LATTICE GAUGE THEORY FOR QCD

Thomas DeGrand
Department of Physics
University of Colorado, Boulder, CO 80309

ABSTRACT

These lectures provide an introduction to lattice methods for nonperturbative studies of Quantum Chromodynamics. Lecture 1 (Ch. 2): Basic techniques for QCD and results for hadron spectroscopy using the simplest discretizations; lecture 2 (Ch. 3): "improved actions"—what they are and how well they work; lecture 3 (Ch. 4): SLAC physics from the lattice—structure functions, the mass of the glueball, heavy quarks and $\alpha_s(M_Z)$, and $B - \bar{B}$ mixing.

1 Introduction

The lattice[1] version of QCD was invented by Wilson[2] in 1974. It has been a fruitful source of qualitative and quantitative information about QCD, the latter especially in the years since Creutz, Jacobs, and Rebbi[3] performed the first numerical simulations of a lattice gauge theory. Lattice methods are presently the only way to compute masses and matrix elements in the strong interactions beginning with the Lagrangian of QCD and including no additional parameters. In the past few years, the quality of many lattice predictions has become very high, and they are beginning to have a large impact in the wide arena of "testing the standard model." My goal in these lectures is to give enough of an overview of the subject that an outsider will be able to make an intelligent appraisal of a lattice calculation when s/he encounters one later on.

The first lecture will describe why one puts QCD on a lattice, and how it is done. This is a long story with a lot of parts, but at the end, I will show you "standard" lattice results for light hadron spectroscopy. The main problem with these calculations is that they are so unwieldy: to get continuum-like numbers requires very large-scale numerical simulations on supercomputers, which can take years to complete (sort of like the high-energy experiments themselves, except that we do not have to stack lead bricks). We would like to reduce the computational burden of our calculations. In Lecture Two, I will describe some of the different philosophies and techniques which are currently being used to invent "improved actions." Some of these methods actually work: some QCD problems can be studied on very large work stations. Finally, in Lecture Three, I will give a survey of recent lattice results for matrix elements, using physics done at SLAC as my unifying theme.

2 Gauge Field Basics

2.1 Beginnings

The lattice is a cutoff which regularizes the ultraviolet divergences of quantum field theories. As with any regulator, it must be removed after renormalization. Contact with experiment only exists in the continuum limit, when the lattice spacing is taken to zero.

We are drawn to lattice methods by our desire to study nonperturbative phenomena. Older regularization schemes are tied closely to perturbative expansions: one calculates a process to some order in a coupling constant; divergences are removed order by order in perturbation theory. The lattice, however, is a nonperturbative cutoff. Before a calculation begins, all wavelengths less than a lattice spacing are removed. Generally, one cannot carry out analytical studies of a field theory for physically interesting parameter values. However, lattice techniques lend themselves naturally to implementation on digital computers, and one can perform more-or-less realistic simulations of quantum field theories, revealing their nonperturbative structure, on present-day computers. I think it is fair to say that little of the quantitative results about QCD which have been obtained in the last decade, could have been gotten without the use of numerical methods.

On the lattice, we sacrifice Lorentz invariance but preserve all internal symmetries, including local gauge invariance. This preservation is important for nonperturbative physics. For example, gauge invariance is a property of the continuum theory which is nonperturbative, so maintaining it as we pass to the lattice means that all of its consequences (including current conservation and renormalizability) will be preserved.

It is very easy to write down an action for scalar fields regulated by a lattice. One just replaces the space-time coordinate x_μ by a set of integers n_μ ($x_\mu = an_\mu$, where a is the lattice spacing). Field variables $\phi(x)$ are defined on sites $\phi(x_n) \equiv \phi_n$. The action, an integral over the Lagrangian, is replaced by a sum over sites

$$\beta S = \int d^4x \mathcal{L} \to a^4 \sum_n \mathcal{L}(\phi_n), \tag{1}$$

and the generating functional for Euclidean Green's functions is replaced by an ordinary integral over the lattice fields

$$Z = \int (\prod_n d\phi_n) e^{-\beta S}. \tag{2}$$

Gauge fields are a little more complicated. They carry a space-time index μ in addition to an internal symmetry index a ($A^a_\mu(x)$) and are associated with a path in space $x_\mu(s)$: a particle traversing a contour in space picks up a phase factor

$$\psi \to P(\exp\ ig \int_s dx_\mu A_\mu)\psi \tag{3}$$

$$\equiv U(s)\psi(x). \tag{4}$$

P is a path-ordering factor analogous to the time-ordering operator in ordinary quantum mechanics. Under a gauge transformation g, $U(s)$ is rotated at each end:

$$U(s) \to g^{-1}(x_\mu(s))U(s)g(x_\mu(0)). \tag{5}$$

These considerations led Wilson[2] to formulate gauge fields on a space-time lattice, as follows:

The fundamental variables are elements of the gauge group G which live on the links of a four-dimensional lattice, connecting x and $x+\mu$: $U_\mu(x)$, with $U_\mu(x+\mu)^\dagger = U_\mu(x)$

$$U_\mu(n) = \exp(igaT^a A^a_\mu(n)) \tag{6}$$

for $SU(N)$ (g is the coupling, a the lattice spacing, A_μ the vector potential, and T^a is a group generator).

Under a gauge transformation, link variables transform as

$$U_\mu(x) \to V(x)U_\mu(x)V(x+\hat{\mu})^\dagger \tag{7}$$

and site variables as

$$\psi(x) \to V(x)\psi(x), \tag{8}$$

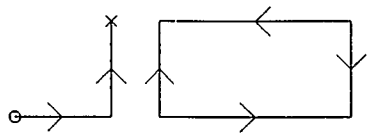
so the only gauge invariant operators we can use as order parameters are matter fields connected by oriented "strings" of U's [Fig. 1(a)]

$$\bar{\psi}(x_1)U_\mu(x_1)U_\mu(x_1+\hat{\mu})\ldots\psi(x_2) \tag{9}$$

or closed-oriented loops of U's [Fig. 1(b)]

$$\mathrm{Tr}\ldots U_\mu(x)U_\mu(x+\hat{\mu})\ldots \to \mathrm{Tr}\ldots U_\mu(x)V^\dagger(x+\hat{\mu})V(x+\hat{\mu})U_\mu(x+\hat{\mu})\ldots. \tag{10}$$

An action is specified by recalling that the classical Yang-Mills action involves the curl of A_μ, $F_{\mu\nu}$. Thus, a lattice action ought to involve a product of U_μ's around some closed contour. There is enormous arbitrariness at this point. We are trying to write down a bare action. So far, the only requirement we want to impose is gauge invariance, and that will be automatically satisfied for actions built of powers of traces of U's around closed loops, with arbitrary coupling constants. If we assume that the gauge fields are smooth, we can expand the link variables in a power series in gaA_μ's. For almost any closed loop, the leading term in the expansion will be proportional to $F^2_{\mu\nu}$. We might want our action to have the same normalization as the continuum action. This would provide one constraint among the lattice coupling constants.



(a) (b)

Figure 1: Gauge invariant observables are either (a) ordered chains ("strings") of links connecting quarks and antiquarks, or (b) closed loops of link variables.

The simplest contour has a perimeter of four links. In $SU(N)$,

$$\beta S = \frac{2N}{g^2}\sum_n\sum_{\mu>\nu}\mathrm{Re\ Tr}\ (1 - U_\mu(n)U_\nu(n+\hat{\mu})U^\dagger_\mu(n+\hat{\nu})U^\dagger_\nu(n)). \tag{11}$$

This action is called the "plaquette action" or the "Wilson action" after its inventor. The lattice parameter $\beta = 2N/g^2$ is often written instead of $g^2 = 4\pi\alpha_s$.

Let us see how this action reduces to the standard continuum action. Specializing to the U(1) gauge group, and slightly redefining the coupling,

$$S = \frac{1}{g^2}\sum_n\sum_{\mu>\nu}\mathrm{Re}\ (1 - \exp(iga[A_\mu(n) + A_\nu(n+\hat{\mu}) - A_\mu(n+\hat{\nu}) - A_\nu(n)])). \tag{12}$$

The naive continuum limit is taken by assuming that the lattice spacing a is small, and Taylor expanding

$$A_\mu(n+\hat{\nu}) = A_\mu(n) + a\partial_\nu A_\mu(n) + \ldots, \tag{13}$$

so the action becomes

$$\beta S = \frac{1}{g^2}\sum_n\sum_{\mu>\nu} 1 - \mathrm{Re}\ (\exp(iga[a(\partial_\nu A_\mu - \partial_\mu A_\nu) + O(a^2)])) \tag{14}$$

$$= \frac{1}{4g^2}a^4\sum_n\sum_{\mu\nu} g^2F^2_{\mu\nu} + \ldots \tag{15}$$

$$= \frac{1}{4}\int d^4x F^2_{\mu\nu} \tag{16}$$

transforming the sum on sites back to an integral.

2.2 Relativistic Fermions on the Lattice

Defining fermions on the lattice presents a new problem: doubling. The naive procedure of discretizing the continuum fermion action results in a lattice model with many more low energy modes than one originally anticipated. Let's illustrate this with free field theory.

The free Euclidean fermion action in the continuum is

$$S = \int d^4x[\bar{\psi}(x)\gamma_\mu\partial_\mu\psi(x) + m\bar{\psi}(x)\psi(x)]. \tag{17}$$

One obtains the so-called naive lattice formulation by replacing the derivatives by symmetric differences: we explicitly introduce the lattice spacing a in the denominator and write

$$S_L^{naive} = \sum_{n,\mu}\bar{\psi}_n\frac{\gamma_\mu}{2a}(\psi_{n+\mu} - \psi_{n-\mu}) + m\sum_n\bar{\psi}_n\psi_n. \tag{18}$$

The propagator is:

$$G(p) = (i\gamma_\mu \sin p_\mu a + ma)^{-1} = \frac{-i\gamma_\mu \sin p_\mu a + ma}{\sum_\mu \sin^2 p_\mu a + m^2a^2}. \tag{19}$$

We identify the physical spectrum through the poles in the propagator, at $p_0 = iE$:

$$\sinh^2 Ea = \sum_j \sin^2 p_j a + m^2a^2. \tag{20}$$

The lowest energy solutions are the expected ones at $p = (0,0,0)$, $E \simeq \pm m$, but there are other degenerate ones, at $p = (\pi,0,0)$, $(0,\pi,0,)$, $\ldots(\pi,\pi,\pi)$. This is a model for eight light fermions, not one.

(a) Wilson Fermions

There are two ways to deal with the doublers. The first way is to alter the dispersion relation so that it has only one low-energy solution. The other solutions are forced to $E \simeq 1/a$ and become very heavy as a is taken to zero. The simplest version of this solution (and almost the only one seen in the literature until recently) is due to Wilson: add a second-derivative-like term

$$S^W = -\frac{r}{2a}\sum_{n,\mu}\bar{\psi}_n(\psi_{n+\mu} - 2\psi_n + \psi_{n-\mu}) \tag{21}$$

to S^{naive}. The parameter r must lie between 0 and 1; $r = 1$ is almost always used and "$r = 1$" is implied when one speaks of using "Wilson fermions." The propagator is

$$G(p) = \frac{-i\gamma_\mu \sin p_\mu a + ma - r\sum_\mu(\cos p_\mu a - 1)}{\sum_\mu \sin^2 p_\mu a + (ma - r\sum_\mu(\cos p_\mu a - 1))^2}. \tag{22}$$

It has one pair of poles at $p_\mu \simeq (\pm im, 0, 0, 0)$, plus other poles at $p \simeq r/a$. In the continuum, these states become infinitely massive and decouple (although decoupling is not trivial to prove).

With Wilson fermions, it is conventional to use not the mass but the "hopping parameter" $\kappa = \frac{1}{2}(ma + 4r)^{-1}$, and to rescale the fields $\psi \to \sqrt{2\kappa}\psi$. The action for an interacting theory is then written

$$S = \sum_n \bar{\psi}_n\psi_n - \kappa\sum_{n\mu}(\bar{\psi}_n(r - \gamma_\mu)U_\mu(n)\psi_{n+\mu} + \bar{\psi}_n(r + \gamma_\mu)U_\mu^\dagger\psi_{n-\mu}). \tag{23}$$

Wilson fermions are closest to the continuum formulation—there is a four-component spinor on every lattice site for every color and/or flavor of quark. Constructing currents and states is just like in the continuum.

However, the Wilson term explicitly breaks chiral symmetry. This has the consequence that the zero bare quark mass limit is not respected by interactions; the quark mass is additively renormalized. The value of κ_c, the value of the hopping parameter at which the pion mass vanishes, is not known a priori before beginning a simulation; it must be computed. This is done in a simulation involving Wilson fermions by varying κ and watching the pion mass extrapolate quadratically to zero as $m_\pi^2 \simeq \kappa_c - \kappa$ ($\kappa_c - \kappa$ is proportional to the quark mass for small m_q). For the lattice person, this is unpleasant since preliminary calculations are required to find "interesting" κ values. For the outsider trying to read lattice papers, it is unpleasant because the graphs in the lattice paper typically list κ, and not quark (or pion) mass, so the reader does not know "where" the simulation was done. Note also that the relation between κ and physical mass changes with lattice coupling β.

(b) Staggered or Kogut-Susskind Fermions

In this formulation, one reduces the number of fermion flavors by using one-component "staggered" fermion fields rather than four-component Dirac spinors. The Dirac spinors are constructed by combining staggered fields on different lattice sites. Staggered fermions preserve an explicit chiral symmetry as $m_q \to 0$

even for finite lattice spacing, as long as all four flavors are degenerate. They are preferred over Wilson fermions in situations in which the chiral properties of the fermions dominate the dynamics—for example, in studying the chiral restoration/deconfinement transition at high temperature. They also present a computationally less intense situation from the point of view of numerics than Wilson fermions, for the trivial reason that there are less variables. However, flavor symmetry and translational symmetry are all mixed together. Construction of meson and baryon states (especially the Δ) is more complicated than for Wilson fermions.[4]

2.3 Enter the Computer

A "generic" Monte Carlo simulation in QCD breaks up naturally into two parts. In the "configuration generation" phase, one constructs an ensemble of states with the appropriate Boltzmann weighting: we compute observables simply by averaging N measurements using the field variables $\phi^{(i)}$ appropriate to the sample

$$\langle\Gamma\rangle \simeq \bar{\Gamma} \equiv \frac{1}{N}\sum_{i=1}^{N}\Gamma[\phi^{(i)}]. \tag{24}$$

As the number of measurements N becomes large, the quantity $\bar{\Gamma}$ will become a Gaussian distribution about a mean value. Its standard deviation is[5]

$$\sigma_\Gamma^2 = \frac{1}{N}\left(\frac{1}{N}\sum_{i=1}^{N}|\Gamma[\phi^{(i)}]|^2 - \bar{\Gamma}^2\right). \tag{25}$$

The idea of essentially all simulation algorithms is that one constructs a new configuration of field variables from an old one. One begins with some simple field configuration and monitors observables while the algorithm steps along. After some number of steps, the value of observables will appear to become independent of the starting configuration. At that point, the system is said to be "in equilibrium" and Eq. (24) can be used to make measurements.

The simplest method for generating configurations is called the Metropolis[6] algorithm. It works as follows: From the old configuration $\{\phi\}$ with action βS, transform the variables (in some reversible way) to a new trial configuration $\{\phi\}'$ and compute the new action $\beta S'$. Then, if $S' < S$, make the change and update all the variables; if not, make the change with probability $\exp(-\beta(S'-S))$.

Why does it work? In equilibrium, the rate at which configurations i turn into configurations j is the same as the rate for the back reaction $j \to i$. The rate of change is (number of configurations) $\times$ (probability of change). Assume for the sake of the argument that $S_i < S_j$. Then the rate $i \to j$ is $N_i P(i \to j)$ with $P(i \to j) = \exp(-\beta(S_j - S_i))$ and the rate $j \to i$ is $N_j P(j \to i)$ with $P(j \to i) = 1$. Thus, $N_i/N_j = \exp(-\beta(S_i - S_j))$.

