition of P^a, i.e.

$$P^a_{\mathrm{imp}} = P^a. \tag{2.25}$$

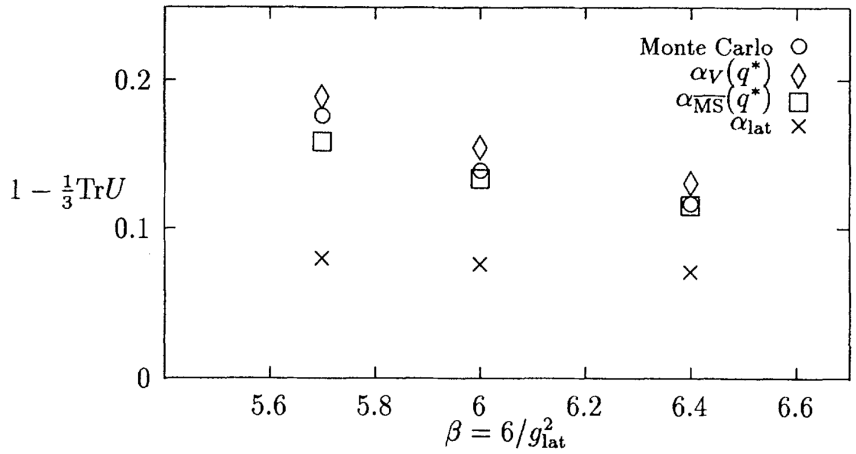

Figure 2.3: Expectation value of the trace of a link, evaluated in the Landau gauge. Results are calculated by MCMC and in first-order perturbation theory, expanding in two "good" parameters (see Ref. [8]) and the lattice coupling. These were carried out for quenched configurations with $\beta \geq 5.7$. Image taken from Ref. [8].

2.2 Tadpole improvement

In the early 1990's, physicists working in lattice field theory encountered a somewhat troubling problem: Many quantities calculated perturbatively in the bare coupling constant $g \equiv g_{\rm lat}$ turned out to differ substantially from MCMC results, even for rather small lattice spacing. An example of the failure was calculated by Lepage and Mackenzie [8], and is shown in Fig. 2.3. In this paper, the authors argue that $g_{\rm lat}$ is a poor expansion parameter, and they give two alternative couplings that converge better.

More pertinent to this discussion of action and interpolator improvement is the origin of this mismatch, which they also investigate. They argue as follows: In lattice perturbation theory one can expand the link variable as

$$U_\mu(x) = \exp\left[igaA^a_\mu(x)T^a\right] = \mathbf{1} + iagA^a_\mu(x) + \mathcal{O}\left(a^2g^2\right). \tag{2.26}$$

Superficially it seems as though we always get some power of ag, but for tadpole diagrams, their divergence in a exactly cancels the power of a coming from the vertex. These diagrams therefore are suppressed by a power of g

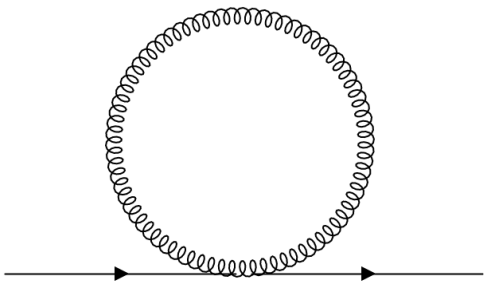

Figure 2.4: A tadpole diagram contributing to the fermion self-energy.

only. An example tadpole is shown in Fig. 2.4. These tadpoles come from the UV part of loop integrals. To obtain a better behaved perturbative expansion, one could imagine factoring out (integrating out) this UV contribution, so that the perturbative expansion would instead read

$$U_\mu(x) \approx u_0 \left(\mathbf{1} + iagA_\mu^{\mathrm{IR},a}(x)T^a\right), \tag{2.27}$$

with the *tadpole factor* u_0 and with the gauge field only including IR modes. Here u_0 has to be a constant[2], and it represents the average value of the link.

In Section 2.1, improvement coefficients were determined using perturbation theory. One can imagine that our coefficients will be better determined taking the influence of tadpoles into account. This is called *tadpole improvement*, and to employ it, we need to determine ${u_0}^3$. One could imagine extracting it from the average link; to do this one must fix to a particular gauge, since a single link is a gauge-dependent quantity that would therefore otherwise vanish. One can also measure it from gauge-invariant constructions such as

$$u_0 = \left(\frac{1}{N_c}\left\langle \mathrm{tr}\, U^{\square}\right\rangle\right)^{1/4}. \tag{2.28}$$

Tadpole improvement then amounts to replacing all links U_μ in lattice expressions with U_μ/u_0.

