

CONF-9707193--

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36.

TITLE: INTRODUCTION TO LATTICE QCD

AUTHOR(S): Rajan Gupta, T-8

SUBMITTED TO: Proceedings of 1997 Les Houches School

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

By acceptance of this article, the publisher recognizes that the U. S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form this contribution, or to allow others to do so, for U.. S. Government purposes.

The Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.. S. Department of Energy.

Los Alamos

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

FORM NO. 836 R4
ST. NO. 2629 5/81

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED



DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

COURSE

INTRODUCTION TO LATTICE QCD

Rajan Gupta

*Theoretical Division, Group T-8, Mail Stop B-285,
Los Alamos National Laboratory
Los Alamos, NM 87545, U. S. A
E-mail: rajan@qcd.lanl.gov*

Photograph of Lecturer

Contents

1. Introduction	5
2. Standard Model of Particle Interactions	7
3. Scope of Lattice QCD Calculations	10
4. Overview of the Lattice Approach	14
5. Connection between Minkowski and Euclidean Field Theory and with Statistical Mechanics	21
6. Formulation of Lattice Gauge Theories	25
6.1. Discrete space-time grid	25
6.2. Lattice transcription of field variables $\psi(x)$ and $A_\mu(x)$:	26
6.3. Discrete Symmetries of the lattice theory	27
6.4. Local Gauge Symmetry	27
7. Simplest Formulation of the Lattice Action	30
7.1. Gauge Action	30
7.2. Fermion Action	31
7.3. The Haar Measure	34
7.4. Gauge Fixing and Gribov Copies	35
8. Continuum limit and Symanzik's Improvement program	37
9. Fermion Doubling Problem in the "naive" action	41
9.1. Generalities of Chiral Symmetry	43
10. Wilson Fermions (WF)	46
10.1. Properties of Wilson Fermions	48
11. Staggered fermions (SF)	53
11.1. Summary	57
12. Improved gauge actions	58
12.1. Lüscher-Weisz Action	58
12.2. Tadpole/Mean-field Improvement	61
12.3. The Renormalization Group and the Renormalized Trajectory	65
12.4. Iwasaki Action	70
12.5. $b = \sqrt{3}$ blocking transformation	71
12.6. $b = 2$ blocking transformation	72
12.7. Classically Perfect Action	72
12.8. Summary	73
13. $O(a)$ improvement in the Dirac Action	74
13.1. The Sheikholeslami-Wohlert (clover) Action	75
13.2. D234c action	76
13.3. NRQCD action	77
13.4. Perfect action	78
13.5. Fermilab action	79
13.6. Summary	79
14. Confinement and Asymptotic Freedom	80

14.1.Asymptotic Freedom	80
14.2.Confinement	84
14.2.1.Wilson Loops	84
14.2.2.Strong Coupling Expansion	88
14.2.3.Wilson/Polyakov Line	91
15.Phase Transitions in the Lattice Theory	93
16.Errors in Lattice Results	98
16.1.Statistical errors	98
16.2.Finite Volume	99
16.3.Finite lattice spacing errors	100
16.4.Chiral Extrapolations in the light quark masses	100
16.5.Discretization of Heavy Quarks	101
16.6. Matching between lattice and continuum scheme (renormalization constants)	102
16.7.Operator mixing	102
16.8.Quenched Approximation (QQCD)	103
17.Lattice Correlators	108
17.1.Two-point meson correlators	110
17.2.Interpolating operators	111
17.3.Three-point correlators	113
17.4.CPS Symmetry	114
18.Lattice Calculations of the Hadron Spectrum	115
18.1.Status of the Light Meson and Baryon Spectrum	118
18.2.Decay Constants	127
18.3.Lattice Calculations of the Glueball Spectrum	128
18.4.Improved action Results on Anisotropic lattices	129
18.5.Mixing between scalar quarkonia and glueballs	131
18.6.Mass Inequalities	133
19.The strong coupling constant α_s	133
20.Quark Masses	136
20.1.Light Quark Masses from Ward Identities	139
20.2.Schrodinger Functional Approach	141
20.3.Role of quark masses in CP violation	141
20.4.Heavy quark masses m_c and m_b	142
References	145

1. Introduction

The goal of the lectures on lattice QCD (LQCD) is to provide an overview of both the technical issues and the progress made so far in obtaining phenomenologically useful numbers. The lectures consist of three parts. My charter is to provide an introduction to LQCD and outline the scope of LQCD calculations. In the second set of lectures, Guido Martinelli will discuss the progress we have made so far in obtaining results, and their impact on Standard Model phenomenology. Finally, Martin Lüscher will discuss the topical subjects of chiral symmetry, improved formulation of lattice QCD, and the impact these improvements will have on the quality of results expected from the next generation of simulations.

QCD is the regnant theory of strong interactions. It is formulated in terms of quarks and gluons which we believe are the basic degrees of freedom that make up hadronic matter. It has been very successful in predicting phenomena involving large momentum transfer. In this regime the coupling constant is small and perturbation theory becomes a reliable tool. On the other hand, at the scale of the hadronic world, $\mu \lesssim 1$ GeV, the coupling constant is of order unity and perturbative methods fail. In this domain lattice QCD provides a non-perturbative tool for calculating the hadronic spectrum and the matrix elements of any operator within these hadronic states from first principles. LQCD can also be used to address issues like the mechanism for confinement and chiral symmetry breaking, the role of topology, and the equilibrium properties of QCD at finite temperature. Unfortunately, these latter topics were not covered at this school, so I will give appropriate references to compensate for their omission.

Lattice QCD is QCD formulated on a discrete Euclidean space time grid. Since no new parameters or field variables are introduced in this discretization, LQCD retains the fundamental character of QCD. Lattice QCD can serve two purposes. First, the discrete space-time lattice acts as a non-perturbative regularization scheme. At finite values of the lattice spacing a , which provides an ultraviolet cutoff at π/a , there are no infinities. Furthermore, renormalized physical quantities have a finite well behaved limit as $a \rightarrow 0$. Thus, in principle, one could do all the standard perturbative calculations using lattice regularization, however, these calculations are far

more complicated and have no advantage over those done in a continuum scheme. Second, the pre-eminent use of transcribing QCD on to a space-time lattice is that LQCD can be simulated on the computer using methods analogous to those used for Statistical Mechanics systems. (A brief review of the connection between Euclidean field theory and Statistical Mechanics is given in Section 5.) These simulations allow us to calculate correlation functions of hadronic operators and matrix elements of any operator between hadronic states in terms of the fundamental quark and gluon degrees of freedom.

The only tunable input parameters in these simulations are the strong coupling constant and the bare masses of the quarks. Our belief is that these parameters are prescribed by some yet more fundamental underlying theory, however, within the context of the standard model they have to be fixed in terms of an equal number of experimental quantities. This is what is done in LQCD. Thereafter all predictions of LQCD have to match experimental data if QCD is the correct theory of strong interactions.

A very useful feature of LQCD is that one can dial the input parameters. Therefore, in addition to testing QCD we can make detailed predictions of the dependence of quantities on α_s and the quark masses. These predictions can then be used to constrain effective theories like chiral perturbation theory, heavy quark effective theory, and various phenomenological models.

My first lecture will be devoted to an overview of the scope of LQCD and to showing that simulations of LQCD are a step by step implementation of field theory. The second lecture will be devoted to explaining the details of how to transcribe the quark and gluon degrees of freedom on to the lattice, and to construct an action that, in the limit of zero lattice spacing, gives continuum QCD. I will also spend some time on issues of gauge invariance, chiral symmetry, fermion doubling problem, designing improved actions, the measure of integration, and gauge fixing.

Numerical simulations of LQCD are based on a Monte Carlo integration of the Euclidean path integral, consequently, the measurements have statistical errors in addition to the systematic errors due to lattice discretization. In order to judge the quality of lattice calculations it is important to understand the origin of these errors, what is being done to quantify them, and finally what will it take to achieve results with a given precision. These issues will be covered in the third lecture along with an elementary discussion of Monte Carlo methods.

The fourth lecture will be devoted to the most basic applications of LQCD – the calculation of the hadron spectrum and the extraction of quark masses and α_s . Progress in LQCD has required a combination of improvements in formulation, numerical techniques, and in computer tech-

nology. My overall message is that current LQCD results, obtained from simulations at $a \approx 0.05 - 0.1$ fermi and with quark masses roughly equal to m_s , have already made an impact on Standard Model phenomenology. I hope that the three sets of lectures succeed in communicating the excitement of the practitioners.

In a short course like this I can only give a very brief description of LQCD. In preparing these lectures I have found the following books [1-4], and reviews [5-8] very useful. In addition, the proceedings of the yearly conferences on lattice field theory [9] are excellent sources of information on current activity and on the status of results. I hope that the lectures at this school and these references will allow any interested reader to master the subject.

2. Standard Model of Particle Interactions

The Standard Model (SM) of particle interactions is a synthesis of three of the four forces of nature. These forces are described by gauge theories, each of which is characterized by a coupling constant as listed below.

STRONG INTERACTIONS $\alpha_s \sim 1$

ELECTROMAGNETIC INTERACTIONS $\alpha_{em} \approx 1/137$

WEAK INTERACTIONS $G_F \approx 10^{-5} \text{ GeV}^{-2}$.

The basic constituents of matter are the six quarks, u, d, s, c, b, t , each of which comes in 3 colors, and the six leptons $e, \nu_e, \mu, \nu_\mu, \tau, \nu_\tau$. The quarks and leptons are classified into 3 generations of families. The interactions between these particles is mediated by vector bosons: the 8 gluons mediate strong interactions, the W^\pm and Z mediate weak interactions, and the electromagnetic interactions are carried by the photon γ . The weak bosons acquire a mass through the Higgs mechanism, and in the minimal SM formulation there should exist one neutral Higgs particle. This has not yet been seen experimentally. The mathematical structure that describes the interactions of these constituents is a local gauge field theory with the gauge group $SU(3) \times SU(2) \times U(1)$. The parameters that characterize the SM are listed below; their origins are in an even more fundamental but as yet undiscovered theory.

Parameters in the Standard Model

Parameters	Number	Comments
Masses of quarks	6	u, d, s light c, b heavy $t = 175 \pm 6 \text{ GeV}$
Masses of leptons	6	e, μ, τ $M_{\nu_e, \nu_\mu, \nu_\tau} = ?$
Mass of W^\pm	1	80.3 GeV
Mass of Z	1	91.2 GeV
Mass of gluons, γ		0 (Gauge symmetry)
Mass of Higgs	1	Not yet seen
Coupling α_s	1	≈ 1 for energies $\lesssim 1 \text{ GeV}$
Coupling α_{em}	1	1/137 (=1/128.9 at M_Z)
Coupling G_F	1	10^{-5} GeV^{-2}
Weak Mixing Angles	3	$\theta_1, \theta_2, \theta_3$
CP Violating phase	1	δ
Strong CP parameter	1	$\Theta = ?$

The question marks indicate that these values are tiny and not yet measured. No mixing angles have been put down for the lepton sector as I have assumed that the neutrino masses are all zero. A status report on experimental searches for neutrino masses and their mixing has been presented by Prof. Sciulli at this school. The structure of weak interactions and the origin of the weak mixing angles has been reviewed by Daniel Treille, while David Kosower covered perturbative QCD.

In the SM, the charged current interactions of the W-bosons with the quarks are given by

$$\mathcal{H}_W = \frac{1}{2} (J_\mu W^\mu + h.c.) \quad (2.1)$$

where the current is

$$J_\mu = \frac{g_w}{2\sqrt{2}} (\bar{u}, \bar{c}, \bar{t}) \gamma_\mu (1 - \gamma_5) V \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (2.2)$$

and the Fermi coupling constant is related to the $SU(2)$ coupling as $G_F/\sqrt{2} = g_w^2/8M_W^2$. V is the 3×3 Cabibbo-Kobayashi-Maskawa (CKM) matrix that has a simple representation in terms of the flavor transformation matrix

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (2.3)$$

This matrix has off-diagonal entries because the eigenstates of weak interactions are not the same as those of strong interactions, *i.e.* the d , s , b quarks (or equivalently the u , c , t quarks) mix under weak interactions.

For 3 generations the unitarity condition $V^{-1} = V^\dagger$ imposes constraints. As a result there are only four independent parameters that can be expressed in terms of 4 angles, $\theta_1, \theta_2, \theta_3$ and the CP violating phase δ . The representation originally proposed by Kobayashi and Masakawa is

$$V_{CKM} = \begin{pmatrix} c_1 & -s_1 c_3 & -s_1 s_3 \\ s_1 c_2 & c_1 c_2 c_3 - s_2 s_3 e^{i\delta} & c_1 c_2 s_3 + s_2 c_3 e^{i\delta} \\ s_1 s_2 & c_1 s_2 c_3 + c_2 s_3 e^{i\delta} & c_1 s_2 s_3 - c_2 c_3 e^{i\delta} \end{pmatrix} \quad (2.4)$$

where $c_i = \cos \theta_i$ and $s_i = \sin \theta_i$ for $i = 1, 2, 3$. A phenomenologically more useful representation of this matrix is the Wolfenstein parameterization (correct to $O(\lambda^4)$)

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & \lambda^3 A(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & \lambda^2 A \\ \lambda^3 A(1 - \rho - i\eta) & -\lambda^2 A & 1 \end{pmatrix} + O(\lambda^4). \quad (2.5)$$

where $\lambda = \sin \theta_c = 0.22$ is known to 1% accuracy, $A \approx 0.83$ is known only to 10% accuracy, and ρ and η are poorly known. The elements V_{td} and V_{ub} of the CKM matrix given in Eq. 2.5 are complex if $\eta \neq 0$, and as a result there exists a natural mechanism for CP-violation in the SM. The phenomenology of the CKM matrix and the determinations of the four parameters have been reviewed by Profs. Buras and Richman at this school [10,11].

A list of processes that are amongst the most sensitive probes of the CKM parameters is shown in Fig. 1. Lattice QCD calculations are beginning to provide amongst the most reliable estimates of the hadronic matrix elements relevant to these processes. Details of these calculations and their impact on the determination of ρ and η parameters will be covered by Prof. Martinelli at this school. My goal is to provide you with the necessary background.

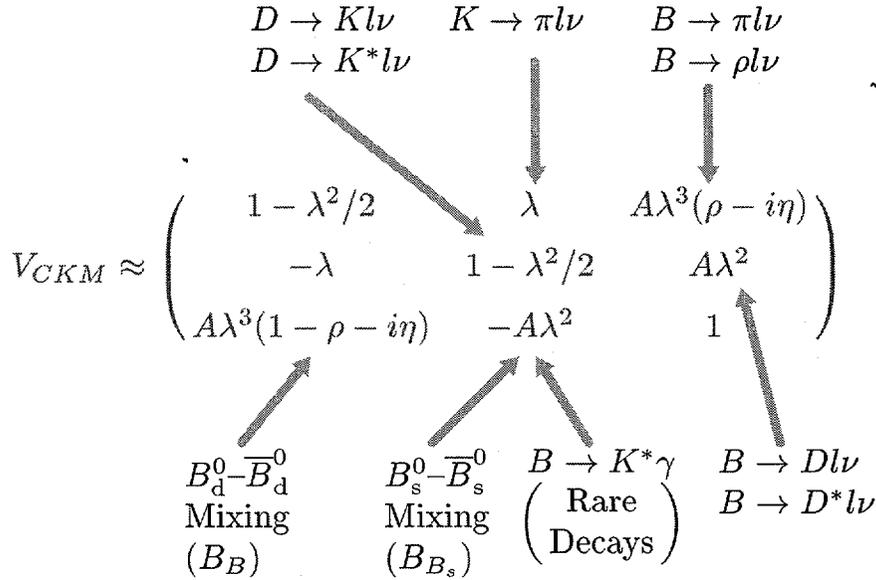


Fig. 1. The CKM matrix in the Wolfenstein parameterization. I show examples of physical processes that need to be measured experimentally and compared with theoretical predictions via the “master” equation to estimate various elements of the CKM matrix.

3. Scope of Lattice QCD Calculations

The primary goal of LQCD is to test whether QCD is the correct theory of strong interactions. Most particle physicists believe that the answer is yes because of its successes in describing processes with large momentum transfer, its mathematical elegance, and uniqueness. However, quantitative confirmation is lacking. The problem with the conventional perturbative approach (expansion in a small parameter) when analyzing hadronic process at scales $\lesssim 1$ GeV is that the strong coupling constant $\alpha_s \sim 1$. Thus, perturbation theory in α_s is not reliable. As a result we cannot calculate the masses of mesons and baryons from QCD even if we are given α_s and the masses of quarks. This is illustrated in Fig. 2 where I show a comparison between the binding energy of a hydrogen atom and that of the proton. It turns out that almost all the mass of the proton is attributed to the strong non-linear interactions of the gluons.

Simulations of lattice QCD have six unknown input parameters. These are the coupling constant α_s and the masses of the up, down, strange, charm and bottom quarks (the top quark is too short lived, 0.4×10^{-24} seconds,

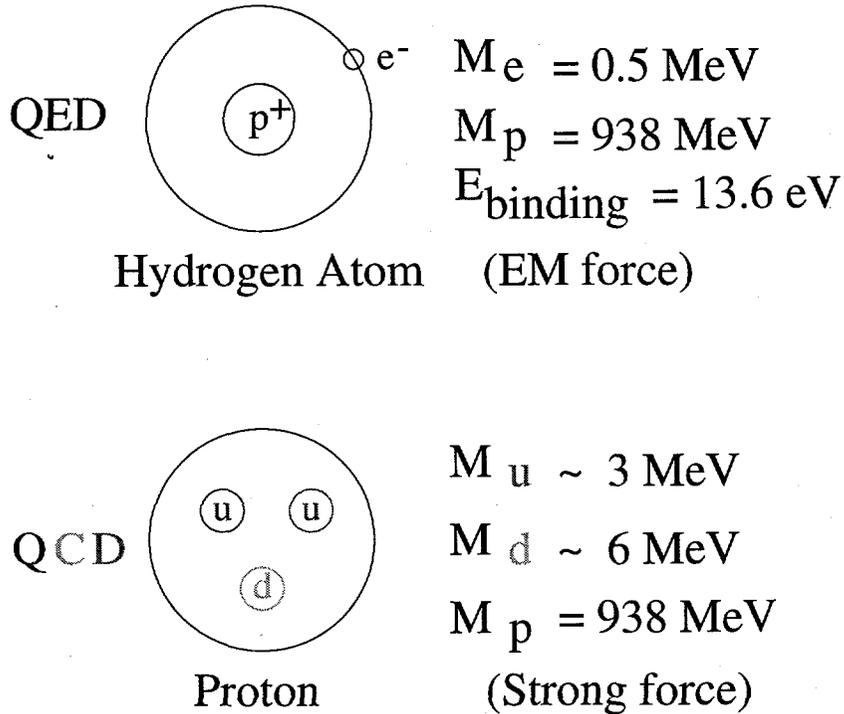


Fig. 2. An illustration of the difference in the binding energy between the electron and the proton in a hydrogen atom (interacting via electromagnetic forces) and the quarks in a proton (strong force).

to form bound states and is thus not suitable for lattice studies). Once these have been fixed in terms of six precisely measured masses of hadrons, thereafter, the masses and properties of the hundreds of other particles made up of these quarks and gluons have to agree with experiments. For example, consider the chiral perturbation theory (χ PT) relation for the pion mass

$$(M_\pi^+)^2 = B_\pi(m_u + m_d) + C_\pi(m_u + m_d)^2 + \dots \quad (3.1)$$

Lattice calculations will provide a check of this relation and fix the constants B_π, C_π, \dots , *i.e.* constrain the chiral lagrangian. In addition, having measured M_π for a number of values of the quark mass, one can invert Eq. 3.1 to determine the current quark masses. I shall call such tests of QCD and the determination of the six unknown parameters of the *SM* a direct application of lattice QCD.

The second area in which perturbative estimates are not reliable is in the calculation of the matrix elements occurring in the weak decays of hadrons. The reason is that the non-perturbative QCD corrections to the basic weak process can be large due to the exchange of soft gluons between the initial and final states. This is illustrated in Fig. 3 for the case of the semi-leptonic decay $D \rightarrow Kl\nu$. In this case the matrix element of the weak interactions Hamiltonian between the initial D meson and final kaon receives large corrections from QCD which cannot be estimated using PQCD.

The result of the experiment is a number, *i.e.* the decay rate. To check whether the Standard Model prediction is in agreement with this number theorists must derive an analytical expression for the same process. This expression, in general, consists of a product of three parts: a combination of the parameters in the SM , matrix elements (ME) of the interaction Hamiltonian between the initial and final states, and kinematic factors. (This is the well-known Fermi's golden rule.) Schematically one can write this "master" equation as follows.

$$Expt.\# = (\text{SM parameters}) (\text{matrix elements}) (\text{kinematic factors}). \quad (3.2)$$

Thus, for each such experimental number one gets a constraint on one particular combination of the SM parameters provided the ME are known. By using several precision experiments and calculating the necessary ME one can over-constrain the model and therefore test it. I shall call lattice calculations of ME an "indirect" probe of Standard Model parameters as one usually needs additional phenomenological/experimental input to extract the unknown parameters via the master equation.

Physicists have been busy calculating and measuring these parameters for the last 25 years, so one already has reasonable estimates. It is, therefore, useful to keep in mind the level of precision that lattice calculations have to aim for to make an impact. Non-lattice estimates, taken from the 1996 Review of Particle Properties and from Prof. Buras's lectures [10], are

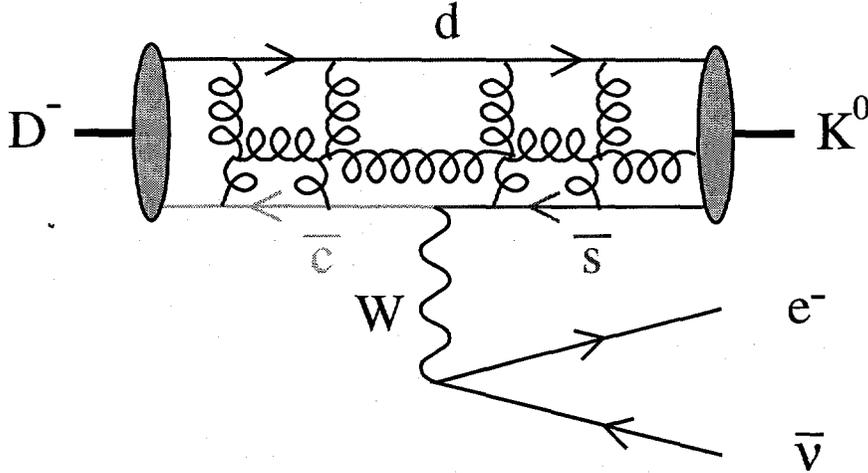


Fig. 3. The Feynman diagram for the semi-leptonic decay $D^- \rightarrow K^0 e^- \bar{\nu}_e$. The QCD corrections are illustrated by the various gluons being exchanged between the initial and final hadrons. The leptonic vertex can be calculated reliably using perturbation theory, whereas the hadronic vertex requires non-perturbative methods.

Parameters	Value	Comments
$m_u(\overline{MS}, 2 \text{ GeV})$	3-4.5 MeV	χ PT, sum-rules
$m_d(\overline{MS}, 2 \text{ GeV})$	5-8 MeV	χ PT, sum-rules
$m_s(\overline{MS}, 2 \text{ GeV})$	100-140 MeV	χ PT, sum-rules
$m_c(\overline{MS}, m_c)$	1.0-1.6 GeV	J/ψ spectra
$m_b(\overline{MS}, m_b)$	4.1-4.6 GeV	Υ spectra
M_{Higgs}		Not Found
$\alpha_s(M_Z)$	0.118 ± 0.003	World average
$ V_{ud} $	0.9736 ± 0.0010	$n \rightarrow pe\nu_e$ and $\mu \rightarrow e\bar{\nu}_e\nu_\mu$
$ V_{us} $	0.2196 ± 0.0023	$K^+ \rightarrow \pi^0 e^+ \nu_e$ and $K_L^0 \rightarrow \pi^\pm e^\mp \nu_e$
$ V_{ub} $	$(3.3 \pm 0.2 \pm 0.4 \pm 0.7)10^{-3}$	$B \rightarrow \pi l\nu$; $B \rightarrow \rho l\nu$
$ V_{cb} $	0.040 ± 0.003	$B \rightarrow X_{cl}\nu$ and $B \rightarrow D^* l\nu$
$ \epsilon $	2.26×10^{-3}	$K^0 \leftrightarrow \bar{K}^0$ mixing
$ \epsilon/\epsilon' $	$0 - 30 \times 10^{-4}$	$K \rightarrow \pi\pi$ decays

Note that only the magnitude of the CKM elements can be determined from experiments. Clearly, the most enticing opportunity for lattice QCD is to help determine the quark masses, α_s , and the essentially unknown Wolfenstein parameters ρ and η . For other elements of V_{CKM} one has reasonable estimates, however here too further progress requires significant improvement in the estimates of the hadronic matrix elements. Phenomenologists are relying on LQCD to provide improved estimates for these.

Let me end this section by mentioning that in addition to validating QCD as the correct theory of strong interactions by reproducing the measured spectrum of mesons, baryons, and glueballs, and the calculation of weak matrix elements, simulations of LQCD will also allow us to

- Investigate the topological structure of the QCD vacuum, and the mechanisms responsible for confinement and spontaneously broken chiral symmetry. For a recent review see [12].

- Calculate hadronic properties like wavefunctions, decay constants, form factors for electromagnetic, weak, and rare decays, and structure functions for deep inelastic scattering. Estimate the pion nucleon sigma term and the proton spin carried by the quarks by calculating the matrix elements $\langle N|\bar{q}q|N \rangle$ and $\langle N|\bar{q}\gamma_\mu\gamma_5q|N \rangle$. Some of these topics will be covered by Prof. Martinelli, for others a good reference is the proceedings of LATTICE conferences [9].

- Analyze the behavior of QCD at high temperatures and address questions like the nature of the phase transition, the transition temperature T_c , the equation of state, and the collective excitations near and above T_c . The topic of QCD at finite temperature has been covered by J-P Blaizot at this school, so I will only list for you my favorite pedagogical reviews of the subject [13]. The lectures by Profs. Martinelli, Lüscher, and me will be confined to understanding and probing the properties of QCD at zero temperature.

4. Overview of the Lattice Approach

LQCD calculations are a non-perturbative implementation of field theory using the Feynman path integral approach. The calculations proceed exactly as if the field theory was being solved analytically had we the ability to do the calculations. The starting point is the partition function in Euclidean space-time

$$Z = \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{-S} \quad (4.1)$$

where S is the QCD action

$$S = \int d^4x \left(\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \bar{\psi} M \psi \right). \quad (4.2)$$

and M is the Dirac operator. The fermions are represented by Grassmann variables ψ and $\bar{\psi}$. These can be integrated out exactly with the result

$$Z = \int \mathcal{D}A_\mu \det M e^{\int d^4x \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right)}. \quad (4.3)$$

The fermionic contribution is now contained in the highly non-local term $\det M$, and the partition function is an integral over only background gauge configurations. One can write the action, after integration over the fermions, as $S = S_{gauge} + S_{quarks} = \int d^4x \left(\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right) - \sum_i \log(\text{Det} M_i)$ where the sum is over the quark flavors, distinguished by the value of the bare quark mass.

It is expedient to define the “quenched” approximation (QQCD) from the very start due to the central role it has played in simulations. It consists of setting $\text{Det} M = \text{constant}$ which corresponds to removing vacuum polarization effects from the QCD vacuum. Details and some consequences of this approximation are discussed in Section 16.8.

Results for physical observables are obtained by calculating expectation values

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int \mathcal{D}A_\mu \mathcal{O} e^{-S}. \quad (4.4)$$

where \mathcal{O} is any given combination of operators expressed in terms of time-ordered products of gauge and quark fields. (Expectation values in the path integral approach correspond to time-ordered correlation functions.) The quarks fields in \mathcal{O} are, in practice, re-expressed in terms of quark propagators using Wick’s theorem for contracting fields. In this way all dependence on quarks as dynamical fields is removed. The basic building block for fermionic quantities is the Feynman propagator,

$$S_F(y, j, b; x, i, a) = (M^{-1})_{x, i, a}^{y, j, b}, \quad (4.5)$$

where M^{-1} is the inverse of the Dirac operator calculated on a given background field. A given element of this matrix $(M^{-1})_{x, i, a}^{y, j, b}$ is the amplitude for the propagation of a quark from site x with spin-color i, a to site-spin-color y, j, b .

So far all of the above is standard field theory. The problem we face in QCD is how to actually calculate these expectation values and how to

extract physical observables from these. I will illustrate the second part first by using as an example the mass and decay constant of the pion.