If you have any interest at all in the techniques I am describing, you should write a little Monte Carlo program to simulate the two-dimensional Ising model. Incidentally, the favorite modern method for pure gauge models is overrelaxation.[7]

One complication for QCD which spin models don't have is fermions. The fermion path integral is not a number and a computer can't simulate fermions directly. However, one can formally integrate out the fermion fields. For n_f degenerate flavors of staggered fermions,

$$Z = \int [dU][d\psi][d\bar{\psi}]\exp(-\beta S(U) - \sum_{i=1}^{n_f}\bar{\psi}M\psi) \tag{26}$$

$$= \int [dU](\det M)^{n_f/2}\exp(-\beta S(U)). \tag{27}$$

(One can make the determinant positive-definite by writing it as $\det(M^\dagger M)^{n_f/4}$.) The determinant introduces a nonlocal interaction among the U's:

$$Z = \int [dU]\exp(-\beta S(U) - \frac{n_f}{4}\mathrm{Tr}\ln(M^\dagger M)). \tag{28}$$

All large-scale dynamical fermion simulations today generate configurations using some variation of the microcanonical ensemble. That is, they introduce momentum variables P conjugate to the U's and integrate Hamilton's equations through a simulation time t

$$\dot{U} = iPU \tag{29}$$

$$\dot{P} = -\frac{\partial S_{eff}}{\partial U}. \tag{30}$$

The integration is done numerically by introducing a timestep Δt. The momenta are repeatedly refreshed by bringing them in contact with a heat bath, and the method is thus called Refreshed or Hybrid Molecular Dynamics.[8]

For special values of n_f (multiples of two for Wilson fermions or of four for staggered fermions), the equations of motion can be derived from a local Hamiltonian, and in that case, Δt systematics in the integration can be removed by an extra Metropolis accept/reject step. This method is called Hybrid Monte Carlo.[9]

The reason for the use of these small timestep algorithms is that for any change in any of the U's, $(M^\dagger M)^{-1}$ must be recomputed. When Eq. (30) is integrated, all

of the U's in the lattice are updated simultaneously, and only one matrix inversion is needed per change of all the bosonic variables.

The major computational problem dynamical fermion simulations face is inverting the fermion matrix M. It has eigenvalues with a very large range—from 2π down to $m_q a$—and in the physically interesting limit of small m_q, the matrix becomes ill-conditioned. At present, it is necessary to compute at unphysically heavy values of the quark mass and to extrapolate to $m_q = 0$. The standard inversion technique today is one of the variants of the conjugate gradient algorithm.[10]

2.4 Taking the Continuum Limit, and Producing a Number in MeV

When we define a theory on a lattice, the lattice spacing a is an ultraviolet cutoff and all the coupling constants in the action are the bare couplings defined with respect to it. When we take a to zero, we must also specify how $g(a)$ behaves. The proper continuum limit comes when we take a to zero holding physical quantities fixed, not when we take a to zero holding the couplings fixed.

On the lattice, if all quark masses are set to zero, the only dimensionful parameter is the lattice spacing, so all masses scale like $1/a$. Said differently, one computes the dimensionless combination $am(a)$. One can determine the lattice spacing by fixing one mass from experiment. Then all other dimensionful quantities can be predicted.

Now imagine computing some masses at several values of the lattice spacing. (Pick several values of the bare parameters at random and calculate masses for each set of couplings.) Our calculated mass ratios will depend on the lattice cutoff. The typical behavior will look like

$$(am_1(a))/(am_2(a)) = m_1(0)/m_2(0) + O(m_1 a) + O((m_1 a)^2) + \dots . \qquad (31)$$

The leading term does not depend on the value of the UV cutoff, while the other terms do. The goal of a lattice calculation (like the goal of almost any calculation in quantum field theory) is to discover the value of some physical observable as the UV cutoff is taken to be very large, so the physics is in the first term. Everything else is an artifact of the calculation. We say that a calculation "scales" if the a-dependent terms in Eq. (31) are zero or small enough that one can extrapolate to $a = 0$, and generically refer to all the a-dependent terms as "scale violations."

We can imagine expressing each dimensionless combination $am(a)$ as some function of the bare coupling(s) $\{g(a)\}$, $am = f(\{g(a)\})$. As $a \to 0$, we must tune the set of couplings $\{g(a)\}$ so

$$\lim_{a\to 0} \frac{1}{a} f(\{g(a)\}) \to \text{constant}. \qquad (32)$$

From the point of view of the lattice theory, we must tune $\{g\}$ so that correlation lengths $1/ma$ diverge. This will occur only at the locations of second- (or higher-) order phase transitions in the lattice theory.

Recall that the β-function is defined by

$$\beta(g) = a\frac{dg(a)}{da} = \frac{dg(a)}{d\ln(1/\Lambda a)}. \qquad (33)$$

(There is actually one equation for each coupling constant in the set. Λ is a dimensional parameter introduced to make the argument of the logarithm dimensionless.) At a critical point, $\beta(g_c) = 0$. Thus, the continuum limit is the limit

$$\lim_{a\to 0}\{g(a)\} \to \{g_c\}. \qquad (34)$$

Continuum QCD is a theory with one dimensionless coupling constant. In QCD, the fixed point is $g_c = 0$, so we must tune the coupling to vanish as a goes to zero.

Pushing this a little further, the two-loop β-function is prescription independent,

$$\beta(g) = -b_1 g^3 + b_2 g^5, \qquad (35)$$

and so if we think that the lattice theory is reproducing the continuum, and if we think that the coupling constant is small enough that the two-loop β-function is correct, we might want to observe perturbative scaling, or "asymptotic scaling," m/Λ fixed, or a varying with g as

$$a\Lambda = \left(\frac{1}{g^2(a)}\right)^{b_2/(2b_1^2)} \exp\left(-\frac{1}{b_1 g^2(a)}\right). \qquad (36)$$

Asymptotic scaling is not scaling. Scaling means that dimensionless ratios of physical observables do not depend on the cutoff. Asymptotic scaling involves perturbation theory and the definition of coupling constants. One can have one without the other. (In fact, one can always define a coupling constant so that one quantity shows asymptotic scaling.)

And this is not all. There are actually two parts to the problem of producing a number to compare with experiment. One must first see scaling. Then one needs

to set the scale by taking some experimental number as input. A complication that you may not have thought of is that the theory we simulate on the computer is different from the real world. For example, a commonly used approximation is called the "quenched approximation": one neglects virtual quark loops but includes valence quarks in the calculation. The pion propagator is the propagator of a $\bar{q}q$ pair, appropriately coupled, moving in a background of gluons. This theory almost certainly does not have the same spectrum as QCD with six flavors of dynamical quarks with their appropriate masses. (In fact, an open question in the lattice community is what is the accuracy of quenched approximation?) Using one mass to set the scale from one of these approximations to the real world might not give a prediction for another mass which agrees with experiment. We will see examples where this is important.

2.5 Spectroscopy Calculations

"In a valley something like a race took place. A little crowd watched bunches of cars, each consisting of two 'ups' and a 'down' one, starting in regular intervals and disappearing in about the same direction. 'It is the measurement of the proton mass,' commented Mr. Strange, 'they have done it for ages. A very dull job, I am glad I am not in the game.'"[11]

Masses are computed in lattice simulations from the asymptotic behavior of Euclidean-time correlation functions. A typical (diagonal) correlator can be written as

$$C(t) = \langle 0|O(t)O(0)|0\rangle. \tag{37}$$

Making the replacement

$$O(t) = e^{Ht}Oe^{-Ht} \tag{38}$$

and inserting a complete set of energy eigenstates, Eq. (37) becomes

$$C(t) = \sum_n |\langle 0|O|n\rangle|^2 e^{-E_n t}. \tag{39}$$

At large separation, the correlation function is approximately

$$C(t) \simeq |\langle 0|O|1\rangle|^2 e^{-E_1 t}, \tag{40}$$

where E_1 is the energy of the lightest state which the operator O can create from the vacuum. If the operator does not couple to the vacuum, then in the limit of large t, one hopes to find the mass E_1 by measuring the leading exponential falloff of the correlation function, and most lattice simulations begin with that measurement. If the operator O has poor overlap with the lightest state, a reliable value for the mass can be extracted only at a large time t. In some cases, that state is the vacuum itself, in which $E_1 = 0$. Then one looks for the next higher state—a signal which disappears into the constant background. This makes the actual calculation of the energy more difficult.

This is the basic way hadronic masses are found in lattice gauge theory. The many calculations differ in important specific details of choosing the operators $O(t)$.

2.6 Recent Results

Today's supercomputer QCD simulations range from $16^3 \times 32$ to $32^3 \times 100$ points and run from hundreds (quenched) to thousands (full QCD) of hours on the fastest supercomputers in the world.

Results are presented in four common ways. Often one sees a plot of some bare parameter versus another bare parameter. This is not very useful if one wants to see continuum physics, but it is how we always begin. Next, one can plot a dimensionless ratio as a function of the lattice spacing. These plots represent quantities like Eq. (31). Both axes can show mass ratios. Examples of such plots are the so-called Edinburgh plot (m_N/m_ρ versus m_π/m_ρ) and the Rome plot (m_N/m_ρ versus $(m_\pi/m_\rho)^2$). These plots can answer continuum questions (how does the nucleon mass change if the quark mass is changed?) or can be used to show (or hide) scaling violations. Plots of one quantity in MeV versus another quantity in MeV are typically rather heavily processed after the data comes off the computer.

Let's look at some examples of spectroscopy, done in the "standard way," with the plaquette gauge action and Wilson or staggered quarks. I will restrict the discussion to quenched simulations because only there are the statistical errors small enough to be interesting to a nonlattice audience. Most dynamical fermion simulations are unfortunately so noisy that it is hard to subject them to detailed questioning.

Figure 2 shows a plot of the rho mass as a function of the size of the simulation, for several values of the quark mass (or m_π/m_ρ ratio in the simulation) and lattice

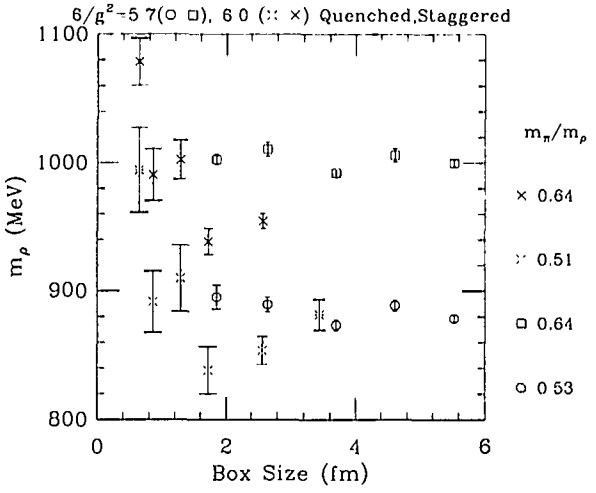


Figure 2: Rho mass vs box size.

spacing ($\beta = 6.0$ is $a \simeq 0.1$ fm and $\beta = 5.7$ is about twice that).[12] This picture shows that if the box has a diameter bigger than about 2 fm, the rho mass is little affected, but if the box is made smaller, the rho is "squeezed" and its mass rises.

Next, we look at an Edinburgh plot, Fig. 3 (Ref. 12). The different plotting symbols correspond to different bare couplings or (equivalently) different lattice spacings. This plot shows large scaling violations: mass ratios from different lattice spacings do not lie on top of each other. We can expose the level of scaling violations by taking "sections" through the plot and plot m_N/m_ρ at fixed values of the quark mass (fixed m_π/m_ρ) versus lattice spacing, in Fig. 4.

Now for some examples of scaling tests in the chiral limit. (Extrapolating to the chiral limit is a whole can of worms on its own, but for now, let's assume we can do it.) Figure 5 shows the nucleon to rho mass ratio (at chiral limit) versus lattice spacing (in units of $1/m_\rho$) for staggered[12] and Wilson[13] fermions. The "analytic" result is from strong coupling. The two curves are quadratic extrapolations to zero lattice spacing using different sets of points from the staggered data set. The burst is from a linear extrapolation to the Wilson data. The reason I show

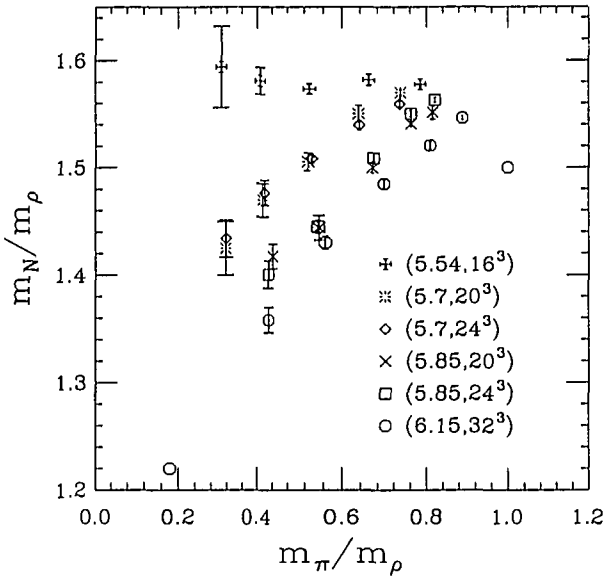


Figure 3: An Edinburgh plot for staggered fermions, from the MILC Collaboration.

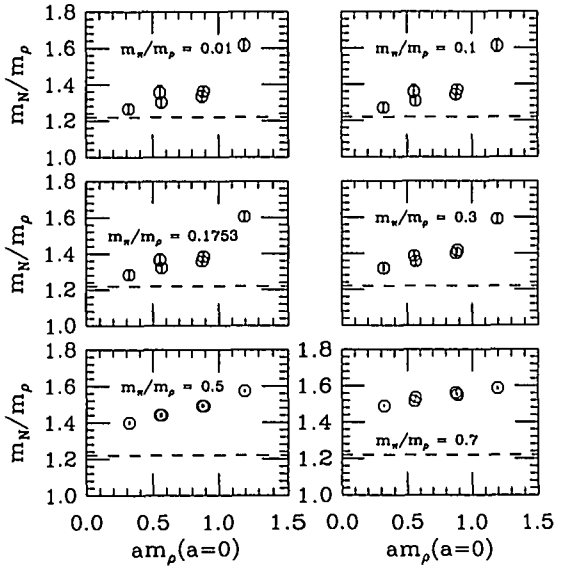


Figure 4: "Sections" through the Edinburgh plot.

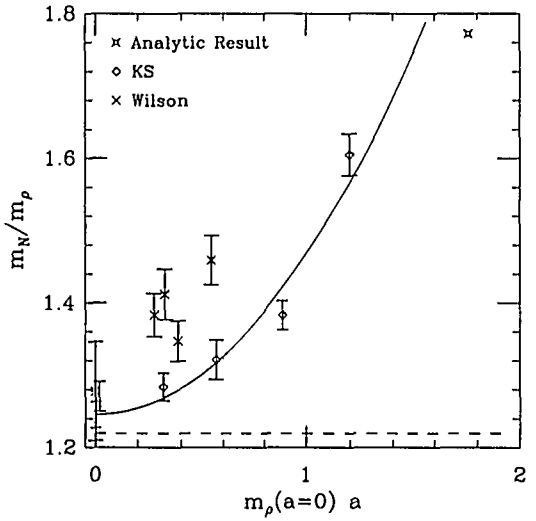


Figure 5: Nucleon to rho mass ratio (at chiral limit) vs lattice spacing (in units of $1/m_\rho$).

this figure is that one would like to know if the continuum limit of quenched spectroscopy "predicts" the real-world N/ρ mass ratio of 1.22 or not. The answer (unfortunately) depends on how the reader chooses to extrapolate.

