Introduction to lattice field theory

Stephen R. Sharpe Physics Department, University of Washington Seattle, WA 98195, USA

ABSTRACT

This is the written version of five lectures given at the Uehling Summer School for Phenomenology and Lattice QCD, University of Washington, June 21 - July 2, 1993. The lectures were initially transcribed by Henning Hoeber.

To appear in

"Phenomenology and Lattice QCD",

Eds. G. Kilcup and S. Sharpe, to be published by World Scientific.

February 1995

Contents

1	Introduction	2
2	Euclidean Field Theory 2.1 The spectrum	2 4
3	Scalar Fields 3.1 The propagator for free field theory	7 8
4	The Transfer Matrix4.1Problem 1—transfer matrix for one-dimensional field theory4.2Transfer matrix in four dimensions4.3Finite Temperature4.4Symmetries of \hat{T} and \hat{H} 4.5Problem 2: the cubic (octahedral) group and its representations	12 12 13 16 17 18
5	Gauge Theories on the Lattice5.1Continuum QCD, a brief overview5.2Discretization5.3Pure Gauge Action5.4Problem 3: the expansion of the plaquette.	 20 20 21 23 25
6	Applications of Lattice Gauge Theories6.1Confinement6.2Problem 4—strong coupling expansion6.3Glueball masses6.4Perturbation Theory and the Continuum Limit6.5Continuum Limit of Pure Gauge Theories6.6Comparison to critical phenomena6.7Relating lattice and continuum coupling constants	25 29 30 32 36 38 38
7	Fermions on the Lattice7.1Generality of the doubling problem7.2Consequences of the doubling problem7.3Wilson fermions	39 43 45 46
8	References	48

1 Introduction

These lectures are intended to provide a basic introduction to the lattice methods used today. They assume a working knowledge of field theory, and in particular of gauge theories and functional integrals, but no previous knowledge of the lattice approach. My aim is to provide the background necessary for understanding the applications of lattice methods, and in particular the subjects discussed in subsequent lectures at this school: phenomenologically useful predictions from QCD (Paul Mackenzie) and finite temperature physics (Akira Ukawa). I focus on the theoretical formulation, and do not discuss numerical methods.

In five lectures one can only cover a limited number of topics. Fortunately, there are now two excellent books available on Lattice Field Theory, by Rothe [1], and by Montvay and Münster [2]. These books take over where Creutz's seminal monograph left off a decade ago [3].

Homework problems were provided with each lecture, and these have either been integrated into the text, or placed at the end of the corresponding section. Some problems work out details not covered in the lectures, while others illustrate new concepts.

2 Euclidean Field Theory

I will focus mostly on the functional integral definition of field theory, since this is the method used in most numerical simulations. Let me begin by reviewing some basic results. All correlation functions can be determined from the partition function (with sources). In Minkowski space this is

$$Z_M = \int [d\phi] \exp\{iS_M(\phi) + \text{source terms}\},\qquad(1)$$

where ϕ denotes a generic field. The Minkowski action S_M is real and thus the integrand is complex. The integral is therefore extremely difficult to evaluate numerically because of cancelations between different regions of configuration space. The practical alternative is to work in Euclidean space where

$$Z_E = \int [d\phi] \exp\{-S_E(\phi_E) + \text{source terms}\}.$$
 (2)

In most theories the Euclidean action is real, and also bounded from below, in which case the exponential can be interpreted as a probability distribution on configuration space. After we discretize the theory, we can compute the integral using Monte-Carlo methods. QCD (without the CP-violating θ -term) is an example of a theory which can be simulated in this way. Examples of theories which cannot be so simulated, because they have complex actions, are QCD at finite chemical potential

(i.e. finite baryon number density) and theories with a chiral representation of fermions (e.g. the electroweak sector of the standard model).*

What we actually calculate in Euclidean space are correlation functions

$$C(x_1^E, x_2^E, \dots, x_n^E) = Z_E^{-1} \int [d\phi_E] e^{-S_E} \phi_E(\tau_1, \vec{x}_1) \phi_E(\tau_2, \vec{x}_2) \dots \phi_E(\tau_n, \vec{x}_n) , \qquad (3)$$

where τ denotes Euclidean "time". To obtain physical quantities we have, in general, to analytically continue the correlation functions back to Minkowski space, $\tau \to it$. It is worthwhile recalling how this is done. The functional integral is constructed to give τ -ordered expectation values

$$C(x_1^E, x_2^E, \dots, x_n^E) = \langle 0 | T[\hat{\phi}(\tau_1, \vec{x}_1) \hat{\phi}(\tau_2, \vec{x}_2) \dots \hat{\phi}(\tau_n, \vec{x}_n)] | 0 \rangle, \qquad (4)$$

where $\hat{\phi}(\tau, \vec{x})$ is the Heisenberg operator corresponding to ϕ . Assuming for simplicity that $\tau_1 > \tau_2 > \ldots \tau_n$, we can use Euclidean time translation (and $\hat{H}|0\rangle = 0$) to rewrite this expression as

$$C(x_1^E, x_2^E, \dots, x_n^E) = \langle 0 | \hat{\phi}(\vec{x}_1) e^{-\hat{H}(\tau_1 - \tau_2)} \hat{\phi}(\vec{x}_2) e^{-\hat{H}(\tau_2 - \tau_3)} \dots e^{-\hat{H}(\tau_{n-1} - \tau_n)} \hat{\phi}(\vec{x}_n)] | 0 \rangle .$$
(5)

Now the factors of τ are explicit, we can perform an inverse Wick-rotation, $\tau \to it$. After some rewriting, this results in the Minkowski time-ordered product

$$\langle 0|T[\hat{\phi}(t_1,\vec{x}_1)\hat{\phi}(t_2,\vec{x}_2)\dots\hat{\phi}(t_n,\vec{x}_n)]|0\rangle.$$
(6)

These time-ordered products are sufficient, via LSZ reduction, to determine all properties of the theory i.e. the spectrum and scattering amplitudes. If we insert operators other than $\hat{\phi}$ in the correlation functions we can also obtain their matrix elements.

This is fine in principle, but in practice we have numerical results for correlation functions on a discrete set of points (either in position space or momentum space), and analytic continuation is, at best, highly problematic. Thus a crucial practical issue is what can be learned directly from Euclidean correlation functions. As I demonstrate shortly, no analytic continuation is needed to extract the low lying spectrum of hadrons or the matrix elements of local operators involving single particles. Nearly all the phenomenologically useful results from lattice studies to date involve such quantities.

First, though, I want to mention a more formal point (see Refs. [2, 4] for further discussion). It is quite possible for a Euclidean-space functional integral to yield well behaved correlation functions, and yet for these functions not to be the analytic continuation of those of a physical Minkowski theory. By a physical theory I mean one having a Hilbert space with positive norm, whose spectrum is bounded from below, and on which Poincaré invariance is implemented by unitary operators.

^{*}In fact, the "fermion doubling problem", to be discussed below, makes it difficult to even formulate chiral theories on the lattice.

In particular there is an hermitian Hamiltonian which generates time translations. Since in the lattice enterprise we begin in Euclidean space, it is important to know under what conditions the corresponding Minkowski theory is physical. This question was studied long ago by Osterwalder and Schrader, who found the following [5]. If the action S_E is Euclidean invariant, and expectation values such as $C(x_1^E, \ldots, x_n^E)$ in Eq. 3 satisfy a property called "reflection positivity", plus some more technical conditions, then there exists a physical Minkowski theory such that the steps leading from Eq. 3 to Eq. 6 are valid. An important example of a theory which does *not* satisfy these conditions is "quenched" QCD—i.e. QCD without internal fermion loops. This is an approximation used in many simulations at present.

I do not have time to discuss reflection positivity here—see, for example, Ref. [2] for a clear exposition. What I will describe (in Chapter 4) is a standard technique for actually constructing the Hilbert space and Hamiltonian operator, using the "transfer matrix". In this way one can see explicitly the passage from Eq. 3 to Eq. 6, in the context of a discretized theory.

2.1 The spectrum

Let me now show how the spectrum can be obtained directly from two-point Euclidean correlation functions, without the need for analytic continuation. Translating one of the points to the origin, we begin with

$$C(x_E) = Z^{-1} \int [d\phi] e^{-S_E} \phi(\tau, \vec{x}) \phi(0) , \qquad (7)$$

$$= \langle 0|T[\hat{\phi}(x_E)\hat{\phi}(0)]|0\rangle.$$
(8)

Assuming $\tau > 0$, and using the full Euclidean translation operator

$$\hat{\phi}(\tau, \vec{x}) = e^{\hat{H}\tau - i\hat{\vec{p}}\cdot\vec{x}}\hat{\phi}(0)e^{-\hat{H}\tau + i\hat{\vec{p}}\cdot\vec{x}}, \qquad (9)$$

where $\hat{\vec{p}}$ is the momentum operator, plus the fact that the vacuum has neither energy nor momentum, we find

$$C(x_E) = \langle 0 | \hat{\phi}(0) e^{-\hat{H}\tau + i\vec{p}\cdot\vec{x}} \hat{\phi}(0) | 0 \rangle .$$
(10)

Inserting a complete set of states, we end up with the spectral decomposition

$$C(x_E) = \sum_{n} |\langle 0|\hat{\phi}(0)|n\rangle|^2 (2E_n V)^{-1} e^{-E_n x_4} e^{i\vec{p}_n \cdot \vec{x}_n} .$$
(11)

Here I am assuming a finite volume V, and using relativistically normalized states

$$\langle \vec{p} | \vec{q} \rangle = 2EV \delta_{p_x q_x} \delta_{p_y q_y} \delta_{p_z q_z} \xrightarrow{V \to \infty} 2E(2\pi)^3 \delta^3(\vec{p} - \vec{q}) , \qquad (12)$$

where $E^2 = |\vec{p}|^2 + m^2$.

<u>Exercise</u>: Show, assuming parity conservation, that for any τ

$$C(x_E) = \sum_{n} |\langle 0|\hat{\phi}(0)|n\rangle|^2 (2E_n V)^{-1} e^{-E_n |\tau|} e^{i\vec{p}_n \cdot \vec{x}_n}$$
(13)

From Eq. 13 we see that we can determine the spectrum directly from the exponential fall-off of the Euclidean correlator. To make things clearer let us project onto $\vec{p} = 0$

$$\int d^3x \, C(x_E) = C(\tau, \vec{p} = 0) = \sum_{\vec{n}, \vec{p} = 0}^{\infty} |\langle 0|\hat{\phi}|n\rangle|^2 (2E_n)^{-1} e^{-E_n|\tau|} \,. \tag{14}$$

If the lightest state produced from the vacuum by $\hat{\phi}$ (call it $|1\rangle$) is a single particle, then, for large $|\tau|$

$$C(\tau, \vec{p} = 0) = |\langle 0|\hat{\phi}|1\rangle|^2 (2m)^{-1} e^{-m|x_4|} + \text{ exponentially suppressed terms.}$$
(15)

Thus one can just read off the mass, along with the associated matrix element to create the state from the vacuum.

Clearly by judiciously choosing the operators in the two point function we can project onto states having different spin-parities and different momenta. In each channel it is simple to extract the energy of the lightest state, but progressively harder to pick out higher energy states, because their contributions are exponentially suppressed. Much effort in lattice simulations goes into fiddling with the operators so as to increase the overlap with the desired states.

Although one does not need to analytically continue, it is nevertheless true that picking out the exponential is equivalent to finding the pole in the propagator. For example, consider the Fourier transform of the contribution of the lightest state to the $\vec{p} = 0$ propagator

$$C(E) = \int d\tau \ e^{iE\tau} \ \frac{e^{-m|\tau|}}{2m} = \frac{1}{2m(m-iE)} + \frac{1}{2m(m+iE)} = \frac{1}{m^2 + E^2} \ .$$
(16)

Analytically continuing to Minkowski energies, $E \rightarrow -iE_0$, we find the usual pole in the propagator

$$C(E_0) = \frac{1}{m^2 - E_0^2} .$$
(17)

One can also extract physical information directly from Euclidean three-point functions. Consider the correlator

$$C_{3}(\tau_{1},\tau_{2},\tau_{3}) = Z^{-1} \int [d\phi] e^{-S_{E}} O_{1}(\tau_{1}) O_{2}(\tau_{2}) O_{3}(\tau_{3})$$

$$= \langle 0 | \hat{O}_{1}(\tau_{1}) e^{-\hat{H}(\tau_{1}-\tau_{2})} \hat{O}_{2}(\tau_{2}) e^{-\hat{H}(\tau_{2}-\tau_{3})} \hat{O}_{3}(\tau_{3}) | 0 \rangle, \qquad (18)$$

where $O_i(\tau_i)$ is a function of the fields at time τ_i , and $\hat{O}_i(\tau_i)$ is the corresponding Heisenberg operator. I have assumed $\tau_1 > \tau_2 > \tau_3$ in the second line. For $\tau_1 - \tau_2$ and $\tau_2 - \tau_3$ large, the correlator behaves as

$$C_3 \propto \langle 0|\hat{O}_1|1\rangle \exp(-E_1|\tau_1 - \tau_2|)\langle 1|\hat{O}_2|3\rangle \exp(-E_3|\tau_2 - \tau_3|)\langle 3|\hat{O}_3|0\rangle , \qquad (19)$$

where $|1\rangle$ and $|3\rangle$ are, respectively, the lightest states created from the vacuum by the operators \hat{O}_1 and \hat{O}_3 . These operators might include a projection onto non-zero spatial momenta, which is why I have written the coefficients in the exponents as energies rather than masses. The creation and destruction matrix elements, together with the energies, can be obtained from Euclidean two-point functions. Thus one can extract $\langle 1|\hat{O}_2|3\rangle$ from C_3 without analytic continuation.

Although it provides a way of thinking about four- and higher point functions, Eq. 5 is not very useful in practice. For example, one cannot use it to extract scattering amplitudes directly from four-point functions. Such amplitudes require analytic continuation: they are real in Euclidean space, yet complex, in general, in Minkowski space.[†] The result does, however, show the close relationship to a Hamiltonian approach, in which one calculates matrix elements like $\langle 0|\hat{O}_1H^n\hat{O}_2|0\rangle$.

[†]There is, however, an elegant indirect method due to Lüscher which uses the volume dependence of two particle energies[6].

3 Scalar Fields

Having understood what we can learn directly Euclidean field theories, I now turn to business of carefully defining them. I begin with the simplest example, the real scalar field. To define a field theory requires regularization. Replacing continuous space-time with a discrete lattice is one option; it corresponds to a (complicated) cut-off in momentum space. The messiness of the cut-off is compensated by the fact that one can perform the Euclidean functional integral for any values of the parameters in the action. In particular, none of the coupling constants need be assumed small, so that we can do non-perturbative calculations.

Although it is not essential, most calculations use lattices with equal spacing (a) in all four directions



Asymmetric lattices with different spacing in space and time can be useful in finite temperature calculations, as discussed in Ukawa's lectures.