Consider the 2-point correlation function, $\langle 0|T[\sum_x \mathcal{O}_f(\mathbf{x}, t)\mathcal{O}_i(\mathbf{0}, 0)]|0\rangle$ with $t > 0$, where the operators \mathcal{O} are chosen to be the fourth component of the axial current $\mathcal{O}_f = \mathcal{O}_i = A_4 = \bar{\psi}\gamma_4\gamma_5\psi$ as these have a large coupling to the pion. This 2-point correlation function gives the amplitude for creating a state with the quantum numbers of the pion out of the vacuum at space-time point 0 by the ‘‘source’’ operator \mathcal{O}_i ; the evolution of this state to the point (\mathbf{x}, t) via the QCD Hamiltonian (or more precisely the transfer matrix to be discussed in Section 5); and finally the annihilation by the ‘‘sink’’ operator \mathcal{O}_f at (\mathbf{x}, t) . The rules of quantum mechanics tell us that \mathcal{O}_i will create a state that is a linear combination of all possible eigenstates of the Hamiltonian that have the same quantum numbers as the pion, *i.e.* the pion, radial excitations of the pion, three pions in $J = 0$ state, \dots . The second rule is that on propagating for Euclidean time t , a given eigenstate with energy E picks up a weight e^{-Et} . Thus, the 2-point function can be written in terms of a sum over all possible intermediate states

$$\langle 0|\sum_x \mathcal{O}_f(\mathbf{x}, t)\mathcal{O}_i(\mathbf{0})|0\rangle = \sum_n \frac{\langle 0|\mathcal{O}_f|n\rangle\langle n|\mathcal{O}_i|0\rangle}{2E_n} e^{-E_n t}. \quad (4.6)$$

To study the properties of the pion at rest we need to isolate this state from the sum over n . To do this, the first simplification is to use the Fourier projection \sum_x as it restricts the sum over states to just zero-momentum states, so $E_n \rightarrow M_n$. (Note that it is sufficient to make the Fourier projection over either \mathcal{O}_i or \mathcal{O}_f .) The second step to isolate the pion, *i.e.* project in the energy, consists of a combination of two strategies. One, make a clever choice of the operators \mathcal{O} to limit the sum over states to a single state (the ideal choice is to set \mathcal{O} equal to the quantum mechanical wave-functional of the pion), and two, examine the large t behavior of the 2-point function where only the contribution of the lowest energy state that couples to \mathcal{O}_i is significant due to the exponential damping. Then

$$\langle 0|\sum_x \mathcal{O}_f(\mathbf{x}, t)\mathcal{O}_i(\mathbf{0})|0\rangle \xrightarrow{t \rightarrow \infty} \frac{\langle 0|\mathcal{O}_f|\pi\rangle\langle \pi|\mathcal{O}_i|0\rangle}{2M_\pi} e^{-M_\pi t}. \quad (4.7)$$

The right hand side is now a function of the two quantities we want since $\langle 0|A_4(\mathbf{p} = 0)|\pi\rangle = M_\pi f_\pi$. In this way, the mass and the decay constant are extracted from the rate of exponential fall-off in time and from the amplitude.

Let me now illustrate how the left hand side of Eq. 4.7 is expressed in terms of the two basic quantities we control in the path integral – the gauge

fields and the quark propagator. Using Wick contractions, the correlation function can be written in terms of a product of two quark propagators S_F ,

$$\begin{aligned} & -\langle 0 | \sum_x \bar{\psi}(x, t) \gamma_4 \gamma_5 \psi(x, t) \bar{\psi}(0, 0) \gamma_4 \gamma_5 \psi(0, 0) | 0 \rangle \\ & \equiv \langle 0 | \sum_x S_F(0; \mathbf{x}, t) \gamma_4 \gamma_5 S_F(\mathbf{x}, t; 0) \gamma_4 \gamma_5 | 0 \rangle. \end{aligned} \quad (4.8)$$

This correlation function is shown in Fig. 4A. To illustrate this Wick contraction procedure further consider using gauge invariant non-local operators, for example take $\mathcal{O} = \bar{\psi}(x, t) \gamma_4 \gamma_5 (\mathcal{P} e^{\int_x^y dz ig A_\mu(z)}) \psi(y, t)$ where \mathcal{P} stands for path-ordered. After Wick contraction the correlation function reads

$$\langle 0 | \sum_x S_F(0; \mathbf{x}, t) \gamma_4 \gamma_5 (\mathcal{P} e^{\int_x^z ig A_\mu}) S_F(\mathbf{z}, t; \mathbf{y}, 0) \gamma_4 \gamma_5 (\mathcal{P} e^{\int_0^y ig A_\mu}) | 0 \rangle. \quad (4.9)$$

and involves both the gauge fields and quark propagators. This correlation function would have the same long t behavior as shown in Eq. 4.6, however, the amplitude will be different and consequently its relation to f_π will no longer be simple. The idea behind strategies for improving the projection of \mathcal{O} on to the pion is to construct a suitable combination of such operators that approximates the pion wave-function. For example, if $\phi(x)$ is the pion wave-functional, then the ‘‘smeared’’ operator $\mathcal{O} = \bar{\psi}(0) \gamma_4 \gamma_5 (\int d^3x \phi(x) \psi(x))$ is the ideal choice.

Finally, it is important to note that to get the 2-point function corresponding to the propagation of the physical pion we have to average the correlation function over gauge configurations as defined in Eq. 4.4, *i.e.* do the functional integral.

An overview of how such calculations of expectation values can be done consists of the following steps.

- *Defining a finite dimensional system:* The Yang-Mills action for gauge fields and the Dirac operator for fermions have to be transcribed onto a discrete space-time lattice in such a way as to preserve all the key properties of QCD – gauge invariance, chiral symmetry, topology, and a one-to-one relation between continuum and lattice fields. This step is the most difficult, and even today we do not have a really satisfactory lattice formulation that is chirally symmetric in the $m_q = 0$ limit and preserves a one-to-one relation between continuum and lattice fields, *i.e.* no doublers. In fact, the Nielsen-Ninomiya theorem states that for a translationally invariant, local, hermitian formulation of the lattice theory one cannot simultaneously have

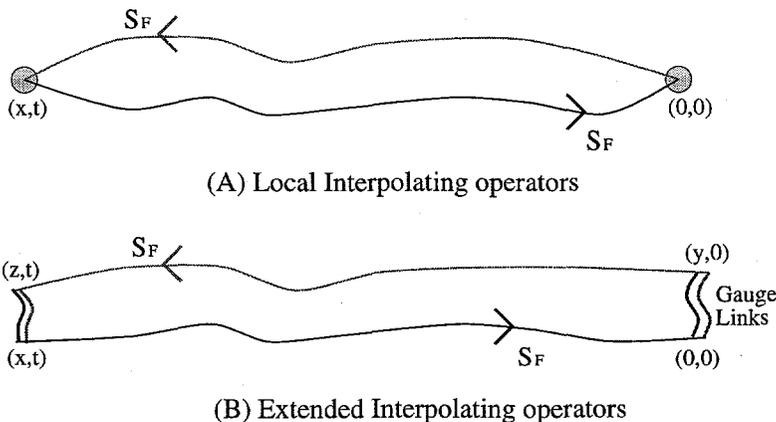


Fig. 4. A schematic of the pion 2-point correlation function for (A) local and (B) non-local interpolating operators.

chiral symmetry and no doublers [14]. One important consequence of this theorem is that, in spite of tremendous effort, there is no practical formulation of chiral fermions on the lattice. For a review of the problems and attempts to solve them see [15–17] and [18].

• *Discretization Errors*: This problem is encountered when approximating derivatives in the action by finite differences,

$$\begin{aligned} \partial_x \phi(x) &\rightarrow \Delta_x \phi(x) = \frac{1}{2a} (\phi(x+a) - \phi(x-a)) \\ &= \frac{1}{2a} (e^{a\partial_x} - e^{-a\partial_x}) \phi(x). \end{aligned} \quad (4.10)$$

As is obvious this introduces discretization errors proportional to the square of the lattice spacing a . These errors can be reduced by either using higher order difference schemes with coefficients adjusted to take into account effects of renormalization, or equivalently, by adding appropriate combinations of irrelevant operators to the action that cancel the errors order by order in a . The various approaches to improving the gauge and Dirac actions are discussed in Sections 12 and 13. In my lectures I will concentrate on the three most frequently used discretizations of the Dirac action – Wilson [19], Sheikholeslami-Wohlert (clover) [20], and staggered [21], which have errors of $O(a)$, $O(\alpha_s a) - O(a^2)$ depending on the value of the coefficient of the clover term, and $O(a^2)$ respectively. Reducing the discretization errors by improving the action and operators increases the reliability of the results, and for dynamical simulations may even prove necessary due to

limitations of computer power. At the same time it is important to note that there exists a well defined procedure for obtaining continuum results with even the simplest formulation, Wilson's original gauge and fermion action [19]. In this case, however, it is necessary to quantify and remove the discretization errors by a reliable extrapolation to $a = 0$.

- *Generating background gauge fields:* The Euclidean action $\mathcal{S} = \int d^4x (\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \text{Tr} \log M)$ for QCD at zero chemical potential is real and bounded from below. Thus $e^{-\mathcal{S}}$ in the path integral is analogous to the Boltzmann factor in the partition function for statistical mechanics systems, *i.e.* it can be regarded as a probability weight for generating configurations. Since \mathcal{S} is an extensive quantity the range of \mathcal{S} is enormous, however, the distribution is very highly peaked about configurations that minimize the free energy. Configurations close to the peak dominate the functional integral. Such important configurations can be generated by setting up a Markov chain in exact analogy to, say, simulations of the Ising model. They are called "importance sampled" as they are generated with probability of occurrence given by the weight $e^{-\mathcal{S}}$. For a discussion of the methods used to update the configurations see [1] or the lectures by Creutz and Sokal in [2].

- *Calculation of the quark propagator on a given background gauge field:* For a given background gauge configuration, the Feynman quark propagator S_F is a matrix labeled by three indices – site, spin, and color. A given element of this matrix gives the amplitude for the propagation of a quark with some spin, color, and space-time point to another space-time point, spin, and color. Operationally, S_F is simply the inverse of the Dirac operator M . Once space-time is made discrete and finite, the Dirac matrix is also finite and its inverse can be calculated numerically. Inverting the Dirac operator is the key computational step in LQCD and consumes over 90% of the CPU cycles. To construct the correlation functions requires S_F only from a given source point to all other sites of the lattice. This corresponds to 12 columns (or rows) of M^{-1} , one each for the 12 spin and color degrees of freedom.

- *Correlation functions:* The correlation functions are expressed as path ordered products of quark propagators and gauge fields using Wick contraction of the operator fields. There are two kinds of correlation functions that are calculated – gauge invariant correlation functions as discussed above for the case of the pion, or those in a fixed gauge.

- *Expectation values:* The expectation value is simply an average of the correlation function evaluated over the set of "importance sampled" configurations. Even on a finite lattice the set of background gauge config-

urations is infinite as the variables are continuous valued. Thus, while it is possible to calculate the correlation functions on specified background gauge configurations, doing the functional integral exactly is not feasible. It is, therefore, done numerically using monte carlo methods.

My recommendation for the simplest way to understand the computational aspects of LQCD calculations is to first gain familiarity with the numerical treatment of any simple statistical mechanics system, for example the Ising model. The differences are: (i) the degrees of freedom in LQCD are much more complicated – SU(3) link matrices rather than Ising spins, and quark propagators on these background configurations evaluated by inverting the Dirac operator; (ii) The action involves the highly nonlocal term $\text{Ln Det } M$ which makes the update of the gauge configurations very expensive; and (iii) the correlation functions are not simple products of spin variables like the specific heat or magnetic susceptibility, but complicated functions of the link variables and quark propagators.

The subtleties arising due to the fact that LQCD is a renormalizable quantum field theory and not a classical statistical mechanics system come into play when analyzing the behavior of the correlation functions as a function of the lattice spacing a , and in the quantum corrections that renormalize the input parameters (quark and gluon fields and their masses) and the composite operators used in the correlation functions. At first glance it might seem that we have introduced an additional parameter, the lattice spacing a , in LQCD, however, as will be shown in Section 14, the coupling g and the cutoff a are not independent quantities but are related by the renormalization group as

$$\Lambda_{QCD} = \lim_{a \rightarrow 0} \frac{1}{a} e^{-1/2\beta_0 g^2(a)} (\beta_0 g^2(a))^{-\beta_1/2\beta_0^2}, \quad (4.11)$$

where Λ_{QCD} is the non-perturbative scale of QCD, and β_0 and β_1 are the first two, scheme independent, coefficients of the β -function. In statistical mechanics systems, the lattice spacing a is a physical quantity – the intermolecular separation. In a quantum field theory, it is simply the ultraviolet regulator that must eventually be taken to zero keeping physical quantities, like the renormalized couplings (Λ_{QCD} and quark masses), spectrum, etc, fixed.

The reason that lattice results are not exact is because in numerical simulations we have to make a number of approximations as discussed in Section 16. The size of the associated uncertainties is dictated by the computer power at hand. They are being improved steadily with computer technology, better numerical algorithms, and better theoretical understand-

ing. I shall return to the issue of errors after discussing the formulation of LQCD.

5. Connection between Minkowski and Euclidean Field Theory and with Statistical Mechanics

This section is a brief recapitulation of technical issues regarding the Euclidean path integral and extracting physics from Euclidean correlation functions. For details see [3,4] or the texts on statistical field theory, for example [22,23].

A D -dimensional Minkowski field theory ($D - 1$ spatial dimensions and one time dimension) is connected to a D -dimensional Euclidean field theory through analytical continuation. Under Wick rotation

$$\begin{aligned} \tilde{x}_0 \equiv t &\rightarrow -ix_4 \equiv -i\tau, \\ p_0 \equiv E &\rightarrow ip_4. \end{aligned} \quad (5.1)$$

The Euclidean convention is

$$\begin{aligned} x_E^2 &= \sum_{i=1}^4 x_i^2 = x^2 - t^2 = -x_M^2, \\ p_E^2 &= \sum_{i=1}^4 p_i^2 = p^2 - E^2 = -p_M^2. \end{aligned} \quad (5.2)$$

The connection between statistical mechanics and a Euclidean field theory in D -dimensions is most transparent when the field theory is quantized using the Feynman path integral approach. This connection forms the basis of numerical simulations using methods common to those used to study statistical mechanics systems. A synopsis of the equivalences between the two is given in Table 1.

Let me start with a quick summary of the Euclidean theory. Consider the generic 2-point correlation function $\Gamma_M(t, \mathbf{k}) = \langle 0 | \int d^3x e^{-ik \cdot x} T[\mathcal{F}(x, t) \mathcal{F}^\dagger(0, 0)] | 0 \rangle$ in Minkowski space. The Heisenberg operator $\mathcal{F}(x, t)$ can be written as

$$\mathcal{F}(x, t) = e^{i\mathcal{H}t - ipx} \mathcal{F} e^{-i\mathcal{H}t + ipx} \quad (5.3)$$

so that

$$\Gamma_M(t, \mathbf{k}) = \langle 0 | \int d^3x e^{-ik \cdot x} \left(e^{i\mathcal{H}t - ipx} \mathcal{F} e^{-i\mathcal{H}t + ipx} \mathcal{F}^\dagger \right) | 0 \rangle \quad (5.4)$$

Table 1

The equivalences between a Euclidean field theory and Classical Statistical Mechanics.

Euclidean Field Theory	Classical Statistical Mechanics
Action	Hamiltonian
unit of action \hbar	units of energy $\beta = 1/kT$
Feynman weight for amplitudes $e^{-S/\hbar} = e^{-\int \mathcal{L} dt/\hbar}$	Boltzmann factor $e^{-\beta H}$
Vacuum to vacuum amplitude $\int \mathcal{D}\phi e^{-S/\hbar}$	Partition function $\sum_{conf.} e^{-\beta H}$
Vacuum energy	Free Energy
Vacuum expectation value $\langle 0 \mathcal{O} 0\rangle$	Canonical ensemble average $\langle \mathcal{O} \rangle$
Time ordered products	Ordinary products
Green's functions $\langle 0 T[\mathcal{O}_1 \dots \mathcal{O}_n] 0\rangle$	Correlation functions $\langle \mathcal{O}_1 \dots \mathcal{O}_n \rangle$
Mass M	correlation length $\xi = 1/M$
Mass-gap	exponential decrease of correlation functions
Mass-less excitations	spin waves
Regularization: cutoff Λ	lattice spacing a
Renormalization: $\Lambda \rightarrow \infty$	continuum limit $a \rightarrow 0$
Changes in the vacuum	phase transitions

where, in dropping the time-ordered product, it is assumed that $t > 0$. Using the invariance of the vacuum under time and space translations and assuming $\mathcal{F}^\dagger = \mathcal{F}$ for brevity, this becomes

$$\begin{aligned} \Gamma_M(t, \mathbf{k}) &= \langle 0 | \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \left(\mathcal{F} e^{-i\mathcal{H}t + i\mathbf{p}\cdot\mathbf{x}} \mathcal{F} \right) | 0 \rangle \\ &= \langle 0 | \delta(\mathbf{p} - \mathbf{k}) \left(\mathcal{F} e^{-i\mathcal{H}t} \mathcal{F} \right) | 0 \rangle. \end{aligned} \quad (5.5)$$

Since the time dependence is explicit, we can Wick rotate to Euclidean space using $t \rightarrow -i\tau$ and keeping \mathcal{H} unchanged

$$\Gamma_E(t) = \langle 0 | \delta(\mathbf{p} - \mathbf{k}) \left(\mathcal{F} e^{-\mathcal{H}\tau} \mathcal{F} \right) | 0 \rangle. \quad (5.6)$$

To get the path integral representation we divide the time interval into N steps $\tau = Na$ and insert a complete set of states $|\phi\rangle\langle\phi|$ at each step.

$$\begin{aligned} \Gamma_E(t) &= \delta(\mathbf{p} - \mathbf{k}) \sum_{\phi_1 \dots \phi_N} \langle 0 | \mathcal{F} | \phi_1 \rangle \langle \phi_1 | e^{-\mathcal{H}a} | \phi_2 \rangle \dots \\ &\quad \dots \langle \phi_{N-1} | e^{-\mathcal{H}a} | \phi_N \rangle \langle \phi_N | \mathcal{F} | 0 \rangle \end{aligned} \quad (5.7)$$

The discussion of how, by this introduction of the complete set of states, one trades the operator form of the quantum mechanical amplitude by

the sum over paths is standard and thus not repeated. On this "lattice" the transfer matrix $\mathcal{T} \equiv \exp(-\mathcal{H}a)$ plays the role of the time translation operator, from which the continuum Hamiltonian is obtained as

$$\mathcal{H} = \lim_{a \rightarrow 0} -\frac{1}{a} \log \mathcal{T}. \quad (5.8)$$

In order for \mathcal{H} to be a self-adjoint Hamiltonian it is necessary that \mathcal{T} is a symmetric, bounded, and positive operator acting on a Hilbert space of states with a positive norm. The necessary condition for this, is the Osterwalder-Schrader *reflection positivity* [24]. That this condition holds for LQCD has been shown by an explicit construction by M. Lüscher [25]. Therefore, the correlation functions in the desired physical Minkowski theory can be obtained by analytical continuation of their Euclidean counterparts.

Assuming that the conditions for a sensible Euclidean theory have been established, the vacuum to vacuum amplitude (also called the generating functional or the partition function), in presence of a source J , is

$$Z[J] = \int d\phi e^{-S_E + J\phi} \quad (5.9)$$

where S_E is the Euclidean action. Correlation functions can be obtained from this in the standard way by differentiating $\log Z[J]$ with respect to J , for example $\langle 0 | \phi(x) \phi(0) | 0 \rangle = \frac{\delta}{\delta J(x)} \frac{\delta}{\delta J(0)} \log Z[J] |_{J=0}$.

Lastly, I summarize the relations between Euclidean and Minkowski quantities. The basic tool is the repeated insertion of a complete set of momentum states in Eq. 5.7. I use the relativistic normalization of states

$$\langle k | p \rangle = 2EL^3 \delta_{k_x p_x} \delta_{k_y p_y} \delta_{k_z p_z} \xrightarrow{L \rightarrow \infty} 2E(2\pi)^3 \delta^3(\mathbf{p} - \mathbf{k}). \quad (5.10)$$

Then, for a spectrum of stable isolated states $|n\rangle$ with masses M_n that couple to an interpolating field \mathcal{F} ,

$$\Gamma_E(\tau) = \sum_n \frac{\langle 0 | \mathcal{F} | n \rangle \langle n | \mathcal{F} | 0 \rangle}{2M_n} e^{-M_n \tau}. \quad (5.11)$$

Now it is easy to obtain the connection between these M_n and the physical spectrum. A fourier transform in Euclidean time followed by a rotation to Minkowski space gives

$$\begin{aligned} \int d\tau e^{ip_4 \tau} \frac{e^{-M_n |\tau|}}{2M_n} &= \frac{1}{2M_n(M_n - ip_4)} + \frac{1}{2M_n(M_n + ip_4)}, \\ &= \frac{1}{M_n^2 + p_4^2}, \end{aligned}$$

$$\xrightarrow{p_4 \rightarrow -iE} \frac{1}{M_n^2 - E^2}. \quad (5.12)$$

The M_n are precisely the location of the poles in the propagator of $|n\rangle$. This is why one can simply short step the necessary analytical continuation as the masses measured from the exponential fall-off are the pole masses of the physical states. For $\tau \rightarrow \infty$ only the lowest state contributes due to the exponential suppression, and this provides a way of isolating the lightest mass state.

This simple connection holds also for matrix elements. Consider, for $\tau_2 \gg \tau_1 \gg 0$, the three-point function

$$\int d^3x d^3y e^{-ik \cdot x} e^{-iq \cdot y} \langle 0 | T[\mathcal{F}_f(x, \tau_2) \mathcal{O}(y, \tau_1) \mathcal{F}_i(0, 0)] | 0 \rangle = \sum_a \sum_b \frac{\langle 0 | \mathcal{F}_f | a \rangle e^{-M_a(\tau_2 - \tau_1)}}{2M_a} \langle a | \mathcal{O} | b \rangle \frac{\langle b | \mathcal{F}_i | 0 \rangle e^{-M_b \tau_1}}{2M_b}. \quad (5.13)$$

The two factors on the left and right of $\langle a | \mathcal{O} | b \rangle$ are each, up to one decay amplitude, the same as the 2-point function discussed in Eq. 5.11. They, therefore, provide, for large $\tau_2 - \tau_1$ and τ_1 , the isolation of the lowest states in the sum over states. These can be gotten from a calculation of the 2-point function. Consequently, $\langle a | \mathcal{O} | b \rangle$ measured in Euclidean simulations is the desired matrix element between these states and no analytic continuation to Minkowski space is necessary.

These arguments break down once the sum over states is not just over stable single particle states. Consider, for example, an \mathcal{F} that is the ρ meson interpolating field. In that case there are also the $2\pi, 4\pi$ intermediate states with a cut beginning at $E = \sqrt{2M_\pi}$. The physical characteristics of ρ meson (mass and width) can be represented by a complex pole in the propagator that, in Minkowski space, lies on the second sheet under the cut. This physical pole can no longer be reproduced by measuring the “energy” given by the exponential fall-off of Euclidean correlation function at large τ [26]. It requires measuring the correlation function for all intermediate states and momenta (all τ), and then doing the fourier transform in τ followed by the analytical continuation $p_4 \rightarrow -iE$.

For the same reason there does not exist a simple way for calculating n-particle scattering amplitudes. A general 2-particle scattering amplitude (4-point correlation function) is complex precisely because the intermediate states involve a sum over resonances and a continuum with all values of momenta. In principle this can be obtained from Euclidean correlation functions but only by measuring them with essentially infinite precision for all possible τ and then doing the analytical continuation. In practice this

is a hopeless possibility. The one exception is the amplitude at threshold which is real. M. Lüscher has shown how the 2-particle scattering amplitudes at threshold can be calculated from the energy shift in a finite volume, the difference between the 2-particle energy and the sum of the two particles in isolation [27–30].

6. Formulation of Lattice Gauge Theories

In 1974 Wilson [31] formulated Euclidean gauge theories on the lattice as a means to study confinement and carry out non-perturbative analysis of QCD. The numerical implementation of the path integral approach requires the following five steps.

- Discretization of space-time.
- The transcription of the gauge and fermion degrees of freedom.
- Construction of the action.
- Definition of the measure of integration in the path integral.
- The transcription of the operators used to probe the physics.

Of these, the construction of the action and the operators is the most intricate. Rather than give you the most recent and improved version of the lattice action, which will be the subject of Martin Lüscher's lectures, I have chosen to follow the original path of discovery.

6.1. Discrete space-time grid

There are a number of possible ways to discretize space-time in 4 Euclidean dimensions. These include hypercubic, body-centered cubic [32], and random [33] lattices. Of these the simplest is the isotropic hypercubic grid with spacing $a = a_S = a_T$ and size $N_S \times N_S \times N_S \times N_T$. Since this grid is used almost exclusively, it will be the only one discussed in these sets of lectures. In the case of the random lattice even the free field analysis becomes difficult [33]. In fact very little exploration of variant grid geometries has actually been done, as there is little expectation of significant physics gains to offset the additional computational complexity. Recently, proposals involving anisotropic lattices – a hypercubic grid with different spacing in space and time directions – have shown promise for reducing the discretization errors. I shall come back to this development in Section 18.3 when discussing the glueball spectrum.

6.2. Lattice transcription of field variables $\psi(x)$ and $A_\mu(x)$:

The lattice transcription of the matter fields $\psi(x)$ is straightforward. The quark field is represented by anticommuting Grassmann variables defined at each site of the lattice. They belong to the fundamental representation of SU(3). The rules for integration over a pair of fermion variables ψ and $\bar{\psi}$ are those standard for Grassmann variables

$$\begin{aligned} \int d\psi d\bar{\psi} &= \int d\psi d\bar{\psi} \psi = \int d\psi d\bar{\psi} \bar{\psi} = 0 \\ \int d\psi \psi &= \int d\bar{\psi} \bar{\psi} = i \\ \int d\psi d\bar{\psi} \psi \bar{\psi} &= 1 \end{aligned} \tag{6.1}$$

Since the fermion action is linear in both ψ and $\bar{\psi}$, these rules can be used to integrate over them, and the path integral reduces to one over only the gauge degrees of freedom. Thus, it turns out that in practice one does not have to worry about transcribing grassmann variables on the lattice and implementing the Pauli exclusion principle.

As an aside I would like to mention that the rules of integration for Grassmann variables are sufficient to do the fermion path integral. Consider $\langle \bar{\psi}(x_1) \dots \bar{\psi}(x_n) \psi(x_1) \dots \psi(x_n) \rangle$. To get a non-zero expectation value one needs to saturate each site with a $\bar{\psi}\psi$. The sites not covered by the fields in the operator have to come from the \mathcal{S}_f^{V-n} term in the expansion $\exp \mathcal{S}_f = 1 + \mathcal{S}_f + \mathcal{S}_f^2/2 + \dots$, where V is the total number of sites. Since $\mathcal{S}_f = \sum_x \bar{\psi}(x)(\not{D} + m)\psi(x)$ is a sum over sites, the number of terms in \mathcal{S}_f^{V-n} poses an exponentially hard combinatorial problem. Furthermore, each term contributing to the connected diagram will contain a path-ordered product of gauge links (coming from the discretized \not{D} term in \mathcal{S}_f . For example, see Eq. 7.5) connecting the sites in the operator, while the fermionic integration is $\pm m^{V-n}$ with the sign given by the number of times the ψ and $\bar{\psi}$ are anti-commuted. You should convince yourself that the fermionic functional integral is the sum over all possible paths touching each site a maximum of one time. Having enumerated all such paths and their weights, the gauge functional integral now requires integrating over each link in each path. This is clearly a very formidable approach, and in practice rarely considered.

The construction of gauge fields is less intuitive. In the continuum, the gauge fields $A_\mu(x)$ carry 4-vector Lorentz indices, and mediate interactions between fermions. To transcribe them Wilson noted that in the continuum

a fermion moving from site x to y in presence of a gauge field $A_\mu(x)$ picks up a phase factor given by the path ordered product

$$\psi(y) = \mathcal{P} e^{\int_x^y ig A_\mu(x) dx_\mu} \psi(x). \quad (6.2)$$

The eight types of gluons that mediate interactions between quarks are written in terms of the matrix $A_\mu(x) \equiv A_\mu^a(x) \cdot \lambda_a/2$, where the group generators λ_a are normalized to $\text{Tr} \lambda_a \lambda_b = 2\delta_{ab}$. Eq. 6.2 suggested that gauge fields be associated with links that connect sites on the lattice. So, with each link Wilson associated a discrete version of the path ordered product

$$U(x, x + \hat{\mu}) \equiv U_\mu(x) = e^{iag A_\mu(x + \frac{\hat{\mu}}{2})}, \quad (6.3)$$

where, for concreteness, the average field A_μ is defined at the midpoint of the link, and U is a 3×3 unitary matrix with unit determinant. Also, the path ordering in eqn. (2.3) specifies that

$$U(x, x - \hat{\mu}) \equiv U_{-\mu}(x) = e^{-iag A_\mu(x - \frac{\hat{\mu}}{2})} = U^\dagger(x - \hat{\mu}, x). \quad (6.4)$$

6.3. Discrete Symmetries of the lattice theory

The symmetry group of the continuum theory – Poincaré invariance – is reduced to a discrete group. On a hypercubic lattice rotations by only 90° are allowed so the continuous rotation group is replaced by the discrete hypercubic group [34]. Translations have to be by at least one lattice unit, so the allowed momenta are discrete

$$k = \frac{2\pi n}{La} \quad n = 0, 1, \dots, L$$

or equivalently

$$k = \pm \frac{2\pi n}{La} \quad n = 0, 1, \dots, L/2.$$

On the lattice momentum is conserved *modulo* 2π .

In addition to the local gauge symmetry and Poincaré invariance, the lattice action is invariant under parity (\mathcal{P}), charge conjugation (\mathcal{C}) and time reversal (\mathcal{T}). The behavior of the field variables under these discrete transformations is given in Table 2.