Another test[14] is the ratio of the rho mass to the square root of the string tension, Fig. 6. Here the diamonds are staggered data and the crosses from the Wilson action. Scaling violations are large, but the eye extrapolates to something close to data (the burst).

Finally, despite Mr. Strange, very few authors have attempted to extrapolate to infinite volume, zero lattice spacing, and to physical quark masses, including the strange quark. One group which did, Butler *et al.*,[13] produced Fig. 7. The squares are lattice data, the octagons are the real world. They look quite similar within errors. Unfortunately, to produce this picture, they had to build their own computer.

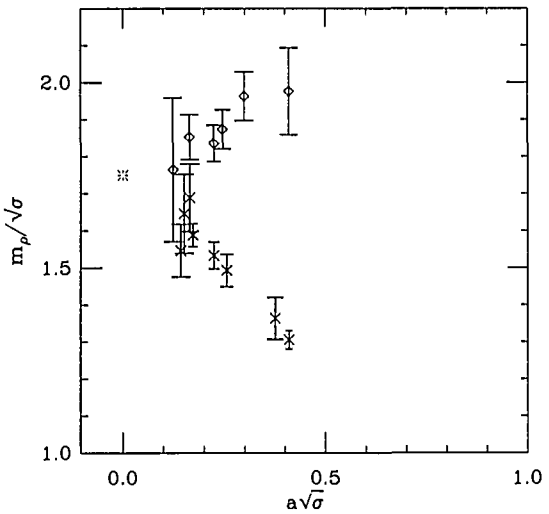


Figure 6: Scaling test for the rho mass in terms of the string tension, with data points labeled as in Fig. 5.

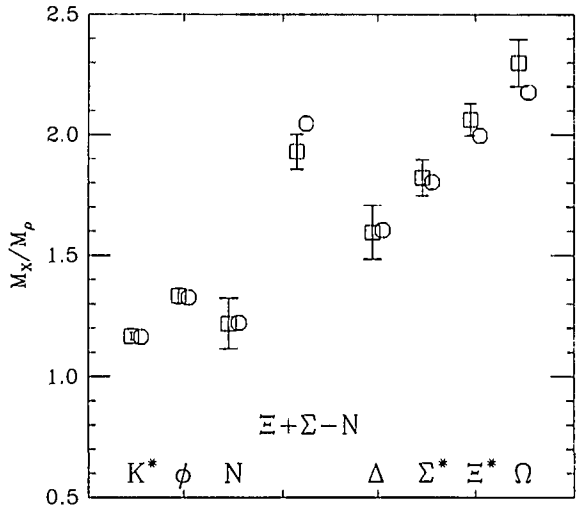


Figure 7: Quenched approximation mass ratios from Ref. 13.

3 Doing a Better Job—Maybe!

The slow approach to scaling presents a practical problem for QCD simulations, since it means that one needs to work at small lattice spacing. This is expensive. The cost of a Monte Carlo simulation in a box of physical size L with lattice spacing a and quark mass m_q scales roughly as

$$(\frac{L}{a})^4(\frac{1}{a})^{1-2}(\frac{1}{m_q})^{2-3}, \tag{41}$$

where the 4 is just the number of sites, the 1–2 is the cost of "critical slowing down"—the extent to which successive configurations are correlated, and the 2–3 is the cost of inverting the fermion propagator, plus critical slowing down from the nearly massless pions. The problem is that one needs a big computer to do anything.

However, all the simulations I described in the last lecture were done with a particular choice of lattice action: the plaquette gauge action, and either Wilson or staggered quarks. While those actions are the simplest ones to program, they

are just particular arbitrary choices of bare actions. Can one invent a better lattice discretization, which has smaller scaling violations?

People are trying many approaches. One could just write down a slightly more complicated action, include some parameters which can be tuned, do a spectroscopy calculation, and see if there is any improvement as the parameters are varied. The problem with this method is that it is like hunting for a needle in a multidimensional haystack—there are so many possible terms to add. One needs an organizing principle.

3.1 Improvement Based on Naive Dimensional Analysis

The simplest idea is to use the naive canonical dimensionality of operators to guide us in our choice of improvement. If we perform a naive Taylor expansion of a lattice operator like the plaquette, we find that it can be written as

$$\begin{aligned} 1-\frac{1}{3}\mathrm{Re\ Tr}U_{plaq} = \; & r_0 \mathrm{Tr}F_{\mu\nu}^2 + a^2[r_1 \textstyle\sum_{\mu\nu} \mathrm{Tr}D_\mu F_{\mu\nu} D_\mu F_{\mu\nu} + \\ & r_2 \textstyle\sum_{\mu\nu\sigma} \mathrm{Tr}D_\mu F_{\nu\sigma} D_\mu F_{\nu\sigma} + \\ & r_3 \textstyle\sum_{\mu\nu\sigma} \mathrm{Tr}D_\mu F_{\mu\sigma} D_\nu F_{\nu\sigma}] + \\ & + O(a^4). \end{aligned} \tag{42}$$

The expansion coefficients have a power series expansion in the coupling, $r_j = A_j + g^2 B_j + \ldots$ and the expectation value of any operator T computed using the plaquette action will have an expansion

$$\langle T(a) \rangle = \langle T(0) \rangle + O(a) + O(g^2 a) + \ldots . \tag{43}$$

Other loops have a similar expansion, with different coefficients. Now the idea is to take the lattice action to be a minimal subset of loops and systematically remove the a^n terms for physical observables order by order in n by taking the right linear combination of loops in the action.

$$S = \sum_j c_j O_j \tag{44}$$

with

$$c_j = c_j^0 + g^2 c_j^1 + \ldots . \tag{45}$$

This method was developed by Symanzik and co-workers[15-17] ten years ago.

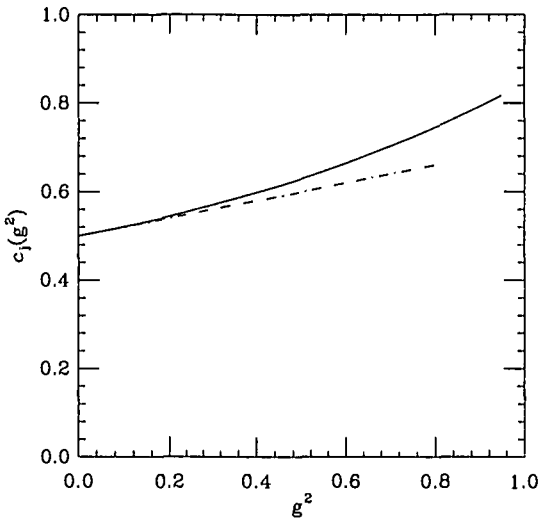


Figure 8: The value of some parameter in a lattice action for which physical observables have no a^n errors. The dotted line is the lowest order perturbative expectation.

To visualize this technique, look at Fig. 8. We imagine parameterizing the coefficients of various terms in the lattice action, which for a pure gauge theory could be a simple plaquette, a 1×2 closed loop, the square of the 1×2 loop, and so on, as some function of g^2. "Tree-level improvement" involves specifying the value of the j-th coefficient $c_j(g^2)$ at $g^2 = 0$. As we move away from $g^2 = 0$, the value of $c_j(g^2)$ for which observables calculated using the lattice action have no errors through the specified exponent n (no a^n errors) will trace out a trajectory in coupling constant space. For small g^2, the variation should be describable by perturbation theory, Eq. (45), but when g^2 gets large, we would not expect that perturbation theory would be a good guide.

The most commonly used "improved" fermion action is the "Sheikholeslami-Wohlert"[18] or "clover" action, an order a^2 improved Wilson action. The original Wilson action has $O(a)$ errors in its vertices, $S_W = S_c + O(a)$. This is corrected by making a field redefinition

$$\psi(x) \to \psi'(x) = \; \psi(x) + \tfrac{ia}{4}\gamma_\mu D_\mu \psi \tag{46}$$

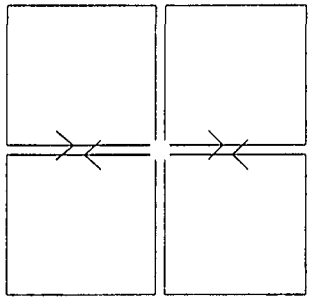


Figure 9: The "clover term."

$$\bar{\psi}(x) \to \bar{\psi}'(x) = \bar{\psi}(x) + \tfrac{ia}{4}\gamma_\mu \bar{\psi} D_\mu, \tag{47}$$

and the net result is an action with an extra lattice anomalous magnetic moment term,

$$S_{SW} - \frac{iag}{4}\bar{\psi}(x)\sigma_{\mu\nu}F_{\mu\nu}\psi(x). \tag{48}$$

It is called the "clover" action because the lattice version of $F_{\mu\nu}$ is the sum of paths shown in Fig. 9.

Studies performed at the time showed that this program did not improve scaling for the pure gauge theory (in the sense that the cost of simulating the more complicated action was greater than the savings from using a larger lattice spacing.) The whole program was re-awakened in the last few years by Lepage and collaborators,[19] and variations of this program give the most widely used "improved" lattice actions.

3.2 Nonperturbative Determination of Coefficients

Although I am breaking chronological order, the simplest approach to Symanzik improvement is the newest. The idea[20] is to force the lattice to obey various desirable identities to some order in a, by tuning parameters until the identities are satisfied by the simulations. That is, we try to find the solid line in Fig. 8 by doing simulations. Then use the action to calculate other things and test to see if scaling is improved. One example is the PCAC relation

$$\partial_\mu A^a_\mu = 2m^a_P + O(a), \tag{49}$$

where the axial and pseudoscalar currents are just

$$A^a_\mu(x) = \bar{\psi}(x)\gamma_\mu\gamma_5\frac{1}{2}\tau^a\psi(x) \tag{50}$$

and

$$P^a(x) = \bar{\psi}(x)\gamma_5\frac{1}{2}\tau^a\psi(x) \tag{51}$$

(τ^a is an isospin index.) The PCAC relation for the quark mass is

$$m \equiv \frac{1}{2}\frac{\langle\partial_\mu A^a_\mu O^a\rangle}{\langle P^a O^a\rangle} + O(a). \tag{52}$$

Now the idea is to take some Symanzik-improved action, with the improvement coefficients allowed to vary, and perform simulations in a little box with some particular choice of boundary conditions for the fields. Parameters which can be tuned include the c_{SW} in the clover term $i/4c_{SW}a\sigma_{\mu\nu}F_{\mu\nu}$ and the ones used for more complicated expressions for the currents

$$A_\mu = Z_A[(1 + b_A am_q)A^a_\mu + c_A a\partial_\mu P^a] \tag{53}$$

$$P^a = Z_P(1 + b_P am_q)P^a. \tag{54}$$

They are varied until the quark mass, defined in Eq. (52), is independent of location in the box or of the boundary conditions. Figures 10 and 11 illustrate what can be done with this tuning procedure. It is still too soon for definitive tests of scaling with this procedure.

3.3 Improving Perturbation Theory

The older version of Symanzik improvement uses lattice perturbation theory to compute the coefficients of the operators in the action. The idea here is to find a new definition of g^2 for which the solid line in Fig. 8 is transformed into a straight line (compare Fig. 12).

Let's make a digression into lattice perturbation theory.[21] It has three major uses. First, we need to relate lattice quantities (like matrix elements) to continuum

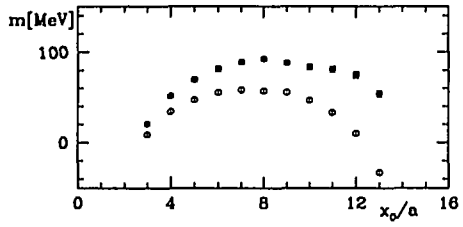


Figure 10: Values of the quark mass as computed from the axial and pseudoscalar currents, using the Wilson action. The open and full symbols correspond to different boundary conditions on the gauge fields.

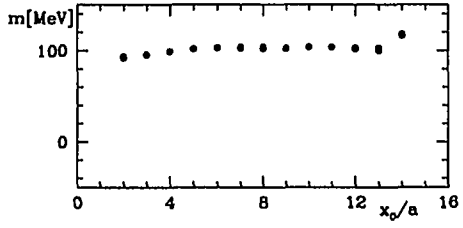


Figure 11: Same as previous figure, but now with improved action and operators.

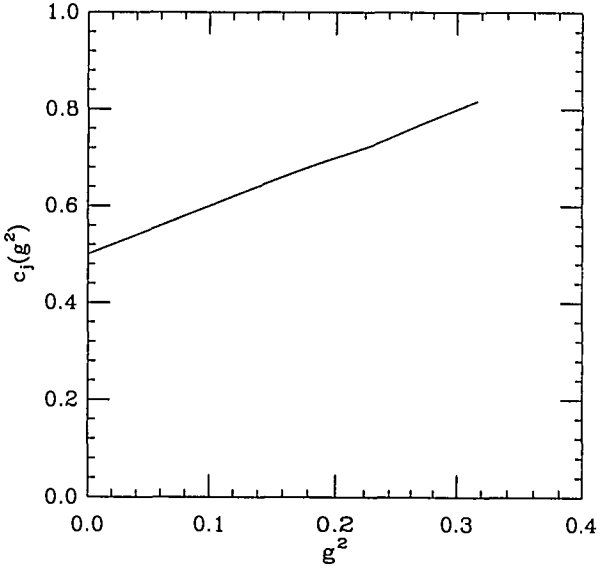


Figure 12: Same as Fig. 8, but now with a redefined coupling constant g_R^2, to make the "improvement line" linear in g_R^2.

ones: $O^{cont}(\mu) = Z(\mu a, g(a))O^{latt}(a)$. This happens because the renormalization of an operator is slightly different in the two schemes. In perturbation theory, Z has an expansion in powers of g^2. Second, we can use perturbation theory to understand and check numerical calculations when the lattice couplings are very small. Finally, one can use perturbative ideas to motivate nonperturbative improvement schemes.[22]

Perturbation theory for lattice actions is just like any other kind of perturbation theory (only much messier). One expands the Lagrangian into a quadratic term and interaction terms, and constructs the propagator from the quadratic terms:

$$\mathcal{L} = A_\mu(x)\rho_{\mu\nu}(x-y)A_\nu(y) + gA^3 + \ldots \quad (55)$$

$$= \mathcal{L}_0 + \mathcal{L}_I. \quad (56)$$

For example, the gluon propagator in Feynman gauge for the Wilson action is

$$D_{\mu\nu}(q) = \frac{g_{\mu\nu}}{\sum_\mu(1-\cos(q_\mu a))}. \quad (57)$$

To do perturbation theory for any system (not just the lattice), one has to do three things: one has to fix the renormalization scheme (RS) (define a coupling), specify the scale at which the coupling is defined, and determine a numerical value for the coupling at that scale. All of these choices are arbitrary, and any perturbative calculation is intrinsically ambiguous.

Any object which has a perturbative expansion can be written

$$O(Q) = c_0 + c_1(Q/\mu, RS)\alpha_s(\mu, RS) + c_2(Q/\mu, RS)\alpha_s(\mu, RS)^2 + \ldots . \quad (58)$$

In perturbative calculations, we truncate the series after a fixed number of terms and implicitly assume that's all there is. The coefficients $c_i(Q/\mu, RS)$ and the coupling $\alpha_s(\mu, RS)$ depend on the renormalization scheme and choice of scale μ. The guiding rule of perturbation theory[21] is "For a good choice of expansion the uncalculated higher order terms should be small." A bad choice has big coefficients.