The continuum Euclidean action for a real scalar field is

$$S = \int d^4x \left[\frac{1}{2}\partial_\mu \phi \,\partial_\mu \phi + V(\phi)\right] \tag{20}$$

where

$$V(\phi) = \frac{1}{2}m_{ph}^{2}\phi^{2} + \lambda\phi^{4} + \dots$$
 (21)

The subscript "ph" refers to the physical mass, to be distinguished from the lattice mass which will be introduced shortly. This action may be discretized in many possible ways. In the "continuum limit", $a \to 0$, the choice should be irrelevant, and we pick the simplest method. Continuous fields are replaced by dimensionless fields on the lattice sites labeled by a vector of integers $n = (n_1, n_2, n_3, n_4)$

$$a\phi(x) \to \phi_n$$
, (22)

and integrals become sums

$$\int d^4x \to a^4 \sum_n \ . \tag{23}$$

For lattice derivatives there are several possibilities

$$a^2 \partial_\mu \phi \rightarrow \Delta^+_\mu \phi_n = \phi_{n+\hat{\mu}} - \phi_n$$
, or (24)

$$\Delta^{-}_{\mu}\phi_n = \phi_n - \phi_{n-\hat{\mu}}, \text{ or }$$
(25)

$$\Delta_{\mu}\phi_{n} = \frac{1}{2}\phi_{n+\hat{\mu}} - \phi_{n-\hat{\mu}} , \dots$$
 (26)

We require the lattice version of $(\partial_{\mu}\phi)^2$ to be positive. The most local choice is

$$\int (\partial_{\mu}\phi)^{2} \to \sum_{n} (\phi_{n+\hat{\mu}} - \phi_{n})^{2} = \sum_{n} (\Delta_{\mu}^{+}\phi_{n})^{2} = \sum_{n} (\Delta_{\mu}^{-}\phi_{n})^{2} .$$
(27)

Sometimes it is useful to rewrite this term as follows

$$\sum_{n,\mu} (\phi_{n+\hat{\mu}} - \phi_n)^2 = -\sum_{n,\mu} \phi_n (\phi_{n+\hat{\mu}} - 2\phi_n + \phi_{n-\hat{\mu}}), \qquad (28)$$

which is the lattice analogue of $\int (\partial_{\mu} \phi)^2 = -\int \phi \Box \phi$. Putting all this together

$$S \to S_L = \sum_{n,\mu} -\frac{1}{2} (\phi_{n+\hat{\mu}} - 2\phi_n + \phi_{n-\hat{\mu}}) \phi_n + \sum_n V_L(\phi_n)$$
(29)

with

$$V_L(\phi_n) = \frac{1}{2} (m_{ph} a)^2 \phi_n^2 + \lambda \phi_n^4 + \dots$$
 (30)

The product $m_{ph}a$ is the dimensionless lattice mass m. Finally, the lattice partition function is

$$Z = \prod_{n} \left[\int_{-\infty}^{+\infty} d\phi_{n} \right] e^{-S_{L}} \equiv \int_{\phi} e^{-S_{L}} .$$
 (31)

It is important to realize the meaning of the arrows in the above equations. The lattice action has the correct continuum limit for "classical" fields, those varying on scales much longer than the lattice spacing. The quantum theory, by contrast, necessarily includes the contribution from "jagged" fields, those that fluctuate on the scale of the lattice spacing. This is true even if one is calculating the response to a smooth external fields, since interactions will couple them to quantum fields of all wavelengths. Thus the continuum limit of the quantum theory need not be described by the classical continuum action. In asymptotically free theories the jagged modes can be studied using perturbation theory, and the continuum limit can be understood, as described below. This is not possible for scalar theories. Indeed, it is almost certain that, whatever the size of the ϕ^4 coupling λ in the lattice Lagrangian, the particles do not interact in the continuum limit. This is called "triviality".

3.1 The propagator for free field theory

To study the properties of the lattice theory, I begin by setting $\lambda = 0$, and calculating the propagator from site n to site p

$$C(n,p) = Z^{-1} \int_{\phi} e^{-S} \phi_n \phi_p \,. \tag{32}$$

For the free theory the action is bilinear in the fields

$$S = \frac{1}{2} \sum_{n,p} \phi_n M_{np} \phi_p \quad , \tag{33}$$

with M real and symmetric

$$M_{np} = -\sum_{\mu=1}^{4} (\delta_{n,p+\hat{\mu}} + \delta_{n,p-\hat{\mu}}) + \delta_{np}(8+m^2) .$$
(34)

Performing the Gaussian integrals we obtain

$$C(n,p) = \frac{[\det M]^{-1/2} (M^{-1})_{np}}{[\det M]^{-1/2}} = (M^{-1})_{np} .$$
(35)

As usual, it is easiest to Fourier transform

$$\phi_n = \int_{-\pi}^{\pi} \frac{d^4k}{(2\pi)^4} e^{ikn} \phi(k) \equiv \int_k e^{ikn} \phi(k) , \quad \phi(-k) = \phi^*(k) , \quad (36)$$

where the second form of the integral is a useful abbreviation. Note that the lattice cuts off the momentum integral. All lattice quantities are periodic in k_{μ} (separately for each μ) with period 2π ; I have chosen the range of integration to be symmetric about the origin. Substituting Eq. 36 into the action we find

$$S = \frac{1}{2} \sum_{n,p} \int_{k,q} e^{ik \cdot n} e^{iq \cdot p} M_{np} \phi(k) \phi(q)$$

$$= \frac{1}{2} \sum_{n} \int_{k} \int_{q} e^{i(k+q) \cdot n} [m^{2} + \sum_{\mu} (2 - e^{iq_{\mu}} - e^{-iq_{\mu}})] \phi(k) \phi(q)$$

$$= \frac{1}{2} \int_{k} \phi(k) \phi(-k) [m^{2} + \sum_{\mu} 4 \sin^{2}(\frac{k_{\mu}}{2})] .$$
(37)
(37)
(37)

M is now diagonal, so the propagator is

$$C(n,p) = M_{np}^{-1} = \int_{k} \frac{e^{ik(n-p)}}{m^2 + \sum_{\mu} 4\sin^2\frac{k_{\mu}}{2}} .$$
(39)

Various features of this result are noteworthy.

• The lattice and continuum propagators differ only in that the momentum is changed

$$k_{\mu} \longrightarrow 2\sin\left(\frac{k_{\mu}}{2}\right) \equiv \hat{k}_{\mu}$$
 (40)

Although the lattice momentum \hat{k}_{μ} is *anti*-periodic when k_{μ} shifts by 2π , what appears in the propagator is its square, which is periodic.

• The lattice propagator is symmetric under discrete Euclidean rotations (e.g. $x \to y, y \to -x$) but not under continuous rotations. One recovers the full symmetry if the continuum limit is taken as follows:

$$m = m_{ph}a \to 0, \quad k = k_{ph}a \to 0 \quad , \tag{41}$$

with the physical mass and momentum, m_{ph} and k_{ph} , held fixed. In this limit

$$\hat{k}_{\mu} \to (k_{ph})_{\mu}a , \ a^2 C(k) \to \frac{1}{(m_{ph})^2 + (k_{ph})^2},$$
(42)

which is the continuum free propagator. This continuum limit is rather trivial, but does illustrate one general point: one must adjust the parameters in the Lagrangian in a particular way to attain the continuum. Here, the lattice mass m must vanish, otherwise one ends up with infinitely heavy continuum scalars.

It is enlightening to calculate the propagator in position space for a given spatial momentum. As shown in the sect. 2.1, this allows one to read off the spectrum. Consider then

$$C(n_{4};\vec{q}) = Z^{-1} \int [d\phi] e^{-S_{L}} \sum_{\vec{n}} e^{-i\vec{q}\cdot\vec{n}} \phi(\vec{n}, n_{4})\phi(0)$$

$$= \sum_{\vec{n}} e^{-i\vec{q}\cdot\vec{n}} \int_{k} \frac{e^{ik_{4}n_{4}} e^{i\vec{k}\cdot\vec{n}}}{m^{2} + \hat{k}_{4}^{2} + \hat{\vec{k}}^{2}}$$

$$= \int_{-\pi}^{\pi} \frac{dk_{4}}{2\pi} \frac{e^{ik_{4}n_{4}}}{m^{2} + \hat{\vec{q}}^{2} + \hat{k}_{4}^{2}}$$

$$(44)$$

The integrand has poles on the imaginary axis. Defining $k_4 = iE$, so that $\hat{k}_4 = 2i \sinh \frac{E}{2}$, the poles are determined by

$$2\sinh\frac{E}{2} = \pm\sqrt{m^2 + \hat{\vec{q}}^2}$$
 (45)

We can perform the integration by forming a closed contour in the complex k_4 plane:



Since the integrand is periodic in k_4 , with period 2π , the two vertical parts of the contour cancel, and so can be added to the original integral. For $n_4 > 0$, one must close the contour in the upper half plane, and one picks up the corresponding residue, leading to

$$C(n_4, \vec{q}) = \frac{e^{-En_4}}{2\sinh E} .$$
(46)

We see that, even with discretized time, the correlator does fall off exponentially. In fact, there is only one exponential, corresponding to the fact that the theory is free, so there is only one state for each momentum. Note that the lattice energy (Eq. 45) differs from the continuum result $E^2 = m^2 + \vec{q}^2$, although the two agree in the continuum limit $(m, \vec{q} \to 0)$.

A very useful concept in lattice calculations is the "correlation length", ξ . This is defined by the rate of exponential fall-off of two point functions in position space: $G(n,p) \sim \exp(-|n-p|/\xi)$, up to powers of the separation |n-p|. Two point functions include all possible values of \vec{q} ; at large distances the smallest value of Econtrols the rate of decay. This is $E_{min} = 2\sinh^{-1}(m/2) \approx m$, so that $\xi \approx 1/m$. One way of thinking about the continuum limit is that ξ must diverge in lattice units, so that the effects of discretization disappear. Sending $m = m_{ph}a \to 0$ indeed makes $\xi \to \infty$.

In an interacting theory, the correlation length(s) are determined, in general, non-perturbatively in terms of the parameters in the lattice action. To find a continuum limit one must discover a place in the space of parameters for which the correlation length diverges.

4 The Transfer Matrix

The nuts-and-bolts way in which one connects lattice Euclidean functional integrals and the corresponding Minkowski theory relies on the transfer matrix, \hat{T} . In other words, using \hat{T} allows one to see exactly what one is calculating when working at finite lattice spacing. In particular, the symmetries of the spectrum are those of T. The transfer matrix is also widely used in statistical mechanics calculations. For these reasons I have chosen to describe its construction, for the simplest case of scalar fields. The generalizations to gauge theories and to theories with fermions are more complicated to work out, but the essential idea is the same.

I begin with a homework problem which serves as a warm-up for the actual analysis.

4.1 Problem 1—transfer matrix for one-dimensional field theory

Consider a 1-dimensional lattice with scalar fields ϕ_n , n = 1, 2, ..., N and periodic boundary conditions $\phi_{N+1} = \phi_1$. Find an operator \hat{T} such that

$$Z = \mathsf{Tr}(\hat{T}^N) \, .$$

Here \hat{T} is an operator acting on the Hilbert space of square integrable functions of one variable, i.e. the usual space for quantum mechanics of a single variable.

A simple way to proceed is as follows.

Step 1: Write the action as

$$S = S_A(\phi_1) + S_B(\phi_1, \phi_2) + S_A(\phi_2) + S_B(\phi_2, \phi_3) + \dots + S_A(\phi_N) + S_B(\phi_N, \phi_1)$$

 S_B contains terms connecting adjacent ϕ 's; it is advantageous to choose S_B to be positive.

Step 2: It is simplest to use a non-normalizable basis $|\phi\rangle$ for the Hilbert-space :

$$\langle \phi | \phi \rangle = \phi | \phi \rangle$$
, $\langle \phi' | \phi \rangle = \delta(\phi' - \phi)$

Introduce a conjugate momentum \hat{p} with $[\hat{p}, \phi] = -i$, and show that

$$e^{-i\hat{p}\Delta}|\phi\rangle = |\phi + \Delta\rangle.$$

With all this in hand, find an operator \hat{T} which satisfies

$$\langle \phi' | \hat{T} | \phi \rangle = e^{S_B(\phi',\phi)} e^{\frac{1}{2}S_A(\phi)} e^{\frac{1}{2}S_A(\phi')}$$

Step 3: Given this definition of \hat{T} show that

$$Z = \operatorname{Tr}(\hat{T}^{N}) \equiv \int d\phi \langle \phi | \hat{T}^{N} | \phi \rangle .$$
(47)

<u>Step 4:</u> Since \hat{T} is the operator which translates by one unit in Euclidean time, it is reasonable to define $\hat{T} = e^{-\hat{H}a}$ (putting back in the lattice spacing). As $a \to 0$, $\hat{H} \to \hat{H}_{cont} + 0(a)$. Find the form of \hat{H}_{cont} (it should be familiar).

4.2 Transfer matrix in four dimensions

Constructing the transfer matrix in $d \ge 2$ differs from the one dimensional case mainly in the need for extra notation. I begin with the partition function

$$Z = \int_{\phi} e^{-S} , \quad S = \sum_{n} \left[\sum_{\mu=1}^{4} \frac{1}{2} (\phi_{n+\mu} - \phi_n)^2 + V(\phi_n) \right] .$$
(48)

It is necessary to work with a lattice having a finite extent in time:



I assume periodic boundary conditions, $\phi_{\vec{n},N+1} = \phi_{\vec{n},1}$, where N is the number of "timeslices". This amounts to working at a finite temperature 1/Na, as discussed below.

The idea is to break the integrand of the functional integral up into two parts: those living on individual timeslices, and those connecting adjacent slices. Since the action includes only nearest neighbor interactions, this is all one needs. Thus I collect all the fields at time τ_i (having all spatial positions) into a set denoted $\{\phi_{\tau_i}\}$, and write the action as[‡]

$$S = S_A(\{\phi_{\tau_N}\}) + S_B(\{\phi_{\tau_N}\}, \{\phi_{\tau_{N-1}}\}) + S_A(\{\phi_{\tau_{N-1}}\}) + S_B(\{\phi_{\tau_{N-1}}\}, \{\phi_{\tau_{N-2}}\})$$

... + S_A(\{\phi_{\tau_1}\}) + S_B(\{\phi_{\tau_1}\}, \{\phi_{\tau_N}\}) (49)

[‡]Note that I am now using τ instead of n_4 to label *dimensionless*, discretized Euclidean time. This saves on subscripts, and should lead to no confusion with my earlier use of τ as dimensionful, continuous Euclidean time.

where

$$S_A(\{\phi_{\tau_i}\}) = \sum_{\vec{n}} \left[V(\phi_{\tau_i,\vec{n}}) + \frac{1}{2} \sum_{j=1,3} (\phi_{\tau_i,\vec{n}+\hat{j}} - \phi_{\tau_i,\vec{n}})^2 \right]$$
(50)

$$S_B(\{\phi_{\tau_1}\},\{\phi_{\tau_2}\}) = \frac{1}{2} \sum_{\vec{n}} (\phi_{\tau_1,\vec{n}} - \phi_{\tau_2,\vec{n}})^2$$
(51)

Next define the transfer function

$$T_{2,1} = T(\{\phi_{\tau_2}\}, \{\phi_{\tau_1}\}) = e^{-\frac{1}{2}S_A(\{\phi_{\tau_2}\})} e^{-S_B(\{\phi_{\tau_2}\}, \{\phi_{\tau_1}\})} e^{-\frac{1}{2}S_A(\{\phi_{\tau_1}\})},$$
(52)

which allows one to express the partition function as

$$Z = \int_{\phi} e^{-S} = \int_{\{\phi_{\tau_1}\}\dots\{\phi_{\tau_N}\}} T_{1,N} T_{N,N-1}\dots T_{2,1}$$
(53)

We now introduce the Hilbert space of square integrable functions at each site \vec{n} with (non-normalizable) basis $|\phi_{\vec{n}}\rangle$, and "position" and momentum operators satisfying

$$\hat{\phi_{\vec{n}}}|\phi_{\vec{n}}\rangle = \phi_{\vec{n}}|\phi_{\vec{n}}\rangle, \quad \langle \phi_{\vec{n}}'|\phi_{\vec{n}}\rangle = \delta(\phi_{\vec{n}}' - \phi_{\vec{n}}), \qquad (54)$$

$$[\hat{\phi}_{\vec{n}}, \hat{\phi}_{\vec{n}'}] = 0 , \quad [\hat{p}_{\vec{n}}, \hat{\phi}_{\vec{n}'}] = -i\delta_{\vec{n}, \vec{n}'} , \quad e^{-i\hat{p}_{\vec{n}}\Delta} |\phi_{\vec{n}}\rangle = |\phi_{\vec{n}} + \Delta\rangle .$$
(55)

The transfer matrix acts on the Hilbert space of direct product states

$$\phi\rangle \equiv \prod_{\vec{n}} \otimes |\phi_{\vec{n}}\rangle \ . \tag{56}$$

The idea is to find an operator $\hat{T}(\hat{\phi}_{\vec{n}},\hat{p}_{\vec{n}})$ satisfying

$$\langle \phi' | \hat{T} | \phi \rangle = T(\{\phi'\}, \{\phi\}), \qquad (57)$$

for then we can write the partition function as

$$Z = \int_{\{\phi_{\tau_1}\}\dots\{\phi_{\tau_N}\}} \langle \phi_{\tau_1} | \hat{T} | \phi_{\tau_N} \rangle \langle \phi_{\tau_N} | \hat{T} | \phi_{\tau_{N-1}} \rangle \dots \langle \phi_{\tau_2} | \hat{T} | \phi_{\tau_1} \rangle$$

$$= \int_{\{\phi_{\tau_1}\}} \langle \phi_{\tau_1} | \hat{T}^N | \phi_{\tau_1} \rangle$$

$$= \operatorname{Tr}(\hat{T}^N).$$
(58)

Here the second line follows from completeness, and the last from the definition of the trace.