As a simple example we present the tree-level, tadpole-improved, one-flavor clover Lagrangian density. We rescale the action by multiplying by

[2]Otherwise we are not guaranteed that $U_\mu(x)$ will work correctly as a gauge connection.
[3]In general u_0 depends on the parameters of the theory.

the hopping parameter

$$2\kappa = \frac{1}{m + 4/a}, \tag{2.29}$$

rescaling the fermion fields by the factor $\sqrt{m + 4/a}$, and rescaling the link variables by the tadpole factor to get

$$\begin{aligned} \mathcal{L}^{\text{tad}}_{\text{clover}} = \; & \bar{\psi}(x)\psi(x) \\ & - \frac{\kappa}{u_0}\frac{1}{a}\bar{\psi}(x) \sum_{\mu=\pm 1}^{\pm 4} (\mathbf{1} - \gamma_\mu)\, U_\mu(x)\delta(x + a\hat{\mu} - y)\, \psi(y) \\ & + \frac{\kappa}{u_0}\frac{c_{\text{SW}}}{u_0^3}\, a^5 \sum_{\mu<\nu} \bar{\psi}(x)\sigma_{\mu\nu}\hat{F}_{\mu\nu}\psi(x). \end{aligned} \tag{2.30}$$

Using this Lagrangian in a simulation is exactly like using the original clover Lagrangian in a simulation; one just has a new hopping parameter $\kappa' = \kappa/u_0$. From eq. (2.20), one finds the tree-level, tadpole-improved Sheikholeslami-Wohlert coefficient to be simply

$$c_{SW} = \frac{1}{u_0^3}. \tag{2.31}$$

One can improve beyond tree-level; for details I refer the reader to Ref. [1].

2.3 Fat links

Typically the gauge connection between two neighboring sites x and y on the lattice is just a single link $U(x, y)$, which is in some sense the most local connection imaginable. One can also relax this locality, so that the gauge connection contains information from a larger region around x and y; for example the connection could depend on a general sum, including many paths connecting x and y. Let's call this sum $\Sigma(x, y)$. Then the gauge connection could be $V(x, y)$, where V is chosen by extremizing $\text{tr}\, V\Sigma^\dagger$. These gauge connections are called *fat links* [9]. We encountered fat links already in Section 1.1, where they were employed to reduce taste-breaking effects, and indeed one of the reasons to use a fat link is to correct hadron spectra. It should not surprise you that fat links modify particle spectra, since they amount to a change of the lattice propagator.

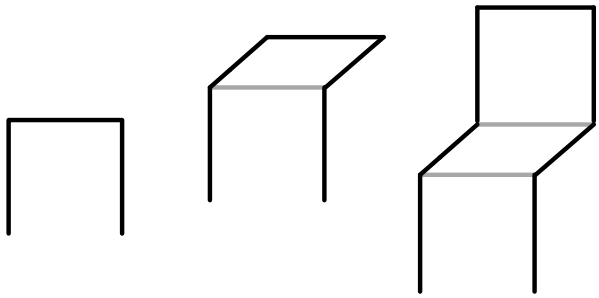

Figure 2.5: Example three-link, five-link, and seven-link constructs that go into the AsqTad fat links. Light grey lines guide the eye.

Replacing a link by some weighted sum connecting x and y is consistent with gauge invariance and hypercube symmetries of the lattice. These are smeared gauge links. Replacing a link by a weighted average that is displaced in directions perpendicular to the direction of the connection is a discretization of higher order derivatives $\partial_\nu^2 A_\mu$ [9] with $\mu \neq \nu$. For example the use of a fat link might amount to a change

$$U_\mu \to \left[1 + \frac{\epsilon a^2 \Delta^2}{n}\right]^n U_\mu, \tag{2.32}$$

where n is a small integer and ϵ is a smearing parameter [10]. This changes the gluon propagator by a factor $\exp[-\epsilon a^2 q^2]$ in momentum space, i.e. UV modes are significantly attenuated. One then sees a connection to Symanzik improvement: By choosing ϵ carefully, one can adjust, e.g., $\mathcal{O}(a^2)$ corrections.

Here we present the AsqTad action [11; 12], a staggered fermion action with nearest- and third-nearest-neighbor interactions

$$\Delta_\mu - \frac{a^2}{6}\Delta_\mu^3, \tag{2.33}$$

i.e. it uses the Naik term. The third-nearest-neighbor uses an ordinary link, but the nearest-neighbor uses the fat link

$$V_\mu = c_1 U_\mu + \sum_\nu \left[w_3 S_{\mu\nu}^{(3)} + \sum_\rho \left(w_5 S_{\mu\nu\rho}^{(5)} + \sum_\sigma w_7 S_{\mu\nu\rho\sigma}^{(7)} \right) + w_L S_{\mu\nu}^{(L)} \right], \tag{2.34}$$