There are many ways to define a coupling: The most obvious is the bare coupling; as we will see shortly, it is a poor expansion parameter. Another possibility is to define the coupling from some physical observable. One popular choice is to use the heavy quark potential at very high momentum transfer to define

$$V(q) \equiv 4\pi C_f \frac{\alpha_V(q)}{q^2}. \quad (59)$$

There are also several possibilities for picking a scale: One can use the bare coupling, then $\mu = 1/a$ the lattice spacing. One can guess the scale or play games just like in the continuum. One game is the Lepage-Mackenzie q^* prescription: find the "typical" momentum transfer q^* for a process involving a loop graph by pulling $\alpha_s(q)$ out of the loop integral and set

$$\alpha_s(q^*)\int d^4q\xi(q) = \int d^4q\xi(q)\alpha_s(q). \quad (60)$$

To find q^*, write $\alpha_s(q) = \alpha_s(\mu) + b\ln(q^2/\mu^2)\alpha_s(\mu)^2 + \ldots$, and similarly for $\alpha_s(q^*)$, insert these expressions into Eq. (60) and compare the $\alpha_s(\mu)^2$ terms, to get

$$\ln(q^*) = \int d^4q \ln(q)\xi / \int d^4q\xi. \quad (61)$$

This is the lattice analog of the Brodsky-Lepage-Mackenzie[23] prescription in continuum PT.

Finally, one must determine the coupling: If one uses the bare lattice coupling, it is already known. Otherwise, one can compute it in terms of the bare coupling:

$$\alpha_{\overline{MS}}(s/a) = \alpha_0 + (5.88 - 1.75\ln s)\alpha_0^2 + (43.41 - 21.89\ln s + 3.06\ln^2 s)\alpha_0^3 + \ldots . \quad (62)$$

Or one can determine it from something one measures on the lattice, which has a perturbative expansion. For example,

$$-\ln\langle\frac{1}{3}\mathrm{Tr}U_{plaq}\rangle = \frac{4\pi}{3}\alpha_P(3.41/a)(1 - 1.185\alpha_P) \quad (63)$$

(to this order, $\alpha_P = \alpha_V$). Does "improved perturbation theory" actually improve perturbative calculations? In many cases, yes: some examples are shown in Fig. 13 from Ref. 22: On the upper left, we see a calculation of the average link in Landau gauge, from simulations (octagons) and then from lowest-order perturbative calculations using the bare coupling (crosses) and α_V and $\alpha_{\overline{MS}}$ (diamonds and squares). In the upper right panel, we see how the lattice prediction of an observable involving the two-by-two Wilson loop depends on the choice of momentum q^*/a (at $\beta = 6.2$, a rather weak value of the bare coupling) in the running coupling constant. The burst is the value of the prescription of Eq. (61). In the lower panel are perturbative predictions, the same observables as a function of lattice coupling. These pictures illustrate that perturbation theory in terms of the bare coupling does not work well, but that using other definitions for couplings, one can get much better agreement with the lattice "data."

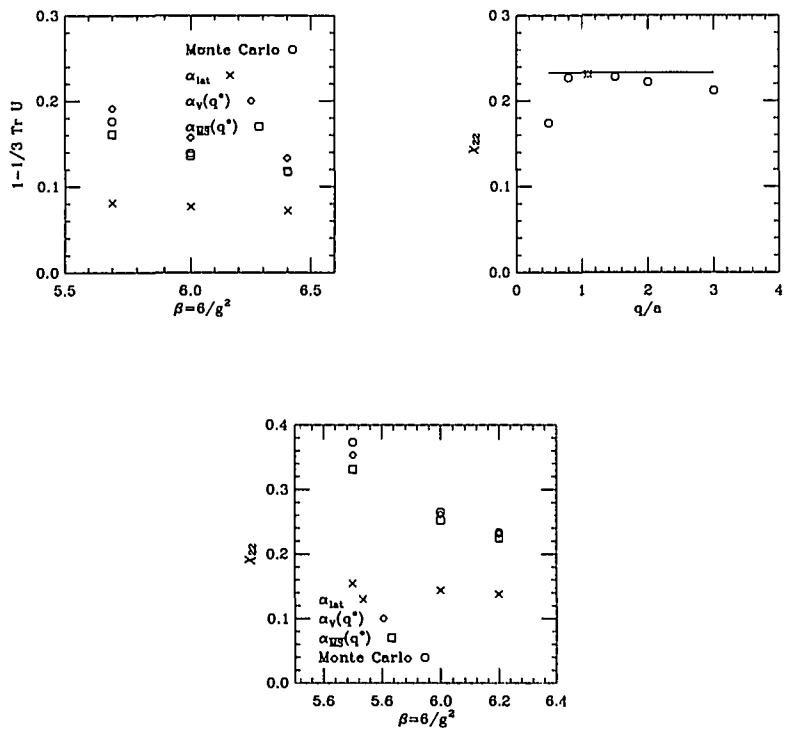


Figure 13: Examples of "improved perturbation theory."

Straight perturbative expansions by themselves for the commonly used lattice actions are typically not very convergent. The culprit is the presence of U_μ's in the action. One might think that for weak coupling, one could expand

$$\bar{\psi}U\psi = \bar{\psi}[1 + igaA + \ldots]\psi \tag{64}$$

and ignore the ..., but the higher order term $\bar{\psi}\frac{1}{2}g^2a^2A^2\psi$ generates the "tadpole graph" of Fig. 14. The UV divergence in the gluon loop $\simeq 1/a^2$ cancels the a^2 in the vertex. The same thing happens at higher order, and the tadpoles add up to an effective $a^0\sum c_n g^{2n}$ contribution. Parisi[21] and later Lepage and Mackenzie[22] suggested a heuristic way to deal with this problem: replace $U_\mu \to u_0(1 + igaA)$ where u_0, the "mean field term" or "tadpole improvement term" is introduced phenomenologically to sum the loops. Then one rewrites the action as

$$S = \frac{1}{g^2u_0^4}\sum \mathrm{Tr}U \leftrightarrow \frac{1}{g_{lat}^2}\sum \mathrm{TrU}, \tag{65}$$

where $g^2 \equiv g_{lat}^2/u_0^4$ is the new expansion parameter. Is $u_0^4 \equiv \langle \mathrm{Tr}U_{plaq}/3\rangle$? This choice is often used; it is by no means unique.

A "standard action" (for this year, anyway) is the "tadpole-improved Lüscher-Weisz[17] action," composed of a 1 by 1, 1 by 2, and "twisted" loop (+x, +y, +z, −x, −y, −z),

$$\beta S = -\beta[\mathrm{Tr}(1\times 1) - \frac{1}{20u_0^2}(1 + 0.48\alpha_s)\mathrm{Tr}(1\times 2) - \frac{1}{u_0^2}0.33\alpha_s\mathrm{Tr}U_{tw}] \tag{66}$$

with $u_0 \equiv \langle \mathrm{Tr}U_{plaq}/3\rangle^{1/4}$ and $3.068\alpha_s \equiv -\ln\langle \mathrm{Tr}U_{plaq}/3\rangle$ determined self-consistently in the simulation.

3.4 Fixed-Point Actions

Let's recall the question we were trying to answer in the previous sections: Can one find a trajectory in coupling constant space, along which the physics has no corrections to some desired order in a^n or $g^m a^n$? Let's take the question one step further: Is there a trajectory in coupling constant space in which there are *no* corrections at all, for any n or m?

To approach the answer, let's think about the connection between scaling and the properties of some arbitrary bare action, which we assume is defined with some UV cutoff a (which does not have to be a lattice cutoff, in principle). The

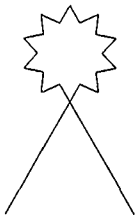


Figure 14: The "tadpole diagram."

action is characterized by an infinite number of coupling constants, $\{c\}$. (Many of them could be set to zero.) When the c's take on almost any arbitrary values, the typical scale for all physics will be the order of the cutoff: $m \simeq 1/a$, correlation length $\xi \simeq a$. There will be strong cutoff effects.

The best way to think about scaling is through the renormalization group.[25] Take the action with cutoff a and integrate out degrees of freedom to construct a new effective action with a new cutoff $a' > a$ and fewer degrees of freedom. The physics at distance scales $r > a$ is unaffected by the coarse-graining (assuming it is done exactly.) We can think of the effective actions as being similar to the original action, but with altered couplings. We can repeat this coarse-graining and generate new actions with new cutoffs. As we do, the coupling constants "flow":

$$S(a, c_j) \to S(a', c'_j) \to S(a'', c''_j) \to \ldots . \tag{67}$$

If under repeated blockings the system flows to a fixed point

$$S(a_n, c_j^n) \to S(a_{n+1}, c_j^{n+1} = c_j^n), \tag{68}$$

then observables are independent of the cutoff a, and in particular, the correlation length ξ must either be zero or infinite.

This can only happen if the original c's belong to a particular restricted set, called the "critical surface." It is easy to see that physics on the critical surface is universal: at long distances, the effective theory is the action at the fixed point, to which all the couplings have flowed, regardless of their original bare values.

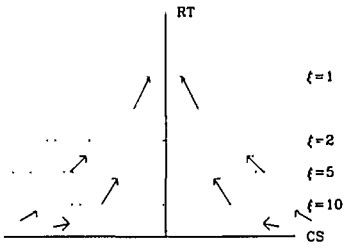


Figure 15: A schematic picture of renormalization group flows along a one-dimensional critical surface, with the associated renormalized trajectory, and superimposed contours of constant correlation length.

In particular, physics at the fixed point is independent of the underlying lattice structure.

But $\xi = \infty$ is not ξ large. Imagine tuning bare parameters close to the critical surface, but not on it. The system will flow towards the fixed point, then away from it. The flow lines in coupling constant space will asymptotically approach a particular trajectory, called the renormalized trajectory (RT), which connects (at $\xi = \infty$) with the fixed point. Along the renormalized trajectory, ξ is finite. However, because it is connected to the fixed point, it shares the scaling properties of the fixed point—in particular, the complete absence of cutoff effects in the spectrum and in Green's functions. (To see this remarkable result, imagine doing QCD spectrum calculations with the original bare action with a cutoff equal to the Planck mass and then coarse-graining. Now exchange the order of the two procedures. If this can be done without making any approximations, the answer should be the same.)

A Colorado analogy is useful for visualizing the critical surface and renormalized trajectory: think of the critical surface as the top of a high mountain ridge. The fixed point is a saddle point on the ridge. A stone released on the ridge will roll to the saddle and come to rest. If it is not released exactly on the ridge, it will roll near to the saddle, then go down the gully leading away from it. For a cartoon, see Fig. 15.

So the ultimate goal of "improvement programs" is to find a true perfect action, without cutoff effects, along the renormalized trajectory of some renormalization group transformation. At present, finding an RT has not been done in a convincing way for any renormalization group transformation. However, an action at the fixed point might also be an improved action, and fixed-point actions really can be constructed and used.

In lattice language, a bare action for QCD is described by one overall factor of $\beta = 2N/g^2$ and arbitrary weights of various closed loops,

$$\beta S = \frac{2N}{g^2} \sum_j c_j O_j. \tag{69}$$

Asymptotic freedom is equivalent to the statement that the critical surface of any renormalization group transformation is at $g^2 = 0$. The location of a fixed point involves some relation among the c_j's.

A direct attack on the renormalized trajectory begins by finding a fixed-point action. Imagine having a set of field variables $\{\phi\}$ defined with a cutoff a. Introduce some coarse-grained variables $\{\Phi\}$ defined with respect to a new cutoff a', and integrate out the fine-grained variables to produce a new action

$$e^{-\beta S'(\Phi)} = \int d\phi e^{-\beta(T(\Phi,\phi)+S(\phi))}, \tag{70}$$

where $\beta(T(\Phi,\phi)$ is the blocking kernel which functionally relates the coarse and fine variables. Integrating Eq. (70) is usually horribly complicated. However, P. Hasenfratz and F. Niedermayer[26] noticed an amazing simplification for asymptotically free theories: Their critical surface is at $\beta = \infty$, and in that limit, Eq. (70) becomes a steepest-descent relation

$$S'(\Phi) = \min_{\phi}((T(\Phi,\phi) + S(\phi)) \tag{71}$$

which can be used to find the fixed-point action

$$S_{FP}(\Phi) = \min_{\phi}((T(\Phi,\phi) + S_{FP}(\phi)). \tag{72}$$

The program has been successfully carried out for $d = 2$ sigma models[26] and for four-dimensional pure gauge theories.[27] These actions have two noteworthy properties: First, not only are they classically perfect actions (they have no a^n scaling violations for any n), but they are also one-loop quantum perfect: that is, as one moves out of the renormalized trajectory,

$$\frac{1}{g^2} S_{RT}(g^2) = \frac{1}{g^2}(S_{FP} + O(g^4)). \tag{73}$$

Physically, this happens because the original action has no irrelevant operators, and they are only generated through loop graphs. Thus, these actions are an extreme realization of the Symanzik program. Second, because these actions are at the fixed point, they have scale invariant classical solutions. This fact can be used to define a topological charge operator on the lattice in a way which is consistent with the lattice action.[28]

These actions are "engineered" in the following way: one picks a favorite blocking kernel, which has some free parameters, and solves Eq. (72), usually approximately at first. Then one tunes the parameters in the kernel to optimize the action for locality and perhaps refines the solution. Now the action is used in simulations at finite correlation length (i.e., do simulations with a Boltzmann factor $\exp(-\beta S_{FP})$). Because of Eq. (73), one believes that the FP action will be a good approximation to the perfect action on the RT; of course, only a numerical test can tell. As we will see in the next section, these actions perform very well. At this point in time, no nonperturbative FP action which includes fermions has been tested, but most of the formalism is there.[29]

3.5 Examples of "Improved" Spectroscopy

I would like to show some examples of the various versions of "improvement" and remind you of the pictures at the end of the last chapter to contrast results from standard actions.

Figure 16 shows a plot of the string tension measured in systems of constant physical size (measured in units of $1/T_c$, the critical temperature for deconfinement), for SU(3) pure gauge theory. In the quenched approximation, with $\sqrt{\sigma} \simeq 440$ MeV, $T_c = 275$ MeV and $1/T_c = 0.7$ fm. Simulations with the standard Wilson action are crosses, while the squares show FP action results[27] and the octagons from the tadpole-improved Lüscher-Weisz action.[30] The figure illustrates that it is hard to quantify improvement. There are no measurements with the Wilson action at small lattice spacing of precisely the same observables that the "improvement people" measured. The best one can do is to take similar measurements (the diamonds) and attempt to compute the $a = 0$ prediction for the observable we measured (the fancy cross at $a = 0$). This attempt lies on a straight line with the FP action data, hinting strongly that the FP action is indeed scaling. The FP action seems to have gained about a factor of three to four in

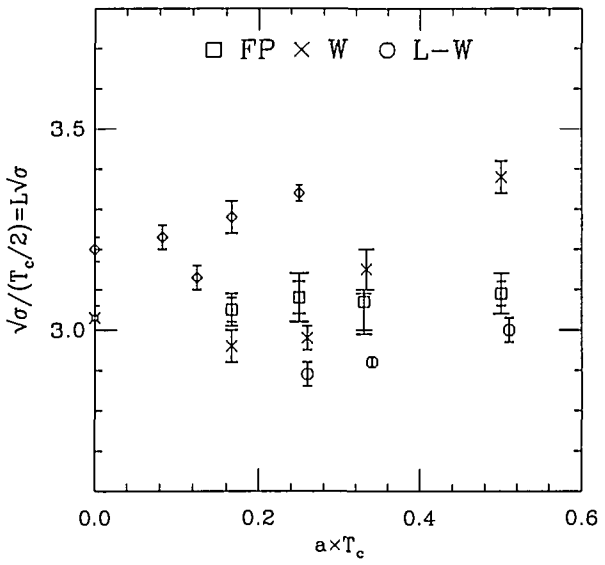


Figure 16: The square root of the string tension in lattices of constant physical size $L = 2/T_c$, but different lattice spacings (in units of $1/T_c$).

lattice spacing, or a gain of $(3-4)^6$ compared to the plaquette action, according to Eq. (41), at a cost of a factor of seven per site because it is more complicated to program. The tadpole-improved Lüscher-Weisz action data lie lower than the FP action data and do not scale as well. As $a \to 0$, the two actions should yield the same result; that is just universality at work. However, there is no guarantee that the approach to the continuum is monotonic.