6.4. Local Gauge Symmetry

The effect of a local gauge transformation $V(x)$ on the variables $\psi(x)$ and U is defined to be

$$\psi(x) \rightarrow V(x)\psi(x)$$

Table 2

The behavior of the gauge and fermion degrees of freedom under the discrete transformations of parity, charge-conjugation, and time-reversal. The charge-conjugation matrix \mathcal{C} satisfies the relation $\mathcal{C}\gamma_\mu\mathcal{C}^{-1} = -\gamma_\mu^T = -\gamma_\mu^*$, and shall be represented by $\mathcal{C} = \gamma_4\gamma_2$. Note that in the Euclidean formulation \mathcal{T} is a linear operator.

	\mathcal{P}	\mathcal{C}	\mathcal{T}
$U_4(\mathbf{x}, \tau)$	$U_4(-\mathbf{x}, \tau)$	$U_4^*(\mathbf{x}, \tau)$	$U_{-4}(\mathbf{x}, -\tau)$
$U_i(\mathbf{x}, \tau)$	$U_{-i}(-\mathbf{x}, \tau)$	$U_i^*(\mathbf{x}, \tau)$	$U_i(\mathbf{x}, -\tau)$
$\psi(\mathbf{x}, \tau)$	$\gamma_4\psi(-\mathbf{x}, \tau)$	$\mathcal{C}\bar{\psi}^T(\mathbf{x}, \tau)$	$\gamma_4\gamma_5\psi(\mathbf{x}, -\tau)$
$\bar{\psi}(\mathbf{x}, \tau)$	$\bar{\psi}(-\mathbf{x}, \tau)\gamma_4$	$-\psi^T(\mathbf{x}, \tau)\mathcal{C}^{-1}$	$\bar{\psi}(\mathbf{x}, -\tau)\gamma_5\gamma_4$

$$\begin{aligned}\bar{\psi}(x) &\rightarrow \bar{\psi}(x)V^\dagger(x) \\ U_\mu(x) &\rightarrow V(x)U_\mu(x)V^\dagger(x + \hat{\mu})\end{aligned}\quad (6.5)$$

where $V(x)$ is in the same representation as the U , i.e., it is an $SU(3)$ matrix. With these definitions there are two types of gauge invariant objects that one can construct on the lattice.

– A string consisting of a path-ordered product of links capped by a fermion and an antifermion as shown in Fig. 5a. A simple example is

$$\text{Tr } \bar{\psi}(x) U_\mu(x) U_\nu(x + \hat{\mu}) \dots U_\rho(y - \hat{\rho}) \psi(y) \quad (6.6)$$

where the trace is over the color indices. I will use the word string for a generalized version; a single spin and color trace as defined in Eq. 6.6, or a product of elementary strings that are path-ordered and do not have a spatial discontinuity. On lattices with periodic boundary conditions one does not need the $\psi, \bar{\psi}$ caps if the string stretches across the lattice and is closed by the periodicity. Such strings are called Wilson/Polyakov lines.

– Closed Wilson loops as shown in Fig. 5b. The simplest example is the plaquette, a 1×1 loop,

$$W_{\mu\nu}^{1 \times 1} = \text{Re Tr } (U_\mu(x) U_\nu(x + \hat{\mu}) U_\mu^\dagger(x + \hat{\nu}) U_\nu^\dagger(x)) \quad (6.7)$$

Unless otherwise specified, whenever I use the term Wilson loops I will assume that the real part of the trace, Re Tr , has been taken. For $SU(N \geq 3)$ the trace of any Wilson loop in the fundamental representation is complex, with the two possible path-orderings giving complex conjugate values. Thus, taking the trace insures gauge invariance, and taking the real part is equivalent to averaging the loop and its charge conjugate.

Exercise: Convince yourself that there do not exist any other independent gauge invariant quantities.

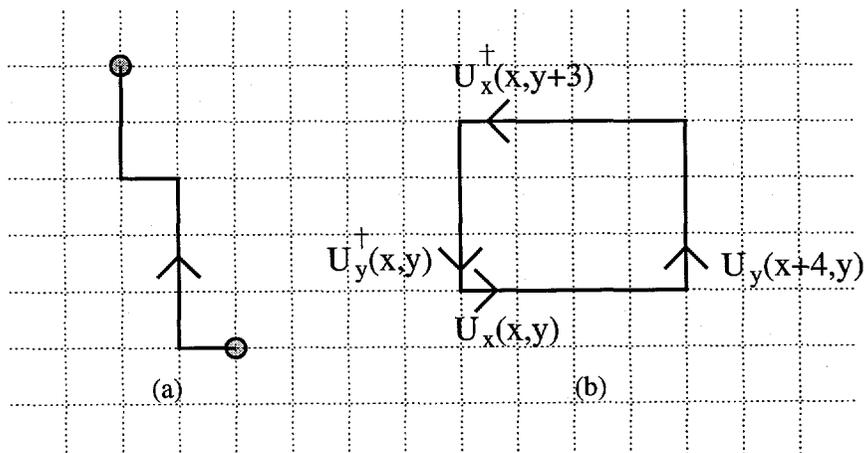


Fig. 5. The two gauge invariant quantities. a) An ordered string of U 's capped by a fermion and an anti-fermion and b) closed Wilson loops.

Exercise: Convince yourself that to preserve gauge invariance one has to formulate the lattice theory in terms of link variables U_μ rather than the A_μ directly.

A gauge invariant action, therefore, has to be built out of loops and strings. However, note that these two basic gauge invariant objects can be of arbitrary size and shape. Furthermore, they can be taken to lie in any representation of color $SU(3)$. The only necessary constraint is that it give the familiar continuum theory in the limit $a \rightarrow 0$. The reasons for wanting to preserve gauge invariance at all a is that otherwise one would have many more parameters to tune (the zero gluon mass, and the equality of the quark-gluon, 3-gluon, and 4-gluon couplings are consequences of gauge symmetry) and there would arise many more operators at any given order in a (these are irrelevant operators, but at finite a give rise to discretization errors as discussed in Section 8). In fact, one would like to preserve as many symmetries of the theory as possible at all values of a and not just in the continuum limit. It turns out to be relatively simple to maintain local gauge invariance, however, maintaining chiral symmetry, which has proven so important in the phenomenology of light quarks, and a one-to-one relation between Dirac and lattice fermion degrees of freedom has eluded us. This is discussed in Section 9. Also, discretization *de facto* breaks continuous rotational, Lorentz, and translational symmetry. The key idea of improved discretization schemes, to be discussed in Sections 8, 12, and

13, will be to implement the breaking of these symmetries by operators that come with sufficiently many powers of a so that their effect is very local, and thus small on physical (long-distance) quantities.

7. Simplest Formulation of the Lattice Action

7.1. Gauge Action

The gauge action can be expressed in terms of closed loops. I will outline the construction for the simpler case of the abelian $U(1)$ model in which the link variables are commuting complex numbers instead of $SU(3)$ matrices. Consider the simplest Wilson loop, the 1×1 plaquette $W_{\mu\nu}^{1 \times 1}$:

$$\begin{aligned} W_{\mu\nu}^{1 \times 1} &= U_\mu(x) U_\nu(x + \hat{\mu}) U_\mu^\dagger(x + \hat{\nu}) U_\nu^\dagger(x) \\ &= e^{iag[A_\mu(x + \frac{\hat{\mu}}{2}) + A_\nu(x + \hat{\mu} + \frac{\hat{\nu}}{2}) - A_\mu(x + \hat{\nu} + \frac{\hat{\mu}}{2}) - A_\nu(x + \frac{\hat{\nu}}{2})]} \end{aligned} \quad (7.1)$$

Expanding about $x + \frac{\hat{\mu} + \hat{\nu}}{2}$ gives

$$\begin{aligned} &\exp \left[ia^2 g (\partial_\mu A_\nu - \partial_\nu A_\mu) + \frac{ia^4 g}{12} (\partial_\mu^3 A_\nu - \partial_\nu^3 A_\mu) + \dots \right] \\ &= 1 + ia^2 g F_{\mu\nu} - \frac{a^4 g^2}{2} F_{\mu\nu} F^{\mu\nu} + O(a^6) + \dots \end{aligned} \quad (7.2)$$

The real and imaginary parts of the loop give

$$\begin{aligned} \text{Re Tr}(1 - W_{\mu\nu}^{1 \times 1}) &= \frac{a^4 g^2}{2} F_{\mu\nu} F^{\mu\nu} + \text{terms higher order in } a \\ \text{Im}(W_{\mu\nu}^{1 \times 1}) &= a^2 g F_{\mu\nu} . \end{aligned} \quad (7.3)$$

So far the indices μ and ν are uncontracted. There are 6 distinct positively oriented plaquettes, $\{\mu < \nu\}$, associated with each site. Summing over μ, ν , and taking care of the double counting by an extra factor of $\frac{1}{2}$, gives

$$\frac{1}{g^2} \sum_x \sum_{\mu < \nu} \text{Re Tr}(1 - W_{\mu\nu}^{1 \times 1}) = \frac{a^4}{4} \sum_x \sum_{\mu, \nu} F_{\mu\nu} F^{\mu\nu} \rightarrow \frac{1}{4} \int d^4 x F_{\mu\nu} F^{\mu\nu} .$$

Thus, to lowest order in a , the expansion of the plaquette gives the continuum action. The steps in the derivation for non-abelian groups are identical and the final expression (up to numeric factors) is the same. The result for the gauge action for $SU(3)$, defined in terms of plaquettes and called the Wilson action, is

$$S_g = \frac{6}{g^2} \sum_x \sum_{\mu < \nu} \text{Re Tr} \frac{1}{3} (1 - W_{\mu\nu}^{1 \times 1}) . \quad (7.4)$$

For historic reasons the lattice calculations are mostly presented in terms of the coupling $\beta \equiv 6/g^2$. Since β is used for many other quantities like $1/kT$ or the β -function, I shall try to be more explicit in the use of notation (g versus β) if there is a possibility for confusion.

Exercise: Repeat the Taylor expansion for the physical case of $SU(3)$. Show that the non-abelian term in $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + g[A_\mu, A_\nu]$ arises due to the non-commuting property of λ matrices. The details of the derivation for $SU(2)$ are given in section 3 of Ref. [8].

There are four important points to note based on the above construction of the lattice action.

- 1) The leading correction is $O(a^2)$. The term $\frac{a^2}{6} F_{\mu\nu} (\partial_\mu^3 A_\nu - \partial_\nu^3 A_\mu)$ is present in the expansion of all planar Wilson loops. Thus at the classical level it can be gotten rid of by choosing an action that is a linear combination of say 1×1 and 1×2 Wilson loops with the appropriate relative strength given by the Taylor expansion (see Section 12).
- 2) Quantum effects will give rise to corrections, *i.e.* $a^2 \rightarrow X(g^2)a^2$ where in perturbation theory $X(g^2) = 1 + c_1 g^2 + \dots$, and will bring in additional non-planar loops. Improvement of the action will consequently require including these additional loops, and adjusting the relative strengths which become functions of g^2 . This is also discussed in Section 12.
- 3) The reason for defining the action in terms of small loops is computational speed and reducing the size of the discretization errors. For example the leading correction to 1×1 loops is proportional to $a^2/6$ whereas for 1×2 loops it increases to $5a^2/12$. Also, the cost of simulation increases by a factor of 2–3.
- 4) The electric and magnetic fields E and B are proportional to $F_{\mu\nu}$. Eq. 7.3 shows that these are given in terms of the imaginary part of Wilson loops.

7.2. Fermion Action

To discretize the Dirac action, Wilson replaced the derivative with the symmetrized difference and included appropriate gauge links to maintain gauge invariance

$$\bar{\psi} \not{D} \psi = \frac{1}{2a} \bar{\psi}(x) \sum_{\mu} \gamma_{\mu} [U_{\mu}(x) \psi(x + \hat{\mu}) - U_{\mu}^{\dagger}(x - \hat{\mu}) \psi(x - \hat{\mu})]. \quad (7.5)$$

It is easy to see that one recovers the Dirac action in the limit $a \rightarrow 0$ by Taylor expanding the U_μ and $\psi(x + \hat{\mu})$ in powers of the lattice spacing a . Keeping only the leading term in a , Eq. 7.5 becomes

$$\begin{aligned} & \frac{1}{2a} \bar{\psi}(x) \gamma_\mu \left[\left(1 + iagA_\mu(x + \frac{\hat{\mu}}{2}) + \dots\right) (\psi(x) + a\psi'(x) + \dots) - \right. \\ & \quad \left. \left(1 - iagA_\mu(x - \frac{\hat{\mu}}{2}) + \dots\right) (\psi(x) - a\psi'(x) + \dots) \right] \\ &= \bar{\psi}(x) \gamma_\mu (\partial_\mu + \frac{a^2}{6} \partial_\mu^3 + \dots) \psi(x) \\ & \quad + i g \bar{\psi}(x) \gamma_\mu \left[A_\mu + \frac{a^2}{2} \left(\frac{1}{4} \partial_\mu^2 A_\mu + (\partial_\mu A_\mu) \partial_\mu + A_\mu \partial_\mu^2 \right) + \dots \right] \psi(x), \end{aligned}$$

which, to $O(a^2)$, is the kinetic part of the standard continuum Dirac action in Euclidean space-time. Thus one arrives at the simplest (called "naive") lattice action for fermions

$$\begin{aligned} \mathcal{S}^N &= m_q \sum_x \bar{\psi}(x) \psi(x) \\ & \quad + \frac{1}{2a} \sum_x \bar{\psi}(x) \gamma_\mu [U_\mu(x) \psi(x + \hat{\mu}) - U_\mu^\dagger(x - \hat{\mu}) \psi(x - \hat{\mu})] \\ & \equiv \sum_x \bar{\psi}(x) M_{xy}^N[U] \psi(y) \end{aligned} \quad (7.6)$$

where the interaction matrix M^N is

$$M_{i,j}^N[U] = m_q \delta_{ij} + \frac{1}{2a} \sum_\mu [\gamma_\mu U_{i,\mu} \delta_{i,j-\mu} - \gamma_\mu U_{i-\mu,\mu}^\dagger \delta_{i,j+\mu}] \quad (7.7)$$

The Euclidean γ matrices are hermitian, $\gamma_\mu = \gamma_\mu^\dagger$, and satisfy $\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}$. The representation I shall use is

$$\gamma = \begin{pmatrix} 0 & i\sigma \\ -i\sigma & 0 \end{pmatrix}, \quad \gamma_4 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (7.8)$$

which is related to Bjorken and Drell conventions as follows: $\gamma_i = i\gamma_{BD}^i$, $\gamma_4 = \gamma_{BD}^0$, $\gamma_5 = \gamma_{BD}^5$. In this representation γ_1, γ_3 are pure imaginary, while $\gamma_2, \gamma_4, \gamma_5$ are real.

The Taylor expansion showed that the discretization errors start at $O(a^2)$. For another simple illustration consider the inverse of the free-field propagator $m + i/a \sum_\mu \gamma_\mu \sin(p_\mu a)$. Set $p = 0$ and rotate to Minkowski

space ($p_4 \rightarrow iE$, i.e. $\sin p_4 a \rightarrow i \sinh Ea$). Then, using the forward propagator (upper two components of γ_4), gives

$$m_q^{pole} a = \sinh Ea \quad (7.9)$$

for the relation between the pole mass and the energy. This shows that, even in the free field case, the continuum relation $E(\mathbf{p} = 0) = m$ is violated by corrections that are $O(a^2)$.

Symmetries: The invariance group of the fermion action under rotations in space and time is the hypercubic group. Full Euclidean invariance will be recovered only in the continuum limit. The action is invariant under translations by a and under \mathcal{P} , \mathcal{C} , and \mathcal{T} as can be checked using Table 2.

The naive action $\bar{\psi}_x M_{xy}^N \psi_y$ has the following global symmetry:

$$\begin{aligned} \psi(x) &\rightarrow e^{i\theta} \psi(x) \\ \bar{\psi}(x) &\rightarrow \bar{\psi}(x) e^{-i\theta} \end{aligned} \quad (7.10)$$

where θ is a continuous parameter. This symmetry is related to baryon number conservation and leads to a conserved vector current. For $m_q = 0$ the action is also invariant under

$$\begin{aligned} \psi(x) &\rightarrow e^{i\theta \gamma_5} \psi(x) \\ \bar{\psi}(x) &\rightarrow \bar{\psi}(x) e^{i\theta \gamma_5} \end{aligned} \quad (7.11)$$

Having both a vector and an axial symmetry in a hard-cutoff regularization would imply a violation of the Adler-Bell-Jackiw theorem. It turns out that while the naive fermion action preserves chiral symmetry it also has the notorious fermion doubling problem as explained in Section 9. The chiral charges of these extra fermions are such as to exactly cancel the anomaly. In fact the analysis of \mathcal{S}^N lead to a no-go theorem by Nielsen-Ninomiya [14] that states that it is not possible to define a local, translationally invariant, hermitian lattice action that preserves chiral symmetry and does not have doublers.

We will discuss two fixes to the doubling/anomaly/chiral symmetry problem. First, include an additional term in the action that breaks chiral symmetry and removes the doublers (Wilson's fix). Second, exploit the fact that the naive fermion action has a much larger symmetry group, $U_V(4) \otimes U_A(4)$, to reduce the doubling problem from $2^d = 16 \rightarrow 16/4$ while maintaining a remnant chiral symmetry (staggered fermions). In both cases one regains the correct anomaly in the continuum limit. Before diving into these fixes, I would like to first discuss issues of the measure of integration and gauge fixing, and give an overview of Symanzik's improvement program.

7.3. The Haar Measure

The fourth ingredient needed to complete the formulation of LQCD as a quantum field theory via the path integral is to define the measure of integration over the gauge degrees of freedom. Note that, unlike the continuum fields A_μ , lattice fields are $SU(3)$ matrices with elements that are bounded in the range $[0, 1]$. Therefore, Wilson proposed an invariant group measure, the Haar measure, for this integration. This measure is defined such that for any elements V and W of the group

$$\int dU f(U) = \int dU f(UV) = \int dU f(WU) \quad (7.12)$$

where $f(U)$ is an arbitrary function over the group. This construction is simple, and has the additional advantage that in non-perturbative studies it avoids the problem of having to include a gauge fixing term in the path integral. This is because the field variables are compact. Hence there are no divergences and one can normalize the measure by defining

$$\int dU = 1. \quad (7.13)$$

For gauge invariant quantities each gauge copy contributes the same amount. Consequently the lack of a gauge fixing term just gives an overall extra numerical factor. This cancels in the calculation of correlation functions due to the normalization of the path integral. Note that gauge fixing and the Fadeev-Papov construction does have to be introduced when doing lattice perturbation theory [36].

The measure and the action are invariant under \mathcal{C} , \mathcal{P} , \mathcal{T} transformations, therefore when calculating correlation functions via the path integral we must average over a configuration U and the ones obtained from it by the application of \mathcal{C} , \mathcal{P} , \mathcal{T} . In most cases this averaging occurs very naturally, for example, Euclidean translation invariance guarantees \mathcal{P} and \mathcal{T} invariance. Similarly, taking the real part of the trace of a Wilson loop is equivalent to summing over U and its the charge conjugate configuration. A cook-book recipe for evaluating the behavior of correlation functions after average over configurations related by the lattice discrete symmetries is given in Section 17.

One important consequence of these symmetries is that only the real part of Euclidean correlation functions has a non-zero expectation value. This places restriction on the quantities one can calculate using LQCD. For example, simulations of LQCD have not been very successful at calculations of scattering amplitudes, which are in general complex [48]. The connection

between Minkowski and Euclidean correlation functions was discussed in Section 5. To get scattering amplitudes what is required is to first calculate Euclidean amplitudes for all values of the Euclidean time τ , followed by a fourier transform in τ , and finally a Wick rotation to Minkowski space. For such an procedure to work requires data with phenomenal accuracy. Unfortunately, simulations give Euclidean correlation functions for a discrete set of τ and with errors that get exponentially magnified in the rotation. The best that has been done is to use Lüscher's method that relates scattering phase shifts to shifts in two particle energy states in finite volume [27,28]. So far this method has been tested on models in 2-4 dimensions, and used to calculate scattering lengths in QCD [29,30].

7.4. Gauge Fixing and Gribov Copies

Gauge invariance and the property of group integration, $\int dU U = 0$, are sufficient to show that only gauge invariant correlation functions have non-zero expectation values. This is the celebrated Elitzur's theorem [37]. It states that a local continuous symmetry cannot be spontaneously broken.

In certain applications (study of quark and gluon propagators [38], 3-gluon coupling [39], hadronic wavefunctions [40], and determination of renormalization constants using quark states [41]) it is useful to work in a fixed gauge. The common gauge choices are axial ($\eta_\mu A_\mu = 0$ where η in Minkowski space is a fixed time-like vector. The common choice is $A_4 = 0$), Coulomb ($\partial_i A_i = 0$), and Landau ($\partial_\mu A_\mu = 0$). Their lattice analogues, assuming periodic boundary conditions, are the following.

Maximal Axial Gauge: On a lattice with periodic boundary conditions one cannot enforce $A_4 = 0$ by setting all links in the time direction to unity as the Wilson line in each direction is an invariant. The gauge freedom at all but one site can be removed as follows. Set all time-like links $U_4 = 1$ on all but one timeslice. On that time-slice all U_4 links are constrained to be the original Wilson line in 4-direction, however, one can set links in \hat{z} direction $U_3 = 1$ except on one z -plane. On that z -plane set U_3 links to the Wilson line in that direction and all links in \hat{y} direction $U_2 = 1$ except on one y -line. On this y -line set all U_2 links to the value of y -lines and links in \hat{x} direction $U_1 = 1$ except on one point. While such a construction is unique, the choice of the time-slice, z -plane, y -line, and x -point and the labeling of x, y, z, t axis are not. Also, this gauge fixing does not respect Euclidean translation invariance and it is not smooth.

Coulomb Gauge: The lattice condition is given by the gauge transfor-

mation $V(x)$ that maximizes the function

$$F[V] = \sum_x \sum_i \text{Re Tr} V(x) \left(U_i(x) V^\dagger(x+i) + U_i^\dagger(x-i) V^\dagger(x-i) \right) \quad (7.14)$$

separately on each time-slice and i runs over the spatial indices 1 – 3.

Landau Gauge: This is defined by the same function as in Eq. 7.14 except that the sum is evaluated on the whole lattice and $i = 1 - 4$.

The maximization condition, Eq. 7.14, corresponds, in the continuum, to finding the stationary points of $F[V] = \|A^V\| = \int d^4x \text{Tr}(A_\mu^V)^2(x)$. What we would like is to find the global minimum corresponding to the “smoothest” fields A_μ , or the copies in case it is not unique. For any gauge configuration there is an infinite set of gauge equivalent configurations obtained by applying the set of transformations $\{V(x)\}$. This set is called the gauge orbit. The first question is whether the gauge fixing procedure gives a unique solution, *i.e.* whether the gauge-fixing algorithm converges to the same final point irrespective of the starting point on the gauge orbit. For the maximal axial gauge, the construction is unique (the configuration $\{U\}$ and $\{VUV^\dagger\}$ for any $\{V\}$ give the same final link matrices), so the answer is clearly yes. For the Coulomb and Landau gauge choices the answer is no – there are a number of solutions to the gauge-fixing condition. These are the famous Gribov copies [42]. For a recent theoretical review of this subject see [43].

The algorithms for fixing to either the Coulomb gauge or the Landau gauge are mostly local. Local algorithms generate, for each site, a gauge matrix $v_i(x)$ that maximizes the sum, defined in Eq. 7.14, of the link matrices emerging from that site. This gauge transformation is then applied to the lattice. These two steps are iterated until the global maximum is found. The ordered product $\prod_i v_i(x)$ gives $V(x)$. For all practical purposes we shall call this end-point of the gauge fixing algorithm a Gribov copy. It is, in most cases and especially for the local gauge-fixing algorithms, the nearest extremum, and not the global minimum of $\|A^V\|$. A set of Gribov copies can be generated by first performing a different random gauge transformation on the configuration and then applying the same gauge-fixing algorithm, or by using different gauge-fixing algorithms. For recent studies of the efficacy of different algorithms see [44,45].

The Monte Carlo update procedure is such that it is highly improbable that any finite statistical sample contains gauge equivalent configurations. So one does not worry about gauge copies. In the gauge-fixed path integral one would *a priori* like to average over only one Gribov copy per configuration and choose this to be the “smoothest”. Finding the $V(x)$ which gives the smoothest configuration has proven to be a non-trivial problem.

The open questions are: (i) how to define the “smoothest” solution of the gauge-fixing algorithm (*i.e.* defining the fundamental modular domain), (ii) how to ascertain that a solution is unique (lies within or on the boundary of the fundamental modular domain), and (iii) how to find it.

The ultimate question is – what is the dependence of the quantities one wants to calculate on the choice of Gribov copy? A number of studies have been carried out [46], however the results and their interpretation are not conclusive. Firstly, it is not known whether the smoothest solution lying within the fundamental modular domain can be found economically for all configurations and over the range of gauge coupling being investigated. In the absence of the smoothest solution it is not clear whether (i) the residual gauge freedom simply contributes to noise in the correlation function, and (ii) one should average over a set of Gribov copies for each configuration to reduce this noise. In my opinion these remain challenging problems that some of you may wish to pursue.

The problem of Gribov copies and the Faddeev-Papov construction (a gauge fixing term and the Faddeev-Papov determinant) resurfaces when one wants to do lattice perturbation theory or when taking the Hamiltonian limit of the path integral via the transfer matrix formalism. This construction is analogous to that in the continuum [36]. Finally, let me mention that in a fixed gauge, the lattice theory has only a BRST invariance [27]. While BRST invariance is sufficient to prove that the theory is renormalizable, the set of operators of any given dimension allowed by BRST invariance is larger than that with full gauge invariance. In the analyses of improvement of the action and operators it is this enlarged set that has to be considered [47].

8. Continuum limit and Symanzik's Improvement program

The Taylor expansion of the simple Wilson action for the gauge and fermion degrees of freedom showed the presence of higher dimensional operators. Since these are suppressed by powers of the lattice spacing, they are irrelevant operators in the language of the renormalization group, *i.e.* they vanish at the fixed point at $a \rightarrow 0$. Another way of expressing the same thing is as follows. The lattice theory with a hard cut-off at π/a can be regarded as an effective theory [49]. Integration of momenta from $\infty \rightarrow \pi/a$ generates effective interactions. Thus, in general, we can write the action as

$$\mathcal{S}_C = \mathcal{S}_L(\pi/a) + \sum_n \sum_i a^n C_i^n \mathcal{A}_i^n(\pi/a) \quad (8.1)$$

where $S_C(S_L)$ is the continuum (lattice) action, and the first sum is over all operators of a given dimension, and the second over all dimensions greater than four. Similarly, the operators used to probe the physics can be written as

$$\mathcal{O}_C = \mathcal{O}_L(\pi/a) + \sum_n \sum_i a^n D_i^n \mathcal{O}_i^n(\pi/a) \quad (8.2)$$

where in this case quantum corrections can induce mixing with lower dimension operators also. The coefficients C_i^n and D_i^n are both functions of the coupling α_s . Parenthetically, one important reason for preserving gauge, Lorentz, and chiral symmetries at finite a is that they greatly restrict the set of possible operators of a given dimension.

Now consider the expectation value of a given operator

$$\begin{aligned} \langle \mathcal{O}_C \rangle &= \int dU \mathcal{O}_C e^{-S_C} \\ &= \int dU \{ \mathcal{O}_L + \sum_n \sum_i a^n D_i^n \mathcal{O}_i^n \} e^{-S_L - \sum_m \sum_i a^m C_i^m \mathcal{A}_i^m} \\ &= \int dU \{ \mathcal{O}_L + \sum_n \sum_i a^n D_i^n \mathcal{O}_i^n \} \{ 1 - \sum_m \sum_i a^m C_i^m \mathcal{A}_i^m + \dots \} e^{-S_L} \\ &= \int dU \mathcal{O}_L e^{-S_L} \\ &\quad + \int dU (\mathcal{O}_L \{ \sum_m \sum_i a^m C_i^m \mathcal{A}_i^m + \dots \}) e^{-S_L} \\ &\quad + \int dU (\sum_n \sum_i a^n D_i^n \mathcal{O}_i^n \{ 1 - \sum_m \sum_i a^m C_i^m \mathcal{A}_i^m + \dots \}) e^{-S_L}. \end{aligned} \quad (8.3)$$

Note that the operators \mathcal{A}_i^m , which appear in the action, are summed over all space-time points. This leads to contact terms in Eq. 8.3 which I shall, for brevity, ignore as they necessitate a proper definition of the operators but do not change the conclusions.

The basis for believing that LQCD provides reliable results from simulations at finite a is that the contributions of the lattice artifacts (last two terms in the last expression) are small, can be estimated, and thus removed. At this point I will only make a few general statements about the calculation of expectation values which will be embellished on in later lectures.

– The contribution of the last two terms in Eq. 8.3 vanishes in the limit $a \rightarrow 0$ unless there is mixing with operators of lower dimension. In such cases

the mixing coefficient has to be determined very accurately (necessitating non-perturbative methods) otherwise the $a \rightarrow 0$ limit is divergent.

– Calculations have to be done at values of a at which contributions of irrelevant terms organized in successive powers in a decrease at least geometrically.

– To achieve continuum results simulations are carried out at a sufficient number of values of a to allow reliable extrapolation to $a = 0$. The “fineness” of the lattice spacing a at which to do the calculations and the number of points needed depends on the size of the corrections and how well we can determine the functional form to use in the extrapolation.