The construction of \hat{T} is similar to the one-dimensional case. The trick is to note that

$$e^{-\frac{1}{2}(\phi'_{\vec{n}} - \phi_{\vec{n}})^{2}} = \int_{-\infty}^{\infty} d\Delta \langle \phi'_{\vec{n}} | e^{-\frac{\Delta^{2}}{2}} e^{-i\hat{p}_{\vec{n}}\Delta} | \phi_{\vec{n}} \rangle$$

$$= \sqrt{2\pi} \langle \phi'_{\vec{n}} | e^{-\frac{1}{2}p_{\vec{n}}^{2}} | \phi_{\vec{n}} \rangle , \qquad (59)$$

so that the following does the job

$$\hat{T} = e^{-\frac{1}{2}S_A(\{\hat{\phi}\})} e^{-\frac{1}{2}\sum_{\vec{n}}\hat{p}_{\vec{n}}^2} e^{-\frac{1}{2}S_A(\{\hat{\phi}\})} (2\pi)^{\frac{L^3}{2}} .$$
(60)

What has been achieved here? We have rewritten the functional integral in terms of an operator \hat{T} which "transfers" information from one time slice to the next. As we will see more clearly below, it corresponds to the Euclidean space time translation operator for one lattice unit, which in the continuum is $\exp(-a\hat{H})$. \hat{T} is of the form $A^{\dagger}A$, and so is hermitean and positive. It thus has real, positive eigenvalues: $\hat{T}|p\rangle = \lambda_p |p\rangle$, $\lambda_p \geq 0$. Labeling the eigenvalues in order of decreasing size[§]

$$\lambda_0 > \lambda_1 > \ldots > 0$$

the partition function becomes

$$Z = \operatorname{Tr} \, \hat{T}^N = \lambda_0^N + \lambda_1^N + \dots \xrightarrow{N \to \infty} \lambda_0^N (1 + (\frac{\lambda_1}{\lambda_0})^N + \dots)$$
(61)

The eigenstate with the *largest* eigenvalue is picked out for large N. This state corresponds in the continuum limit to that having the smallest eigenvalue of \hat{H} . It is the lattice vacuum state, $|0\rangle$.

The significance of \hat{T} can be seen more clearly by applying the above construction to correlation functions. Consider the two-point function

$$C_2(n_1, n_2) = Z^{-1} \int_{\phi} e^{-S_L} \phi_{\tau_2, \vec{n_2}} \phi_{\tau_1, \vec{n_1}} \qquad (\tau_2 > \tau_1)$$
(62)

which can be rewritten in terms of the transfer matrix

$$C_2(n_1, n_2) = Z^{-1} \int_{\phi} \hat{T}_{1,N} \dots \hat{T}_{\tau_2+1,\tau_2} \phi_{\tau_2, \vec{n}_2} \hat{T}_{\tau_2, \tau_2-1} \dots \hat{T}_{\tau_1+1,\tau_1} \phi_{\tau_1, \vec{n}_1} \hat{T}_{\tau_1, \tau_1-1} \dots \hat{T}_{2,1}.$$
(63)

Using

$$\langle \phi' | \hat{\phi}_{\vec{n}} \hat{T} | \phi \rangle = \phi'_{\vec{n}} T(\{\phi'\}, \{\phi\})$$
(64)

we find

$$C_{2}(n_{1}, n_{2}) = Z^{-1} \int_{\phi} \langle \phi_{\tau_{1}} | \hat{T} | \phi_{\tau_{n}} \rangle \dots \langle \phi_{\tau_{2}} | \hat{\phi}_{\vec{n}_{2}} \hat{T} | \phi_{\tau_{2}-1} \rangle \dots \langle \phi_{\tau_{1}} | \hat{\phi}_{\vec{n}_{1}} \hat{T} | \phi_{\tau_{1}-1} \rangle \dots \langle \phi_{\tau_{2}} | \hat{T} | \phi_{\tau_{1}} \rangle$$

$$= \operatorname{Tr} \{ \hat{\phi}_{\vec{n}_{2}} \hat{T}^{\tau_{2}-\tau_{1}} \hat{\phi}_{\vec{n}_{1}} \hat{T}^{N-\tau_{2}+\tau_{1}} \} / \operatorname{Tr} \{ \hat{T}^{N} \}$$

$$= \operatorname{Tr} \{ \hat{\phi}_{\vec{n}_{2}} \overline{T}^{\tau_{2}-\tau_{1}} \hat{\phi}_{\vec{n}_{1}} \overline{T}^{N-\tau_{2}+\tau_{1}} \} / \operatorname{Tr} \{ \overline{T}^{N} \} .$$
(65)

In the last step I have rescaled the transfer matrix so that the eigenvalues lie between 0 and 1: $\hat{}$

$$\overline{T} = \frac{\overline{T}}{\lambda_0}$$
 has eigenvalues $1 > \overline{\lambda}_1 = \frac{\lambda_1}{\lambda_0} > \dots \ge 0$. (66)

[§]For a finite system, with a potential $V(\phi)$ bounded from below, there is a finite maximum eigenvalue.

Taking the limit $N \to \infty$, keeping $\tau_2 - \tau_1$ fixed, selects the "vacuum" in both numerator and denominator since $\overline{T}^N |0\rangle = 1$ while $\overline{T}^N |1\rangle = \overline{\lambda}^N |1\rangle \xrightarrow{N \to \infty} 0$. Thus

$$C_2(n_1, n_2) \xrightarrow{N \to \infty} \langle 0 | \hat{\phi}_{\vec{n}_2} \overline{T}^{\tau_2 - \tau_1} \hat{\phi}_{\vec{n}_1} | 0 \rangle .$$
(67)

Comparing to the continuum expression (Eq. 5)

$$C_2 = \langle 0 | \hat{\phi}_{\vec{n}_2} e^{-\hat{H}(\tau_2 - \tau_1)a} \hat{\phi}_{\vec{n}_1} | 0 \rangle, \qquad (68)$$

we see that the correspondence is $\overline{T} = e^{-\hat{H}a}$. Clearly the same applies to multi-point correlation functions. Thus it makes sense to *define* a lattice Hamiltonian by

$$a\hat{H} \equiv -\ln\overline{T} \ . \tag{69}$$

 \hat{H} has the properties that we want for a Hamiltonian: it is hermitean, and its smallest eigenvalue is 0

$$\hat{H}|0\rangle = E_0|0\rangle \qquad E_0 = 0 , \qquad (70)$$

$$\hat{H}|p\rangle = E_p|p\rangle \quad E_p = -\ln(\bar{\lambda}_p) > 0 \quad .$$

$$\tag{71}$$

In addition, it acts on a Hilbert space with positive norm. Thus we have succeeded in constructing the discretized version of the operator expression Eq. 5.

The actual form of H is horribly non-local, because the different exponents in Eq. 60 do not commute, so that taking the logarithm is not simple. In the classical continuum limit (i.e. when we assume that the operators are smoothly varying), however, \hat{H} indeed goes over into the continuum Hamiltonian

$$\hat{H} \to \int d^3x [\frac{1}{2}\hat{p}_{ph}^2 + \frac{1}{2}(\partial_i\hat{\phi}_{ph})^2 + \frac{1}{2}m_{ph}^2\hat{\phi}_{ph}^2 + \lambda\hat{\phi}_{ph}^4 + \dots] + \text{constant}$$
(72)

where $\hat{\phi} = a \hat{\phi}_{ph}$ and $\hat{p} = a^2 \hat{p}_{ph}$, and I have used the potential of Eq. 30.

An alternative continuum limit uses asymmetric lattices, in which we take the spacing in the time direction to zero, while keeping the spatial lattice spacing fixed. In this way one obtains a discretized Hamiltonian theory, which can be studied using various approximation schemes. This is discussed in the texts.

4.3 Finite Temperature

We can now see the significance of working on a lattice of finite time extent. The partition function is

$$Z = \operatorname{Tr} \, \hat{T}^N \propto \operatorname{Tr} e^{-Na\hat{H}} \,. \tag{73}$$

But this is nothing other than the partition function describing the theory in equilibrium at a finite temperature $\beta = 1/kT = Na$. The longer the lattice in physical units, the smaller the temperature. Inevitably, numerical simulations correspond to a non-zero temperature, although it can be made very small in practice. For further discussion see the lectures by Ukawa.

4.4 Symmetries of \hat{T} and \hat{H}

As with any quantum mechanical problem, it is useful to study the symmetries of the Hamiltonian. On the lattice, it makes more sense to phrase the discussion in terms of the symmetries of the transfer matrix (\hat{T} or \overline{T} , it makes no difference). Given that $a\hat{H} = -\ln \overline{T}$, \hat{H} and \hat{T} have the same symmetries anyway.

Given the definition of \hat{T} (Eqs. 57 and 52). its symmetries are those of the functions S_A and S_B . These are

- A: discrete translations,
- B: 3-d rotations in the cubic group,
- C: spatial inversions.

To each symmetry transformation g there is a corresponding unitary operator U(g) which represents the transformation, and which commutes with \hat{T}

$$\langle \alpha \rangle \to |\alpha \rangle_g = U(g) |\alpha \rangle, \quad \hat{T} \to U(g) \hat{T} U(g)^{-1} = \hat{T}.$$
 (74)

Thus if $|p\rangle$ is an eigenstate of \hat{T} with eigenvalue λ_p , then so is $U(g)|p\rangle$. Eigenstates can be labeled not only by their "energies" (here λ_p) but also by the representations of the symmetry groups in which they lie. The labels corresponding to the above symmetries are

- A: momentum \vec{k} , satisfying $\pi < k_i \leq \pi$,
- B: representations of the cubic group—the discrete analog of angular momentum (see problem 2 below),¶
- C: parity, $P = \pm 1$.

In a general theory there will also be internal symmetries, with corresponding representations. For example, if the field is complex there is a U(1) symmetry, and states are labeled by the corresponding charge.

What use is this classification of states? To see its utility, consider the expression for the two-point function in the limit of infinite time extent (cf. Eq. 67)

$$C_2(\tau_2 - \tau_1) = \langle 0 | \widehat{\mathcal{O}}^{\dagger} \overline{T}^{\tau_2 - \tau_1} \widehat{\mathcal{O}} | 0 \rangle .$$
(75)

Here I have considered a general operator, $\hat{\mathcal{O}}$, composed of the fields in the theory. The point is that we can construct this operator so that it belongs in a definite

[¶]I am being sloppy here. Rotations and translations do not commute. Only for $\vec{k} = 0$ can one label the states by representations of the full cubic group. For $\vec{k} \neq 0$, one must wheel out the technology for representing semi-direct products.

irreducible representation of the symmetry group, e.g. having definite momentum (e.g. $\hat{\mathcal{O}}(\vec{p}=0) = \sum_{\vec{n}} \hat{\phi}_{\vec{n}}$). Inserting a complete set of eigenstates of \hat{T} , one finds

$$C_2(\tau_2 - \tau_1) \propto \sum_{p=0}^{\infty} |\langle p | \hat{\mathcal{O}} | 0 \rangle|^2 \bar{\lambda}_p^{\tau_2 - \tau_1} .$$
(76)

Only states with the same symmetries as \mathcal{O} contribute to this sum, since the vacuum lies in the trivial representation of the symmetry group. For large $\tau_2 - \tau_1$ the "lightest" such state (that with the largest λ_p) dominates the sum. Thus it is straightforward, in principle, to calculate the energy of the lightest state in each representation of the symmetry group. This allows one, for example, to study the lattice dispersion relation by calculate energies, E, at various momenta, \vec{k} .

Thus a lattice practitioner must know something about discrete groups. The most useful is the cubic group, the discretization of SO(3), whose representations are the discrete version of angular momentum states. A nice discussion of this group, its representations, and their relation to those of SO(3) is given by Mandula, Zweig and Govaerts [7]. The following problem concerns this group.

4.5 Problem 2: the cubic (octahedral) group and its representations

This is the group of 3-d rotations (not including reflections), which you can think of as proper rotations of a cube. The problem is this: *construct the character table* of this group, and display examples of representations.

Here are some useful facts about discrete groups:

- Elements fall into conjugacy classes. Given two elements of a conjugacy class, a and b, one can always find a group element, g, such that $gag^{-1} = b$.
- The number of irreducible representations ("irreps") equals the number of conjugacy classes.
- For each element of each irrep there is a "character", $\chi(g)$, given by the trace of the matrix representing the element. Thus, for example, the character of the identity element is always the dimension of the irrep. It follows from the cyclicity of the trace that all elements in a conjugacy class have the same character.
- Character orthonormality:

$$\sum_{\text{classes } c} \eta_c \chi^{r_1}(c) \chi^{r_2}(c) = \delta^{r_1 r_2} N,$$

where

$$\eta_c$$
 = number of elements in class
 N = number of elements in group
 $\chi^r(c)$ = character of elements of class c in irrep r

• Given a reducible representation with character $\chi(c)$, the number of times an irreducible representation appears is

$$\frac{1}{N}\sum_{c}\eta_{c}\chi(c)\chi^{r}(c)$$

• Another useful result is:

$$\sum_{\text{irreps } r} \eta_c \chi^r(c) \chi^r(c') = \delta_{cc'} N$$

for each pair of classes c and c'. Picking c = c' = I, this gives $\sum_r d_r^2 = N$, where d_r is the dimension of the representation r.