Figure 17 shows the heavy quark-antiquark potential in SU(2) gauge theory, where $V(r)$ and r are measured in the appropriate units of T_c, the critical temperature for deconfinement. The Wilson action is on the left and an FP action is on the right. The vertical displacements of the potentials are just there to separate them. Notice the large violations of rotational symmetry in the Wilson action data when the lattice spacing is $a = 1/2T_c$, which are considerably improved in the FP action results.

Next, we consider nonrelativistic QCD. A comparison of the quenched charmonium spectrum from Ref. 19 using data from Ref. 31 is shown in Fig. 18.

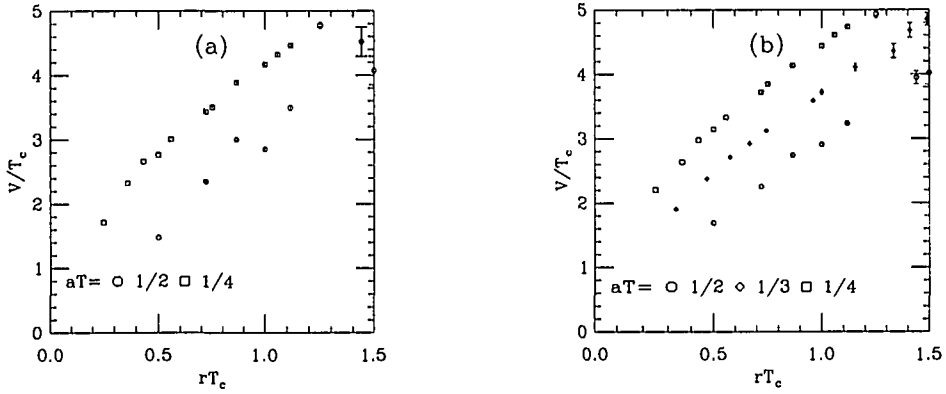


Figure 17: The heavy quark potential in SU(2) pure gauge theory measured in units of T_c. (a) Wilson action, and (b) an FP action.

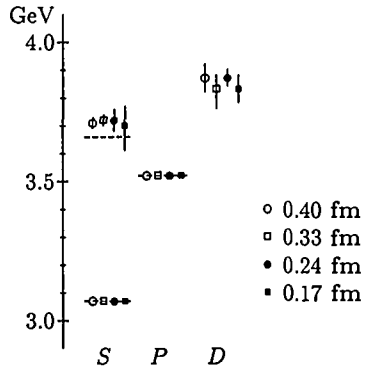


Figure 18: S, P, and D states of charmonium computed on lattices with: a = 0.40 fm (improved action, $\beta_{plaq} = 6.8$); $a = 0.33$ fm (improved action, $\beta_{plaq} = 7.1$); $a = 0.24$ fm (improved action, $\beta_{plaq} = 7.4$); and $a = 0.17$ fm (Wilson action, $\beta = 5.7$, from Ref. [31]), from Ref. 19. The dashed lines indicate the true masses.

When the tadpole-improved L-W action is used to generate gauge configurations, the scaling window is pushed out to $a \simeq 0.4$ fm for these observables.

Now we turn to tests of quenched QCD for light quarks. The two actions which have been most extensively tested are the S-W action, with and without tadpole improvement, and an action called the D234(2/3) action, a higher-order variant of the S-W action.[32] Figures 19 and 20 are the analogs of Figs. 5 and 6. Diamonds[33] and plusses[34] are S-W actions, ordinary and tadpole-improved, squares are the D234(2/3) action. They appear to have about half the scaling violations as the standard actions but they don't remove all scaling violations. It's a bit hard to quantify the extent of improvement from these pictures because a chiral extrapolation is hidden in them. However, one can take one of the "sections" of Fig. 4 and overlay the new data on it, Fig. 21. It looks like one can double the lattice spacing for an equivalent amount of scale violation. However, the extrapolation in a is not altogether clear. Figure 22 is the same data as Fig. 21, only plotted versus a^2, not a. All of the actions shown in these figures are supposed to have $O(a^2)$ (or better) scaling violations. Do the data look any straighter in Fig. 22 than in Fig. 21?

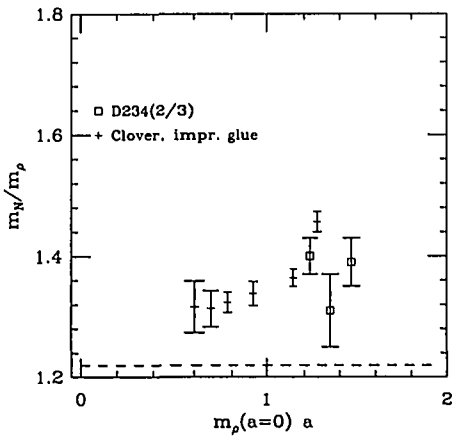


Figure 19: Nucleon to rho mass ratio (at chiral limit) vs lattice spacing (in units of $1/m_\rho$).

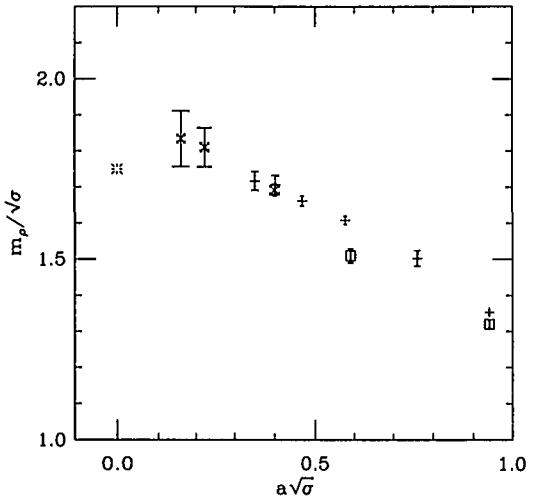


Figure 20: Rho mass scaling test with respect to the string tension.

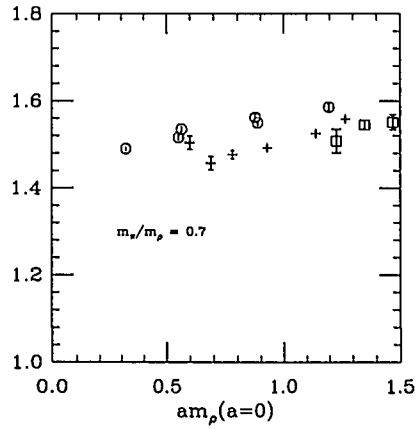


Figure 21: m_N/m_ρ vs am_ρ at fixed quark mass (fixed m_π/m_ρ). Interpolations of the S-W and D234(2/3) data were done by me.

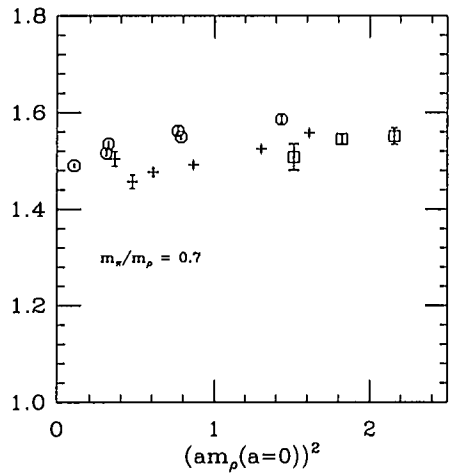


Figure 22: m_N/m_ρ vs $(am_\rho)^2$ at fixed quark mass (fixed m_π/m_ρ).

3.6 The Bottom Line

At the cost of enormous effort, one can do fairly high-precision simulations of QCD in the quenched approximation with standard actions. The actions I have shown you appear to reduce the amount of computation required for pure gauge simulations from supercomputers to very large work stations, probably a gain of a few hundreds. All of the light quark data I showed actually came from supercomputers. According to Eq. (41), a factor of two in the lattice spacing gains a factor of 64 in speed. The cost of either of the two improved actions I showed is about a factor of eight to ten times the fiducial staggered simulation. Improvement methods for fermions are a few years less mature than ones for pure gauge theory, and so the next time you hear a talk about the lattice, things will have changed for the better (maybe).

4 SLAC Physics from the Lattice

One of the major goals of lattice calculations is to provide hadronic matrix elements which either test QCD or can be used as inputs to test the Standard Model. In many cases, the lattice numbers have uncertainties which are small enough that they are interesting to experimentalists. I want to give a survey of lattice calculations of matrix elements, and what better way at this summer school, than to recall science which was done here at SLAC, as the framework?

4.1 Generic Matrix Element Calculations

Most of the matrix elements measured on the lattice are expectation values of local operators composed of quark and gluon fields. The mechanical part of the lattice calculation begins by writing down some Green's functions which contain the local operator (call it $J(x)$) and somehow extracting the matrix element. For example, if one wanted $\langle 0|J(x)|h\rangle$, one could look at the two-point function

$$C_{JO}(t) = \sum_x \langle 0|J(x,t)O(0,0)|0\rangle. \tag{74}$$

Inserting a complete set of correctly normalized momentum eigenstates

$$1 = \frac{1}{L^3}\sum_{A,\vec{p}} \frac{|A,\vec{p}\rangle\langle A,\vec{p}|}{2E_A(p)} \tag{75}$$

and using translational invariance and going to large t gives

$$C_{JO}(t) = e^{-m_A t}\frac{\langle 0|J|A\rangle\langle A|O|0\rangle}{2m_A}. \tag{76}$$

A second calculation of

$$C_{OO}(t) = \sum_x \langle 0|O(x,t)O(0,0)|0\rangle = e^{-m_A t}\frac{|\langle 0|O|A\rangle|^2}{2m_A} \tag{77}$$

is needed to extract $\langle 0|J|A\rangle$ by fitting two correlators with three parameters.

Similarly, a matrix element $\langle h|J|h'\rangle$ can be gotten from

$$C_{AB}(t,t') = \sum_x \langle 0|O_A(t)J(x,t')O_B(0)|0\rangle \tag{78}$$

by stretching the source and sink operators O_A and O_B far apart on the lattice, letting the lattice project out the lightest states, and then measuring and dividing out $\langle 0|O_A|h\rangle$ and $\langle 0|O_B|h\rangle$.

These lattice matrix elements are not yet the continuum matrix elements. The lattice is a UV regulator, and changing from the lattice cutoff to a continuum regulator (like $\overline{MS}$) introduces a shift

$$\langle f|O^{cont}(\mu = 1/a)|i\rangle_{\overline{MS}} = a^D(1 + \frac{\alpha_s}{4\pi}(C_{\overline{MS}} - C_{latt}) + \ldots)\langle f|O^{latt}(a)|i\rangle + O(a) + \ldots. \tag{79}$$

The factor a^D converts the dimensionless lattice number to its continuum result. The $O(a)$ corrections arise because the lattice operator might not be the continuum operator: $df/dx = (f(x+a) - f(x))/a + O(a)$. The C's are calculable in perturbation theory, and the "improved perturbation theory" described in the last section is often used to reduce the difference $C_{\overline{MS}} - C_{latt}$.

4.2 Structure Functions

In the beginning, there was deep inelastic scattering. The lattice knows about structure functions through their moments:

$$\int_0^1 dx x^{n-2}F_2(x,Q^2) \equiv M_n(Q^2) = \sum c_n^f v_n^{(f)} \tag{80}$$

has a representation in terms of matrix elements of fairly complicated quark (for the nonsinglet structure function) or gluon bilinears: for quarks

$$\mathcal{O}^{(q)}_{\mu_1\cdots\mu_n} = \left(\frac{i}{2}\right)^{n-1}\bar{q}\gamma_{\mu_1}\overleftrightarrow{D}_{\mu_2}\cdots\overleftrightarrow{D}_{\mu_n}q \tag{81}$$

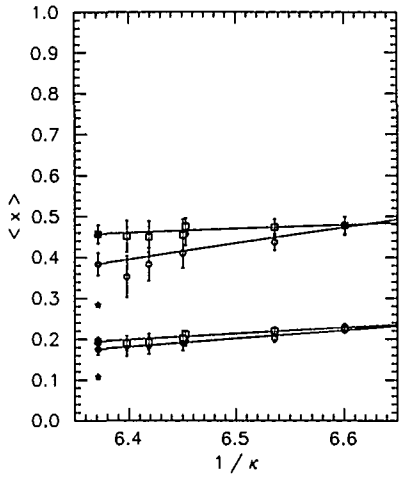


Figure 23: $\langle x\rangle$ for the proton (MOM scheme) from Ref. 35. The circles (boxes) correspond to different choices of lattice operators. The upper (lower) band of data represents the results for the up- (down-) quark distribution.

$$\sum_s \langle p,s|\mathcal{O}^{(f)}_{\{\mu_1\ldots\mu_n\}}|p,s\rangle \simeq v_n^{(f)}\langle p_{\mu_1}\ldots p_{\mu_n} + \ldots\rangle. \tag{82}$$

D is a lattice covariant derivative, which is approximated by a finite difference. The Wilson coefficients $c_{i,n}^{(f)}$ are calculated in perturbation theory and depend on μ^2/Q^2 as well as on the coupling constant $g(\mu)$. The lattice calculation is done by sandwiching the operator in Eq. (78).

It is presently possible to calculate the two lowest moments of the proton structure function on the lattice. Two groups[35,36] presented results at this year's lattice conference. Figure 23 shows $\langle x\rangle$ from Ref. 35, a calculation done in quenched approximation. In this picture, the massless quark limit is the left edge of the picture. There are several different lattice operators which serve as discretizations of the continuum operator, and the figure shows two possibilities.

Unfortunately, the calculation is badly compromised by the quenched approximation. It shows $\langle x\rangle_u = 0.38$ and $\langle x\rangle_d = 0.19$, while in the real world, we expect about 0.28 and 0.10, respectively. In the computer, there are no sea quarks, and their momentum is obviously picked up (at least partly) by the valence quarks.

One nice feature about the lattice calculation is that the spin structure function can be calculated in essentially the same way; the operator $\mathcal{O}$ has an extra

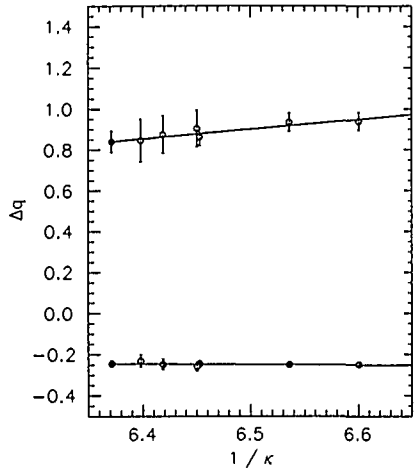


Figure 24: Δu (upper values) and Δd (lower values) for the proton from Ref. 35.

gamma-5 in it. Reference 35 computed $\Delta u = 0.84$ and $\Delta d = -0.24$ (in contrast to 0.92 and -0.34 in the real world). A plot versus quark mass is shown in Fig. 24.

There is no problem in principle which prevents extending these calculations to full QCD (with dynamical sea quarks). It will probably be very expensive to push beyond the lowest moments.

4.3 Heavy Quark Physics

Then there was the November revolution. Twenty years later, systems with one or more heavy quarks remain interesting objects for study. The lattice is no exception. Many groups study spectroscopy, decay amplitudes, form factors, etc., with the goal of confronting both experiment and analytic theoretical models.

There are several ways to study heavy quarks on the lattice. If the quark has infinite mass (the "static limit"), its propagator is very simple: the quark is confined to one spatial location, and as it evolves in time, its color "twinkles." The propagator is just a product of link matrices going forward in time.

One can simulate nonrelativistic quarks directly on the lattice.[31] This has evolved into one of the most successful (and most elaborate) lattice programs.