– To improve the result to any given order in a , one has to improve both the action and the operators to the same order. Such an approach, based on removing lattice artifacts organized as a power series expansion in a , is called the Symanzik improvement program.

Eq. 8.3 shows clearly that there is fair degree of flexibility in constructing LQCD. We are free to add any irrelevant operator with a sensible strength and still recover QCD in the continuum limit. The only difference between various constructions of the action is the approach to the continuum limit. What we would like to do is improve the action and operators so as to get continuum results on as coarse lattices as possible. On the other hand improvement requires adding irrelevant terms to both the action and the operators. This increases the complexity of the calculations and the simulation time. Thus, the strategy for minimizing discretization errors is a compromise – optimize between simplicity of the action and operators (defined in terms of the cost of simulation, code implementation, and analyses) and the reduction of the discretization errors evaluated at some given fixed value of the lattice spacing, say $a \approx 0.1$ fermi. The remaining errors, presumably small, can then be removed by extrapolating the results obtained from simulations at a few values of a to $a = 0$.

The number of floating points operations needed in the simulations of LQCD scale roughly as L^6 in the quenched approximation and L^{10} with dynamical fermions. Current lattice simulations of QCD are done for lattice spacing a varying in the range $2 \text{ GeV} \leq a^{-1} \leq 5 \text{ GeV}$. For these parameter values the corrections due to the $O(a)$, $O(ma)$, $O(pa)$, $O(a^2)$, ... terms in the Wilson action are found to be large in many observables. The importance of improving the lattice formulation (action and operators) is therefore self-evident, especially for full QCD. Thus, the search for such actions is a hot topic right now. Before discussing these let me first summarize what one hopes to gain by, and the criteria by which to judge, such an improvement program.

1. Improved scaling, *i.e.* discretization errors are reduced. Consequently one can work on coarser lattices (smaller $L = N_S$) for a given accuracy. Even a factor of 2 reduction in L translates into $2^{10} \approx 1000$ in computer time!
2. Better restoration of rotational and internal symmetries like chiral, staggered flavor, *etc.*.
3. The trajectory along which the continuum limit is taken should not pass close to extraneous phase transitions that are artifacts of the lattice discretization (see Section 15 for a discussion of phase transitions in the lattice theory). The problem with such transitions is that in their vicinity the desired continuum scaling behavior can be modified for certain observables, *i.e.* the artifact terms in Eq. 8.3 may not have a geometric convergence in a and may thus be large.
4. The study of topology on the lattice is particularly sensitive to short distance fluctuations that are artifacts of discretizing the gauge action. For the Wilson gauge action one finds that the topological charge, calculated using Lüscher's geometric method [50], jumps from $0 \rightarrow 1$ at a core radius of the instanton of $\rho_c a \approx 0.7$, whereas the action of such instantons, S_{ρ_c} , is significantly less than the classical value $S^{cl} = 8\pi^2/\beta$. The entropy factor for such small instantons with $S < S^{cl}$ overwhelms the Boltzmann suppression, leading to a divergent topological susceptibility in the continuum limit. These short distance artifacts are called dislocations and it is not *a posteriori* obvious how to separate the physical from the unphysical when $\rho \approx \rho_c$. It is, therefore, very important to start with a lattice formulation in which such artifacts are excluded, and there exist only physical instantons, *i.e.* those with mean core size that has the correct scaling behavior with g . One finds that dislocations are suppressed by improved actions, *i.e.* improving the gauge action increases S_{ρ_c} [51] and better definition of the topological charge [52].
5. Lattice perturbation theory becomes increasingly complicated as more and more irrelevant operators are added to the action. Even though, with the advent of non-perturbative methods for determining renormalization constants, the reliance on lattice perturbation theory has decreased, perturbative analyses serve as a valuable guide and check, so we would like to retain this possibility. What one expects for an improved action is that the lattice and continuum value of α_s , and the coefficients of the series expansion in α_s for a given observable, are much

closer. In that case a much more reliable matching between the two theories is possible. At present the 1-loop matching between the lattice and continuum theories is one of the larger sources of uncertainties.

6. The Adler-Bell-Jackiw (ABJ) Anomaly: In the continuum the flavor singlet axial current has an anomaly. This shows up as part of the UV regularization scheme. In the Pauli-Villars scheme one introduces a heavy mass as the cutoff. This breaks γ_5 invariance, while in dimensional regularization, γ_5 is not well-defined away from 4 dimensions. Naive fermions, as shown below, preserve the γ_5 invariance but at the expense of flavor doubling. A consequence of the doublers is a pairwise cancellation and no net anomaly. A viable lattice theory should reproduce the ABJ anomaly.

Before discussing such “improved” actions, it is instructive to revisit the construction of the naive Dirac action to highlight its problems.

9. Fermion Doubling Problem in the “naive” action

The problem with the naive discretization of the Dirac action, specified by the operator M^N in Eq. 7.7, is that in the continuum limit it gives rise to $2^d = 16$ flavors rather than one. This doubling problem is readily demonstrated by the inverse of the free field propagator (obtained by taking the Fourier transform of the action with all $U_\mu(x) = 1$).

$$S^{-1}(p) = m_q + \frac{i}{a} \sum_{\mu} \gamma_{\mu} \sin p_{\mu} a \quad (9.1)$$

which has 16 zeros within the Brillouin cell in the limit $m_q \rightarrow 0$. Defining the momentum range of the Brillouin cell to be $\{-\pi/2, 3\pi/2\}$, the zeros lie at $p_{\mu} = 0$ and π . As discussed later, this proliferation holds under very general conditions specified by the Nielsen-Ninomiya theorem [14], and is intimately related to simultaneously preserving chiral symmetry. The inclusion of the gauge fields does not solve the doubling problem.

Before discussing partial fixes, it is useful to investigate the properties of these extra zero modes under chiral transformations. Let us define a set of 16 4-vectors $\Pi^A = \{(0, 0, 0, 0), (\pi, 0, 0, 0), \dots, (\pi, \pi, \pi, \pi)\}$ with $A = \{1 \dots 16\}$, and consider the expansion of the massless propagator about these points. Then

$$S^{-1}(p, m = 0) = \frac{i}{a} \sum_{\mu} \gamma_{\mu} \sin p_{\mu} a$$

Table 3

The chiral sign and degeneracy factor for the doublers.

A	Degeneracy	X^A
(0, 0, 0, 0)	1	+1
(1, 0, 0, 0)	4	-1
(1, 1, 0, 0)	6	+1
(1, 1, 1, 0)	4	-1
(1, 1, 1, 1)	1	+1

$$\begin{aligned}
&= \frac{i}{a} \sum_{\mu} \gamma_{\mu} \sin(\Pi^A + k)_{\mu} a \\
&= \frac{i}{a} \sum_{\mu} \gamma_{\mu} \mathcal{S}_{\mu}^A \sin k_{\mu} a \\
&\equiv \frac{i}{a} \sum_{\mu} \tilde{\gamma}_{\mu} \sin k_{\mu} a, \tag{9.2}
\end{aligned}$$

where $\mathcal{S}_{\mu}^A = \{+1, -1\}$ depending on whether the μ component of Π^A is $\{0, \pi\}$. In the last expression a new representation of gamma matrices has been defined by the similarity transformation $\tilde{\gamma}_{\mu} = Y_A \gamma_{\mu} Y_A^{\dagger}$ where

$$Y_A = \prod_{\mu} (\gamma_{\mu} \gamma_5)^{n_{\mu}^A} \tag{9.3}$$

and n_{μ}^A are $\{0, 1\}$ depending on whether the momentum expansion in the μ direction is about 0 or π . Now, $\tilde{\gamma}_5 = Y_A \gamma_5 Y_A^{\dagger} = \mathcal{S}_1^A \mathcal{S}_2^A \mathcal{S}_3^A \mathcal{S}_4^A \gamma_5 \equiv X^A \gamma_5$, with the signs X^A given in table 3. Thus, the sixteen species break up into two sets of 8 with chiral charge ± 1 , and render the theory anomaly-free. Consequently, this "naive" discretization is phenomenologically not acceptable. The presence of doublers can, in fact, be traced to a larger flavor symmetry of the naive fermion action as I now discuss.

In the continuum the mass-less Dirac action is invariant under the $U(1)_V \otimes U(1)_A$ flavor transformations

$$\begin{aligned}
\psi_L &\rightarrow V_L \psi_L, & \bar{\psi}_L &\rightarrow \bar{\psi}_L V_L^{\dagger} \\
\psi_R &\rightarrow V_R \psi_R, & \bar{\psi}_R &\rightarrow \bar{\psi}_R V_R^{\dagger}.
\end{aligned} \tag{9.4}$$

Since the left and right handed fields are only coupled through the mass term, the vector symmetry $V_L = V_R = e^{i\theta}$ holds for all m , whereas the

axial symmetry $V_L = V_R^\dagger = e^{i\theta}$ (or equivalently $V_L = V_R = e^{i\theta\gamma_5}$) holds only for $m = 0$. The naive fermion action \mathcal{S}_N for a single flavor has a much larger symmetry group, $U(4)_V \otimes U(4)_A$, under

$$\begin{aligned}\psi(x) &\rightarrow \Gamma_x \left(\frac{1}{2}(1 - \epsilon(x)\gamma_5)V_L + \frac{1}{2}(1 + \epsilon(x)\gamma_5)V_R \right) \Gamma_x^\dagger \psi(x) \\ \bar{\psi}(x) &\rightarrow \bar{\psi}(x) \Gamma_x \left(V_L^\dagger \frac{1}{2}(1 + \epsilon(x)\gamma_5) + V_R^\dagger \frac{1}{2}(1 - \epsilon(x)\gamma_5) \right) \Gamma_x^\dagger\end{aligned}\quad (9.5)$$

if one considers translations by $2a$. Γ_x is given in Eq. 11.1, the phase $\epsilon(x)$ in Eq. 11.4, and $\Gamma_x \gamma_5 \Gamma_x^\dagger = \epsilon(x)\gamma_5$. Here $V_{L,R}$ are general 4×4 unitary matrices acting on the spinor index of the fermions. Even though there is, in Eq. 9.5, a dependence on x through $\epsilon(x)$, this is a global symmetry as specifying it at one point fixes it at all points.

To see how doubling arises, consider the vector ($V_L = V_R$) discrete subgroup with the thirty two elements $\{\pm Y_A\}$ defined in Eq. 9.3. Eq. 9.2 shows that the Y_A shift the momenta by π/a along the direction specified by A ,

$$Y_A S_F(k) Y_A^\dagger = S_F(k + \frac{\pi}{a} A). \quad (9.6)$$

Since this is a similarity transform, the sixteen regions of the Brillouin zone related by this transformation are physically equivalent. This analysis, which relied only on the spinor structure of the action, makes it easy to see why the doublers exist even for the interacting theory. In Section 11 this symmetry will be exploited to construct staggered fermions. In this formulation the sixteen doublers are reduced to four, and at the same time a $U(1)_V \otimes U(1)_\epsilon$ symmetry (corresponding to when $V_{L,R}$ are just phases) is retained. The $U(1)_\epsilon$ plays the role of a chiral symmetry at finite a and in the continuum limit becomes the $U(1)_A$.

9.1. Generalities of Chiral Symmetry

In the massless limit of the continuum theory, the flavor symmetry at the classical level is $U_V(1) \times U_A(1) \times SU_V(n_f) \times SU_A(n_f)$, where n_f is the number of light fermions. The unbroken $U_V(1)$ gives rise to baryon number conservation. The $U_A(1)$ is broken by instanton contributions and the flavor singlet axial current is anomalous. The $SU_A(n_f)$ is spontaneously broken and the quark condensates acquire a non-zero expectation value in the QCD vacuum. Associated with this spontaneously broken symmetry are $(n_f^2 - 1)$ goldstone bosons, *i.e.* the pions. Chiral symmetry has played a

very useful role in the continuum theory. For example, one can (i) classify operators under distinct representations of $SU_L(n_f) \times SU_R(n_f)$, and chiral symmetry prohibits mixing between operators in different representations, (ii) derive relations between matrix elements of a given operator between states with different number of pions in the final state (for example $K \rightarrow \pi\pi \leftrightarrow K \rightarrow \pi\pi\pi$), (iii) exploit the operator Ward identity, and (iv) derive relations between correlation functions, *i.e.* the Ward-Takahashi identities. It is therefore desirable to preserve chiral symmetry at finite a and not just in the continuum limit. To determine if this is possible one needs to understand the relationship between doublers and chiral symmetry in the lattice theory.

The chiral symmetry defined in Eq. 7.11 is realized by the lattice theory provided $\gamma_5 M + M \gamma_5 = 0$. It is easy to verify that the naïve lattice action satisfies the hermiticity property

$$\gamma_5 M \gamma_5 = M^\dagger \quad (9.7)$$

Thus, in the massless limit, the Euclidean lattice action has to be anti-hermitian for chiral symmetry, Eq. 7.11, to hold. This is true of M^N given in Eq. 7.7, and was accomplished by simply taking the symmetric difference for the derivative. The important question is – what are the general conditions under which a lattice theory with this realization of chiral symmetry exhibits doublers. The answer is given by the Nielsen-Ninomiya theorem [14]. Consider a generalized action such that $S^{-1}(p, m = 0) = iF(p)$. Then, if

- the function $F(p)$ is periodic in momentum space with period $2\pi/a$. (This is a consequence of *translation invariance*.)
- the lattice momenta are continuous in the range $\{0, 2\pi\}$. This is true in the $L \rightarrow \infty$ limit at all values of the lattice spacing.
- $F(p)$ is continuous in momentum space. This is guaranteed if the interactions are local.
- $F(p) \rightarrow p_\mu \gamma_\mu$ for small p_μ and as $a \rightarrow 0$ to match the continuum theory.
- the action has chiral symmetry, *i.e.* it satisfies the hermiticity property $\gamma_5 M \gamma_5 = M^\dagger$ and the massless Dirac action is anti-hermitian.

the theory will have doublers. It is easy to understand this from the schematic one dimensional plot of $F(p)$ shown in Fig. 6 which highlights the fact that $F(p)$ has to have an even number of zeros under the above

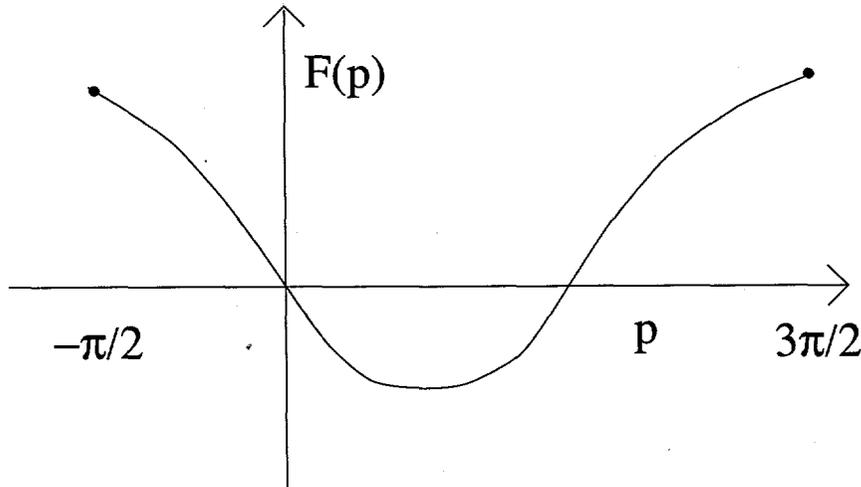


Fig. 6. A schematic 1-d representation of the massless inverse propagator $F(p)$.

conditions. The \pm slopes imply that one state is left moving while the other is right moving. In 1-dimension the direction is the same as handedness, consequently left-handed and right-handed fields come in inseparable pairs, and together form a Dirac fermion.

The key consequence of the Nielsen-Ninomiya theorem is that one cannot define a lattice theory with chiral fermions. For example if one writes down M^N for a left-handed Dirac fermion, then its doubler will be right-handed but have the same transformation under $L \otimes R$. Consequently, the two "states" coexist to form a Dirac fermion. Assuming that the Nielsen-Ninomiya theorem is air-tight, then how does one proceed to (i) define a lattice theory without doublers and which recovers chiral symmetry in the continuum limit, and (ii) define chiral gauge theories on the lattice?

The answer to the first question will be given by explicit reconstruction of Wilson's and staggered fixes. An alternate very exciting new development, which I will not do justice to, is the resurrection [53] of the Ginsparg-Wilson condition for a chirally symmetric formulation [54]

$$\gamma_5 D + D \gamma_5 = a D \gamma_5 D. \quad (9.8)$$

Lüscher has recently presented a clear analysis of how this condition allows the construction of a lattice theory without doublers and with chiral symmetry, avoids the NN theorem, and gives the ABJ anomaly [55]. Unfortunately, the bottleneck with this alternative is the lack of a closed form

expression for the lattice Dirac action that is sufficiently local to make simulations feasible.

The second question – construction of chiral gauge theories – is still unresolved. For those interested in this very important unsolved problem, a good starting point are the reviews [15–17] which contain some of the recent efforts.

10. Wilson Fermions (WF)

In Section 8, it was shown that the lattice action is not unique. One has the freedom to add an arbitrary number of irrelevant operators to the action, as these do not change the continuum limit. Wilson's solution to the doubling problem was to add a dimension five operator $ar\bar{\psi}\square\psi$, whereby the extra fifteen species at $p_\mu = \pi$ get a mass proportional to r/a [56]. The Wilson action is

$$\begin{aligned}
A^W &= m_q \sum_x \bar{\psi}(x)\psi(x) \\
&+ \frac{1}{2a} \sum_{x,\mu} \bar{\psi}(x)\gamma_\mu[U_\mu(x)\psi(x+\hat{\mu}) - U_\mu^\dagger(x-\hat{\mu})\psi(x-\hat{\mu})] \\
&- \frac{r}{2a} \sum_{x,\mu} \bar{\psi}(x)[U_\mu(x)\psi(x+\hat{\mu}) - 2\psi(x) + U_\mu^\dagger(x-\hat{\mu})\psi(x-\hat{\mu})] \\
&= \frac{(m_q a + 4r)}{a} \sum_x \bar{\psi}(x)\psi(x) \\
&+ \frac{1}{2a} \sum_x \bar{\psi}[(\gamma_\mu - r)U_\mu(x)\psi(x+\hat{\mu}) - (\gamma_\mu + r)U_\mu^\dagger(x-\hat{\mu})\psi(x-\hat{\mu})] \\
&\equiv \sum_{x,y} \bar{\psi}_x^L M_{xy}^W \psi_y^L \tag{10.1}
\end{aligned}$$

where the interaction matrix M^W is usually written as

$$M_{x,y}^W[U]a = \delta_{xy} - \kappa \sum_\mu [(r-\gamma_\mu)U_{x,\mu}\delta_{x,y-\mu} + (r+\gamma_\mu)U_{x-\mu,\mu}^\dagger\delta_{x,y+\mu}] \tag{10.2}$$

with the rescaling

$$\begin{aligned}
\kappa &= 1/(2m_q a + 8r) \\
\psi^L &= \sqrt{m_q a + 4r} \psi = \psi/\sqrt{2\kappa}. \tag{10.3}
\end{aligned}$$

In this form the Wilson action has the following interpretation as a Statistical Mechanics model. The local term acts like an inertia keeping the fermion at the same site while the non-local term makes the fermion hop to the nearest neighbor site with strength κ . During this hop the fermion acquires a twist by $(\gamma_\mu - r)$ in spin space and by U_μ in color space. The case $r = 1$ is special as the spin twist $1 \pm \gamma_\mu$ is a projection operator of rank 2,

$$\begin{aligned} \left(\frac{1 \pm \gamma_\mu}{2}\right)^2 &= \left(\frac{1 \pm \gamma_\mu}{2}\right) \\ \text{Tr}\left(\frac{1 \pm \gamma_\mu}{2}\right) &= 2. \end{aligned} \quad (10.4)$$

Later we shall see that for $r = 1$ the dispersion relation has a single branch, *i.e.* there are no doublers in the free field limit.

From Eq. 10.3, one notes that for the free theory the quark mass is given in terms of the lattice parameters κ and r as

$$m_q a = \frac{1}{2\kappa} - 4r \equiv \frac{1}{2\kappa} - \frac{1}{2\kappa_c} \quad (10.5)$$

with a zero at $\kappa = \kappa_c \equiv 1/8r$. For the interacting theory ($U_\mu(x) \neq 1$) we will continue to define $m_q a = 1/2\kappa - 1/2\kappa_c$ with the proviso that κ_c depends on a . The renormalization of κ_c independent of κ implies that the quark mass has both multiplicative and additive renormalizations. This is a consequence of the explicit breaking of chiral symmetry by the irrelevant term proportional to r . Wilson's fix for doublers comes with a high price tag - a hard breaking of chiral symmetry at $O(a)$.

The free propagator in momentum space for Wilson fermions is

$$S_F(p) = M_W^{-1}(p) = \frac{a}{1 - 2\kappa \sum_\mu (r \cos p_\mu a - i\gamma_\mu \sin p_\mu a)}. \quad (10.6)$$

Due to the dimension 5 operator (terms proportional to r) the 15 extra states at $p_\mu a = \pi$ get masses of order $2rn/a$, where n is the number of components of $p_\mu a = \pi$.

The operator M satisfies the following relations:

$$\begin{aligned} \gamma_5 M_W^\dagger \gamma_5 &= M_W, \\ \gamma_5 S_F^\dagger(x, y) \gamma_5 &= S_F(y, x), \\ M_W^\dagger(\kappa, r) &= M_W(-\kappa, -r). \end{aligned} \quad (10.7)$$

The first two state that the "hermiticity" property of M^N is preserved. The second equation relates a quark propagator from $x \rightarrow y$ to the antiquark

propagator from $y \rightarrow x$. This important identity, called hermiticity or γ_5 invariance, leads to a significant simplification in numerical calculations as discussed in Section 17.1. The adjoint in S_F^\dagger is with respect to the spin and color indices at each site. The last relation shows that M^W is not anti-hermitian due to the Wilson r term.

An analysis similar to Eq. 7.9 for the naive fermion action, shows that the pole mass derived from Eq. 10.6 is different from the bare mass and given by

$$m_q^{pole} a = r(\cosh Ea - 1) + \sinh Ea. \quad (10.8)$$

This shows, as expected, that the discretization corrections in spectral quantities occur at $O(a)$.

10.1. Properties of Wilson Fermions

I give a brief summary of the important properties of Wilson fermions.

- The doublers are given a heavy mass, $2rn/a$, and decouple in the continuum limit.
- Chiral symmetry is broken explicitly. The derivation of axial Ward identities, using the invariance under the transformation given in Eq. 7.11, have the generic form

$$\left\langle \frac{\partial \delta \mathcal{S}}{\partial \theta} \mathcal{O} \right\rangle = \left\langle \frac{\partial \delta \mathcal{O}}{\partial \theta} \right\rangle. \quad (10.9)$$

For WF, the variation of the action under an axial transformation is

$$\frac{\partial \delta \mathcal{S}}{\partial \theta} = \partial_\mu A_\mu - 2mP + raX, \quad (10.10)$$

where X is an additional term coming from the variation of the Wilson r term [57]. Thus, in general, all relations based on axial WI will have corrections at order ra and involve mixing with operators that would normally be absent due to the chiral symmetry.

- Calculation of the matrix element $\langle 0 | T[A_\mu V_\nu V_\rho] | 0 \rangle$ shows that the ABJ anomaly is correctly reproduced in the continuum limit [58]. Karsten and Smit show, by an explicit 1-loop calculation, that while each of the extra 15 states contribute terms proportional to r , the total contribution from all sixteen states is independent of r and equals the ABJ anomaly.

- The quark mass gets both additive and multiplicative renormalization.
- The zero of the quark mass is set by κ_c in Eq. 10.5. There are two ways to calculate κ_c at any given a . (i) Assume the chiral relation $M_\pi^2 \propto m_q$, calculate the pion mass as a function of $1/2\kappa$, and extrapolate it to zero. The value of κ at which the pion becomes massless is, by definition, κ_c . (ii) Calculate the quark mass through the ratio $\langle \partial_\mu A_\mu(x) P(0) \rangle / \langle P(x) P(0) \rangle$ (based on the axial Ward identity) as a function of $1/2\kappa$ and extrapolate to zero. The two estimates can differ by corrections of $O(a)$. Both these calculations requires a statistical average over gauge configurations and an extrapolation in $1/2\kappa$. Consequently, on an individual configuration the zero mode of the matrix M will not occur at κ_c but will have a distribution about κ_c .

In the full theory $\det M$ suppresses configurations with zero modes, however, this protection is lost in QCD. Nevertheless, the problem of zero modes at $\kappa < \kappa_c$ on certain configurations, called “exceptional” configurations, needs to be addressed in both theories.

Near these zero-modes the quark propagator becomes singular even though one is in the physical region $\kappa < \kappa_c$. Consequently, in a large enough statistical sample one will hit “exceptional” configurations, and hadronic correlations functions on these will show fluctuations by orders of magnitude (depending on small the quark mass is) from the rest. Since this pathology is a feature of Wilson-like actions and not of the gauge configurations themselves, one cannot simply remove these configurations from the statistical sample. The problem of such zero-modes increases with decreasing quark mass and β . Two practical approaches used so far have been to (i) simulate at $m_q > m_q^*$ where m_q^* is sufficiently large such that the spread in the position of the zero modes does not give rise to large fluctuations in the correlation function, and (ii) for fixed m_q increase the lattice volume as it decreases the probability of hitting exceptional configurations. These fixes do not help if one is interested in simulating QCD near the chiral limit, *i.e.*, for physical light quarks. For this new ideas are needed. A more detailed analysis of this problem and a possible solution is given in [59,60].

- The spin and flavor degrees of freedom are in one-to-one correspondence with continuum Dirac fermions. Thus, the construction of interpolating field operators is straightforward, for example, $\bar{\psi}\gamma_5\psi$ and $\bar{\psi}\gamma_i\psi$ are interpolating operators for pseudoscalar and vector mesons just as in the continuum (see Section 17).
- The Wilson term changes the discretization errors to $O(a)$.

- **Feynman Rules** and details of gauge fixing for the plaquette gauge action and Wilson fermions are given in [36].
- **Symmetries of S_F** : The transformation of the propagator under the discrete symmetries is the same as for naive fermions. For a given background configuration U these are

$$\begin{aligned}
\mathcal{P}: \quad S_F(x, y, [U]) &\rightarrow \gamma_4 S_F(x^P, y^P, [U^P]) \gamma_4 \\
\mathcal{T}: \quad S_F(x, y, [U]) &\rightarrow \gamma_4 \gamma_5 S_F(x^T, y^T, [U^T]) \gamma_5 \gamma_4 \\
\mathcal{C}: \quad S_F(x, y, [U]) &\rightarrow \gamma_4 \gamma_2 S_F^T(x, y, [U^C]) \gamma_2 \gamma_4 \\
\mathcal{H}: \quad S_F(x, y, [U]) &\rightarrow \gamma_5 S_F^\dagger(y, x, [U]) \gamma_5 \\
\mathcal{CPH}: \quad S_F(x, y, [U]) &\rightarrow \mathcal{C} \gamma_4 S_F^\dagger(y^P, x^P, [U^C]) \gamma_4 \mathcal{C}^{-1}
\end{aligned} \tag{10.11}$$

These relations are very useful in determining the properties of correlation functions (real versus imaginary, even or odd in \mathbf{p} , *etc.*) as discussed in Section 17.

- The dispersion relation, Eq. 10.8, has the following two solutions

$$e^{Ea} = \frac{(ma + r) \pm \sqrt{(1 + 2mar + m^2a^2)}}{1 + r} \tag{10.12}$$

For $r = 0$ and m small (naive fermion action), the two solutions are

$$\begin{aligned}
E &= m + O(m^2a^2) \\
E &= -m + \frac{i\pi}{a} + O(m^2a^2),
\end{aligned} \tag{10.13}$$

where the second solution corresponds to the pole in the Euclidean propagator at $p_4 = \pi$ and is the temporal doubler. For $r = 1$ this doubler is removed, as expected, and the single root $Ea = \log(1 + ma)$ is associated with the physical particle. For a general $r \neq 1$ the second root is given by

$$E = -m + \frac{1}{a} \left(i\pi + \log \frac{(1-r)}{1+r} \right) + \dots \tag{10.14}$$

Such branches which are not associated with the physical state or with doublers are called ghost branch. I use a very restrictive definition of a doubler -- a solution which has $|E| = 0$ for $m = \mathbf{p} = 0$. Again

from Eq. 10.14 it is clear that this ghost branch is pushed to infinity for $r = 1$. In general, if the action couples $(n + 1)$ time slices, then the dispersion relation has n branches. The roots can be real, complex-conjugate pairs, and/or have an imaginary part $i\pi/a$ as illustrated above. Such additional branches can be a nuisance for improved actions, for example the $D234$ action discussed in Section 13.2. The goal is to obtain a doubler and ghost free theory. This can be done by a combination of tuning the parameters (for example choosing $r = 1$ in the Wilson case) and/or going to anisotropic lattices $a_t \ll a_s$ as advocated by Lepage and collaborators [107].