- In the following I outline a possible approach to this problem.
- Step 1: Enumerate and classify elements of the group. Ref. [7] uses one method; here is an alternative suggestion.

Elements of the group can be generated by one 90° rotation R, e.g.

$$R_{xy}: [\hat{x} \to \hat{y}, \hat{y} \to -\hat{x}, \hat{z} \to \hat{z}]$$

and S, a rotation of the cube about a body diagonal by 120 degrees (say $\hat{x} \rightarrow \hat{y}, \hat{y} \rightarrow \hat{z}, \hat{z} \rightarrow \hat{x}$). Clearly, $R^4 = S^3 = I$. Show also that RSRS = I.

These relations define the group—all sequences of R and S can be simplified into a finite number of elements using them, e.g., $RSR = S^2$. To enumerate the elements it is easier, rather than using brute force, to collect them in conjugacy classes, since each element of a class carries out a similar type of transformation. What are the different types of elements? How many of each type are there ?

<u>Step 2</u>: One way to construct the character table is to invent representations, convince oneself that they are irreducible and calculate the characters. The character formulae given above act as a check.

A good choice of starting reps are those of the proper 3-d rotation group SO(3). These reps are labeled by J = 0, 1, 2, ... and have dimension 2J + 1. For example, the J = 0 rep is 1-dimensional and invariant under all transformations, so $\chi(c) = 1$, $\forall c$. This is an irrep of the cubic group too—the identity representation I (also called A_1).

Good luck!

5 Gauge Theories on the Lattice

In this lecture I explain how to discretize gauge theories, and in particular QCD. While there are interesting non-perturbative questions associated with the $SU(2)_L \times U(1)$ part of the standard model, e.g. the nature of its finite temperature phase transition, the most important practical application of lattice methods is to QCD. The low-energy phenomena of QCD, in particular confinement and chiral symmetry breaking, are non-perturbative, and the lattice is the only method available for studying them from first principles. A crucial aspect of lattice regularization is that it maintains gauge invariance, for this guarantees that the theory is unitary.

5.1 Continuum QCD, a brief overview

The continuum action is given by

$$S_E = -\sum_{q=u,d,s,c,b,\dots} \int_x \bar{q} (D + m_q) q + \frac{1}{2} \int_x \text{Tr}(F_{\mu\nu}F_{\mu\nu}) , \qquad (77)$$

where the integrals run over Euclidean space. The covariant derivative is

$$D_{\mu} = \partial_{\mu} - igA_{\mu} , \qquad (78)$$

in which the gauge fields are collected into a matrix $A_{\mu} = A^{a}_{\mu}T^{a}$, with T^{a} the generators of the SU(3) Lie Algebra

$$[T^a, T^b] = i f^{abc} T^c , \qquad \operatorname{tr}(T^a T^b) = \frac{1}{2} \delta_{ab} .$$
(79)

The quark fields are color triplets, with an implicit color index. Finally, the gauge field strength is

$$F_{\mu\nu} = F^{a}_{\mu\nu}T^{a} = \frac{i}{g}[D_{\mu}, D_{\nu}] = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} - ig[A_{\mu}, A_{\nu}] .$$
(80)

A local SU(3) gauge transformations is described by a space-time dependent element $V(x) \in SU(3)$ $(V^{-1} = V^{\dagger}, \det(V) = 1)$:

$$q(x) \to V(x)q(x) , \quad \bar{q}(x) \to \bar{q}(x)V^{-1}(x) , \qquad (81)$$

$$A_{\mu}(x) \to V(x)A_{\mu}(x)V^{-1}(x) + \frac{i}{g}V(x)\partial_{\mu}V^{-1}(x),$$
 (82)

$$F_{\mu\nu}(x) \to V(x)F_{\mu\nu}(x)V^{-1}(x),$$
 (83)

$$[D_{\mu}q](x) \to V(x)[D_{\mu}q](x)$$
 . (84)

Given the last two lines, it is simple to see that S_E is invariant.

It is useful to introduce the path-ordered integrals

$$L(x,y) = P \exp\{ig \int_{y}^{x} dz_{\mu} A_{\mu}(z)\},$$
(85)

which are to be thought of as going from y to x. The ordering is such that, for example, $A_{\mu}(x)$ is always to the left of $A_{\mu}(y)$. The reverse ordering is obtained by hermitian conjugation, i.e. $L(y,x) = L(x,y)^{\dagger}$. The L's can be built from a product of infinitessimal steps, $(1 + igdz_{\mu}A_{\mu}(x'))$, along the path. Using the gauge transformations listed above, it is easy to see that these quantities transform as

$$(1 + igdz_{\mu}A_{\mu}(x')) \to V(x' + dz_{\mu})(1 + igdz_{\mu}A_{\mu}(x'))V^{-1}(x') + O(dz_{\mu}^{2}).$$
(86)

It follows that the gauge transformation properties depend only on the end points of L, and not on the path of integration

$$L(x,y) \longrightarrow V(x)L(x,y)V^{-1}(y) .$$
(87)

They thus transport the gauge rotation from one point to another, such that the quantity $\overline{q}(x)L(x,y)q(y)$ is gauge invariant:

$$\overline{q}(x)L(x,y)q(y) \rightarrow \overline{q}(x)V^{-1}(x)V(x)L(x,y)V^{-1}(y)V(y)q(y) = \overline{q}(x)L(x,y)q(y).$$
(88)

Another gauge invariant quantity is the trace of the path-ordered integral around any closed path

$$Tr[L(x,x)] \longrightarrow Tr[V(x)L(x,x)V^{-1}(x)] = Tr[L(x,x)]$$
(89)

These objects are called Wilson loops.

5.2 Discretization

With these quantities in hand, we can now construct a gauge invariant lattice version of QCD. Well, not quite. It turns out that discretizing fermions presents problems unrelated to gauge invariance, problems which I discuss in the last lecture, In this lecture I avoid these problems by replacing the quarks with **scalar** colour triplets, ϕ . These are necessarily complex, and their continuum action is

$$S_{cont} = \int_x -\phi^{\dagger} D_{\mu} D_{\mu} \phi + P(\phi^{\dagger} \phi) , \qquad P(y) = m^2 y + \lambda y^2 + \dots$$
 (90)

Under gauge transformations they behave like quarks

$$\phi^{\dagger}(x) \to \phi^{\dagger}(x)V^{-1}(x), \quad \phi(x) \to V(x)\phi(x).$$
 (91)

To construct a lattice theory, one cannot simply place quark and gauge fields on the sites of the lattice and discretize the derivatives appearing in S_E . Instead, the gauge fields, which transmit information about gauge transformations from one position to another, live on the "links" or "bonds" connecting the sites. I will choose the lattice to be hypercubical, since this is the form most easily studied numerically.



Figure 1: Notation for lattice quantities. n is a vector of integers.

The notation for the sites and links on the lattice is shown in Fig. 1. Note that I am using $n + \mu$ instead of $n + \hat{\mu}$, to denote the point one lattice spacing forward from n in the μ -direction.

We begin with the free action for the scalars, which is discretized by a simple generalization of the result for a real scalar field (Eqs. 28 and 29)

$$S_{lat}^{0} = \sum_{n} \sum_{\mu} [\phi_{n+\mu}^{\dagger} - \phi_{n}^{\dagger}] [\phi_{n+\mu} - \phi_{n}] + \sum_{n} P(\phi_{n}^{\dagger}\phi_{n}).$$
(92)

Next we introduce the gauge fields. The discrete version of the gauge transformation function V(x) is V_n , a site dependent SU(3) matrix. Under a gauge transform

$$\phi_n^{\dagger} \to \phi_n^{\dagger} V_n^{-1}, \quad \phi_n \to V_n \phi_n, \quad \phi_n^{\dagger} \phi_n \to \phi_n^{\dagger} \phi_n.$$
 (93)

Thus the potential term in Eq. 92 is invariant, while the kinetic term is not.

To fix this up, we introduce elements of SU(3), $U_{n,\mu}$, associated with the link from site n to site $n+\mu$, and corresponding to the continuum line integral backwards along the link:

$$U_{n,\mu} \sim L(an, an + a\hat{\mu})$$

= $P \exp\left[ig \int_{an+a\hat{\mu}}^{an} dz_{\mu} A_{\mu}(z)\right]$ (94)

$$= 1 - igaA_{\mu}(n + \frac{1}{2}\hat{\mu}) + O(a^2) .$$
(95)

Figure 1 shows examples of the diagrammatic representation of these link matrices as arrows. We associate $U_{n,\mu}^{\dagger}$ with the link from $n + \mu$ to n, in correspondence with the continuum result $L(x, y) = L(y, x)^{\dagger}$. The " \sim " in the first line of Eq. 94 means "corresponds to". The vagueness here is deliberate—once we put the theory on the lattice there are no gauge fields A_{μ} : they are replaced by the U's. The expansion in the last line (which assumes the straight line path) is useful, however, for thinking about what the U's mean, and also for taking the classical continuum limit.

The gauge transformation properties of the U's are taken to be the same as those of the corresponding L's

$$U_{n,\mu} \to V_n U_{n,\mu} V_{n+\mu}^{\dagger}, \quad U_{n,\mu}^{\dagger} \to V_{n+\mu} U_{n,\mu}^{\dagger} V_n^{\dagger}.$$
(96)

Thus $U_{n,\mu}\phi_{n+\mu}$ transforms in the same way as ϕ_n ,

$$U_{n,\mu}\phi_{n+\mu} \to V_n U_{n,\mu} V_{n\mu}^{\dagger} V_{n+\mu} \phi_{n+\mu} = V_n U_{n,\mu} \phi_{n+\mu} , \qquad (97)$$

and $\phi_{n+\mu}^{\dagger} U_{n,\mu}^{\dagger}$ transforms as does ϕ_n^{\dagger} . With these relations we can construct a gauge invariant lattice action

$$S_{lat} = \sum_{n} \sum_{\mu} [\phi_{n+\mu}^{\dagger} U_{n,\mu}^{\dagger} - \phi_{n}^{\dagger}] [U_{n,\mu} \phi_{n+\mu} - \phi_{n}] + \sum_{n} P(\phi_{n}^{\dagger} \phi_{n}) .$$
(98)

It is instructive to check that the kinetic part of this action has the correct continuum limit. Using the result of Eq. 95 we find

$$U_{n,\mu}\phi_{n+\mu} - \phi_n \sim a^2 [\partial_\mu \phi] (an + \frac{1}{2}a\mu) - iga^2 A_\mu (an + \frac{1}{2}a\mu)\phi(an + a\mu) = a^2 [D_\mu \phi] (an + \frac{1}{2}a\mu) + 0(a^3)$$
(99)

Combining this with its hermitian conjugate, we regain the kinetic term in continuum action, $\int |D_{\mu}\phi|^2$.

5.3 Pure Gauge Action

We can construct a lattice version of the pure gauge action using the smallest Wilson loop, that around an elementary square or "plaquette"

$$P_{\mu\nu}^{\dagger} = U_{n,\mu} U_{n+\mu,\nu} U_{n+\nu,\mu}^{\dagger} U_{n,\nu}^{\dagger} .$$
 (100)

The geometry is illustrated here.



It is reasonable that such a loop is related to $F_{\mu\nu}$, because the field strength is the curvature associated with the connection A_{μ} . In any case, using the correspondence given above for the U's, and after some algebra, one finds that the classical continuum limit of the plaquette is

$$P_{\mu\nu} = 1 + iga^2 F_{\mu\nu} - \frac{g^2}{2}a^4 F_{\mu\nu}^2 + ia^3 G_{\mu\nu} + ia^4 H_{\mu\nu} + 0(a^5), \qquad (101)$$

where $H_{\mu\nu}$ and $G_{\mu\nu}$ and are hermitian (see Problem 3). Thus one can use the $\mu\nu$ plaquette as a discretized version of the corresponding component of the field strength, $F_{\mu\nu}$. If we take the trace, so as to get a gauge invariant quantity, we find

Re Tr
$$P_{\mu\nu} = N_c - \frac{g^2}{2} a^4 \operatorname{Tr}(F_{\mu\nu}^2) + 0(a^6)$$
, (102)

where $N_c = 3$ is the number of colors. We then have

$$\int d^4x \sum_{\mu\nu} \frac{1}{2} \operatorname{Tr} F_{\mu\nu} F_{\mu\nu} \sim \sum_{\Box} \frac{2}{g^2} (N_c - \operatorname{ReTr} \Box) .$$
(103)

The factor of 2 arises because of the mismatch between the number of plaquettes per site, 6, and the number of terms in the sum $\sum_{\mu\nu}$, 12.

Standard notation in the field is to replace the coupling constant with $\beta = \frac{2N_c}{g^2}$, so that

$$S_g = -\beta \sum_{\Box} \frac{\text{ReTr}\Box}{N_c} + \text{ irrelevant constant }.$$
(104)

This is called the "Wilson (gauge) action". Using the symbol β is unfortunate, as it leads to possible confusion with 1/kT, but it is a notation that is well and truely entrenched.

It is important to realize that there is nothing special about using the smallest loop to define the action. Any loop, e.g. a 1×2 rectangle, contains a term in its expansion proportional to a linear combination of components of $(F_{\mu\nu})^2$. By taking an appropriate combination of loops we can obtain the continuum action as $a \to 0$. The advantage of a small loop is that corrections proportional to powers of the lattice spacing are typically smaller than with a larger loop.

To complete the definition of the theory I need to specify the measure. Each link variable is integrated with the Haar measure over the group manifold. This measure satisfies (V and W are arbitrary group elements)

$$\int dUF(U) = \int dUF(UV) = \int dUF(WU).$$
(105)

Given this, it is simple to see that the functional integral

$$Z_{\text{gauge}} = \int \prod_{\text{links}} dU_{n,\mu} \prod_{n} (d\phi_n d\phi_n^{\dagger}) \exp\left(\beta \sum_{\Box} \frac{1}{N_c} \text{Re } \operatorname{Tr} \Box - S_{lat}(\phi^{\dagger}, \phi, U)\right)$$
(106)

is gauge invariant. Note that the U matrices live on a group manifold with finite volume, in contrast to the infinite range of A_{μ} in the continuum.

What has been accomplished here is a non-perturbative, gauge invariant regularization of gauge theories coupled to scalars. What has been sacrificed is full Euclidean invariance: rotations and translations. The hope is that, as one approaches the continuum limit, these symmetries are restored.

5.4 Problem 3: the expansion of the plaquette.

Calculate the expansion of the plaquette $P_{\mu\nu}$, and its trace, in powers of a.

Step 1: The calculation is simplified if one defines the continuum field by

$$U_{n,\nu} = \exp\left[-iagA_{\nu}(an+\frac{1}{2}a\nu)
ight], \quad \text{etc.}$$

and expands the continuum fields above the center of the plaquette. Writing $P_{\mu\nu} = \exp(X)$ calculate X explicitly through $O(a^2)$. What are the properties of the higher order terms?

- <u>Step 2</u>: Consider Re Tr $P_{\mu\nu}$. Expand up to quartic terms in a. Show that $0(a^3, a^4)$ terms in X do not contribute at this order.
- <u>Step 3</u>: Show that $\operatorname{Tr}(P_{\mu\nu})$ has an expansion in even powers of a, so that the corrections to the a^4 terms from step 2 are of $O(a^6)$.