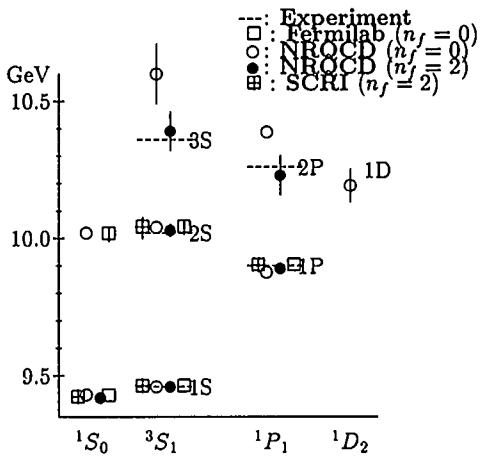


Figure 25: Υ spectrum.

The idea is to write down lattice actions which are organized in an expansion of powers of the quark velocity and to systematically keep all the terms to some desired order. For example, one might write

$$S = \psi^\dagger [iD_t - \frac{D^2}{2m} + \vec{\mu} \cdot \vec{B} + \ldots]\psi \tag{83}$$

including kinetic and magnetic moment terms, suitably (and artistically) discretized. Tadpole-improved perturbation theory is heavily used to set coefficients. Figures 25 and 26 show the Upsilon spectrum and its hyperfine splittings from various NRQCD calculations (from a recent summary by Shigemitsu[37]).

The main shortcoming of nonrelativistic QCD is, of course, that when the quark mass gets small, the nonrelativistic approximation breaks down. For charmonium, $v/c \simeq 0.3$, so the method is less safe for this system than for the Upsilon.

Finally, one can take relativistic lattice quarks and just make the mass heavy. If the quark mass gets too heavy ($ma \simeq 1$), lattice artifacts dominate the calculation. For Wilson fermions, the dispersion relation breaks down: $E(p) \simeq m_1 + p^2/2m_2$ where $m_2 \neq m_1$. The magnetic moment is governed by its own different mass, too.

Another signal of difficulty is that all these formulations have their own pattern of scale violations. That is, nonrelativistic quarks and Wilson quarks approach

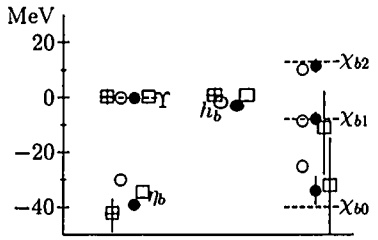


Figure 26: Υ spin splittings: Symbols have the same meaning as in Fig. 25.

their $a \to 0$ limits differently. This is often described in the literature by the statement that "the lattice spacing is different for different observables." For example, in one data set,[38] the inverse lattice spacing (in MeV) is given as 2055 MeV from fitting the heavy $q\bar{q}$ potential, 2140 MeV from the rho mass, 1800 MeV from the proton mass, and 2400 MeV from the Upsilon spectrum. These simulations are just sitting in the middle of figures like Fig. 5 with only one point, trying to guess where the left-hand edge of the picture will be. This is a problem for calculations of B meson and baryon spectroscopy, where the heavy quarks might be treated nonrelativistically and the light quarks are relativistic. What observable should be used to set the overall scale?[39]

One of the major uses of heavy quark systems by the lattice community is to try to calculate the strong coupling constant at $Q^2 = M_Z^2$. This topic deserves its own section.

4.4 $\alpha_s(M_Z)$

Now we are at the SLC and LEP. For some time now, there have been claims that physics at the Z pole hints at a possible breakdown in the Standard Model.[40] A key question in the discussion is whether or not the value of $\alpha_{\overline{MS}}$ inferred from the decay width of the Z is anomalously high relative to other determinations of the strong coupling (which are usually measured at lower Q and run to the Z pole).

The most recent analysis of $\alpha_s(M_Z)$ I am aware of is due to Erler and Langacker.[41] Currently, $\alpha_{\overline{MS}}^{lineshape} = 0.123(4)(2)(1)$ for the Standard Model Higgs mass range, where a first, second, and third uncertainty is from the inputs, Higgs mass, and estimate of α_s^4 terms, respectively. The central Higgs mass is assumed to be 300 GeV, and the second error is for $M_H = 1000$ GeV (+), 60 GeV (-). For the SUSY Higgs mass range (60-150 GeV), one has the lower value $\alpha_{\overline{MS}} = .121(4)(+1-0)(1)$. A global fit to all data gives 0.121(4)(1). Hinchcliffe in the same compilation quotes a global average of 0.118(3).

The lattice can contribute to this question by predicting $\alpha_{\overline{MS}}$ from low-energy physics. The basic idea is simple: The lattice is a (peculiar) UV cutoff. A lattice mass $\mu = Ma$ plus an experimental mass M give a lattice spacing $a = \mu/M$ in fm. If one can measure some quantity related to α_s at a scale $Q \simeq 1/a$, one can then run the coupling constant out to the Z.

The best (recent) lattice number, from Shigemitsu's Lattice '96 summary talk,[37] is

$$\alpha_{\overline{MS}}(M_Z) = 0.1159(19)(13)(19), \tag{84}$$

where the first error includes both statistics and estimates of discretization errors, the second is due to uncertainties from the dynamical quark mass, and the third is from conversions of conventions. The lattice number is about one standard deviation below the pure Z-physics number. Lattice results are compared to other recent determinations of $\alpha_{\overline{MS}}(Z)$ in Fig. 27, a figure provided by P. Burrows.[42]

Two ways of calculating $\alpha_s(M_Z)$ from lattice have been proposed: The first is the "small loop method."[43] This method uses the "improved perturbation theory" described in Chap. 2: One assumes that a version of perturbation theory can describe the behavior of short-distance objects on the lattice: in particular, that the plaquette can be used to define $\alpha_V(q = 3.41/a)$. With typical lattice spacings now in use, this gives the coupling at a momentum $Q_0 = 8 - 10$ GeV. One then converts the coupling to $\alpha_{\overline{MS}}$ and runs out to the Z using the (published) three-loop beta function.[44]

Usually, the lattice spacing is determined from the mass splittings of heavy $Q\bar{Q}$ states. This is done because the mass differences between physical heavy quark states are nearly independent of the quark mass—for example, the S-P mass splitting of the ψ family is about 460 MeV, and it is about 440 MeV for the Υ. A second reason is that the mass splitting is believed to be much less sensitive to sea quark effects than light quark observables, and one can estimate the effects of sea quarks through simple potential models. The uncertainty in the lattice spacing is three to five per cent, but systematic effects are much greater (as we will see below).

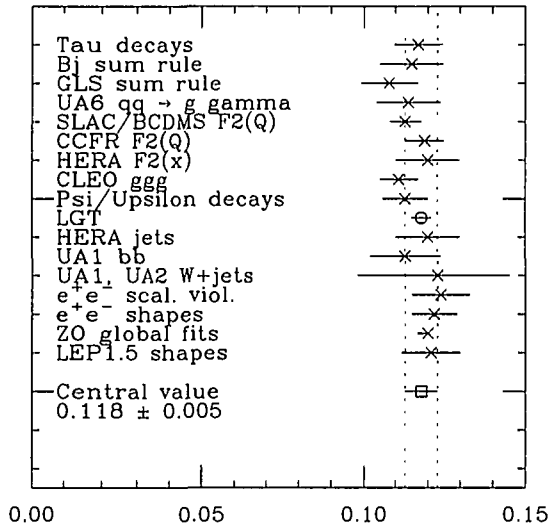


Figure 27: Survey of $\alpha_{\overline{MS}}(M_Z)$ from Ref. 42.

The coupling constant comes from Eq. (63). The plaquette can be measured to exquisite accuracy (0.01 per cent is not atypical) and so the coupling constant is known essentially without error. However, the scale of the coupling is uncertain (due to the lattice spacing).

The next problem is getting from lattice simulations, which are done with $n_f = 0$ (quenched) or $n_f = 2$ (but unphysical sea quark masses) to the real world of $n_f = 3$. Before simulations with dynamical fermions were available, the translation was done by running down in Q to a "typical gluonic scale" for the psi or the upsilon (a few hundred MeV) and then matching the coupling to the three-flavor coupling (in the spirit of effective field theories). This produced a rather low $\alpha_s \simeq 0.105$. Now we have simulations at $n_f = 2$ and can do better. Recall that in lowest order

$$\frac{1}{\alpha_s} = (\frac{11 - \frac{2}{3}n_f}{4\pi}) \ln \frac{Q^2}{\Lambda^2}. \tag{85}$$

One measures $1/\alpha_s$ in two simulations, one quenched, the other at $n_f = 2$, runs one measurement in Q to the Q of the other, then extrapolates $1/\alpha$ linearly in n_f to $n_f = 3$. Then one can convert to $\overline{MS}$ and run away.

Pictures like Fig. 27 are not very useful when one wants to get a feel for the errors inherent in the lattice calculation. Instead, let's run our expectations for $\alpha_s(M_Z)$ down to the scale where the lattice simulations are done, and compare. Figure 28 does that. The squares are the results of simulations of charmed quarks and the octagons are from bottom quarks, both with $n_f = 0$. The crosses and diamond are $n_f = 2$ bottom and charm results. (The bursts show upsilon data when the 1S-2S mass difference gives a lattice spacing.) Note the horizontal error bars on the lattice data. Finally, the predicted $n_f = 3$ coupling α_P is shown in the fancy squares, with error bars now rotated because the convention is to quote an error in α_s. The lower three lines in the picture (from top to bottom) are $\alpha_{\overline{MS}}(M_Z) = 0.118$, 0.123, and 0.128 run down and converted to the lattice prescription.

The two top lines are predictions for how quenched α should run.

Now for the bad news. All of the $n_f = 2$ data shown here were actually run on the same set of configurations. The bare couplings are the same, but the lattice spacings came out different. What is happening is that we are taking calculations at some lattice spacing and inferring continuum numbers from them, but the lattice predictions have scale violations which are different. (The Υ calculations use nonrelativistic quarks, the ψ calculations use heavy Wilson quarks.) Notice also that the bottom and charm quenched lattice spacings are different. This discrepancy is thought to be a failure of the quenched approximation: the characteristic momentum scale for binding in the ψ and Υ are different, and because n_f is not the real world value, α runs incorrectly between the two scales. Said differently, in the quenched approximation, the spectrum of heavy quark bound states is different from the real world.

There is a second method of determining a running coupling constant which actually allows one to see the running over a large range of scales. It goes by the name of the "Schrödinger functional," (referring to the fact that the authors study QCD in a little box with specified boundary conditions) but "coupling determined by varying the box size" would be a more descriptive title. It has been applied to quenched QCD but has not yet been extended to full QCD,[45] and so it has not yet had an impact on phenomenology. This calculation does not use perturbation theory overtly. For a critical comparison of the two methods, see Ref. 46.

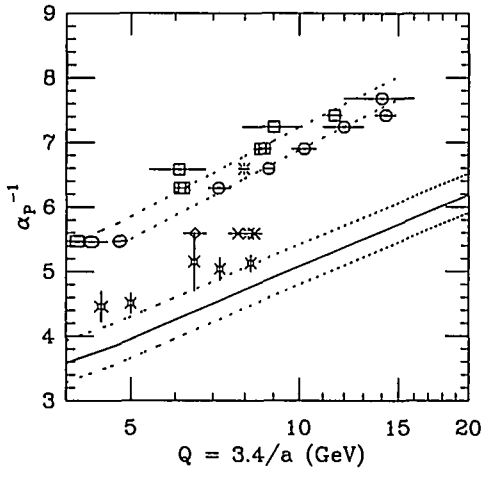


Figure 28: Survey of $\alpha_{\overline{MS}}(Q)$ at the scale where lattice simulations are actually done.

4.5 Glueballs

I have been hearing people talk about glueballs from psi decay for almost 20 years. Toki[47] has summarized the experimental situation for glueballs. What do theorists expect for a spectrum? The problem is that any nonlattice model requires making uncontrolled approximations to get any kind of an answer: there are no obvious zeroth order solutions with small expansion parameters. The lattice is the only game in town for a first-principles calculation.

People have been trying to measure the masses of the lightest glueballs (the scalar and the tensor) using lattice simulations for many years. The problem has proven to be very hard, for several reasons.

Recall how we measure a mass from a correlation function [Eq. (40)]. The problem with the scalar glueball is that the operator O has nonzero vacuum expectation value, and the correlation function approaches a constant at large t:

$$\lim_{t\to\infty} C(t) \to |\langle 0|O|\vec{p}=0\rangle|^2 \exp(-mt) + |\langle 0|O|0\rangle|^2. \tag{86}$$

The statistical fluctuations on $C(t)$ are given by Eq. (25), and we find after a short calculation that

$$\sigma \to \frac{C(0)}{\sqrt{N}}. \tag{87}$$

Thus, the signal-to-noise ratio collapses at large t like $\sqrt{N}\exp(-mt)$ due to the constant term.

A partial cure for this problem is a good trial wave function O. While in principle the plaquette itself could be used, it is so dominated by ultraviolet fluctuations that it does not produce a good signal. Instead, people invent "fat links" which average the gauge field over several lattice spacings, and then make interpolating fields which are closed loops of these fat links. The lattice glueball is a smoke ring.

The second problem is that lattice actions can have phase transitions at strong or intermediate coupling, which have nothing to do with the continuum limit, but mask continuum behavior.[48] As an example of this, consider the gauge group $SU(2)$, where a link matrix can be parameterized as $U = 1\cos\theta + i\vec{\sigma}\cdot\vec{n}\sin\theta$, so $\mathrm{Tr}U = 2\cos\theta$. Now consider a generalization of the Wilson action $-S = \beta\mathrm{Tr}U + \gamma(\mathrm{Tr}U)^2$. (This is a mixed fundamental-adjoint representation action.) At $\gamma\to\infty$, $\mathrm{Tr}U \to \pm 1$ and the gauge symmetry is broken down to $Z(2)$. But $Z(2)$ gauge theories have a first-order phase transition. First-order transitions are stable

under perturbations, and so the phase diagram of this theory, shown in Fig. 29, has a line of first-order transitions which terminate in a second-order point. At the second-order point, some state with scalar quantum numbers becomes massless. However, now imagine that you are doing Monte Carlo along the $\gamma = 0$ line, that is, with the Wilson action. When you come near the critical point, any operator which couples to a scalar particle (like the one you are using to see the scalar glueball) will see the nearby transition and the lightest mass in the scalar channel will shrink. Once you are past the point of closest approach, the mass will rise again. Any scaling test which ignores the nearby singularity will lie to you.

This scenario has been mapped out for $SU(3)$, and the place of closest approach is at a Wilson coupling corresponding to a lattice spacing of 0.2 fm or so, meaning that very small lattice spacings are needed before one can extrapolate to zero lattice spacing. A summary of the situation is shown in Fig. 30 (Ref. 49). Here the quantity r_0 is the "Sommer radius,"[50] defined through the heavy quark force, $F(r) = -dV(r)/dr$, by $r_0^2 F(r_0) = -1.65$. In the physical world of three colors and four flavors, $r_0 = 0.5$ fm.

Finally, other arguments suggest that a small lattice spacing or a good approximation to an action on an RT are needed to for glueballs: the physical diameter of the glueball, as inferred from the size of the best interpolating field, is small, about 0.5 fm. Shäfer and Shuryak[51] have argued that the small size is due to instanton effects. Most lattice actions do bad things to instantons at large lattice spacing.[28]

Two big simulations have carried calculations of the glueball mass close to the continuum limit: the UKQCD Collaboration[52] and a collaboration at IBM which built its own computer.[53] (The latter group is the one with the press release last December announcing the discovery of the glueball.) Their predictions in MeV are different, and they each favor a different experimental candidate for the scalar glueball (the one which is closer to their prediction, of course). It is a useful object lesson because both groups say that their lattice numbers agree before extrapolation, but they extrapolate differently to $a = 0$.