- **Conserved Vector Current:** The Wilson action is invariant under the global transformation defined in Eq. 7.10. To derive the associated conserved current (for degenerate masses) we use the standard trick of calculating the variation of the action under a space-time dependent $\theta(x)$ and associating the coefficient of $\theta(x)$ in δS with $\partial_\mu J_\mu$. The change in the action under such an infinitesimal transformation is

$$\begin{aligned} \delta S = & \kappa \sum_{x,\mu} \bar{\psi}(x)(\gamma_\mu - r)U_\mu(x)\psi(x + \hat{\mu}) \exp(i\theta(x) - i\theta(x + \mu)) \\ & - \kappa \sum_{x,\mu} \bar{\psi}(x + \mu)(\gamma_\mu + r)U_\mu^\dagger(x)\psi(x) \exp(i\theta(x + \mu) - i\theta(x)) \end{aligned}$$

which, to first order in θ , is

$$\begin{aligned} - \sum_{x,\mu} & [\bar{\psi}(x)(\gamma_\mu - r)U_\mu(x)\psi(x + \hat{\mu}) + \\ & \bar{\psi}(x + \mu)(\gamma_\mu + r)U_\mu^\dagger(x)\psi(x)] \left[i \frac{\partial \theta}{\partial x_\mu} \right]. \end{aligned}$$

The conserved current, obtained after integration by parts, is

$$V_\mu^c = \bar{\psi}(x)(\gamma_\mu - r)U_\mu(x)\psi(x + \hat{\mu}) + \bar{\psi}(x + \mu)(\gamma_\mu + r)U_\mu^\dagger(x)\psi(x). \quad (10.15)$$

V_μ^c is hermitian and reduces to the symmetrized version of the 1-link vector current for $r = 0$. In many applications like decay constants one uses the local (flavor) currents defined by

$$\begin{aligned} 2V_\mu(x) &= \bar{\psi}(x)\gamma_\mu \frac{\lambda^a}{2} U_\mu(x)\psi(x + \hat{\mu}) + h.c. \\ 2A_\mu(x) &= \bar{\psi}(x)\gamma_5 \gamma_\mu \frac{\lambda^a}{2} U_\mu(x)\psi(x + \hat{\mu}) + h.c., \end{aligned} \quad (10.16)$$

which are not conserved, and consequently have associated non-trivial renormalization factors Z_V and Z_A which have to be included when calculating matrix elements.

- **The chiral condensate:** As explained in [57], the chiral condensate $\langle \bar{\psi}\psi \rangle$ is not simply related to the trace of the Wilson quark propagator $\langle S_F(0,0) \rangle$. The breaking of chiral symmetry by the r term introduces contact terms that need to be subtracted non-perturbatively from $S_F(0,0)$. This has not been proven practical. The methods of choice are to use either the continuum Ward Identity

$$\langle \bar{\psi}\psi \rangle^{\text{WI}} \equiv \langle 0 | S_F(0,0) | 0 \rangle = \lim_{m_q \rightarrow 0} m_q \int d^4x \langle 0 | P(x) P(0) | 0 \rangle, \quad (10.17)$$

or the lattice version of the Gell-Mann, Oakes, Renner relation [61]

$$\langle \bar{\psi}\psi \rangle^{\text{GMOR}} = \lim_{m_q \rightarrow 0} -\frac{f_\pi^2 M_\pi^2}{4m_q}. \quad (10.18)$$

A comparison of the efficacy of the two methods is discussed in [62].

- **Operator Mixing:** There is, in general, a mixing of operators with different chirality in the Callan Symanzik equations. This is true even at κ_c and is a consequence of the explicit breaking of chiral symmetry. This mixing of operators poses serious problems in the calculation of matrix elements of the weak interactions Hamiltonian, *i.e.*, the lattice analyses of Ward identities and obtaining amplitudes with the correct chiral behavior is more involved. For example the chiral behavior of matrix elements of the tree-level 4-fermion operators (relevant to the extraction of B_K , B_7 , and B_8) is not the same as in the continuum theory. The reason being that the Ward Identities associated with the spontaneously broken $SU_A(n_f)$ receive corrections at $O(a)$ in the lattice theory. This issue will be discussed in more detail by Prof. Martinelli.

In Section 13 I will discuss how the $O(a)$ errors in WF can, with very little extra effort, be reduced to $O(a^2)$ or higher. For the time being let me use this drawback of Wilson fermions to motivate the formulation of staggered fermions. This formulation has $O(a^2)$ errors, retains enough chiral symmetry to give the correct chiral behavior for amplitudes involving Goldstone pions, but at the expense of a 4-fold increase in the number of flavors.

11. Staggered fermions (*SF*)

The 16-fold doubling problem of the naive fermion action given in Eq. 7.6 can be reduced to 4 by the trick of spin-diagonalization [21]. Staggered fermions χ are defined by the transformation

$$\psi(x) = \Gamma_x \chi(x) \quad \bar{\psi}(x) = \bar{\chi}(x) \Gamma_x^\dagger \quad \Gamma_x = \gamma_1^{x_1} \gamma_2^{x_2} \gamma_3^{x_3} \gamma_4^{x_4}. \quad (11.1)$$

In terms of χ the action can be written as [21]

$$\begin{aligned} \mathcal{S}_S &= m_q \sum_x \bar{\chi}(x) \chi(x) + \frac{1}{2} \sum_{x,\mu} \bar{\chi}_x \eta_{x,\mu} (U_{\mu,x} \chi_{x+\hat{\mu}} - U_{\mu,x-\hat{\mu}}^\dagger \chi_{x-\hat{\mu}}) \\ &\equiv \sum_{x,y} \bar{\chi}(x) M_{xy}^S \chi(y) \end{aligned} \quad (11.2)$$

with the matrix M^S given by

$$M^S[U]_{x,y} = m_q \delta_{xy} + \frac{1}{2} \sum_{\mu} \eta_{x,\mu} [U_{\mu,x} \delta_{x,y-\hat{\mu}} - U_{x-\hat{\mu},\mu}^\dagger \delta_{x,y+\hat{\mu}}]. \quad (11.3)$$

The γ matrices are replaced by the phases $\eta_{x,\mu}$. It is convenient to define the following phase factors

$$\begin{aligned} \eta_{x,\mu} &= (-)^{\sum_{\nu < \mu} x_\nu} \\ \zeta_{x,\mu} &= (-)^{\sum_{\nu > \mu} x_\nu} \\ \epsilon_x &= (-)^{x_1+x_2+x_3+x_4} \\ \mathcal{S}_R(x) &= \frac{1}{2} (1 + \zeta_\mu \zeta_\nu - \eta_\mu \eta_\nu + \zeta_\mu \zeta_\nu \eta_\mu \eta_\nu). \end{aligned} \quad (11.4)$$

From Eqs. 11.2,11.3 it is easy to see that the different spin components of χ are decoupled as the phase factor $\eta_{x,\mu}$ depends only on the site index and direction and do not have a spinor index. One can therefore drop the spin index on χ leaving only color degrees of freedom at each site. This reduces the original 2^d -fold degeneracy of naive fermions by a factor of four. The mass term in Eq. 11.3 is hermitian, while the \not{D} term is anti-hermitian, which as I discussed before is required to realize chiral symmetry.

The *SF* action has translation invariance under shifts by $2a$ due to the phase factors $\eta_{x,\mu}$. Thus, in the continuum limit, a 2^4 hypercube is mapped to a single point, and the 16 degrees of freedom reduce to 4 copies of Dirac fermions. That is, for every physical flavor, the staggered discretization produces a 4-fold degeneracy denoted as the staggered flavor. At finite a the gauge interactions break this flavor symmetry and the 16 degrees of freedom in the hypercube are a mixture of spin and flavor. This is one

of the major drawbacks of staggered fermions. Construction of operators and their interpretation in terms of spin and flavor is non-trivial as I show below for the case of bilinears. I begin the discussion of staggered fermions by listing its symmetries. I will then outline the construction of bilinear operators and end with a discussion of Ward identities.

Symmetries of SF: The analogue of γ_5 invariance for SF is

$$\begin{aligned}\epsilon_x M_S^\dagger(x, y) \epsilon_y &= M_S(x, y), \\ \epsilon_x S_F^\dagger(x, y) \epsilon_y &= S_F(y, x).\end{aligned}\tag{11.5}$$

and the symmetries of this action are [63]

Translations	S_μ	:	$\chi(x) \rightarrow \zeta_\mu(x) \chi(x + \mu)$
Rotations	$R_{\mu\nu}$:	$\chi(x) \rightarrow S_R(R^{-1}x) \chi(R^{-1}x)$
Inversion	I	:	$\chi(x) \rightarrow \eta_4(x) \chi(I^{-1}x)$
Charge Conjugation	C	:	$\begin{cases} \chi(x) \rightarrow \epsilon(x) \bar{\chi}(x), \\ \bar{\chi}(x) \rightarrow -\epsilon(x) \chi(x) \end{cases}$
Vector(Baryon#)	$U(1)_V$:	$\begin{cases} \chi(x) \rightarrow e^{i\theta} \chi(x), \\ \bar{\chi}(x) \rightarrow e^{-i\theta} \bar{\chi}(x) \end{cases}$
Axial	$U(1)_A$:	$\begin{cases} \chi(x) \rightarrow e^{i\theta\epsilon(x)} \chi(x), \\ \bar{\chi}(x) \rightarrow e^{i\theta\epsilon(x)} \bar{\chi}(x) \end{cases}$

where the phases η, ζ, ϵ and S_R are defined in Eq. 11.4. It is the last invariance, the $U(1)_A$, or also called the $U(1)_\epsilon$ invariance, that becomes a flavor non-singlet axial symmetry in the continuum limit, and is the main advantage of staggered fermions.

Operators with Staggered fermions: There are two equivalent ways of understanding flavor identification of staggered fermions: (i) momentum space [64], and (ii) position space [65,66]. Here I review the position space approach.

For the purposes of constructing hadronic operators it is convenient to represent the 16 degrees of freedom in a hypercube by a hypercube field Q , which in the continuum limit represents the four spin components of four degenerate flavors [67]. This is done by dividing the lattice into hypercubes identified by a 4-vector y , *i.e.* y is defined on a lattice of spacing $2a$, and the set of 16 vectors $\{A\}$ to points within the hypercube. These are called hypercube vectors, and addition of two such vectors is defined modulo 2, $C_\mu = (A_\mu + B_\mu)_{mod 2}$. Then, in the free field limit (setting gauge links $U_{x,\mu} = 1$ with the understanding that for the interacting theory non-local operators are to be connected by a path ordered product of links that

makes them gauge invariant, or these operators are to be evaluated in a fixed gauge)

$$Q(y) = \frac{1}{N_f} \sum_{\{A\}} \Gamma_A \chi(y+A) \quad (11.6)$$

where $N_f = 4$ and

$$\Gamma_A = \gamma_1^{A_1} \gamma_2^{A_2} \gamma_3^{A_3} \gamma_4^{A_4}. \quad (11.7)$$

A general bilinear operators can be written as

$$\begin{aligned} \mathcal{O}_{SF} &= \text{Tr}(\bar{Q} \gamma_S Q \gamma_F^\dagger) \\ &= \sum_{A,B} \bar{\chi}(y+A) \text{Tr}(\Gamma_A^\dagger \gamma_S \Gamma_B \xi_F) \chi(y+B) \\ &= \sum_{A,B} \bar{\chi}(y+A) (\gamma_S \otimes \xi_F)_{AB} \chi(y+B) \\ &= \bar{Q} (\gamma_S \otimes \xi_F) Q. \end{aligned} \quad (11.8)$$

where the matrices γ_S determines the spin and ξ_F the flavor of the bilinear. Each of these can be one of the standard sixteen elements of the Clifford algebra. In the last expression, the field Q is expressed as a 16 component vector. It is easy to see from Eq. 11.8 that local operators are given by $\gamma_S = \xi_F$. The identification of meson operators is as follows [65,68,69].

$$\begin{aligned} V_\mu &= \bar{Q} (\gamma_\mu \otimes \xi_F) Q \\ A_\mu &= \bar{Q} (\gamma_\mu \gamma_5 \otimes \xi_F) Q \\ P &= \bar{Q} (\gamma_5 \otimes \xi_F) Q \\ S &= \bar{Q} (1 \otimes \xi_F) Q. \end{aligned} \quad (11.9)$$

In the continuum the mesons lie in the flavor SU(4) representations **1** and **15**, while on the lattice these representations break into many smaller ones. For example, the continuum **15**-plet of pions breaks into seven lattice representations (four 3-d and three 1-d).

$$\begin{aligned} \mathbf{15} &\rightarrow (\gamma_5 \otimes \gamma_i) \oplus (\gamma_5 \otimes \gamma_i \gamma_4) \oplus (\gamma_5 \otimes \gamma_i \gamma_5) \oplus (\gamma_5 \otimes \gamma_i \gamma_4 \gamma_5) \oplus \\ &\quad (\gamma_5 \otimes \gamma_4) \oplus (\gamma_5 \otimes \gamma_5) \oplus (\gamma_5 \otimes \gamma_4 \gamma_5), \\ \mathbf{1} &\rightarrow (\gamma_5 \otimes 1). \end{aligned} \quad (11.10)$$

In the first line, $i = 1, 2, 3$, so the representations are 3-dimensional. The other representations are 1-dimensional. Similarly, the continuum rho **15**-plet (times three spin components) breaks into eleven representations (four

6-d and seven 3-d).

$$\begin{aligned}
\mathbf{15} &\rightarrow (\gamma_i \otimes \gamma_j) \oplus (\gamma_i \otimes \gamma_j \gamma_4) \oplus (\gamma_i \otimes \gamma_j \gamma_5) \oplus (\gamma_i \otimes \gamma_j \gamma_4 \gamma_5) \oplus \\
&\quad (\gamma_i \otimes \gamma_i) \oplus (\gamma_i \otimes \gamma_i \gamma_4) \oplus (\gamma_i \otimes \gamma_i \gamma_5) \oplus (\gamma_i \otimes \gamma_i \gamma_4 \gamma_5) \oplus \\
&\quad (\gamma_i \otimes \gamma_4) \oplus (\gamma_i \otimes \gamma_5) \oplus (\gamma_i \otimes \gamma_4 \gamma_5), \\
\mathbf{1} &\rightarrow (\gamma_i \otimes \mathbf{1}).
\end{aligned} \tag{11.11}$$

Of the sixteen pions, the Goldstone pion has flavor ξ_5 . The remaining 15 pions are given by the other 15 members of the Clifford algebra and are heavier due to the staggered flavor symmetry breaking, and their masses do not vanish at $m_q = 0$. In fact the mass for each representation can be different. The degree to which the staggered flavor symmetry is broken can be quantified, for example, by comparing the mass of the Goldstone pion versus those in the other representations.

The above examples make it clear that the mixing of spin and flavor complicates the analysis and is a major disadvantage of staggered fermions. I will not go into details of calculations using staggered fermions. A useful list of references is as follows. The construction of baryon operators is given in [70]. For an illustrative calculations of the spectrum see [69]. The transcription of 4-fermion operators arising in effective weak Hamiltonian and their matrix elements within hadronic states are discussed in [63,71].

Chiral Symmetry and Ward Identities: The staggered fermion action can be written in terms of the 16 component fields Q , defined in Eq. 11.6, as [4]

$$\begin{aligned}
\mathcal{S}_S &= m_q \sum_y \bar{Q}(y) (\mathbf{1} \otimes \mathbf{1}) Q(y) \\
&+ \sum_{y,\mu} \bar{Q}(y) [(\gamma_\mu \otimes \mathbf{1}) \partial_x + (\gamma_5 \otimes \xi_\mu \xi_5) \square_\mu] Q(y).
\end{aligned} \tag{11.12}$$

This action, for $m_q = 0$, is invariant under the following abelian transformation

$$\begin{aligned}
Q(y) &\rightarrow e^{i\theta(\gamma_5 \otimes \xi_5)} Q(y) \\
\bar{Q}(y) &\rightarrow \bar{Q}(y) e^{i\theta(\gamma_5 \otimes \xi_5)}.
\end{aligned} \tag{11.13}$$

Thus one preserves a non-trivial piece of the original $U(4) \times U(4)$ symmetry (called $U(1)_\epsilon$ or $U(1)_A$) whose generator is $\gamma_5 \otimes \xi_5$ and the associated Goldstone boson operator is $\bar{Q}(\gamma_5 \otimes \xi_5)Q$. This axial $U(1)_A$ symmetry, along with the $U(1)_V$, allows, for appropriate transcription of operators, Ward Identities for lattice amplitudes similar to those in the continuum.

The best studied example of this is the lattice calculation of the $K^0\bar{K}^0$ mixing parameter B_K [72].

The basic tools for constructing the Ward Identities are exemplified by the following two relations involving quark propagators and the zero 4-momentum insertion of the pseudo-Goldstone operator $\bar{\chi}(x)\chi(x)(-1)^x$ and the scalar density $\bar{\chi}(x)\chi(x)$ [63]

$$(-1)^{x_f} G_2(x_f, x_i) + G_1(x_f, x_i)(-1)^{x_i} = (m_1 + m_2) \sum_x G_1(x_f, x)(-1)^x G_2(x, x_i) \quad (11.14)$$

$$G_2(x_f, x_i) - G_1(x_f, x_i) = (m_1 - m_2) \sum_x G_1(x_f, x) G_2(x, x_i) \quad (11.15)$$

where $G_i(x_f, x_i)$ is the propagator from x_i to x_f for a quark of mass of mass m_i . The derivation of these identities, using the hopping parameter expansion, is given in Ref. [63]. Note that these relations are valid for all a and at finite m . An example of the usefulness of these relations is in extracting the chiral condensate as shown below. Eq. 11.14 gives

$$\begin{aligned} \langle \bar{\chi}\chi \rangle &= G(x_i, x_i) = m \sum_x (-1)^{-x_i} G(x_i, x)(-1)^x G(x, x_i) \\ &= m \sum_x |G(x_i, x)|^2, \end{aligned} \quad (11.16)$$

where the last term on the right is the zero 4-momentum sum of the pion correlator. Similarly Eq. 11.15 gives

$$\frac{\partial \langle \bar{\chi}\chi \rangle}{\partial m} = \sum_x G(x_i, x) G(x, x_i) = \sum_x |G(x_i, x)|^2 (-1)^{x-x_i}. \quad (11.17)$$

Combining these two equations one can extract the value of the condensate in the chiral limit

$$\langle \bar{\chi}\chi \rangle(m=0) = \left(1 - m \frac{\partial}{\partial m}\right) \langle \bar{\chi}\chi \rangle = \sum_{\text{even}} |G_1(x_i, x)|^2 \quad (11.18)$$

by summing the Goldstone pion's 2-point function over sites with even $(x - x_i)$.

11.1. Summary

Even though most of the simulations done until recently have used either the staggered or Wilson formulation of fermions, neither is satisfactory.

They have their relative advantages and disadvantages which have been evaluated on a case by case basis. The rule of thumb is that in the extraction of observables where chiral symmetry plays an essential role and the external states are Goldstone bosons, staggered fermions do better. Otherwise Wilson fermions are preferred due to their correspondence with Dirac fermions in terms of spin and flavor. The focus of current research is to improve the Wilson formulation and thus have the best of both worlds. This is being done in a variety of ways as discussed in Section 13.

12. Improved gauge actions

There have been a number of methods proposed to improve the gauge action beyond $O(a^2)$. These include the Symanzik approach, mean-field improvement, and renormalization group inspired perfect actions. Let me begin this discussion by reviewing the Symanzik program.

12.1. Lüscher-Weisz Action

The leading order term in the expansion of all Wilson loops is $\mathcal{O}^{(4)} = \sum_{\mu\nu} F_{\mu\nu} F_{\mu\nu}$ and corrections begin at $O(a^2)$ as there are no dimension 5 operators. Thus, any lattice action written as a linear combination of Wilson loops will have the correct continuum limit with corrections at $O(a^2)$. There are three dimension 6 operators, which in continuum notation are:

$$\begin{aligned}\mathcal{O}_1^{(6)} &= \sum_{\mu,\nu} \text{Tr} \left(D_\mu F_{\mu\nu} D_\mu F_{\mu\nu} \right), \\ \mathcal{O}_2^{(6)} &= \sum_{\mu,\nu,\rho} \text{Tr} \left(D_\mu F_{\nu\rho} D_\mu F_{\nu\rho} \right), \\ \mathcal{O}_3^{(6)} &= \sum_{\mu,\nu,\rho} \text{Tr} \left(D_\mu F_{\mu\rho} D_\nu F_{\nu\rho} \right).\end{aligned}\tag{12.1}$$

Also there are only 3 six-link loops that one can draw on the lattice. These, as shown in Fig. 7, are the planar $\mathcal{L}_1^{(6)}$, twisted $\mathcal{L}_2^{(6)}$ and the L shaped $\mathcal{L}_3^{(6)}$ respectively. Thus, classical improvement of the lattice action, *i.e.* removing the $O(a^2)$ term, can be achieved by taking a linear combination of the plaquette and these three six-link loops. Each of these loops has the expansion

$$\mathcal{L} = r^{(4)} \mathcal{O}^{(4)} + r_1^{(6)} \mathcal{O}_1^{(6)} + r_2^{(6)} \mathcal{O}_2^{(6)} + r_3^{(6)} \mathcal{O}_3^{(6)} + \dots,\tag{12.2}$$

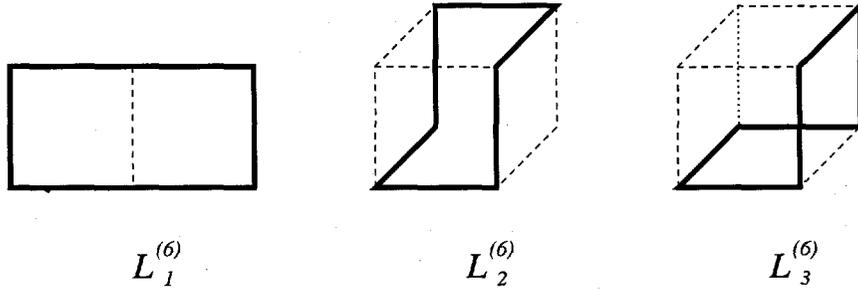


Fig. 7. The three types of six link loops.

Table 4

The coefficients of the continuum operators of dimension 4 and 6 in the classical expansion of Wilson loops with 4 and 6 links.

Loop	$r^{(4)}$	$r_1^{(6)}$	$r_2^{(6)}$	$r_3^{(6)}$
$\mathcal{L}^{(4)}$	$-\frac{1}{4}$	$\frac{1}{24}$	0	0
$\mathcal{L}_1^{(6)}$	-2	$\frac{5}{6}$	0	0
$\mathcal{L}_2^{(6)}$	-2	$-\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$\mathcal{L}_3^{(6)}$	-4	$\frac{1}{6}$	0	$\frac{1}{2}$

and Lüscher and Weisz have shown an elegant way of calculating the expansion coefficients $r_\alpha^{(d)}$ [73]. Their results are summarized in Table 4.

The lattice gauge action can be written as

$$\mathcal{S}_g = \frac{6}{g^2} \left\{ c^{(4)}(g^2) \mathcal{L}^{(4)} + \sum_{i=1,3} c_i^{(6)}(g^2) \mathcal{L}_i^{(6)} \right\} \quad (12.3)$$

in terms of the plaquette and these three 6-link loops. Using the results in Table 4 one gets the normalization condition

$$c^{(4)}(g^2) + 8c_1^{(6)}(g^2) + 8c_2^{(6)}(g^2) + 16c_3^{(6)}(g^2) = 1 \quad (12.4)$$

such that in the continuum limit the action reduces to $F_{\mu\nu}F_{\mu\nu}/4$. Also, from Table 4 one can see that tree level improvement can be obtained by the choice

$$c^{(4)} + 20c_1^{(6)} = 0; \quad c_2^{(6)} = 0; \quad c_3^{(6)} = 0. \quad (12.5)$$

This removes $\mathcal{O}_1^{(6)}$ by cancelling the contribution from the two planar loops. Of the three dimension six operators, $\mathcal{O}_1^{(6)}$ alone breaks Lorentz (rotational) invariance. Thus, rotational invariance of the heavy $q\bar{q}$ potential can be used to non-perturbatively tune $c_1^{(6)}$ and to compare the various improved actions as discussed later.

To include quantum effects, Curci, Menotti, and Paffuti did a 1-loop calculation. They imposed the criteria that all Wilson loops be improved at the leading logarithm level [74,75]. This corresponds to improving the string tension. (To kill the $O(a^2)$ terms at all length scales instead of only in the long distance heavy quark potential requires including the finite pieces in the 1-loop calculations. The fully improved action, as we shall see in the final Lüscher-Weisz form, will lead to $O(g^2)$ corrections to $c^{(4)}$ and $c_1^{(6)}$ and the need for $\mathcal{O}_2^{(6)}$.) The result of their calculation required the improved action to have the same form as given in Eq. 12.5. This agreement with the classical result is fortuitous. Symanzik's on-shell improvement program, in general, prescribes that one has to derive three relations to fix the $c_i^{(6)}$ such that on-shell quantities do not have $O(a^2)$ corrections. Lüscher and Weisz, in fact, found that improving all large loops produced only two constraints at tree-level [76]

$$\begin{aligned} c^{(4)} + 20c_1^{(6)} - 4c_2^{(6)} + 4c_3^{(6)} &= 0 \\ c_2^{(6)} &= 0. \end{aligned} \tag{12.6}$$

Thus, they were left with a 1-parameter family of improved actions. The reason for this freedom is the presence of a redundant operator $\mathcal{O}_3^{(6)}$ generated by the field redefinition $A_\mu \rightarrow A_\mu + a^2 g^2 d(g) D_\nu F_{\nu\mu}$ [73,76]. Since one has the liberty to choose the coefficient of an interaction that is part of a redundant operator without affecting on-shell improvement, the convenient choice $c_3^{(6)} = 0$ simplifies the action and again leads us back to Eq. 12.5. To summarize, Eqs. 12.6 are the most general set of constraints that can be derived based on improving large Wilson loops. The criteria employed by Curci *et.al.*, improving the large distance part of $q\bar{q}$ potential, is a subset. Since this involves a spectral quantity one can use the redundant operator to get the simpler form Eq. 12.5.

The relation $c_3^{(6)} = 0$ holds both at tree and 1-loop level as it comes from exploiting the redundant operator. Combining this with the first constraint in Eq. 12.6 and the normalization condition Eq. 12.4 leads to a single condition, which at tree-level is

$$c_1^{(6)} - c_2^{(6)} = -\frac{1}{12}, \tag{12.7}$$

and is all that can be gotten from Wilson loops. To get a second condition Lüscher and Weisz had to set up other spectral quantities to be improved. Failing to find a suitable condition in the real theory, they looked at a situation in which the gluons are massive and appear as asymptotic states. Using the mass and scattering amplitude of these states they derived the 1-loop improved action [76], which goes under the name Lüscher-Weisz action and has leading correction at $O(g^4 a^4)$,

$$\begin{aligned} c_0^{(4)}(g^2) &= \frac{5}{3} + 0.2370g^2 \\ c_1^{(6)}(g^2) &= -\frac{1}{12} - 0.02521g^2 \\ c_2^{(6)}(g^2) &= -0.00441g^2 \\ c_3^{(6)}(g^2) &= 0. \end{aligned} \tag{12.8}$$

In this form for the c_i , the normalization condition, Eq. 12.4, is already incorporated.

The bottom line on the Lüscher-Weisz action is that since the coefficient $c_2^{(6)}$ is small, the action, to a very good approximation, can be improved by keeping just $\mathcal{L}^{(4)}$ and $\mathcal{L}_1^{(6)}$ as in the classical case. This 1-loop result can be mean-field improved without much effort as discussed next.

12.2. Tadpole/Mean-field Improvement

The representation of gauge fields by the exponentiated form

$$U_\mu(x) = e^{iagA_\mu(x)} = 1 + iagA_\mu(x) - \frac{a^2g^2}{2}A_\mu^2(x) + \dots \tag{12.9}$$

gives rise to local quark-gluon vertices with one, two, ... gluons. Of these vertices, all but the lowest are lattice artifacts. Contracting the two gluons in $a^2g^2A_\mu^2(x)/2$ gives rise to tadpole diagrams. Parisi, and Lepage and Mackenzie [77,78] recognized that these artifacts, though naïvely suppressed by powers of a , are in fact only suppressed by powers of g^2 due to the ultraviolet divergences generated by the tadpole loops. They reasoned that these tadpole contributions were responsible for (i) the poor match between short distance quantities (like the link in Landau gauge, small Wilson loops, κ_c , ...) and their perturbative estimates, and (ii) large coefficients in the perturbative lattice expansions. They, therefore, proposed a mean-field improvement (also called tadpole improvement(TI)) to get rid of these artifacts.

The mean-field improvement starts by assuming that the lattice fields can be split into a UV and IR parts, and that the UV part should be integrated out [77,78], *i.e.*

$$e^{iagA_\mu(x)} = e^{iag(A_\mu^{IR}(x)+A_\mu^{UV}(x))} = u_0 e^{iagA_\mu^{IR}(x)} \equiv u_0 \tilde{U}_\mu(x). \quad (12.10)$$

The rescaling of links by an overall constant factor $u_0 < 1$ keeps the theory gauge invariant. Under this rescaling (each $U \rightarrow u_0 \tilde{U}$) the Wilson gauge and fermion action becomes

$$S^W \rightarrow \sum_x \left[\beta u_0^4 \mathcal{W}_{11} + \bar{\psi} \tilde{\psi} - \kappa u_0 \bar{\psi} \tilde{D} \tilde{\psi} \right]. \quad (12.11)$$

Thus, a simplified recipe for mean-field improvement is to use the better expansion parameters $\tilde{g}^2 = g^2/u_0^4$, $\tilde{\kappa} = \kappa u_0$, $\tilde{U} = U/u_0$, and $\tilde{\psi} = \psi^{cont}/\sqrt{2\kappa u_0}$. Note that in terms of the tilde variables the theory looks the same as the original theory except that all links are scaled by u_0 . Lepage and Mackenzie [78], and subsequent work, have shown that mean-field improvement of 1-loop estimates gives, in many cases, results that are very close to non-perturbative estimates, and the expansion coefficients are much smaller. Tadpole improvement is, therefore, a useful thing to do, especially since it requires negligible extra effort once the 1-loop calculation has been done. The only limitation is that it does not work equally well in all cases and there is no real understanding of why it does not work well in some cases. To illustration tadpole improvement I will apply it to the Lüscher-Weisz action.