6 Applications of Lattice Gauge Theories

Before proceeding to study fermions, and thus to a lattice version of QCD including quarks, I will discuss what one can learn from pure SU(3) gauge theory. This is a non-trivial theory in its own right, sharing some properties in common with QCD. In particular, its spectrum consists of massive glueballs, in which the gluons are confined by their self-interactions. First, though, let me discuss one of the classic results of lattice gauge theory ...

6.1 Confinement

Long ago, Wilson introduced a test for confinement in gauge theories. Imagine that there is a very heavy quark, and a corresponding antiquark, separated by a distance R. Calculate the energy V(R) of the pair as a function of R. If there is confinement, then V(R) grows monotonically with R for large R. The picture is that there is (color) electric flux between the quark and antiquark, which does not spread out like in QED, but is forced into a "string":



This picture implies linear growth $V(R) \propto \sigma R$ for large R, where σ is the "string tension".

The heaviness of the quark plays three roles. First, it means that the quark does not move—it has infinite inertia—so it makes sense to consider a static potential. Second, $\bar{q}q$ pair creation is suppressed (by terms of $O(1/m^2)$), so that one need not consider loops of the heavy quark. Such quark loops would, in fact, cause the string to break, and thus remove the possibility for a clean test of confinement. And, third, at leading order in 1/m, the coupling of quark to gluons does not involve its spin. The quark acts as a static scalar color source. This means that one can just as well use scalar quarks to extract V(R).

To calculate the potential consider the correlator (written in continuum language for the moment)



The expectation value indicates the functional integral over gauge and scalar fields, and the thick lines with arrows represent the line integrals L. Thus this is a gauge invariant correlator. I have introduced *two* heavy scalar fields ϕ and ψ , for reasons which will be clear shortly. And, finally, I have used t instead of τ to represent Euclidean time—as we will remain in Euclidean space henceforth, the distinction is no longer necessary.

This correlator is just a more complicated version of the two-point function considered earlier, Eq. 8; instead of a scalar field there is a quark-antiquark operator, joined by a line of gauge fields. The analysis of the two-point function remains the same. For large t, it is dominated by the lightest state created from the vacuum by the quark-antiquark operator. The sum over \vec{x} projects onto states at rest, although as we will see, this sum is unnecessary as the only $\vec{x} = 0$ contributes when $m \to \infty$. If the lightest state has energy E, then $C(t) \propto \exp(-Et)$. But E is nothing other than V(R). Thus, for large enough R, we expect

$$C(t) \xrightarrow{t \to \infty} e^{-tV(R)} \xrightarrow{R \to \infty} e^{-tR\sigma} = e^{-[(\operatorname{area}) \times \sigma]}, \qquad (107)$$

where the area is that of the rectangle formed by the fields when $\vec{x} = 0$. This is (almost) the famous area law criterion for confinement.

The most straightforward way to proceed is to do the functional integral over the scalar fields in the continuum and then discretize the result. Let me sketch the argument. Imagine that the correlator is in Minkowski space. In the limit $m \to \infty$, one can use the heavy quark effective theory discussed in the lectures of Isgur. Quarks, whether scalars or fermions, just maintain their velocities. Now the field ϕ^{\dagger} creates quarks with all velocities, but what we are interested in, to obtain V(R), is those with $\vec{v} = 0$. So what we really mean by ϕ is $\phi_{\vec{v}=0}$. Given that the quarks are static, only the $\vec{x} = 0$ term in the sum contributes. The propagator for a static quark at site \vec{R} in a background gauge field is $\exp(-imt)P\exp[ig\int_0^t dt'A_0(\vec{R},t')]$, i.e. the line integral $L(\vec{x},t;\vec{x},0)$ (along a straight path) up to a kinematical factor which is independent of R and thus irrelevant to the determination of σ . The antiquark propagator is proportional to the line integral in the opposite direction L(0,0;0,t). Putting these together with the line integrals in the operators, one obtains a rectangular Wilson loop. Wick rotation yields a Euclidean Wilson loop. This is then simple to discretize in terms of the U matrices.

I have chosen to obtain this result by a more circuitous route, since doing so allows me to introduces a useful technique, the hopping parameter expansion. We discretize the correlator *before* integrating over the scalars. Assuming that \vec{R} is a lattice vector lying along the positive j-th direction, we have

$$C(t) = \sum_{\vec{n}} \frac{1}{Z} \int [dU] e^{-S_G} \int [d\phi] [d\phi^{\dagger}] [d\psi] [d\psi^{\dagger}] e^{-S_{\phi} - S_{\psi}} \times (\phi^{\dagger}_{\vec{n},t} U_{(\vec{n},t)j} \dots \phi_{\vec{n}+\vec{R},t}) (\psi^{\dagger}_{\vec{R},0} \dots U^{\dagger}_{0,j} \phi_0) .$$
(108)

It is useful to rescale the scalar fields

$$\phi \to m\phi, \quad \psi \to m\psi.$$
 (109)

for then the scalar action becomes

$$S_{\phi} = \sum_{np} \phi_n^{\dagger} K_{np} \phi_p + \frac{\lambda}{m^4} (\phi'^{\dagger} \phi)^2$$
(110)

where

$$K_{np} = \delta_{np} + \frac{1}{m^2} \sum_{\mu} [2\delta_{np} - U_{n,\mu}\delta_{n+\mu,p} - U_{n,\mu}^{\dagger}\delta_{n,p+\mu}].$$
(111)

We see that the interaction term is suppressed by $1/m^4$; it can be ignored as $m \to \infty$. It is useful to rewrite the kernel as

$$m^{2}K_{np} = \delta_{np}(m^{2} + 8)(1 - \frac{H_{np}}{m^{2} + 8}), \qquad (112)$$

where H_{np} is called the "hopping matrix" as it hops fields from one site to another

$$H_{np} = \sum_{\mu} U_{n,\mu} \delta_{n+\mu,p} + U_{n,\mu}^{\dagger} \delta_{n,p+\mu} \,.$$
(113)

Now that the scalars are free, we can integrate them out using

$$\int [d\phi] [d\phi^{\dagger}] [d\psi] [d\psi^{\dagger}] e^{-S_{\phi} - S_{\psi}} \psi^{\dagger}_{n_{1}} \phi_{n_{2}} \psi_{n_{3}} \phi^{\dagger}_{n_{4}} = G^{\psi}_{n_{3}n_{1}} G^{\phi}_{n_{4}n_{2}} \int [d\phi] [d\phi^{\dagger}] [d\psi] [d\psi^{\dagger}] e^{-S_{\phi} - S_{\psi}}$$
(114)

where G denotes the propagator. The reason for having two fields ϕ and ψ is now apparent: there is only one Wick contraction. The propagator can be obtained by a "hopping expansion"

$$G = \frac{K^{-1}}{m^2} = \frac{1}{m^2 + 8} \left(1 + \frac{H}{m^2 + 8} + \frac{H^2}{(m^2 + 8)^2} + \dots \right).$$
(115)

It turns out that similar expansions apply for quark propagators, as we will see below. Each "hop" costs a factor of $\frac{1}{m^2}$ (the "8" is irrelevant when $m \to \infty$) and gives a matrix U or U^{\dagger} . The minimum cost is achieved by hopping along the shortest path. This picks out $\vec{n} = 0$ in the sum in Eq. 108, and gives rise to a discretized Wilson loop



up to corrections suppressed by powers of $1/m^2$. The subscript U indicates that only the functional integral over gauge fields remains.

Recall that if there is confinement, we expect $C(t) \propto \exp(-Rt\sigma)$ for large t and R. The only way this can happen is if the expectation value of the $R \times t$ Wilson loop falls off in this way. This is Wilson's area law test for confinement.

In fact, one can show that lattice gauge theories do confine in the strong coupling limit, $\beta \to 0$ or equivalently $g^2 \to \infty$. One can calculate expectation values of Wilson loops analytically, in an expansion in powers of β . One finds (problem 4)

$$\sigma = -\ln\left(\frac{\beta}{2N_c^2}\right) + 0(\beta).$$
(116)

This cannot be used as a direct indication of what happens to the continuum theory, because, as I discuss below, the continuum limit occurs for weak coupling, $g^2 \rightarrow 0$. The strong coupling expansion has a finite radius of convergence, and present attempts to extrapolate using tricks such as Padé approximants break down at couplings $g^2 \ll 1$. Thus the only way to test for confinement near the continuum



Figure 2: The heavy quark potential, in lattice units. The horizontal scale is in units of lattice spacing, which is $a \approx 0.06$ fm. The short distance points have been corrected for lattice artifacts using the lattice Coulomb propagator.

limit is to calculate the expectation value of Wilson loops numerically. I give an example of the results for the potential, from Ref. [8], in Fig. 2. The linear rise in V is clear, starting at about 0.5 fm.

The behavior of non-abelian (and thus asymptotically free) gauge theories is quite different from that of abelian theories, the simplest example of which has the gauge group U(1). In strong coupling, the result of Eq. 116 applies also for U(1), with $N_c = 1$, Thus the theory confines. By contrast, the continuum U(1) theory involves non-interacting photons, and does not confine. How are these two limits reconciled? It turns out that, unlike for non-abelian theories, there is a (first order) phase transition at finite g. Indeed, for a certain choice of the action this can be proved to exist. At the transition, one goes from a confining to a non-confining phase. The issue with non-abelian groups is whether there is a similar transition, and the evidence strongly suggests that there is not.

6.2 Problem 4—strong coupling expansion

Calculate the string tension σ in pure gauge SU(3) theory in the strong coupling limit $(\beta \to 0)$, i.e., evaluate

$$\langle W(R,T)\rangle = \frac{\int \prod_{n,\mu} dU_{n,\mu} \exp\left[\frac{\beta}{N_c} \sum_{\Box} \mathsf{ReTr}\Box\right] W(R,T)}{\int \prod_{n,\mu} dU_{n,\mu} \exp\left[\frac{\beta}{N_c} \sum_{\Box} \mathsf{ReTr}\Box\right]}$$
(117)

where W(R,T) is an $R \times T$ Wilson loop, and use

$$\ln \langle W(R,T) \rangle \xrightarrow{R,T \to \infty} -\sigma R T + c_1 R + c_2 T + c_3 + \dots$$
(118)

To do this you will need the integrals (valid for $SU(N_c \ge 3)$)

- $\int dU(1) = 1$
- $\int dU(U_{ij}) = 0$
- $\int dU(U_{ij}U_{kl}) = 0 = \int dU(U_{ij}^{\dagger}U_{kl}^{\dagger})$
- $\int dU(U_{ij}U_{kl}^{\dagger}) = \frac{1}{N_c}\delta_{jk}\delta_{il}$

The lattice result is for the dimensionless string tension

$$\sigma(g^2) = a^2 \sigma_{\rm phys} \,.$$

This equation has corrections of higher order in a^2 , which come from discretization errors, and which are not small at strong coupling. Nevertheless, for purposes of illustration, we will ignore them. Then, as we vary g^2 we can adjust a such that σ_{phys} is constant, which is what we want in the continuum limit (see below). Using this method, calculate the β -function, i.e. $\frac{dg^2}{d\ln a}$.

6.3 Glueball masses

The spectrum of pure gauge theory has been studied extensively, for gauge groups SU(2) and SU(3), using numerical methods. The low lying spectrum close to the continuum limit has been established with small errors, as shown in Fig. 3 (Ref. [9]). One can only extract ratios of dimensionful quantities; here, the masses are given (the left hand scale) in units of the square root of the string tension. The J^{PC} of the corresponding continuum states is noted on the plots.

The spectrum is obtained from two point functions such as

$$C(t) = \frac{1}{Z} \int \prod_{n,\mu} (dU_{n,\mu}) e^{-S_{gauge}} \sum_{\vec{n}} \operatorname{Tr} P_{kl}(t,\vec{n}) \operatorname{Tr} P_{ij}(0),$$
(119)

where P is a plaquette, and $1 \le i, j, k, l \le 3$. Proceeding as above yields

$$C(t) = \sum_{g} \frac{1}{2m_g V} \langle 0|\hat{P}_{kl}|g, \vec{k} = 0 \rangle \langle g, \vec{k} = 0|\hat{P}_{ij}|0\rangle e^{-m_g|t|} .$$
(120)

I do not have time to construct the transfer matrix for gauge theories. The method is similar to that for scalars, yielding operators $\hat{U}_{n,i}$ (and corresponding momenta) which act on the Hilbert space of square integrable functions on the group. The operators \hat{P}_{ij} are made up of operators $\hat{U}_{n,i}$ just as in Eq. 100, All that matters here



Figure 3: Spectrum of pure gauge SU(3) theory at $\beta = 6.4$. Open symbols represent upper limits.

is that the symmetries of the \hat{P} 's are the same as those of the P's. By combining plaquettes appropriately one can project onto different representations of the cubic group. The simplest examples are

scalar glueball
$$\sim P_{12} + P_{21} + P_{13} + P_{31} + P_{23} + P_{32}$$
, (121)

transforming in the identity, or A_1 , representation, and

tensor glueball ~
$$\operatorname{\mathsf{Re}}(P_{12} - P_{13})$$
, (122)

transforming as part of the two dimensional E irrep. For more examples see Ref. [7].

Why have I called these "scalar" (i.e. $J^P = 0^+$) and "tensor" (2⁺) glueball operators? The cubic group is a subgroup of the continuum rotation group, SO(3). The representations of SO(3) (J = 0, 1, 2, ...) form, in general, reducible representations of its cubic subgroup. For example, the J = 2 representation breaks up into $E + T_2$, where T_2 is a three dimensional irrep of the cubic group.^{||} Conversely, each

 $^{\|}$ There is a parity symmetry on the lattice just as in the continuum, so all states come with an additional parity label. The scalar and tensor glueball operators have positive lattice parity, and thus must correspond to positive parity continuum glueballs.

cubic group irrep receives contributions from an infinite tower of SO(3) irreps: e.g. $A_1 \sim J = 0, 4, 6, \ldots, E \sim J = 2, 4, 5, 6, \ldots$, and $T_2 \sim J = 2, 3, 4, 5, 6, \ldots$. Thus, if we measure a two point correlator of the "scalar" operator of Eq. 121, and extract the mass of the lightest particle, we do not know a priori whether it corresponds in the continuum to a particle with spin-0 or with spin-4, etc. We must assume that the lightest particle has the lowest spin. We can test this assumption by comparing the results in different channels. For example, if the lightest states in the E and T_2 irreps are different polarizations of the same tensor glueball in the continuum limit, their masses must converge as $a \to 0$. This indeed seems to happen.

<u>Exercise</u>: Show that in the strong coupling limit the scalar glueball mass is

$$m_g = -4 \ln(\frac{\beta}{2N_c^2})$$
 (123)

Use this to extract a β -function. How does this compare to that obtained from the string tension?

6.4 Perturbation Theory and the Continuum Limit

It is important to understand the stucture of lattice perturbation theory (i.e. the expansion about $g^2 = 0$), for several reasons. First, it tells one how to take the continuum limit. Second, it is a crucial part of some phenomenological applications of lattice QCD, those in which one calculates the matrix elements of external operators (e.g. those of the electroweak effective Hamiltonian) between hadronic states. Third, it allows one to relate lattice and continuum coupling constants, and in this way use non-perturbative calculations of the spectrum to predict the value of $\alpha_S(M_Z)$ measured in high energy perturbative processes such as jet cross sections.