The UKQCD group sees that the ratio $m(0^{++})/\sqrt{\sigma}$ can be well-fitted with a form $b + ca^2\sigma$ (σ is the string tension) and a fit of this form to the lattice data of both groups gives $m(0^{++})/\sqrt{\sigma} = 3.64 \pm 0.15$. To turn this into MeV, we need σ in MeV units. One way is to take $m_\rho/\sqrt{\sigma}$ and extrapolate that to $a = 0$ using $b + ca\sqrt{\sigma}$. Averaging and putting 770 MeV for m_ρ, one gets $\sqrt{\sigma} = 432 \pm 15$ MeV,

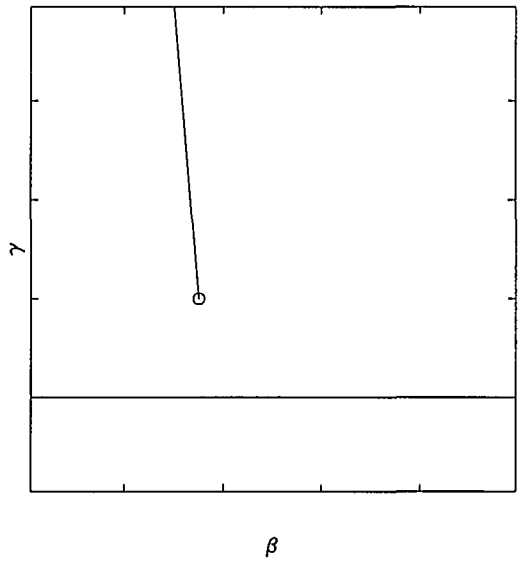


Figure 29: Phase transitions in the fundamental-adjoint plane.

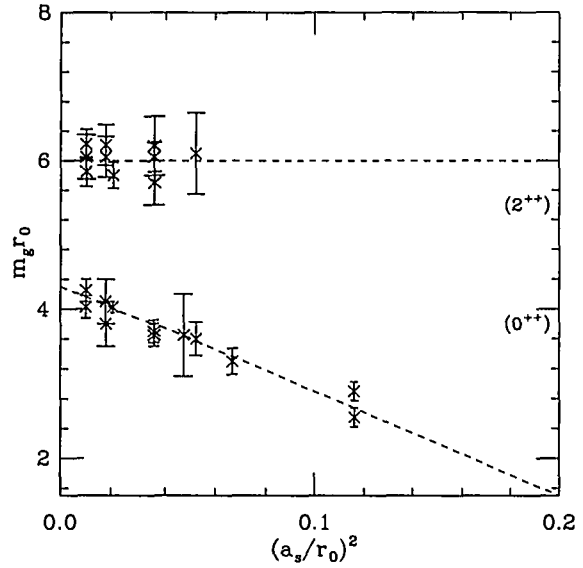


Figure 30: Glueball mass vs r_0 with the Wilson action, from a summary picture in Ref. 49.

which is consistent with the usual estimate (from extracting the string tension from the heavy quark potential) of about 440 MeV. Using the total average, they get $m(0^{++}) = 1572 \pm 65 \pm 55$ MeV where the first error is statistical and the second comes from the scale.

The IBM group, on the other hand, notices that $m_\rho a$ and $m_\phi a$ scale asymptotically, and uses the phi mass to predict quenched $\Lambda_{\overline{MS}}$, then extrapolates $m(0^{++})/\Lambda = A + B(a\Lambda)^2$. They get 1740(41) MeV from their data; when they analyze UKQCD data, they get 1625(94) MeV; and when they combine the data sets, they get 1707(64) MeV.

A neutral reporter could get hurt here. It seems to me that the lattice prediction for the scalar glueball is 1600 ± 100 MeV, and that there are two experimental candidates for it, the $f_0(1500)$ and the $f_J(1710)$.

Masses are not the end of the story. The IBM group has done two interesting recent calculations related to glueballs, which strengthen their claim that the $f_J(1710)$ is the glueball.

The first one of them[54] was actually responsible for the press release. It is a calculation of the decay width of the glueball into pairs of pseudoscalars. This is done by computing an unamputated three-point function on the lattice, with an assumed form for the vertex, whose magnitude is fitted. The result is shown in Fig. 31. The octagons are the results of the simulation, and the diamonds show interpolations in the quark mass. The "experimental" points (squares) are from a partial wave analysis of isoscalar scalar resonances by Longacre and Lindenbaum.[55]

The response of a member of the other side is that the slope of the straight line that one would put through the three experimental points is barely, if at all, compatible with the slope of the theoretical points. Since they argue theoretically for a straight line, the comparison of such slopes is a valid one.

If one of the experimental states is not a glueball, it is likely to be a $^3P_0$ orbital excitation of quarks. Weingarten and Lee[56] are computing the mass of this state on the lattice and argue that it is lighter than 1700 MeV; in their picture, the $f_0(1500)$ is an $s\bar{s}$ state. I have now said more than I know and will just refer you to recent discussions of the question.[57]

Both groups predict that the 2^{++} glueball is at about 2300 MeV.

Can "improved actions" help the situation? Recently, Peardon and Morningstar[49] implemented a clever method for beating the exponential signal-to-noise ratio: make the lattice spacing smaller in the time direction than in the space

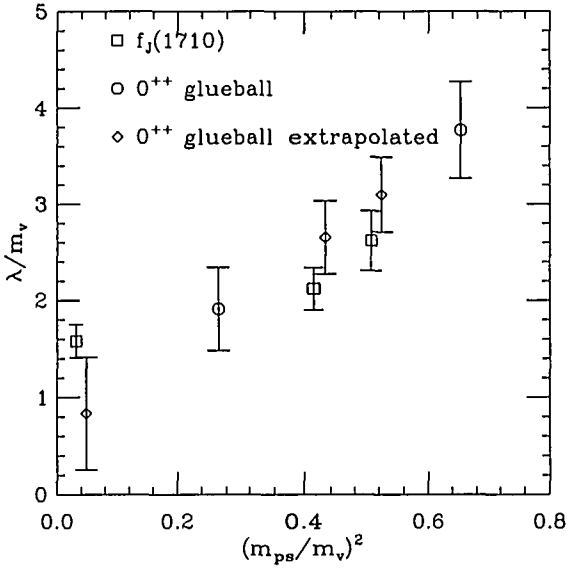


Figure 31: Scalar glueball decay couplings from Ref. 54.

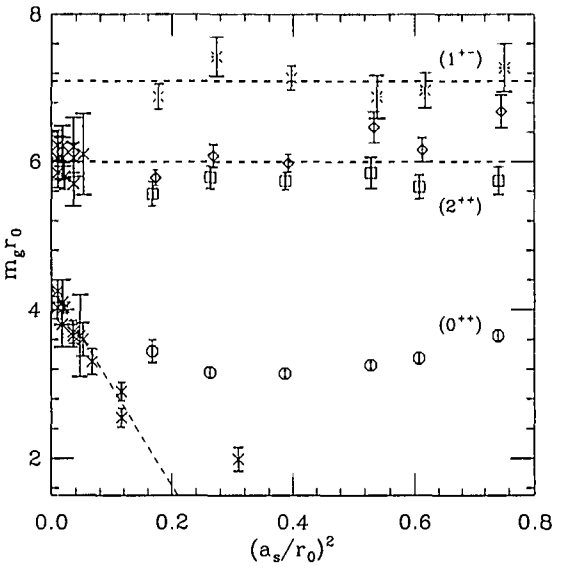


Figure 32: Glueball mass vs r_0 from Ref. 49, including large lattice spacing data.

direction. Then the signal, which falls like $\exp(-ma_t L_t)$ after L_t lattice spacings, dies more slowly because a_t is reduced. Their picture of the glueball mass versus r_0 is shown in Fig. 32. They are using the tadpole-improved Lüscher-Weisz action. The pessimist notes the prominent dip in the middle of the curve; this action also has a lattice-artifact transition (somewhere); the optimist notes that the dip is much smaller than for the Wilson action, and then the pessimist notes that there is no Wilson action data at large lattice spacing to compare. I think the jury is still out.

4.6 The B Parameter of the B Meson

And finally we are at BaBar. $\bar{B} - B$ mixing is parameterized by the ratio

$$x_d = \frac{(\Delta M)_{b\bar{d}}}{\Gamma_{b\bar{d}}} = \tau_{b\bar{d}} \frac{G_F^2}{6\pi^2} \eta_{QCD} F(\frac{m_t^2}{m_W^2}) |V_{tb}^* V_{td}|^2 b(\mu) \{\frac{3}{8} \langle \bar{B}|\bar{b}\gamma_\rho(1-\gamma_5)d\bar{b}\gamma_\rho(1-\gamma_5)d|B\rangle\}. \quad (88)$$

Experiment is on the left, theory on the right. Moving into the long equation from the left, we see many known (more or less) parameters from phase space integrals

or perturbative QCD calculations, then a combination of CKM matrix elements, followed by a four-quark hadronic matrix element.[58] We would like to extract the CKM matrix element from the measurement of x_d (and its strange partner x_s). To do so, we need to know the value of the object in the curly brackets, defined as $3/8 M_{bd}$ and parameterized as $m_B^2 f_{B_d}^2 B_{b_d}$ where B_{b_d} is the so-called B parameter, and f_B is the B-meson decay constant

$$\langle 0|\bar{b}\gamma_0\gamma_5 d|B\rangle = f_B m_B. \tag{89}$$

Naive theory, which is expected to work well for the B system, suggests that $B_B = 1$ to good accuracy. Of course, the stakes are high and a good determination of M_{bd} is needed to test the Standard Model. The lattice can do just that.

In Eq. (88) $b(\mu)$, the coefficient which runs the effective interaction down from the W-boson scale to the QCD scale μ, and the matrix element $M(\mu)$ both depend on the QCD scale, and one often sees the renormalization group invariant quantities $\hat{M}_{bd} = b(\mu)M_{bd}(\mu)$ or $\hat{B}_{bd} = b(\mu)B_{bd}(\mu)$ quoted in the literature.

Decay constants probe very simple properties of the wave function: in the nonrelativistic quark model,

$$f_M = \frac{\psi(0)}{\sqrt{m_M}}, \tag{90}$$

where $\psi(0)$ is the $\bar{q}q$ wave function at the origin. For a heavy-quark (Q) light-quark (q) system, $\psi(0)$ should become independent of the heavy quark's mass as the Q mass goes to infinity, and in that limit, one can show in QCD that $\sqrt{m_M} f_M$ approaches a constant.

One way to compute the decay constant is to put a light quark and a heavy quark on the lattice and let them propagate. It is difficult to calculate f_B directly on present day lattices with relativistic lattice fermions because the lattice spacing is much greater than the b quark's Compton wavelength (or the UV cutoff is below m_b). In this limit, the b quark is strongly affected by lattice artifacts as it propagates. However, one can make m_b infinite on the lattice and determine the combination $\sqrt{m_B} f_B$ in the limit. Then one can extrapolate down to the B mass and see if the two extrapolations up and down give the same result. (Nonrelativistic b quarks can solve this problem in principle, but the problem of setting the lattice spacing between light and nonrelativistic quarks has prevented workers from quoting a useful decay constant from these simulations.)

Among the many lattice decay constant calculations, the one of Ref. 59 stands out in my mind for being the most complete. These authors did careful quenched

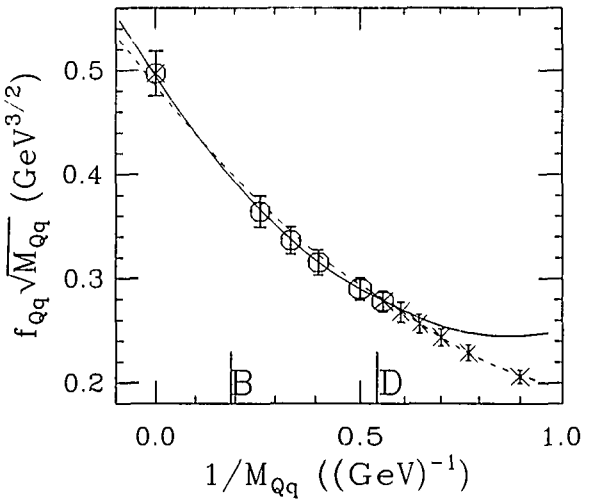


Figure 33: Pseudoscalar meson decay constant vs $1/M$, from Ref. 59.

simulations at many values of the lattice spacing, which allows one to extrapolate to the continuum limit by brute force. They have also done a less complete set of simulations which include light dynamical quarks, which should give some idea of the accuracy of the quenched approximation.

The analysis of all this data is quite involved. One begins with a set of lattice decay constants measured in lattice units, from simulations done with heavy quarks which are probably too light and light quarks which are certainly too heavy. One has to interpolate or extrapolate the heavy quark masses to their real world values, extrapolate the light quarks down in mass to their physical values, and finally try to extrapolate to $a \to 0$. It is not always obvious how to do this. Complicating everything are the lattice artifacts in the fermion and gauge actions, and the lattice-to-continuum renormalization factors as in Eq. (79).

The (still preliminary) results of Ref. 59 are shown in Figs. 33 and 34. The $N_f = 2$ dynamical fermion data in Fig. 34 have moved around a bit in the past year and may not have settled down yet.

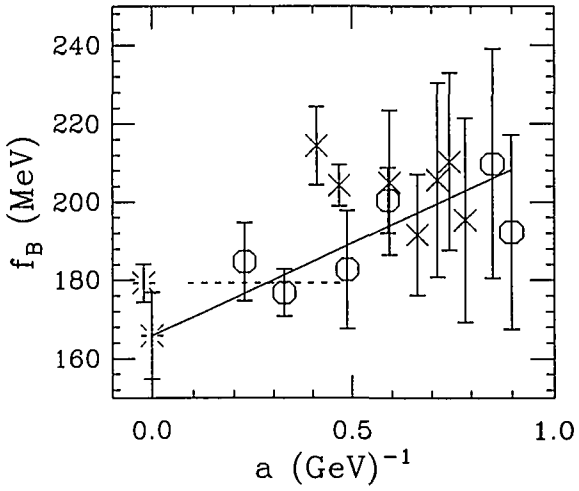


Figure 34: f_B vs a from Ref. 59. Octagons are quenched data; crosses, $N_F = 2$. The solid line is a linear fit to all quenched points; the dashed line is a constant fit to the three quenched points with $a < 0.5$ GeV^{-1}. The extrapolated values at $a = 0$ are indicated by bursts. The scale is set by $f_\pi = 132$ MeV throughout.

The numerical results of Ref. 59 are:

$$f_B = 166(11)(28)(14) \quad f_D = 196(9)(14)(8)$$
$$f_{B_s} = 181(10)(36)(18) \quad f_{D_s} = 211(7)(25)(11)$$
$$\frac{f_{B_s}}{f_B} = 1.10(2)(5)(8) \quad \frac{f_{D_s}}{f_D} = 1.09(2)(5)(5),$$

where the first error includes statistical errors and systematic effects of changing fitting ranges; the second, other errors within the quenched approximation; the third, an estimate of the quenching error. Decay constants are in MeV.

Note that the error bars for the B system are small enough to be phenomenologically interesting. The Particle Data Group's[60] determination of CKM matrix elements, which does not include this data, says, "Using $\hat{B}_{B_d} f_{B_d}^2 = (1.2 \pm 0.2)(173 \pm 40$ MeV$)^2$..., $|V_{tb}^* V_{td}|^2 = 0.009 \pm 0.003$, where the error bar comes primarily from the theoretical uncertainty in the hadronic matrix elements."

Can we trust these numbers? Lattice calculations have been predicting $f_{D_s} \simeq 200$ MeV for about eight years. The central values have changed very little, while the uncertainties have decreased. So far, four experiments have reported measurements of this quantity. The most recent is the Fermilab E653 Collaboration[61] $f_{D_s} = 194(35)(20)(14)$ MeV. The older numbers, with bigger errors, were 238(47)(21)(43) from WA75 (1993) (Ref. 62); 344(37)(52)(42) from CLEO (1994) (Ref. 63); 430 (+150 −130)(40) from BES (1995) (Ref. 64).