There are two common choices of the tadpole factor u_0 : (i) the fourth root of the plaquette, and (ii) the expectation value of the link in Landau gauge. So far most calculation have used the first choice, whereas Lepage and collaborators now advocate the link in Landau gauge as a better choice [79]. To implement tadpole improvement requires both the 1-loop expressions for these variables and their non-perturbative values. For the Wilson action

$$\begin{aligned} u_0 &= \langle \mathcal{W}_{11} \rangle^{1/4} = 1 - \frac{1}{12} g^2 \\ u_0 &= \langle U \rangle_{LG} = 1 - 0.0772 g^2 \end{aligned} \quad (12.12)$$

and for the Lüscher-Weisz action, Eq. 12.5, it is [80]

$$u_0 = \langle \mathcal{W}_{11} \rangle^{1/4} = 1 - \left(\frac{1}{2}\right) \left(\frac{4}{3}\right) \left(\frac{0.366262}{4}\right) g^2 = 1 - 0.061044 g^2. \quad (12.13)$$

Denoting the 1-loop tadpole factor as $u_0 = 1 - Xg^2$, the tadpole-improved Lüscher-Weisz (TILW) action is [81]

$$\begin{aligned} S_{TILW} = \beta_{eff} & \left(\mathcal{L}^{(4)} - \frac{1}{20u_0^2} (1 + (0.1602 - 2X)g^2) \mathcal{L}_1^{(6)} \right. \\ & \left. - \frac{1}{u_0^2} 0.00264 g^2 \mathcal{L}_2^{(6)} \right) \end{aligned} \quad (12.14)$$

The non-perturbative value of u_0 has to be determined self-consistently as the relative co-efficients in the action themselves depend on u_0 . Note that $\mathcal{L}_2^{(6)}$, which comes in only as a quantum effect, has, after TI, a coefficient is reduced to a negligible value.

An important point to note about the TILW action is the strength of $\mathcal{L}_1^{(6)}$. Since $u_0 = 0.85 - 0.90$ for lattice parameters currently in use, the strength of $\mathcal{L}_1^{(6)}$ is larger than the classical value $1/20$, i.e tadpole improvement suggests an over-correction. Later on we shall see that this over-correction is also characteristic of improved actions based on renormalization group methods.

A simplified description of TI applied to the above example is that the coefficient of $\mathcal{L}_1^{(6)}$ was multiplied by $u_0^2(pert)/u_0^2(non - pert)$, the relative factor of two extra links between the plaquette and $\mathcal{L}_1^{(6)}$. The factor $u_0^2(pert)$, which differs from $u_0^2(non - pert)$ at $O(g^4)$, is absorbed in the 1-loop expansion. This, as is obvious, makes the coefficient of g^2 small. The correction is, therefore, dominated by $u_0(non - pert)$ which can be measured with very high precision. There, however, is the residual uncertainty due to the freedom in the choice of u_0 . To summarize, the underlying philosophy of TI is that the large tadpole contributions are present in the perturbative expansion of all quantities, and can be absorbed by factoring $u_0(pert)$ raised to the appropriate power depending on the number of links. This makes the expansion coefficients small. The modified perturbative expansion is then in terms of an improved $\tilde{g}^2 = g^2/u_0^4$, and the correction is dominated by overall factors of $u_0(non - pert)$, which should be chosen to be the expectation value of the link in Landau gauge.

One of the first tests to convincingly demonstrate improvement using the TILW action was restoration of rotational invariance in the $q\bar{q}$ potential. The results from [82] are reproduced in Fig. 8. Since then tests of improvement using the scaling of the finite temperature transition, T_c , and topological susceptibility have been carried out. These results, as will be discussed later, all make a strong case for improvement.

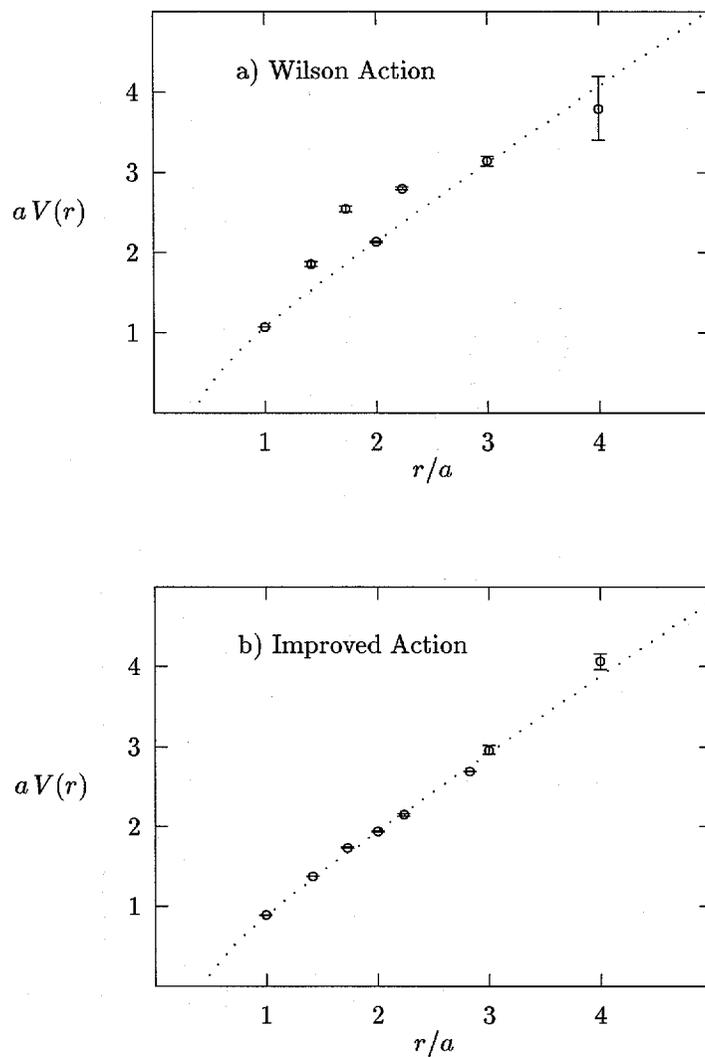


Fig. 8. Static-quark potential computed on 6^4 lattices with $a \approx 0.4$ fm using the Wilson action and the TILW action. The dotted line is the standard infrared parameterization for the continuum potential, $V(r) = Kr - \pi/12r + c$, adjusted to fit the on-axis values of the potential. Figure reproduced from [82].

12.3. The Renormalization Group and the Renormalized Trajectory

This is a brief summary of Wilson's formulation of the renormalization group and the concept of the renormalized trajectory. For details see [83–90]. I will start with a discussion in momentum space. Consider the effective theory formulated at some scale Q in terms of the generalized action

$$S(Q) = \sum_{\alpha} K_{\alpha}^{(0)}(Q) S_{\alpha} \quad (12.15)$$

where S_{α} are all possible operators (interactions) and $K_{\alpha}^{(0)}$ are the corresponding coupling constants. Now consider an exact renormalization group transformation (RGT) \mathcal{T}_b that integrates all momenta between Q and $Q/b \equiv Q/2$. Here $b = 2$ is the scale factor of the transformation. The renormalized theory at scale $Q/2$ is

$$S(Q/2) = \sum_{\alpha} K_{\alpha}^{(1)}(Q/2) S_{\alpha} \quad (12.16)$$

where $K_{\alpha}^{(1)}$ are analytical functions of $K_{\alpha}^{(0)}$,

$$K_{\alpha}^{(1)} = \mathcal{T}_b[K_{\alpha}^{(0)}] . \quad (12.17)$$

The sequence of theories defined by repeated use of \mathcal{T}_b defines a flow trajectory in the space of couplings. The physical theory is preserved under each \mathcal{T}_b , and thus along the whole trajectory. Correlation lengths, measured in lattice units, change trivially $\xi/a \rightarrow \xi/ab$. Fixed points of this transformation are defined by

$$K_{\alpha}^{(*)} = \mathcal{T}_b[K_{\alpha}^{(*)}] , \quad (12.18)$$

i.e. points at which the theory reproduces itself under \mathcal{T}_b . Since all correlation lengths (momenta) scale by b under \mathcal{T}_b , therefore at the fixed points

$$\xi|_{\text{fixed point}} = 0 \quad \text{or} \quad \infty . \quad (12.19)$$

Fixed points with $\xi = 0$ are trivial fixed points. Examples of these are the $T = 0$ and $T = \infty$ limit of most statistical mechanics models. We are interested in non-trivial fixed points with $\xi = \infty$. At these points the correlation length $\xi/a = 1/Ma$, measured in units of lattice spacing, is infinite. This allows us to take the continuum limit of lattice field theories — set $a = 0$ holding M fixed.

Now consider a point $\{K_{\beta}\}$ close to the fixed point $\{K_{\beta}^{*}\}$ and a RGT with an infinitesimal b . Under the assumption that there are no singularities in

the space of coupling constants, the new couplings $K'_\alpha(K_\beta)$ generated by \mathcal{T}_b can be written as a Taylor expansion about the fixed point

$$K'_\alpha(K_\beta) = K_\alpha^* + \left. \frac{\partial K'_\alpha}{\partial K_\beta} \right|_{K^*} \Delta K_\beta + \dots \quad (12.20)$$

In that case one can define a "linear region" in the neighborhood of $\{K_\beta^*\}$ in which

$$\Delta K'_\alpha = \left. \frac{\partial K'_\alpha}{\partial K_\beta} \right|_{K^*} \Delta K_\beta \quad (12.21)$$

where $\Delta K_\beta = K_\beta - K_\beta^*$. The linearized transformation matrix (also called the stability matrix)

$$\mathcal{T}_{\alpha\beta} \equiv \left. \frac{\partial K'_\alpha}{\partial K_\beta} \right|_{K^*} \quad (12.22)$$

controls the RG flows locally. In field theory discussions, the stability matrix $\mathcal{T}_{\alpha\beta}$ is called the β -function. Note that, because of the truncation of Eq. 12.20 all statements based on the analysis of $\mathcal{T}_{\alpha\beta}$ have the implicit assumption that they are approximately true in the linear region close to the fixed point and for small b , and become exact only as $\{K_\beta\} \rightarrow \{K_\beta^*\}$.

Wilson formulated the real space renormalization group as an alternate to the momentum space renormalization group discussed above. This approach is characterized by a constrained integral over the position dependent original degrees of freedom s to produce an effective theory,

$$e^{K'_\alpha S_\alpha(s')} = \int \mathcal{D}s \mathcal{T}(s', s) e^{K_\alpha S_\alpha(s)}. \quad (12.23)$$

Here $\mathcal{T}(s', s)$ is the constraint that specifies how the s in some local neighborhood (called the block cell) are to be averaged and replaced by an effective s' . A necessary condition on $\mathcal{T}(s', s)$, that preserves the partition function, is

$$\int \mathcal{D}s' \mathcal{T}(s', s) = 1. \quad (12.24)$$

One can see that if the block cell in D -dimensions is b^D , and the constraint $\mathcal{T}(s', s)$ is equivalent to integrating momenta in the range $Q/2 - Q$, then the two approaches are equivalent. In general, an exact integration of the upper half momenta is not feasible and $\mathcal{T}(s', s)$ is constructed as some sensible "average" of the original spins in the block cell. Other than the condition Eq. 12.24, the definition of a sensible "average" is that it

preserves the physical properties of the theory. Because of this dual description, momentum versus real space formulations, the RGT is also called a blocking transformation — a usage more common in statistical mechanics and numerical implementations.

A brief summary of the properties of the matrix $\mathcal{T}_{\alpha\beta}$, evaluated using either the momentum or real space formulations, and its associated eigenvalues Λ_i and eigenvectors Φ_i obtained by solving $\mathcal{T}\Phi_i = \Lambda_i\Phi_i$ is as follows.

1. $\mathcal{T}_{\alpha\beta}$ is, in general, not a symmetric matrix. Its eigenvalues are, however, real.
2. The eigenvalues depend on the scale factor b . Since the theory obtained after two successive transformations, $b \otimes b$, should be the same as the one obtained after a single scale change by b^2 , the general form of the eigenvalues is $\Lambda_i = b^{\lambda_i}$ where the λ_i are independent of \mathcal{T}_b and are called the critical exponents.
3. The eigenvectors Φ_i are by definition the linear scaling fields. Consider a small deviation of size u from the fixed point along a given eigenvector Φ . Then under a RGT, the renormalized value of the coupling, u' , is given by

$$u'\Phi = \mathcal{T}_b u\Phi = b^{\lambda_i} u\Phi \quad (12.25)$$

Note the crucial use of the assumption of linearity in deducing this simple scaling of u . The eigenvectors depend on the precise form of \mathcal{T}_b .

4. Eigenvalues $\Lambda_i > 1$ lead to flows away from the fixed point as deviations along these Φ_i grow under \mathcal{T}_b . These eigenvalues and the corresponding eigenvectors are called relevant.
5. Operators corresponding to eigenvalues $\Lambda_i < 1$ die out geometrically with the number of blocking steps. These operators do not contribute to the long distance properties of the system and are consequently called irrelevant. The associated exponents λ_i control corrections to scaling.
6. Eigenvalues that are exactly unity are called marginal. Deviations along marginal directions do not change under RGT. To ascertain that an eigenvalue is truly marginal, one has to go beyond the linear approximation. Typically marginal operators develop logarithmic corrections at higher order. A quick examination of the β -function of asymptotically free field theories should convince you that g is a marginal coupling at tree-level and becomes relevant due to quantum corrections.

7. There is an additional class of eigenvectors, called redundant operators, that are not physical. Redundant operators are manifestation of the invariance of the action under a redefinition of the fields, *i.e.*, they define relations between the different operators such as the equations of motion. The corresponding eigenvalues can be < 1 , $= 1$, or > 1 . These eigenvalues depend on the choice of the *RGT*, *i.e.* different RGT with the same or different scale factor b will give different λ_i . Since different non-linear RGT correspond to different rescaling of the fields, redundant operators are specific to the transformation. In fact, one way to isolate them is to repeat the calculation of $\mathcal{T}_{\alpha\beta}$ with a different *RGT*; the exponents λ_i that change with the RGT correspond to redundant operators. Even though they carry no physics information it is important to isolate these operators, especially in numerical calculations. For example, when improving the action one wants to use all possible redundant operators to simplify the action and thus minimize the computational cost.
8. Under certain conditions $\mathcal{T}_{\alpha\beta}$ is block diagonal, *i.e.* the couplings break up into sub-sets that are closed under the *RGT*. For example, the spin flip symmetry $s \rightarrow -s$ of the Ising Hamiltonian at zero magnetic field causes the interactions to break up into odd and even sectors. One can, therefore, calculate $\mathcal{T}_{\alpha\beta}$ for each sector separately as the eigenvalue spectrum is independent.

In this generalized space of couplings $\{K_\alpha\}$, there exist points, called critical points $\{K_\alpha^c\}$, with $\xi = \infty$. Starting from a critical point, a RGT produces another since $\xi' = \xi/b = \infty$. Thus the set of critical points define a hypersurface in the infinite dimensional space $\{K_\alpha\}$. The RG flows on this surface can (a) meander randomly, (b) go to some limit cycle or a strange attractor, or (c) converge to a fixed point $\{K_\alpha^*\}$. We shall only be concerned with possibility (c). At the fixed points the renormalized couplings are exactly equal to the original couplings, $K'_\alpha = K_\alpha$, and the theory reproduces itself at all length scales. This is in distinction to critical points where only the long distance behavior is reproduced. The location of the fixed point depends on the RGT and in general a given RGT will have more than one fixed point with $\xi = \infty$. Each such fixed point has a basin of attraction *i.e.* the set of $\{K_\alpha^c\}$ that converge to it under the RGT. This hypersurface is thus orthogonal to all relevant eigenvectors of \mathcal{T}_b . The basin of attraction defines the universality class since the long distance behavior of all theories is the same and governed by the fixed point theory.

The flow along a relevant scaling field, whose end-point is the fixed point is called the renormalized trajectory (RT). Just as the fixed point

is an attractor of all critical points in its universality class, the RT is the attractor of all flows terminating in the critical points that lie in the basin of attraction of the fixed point. This is illustrated in Fig. 9. For simplicity I have assumed that there is only one relevant coupling, as is the case for QCD. Along the RT there are no scaling violations as, by construction, all the irrelevant scaling fields are zero. Thus, simulations done using an action along the exact RT, will reproduce the continuum physics without discretization errors. The RT is thus the holy grail of attempts to improve the lattice formulation, *i.e.* the perfect action.

Simulations on finite lattices are done in a region of coupling space where all $\xi \ll L$, and not at the fixed point or at the critical points. Current typical values lie in the range $\xi/a = 1/\sqrt{\sigma}a = 5 - 15$. Even for quenched simulations this range will not change in the near future. Points in this region lie on flows that terminate at critical points that may lie quite far from the fixed point. Corrections to scaling along these flows at $\xi \approx 10$ may, therefore, be large. One can reduce these corrections by adjusting the action to lie on another flow that starts closer to the fixed point and at $\xi \approx 10$ is closer to the RT. Thus, to improve the action a non-perturbative approach is to estimate the RT. Since flows are attracted by the RT, its location can be estimated by starting with even the simple plaquette action and following the flows generated under some blocking transformation. Evaluating the blocked action along the flow is therefore the basis of all renormalization group methods for improving the action. Methods used for determining the blocked action are reviewed in [86].

A standard assumption underlying the validity of the renormalization group is that the fixed point action is local (the strength of the couplings fall off exponentially with the size of the Wilson loops and the separation between points that are coupled by the Dirac operator), nevertheless, a local action still involves an "infinite" number of couplings. Keeping all these when evaluating the action generated under blocking, and in simulations, is impractical. So one is forced to truncate the action to a finite (small) number of interactions. Also, the RT at $\xi \approx 10$ may be sufficiently different in form from that near the fixed point. Consequently, the questions one has to face when designing improved actions are (i) what is the minimal set of terms to keep, (ii) are all the redundant interactions removed from this set, and (iii) how to compare such improved actions evaluated using different RGT. Also note that the projection of the RT on to a truncated subspace may not be the action with the smallest scaling violations that one can design in that space. I shall explore attempts to answer these questions next.

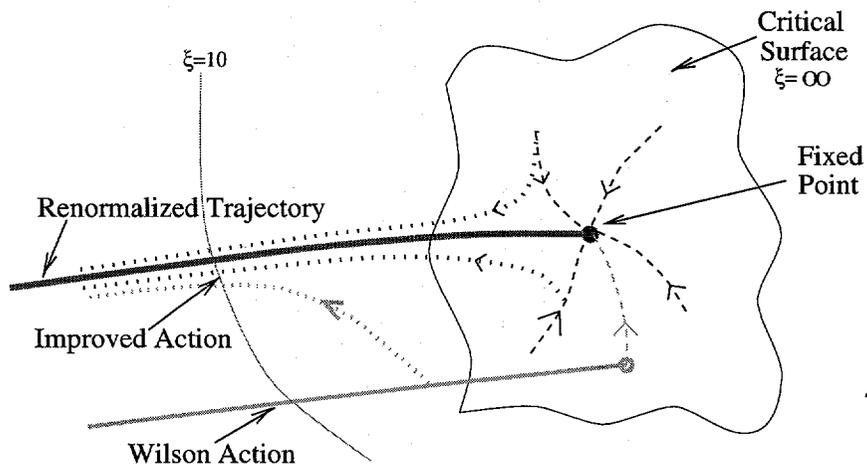


Fig. 9. A schematic of the critical hypersurface ($\xi = \infty$), the fixed point, the renormalized trajectory (RT) flowing towards the infrared under blocking transformations, and flows attracted by them. Also shown is the simple Wilson action and flows starting from it.

I will summarize current activity in designing improved gauge actions by reviewing those that are being used to calculate phenomenologically relevant quantities. Unfortunately, there is no standard notation used by different collaborations. To facilitate comparison I will therefore try to write them in a common form as well as maintain contact with the original notation. Wilson loops will be denoted by $W_{\mu\nu}^i$ where the index i will be used for both the type of loop and its representation. Unless otherwise stated, it will be implicit that the real part of the trace of Wilson loops is to be taken. The $W_{\mu\nu}^i$ are normalized such that the maximum value of the trace is unity. As before β is the overall coupling and the c_i give the relative strengths of different loops.

12.4. Iwasaki Action

Iwasaki studied block transformations in perturbation theory. His criterion was to obtain a gauge action that after a few blockings comes close to the RT. His *Ansatz* is [91]

$$\mathcal{S}_{Iwasaki} = \beta \sum_{\mu,\nu} \sum_x \left(W_{\mu\nu}^{1 \times 1} - 0.0907 W_{\mu\nu}^{1 \times 2} \right) \quad (12.26)$$

where $W^{1 \times 2}$ is the rectangular loop $\mathcal{L}_1^{(6)}$ shown in Fig. 7. (Iwasaki used the notation $(1 - 8c_0)W_{\mu\nu}^{1 \times 1} + c_0W_{\mu\nu}^{1 \times 2}$ with $c_0 = -0.331$ which is the same as Eq. 12.26 with a rescaled β .) This action shows better rotational invariance and scaling of $T_c/\sqrt{\sigma}$. It is being pursued by the JLQCD collaboration [92].

12.5. $b = \sqrt{3}$ blocking transformation

The $\sqrt{3}$ blocking transformation has certain very nice properties for 4-dimensional gauge theories in addition to being a smaller scaling factor than the conventional $b = 2$ [89]. The block cell consists of a site and its eight neighbors. As a result the transformation has much higher rotational symmetry for both fermions and gauge fields [89]. Gupta and Patel have studied the flow under blocking using Monte Carlo Renormalization Group methods. They investigated the effect of (i) truncation and (ii) variation in the c_i as a function of β [87]. For phenomenological studies they approximate the RT by a linear trajectory in a four parameter space

$$\begin{aligned} S_{\sqrt{3}} = \beta \operatorname{Re}\{ & c_F^{1 \times 1} \sum \operatorname{Tr} W^{(1 \times 1)} \\ & + c_F^{1 \times 2} \sum \operatorname{Tr} W^{(1 \times 2)} \\ & + c_6^{1 \times 1} \sum [\frac{3}{2} (\operatorname{Tr} W^{(1 \times 1)})^2 - \frac{1}{2} \operatorname{Tr} W^{(1 \times 1)}] \\ & + c_8^{1 \times 1} \sum [\frac{9}{8} |\operatorname{Tr} W^{(1 \times 1)}|^2 - \frac{1}{8}] \}, \end{aligned} \quad (12.27)$$

with

$$c_6^{1 \times 1}/c_F^{1 \times 1} = -0.12; \quad c_8^{1 \times 1}/c_F^{1 \times 1} = -0.12; \quad c_F^{1 \times 2}/c_F^{1 \times 1} = -0.04. \quad (12.28)$$

This particular choice of couplings (truncation of the action) was made for the following reasons. The coupling for the plaquette in the adjoint representation, $c_8^{1 \times 1}$, was included to avoid the phase structure in the fundamental-adjoint plane (see Section 15); that in the 6-dimensional representation, $c_6^{1 \times 1}$, since it differs from $c_8^{1 \times 1}$ only at $O(a^4)$ and higher, and the 1×2 rectangle in the fundamental representation, $c_F^{1 \times 2}$, as it is the simplest loop that kills the $O(a^2)$ terms *à la* Symanzik.

Classical expansion, see Eq. 7.3, of the loops gives the normalization condition

$$\left[c_F^{1 \times 1} + \frac{5}{2} c_6^{1 \times 1} + \frac{9}{4} c_8^{1 \times 1} \right] + 8c_F^{1 \times 2} = 1, \quad (12.29)$$

where I have lumped together the first three terms coming from $W^{(1 \times 1)}$ as the effective contribution of the plaquette. Using the ratios given in

Eq. 12.28, one gets a very high weight $c^{(4)} = 1/0.11$, which as shown in [51] is correlated with a suppression of dislocations versus physical instantons. Similarly, the coefficient of the leading discretization correction $\mathcal{O}_1^{(6)}$, using Table 4, is

$$\frac{1}{24} \left[c_F^{1 \times 1} + \frac{5}{2} c_6^{1 \times 1} + \frac{9}{4} c_8^{1 \times 1} \right] + \frac{5}{6} c_F^{1 \times 2}. \quad (12.30)$$

For the ratios given in Eq. 12.28 this is not zero, but over-corrected by the proposed strength of $W^{(1 \times 2)}$. At $1/a \approx 2$ GeV, this over-correction is larger than that for the tadpole improved Lüscher-Weisz *Ansatz*, similar to $\mathcal{S}_{Iwasaki}$, and smaller than that in $\mathcal{S}_{QCDTARO}$.

Early spectrum studies and measurement of the glueball spectrum [93] to test improvement were inconclusive. In hindsight the failure to quantify improvement is due to poor statistics and not looking at quantities like rotational invariance at strong coupling, where the lattice artifacts show up clearly for the simple Wilson action. Recently this action has also been used to study lattice topology and the results show that short distance artifacts are suppressed [51].

12.6. $b = 2$ blocking transformation

The QCDTARO collaboration has also investigated the flow of the action under blocking [94,95]. They use a $b = 2$ blocking transformation and evaluate the couplings on blocked lattices using Schwinger-Dyson equations. Their overall conclusion is that the blocked action, to a very good approximation, can be represented by the two couplings $W^{1 \times 1}$ and $W^{1 \times 2}$. Their estimate of the RT cannot be approximated by a linear relation between the coefficients. An approximate *Ansatz* in the range of couplings simulated is

$$\mathcal{S}_{QCDTARO} = \beta \sum_i \sum_{\mu, \nu} \sum_x \left(W_{\mu\nu}^{1 \times 1} - 0.11 W_{\mu\nu}^{1 \times 2} \right) \quad (12.31)$$

Note that the coefficient of $W^{1 \times 2}$ is larger than the tadpole improved Luscher-Weisz, Iwasaki, or the $b = \sqrt{3}$ actions.

12.7. Classically Perfect Action

Hasenfratz and Niedermayer showed that one can derive the classically perfect action for asymptotically free-field theories using a saddle-point integration of the RGT about $g = 0$ [96,97]. They use the $b = 2$ blocking transformation proposed by Swendsen [98]. The practical issue here is

Table 5

The current coefficients of the truncated "perfect" action with 4 and 6 link loops.

Loop	c_1	c_2	c_3	c_4
1×1	3.248	-1.580	0.1257	0.0576
Twisted 6 link	-0.2810	0.0051	0.0049	-0.0096

whether the saddle-point integration gives an action that can be approximated by a few local terms. The application of this method to $SU(3)$ and the tuning of the truncated version is still under investigation [99]. The current version of the truncated classical fixed-point action (CFP) is [100]

$$\mathcal{S}_{CFP} = \beta \sum_{i,x} \{c_1^i (1 - \text{Re Tr } W^i) + \dots + c_4^i (1 - \text{Re Tr } W^i)^4\}. \quad (12.32)$$

where the sum over i involves two kinds of loops, the plaquette and the 6-link loop that twists around a 3-cube ($\mathcal{L}_2^{(6)}$ in Fig. 7). Higher powers of the loops arise naturally in the saddle-point integration. Since the strengths decrease rapidly, the authors choose to truncate at the fourth power. The choice of the particular 6-link loop is motivated by the desire to keep the action local in time and to help improve rotational invariance. The values of coefficients $c_1 - c_4$ for the two types of loops are given in Table 5.

At this point full control over stability with respect to truncation has not been achieved. The second issue is whether this *Ansatz*, obtained at $a = 0$, continues to be a useful parameterization at $1/a \sim 2$ GeV where simulations are done, *i.e.* is the renormalized trajectory approximately linear. To test this requires a non-perturbative tuning about $1/a \sim 2$ GeV using the same $b = 2$ blocking transformation used in the saddle-point integration. The calculations necessary to do this tuning have not yet been initiated. As a result this *Ansatz* is considered as work in progress. Nevertheless, the tests for improvement in the scaling of T_c , string tension, topological susceptibility, and in the restoration of rotational symmetry are already very encouraging [100,52].

12.8. Summary

The question – which is the most improved gauge action – is still far from settled. As should be obvious, the drawback of the renormalization group improved actions is that the expansion is in the size of the loops or the dimensionality of the representation and not in the dimensionality of the

irrelevant operators. Since each loop contributes to operators of all dimensions it is not clear how to parameterize the residual error ($O(a^2) + O(a^4)$, or $O(a^4)$, or higher). Also, implementing the equations of motions (removing the freedom associated with redundant directions) has not been easy. So there is not yet a consensus on the obvious minimal set of operators that should be kept. Different operators have been added based on (i) their importance in perturbative expansions, (ii) size of the coefficients in blocked actions (keep in mind that the freedom of the choice of the blocking transformation and consequently the associated RT effects these coefficients), and (iii) specific incorporation of non-perturbative effects; for example the inclusion of a negative adjoint coupling helps avoid the phase structure in the fundamental-adjoint plane (see Section 15), and adding loops mimicking the $F^{\mu\nu}\tilde{F}_{\mu\nu}$ operator help give scale invariant instantons [100,101]. One hopes that a more precise method for tuning the gauge action will emerge in the future. Fortunately, in spite of this ambiguity, the situation is very encouraging. The improvement seen in the restoration of rotational invariance (measured using $q\bar{q}$ potential as shown in Fig. 8), glueball spectrum, and finite temperature studies is remarkable. To a good approximation, most of these actions are very well represented by adding just the $W^{(1\times 2)}$ term with a coefficient larger than the value suggested by tadpole improvement, and keeping a few higher representations of the plaquette. Such a truncated space keeps the computational problem tractable.