Recall the form of the pure gauge lattice action

$$S_{gauge} = -\beta \sum_{\Box} \frac{\operatorname{\mathsf{ReTr}} P_{\Box}}{N_c} = -\frac{2}{g^2} \sum_{\Box} \operatorname{\mathsf{ReTr}} P_{\Box} .$$
(124)

We will expand in powers of $g^2 = 0$. In this limit the plaquettes all tend to the unit matrix, $P \to I$, so as to maximize the trace and thus minimize the action. This is strictly true only in finite volume—a point I return to below. But $P \to I$ does not imply that $U_{n,\mu} = I$, only that the U's must be gauge equivalent to the identity configuration. For example, if we start with $U_{n,\mu} = I$ everywhere, and do a local gauge transformation, we will find $U_{n,\mu} = V_n V_{n,n+\mu}^{-1}$, which can be as far as we want from the identity. The plaquettes, which are gauge invariant, of course remain at P = I. We would like to expand the U's about I, but to do this we must fix the gauge, just as in the continuum. This is in contrast to numerical simulations (and the strong coupling expansion) which do not require gauge-fixing.

One choice of gauge-fixing condition is to maximize

$$\sum_{n,\mu} \operatorname{ReTr}(U_{n,\mu}) . \tag{125}$$

This brings the U's as close to the identity as possible on average. It is the lattice Landau gauge. In an actual simulation at finite g^2 , there are many maxima of this function—which is nothing other than the Gribov ambiguity on the lattice. This makes it problematic to use this condition in a non-perturbative simulation a subject receiving considerable attention at present [10]. As in the continuum, however, there is no Gribov ambiguity in perturbation theory.

Assuming we have fixed the gauge appropriately, we can expand the links about the identity

$$U_{n,\mu} = e^{-igA_{\mu}(n)} = 1 - igA_{\mu}(n) - \frac{1}{2}g^2A_{\mu}^2(n) + \dots, \qquad (126)$$

where the first equality defines A_{μ} . For notational simplicity, I have set a = 1 and put the field A_{μ} at the site n and not at $n + \frac{\mu}{2}$. Using this expansion, we find (essentially a repeat of the derivation of the lattice action)

$$S_{gauge} = \sum_{n,\mu} \frac{1}{2} \operatorname{Tr}[(\Delta^{+}_{\mu} A_{\nu}(n) - \Delta^{+}_{\nu} A_{\mu}(n))(\Delta^{+}_{\mu} A_{\nu}(n) - \Delta^{+}_{\nu} A_{\mu}(n))] + 0(A^{3})$$

= $S_{A^{2}} + 0(A^{3}),$ (127)

where

$$\Delta^{+}_{\mu}A_{\nu}(n) = A_{\nu}(n+\mu) - A_{\nu}(n)$$
(128)

Using "summation by parts"

$$\sum_{n} g(n) \Delta_{\mu}^{+} h(n) = -\sum_{n} (\Delta_{\mu}^{-} g(n)) h(n) , \qquad (129)$$

where

$$\Delta_{\mu}^{-}g(n) = g(n) - g(n-\mu), \quad \Delta_{\mu}^{+}\Delta_{\mu}^{-}g(n) = \Delta_{\mu}^{-}\Delta_{\mu}^{+}g(n), \quad (130)$$

we find

$$S_{A^2} = \sum_{n} \operatorname{Tr}\left[\underbrace{(-\Delta_{\nu}^{-}A_{\nu})(\Delta_{\mu}^{-}A_{\mu})}_{\text{remove by gauge-fixing}} - A_{\nu} \underbrace{\Delta_{\mu}^{-}\Delta_{\mu}^{+}}_{\Delta_{\mu}^{-}\Delta_{\mu}^{+}} A_{\nu}\right].$$
(131)

To remove the unwanted term (which, as in the continuum, makes the quadratic term in the action non-invertible), we follow the lattice version of the Fadeev-Popov procedure. We assume a gauge condition of the form

$$f_n(U) = \alpha_n, \quad \forall n \,, \tag{132}$$

where f_n is a function only of the link matrices adjacent to the site n, i.e. those U's which are rotated by the gauge transformation matrix V_n . A useful quantity is the Jacobian

$$J_n(U) = \left| \det\left[\frac{\partial f_n}{\partial V_n}\right](U) \right|.$$
(133)

We consider the functional

$$I(V,\alpha) = \int \prod_{n,\mu} dU_{n,\mu} e^{-S_{gauge}(U)} \prod_{n} \left[\delta(f_n(U^{V_n}) - \alpha_n) J_n(U^{V_n}) \right]], \qquad (134)$$

where U^V means U after gauge rotation by V. Despite appearances, $I(V,\alpha)$ does not depend upon V. This is because $dU = dU^V$ and $S_{gauge}(U^V) = S_{gauge}(U)$, so we can change variables to $U' = U^V$ and remove all reference to V. Furthermore, the functional is constructed to satisfy

$$\int \prod_{n} DV_n I(V,\alpha) = \int \prod_{n,\mu} dU_{n,\mu} e^{-S_{gauge}} = Z.$$
(135)

Since we use a normalized group integration measure $(\int dU = 1)$, it follows that

$$Z = I(V, \alpha_n) \text{ for any } V, \alpha_n$$

$$\propto \prod_n \int d\alpha_n e^{-\mathsf{Tr}(\alpha_n^2)} I(1, \alpha_n)$$

$$= \int \prod_{n,\mu} dU_{n,\mu} e^{-S_{gauge}} \underbrace{e^{-\sum_n \mathsf{Tr}[f_n(U)^2]}}_{\text{gauge-fixing}} \underbrace{\prod_n J_n(U)}_n I(136)$$

The product of jacobians can be written in more familiar form as

$$\prod_{n} J_{n}(U) = \det \left[\frac{\partial f}{\partial V} \right] .$$
(137)

Now we return to the choice of gauge condition. It is convenient to write it in terms of A's. The simplest and most commonly used choice is

$$f_n = \sum_{\mu} \Delta_{\mu}^- A_{\mu}(n) \,, \tag{138}$$

the discrete version of $\partial_{\mu}A_{\mu}$. As advertized above, this does involve only the links emanating from the site *n*. The resulting gauge-fixing term cancels the unwanted part of S_{A^2}

$$S_{A^{2}} + S_{g.f.} = -\sum_{n\mu\nu} \operatorname{Tr}[A_{\nu}\Delta_{\mu}^{-}\Delta_{\mu}^{+}A_{\nu}]$$
(139)

This is the lattice version of Feynman gauge.

Exercise: Show that the lattice gluon propagator in Feynman gauge is

$$\frac{\delta_{\mu\nu}\delta_{ab}}{4\sum_{\mu}\sin^2\left(\frac{k_{\mu}}{2}\right)}\,,\tag{140}$$

where a, b = 1, 8 are color indices. The denominator is the lattice version of k^2 .

If instead we take $f_n = \gamma \sum_{\mu} \Delta_{\mu}^- A_{\mu}(n)$, and send $\gamma \to \infty$, we obtain lattice Landau gauge. If you think about the Fadeev-Popov procedure, you will see that this limit amounts to setting $\sum_{\mu} \Delta_{\mu}^- A_{\mu}(n) = 0$. This is nothing other than the differential form of the condition Eq. 125, expressed in terms of A's, where we keep only the O(a) term. To complete the construction of lattice perturbation theory, we need the expansion of the measure

$$dU_{n,\mu} \propto \prod_{\alpha=1,8} dA^{\alpha}_{\mu} \times (1 + \text{higher order terms, not needed here}).$$
 (141)

Thus calculations proceed just as in the continuum, except that (i) integrals are automatically regularized in the ultraviolet because (reinserting the a)

$$\int \frac{d^4k}{(2\pi)^4} \to \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \frac{d^4k}{(2\pi)^4} \,, \tag{142}$$

and (ii) the form of the propagators and vertices are altered. In particular, there is an infinite sequence of vertices coming from the expansion of S_{gauge} in powers of A



and a similar infinite tower of vertices involving ghosts. We could do all the familiar calculations of perturbative QCD using lattice regularized perturbation theory instead of, say, dimensional regularization. Power counting for UV divergences works as usual [11]. It would just be very messy.

Let me consider, schematically, the calculation of the β -function. To do this we calculate $g_R(p)$, the renormalized three-point function with a definite choice of external momenta having magnitude $\sim p$. At one-loop the graphs are



where the last diagram (the "tadpole diagram") is specific to the lattice. The diagrams are finite—ultra-violet divergences are cut-off by the lattice spacing, infrared divergences by the external momenta. The result has the form

$$g_R(p) = g\left\{1 + g^2[-\beta_0 \ln(ap) + C^L + 0(a^2 p^2 \ln(ap))] + 0(g^4)\right\}.$$
 (143)

This should be reliable as long as $\pi/a \gg p \gg \Lambda_{\text{QCD}}$. The upper limit must be satisfied so that cut-off effects (such as that proportional to $(ap)^2$ in Eq. 143) are small. The lower limit ensures that g(p) is small enough that perturbation theory can be used.

The form of Eq. 143 is the same for all regulators— in particular, the constant β_0 is the universal first coefficient of the β -function

$$\beta_0 = \frac{1}{3(16\pi^2)} (11N_c - 2N_f) \,. \tag{144}$$

It governs the way in which $g_R(p)$ decreases as p increases

$$\frac{dg_R(p)}{d\ln p}\Big|_{a,g} = -\beta_0 g^3 [1 + 0(g^2) + 0(a^2)].$$
(145)

What differs between regulators is the value of the constant. In particular, the tadpole diagram, which is specific to the lattice, contributes only to C_L and not to β_0 . This is because the tadpole loop is quadratically and not logarithmically divergent: $a^2 \int \frac{d^4k}{k^2} \propto \pi^2$, where the factor of a^2 comes from the vertex.

6.5 Continuum Limit of Pure Gauge Theories

We can use the perturbative result Eq. 143 to understand the "quantum" continuum limit. We want to take this limit in such a way that physical quantities, evaluated at definite physical momenta, remain fixed. In perturbation theory, $g_R(p)$ is such a "physical" quantity. To keep it fixed, we must vary the bare lattice coupling g with lattice spacing:

$$\frac{dg_R(p)}{d\ln a} = 0 \quad \Rightarrow \quad \frac{dg}{d\ln a} = \beta_0 g^3 [1 + 0(g^2) + 0(a^2)] \\
\Rightarrow \quad \frac{d(\frac{1}{g^2})}{d\ln a} = -2\beta_0 [1 + 0(g^2) + 0(a^2)] \\
\Rightarrow \quad \frac{1}{g^2} = -2\beta_0 \ln(a\Lambda_{lat}) [1 + 0(g^2) + 0(a^2)], \quad (146)$$

where Λ_{lat} is the integration constant. If g^2 is small enough that we can trust this calculation, it tells us how a must be varied with g^2 : $a\Lambda_{lat} = \exp(-1/2\beta_0 g^2)$.

Exercise: show that including the next order term in the β -function

$$\frac{dg}{d\ln a} = \beta_0 g^3 + \beta_1 g^5 + O(g^7), \qquad (147)$$

leads to

$$a\Lambda_{lat} = \exp\left[-\frac{1}{2\beta_0 g^2}\right] \left(g^2\beta_0\right)^{-\frac{\beta_1}{2\beta_0^2}} [1+0(g^2)].$$
(148)

(The inclusion of $\beta_0^{-\beta_1/2\beta_0^2}$ on the right hand side is a convention.) Here I am assuming that $O(a^2)$ terms can be ignored.

There are a number of important features of this result.

- The integration constant Λ_{lat} is not determined—i.e. we do not know a priori what value of a is associated with, say, $g^2 = 1$. This must be determined by comparison with experiment. It turns out that (in the quenched approximation discussed in Mackenzie's lectures) $a(g^2 = 1) \approx 0.1$ fm.
- Identical equations define a Λ -parameter in any regularization scheme, e.g. there is a Λ_{MOM} , a $\Lambda_{\overline{\text{MS}}}$, etc. These all serve the same purpose, that of specifying what the coupling constant is in the given scheme at a particular scale. We can use perturbation theory to relate the Λ -parameters (i.e. the coupling constants) in different schemes. See below.
- Numerical simulations are restricted to a range of lattice spacings which is roughly 0.2 fm > a > 0.05 fm at present: $O(a^2)$ errors become too large above the upper limit, while the lattice volume ($N_s a$, where N_s is the number of points across the lattice) becomes too small to contain hadrons below the lower limit. The rapid decrease of a as $g \to 0$ means that this corresponds to a very small range of g^2 . In the quenched approximation, it turns out to be $5.7 < \beta < 6.5$

We expect the pure gauge theory to have a spectrum of glueballs, with "physical" masses $m_{g,n}$. (Physical is in quotes as the real world is not a pure gauge theory, though we can imagine that it might have been.) We calculate these masses in lattice units (numerically, say), at a number of values of g^2 . To have a continuum limit they should behave as

$$m_{lat,n} = am_{g,n} = \frac{m_{g,n}}{\Lambda_{lat}} \exp\left[-\frac{1}{2\beta_0 g^2}\right] \left(g^2 \beta_0\right)^{-\frac{\beta_1}{2\beta_0^2}} \left[1 + 0(g^2) + O(a^2)\right],$$
(149)

where I have reinserted the expected $O(a^2)$ corrections. This equation is quite remarkable. The masses $m_{lat,n}$ are non-perturbative—the RHS of the equation has an essential zero at $g^2 = 0$ —but, using perturbation theory, we can predict how the masses decrease as $g \to 0$. Strictly speaking, what we are assuming here is that there is a continuum limit in which both perturbative quantities (such as $g_R(p)$) and non-perturbative quantities simultaneously have well defined limits. Another way of saying this is that all dimensionful quantities must be proportional to Λ_{lat} , as there are no other scales in the theory. Aside from corrections which fall as $O(a^2)$, the output from simulations is a set of pure numbers c_n

$$\frac{m_{g,n}}{\Lambda_{lat}} = c_{glue,n} \qquad \frac{\sqrt{\sigma}}{\Lambda_{lat}} = c_{\sigma} .$$
(150)

I should mention that there is a small segment of the lattice community which does not accept the above[12]. It is logically possible that there is no confinement and that $c_{\sigma} = c_{glue} = 0$, etc. How would this happen? We know that there is confinement for strong coupling, and numerical simulations extend this result up to $\beta \sim 6.5$. Furthermore, in the region $\beta \sim 6 - 6.5$ the expected dependence of masses on g^2 (Eq. 149) has been verified (as long as one implements perturbation theory correctly[13]). Nevertheless, it remains possible that there is a phase transition at weaker coupling, beyond which the $c_n = 0$.

There is a potential confusion that I wish to dispel. If we calculate σ and $m_{g,n}$ in perturbation theory, we find that they vanish, to all orders in g^2 . This is consistent with Eq. 149, as the result is non-perturbative. But how does perturbation theory fail to get the correct result? The point is that perturbation theory assumes the link matrices $U_{n,\mu}$ can all be rotated to lie close to the identity. This is *false* on scales of $\xi = \frac{1}{m} \propto \exp(1/2\beta_0 g^2)$ or greater. Once we get out to these length scales, important non-perturbative fluctuations are occurring, those that build up the hadrons. This is not important for $g_R(p)$, however, as long as $p \gg m$, for then $g_R(p)$ is only sensitive to distances much shorter than 1/m.