where the link constructs are given by

$$\begin{aligned} S^{(3)}_{\mu\nu}(x) &= U_\nu(x)U_\mu(x+a\hat{\nu})U^\dagger_\nu(x+a\hat{\mu}), \\ S^{(5)}_{\mu\nu\rho}(x) &= U_\nu(x)S^{(3)}_{\mu\rho}(x+a\hat{\nu})U^\dagger_\nu(x+a\hat{\mu}), \\ S^{(7)}_{\mu\nu\rho\sigma}(x) &= U_\nu(x)S^{(5)}_{\mu\rho\sigma}(x+a\hat{\nu})U^\dagger_\nu(x+a\hat{\mu}), \\ S^{(L)}_{\mu\nu}(x) &= U_\nu(x)S^{(3)}_{\mu\nu}(x+a\hat{\nu})U^\dagger_\nu(x+a\hat{\mu}). \end{aligned} \tag{2.35}$$

Examples of these constructs are shown in Fig. 2.5. The tadpole-improved coefficients are

$$c_1 = 8w_5 = 48w_7 = \frac{1}{8}w_L = -\frac{1}{16} \qquad \text{and} \qquad w_3 = -\frac{5}{16}. \tag{2.36}$$

References

[1] T. DeGrand and C. DeTar. *Lattice methods for quantum chromodynamics*. World Scientific, 2006. ISBN 978-981-256-727-7.

[2] C. Gattringer and C. B. Lang. *Quantum Chromodynamics on the Lattice*. Springer, Berlin, 2010. ISBN 978-3-642-01849-7.

[3] K. Symanzik. Continuum limit and improved action in lattice theories: (i). principles and ϕ^4 theory. *Nucl. Phys. B*, 226(1):187–204, 1983. doi: https://doi.org/10.1016/0550-3213(83)90468-6. URL http://www.sciencedirect.com/science/article/pii/0550321383904686.

[4] K. Symanzik. Continuum limit and improved action in lattice theories: (II). O(N) non-linear sigma model in perturbation theory. *Nucl. Phys. B*, 226(1):205 – 227, 1983. doi: https://doi.org/10.1016/0550-3213(83)90469-8. URL http://www.sciencedirect.com/science/article/pii/0550321383904698.

[5] G. Curci, P. Menotti, and G. Paffuti. Symanzik's improved lagrangian for lattice gauge theory. *Physics Letters B*, 130(3):205–208, 1983-10. ISSN 03702693. doi: 10.1016/0370-2693(83)91043-2. URL https://linkinghub.elsevier.com/retrieve/pii/0370269383910432.

[6] M. Lüscher and P. Weisz. On-shell improved lattice gauge theories. *Comm. Math. Phys.*, 97(1):59–77, 1985. URL https://projecteuclid.org:443/euclid.cmp/1103941978.

[7] B. Sheikholeslami and R. Wohlert. Improved continuum limit lattice action for QCD with Wilson fermions. *Nucl. Phys. B*, 259:572–576, 1985. doi: 10.1016/0550-3213(85)90002-1. URL https://doi.org/10.1016/0550-3213(85)90002-1.

[8] G. P. Lepage and P. B. Mackenzie. Viability of lattice perturbation theory. 48(5):2250–2264, 1993. ISSN 0556-2821. doi: 10.1103/PhysRevD.48.2250. URL https://link.aps.org/doi/10.1103/PhysRevD.48.2250.

[9] T. Blum, C. DeTar, S. Gottlieb, K. Rummukainen, U. M. Heller, J. E. Hetrick, D. Toussaint, R. L. Sugar, and M. Wingate. Improving flavor symmetry in the kogut-susskind hadron spectrum. 55(3):R1133–R1137, 199. ISSN 0556-2821, 1089-4918. doi: 10.1103/PhysRevD.55.R1133. URL https://link.aps.org/doi/10.1103/PhysRevD.55.R1133.

[10] P. Lepage. Perturbative improvement for lattice QCD: An Update. *Nucl. Phys. B (Proc. Suppl.)*, 60:267–278, 1998. doi: 10.1016/S0920-5632(97)00489-1. URL https://doi.org/10.1016/S0920-5632(97)00489-1.

[11] K. Orginos, D. Toussaint, and R. L. Sugar. Variants of fattening and flavor symmetry restoration. 60(5):054503, 1999. ISSN 0556-2821, 1089-4918. doi: 10.1103/PhysRevD.60.054503. URL https://link.aps.org/doi/10.1103/PhysRevD.60.054503.

[12] G. P. Lepage. Flavor-symmetry restoration and symanzik improvement for staggered quarks. 59(7):074502, 1999. ISSN 0556-2821, 1089-4918. doi: 10.1103/PhysRevD.59.074502. URL https://link.aps.org/doi/10.1103/PhysRevD.59.074502.

Index