The bottom line is that even though there is not a commonly accepted improved gauge action, the improvements seen with any of the above actions are significant. At present the more pressing challenge, in terms of the size of the errors and the computational cost, is improving the Dirac discretization. This is discussed next.

13. $O(a)$ improvement in the Dirac Action

The first attempt to improve the Wilson fermion action by removing the $O(a)$ corrections was by Hamber and Wu [102]. They added a two link term to the Wilson action

$$\mathcal{S}_{HW} = A_W + \frac{r\kappa}{4a} \sum_{x,\mu} \bar{\psi}^L(x) \left[U_\mu(x) U_\mu(x + \hat{\mu}) \psi^L(x + 2\hat{\mu}) + U_\mu^\dagger(x - \hat{\mu}) U_\mu^\dagger(x - 2\hat{\mu}) \psi^L(x - 2\hat{\mu}) \right]$$

$$\equiv \sum_{x,y} \bar{\psi}_L(x) M_{xy}^{HW} \psi_L(y). \quad (13.1)$$

It is easy to check by Taylor expansion that, at the classical level, the $O(a)$ artifact in \mathcal{S}_W is cancelled by the HW term. One can include quantum corrections by calculating the relative coefficient in perturbation theory. It turns out that this action has not been pursued – in the 1980's due to the lack of computer power, and in the 1990's due to the merits of the Sheikholeslami-Wohlert (called SW or the clover) action [20].

13.1. The Sheikholeslami-Wohlert (clover) Action

SW proposed adding a second dimension five operator to \mathcal{S}_W , the magnetic moment term, to remove the $O(a)$ artifacts without spoiling Wilson's fix for doublers. The resulting SW action is

$$\mathcal{S}_{SW} = \mathcal{S}_W - \frac{iaC_{SW}\kappa r}{4} \bar{\psi}(x) \sigma_{\mu\nu} F_{\mu\nu} \psi(x). \quad (13.2)$$

The fact that this action is $O(a)$ improved for on-shell quantities cannot be deduced from a Taylor expansion. To show improvement requires an enumeration of all dimension 5 operators, the use of equations of motion to remove some of them, and incorporating the remaining two ($m^2 \bar{\psi}\psi$ and $mF^{\mu\nu}F_{\mu\nu}$) by renormalizing $g^2 \rightarrow g^2(1 + b_g ma)$ and $m \rightarrow m(1 + b_m ma)$. The construction of the SW action has been explained in detail by Prof. Lüscher at this school, so instead of repeating it, I direct you to his lectures. I will simply summarize results of tuning C_{SW} using both perturbative and non-perturbative methods. A list of perturbatively improved values of C_{SW} that have been used in simulations is as follows.

$C_{SW} = 1$	tree-level improvement
$C_{SW} = 1 + 0.2659(1)g_0^2$	1-loop improved [103]
$C_{SW} = 1/u_0^3$	tadpole improved
$C_{SW} = 1/u_0^3(1 + 0.0159g_f^2)$	1-loop tadpole improved

In the 1-loop TI result u_0 was chosen to be the fourth root of the plaquette. As an exercise convince yourself that using the expectation value of the link in Landau gauge gives a value closer (1.6) to the non-perturbatively improved result (1.77 at $g = 1$) given below. For each of the above values of C_{SW} , the errors are reduced from $O(a)$ (Wilson action) to at least $O(\alpha_s a)$.

The recent exciting development, also reviewed by Lüscher, is that the axial Ward identity C_{SW} can be used to non-perturbatively tune C_{SW} and remove all $O(a)$ artifacts provided one simultaneously improves the coupling g^2 , the quark mass m_q , and the currents [104]. Estimates of

C_{SW} have been made at a number of values of g for the quenched theory with Wilson's gauge action [104,105]. These data can be fit by [105]

$$C_{SW} = \frac{1 - 0.6084g^2 - 0.2015g^4 + 0.03075g^6}{1 - 0.8743g^2}. \quad (13.3)$$

for $6/g^2 > 5.7$.

The advantages of the SW action is that it is local and leaves perturbation theory tractable [106]. Adding the clover term is only a $\sim 15\%$ overhead on Wilson fermion simulations. An example of improvement obtained using the SW versus Wilson action for fixed quark mass ($M_\pi/M_\rho = 0.7$) is that the deviations from the $a = 0$ values in $M_N/\sqrt{\sigma}$ and $M_{vector}/\sqrt{\sigma}$ at $a = 0.17$ fermi are reduced from $30 - 40\% \rightarrow \sim 3 - 5\%$ [105]. Further tests and evaluation of the improvement are being done by a number of collaborations. I expect a number of quenched results to be reported at the LATTICE 98 conference, and these should appear on the e-print archive by late 1998.

13.2. D234c action

As mentioned before, there are two sources of discretization errors: (i) approximating derivatives by finite differences, and (ii) quantum corrections, for example those due to modes with momentum $> \pi/a$ that are absent from the lattice theory. Alford, Klassen, and Lepage have proposed improved gauge and Dirac actions that reduce (i) by classical improvement and (ii) by tadpole improvement [107,79]. Their current proposal for isotropic lattices, called the **D234c action** as it involves second, third, and fourth order derivatives, is

$$M_{D234c} = m(1 + 0.5ram) + \sum_{\mu} \left\{ \gamma_{\mu} \Delta_{\mu}^{(1)} - \frac{C_3}{6} a^2 \gamma_{\mu} \Delta_{\mu}^{(3)} \right\} \\ + r \sum_{\mu} \left\{ -\frac{1}{2} a \Delta_{\mu}^{(2)} - \frac{C_F}{4} a \sum_{\nu} \sigma_{\mu\nu} F_{\mu\nu} + \frac{C_4}{24} a^2 \Delta_{\mu}^{(4)} \right\}. \quad (13.4)$$

Here $\Delta_{\mu}^{(n)}$ is the n^{th} order lattice covariant derivative (symmetric and of shortest extent). The term proportional to C_3 is precisely what is needed to kill the leading discretization error in $\Delta_{\mu}^{(1)}$ as shown in Eq. 7.6. The terms proportional to r are generated by a field redefinition and thus represent a redundant operator. They, therefore leave unaltered the $O(a)$ improvement of the naive action for on-shell quantities. The three terms in it are respectively the Wilson term, that is again used to remove the

doublers, clover term (for $F_{\mu\nu}$ they use an $O(a^2)$ improved version) as in the SW action needed to remove the $O(a)$ error introduced by the Wilson fix, and an a^3 correction. To include quantum corrections to the classical values $C_3 = C_F = C_4 = 1$ they resort to tadpole improvement, *i.e.*, they divide all links appearing in M_{D234c} by u_0 , which is taken to be the mean link in Landau gauge.

Quantum corrections generally also induce additional operators, however, in perturbation theory these start at $O(\alpha_s)$. After tadpole improvement the coefficients of α_s for such operators are all expected to be small, so one hopes that their contributions can be neglected. The philosophy/motivation is analogous to dropping the small $\mathcal{L}_2^{(6)}$ term in the tadpole improved Luscher-Weisz pure gauge action given in Eq. 12.14.

For two of the coefficients, C_F and C_3 , the authors have carried out non-perturbative tests of the accuracy of tadpole improvement [79]. Since only the operator $\Delta_\mu^{(3)}$ breaks rotational symmetry at $O(a^2)$, therefore C_3 can be tuned non-perturbatively by studying rotational invariance of dispersion relations. Similarly, the clover term and the Wilson term break chiral symmetry at $O(a)$, thus their relative coefficient C_F can be tuned by studying the axial Ward identity. Results of their tests show that errors in the spectrum and dispersion relations for momenta as high as $p = 1.5/a$ are $\approx 10\%$ even at $a = 0.4$ fermi! The authors are extending their analyses of the $D234$ action to anisotropic lattices which, as will be discussed in Section 18.4, have already yielded rich dividends in the calculation of glueball masses.

13.3. NRQCD action

In the simulation of heavy quarks one has to worry about $O(m_H a)$ errors in addition to $O(pa)$ and $O(\Lambda_{QCD} a)$. For charm and bottom quarks these errors are $O(1)$ in the range of lattice spacings one can reasonably simulate. Lepage and collaborators, recognizing that the heavy quark in heavy-heavy or heavy-light systems is, to a good approximation, non-relativistic, developed a non-relativistic formulation (NRQCD) valid for $m_H a > 1$ [108]. In this approach, the action for the two-component Pauli spinors is obtained by a Foldy-Wouthuysen transformation, and relativistic and discretization corrections are added as a systematic expansion in $1/M$ and p/M . One version of the NRQCD action is [109]

$$\mathcal{L} = \bar{\psi}_t \psi_t \quad (13.5)$$

$$- \bar{\psi}_t \left(1 - \frac{a\delta H}{2} \right)_t \left(1 - \frac{aH_0}{2n} \right)_t U_4^\dagger \left(1 - \frac{aH_0}{2n} \right)_{t-1} \left(1 - \frac{a\delta H}{2} \right)_{t-1} \psi_{t-1},$$

where ψ is the two-component Pauli spinor and H_0 is the nonrelativistic kinetic energy operator,

$$H_0 = -\frac{\Delta^{(2)}}{2M_0}. \quad (13.6)$$

δH includes relativistic and finite-lattice-spacing corrections,

$$\begin{aligned} \delta H = & -\frac{gc_1}{2M_0} \boldsymbol{\sigma} \cdot \mathbf{B} \\ & + \frac{igc_2}{8(M_0)^2} (\boldsymbol{\Delta} \cdot \mathbf{E} - \mathbf{E} \cdot \boldsymbol{\Delta}) - \frac{gc_3}{8(M_0)^2} \boldsymbol{\sigma} \cdot (\boldsymbol{\Delta} \times \mathbf{E} - \mathbf{E} \times \boldsymbol{\Delta}) \\ & - \frac{c_4(\Delta^{(2)})^2}{8(M_0)^3} + \frac{c_5 a^2 \Delta^{(4)}}{24M_0} - \frac{c_6 a (\Delta^{(2)})^2}{16n(M_0)^2}. \end{aligned} \quad (13.7)$$

Here \mathbf{E} and \mathbf{B} are the electric and magnetic fields, $\boldsymbol{\Delta}$ and $\Delta^{(2)}$ are the spatial gauge-covariant lattice derivative and laplacian, and $\Delta^{(4)}$ is a lattice version of the continuum operator $\sum D_i^4$. The parameter n is introduced to remove instabilities in the heavy quark propagator due to the highest momentum modes [108]. Again, the tree-level coefficients $c_i = 1$ can be tadpole-improved by rescaling all the link matrices by u_0 .

This formulation has been used very successfully to calculate α_s from the epsilon spectrum [110] as discussed in Section 19, the heavy-heavy and heavy-light spectrum [111], and the heavy-light decay constants f_B and f_{B_s} [109]. The drawback of NRQCD is that it is an effective theory and one cannot take $a \rightarrow 0$. In fact, to simulate charm and bottom quarks, one has to work in a narrow window $1 \lesssim 1/a \lesssim 2.5$ GeV to keep $m_H a > 1$. The good news is that current data [110,111,109] show that the discretization and relativistic corrections are small and under control with the terms included in Eq. 13.7.

13.4. Perfect action

The development of a perfect Dirac action is not complete. The method is same as used for the gauge action and discussed in [90]. The bottleneck is that the saddle point integration gives an action that even in the free field limit involves a number of sites, and the gauge connections between any pair of sites is the average of a large number of paths [112]. First tests of such an action truncated to a 3^4 hypercube have been carried out by Degrand and collaborators [100]. From these it is not clear whether the improvement, compared to other discretization schemes, justifies the extra cost of simulating such an action. Hopefully better ways to tune and test such actions will be formulated soon.

13.5. Fermilab action

The Fermilab group has proposed a "Heavy Wilson" fermion action [113]. This action is meant to interpolate between the light $O(a)$ improved SW-clover action and the NRQCD action for heavy quarks. The goal is to remove not only $O(a)$ discretization corrections but also take into account all powers of ma errors so that charmonium and bottomonium systems can be simulated with the same reliability as light-light physics. This improvement is especially targeted towards improving heavy-light matrix elements. So far no tests have been reported with the proposed action, therefore I will not reproduce the action here but refer interested readers to [113] for details.

13.6. Summary

In analogy to discussions of improved gauge actions, having seen tangible evidence for improvement in quenched lattice data with each of the above formulations, the question is, which action is the most efficient for a comprehensive study? The deciding factor, in my opinion, will be dynamical simulations since the prefactor and the scaling exponent ($L^{\sim 10}$) makes the cost of update of background configurations prohibitive without improvement on today's computers. Dynamical simulations using improved actions have just begun, so it will be a few years before enough data can be generated to determine the advantages and disadvantages of different approaches.

Finally, I should mention a recent novel approach -- domain wall fermions -- for improving the chiral properties of Wilson fermions [114]. In this approach the system (domain) is 5-dimensional and the fermions live on 4-dimensional surfaces that make up the opposite boundaries (walls). The 4-dimensional gauge fields on the two walls are identical, whereas the fermion zero-modes have opposite chirality. The overlap between these modes is exponentially suppressed by the separation in the fifth dimension. Consequently, the chiral symmetry breaking contributions are also suppressed, and results with good chiral properties for vector like theories like QCD can be obtained by adding the contribution from the two domain walls. First simulations of weak matrix elements show results with good chiral properties and in general look extremely promising [115]. Since I do not do justice to this exciting approach in these lectures, interested readers should consult [115] and references therein.

14. Confinement and Asymptotic Freedom

For any theory to provide a successful description of strong interactions it should simultaneously exhibit the phenomena of confinement at large distances and asymptotic freedom at short distances. Lattice calculations support the hypothesis that for non-abelian gauge theories the two domains are analytically connected, and confinement and asymptotic freedom co-exist. Similarly, one way to show that QCD is the correct theory of strong interactions is that the coupling extracted at various scales (using experimental data or lattice simulations) is unique in the sense that its variation with scale is given by the renormalization group. The data for α_s is reviewed in Section 19. In this section I will discuss what these statements mean and imply.

14.1. Asymptotic Freedom

The term asymptotic freedom implies that the running coupling $g \rightarrow 0$ as the momentum scale of the probe $\mu \rightarrow \infty$. This is characterized by the β -function

$$\mu \frac{\partial g}{\partial \mu} = -\beta(g) = -(\beta_0 g^3 + \beta_1 g^5 + \dots) \quad (14.1)$$

where $\beta(g)$ is positive for non-abelian gauge groups. In this case $g = 0$ is a UV stable fixed point of the theory, and about this point the theory can be analyzed using standard perturbative expansions. This novel behavior was confirmed by 't Hooft, Politzer, and by Gross and Wilczek who showed that the perturbative β -function satisfies Eq. 14.1 [116], *i.e.*, for N colors and n_f active flavors

$$\begin{aligned} \beta_0 &= \left(\frac{11N - 2n_f}{3} \right) / 16\pi^2, \\ \beta_1 &= \left(\frac{34N^2}{3} - \frac{10Nn_f}{3} - \frac{n_f(N^2 - 1)}{N} \right) / (16\pi^2)^2. \end{aligned} \quad (14.2)$$

These two leading terms in the expansion of $\beta(g)$ are gauge and regularization scheme invariant. From these it is easy to see that $\beta(g)$ is positive for $n_f < 8$. This result was essential in establishing QCD as the theory of strong interactions for it explained existing experimental data which showed that the strength of strong interactions decreases as the momentum exchanged in a process increases.

Asymptotic freedom, Eq. 14.1, implies that QCD dynamically generates a mass scale. Integrating Eq. 14.1 from momentum scale μ_1 to μ_2 with

$\mu_2 > \mu_1$ and keeping only the β_0 term to simplify the standard calculation, gives

$$\frac{1}{2\beta_0 g^2(\mu_2)} - \frac{1}{2\beta_0 g^2(\mu_1)} = \log \frac{\mu_2}{\mu_1}, \quad (14.3)$$

i.e. the coupling constant of non-abelian gauge theories depends logarithmically on the momentum scale of the process. Rewriting Eq. 14.3

$$\begin{aligned} \frac{1}{2\beta_0 g^2(\mu)} - \log \mu &= \text{constant} \\ \Rightarrow \exp \left\{ \frac{1}{2\beta_0 g^2(\mu)} \right\} &= \frac{\mu}{\Lambda_{QCD}} \\ \Rightarrow \alpha_s(\mu) = \frac{g^2(\mu)}{4\pi} &= \frac{1}{8\pi\beta_0 \log \frac{\mu}{\Lambda_{QCD}}} \end{aligned} \quad (14.4)$$

introduces Λ_{QCD} , the invariant scale of the theory with dimensions of mass. Thus QCD, a theory with a dimensionless coupling constant and no intrinsic mass scale in the absence of quark masses, dynamically generates a mass scale. This happens because in order to specify g one has to also specify a momentum scale at which it is defined. Extending the above analysis to include β_1 in Eq. 14.1 gives

$$\Lambda_{QCD} = \lim_{\mu \rightarrow \infty} \mu \left(\frac{1}{\beta_0 g^2(\mu)} \right)^{\frac{\beta_1}{2\beta_0^2}} \exp \left[- \frac{1}{2\beta_0 g^2(\mu)} \right] \equiv \mu f_p(g(\mu)) \quad (14.5)$$

This 2-loop definition of Λ_{QCD} is not unique; the value of Λ_{QCD} depends on the precise relation between g and μ . However, once the value of Λ is determined in one scheme it can be related to that in any other perturbative scheme. For example, in the lattice regularized theory $\Lambda_{lattice}$ is also defined by Eq. 14.5 but with μ replaced by $1/a$. Then to 1-loop

$$\frac{\Lambda_{QCD}}{\Lambda_{lattice}} = \mu a \exp \left\{ - \frac{1}{2\beta_0} \left[\frac{1}{g^2(\mu)} - \frac{1}{g^2(a)} \right] \right\}. \quad (14.6)$$

In perturbation theory the two coupling constants are related as

$$g^2(\mu) = g^2(a) \left\{ 1 - \beta_0 g^2(a) \left(\log(\mu a)^2 - \log C^2 \right) + O(g^4) \right\} \quad (14.7)$$

By substituting Eq. 14.7 into Eq. 14.6 one finds

$$\Lambda_{QCD} = C \Lambda_{lattice} \quad (14.8)$$

i.e. the two constants, Λ_{QCD} and $\Lambda_{lattice}$, are related by a multiplicative constant. To calculate C requires knowing the finite part of the coupling

Table 6

The relation between Λ_{MOM} and $\Lambda_{\overline{MS}}$ and Λ_{latt} as a function of the number of active flavors. The results for $\Lambda_{MOM}/\Lambda_{latt}$ with g defined by the triple gluon vertex are in Feynman gauge.

n_f	0	1	2	3	4
$\Lambda_{\overline{MS}}/\Lambda_{latt}$	28.8	34.0	41.1	51.0	65.5
$\Lambda_{MOM}/\Lambda_{latt}$	83.4	89.4	96.7	105.8	117.4

constant renormalization to 1-loop in both the lattice and continuum regularization schemes [117,36]. The results are listed in Table 6 for Λ_{MOM} and $\Lambda_{\overline{MS}}$. It is important to note that even though the above definition and estimate of Λ_{QCD} is made using perturbation theory, it is intrinsically a non-perturbative quantity.

All dimensionful quantities in lattice simulations are measured in units of the lattice spacing, for example one measures Ma and not M . Such dimensionless quantities vary with g as

$$M_i a = c_i f_i(g) \equiv c_i \Lambda_{non-pert}^i a. \quad (14.9)$$

where the c_i are constants representing the continuum value and f_i are scaling functions. Lattice correlation lengths $\xi_i \equiv 1/M_i a$ diverge as the fixed point at $g = 0$ is approached. This divergence is an artifact – it is the unit of measurement a that is going to zero and not M_i . This is precisely what one wants to have happen to get rid of the scaffold (lattice) erected to regularize the theory. Non-perturbative renormalization consists of taking the continuum limit holding some physical quantity M_i fixed while allowing $a \rightarrow 0$ according to $f_i(g)$. The particular quantity M_i (string tension, or nucleon mass, *etc.*) held fixed defines the renormalization condition, and the constraint between a and g is the scaling relation.

– **Scaling:** The hypothesis of scaling is that all $f_i(g)$ converge to a universal scaling function $f(g) = \Lambda_{non-pert} a$. The renormalization condition is then independent of the state used to define it. For $g < g_{scaling}$ each lattice quantity is characterized by the non-perturbative number c_i and all dimensionless ratios of physical quantities are independent of g .

– **Asymptotic Scaling:** Close to the fixed point at $g = 0$, perturbation theory is apposite and all $f_i(g) \rightarrow f_{pert}(g) \rightarrow \Lambda_{QCD} a$.

Lattice simulations, therefore, provide continuum results to within some error tolerance from simulations done in a neighborhood of $g = 0$. In this region scaling holds. For a fixed error criteria, the extent of this scaling window can be enlarged by improving the action and operators as discussed

in Section 8. For simulations done outside the scaling window it is imperative that a reliable extrapolation to $a = 0$ be done. Obviously, one way to get an upper bound on the errors is to compare the measured scaling function with $f_{pert}(g)$. The good news is that simulations in the past few years show that corrections to scaling can be reduced to a few percent level for lattice spacings with $5 \lesssim \xi/a \lesssim 10$ (where $\xi/a = 1/\sqrt{\sigma a}$). This corresponds to $2 \lesssim 1/a \lesssim 4$ GeV. Consequently, realistic results with physical light quarks can be obtained on lattices of side $\sim 75 - 150$, and for many quantities, like those involving the heavy quarks, much smaller lattices are sufficient. The goal of the lattice approach is twofold. First, improve the discretization of the action and operators to extend the scaling region to stronger coupling, and second do simulations at a number of points within this region and remove the remaining errors by extrapolation to $a = 0$.

An implication of the above discussion is that if all states in QCD have zero mass in the continuum limit then $g = 0$ would be a trivial fixed point as the renormalized coupling g_R would also be zero. On the other hand one can ask whether some of the states of QCD can be massless while the others are massive? It is interesting to note that we can actually say something about this aspect of the spectrum of QCD once the scaling behavior is fixed by Eq. 14.9. In fact, under the following two assumptions, the pure gauge sector of QCD will have a mass gap:

- There are no zero mass states at any finite non-zero value of g , *i.e.* the dimensionless quantity Ma does not vanish at any non-zero value of the lattice spacing a in a region that lies in the same thermodynamic phase as the continuum limit at $a = 0$.
- There exists only one relevant coupling g in QCD, and corresponding to it a single universal scaling function defined by say Eq. 14.5. In this case the scaling behavior of all observables is fixed and all mass-ratios have to stay finite as the continuum limit is taken.

These two assumptions are borne out by present numerical data, *i.e.* there is no evidence for a massless state in the pure gauge sector. Also there is no indication of a phase boundary separating the region where simulations have been done and $g = 0$. Therefore, combining lattice data with renormalization group scaling, QCD predicts that the lightest glueball state is massive.

This prediction does not change with the introduction of n_f flavors of massless quarks into the theory. Spontaneously broken chiral symmetry will give rise to $n_f^2 - 1$ massless pions in the spectrum. To take the chiral limit one has to tune the quark masses to zero to define the physical world. This tuning has to be done at each value of a . Thus one cannot form mass-

ratios with respect to these Goldstone states as the corresponding $c_i(g)$ are tuned to zero.

To conclude, once we have verified that scaling (perturbative or non-perturbative) exists then the only unknowns needed to predict the spectrum of QCD are the constants c_i . Asymptotic freedom simplifies QCD by providing an analytical prediction of the universal scaling behavior as $g \rightarrow 0$. Current data show that the coefficients c_i , which are intrinsically non-perturbative, can be extracted using LQCD from simulations done at $a \sim 0.1$ fermi.

14.2. Confinement

There are two ways one can test whether a theory confines. One, demonstrate that the free energy of an isolated charge is infinite; two, show that the potential energy between two charges grows with distance. The first test shows that it takes infinite energy to add an isolated test charge to the system, while the second shows that it requires infinite energy to separate two charges by infinite distance. I shall show that the first is probed by the expectation value of the Wilson line $\langle \mathcal{L} \rangle$ in the pure gauge theory, while the second by the expectation value of Wilson loops $\langle \mathcal{W} \rangle$ or by the correlation function $\langle \mathcal{L}(\tau) \mathcal{L}^\dagger(0) \rangle$.

14.2.1. Wilson Loops

The extra action associated with an external charge placed in a gauge field is $S_J = \int d^4x J_\mu^a A_\mu^a$. For a point charge $J(x) = \delta^4(x)$, S_J is given by the path ordered integral of the gauge field along the world line of the charge. Thus

$$\langle \mathcal{W} \rangle = e^{-S_J} \sim e^{-V(R)T} \quad (14.10)$$

for a $R \times T$ rectangular Wilson loop. One can regard the expectation value of a Wilson loop as the creation of a $q\bar{q}$ pair at time $T = 0$ at point $R/2$, separated instantaneously to R and 0 , allowed to evolve for time T , and then allowed to annihilate. Amongst the many terms (instantaneous creation, annihilation, ...) that contribute to the action for this procedure is that due to the potential between two static charges separated by distance R and integrated for time T , i.e. $V(R)T$. Thus

$$V(R) = \lim_{T \rightarrow \infty} -\frac{1}{T} \log \langle \mathcal{W}(R, T) \rangle. \quad (14.11)$$

The simplest *Ansatz* for a confining potential is the Cornell potential

$$V(R) = V_0 + \sigma R - \frac{\alpha}{R}, \quad (14.12)$$

where σ is the string tension. At large distances σR dominates, while at short distances it is the Coulomb term α/R that is relevant. Such a potential, therefore, simultaneously exhibits confinement and asymptotic freedom. For such linear potentials large Wilson loops show an area law, $\log\langle\mathcal{W}(R, T)\rangle \propto RT$, and thus probe confinement. A state-of-the-art calculation by the Wuppertal collaboration [118] of $V(R)$ for SU(3) gauge theory is shown in Fig. 10. It shows the expected behavior – a linear rise at large R and a Coulomb behavior at small R .

One can also extract the potential from the correlation function $\langle\mathcal{L}(\mathcal{R})\mathcal{L}^\dagger(0)\rangle$ of Wilson lines, defined in Section 14.2.3, pointing in the T direction and separated by distance R . It is easy to see that this correlation function is equivalent to a Wilson loop with the sides at $T = 0, N_T$ removed. The Wilson lines are closed due to the periodic boundary conditions on the gauge links and thus individually gauge invariant. This correlation function measures the action due to two static test charges separated by distance R .

Measurements of the potential provide a way to set the scale of lattice calculations. One approach is to use σ . The string tension can be extracted from the slope of $V(R)$ at large R or from the asymptotic behavior of Creutz ratios [120]

$$\sigma = \lim_{\substack{T \rightarrow \infty \\ R \rightarrow \infty}} -\log \left(\frac{\langle\mathcal{W}(R, T)\rangle\langle\mathcal{W}(R+1, T+1)\rangle}{\langle\mathcal{W}(R+1, T)\rangle\langle\mathcal{W}(R, T+1)\rangle} \right). \quad (14.13)$$

It can be measured very accurately in lattice simulations, however, its physical value is either inferred from the Regge slope or from potential models. Regge phenomenology gives the relation $M_l^2(l) = l/\alpha' + c$ where M_l is the mass of a state on the Regge trajectory with orbital angular momentum l . To get a relation between σ and α' two approaches are used: (1) string models which give $\sigma = 1/2\pi\alpha'$, and (ii) potential models which give $\sigma = 1/8\alpha'$ [123]. The value of α' depends on the quark mass. For example α' from the light hadrons, $c\bar{c}$, and $b\bar{b}$ is $0.85 \rightarrow 0.85/2 \rightarrow \approx 0.85/4.5 \text{ GeV}^{-2}$. It is customary to choose the value $\alpha' = 0.85 \text{ GeV}^{-2}$ from the light sector as it is expected to be most sensitive to the large distance part of the potential. Combining this with the string *Ansatz* $\sigma = 1/2\pi\alpha'$ yields $\sqrt{\sigma} \approx 440 \text{ MeV}$ [124,125]. A survey of results from potential models can be found in [126] and one finds σ varies in the range $0.18 - 0.22 \text{ GeV}^2$. In most lattice calculations the value $\sqrt{\sigma} = 440 \text{ MeV}$ is used, however, one should be aware of the significant uncertainty in its extraction.