6.6 Comparison to critical phenomena

There is a large overlap between the analyses of critical phenomena and the continuum limit of lattice theories. As a critical temperature T_c is approached, the correlation length, defined by the two point function of some operator,

$$\langle \mathcal{O}(x)\mathcal{O}(0)\rangle \sim e^{-\frac{x}{\xi}},$$
 (151)

diverges as $\xi = |T - T_c|^{-\nu}$. The correlation-length exponent ν is one of a number of critical exponents. If there are no other scales ("relevant parameters") all correlation lengths are proportional to ξ . The lattice granularity becomes irrelevant as $\xi \to \infty$, corresponding to a continuum limit.

In this way of looking at the continuum limit, we hold the lattice spacing fixed, and adjust the coupling T such that the length scale of physical quantities diverges. This is not the way we are used to thinking about the continuum limit of lattice theories. Instead we keep physical sizes fixed, and imagine reducing the lattice spacing. The viewpoints are, however, entirely equivalent.

The detailed form of the critical behavior does differ for a gauge theory. First, there is the trivial change of using replacing T with g^2 . Second, we know that $g_c = 0$. But most importantly, the power law divergence is replaced by

$$\xi = 1/m \propto \exp[1/2\beta_0 g^2], \qquad (152)$$

i.e. an essential singularity.

6.7 Relating lattice and continuum coupling constants

An important application of perturbation theory is to relate coupling constants in different schemes. I will discuss how this works for an SU(3) pure gauge theory—the generalization to QCD involves simply changing some numerical factors. I am poaching somewhat on the subject matter of Paul Mackenzie's lectures, but I can't resist as the result is one of the present triumphs of lattice QCD.

Physical quantities must be independent of the regularization used to define the theory. Thus, if we calculate the renormalized coupling $g_R(p)$ in the MS scheme,

$$g_R(p) = g_{\overline{\text{MS}}} \left\{ 1 + g_{\overline{\text{MS}}}^2(\mu) [\beta_0 \ln(\frac{\mu}{p}) + C^{\overline{\text{MS}}}] + 0(g^4) \right\} , \qquad (153)$$

and equate it with the lattice result, Eq. 143, we find the relation between the couplings in the two schemes

$$g = g_{\overline{\text{MS}}}(\mu) \left\{ 1 + g^2 [\underbrace{(C^{\overline{\text{MS}}} - C^L)}_{\Delta C = -0.234} + \beta_0 \ln(\mu a)] + 0(g^4) + O(a^2) \right\}.$$
 (154)

At this order, one can use equally well use g^2 or $g_{\overline{MS}}^2$ in the correction term. This correction is quite large—if we take g = 1, a typical value in present simulations, corresponding to $1/a \approx 2 \,\mathrm{GeV}$, the lattice value of q^2 is 47% smaller than $g_{\overline{\text{MS}}}(\mu = 1/a)^2$. The size of this correction is well understood [13]—it is mainly due to the tadpole diagram. This large correction means that only one of the coupling constants can be a good expansion parameter for quantities involving momentum flows of $p \sim 2 \,\text{GeV}$. Experience with perturbative QCD indicates that expansions in $\alpha_{\overline{\text{MS}}} = g_{\overline{\text{MS}}}^2/4\pi$ work well for jet cross sections and other such quantities. Thus $\alpha_{lat}(a=1/p) = g^2(a=1/p)/4\pi$ will be a poor expansion parameter for such processes, and is likely to be poor expansion parameter in general. This is true in practice— 1- and 2-loop perturbative results for small Wilson loops, expressed in terms of α_{lat} , disagree significantly with results obtained from numerical simulations. Lepage and Mackenzie have shown, however, that the perturbative results work well if reexpressed in terms of $\alpha_{\overline{\text{MS}}}[13]$. See Mackenzie's lectures for more details.

Using an improved form of Eq. 154 suggested by Ref. [13], one can convert reliably from α_{lat} to $\alpha_{\overline{MS}}$. If one has established the lattice spacing a by comparing a physical quantity such as f_{π} to its lattice value, $a = f_{\pi}^{lat}/f_{\pi}^{phys}$, then the outcome is a *prediction* for $\alpha_{\overline{MS}}$ at a known physical scale. This can then be run to any other scale using the renormalization group. The latest result is |14|

$$\alpha_{\overline{\rm MS}}^{(5)}(m_Z) = 0.115 \pm 0.002\,, \tag{155}$$

where the error is claimed to account for all systematic and statistical effects. This is a very impressive result, and is consistent with the latest world average obtainde from comparisons of high-energy experiments with perturbative expansions [15] $\alpha_{\overline{\text{MS}}}^{(5)}(m_Z) =$ 0.117 ± 0.005 . This is a nice demonstration that QCD works simultaneously in the perturbative and non-perturbative regimes.

7 Fermions on the Lattice

Fermions are notoriously difficult to discretize in a satisfactory way, because of the so-called "doubling" problem. I devote the last lecture to an explanation of this problem, and a brief discussion of possible resolutions. I mainly focus on free fermions, because most aspects of the problem can be understood without coupling them to gauge fields. I work entirely in Euclidean space—it is worth noting, however, that the problem cannot be overcome by working in the Hamiltonian formulation where one discretizes space but not time.

Let me begin with a reminder of the Euclidean-space fermion action. In Minkowskispace the action is

$$S_M = \int_x \overline{\psi}(i\partial \!\!\!/ - m_{phys})\psi, \quad \text{where} \quad \overline{\psi} = \psi^{\dagger}\gamma_M^0, \quad \{\gamma_M^{\mu}, \gamma_M^{\nu}\} = 2g^{\mu\nu} \quad (156)$$

 S_M is hermitian because the Dirac matrices satisfy $\gamma_M^{\mu \dagger} \gamma_M^{0 \dagger} = \gamma_M^0 \gamma_M^{\mu}$. Now go to Euclidean space by the Wick rotation $x_0 \to -ix_4$, so that

$$S_{M} = \int_{x} d^{4}x_{M} \overline{\psi} (i\gamma_{M}^{0} \frac{\partial}{\partial x^{0}} + i\gamma_{M}^{i} \frac{\partial}{\partial x^{i}} - m_{phys})\psi \longrightarrow$$
$$S_{E} = -\int d^{4}x_{E} \overline{\psi} (\gamma_{E}^{4} \frac{\partial}{\partial x^{4}} + \gamma_{E}^{i} \frac{\partial}{\partial x^{i}} + m_{phys})\psi$$
(157)

where the Euclidean Dirac matrices are

$$\gamma_E^4 = \gamma_M^0 \,, \quad \gamma_E^i = -i\gamma_M^i \,, \quad \{\gamma_E^\mu, \gamma_E^\nu\} = 2\delta^{\mu\nu} \,, \quad \gamma_E^\mu = (\gamma_E^\nu)^\dagger \,. \tag{158}$$

Thus the Euclidean action is $S_E = -\int \overline{\psi} (\partial \!\!\!/_E + m_{phys}) \psi$. From now on I will drop the subscript E.

In the functional integral representation for the Euclidean partition function, fermions are Grassman variables, and we must treat ψ and $\overline{\psi}$ as independent fields. The rules of Grassman integration then yield

$$Z = \int [d\psi] [d\overline{\psi}] \, \exp[\int \overline{\psi}(\partial \!\!\!/ + m_{phys})\psi] = \det(\partial \!\!\!/ + m_{phys}), \qquad (159)$$

$$G(x,y) = -Z^{-1} \int [d\psi] [d\overline{\psi}] \exp[\int \overline{\psi}(\partial \!\!\!/ + m_{phys})\psi] \psi(x)\overline{\psi}(y) = [\frac{1}{\partial \!\!\!/ + m_{phys}}]_{xy}.$$
 (160)

The appearance of the determinant in the numerator, rather than the denominator as for scalar fields, corresponds to the minus sign for fermion loops. For a general Greens function the anticommuting nature of Grassman variables ensures the correct relative sign between different Wick contractions.

Now to the issue of discretization. We place fermions and antifermions on sites

$$a^{3/2}\psi(x) \to \psi_n , \quad a^{3/2}\overline{\psi}(x) \to \overline{\psi}_n ,$$
 (161)

Possible options for the derivative are

$$a^{5/2}\partial_{\mu}\psi(x) \to \begin{cases} \Delta^{+}_{\mu}\psi_{n} = \psi_{n+\mu} - \psi_{n} & (A), \\ \Delta^{-}_{\mu}\psi_{n} = \psi_{n} - \psi_{n-\mu} & (B), \\ \Delta_{\mu}\psi_{n} = \frac{1}{2}(\Delta^{+} + \Delta^{-})\psi_{n} = \frac{1}{2}(\psi_{n+\mu} - \psi_{n-\mu}) & (C). \end{cases}$$
(162)

When discretizing the scalar kinetic term $(|\partial_{\mu}\phi|^2)$, (A) and (B) are equivalent, and preferable to (C), because they are more local. For fermions, note that the Euclidean ∂ is **anti-**hermitean. This is a property we wish to preserve, as it traces back to the Hermiticity of the Hamiltonian. This eliminates options (A) and (B), since^{**}

$$(\Delta_{\mu}^{+})^{\dagger} = -(\Delta_{\mu}^{-}).$$
(163)

If we insist on nearest neighbors, we are forced to use (C), which leads to

$$\int \overline{\psi}(\partial \!\!\!/ + m_{phys})\psi \longrightarrow \sum_{n,\mu} \frac{1}{2}\overline{\psi}_n \gamma_\mu (\psi_{n+\mu} - \psi_{n-\mu}) + \sum_n m\overline{\psi}_n \psi_n = -S_N , \qquad (164)$$

where the lattice mass is $m = m_{phys}a$. This straightforward discretization gives rise to what are called "naive" lattice fermions.

To study naive fermions we look at the two-point function, $G = 1/(\partial + m)$. As in the continuum, ∂ is diagonal in momentum space. Introducing

$$\psi_n = \int_{-\pi}^{\pi} \frac{d^4k}{(2\pi)^4} e^{ikn} \,\psi(k) \quad \text{and} \quad \overline{\psi}_n = \int_{-\pi}^{\pi} \frac{d^4k}{(2\pi)^4} e^{-ikn} \,\overline{\psi}(k) \tag{165}$$

we find $\left(\int_k \equiv \int_{-\pi}^{\pi} d^4 k / (2\pi)^4\right)$

$$-S_N = \int_k \overline{\psi}(k) (i \sum_{\mu} s_{\mu} \gamma_{\mu} + m) \psi(k) .$$
(166)

Note that the discrete form of k_{μ} is $s_{\mu} = \sin k_{\mu}$, rather than the $2 \sin \frac{k_{\mu}}{2}$ we found with scalars. Thus the propagator in momentum space is

$$G(k) = \frac{1}{i\not s + m} = \frac{-i\not s + m}{s^2 + m^2}.$$
(167)

If we take the continuum limit with fixed physical mass and momenta, then $k = k_{phys}a \rightarrow 0$ and $m = m_{phys}a \rightarrow 0$. We can expand the sine, $s_{\mu} = ak_{\mu,phys}(1 + O(a^2))$, yielding

$$aG(k) \approx \frac{-i\gamma_{\mu}k_{\mu,phys} + m_{phys}}{k_{phys}^2 + m_{phys}^2} .$$

$$(168)$$

This has a pole at $k_{phys}^2 = -m_{phys}^2$, representing the fermion that we expected to find.

Now we come to doubling. The lattice momentum function s_{μ} vanishes for $k_{\mu} = \pi$ as well as $k_{\mu} = 0$. In the neighborhood of the momentum $(\pi, 0, 0, 0)$, if we define new variables by $k'_{1} = \pi - k_{1}$, $k'_{i} = k_{i}$, i = 2 - 4, then

$$G(k') \approx \frac{-i\sum_{\mu} k'_{\mu}\gamma'_{\mu} + m}{k'^2 + m^2}$$
 (169)

^{**}One can see that (A) and (B) are unphysical by noting that they correspond to propagation only forwards or backwards, respectively, but not in both directions. They cannot yield a Lorentz invariant Minkowski theory.

To bring the propagator into the standard continuum form, I have introduced new gamma-matrices, $\gamma'_1 = -\gamma_1$, $\gamma'_i = \gamma_i$, i = 2 - 4, unitarily equivalent to the standard set

$$\gamma_{\rho}' = (\gamma_{\mu}\gamma_{5})\gamma_{\rho}(\gamma_{\mu}\gamma_{5})^{\dagger} \quad (\text{no sum on } \mu).$$
(170)

Equation 169 shows that there is a second pole, at $k^{\prime 2} = -m^2$, which also represents a continuum fermion. This is our first "doubler".^{††}

The saga continues in an obvious way: s^2 vanishes if each of the four components of k_{μ} equals 0 or π . There is a pole near each of these 16 possible positions. Our single lattice fermion turns out to represent 16 degenerate states.

To further illuminate the doublers let us Fourier transform the propagator back to Euclidean time

$$G(\vec{k}, n_4) = \int_{-\frac{\pi}{2}}^{\frac{3\pi}{2}} \frac{dk_4}{2\pi} e^{ik_4 n_4} \frac{-i\not\!\!\!/ + m}{s^2 + m^2}$$
(171)

To evaluate the integral we first locate the poles. These occur when $-s_4^2 = \vec{s} \cdot \vec{s} + m^2$. Thus $\sin(k_4)$ is pure imaginary, implying

$$k_4 = n\pi + iE$$
, $n = \text{integer}$, $E = \pm \sinh^{-1}[\sqrt{\vec{s} \cdot \vec{s} + m^2}]$. (172)

The relevant poles are thus as shown here

 $-\frac{\pi}{2} \otimes \otimes \frac{3\pi}{2}$

We can close the contour as shown because of the periodicity of the integrand. If $n_4 > 0$ we pick up the two upper poles, while if $n_4 < 0$ we pick up the lower poles. A little work leads to the result

A little work leads to the result

$$G(\vec{k}, n_4) = \frac{m \pm \sinh [E\gamma_4] - i\vec{\gamma} \cdot \vec{s}}{\sinh [2E]} + \frac{m \mp \sinh [E\gamma_0] - i\vec{\gamma} \cdot \vec{s}}{\sinh [2E]} (-1)^{n_4} e^{-E|n_4|}, \qquad (173)$$

where the +(-) sign corresponds to $n_4 > 0$ (< 0). To interpret this result, recall the expression for the two point function in terms of the transfer matrix, Eq. 67.

^{††}Note that due to the periodicity of the lattice one can shift integration in momentum space from $\int_{-\pi}^{\pi}$ to $\int_{-\pi/2}^{3\pi/2}$, so there is no problem of k' lying near the boundary.

One can derive an analogous result for fermions

$$G(\vec{k}, n_4) = \langle 0 | \psi(\vec{k}) \overline{T}^{n_4} \hat{\psi}_0 | 0 \rangle \propto \sum_p |\langle 0 | \psi(\vec{k}) | p \rangle|^2 \lambda_p^{n_4} , \qquad (174)$$

where I have chosen $n_4 > 0$, and λ_p are the eigenvalues of the transfer matrix. The two terms in Eq. 173 thus correspond to two states. This is the doubling in the time direction. Note, however, that $\lambda_p = \pm \exp(-E)$, so the transfer matrix is not positive, and we cannot define a Hamiltonian by $\hat{H}a = -\ln \overline{T}$. We can overcome this problem by noting that the transfer matrix for two steps in the time direction is positive, so that a sensible definition of a hermitian Hamiltonian is

$$\overline{T}^2 = e^{-2\hat{H}a} \,. \tag{175}$$

In this case our two states have the same energy E.