Now back to the mixing problem. On the lattice, one could measure the decay constants and B parameter separately and combine them after extrapolation, or measure M directly and extrapolate it. In principle, the numbers should be the same, but in practice, they will not be.

A recent calculation illustrates this point.[65] The authors computed the four fermion operators M_{bs} and M_{bd} directly on the lattice. Figure 35 shows the behavior of M as a function of hadron mass at one of their lattice spacings.

The ratio $r_{sd} = M_{bs}/M_{bd}$ is presumably much less sensitive to lattice spacing or to quark mass extrapolation. The authors' result for the lattice spacing dependence of this ratio is shown in Fig. 36, along with an extrapolation to zero lattice spacing. They find $1.54 \pm .13 \pm .32$ from their direct method, compared to $r_{sd} \approx 1.32 \pm .23$ from separate extrapolations of the decay constants and the B parameter. (They measure equal B parameters for strange and nonstrange B mesons, $B(\mu) = 1.02(13)$ for $\mu = 2$ GeV.)

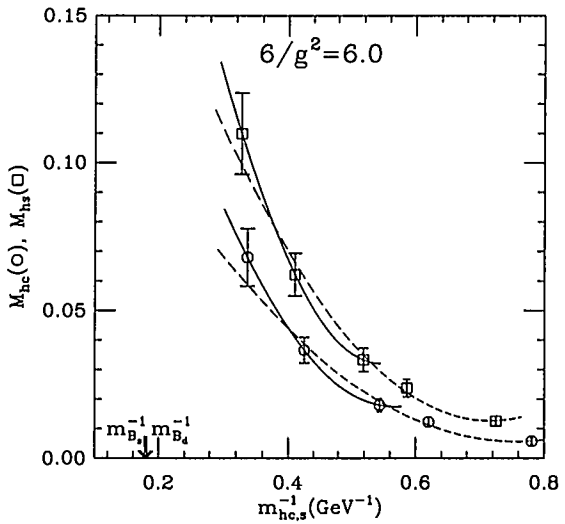


Figure 35: M_{hc} (octagons) and M_{hs} (squares) as a function of the inverse heavy-down (strange) meson mass, at $\beta = 6.0$. The dashed line shows the effect of the lightest points on the fit.

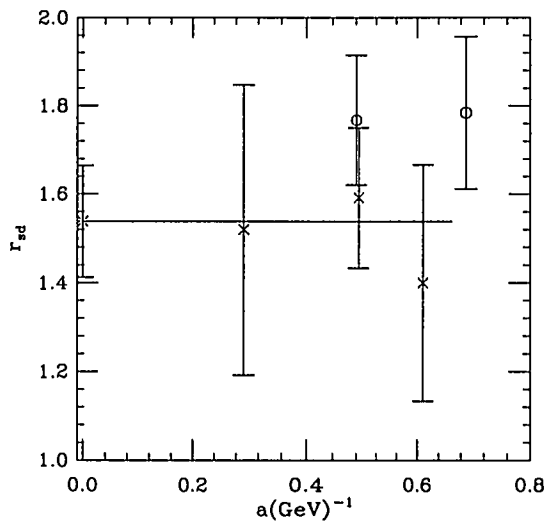


Figure 36: The SU(3) flavor breaking ratio M_{bs}/M_{bd} vs the lattice spacing a. The points denoted by crosses were used in the fit (solid line). The burst shows the extrapolation to $a = 0$.

It looks like $SU(3)$ breaking is fairly large, and if that is so, it looks like the parameter x_s, the strange analog of Eq. (88), might be about 20, unmeasurably large.[66]

5 Conclusions

Lattice methods have arrived. There are so many lattice calculations of different matrix elements that it is impossible to describe them all, and in many cases, the quality of the results is very high. One can see plots showing extrapolations in lattice spacing which show that the control of lattice spacing has become good enough to make continuum predictions with small uncertainties. Calculations with dynamical fermions and a small lattice spacing are still nearly impossibly expensive to perform, and "quenching" remains the dominant unknown in all lattice matrix element calculations.

There are two major tasks facing lattice experts. I believe that all the people in our field would agree that the first problem is to reduce the computational burden, so that we can do more realistic simulations with smaller computer resources. I have illustrated several of the approaches people are using to attack this problem. I believe that some of them have been shown to be successful and that "improvement" will continue to improve.

There is a second question for lattice people, which I have not discussed, but I will mention at the end: Is there a continuum phenomenology of light hadron structure or confinement, which can be justified from lattice simulations? The motivation for asking this question is that there are many processes which cannot be easily addressed via the lattice, but for which a QCD prediction ought to exist. For examples of such questions, see the talk of Bjorken in this conference,[67] or Shuryak's article.[11] Few lattice people are thinking about this question. Part of the lattice community spends its time looking for "structure" in Monte Carlo-generated configurations of gauge fields: instantons, monopoles, etc. This effort is not part of the mainstream because the techniques either involve gauge fixing (and so it is not clear whether what is being seen is just an artifact of a particular gauge), or they involve arbitrary decisions during the search (perform a certain number of processing operations, no more, no less). To answer this question requires new ideas and a controlled approach to simulation data. Will an answer be found?

Acknowledgments

I would like to thank M. Alford, C. Bernard, T. Blum, P. Burrows, S. Gottlieb, A. Hasenfratz, P. Hasenfratz, U. Heller, P. Langacker, P. Lepage, P. Mackenzie, J. Negele, F. Niedermayer, J. Shigemitsu, J. Simone, R. Sommer, R. Sugar, M. Teper, D. Toussaint, D. Weingarten, U. Wiese, and M. Wingate for discussions, figures, and correspondence. I would also like to thank Brian Greene, the organizer of the 1996 TASI Summer School for asking me to lecture there: it means that I had already written most of the material for these lectures well in advance of the deadline! I would also like to thank the Institute for Theoretical Physics at the University of Bern for its hospitality, where these lectures were written. Finally, it was a pleasure to be able to visit SLAC again, talk to many of my old friends, and present my subject to a wider audience. This work was supported by the U. S. Department of Energy.

References

[1] Some standard reviews of lattice gauge theory are M. Creutz, *Quarks, Strings, and Lattices* (Cambridge, 1983); *Quantum Fields on the Computer,* (World Scientific, 1992) edited by M. Creutz; I. Montvay and G. Munster, *Quantum Fields on a Lattice* (Cambridge, 1994). The lattice community has a large annual meeting, and the proceedings of those meetings (Lattice XX, published so far by North Holland) are the best places to find the most recent results.

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

[3] M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rev. Lett. **42**, 1390 (1979).

[4] M. Golterman and J. Smit, Nucl. Phys. B **255**, 328 (1985).

[5] G. P. Lepage, in *From Actions to Answers–Proceedings of the 1989 Theoretical Advanced Summer Institute in Particle Physics*, edited by T. DeGrand and D. Toussaint (World Scientific, 1990).

[6] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[7] F. Brown and T. Woch, Phys. Rev. Lett. **58**, 2394 (1987); M. Creutz, Phys. Rev. D **36**, 55 (1987). For a review, see S. Adler in the *Proceedings of the 1988 Symposium on Lattice Field Theory*, edited by A. Kronfeld and P. Mackenzie [Nucl. Phys. B (Proc. Suppl.) 9 (1989)].

[8] H. C. Andersen, J. Chem. Phys. **72**, 2384 (1980); S. Duane, Nucl. Phys. B **257**, 652 (1985); S. Duane and J. Kogut, Phys. Rev. Lett. **55**, 2774 (1985); S. Gottlieb, W. Liu, D. Toussaint, R. Renken, and R. Sugar, Phys. Rev. D **35**, 2531 (1987).

[9] S. Duane, A. Kennedy, B. Pendleton, and D. Roweth, Phys. Lett. B **194**, 271 (1987).

[10] For a recent review, see A. Frommer, talk presented at Lattice '96, hep-lat/9608074.

[11] E. Shuryak, "Why is a nucleon bound?" hep-ph/9603354, to appear in Festschrift for Gerry Brown's 70.

[12] The MILC Collaboration, presented at Lattice '96, hep-lat/9608012.

[13] F. Butler *et al.*, Phys. Rev. Lett. **70**, 2849 (1993).

[14] This plot was kindly provided by R. Sommer.

[15] K. Symanzik, in *Recent Developments in Gauge Theories*, edited by G. 't Hooft *et al.* (Plenum, New York, 1980), p. 313; in *Mathematical Problems in Theoretical Physics*, edited by R. Schrader *et al.* (Springer, New York, 1982); Nucl. Phys. B **226**, 187, 205 (1983).

[16] P. Weisz , Nucl. Phys. B **212**, 1 (1983); M. Lüscher and P. Weisz, Comm. Math. Phys. **97**, 59 (1985).

[17] M. Lüscher and P. Weisz, Phys. Lett. B **158**, 250 (1985).

[18] B. Sheikholeslami and R. Wohlert, Nucl. Phys. B **259**, 572 (1985).

[19] M. Alford, W. Dimm, G. P. Lepage, G. Hockney, and P. Mackenzie, Phys. Lett. B **361**, 87 (1995); G. P. Lepage, to appear in the *Proceedings of Lattice '95*, and his Schladming lectures, hep-lat/9607076.

[20] K. Jansen *et al.*, Phys. Lett. B **372**, 275 (1996) (hep-lat/9512009); M. Lüscher *et al.*, hep-lat/9605038, hep-lat/9608049.

[21] C. Morningstar, to appear in the *Proceedings of Lattice '95*, hep-lat/9509073.

[22] G. P. Lepage and P. Mackenzie, Phys. Rev. D **48**, 2250 (1993).

[23] S. Brodsky, G. P. Lepage, and P. Mackenzie, Phys. Rev. D **28**, 228 (1983).

[24] G. Parisi, in *High Energy Physics–1980, XX Int. Conf.*, edited by L. Durand and L. G. Pondrom (AIP, New York, 1981).

[25] The classic reference is K. Wilson and J. Kogut, Phys. Repts. **12** C, 77 (1974).

[26] P. Hasenfratz and F. Niedermayer, Nucl. Phys. B **414**, 785 (1994); P. Hasenfratz, Nucl. Phys. B (Proc. Suppl.) **34**, 3 (1994); F. Niedermayer, *ibid.*, 513; A. Farchioni, P. Hasenfratz, F. Niedermayer, and A. Papa, Nucl. Phys. B **454**, 638 (1995).

[27] T. DeGrand, A. Hasenfratz, P. Hasenfratz, and F. Niedermayer, Nucl. Phys. B **454**, 587 (1995); Nucl. Phys. B **454**, 615 (1995).

[28] T. DeGrand, A. Hasenfratz, and D. Zhu, COLO-HEP-369, hep-lat/9603015.

[29] U.-J. Wiese, Phys. Lett. B **315**, 417 (1993); W. Bietenholz, E. Focht, and U.-J. Wiese, Nucl. Phys. B **436**, 385 (1995); W. Bietenholz and U.-J. Wiese, MIT preprint, CTP 2423 (1995).

[30] D. Bliss *et al.*, hep-lat/9605041.

[31] C. Davies *et al.*, Phys. Rev. D **52**, 6519 (1995); Phys. Lett. B **345**, 42 (1995); Phys. Rev. D **50**, 6963 (1994); Phys. Rev. Lett. **73**, 2654 (1994); and G. P. Lepage *et al.*, Phys. Rev. D **46**, 4052 (1992).

[32] M. Alford, T. Klassen, and G. P. Lepage, presentations at Lattice '95, hep-lat/9509087, and in Lattice '96, hep-lat/9608113.

[33] UKQCD Collaboration, presentation at Lattice '96, hep-lat/9608034.

[34] S. Collins, R. Edwards, U. Heller, and J. Sloan, presentation at Lattice '96, hep-lat/9608021.

[35] M. Göckeler *et al.*, talk presented at Lattice '95, hep-lat/9608046, plus references therein.

[36] R. C. Brower *et al.*, talk presented at Lattice '95, hep-lat/9608069.

[37] J. Shigemitsu, talk at Lattice '96, hep-lat/9608058.

[38] Compare M. Wingate *et al.*, Nucl. Phys. B (Proc. Suppl.) **42**, 373 (1995).

[39] See S. Collins *et al.*, preprint FSU-SCRI-96-43, hep-lat/9607004.

[40] For one example of this idea, see M. Shifman, Int. J. Mod. Phys. A **11**, 3195 (1996), hep-ph/9511469.

[41] L. Montonet *et al.*, Phys. Rev. D **50**, 1173 (1994) and 1995 off-year partial update for the 1995 edition available on the PDG WWW pages (URL: http://pdg.lbl.gov/).

[42] P. Burrows, in these proceedings.

[43] A. El-Khadra, G. Hockney, A. Kronfeld, and P. Mackenzie, Phys. Rev. Lett. **69**, 729 (1992).

[44] G. Rodrigo and A. Santamaria, Phys. Lett. B **313**, 441 (1993).

[45] M. Lüscher, P. Weisz, and U. Wolff, Nucl. Phys. B **359**, 221 (1991); M. Lüscher, R. Narayanan, P. Weisz, and U. Wolff, Nucl. Phys. B **384**, 168 (1992); M. Lüscher, R. Sommer, U. Wolff, and P. Weisz, Nucl. Phys. B **389**, 247 (1993); M. Lüscher, R. Sommer, P. Weisz, and U. Wolff, Nucl. Phys. B **413**, 491 (1994).

[46] P. Weisz, in the *Proceedings of Lattice '95*, hep-lat/9511017.

[47] For a review of the experimental situation, see W. Toki, in these proceedings.

[48] G. Bhanot and M. Creutz, Phys. Rev. D **24**, 3212 (1981); for recent work, see U. Heller, Phys. Lett. B **362**, 123 (1995) (hep-lat/9508009).

[49] C. Morningstar and M. Peardon, contribution to Lattice '96, hep-lat/9608050.

[50] R. Sommer, Nucl. Phys. B **411**, 839 (1994).

[51] T. Schäfer and E. Shuryak, Phys. Rev. Lett. **75**, 1707 (1995).

[52] G. Bali *et al.*, Phys. Lett. B **309**, 378 (1993).

[53] H. Chen, J. Sexton, A. Vaccarino, and D. Weingarten, Nucl. Phys. B (Proc. Suppl.) **34**, 347 (1994).

[54] J. Sexton, A. Vaccarino, and D. Weingarten, Phys. Rev. Lett. **75**, 4563 (1995).

[55] S. Lindenbaum and R. Longacre, Phys. Lett. B **274**, 492 (1992).

[56] W. Lee and D. Weingarten, contribution to Lattice '96, hep-lat/9608071.

[57] For the UKQCD/$f_0(1500)$ picture, see F. Close and M. Teper, RAL-96-040, contribution to the 1996 International Conference on High Energy Physics, C. Amsler and F. Close, Phys. Lett. B **353**, 385 (1995) and Phys. Rev. D **53**, 295 (1996); while for the IBM/$f_J(1710)$ picture, see D. Weingarten, contribution to Lattice '96, hep-lat/9608070.

[58] For a nice derivation of this formula, see the lectures of J. Rosner in the *Proceedings of the 1990 Theoretical Advanced Study Institute*, edited by M. Cvetič and P. Langacker (World Scientific, 1991).

[59] C. Bernard and the MILC Collaboration, presentation at Lattice '96, hep-lat/9608092.

[60] R. M. Barnett *et al.* , Phys. Rev. D **54**, 1 (1996).

[61] K. Kodama *et al.*, DPNU-96-33, hep-ex/9696017.

[62] S. Aoki *et al.*, Prog. Theor. Phys. **89**, 1 (1993).

[63] D. Acosta *et al.*, Phys. Rev. D **49**, 5690 (1994).

[64] J. Z. Bai *et al.*, Phys. Rev. Lett. **74**, 4599 (1995).

[65] C. Bernard, T. Blum, and A. Soni, presentation at Lattice '96, hep-lat/9609005.

[66] J. Flynn, talk presented at Lattice '96, hep-lat/9610010.

[67] J. D. Bjorken, in these proceedings.