The remaining octupling in the space directions is hidden in the expression for E. E has a minimum $(\sinh(E)_{\min} = m)$, when $\vec{s} = 0$, which occurs for eight values of \vec{k} in the integration range: $\vec{k} = (0, 0, 0)$, $(\pi, 0, 0)$, etc. Each of these corresponds to a fermion at rest, whose mass, in the continuum limit is $m/a = m_{phys}$.

7.1 Generality of the doubling problem

It is not, in fact, the replication of fermions which is the hard part of the problem, but rather the way in which the chiralities of the states work out. If m = 0, then we can introduce a chiral projection into the action

$$\gamma_{\mu} \to \gamma_{\mu}^{L} = \gamma_{\mu} (1 + \gamma_{5})/2 , \qquad (176)$$

which in the continuum restricts one to left-handed (LH) fields. On the lattice, the pole near k = 0 is then LH. The second pole I uncovered, however, represents a RH field. This is plausible, because

$$\gamma_{5}' = \gamma_{1}' \gamma_{2}' \gamma_{3}' \gamma_{4}' = -\gamma_{1} \gamma_{2} \gamma_{3} \gamma_{4} = -\gamma_{5} \quad \Rightarrow \quad (1+\gamma_{5}) = (1-\gamma_{5}') \,. \tag{177}$$

To actually show this one must consider the coupling to external currents.

Extending this analysis, it is easy to see that the chirality flips sign for each of the components of k that is near π . Thus one ends up with eight LH and eight RH fermions. This means that, when one introduces gauge fields (to be discussed below), one always obtains a "vector" representation of fermions, i.e. one in which LH and RH fields lie in the same representation of the gauge group.

How general is this result? Karsten and Smit have shown that LH and RH fermions always come in pairs [16], provided

is discretized into an <u>antihermitean</u> operator, so that its eigenvalues are imaginary;



Figure 4: Possible forms of the function F(k).

- the interactions are <u>local</u>, which implies that the propagator is continuous in momentum space;
- the space-time <u>volume is infinite</u>, which implies that momentum space is continuous;
- the action is translationally invariant, from which follows that momentum space is periodic (a torus with period 2π in each direction).

This result is easy to understand in 1 dimension. The propagator is of the form

$$G^{-1} = i\gamma_1 F(k) \tag{178}$$

where $k = k_1$ is the single momentum variable, and F is real, continuous function of k with period 2π . We are interested in the poles of G, and thus the zeroes of F(k). Possible forms for F are shown in Fig. 4. It is clear that, if there are only first order zeroes, there must be an even number in the interval $[-\pi/2, 3\pi/2]$. Furthermore, they come in pairs with opposite slopes. Following the above discussion, the slope corresponds to the chirality, so one always has an equal number of LH and RH fermions. The only alternative is to have a higher order zero, e.g. a zero with $\frac{\partial F}{\partial k} = 0$. This, however, gives a double pole in G, which does not correspond to a physical particle.

Returning to four dimensions, it is worth noting that one can truly reduce the issue to a doubling problem. Wilczek has given an example with only two states, at the price of using an action which breaks Euclidean rotation invariance [17].

7.2 Consequences of the doubling problem

What are the consequences of Karsten and Smit's result?

- One cannot discretize a chiral gauge theory, i.e. one in which the LH and RH fermions lie in different representations of the gauge group. These theories are well defined perturbatively, because one chooses the representations so that triangle anomalies cancel. But no satisfactory non-perturbative regulator exists, and the Karsten-Smit result rules out a simple lattice implementation. This means that one cannot discretize the electroweak sector of the standard model, which is chiral. If one tried, each (e_L, ν_L) doublet, for example, would come with an (e_R, ν_R) partner, which is not part of the standard model.
- Lattice regularization automatically takes care of the fact that theories with anomalous chiral representations of fermions (e.g. $SU(N_c)$ with a single left-handed fermion) cannot be defined.
- One cannot discretize QCD with (n_f) massless quarks, in the following sense. Such a theory should have an $SU(n_f)_L \times SU(n_f)_R$ chiral symmetry, under which the LH and RH quarks rotate with independent phases. But the lattice fermions are all begotten of the same lattice field, and so cannot be rotated independently.

In summary, then, the lattice theory lacks chiral symmetries.

Can we evade the general result? Can we simulate chiral theories? Can we simulate QCD with massless quarks? Much effort has been devoted to these questions. Notable among the attempts are:

- Avoid the Karsten-Smit result using a random lattice [18]. This breaks translation invariance, and thus the necessity of periodicity of the propagator. This idea is very difficult to analyze; even free fermions must be studied numerically. Little progress has been made—see Ref. [19] for a recent study.
- Use a non-local derivative, which allows the Fourier transform to be discontinuous, so, for example, one need have only one zero in F(k). An example is the "SLAC derivative" [20]. This fails when one introduces interactions with gauge fields—the doublers reappear because of the non-locality of the interactions[21].
- One can explicitly break chiral symmetry right from the start, and aim to recover it only in the continuum limit. This is, after all, what one does with the rotations and translations. For fermions in vector representations, this is the approach originally taken by Wilson, which I discuss in more detail below. For chiral theories, this is the approach advocated by the Rome group, and involves breaking the gauge symmetry at finite lattice spacing [22]. The approach has been shown to work in low-order perturbation theory. What

one is really interested in, however, is a non-perturbative simulation, and the theory is too complicated to simulate at present.

• The most exotic and interesting proposals are the descendants of Kaplan's "domain-wall fermions" [24], which all involve an infinite number of extra regulator fields. For a summary see Ref. [25]; the methods appear to work if there truely is an infinite number of fields. As for practical methods (necessarily restricted to finite numbers of fields), things are not yet clear. It does appear that one can simulate QCD with massless quarks, maintaining a genuine chiral symmetry. The question of whether the method is practical for chiral theories is being hotly debated.

It is worth noting that, even if a viable method for discretizing chiral fermions on finite lattices is developed, simulations will be hampered by the fact that the action is complex for such theories. This means that the factor $\exp(-S)$ in the functional integral cannot be interpreted as a probability.

What does one do if one wants to simulate QCD? In practice, one either gives up on chiral symmetry entirely, and uses Wilson fermions, which I discuss in the next subsection, or one uses "staggered" fermions. One can show that naive fermions brake up into four sets of four Dirac fermions, of which three can be ignored. The result is staggered fermions, which correspond to four degenerate fermions in the continuum limit. Such a theory, if m = 0, would have an $SU(4)_L \times SU(4)_R$ chiral symmetry. At finite lattice spacing, this is broken down to a flavor non-singlet axial U(1) symmetry. This is not much, but it is enough to guarantee that m is only multiplicatively, and not additively, renormalized. It is also important when calculating matrix elements which are constrained by chiral symmetry, such as $K \to \pi\pi$ amplitudes. Indeed, staggered fermions are the method of choice to study such quantities. For futher details of staggered fermions, see Ukawa's lectures, or Ref. [26].

7.3 Wilson fermions

I end these lectures with a description of the fermions used in most present simulations of QCD.

The simple way to understand why doublers occur is to note that the lattice derivative $\Delta^{\mu}\psi_n = \psi_{n+\mu} - \psi_{n-\mu}$ is small both for functions that are smooth, and for those that alternate in sign but are otherwise smooth. By contrast the bosonic derivative

$$\Delta^{+}_{\mu}\Delta^{-}_{\mu}\psi_{n} = \psi_{n+\mu} - 2\psi_{n} + \psi_{n-\mu}$$
(179)

is small only for smooth functions. Thus we try adding the "Wilson term"

$$S_W = \sum_{n\mu} \frac{r}{2} \overline{\psi}_n \Delta^+_\mu \Delta^-_\mu \psi_n \tag{180}$$

to the action, yielding

$$S_F = -\sum_n \overline{\psi}_n (m + \sum_\mu \gamma_\mu \Delta_\mu) \psi_n + S_W.$$
(181)

Exercise: show that the momentum space propagator is

$$G(k) = \frac{1}{i \not s + m + \frac{r}{2} \hat{k}^2}$$
(182)

$$= \frac{-i\not + (m - \frac{r}{2}\hat{k}^2)}{s^2 + (m + \frac{r}{2}\hat{k}^2)^2}, \qquad (183)$$

where $s_{\mu} = \sin(k_{\mu})$, and $\hat{k} = 2\sin(k_{\mu}/2)$.

If $k_{\mu} = \pi$, then $s_{\mu} = 0$, which is the cause of doubling, but $\hat{k_{\mu}} = 2$. Thus the would-be doubler poles picks up an effective mass $m_{\text{eff}} = m + 2rn$, where *n* is the number of components of k_{μ} close to π . If one keeps *r* finite in the continuum limit, when $m = m_{phys}a \to 0$, the effective lattice masses of the doublers stay finite, and so the effective physical masses become infinite. Thus only the single Dirac fermion corresponding to the pole near $\vec{k} = 0$ survives in the continuum limit.

Exercise: confirm this discussion by looking at the propagator as a function of Euclidean time. Take r = 1, which simplifies the calculation, and is the value used in most numerical simulations. One reason for this is that, for r = 1, one can derive a hermitean positive transfer matrix [27].

The drawback with Wilson fermions is that chiral symmetry is explicitly broken by the Wilson term, even when m = 0. This symmetry places important constraints on matrix elements involving plons, kaons and η 's, constraints which are therefore absent on the lattice. This makes it difficult to calculate some of these matrix elements, and for these one can do better with staggered fermions. The symmetry is regained in the continuum limit, because S_W vanishes: $S_W \sim a \int \overline{\psi} \Box \psi$.

It is straightforward to make the action gauge invariant by inserting appropriate link matrices in the derivatives

$$\overline{\psi}_n \psi_{n+\mu} \longrightarrow \overline{\psi}_n U_{n,\mu} \psi_{n+\mu} , \quad \overline{\psi}_n \psi_{n-\mu} \longrightarrow \overline{\psi}_n U_{n-\mu,\mu}^{\dagger} \psi_{n-\mu} .$$
(184)

For r = 1, the total gauged fermion action is thus (in d dimensions)

$$S_{F} = -\sum_{n} \overline{\psi}_{n} \psi_{n}(m+d) + \sum_{n} [\overline{\psi}_{n}(\frac{1-\gamma_{\mu}}{2})U_{n,\mu}\psi_{n+\mu} + \overline{\psi}_{n}(\frac{1+\gamma_{\mu}}{2})U_{n-\mu,\mu}^{\dagger}\psi_{n-\mu}]$$

$$= -\sum_{n} \overline{\psi}_{n}'\psi_{n}' + \kappa \sum_{n} [\overline{\psi}'(1-\gamma_{\mu})U_{n,\mu}\psi_{n+\mu}' + \overline{\psi}_{n}'(1+\gamma_{\mu}')U_{n-\mu,\mu}^{\dagger}\psi_{n-\mu}']$$
(185)

where I have introduced a rescaled field and a hopping parameter κ

$$\psi'_{n} = \sqrt{(m+d)}\psi_{n}, \quad \kappa = \frac{1}{2(m+d)}.$$
 (186)

The last form of the action is that used in most simulations. The quark mass is specified indirectly by $m = 1/(2\kappa) - d$. Large mass corresponds to $\kappa \to 0$, and in this limit one can calculate propagators using a hopping parameter expansion very similar to that discussed above for scalars.

The quark mass m vanishes when $\kappa = \kappa_c = 1/2d$, so this is the critical value to which κ should be tuned to take the continuum limit with free fermions. Gauge interactions additively renormalize κ_c —for finite a one must determine κ_c from the simulation itself. This is typically done by finding the value at which the pion mass vanishes, since $m_{\pi}^2 \propto m_q$. This renormalization is an example of the Applequist-Carrazone decoupling theorem. In the presence of interactions, one can decouple the doublers, but they cause finite renormalizations in the parameters of the remaining effective action. There is no reason why m (and thus κ) should not be renormalized additively, since there is no chiral symmetry when $m \to 0$.

8 References

- 1. H.J. Rothe, Lattice Gauge Theories (World Scientific, Singapore, 1992).
- I. Montvay and G. Münster, *Quantum Fields on a Lattice* (Cambridge University Press, Cambridge, 1994).
- 3. M. Creutz, *Quarks, Gluons and Lattices* (Cambridge University Press, Cambridge, 1983).
- 4. G. Parisi, "Statistical Field Theory", Addison-Wesley (1988).
- 5. K. Osterwalder and R. Schrader, Comm. Math. Phys. 42 (1975) 281.
- M. Lüscher, Commun. Math. Phys., 104 (1986) 177; 105 (1986) 153; Nucl. Phys. B354 (1991) 531.
- 7. J. Mandula, G.Zweig and J.Govaerts, Nucl. Phys. **B228** (1983) 91.
- 8. G.S. Bali and K. Schilling, Phys. Rev. D47 (1993) 661.
- 9. G. S. Bali et al.(UKQCD), Phys. Lett. **309B** (1993) 378.
- For example, D. Zwanziger, Nucl. Phys. B412 (1994) 657, P. van Baal, Nucl. Phys. B369 (1992) 259.
- 11. T. Reisz, Commm. Math. Phys. 116 (1988) 81, 573; ibid, 117 (1988) 79, 639.
- 12. A. Patrascioiu and Erhard Seiler, "Superinstantons and the reliability of perturbation theory in nonabelian models", hep-lat/9311019.
- 13. G.P. Lepage and P.B. Mackenzie, *Phys. Rev.* **D48** (1993) 2250.
- J. Shigemitsu, "Lattice gauge theory: status report 1994", 27th International Conference on High Energy Physics, Glasgow, Scotland, 20-27 July 1994, hepph/9410212.
- B.R. Webber, "QCD and jet physics", 27th International Conference on High Energy Physics, Glasgow, Scotland, 20-27 July 1994, hep-ph/9410268.
- L. H. Karsten and J. Smit, Nucl. Phys. B183 (1981) 103. The corresponding result for Hamiltonian theories was demonstrated in H. B. Nielsen and N. Ninomiya, Nucl. Phys. B185 (1981) 20, ibid., B193 (1981) 173.

- 17. F. Wilczek, Phys. Rev. Lett. 57 (1986) 2617.
- 18. N.H. Christ, R. Friedberg and T.D. Lee, Nucl. Phys. **B202** (1982) 89.
- 19. B. Alles *et al.* "Continuum limit of field theories regularized on a random lattice", hep-lat/9411014.
- 20. S.D. Drell, M. Weinstein and S. Yankielowicz, Phys. Rev. D14 (1976) 1627.
- 21. L.H. Karsten and J. Smit, *Phys. Lett.* **85B** (1979) 100.
- 22. A. Borelli et al., Nucl. Phys. **B333** (1990) 335.
- 23. L. Susskind, Phys. Rev. **D16** (1977) 3031.
- 24. D. B. Kaplan, Phys. Lett. 288B (1992) 342.
- 25. K. Jansen, "Domain Wall Fermions and Chiral Gauge Theories", heplat/9410018.
- 26. S. Sharpe, in *Standard Model, Hadron Phenomenology and Weak Decays on the Lattice*, ed. G. Martinelli, (World Scientific, Singapore, in press).
- M. Lüscher, Commun. Math. Phys. 54 (1977) 283;
 M. Creutz, Phys. Rev. D35 (1987) 1460.