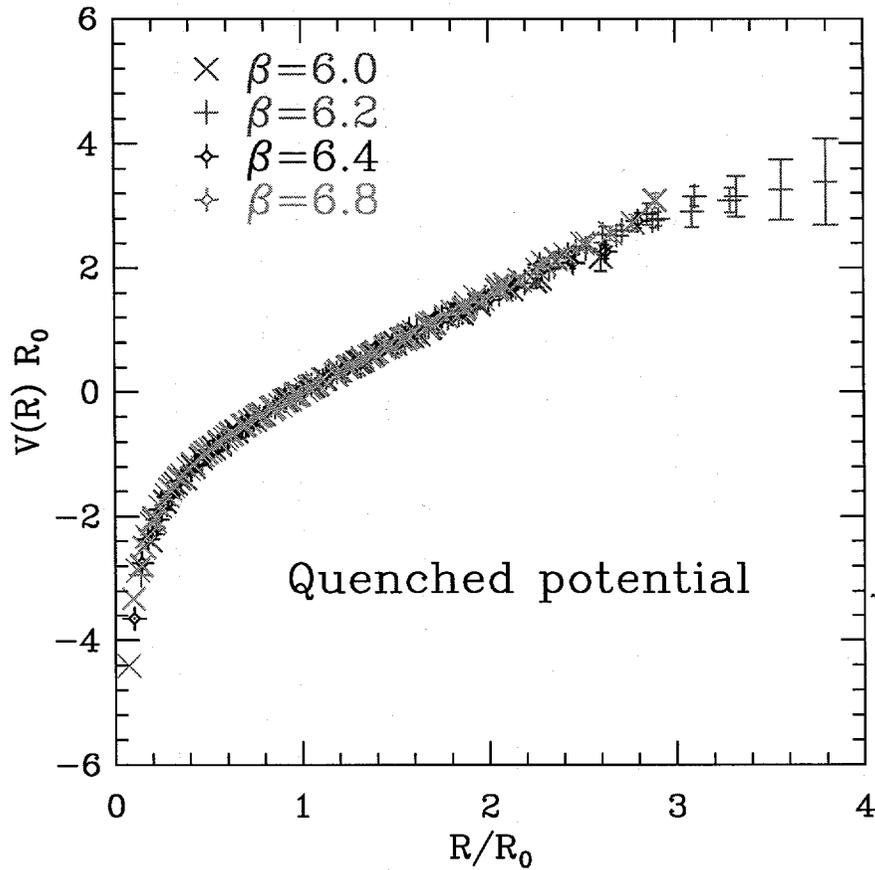


Fig. 10. The static $q\bar{q}$ potential in the quenched approximation obtained by the Wuppertal collaboration [118]. The data at $\beta = 6.0, 6.2, 6.4$ and 6.8 has been scaled by R_0 defined in Eq. 14.14, and normalized such that $V(R_0) = 0$. The collapse of the different sets of data on to a single curve after the rescaling by R_0 is evidence for scaling. The linear rise at large R implies confinement.

An even better quantity, proposed by Sommer [127], is the distance scale R_0 defined through the force between two static charges

$$R^2 \left. \frac{\partial V(R)}{\partial R} \right|_{R=R_0} = 1.65 \quad (14.14)$$

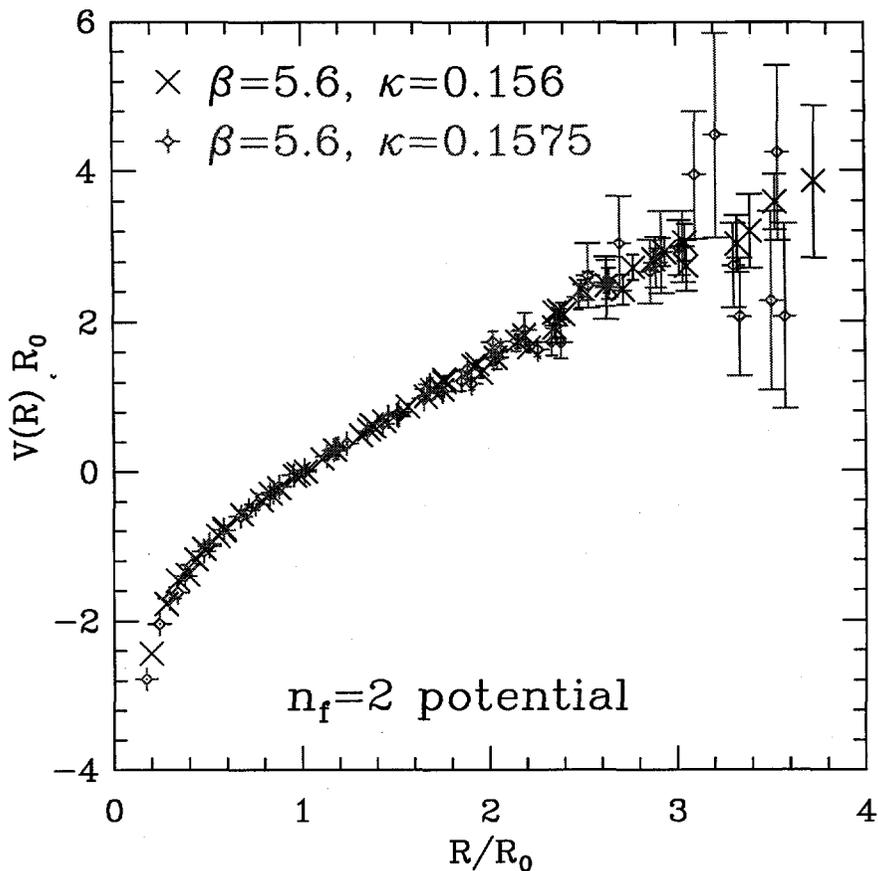


Fig. 11. The static $q\bar{q}$ potential with two flavors of dynamical fermions obtained by the Wuppertal collaboration [119]. The dynamical quark mass in the $\beta = 6.0$, $\kappa = 0.156$ simulation is $\sim 2m_s$ while that at $\kappa = 0.1575$ is $\sim m_s$. The rest is same as in Fig. 10. These data lie lower than the quenched points, however, for these dynamical quark masses there is no evidence yet of string breaking at large R .

The Cornell and Richardson potentials, which have been very successful in predicting levels of charmonium and Upsilon system, yield

$$R_0 \approx 0.49 \text{ fermi} \quad (14.15)$$

Setting the scale using R_0 has the following advantages: (i) its value is not very sensitive to the choice of the phenomenological potential, (ii) it is defined at an intermediate distance where phenomenological potentials are

highly constrained by the $c\bar{c}$ and $b\bar{b}$ spectra, (iii) lattice measurements of $V(R)$ at these intermediate distances can be obtained with high precision, and (iv) this construction is valid for both full and quenched QCD. The status of the current quenched data from three collaborations is shown in Fig. 12. The data agree, and it turns out that over this same range of β the relation $R_0 = 1.18/\sqrt{\sigma}$ holds to a very good approximation. (The σ is extracted from the large R behavior of the $q\bar{q}$ potential as discussed above.) Thus, in subsequent analyses of lattice data I will assume this relation when using either σ or R_0 to set the scale.

For full QCD, the linear rising potential is screened. The string between the $q\bar{q}$ breaks when the energy is large enough to pop an additional $q\bar{q}$ out of the vacuum and create two mesons. This phenomena of screening can be tested by plotting $V(R)$ for different values of the dynamical quark mass and observing the flattening of the linear rise. The value of R/R_0 at which the flattening should occur should decrease with m_q . A recent analyses, again by the Wuppertal collaboration, is shown in Fig. 11 [119]. The data show no flattening – presumably because the sea quark masses used in the update are not light enough to cause string breaking at these R/R_0 .

14.2.2. Strong Coupling Expansion

Confinement, for $\beta \rightarrow 0$, can be demonstrated using strong coupling expansions. I will illustrate this technique by two examples – string tension and the 0^{++} glueball mass.

The strong coupling expansion for $SU(N)$ gauge theory depends on the following identities for integration over link matrices

$$\begin{aligned} \int dg &= 1 \\ \int dg U_{ij} &= 0 \\ \int dg U_{ij}^\dagger &= 0 \\ \int dg U_{ij} U_{kl} &= 0 \\ \int dg U_{ij}^\dagger U_{kl}^\dagger &= 0 \\ \int dg U_{ij} U_{kl}^\dagger &= \frac{1}{N} \delta_{il} \delta_{jk} \end{aligned}$$

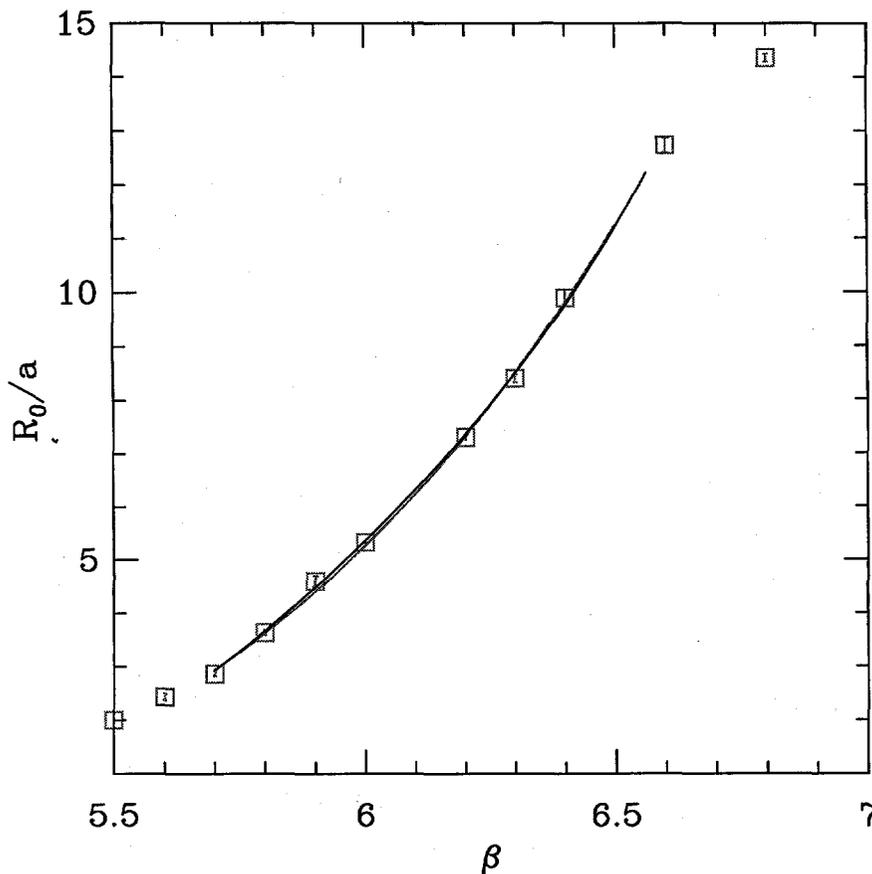


Fig. 12. The scale R_0 extracted from the static $q\bar{q}$ potential. The data in purple are from [118,119,121] and the red line is a fit to these data. The black line is the fit $\log(a/R_0) = -1.6805 - 1.7139(\beta - 6) + 0.8155(\beta - 6.0)^2 - 0.6667(\beta - 6)^3$ to the data by the Alpha collaboration [122].

$$\int dg U_{i_1 j_1} \dots U_{i_N j_N} = \frac{1}{N} \epsilon_{i_1 \dots i_N} \epsilon_{j_1 \dots j_N}. \quad (14.16)$$

The essential point is that the group integration gives a non-zero result only if each link occurs in a combination from which a color singlet can be formed. Eq. 14.16 shows this for the identity, "meson" and "baryon" configuration of links.

The expectation value of the Wilson loop, at small β (large g) for the

plaquette action, can be expanded as follows.

$$\begin{aligned}\langle \mathcal{W}_{RT} \rangle &= \frac{1}{Z} \int dU \mathcal{W}_{RT} e^{\beta/2N(\mathcal{W}_{11}^+ + \mathcal{W}_{11}^-)} \\ &= \frac{1}{Z} \int dU \mathcal{W}_{RT} (1 + \beta/2N(\mathcal{W}_{11}^+ + \mathcal{W}_{11}^-) + \dots)\end{aligned}\quad (14.17)$$

where \mathcal{W}_{11}^+ , \mathcal{W}_{11}^- are the two orientations of the plaquette and the trace over color indices in each loop is implicit. An inspection of Fig. 13A shows that the first non-zero contribution to the integral occurs when the loop \mathcal{W}_{RT} is tiled by elementary plaquettes with the right orientation. Each such plaquette brings a factor of $\beta/2N$ from the expansion and another factor of $1/N$ from the integration. (Note that SU(2) is special and the combined factor is β/N^2 as the two orientations of the loop have the same value, *i.e.* the trace is real.) Thus

$$\langle \mathcal{W}_{RT} \rangle = \frac{\beta^{RT}}{2N^2} (1 + \dots)\quad (14.18)$$

where the leading corrections come from replacing any given tile with a pillbox. Now using Eqs. 14.10 and 14.12 gives

$$\sigma = -\log \frac{\beta}{2N^2} + O(\beta).\quad (14.19)$$

Thus all gauge theories, including U(1), confine in the strong coupling limit. What distinguishes U(1) from non-abelian gauge theories is that the U(1) gauge theory has a phase transition at $g^2 \sim 1$ (see Section 15), and the physical theory (electrodynamics) lies in the weak coupling region which is analytically disconnected from the confining strong coupling limit.

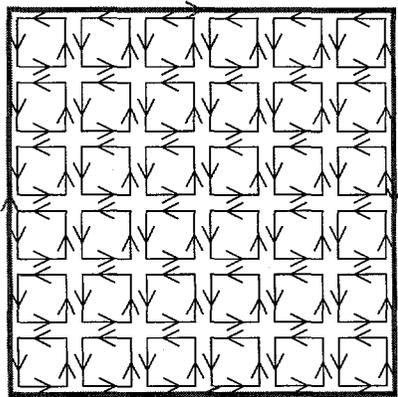
As a second application of the strong coupling expansion consider the calculation of 0^{++} glueball mass. The minimum tiling of the connected plaquette-plaquette correlation function is the rectangular tube consisting of four sides of size $1 \times T$ as shown in Fig. 13B. Thus

$$\langle \mathcal{W}_{11}(T) \mathcal{W}_{11}(0) \rangle \sim e^{-M_{0^{++}}T} = \left(\frac{\beta}{2N^2} \right)^{4T} (1 + \dots),\quad (14.20)$$

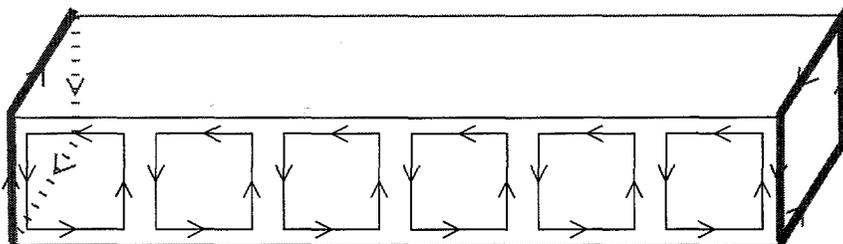
and

$$M_{0^{++}} = -4 \log \frac{\beta}{2N^2} + \dots\quad (14.21)$$

The fact that the lowest order result $M_{0^{++}}/\sqrt{\sigma} = 4$ is very close to present day estimates (see Section 18.3) is fortuitous.



(A) Minimum tiling of a 6x6 Wilson loop.



(B) Tiling of one face of a plaquette-plaquette correlation function.

Fig. 13. Examples of minimum tiling of (A) a 6×6 Wilson loop, and (B) the plaquette-plaquette correlation function.

Exercise: What is the minimal tiling for the 0^{++} correlation function built out of a $R \times R$ Wilson loop. Show that Eq. 14.20 holds independent of the size of the loop.

14.2.3. Wilson/Polyakov Line

The Wilson line, pointing in say the time direction, is defined by the path ordered product

$$\langle \mathcal{L} \rangle(x, y, z) = \text{Tr} \mathcal{P} \prod_{t=1}^{N_T} U_4(x, y, z, t) \quad (14.22)$$

It is gauge invariant due to periodic boundary conditions on gauge fields, $U_4(x, y, z, N_T + 1) = U_4(x, y, z, 1)$. Repeating the above arguments made for the Wilson loop one finds that the free energy of an isolated quark is given by the Wilson/Polyakov line

$$\langle \mathcal{L} \rangle \sim e^{-F_q N_T}. \quad (14.23)$$

Thus, the possibility of having isolated charges in the theory requires F_q be finite.

In addition to gauge invariance, $SU(N)$ gauge theories have a global $Z(N)$ invariance. It consists of multiplying all links on a given time slice by the same element z of $Z(N)$ without changing the action. Under this transformation $\langle \mathcal{L} \rangle \rightarrow z \langle \mathcal{L} \rangle$. Such a global symmetry can be spontaneously broken. Thus we have two possibilities

$$\begin{array}{lll} \text{Broken :} & \langle \mathcal{L} \rangle \neq 0 & \text{Deconfined } (F_q \text{ finite}) \\ \text{Unbroken :} & \langle \mathcal{L} \rangle = 0 & \text{Confined } (F_q \text{ infinite}) \end{array} \quad (14.24)$$

Thus, $\langle \mathcal{L} \rangle$ is an order parameter, and tests for confinement in pure gauge theories. It is also used to probe the location and order of the finite temperature transition – a discontinuity in $\langle \mathcal{L} \rangle$ signals a first order transition, whereas a continuous change implies second order.

Svetitsky and Yaffe [128] argued that for the purposes of determining the order of the finite temperature transition the relevant effective theory is a 3-dimensional spin model of short-ranged interactions (between Wilson lines reduced to spin variables) with a $Z(N)$ symmetry. Then using the universality classes known from renormalization group analyses, they predicted that the transition for $SU(2)$ should be second order (Ising class), whereas for $SU(N \geq 3)$ it should be first order (Pöotts class). Numerical results are consistent with these predictions, in particular for $SU(2)$ the critical exponents have been determined using standard techniques of Statistical Mechanics and found to agree with those of the 3-dimensional Ising model [129].

In the presence of dynamical quarks, it is easy to see that the discretized Dirac action is not invariant under the $Z(N)$ symmetry. Consequently, $\langle \mathcal{L} \rangle \neq 0$ for all values of β , so $\langle \mathcal{L} \rangle$ is no longer an order parameter. Nevertheless, changes (discontinuous or continuous) in it are used to search for the location of the finite temperature phase transition and its order.

15. Phase Transitions in the Lattice Theory

The lattice theory with a finite cut-off $\mu = \pi/a$ can be regarded as an effective theory. The integration over momenta in the range $\{\pi/a, \infty\}$ renormalizes the couplings and generates additional effective interactions (see Section 8 and the lectures by M. Lüscher for a detailed discussion of this approach). Thus, one can regard the lattice theory as a point in an infinite dimensional space of couplings, and taking the continuum limit as a flow in this space to the critical point at $a = 0$ that defines the physical theory. In this way of thinking about LQCD as a Statistical Mechanics system one is automatically lead to the questions

- What is the phase diagram in this extended coupling constant space?
- Are there other fixed points, and if so what is the nature of the theory at those points?
- What are the order parameters that characterize different phases?

The starting point of my discussion of these questions will be the pure gauge theory. As before, the generalized form of the gauge action is

$$S_g = \beta \sum_{i,j} \sum_{\mu,\nu} \sum_x c_{i,j} \text{ReTr } W_{\mu\nu}^{i,j} \quad (15.1)$$

where the sum over i in $W_{\mu\nu}^{i,j}$ is over all possible Wilson loops and j is over all representations of the loop. β is the overall coupling and the $c_{i,j}$ give the relative strengths of the different loops.

The three simplest gauge invariant probes of the phase diagram are (gauge non-invariant quantities have zero expectation values as shown by Elitzur [37])

- Wilson loops: These, as discussed in Section 14.2.1, contain information on the potential between static quarks.
- Wilson/Polyakov lines. These measure the free energy of an isolated quark.
- The chiral condensate $\langle \bar{\psi}\psi \rangle$. This provides information on vacuum alignment under spontaneous chiral symmetry breaking.

LQCD, in the limit $g \rightarrow \infty$, can be analyzed using strong coupling expansions as discussed in Section 14.2.2. Using these it was shown that, in the limit $g \rightarrow \infty$, the expectation value for Wilson loops has an area law for all gauge groups ($U(1)$ and $SU(N)$). Since electrostatics ($U(1)$) does not confine, this region of coupling space cannot correspond to the physical theory. In other words, there should exist a phase transition separating the strong coupling and weak coupling phases of non-confining theories like $U(1)$. This has been established for $U(1)$, both analytically [130] and by

Monte Carlo simulations [131]. For a theory like QCD, with both confinement and asymptotic freedom, we need to know if the two regions of parameter space are analytically connected. Analytical methods like strong and weak coupling expansions can be used provided the range of validity of the expansions gives a sufficiently large overlap, otherwise non-perturbative methods are necessary.

To address these questions for SU(3) I show, in Fig. 14, the behavior of square loops, $\{\text{Re Tr } W^{i \times i}\}$, as a function of the coupling β for the simplest action – Wilson’s plaquette action. Also shown for comparison are the strong coupling expansion to $O(\beta^5)$ [133], the weak coupling expansion to $O(g^4)$ [134] and its tadpole improved (TIPT) version. These numerical data show that there is a smooth though sharp transition between the weak coupling and strong coupling phases, *i.e.* no phase transitions. The crossover takes place in the window $5 \lesssim \beta \sim 6$, with the larger loops becoming perturbative at slightly higher β . Both strong coupling and weak coupling expansions fail in this region (for TIPT this is evident only for large loops), so numerical methods are essential. Such an analysis was the basis of the pioneering work of Creutz, and Creutz, Jacobs, and Rebbi [132], who showed that the string tension changes from its strong to weak coupling behavior without a phase transition. Their calculations yielded the first clue that it may be possible to get continuum physics with $O(20\%)$ errors from simulations at $\beta \gtrsim 6.0$.

The TIPT result shown in Fig. 14 is obtained as follows. The mean-field improved Wilson loop expectation value is defined to be

$$\langle W^{i \times i} \rangle = \langle \widetilde{W}^{i \times i} \rangle u_0^{4i}, \quad (15.2)$$

where for u_0 I use the non-perturbative value obtained from the plaquette as it is more readily available. The improved expansion for $\langle \widetilde{W}^{i \times i} \rangle$, after absorbing the tadpoles using the plaquette, is given by

$$\langle \widetilde{W}^{i \times i} \rangle = \frac{\langle W^{i \times i} \rangle}{\langle W^{1 \times 1} \rangle^i}, \quad (15.3)$$

where both terms on the right are first expanded in terms of $\tilde{g}^2 = g^2/u_0^4(\text{pert}) = g^2/\langle W^{1 \times 1} \rangle_{\text{pert}}$ to order \tilde{g}^4 , and then the ratio is truncated at order \tilde{g}^4 . It is obvious that using u_0 from the plaquette guarantees a perfect fit for $\langle W^{1 \times 1} \rangle$, however the improvement in even the $\langle W^{4 \times 4} \rangle$, is remarkable.

The first study of the phase diagram in the generalized coupling constant space was done using an action constructed from the plaquette in both the

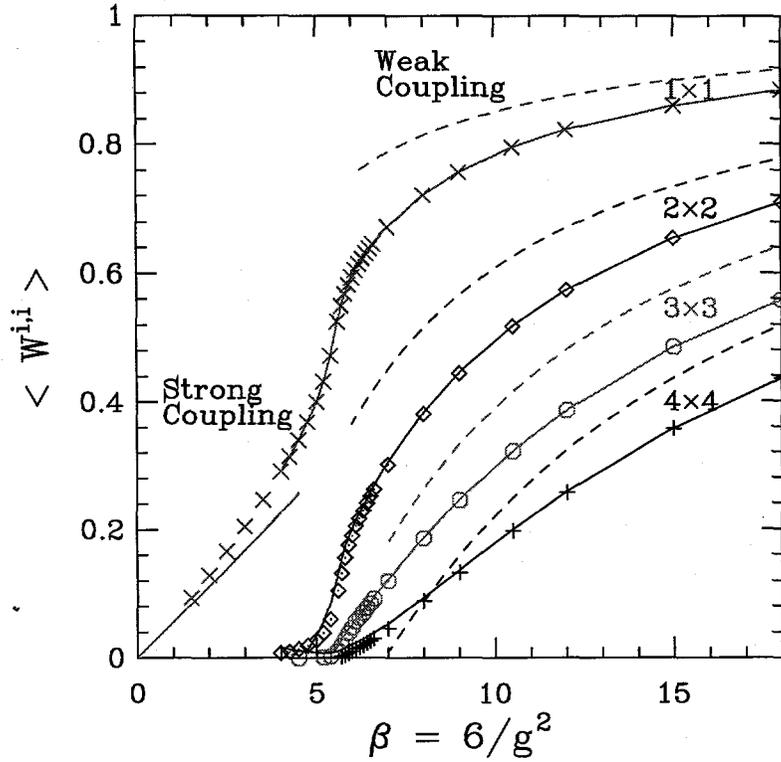


Fig. 14. The behavior of the expectation value of square Wilson loops as a function of coupling β . The data, show a sharp crossover between $5 \lesssim \beta \lesssim 6.0$. Result of the strong coupling expansion to $O(\beta^5)$ and weak coupling expansion to $O(g^4)$ are shown by the dashed line. Results of tadpole improved weak coupling expansion (TIPT) are shown by the solid lines. Since the improved coupling is defined in terms of the 1×1 loop, $g_I^2 = g^2 / \langle W^{1 \times 1} \rangle$, the agreement between the data and TIPT is a trivial result of the construction as explained in the text. The improvement in larger loops, however, demonstrates the efficacy of TIPT.

fundamental and adjoint representations of the $SU(N)$ gauge theory [135]

$$S = \beta_F \sum \frac{1}{N} \text{Re Tr } W_{1 \times 1} + \beta_A \sum \frac{1}{N^2} |\text{Tr } W_{1 \times 1}|^2. \quad (15.4)$$

For this action, the effective bare coupling, obtained using a Taylor expansion, is

$$\beta_{eff} \equiv \frac{2N}{g_{eff}^2} = \beta_F + 2\beta_A. \quad (15.5)$$

Thus simulations can be done in the negative β_A region as long as $\beta_F > 2|\beta_A|$. This is a classical condition that avoids the weak coupling singularity.

The resulting lines of first order bulk transitions, along with the location of the end point for three gauge groups, are shown in Fig. 15. The point at $\beta_A = \infty$ is the first order transition in the $Z(N)$ gauge theory, while that at $\beta_F = 0$ corresponds to that in the $SO(N)$ theory. The location of the third end-point with respect to the β_F axis depends on N . Current estimates of the location are [136–138]

$$\begin{aligned} \beta_F = 1.22 \quad \beta_A = 1.25 \quad & \text{SU}(2) \\ \beta_F = 4.00(7) \quad \beta_A = 2.06(8) \quad & \text{SU}(3) \\ \beta_F \sim 12 - 15 \quad \beta_A = (-1) - (-5) \quad & \text{SU}(4) \end{aligned} \quad (15.6)$$

The fact that for $N \leq 3$ this point lies above the β_F axis is consistent with the observation that there is no discontinuity in the behavior of Wilson loop data as shown in Fig. 14. If one were to take the continuum limit along a line crossing the transition line above this point, for example along the dotted line A in Fig. 15, then there would be discontinuities in Wilson loops (and consequently in the string tension), specific heat, and glueball masses at the point of intersection with the transition line. At the end-point X the specific heat diverges and consequently the 0^{++} glueball mass goes to zero. (Note that the specific heat is the volume integral of the 0^{++} glueball correlation function with the plaquette as the interpolating operator. Thus the divergence in the specific heat implies a zero in the 0^{++} glueball mass.) The string tension and masses of glueball states with other spin-parity quantum numbers are non-zero and continuous at this point. Such phase transitions are lattice artifacts. For example, taking the continuum limit at X , *i.e.* setting $a = 0$, would give a free theory as all states other than the 0^{++} glueball become infinitely heavy.

One can take the continuum limit for the $SU(3)$ gauge theory along any trajectory like A or W or B . In case A it is necessary to keep $\beta_F > \beta_F^*$ to avoid the lattice artifacts. Along W , while there are no singularities, the artifact X could, around $\beta_F = 5.7$, cause significant deviations from the physics of the relevant fixed point at $g = 0$. Lattice data verify this scenario – the non-perturbative β -function measured from observables other than

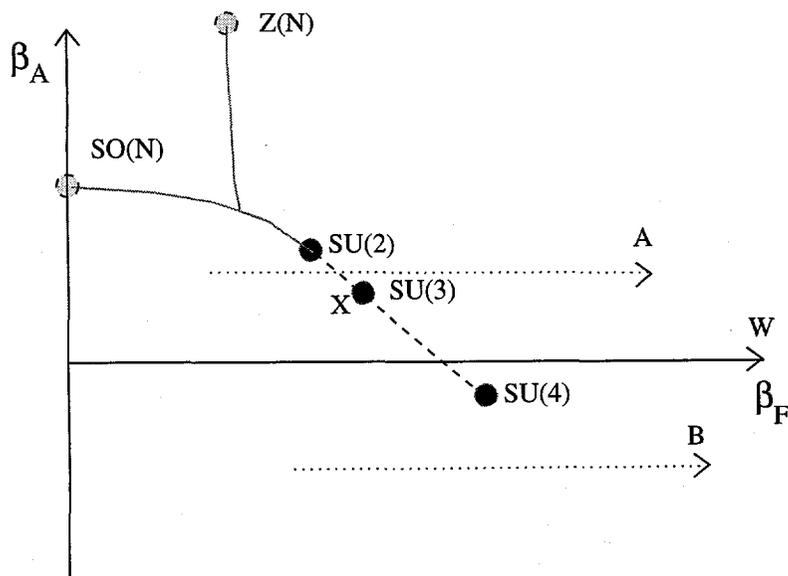


Fig. 15. The phase diagram of $SU(N)$ gauge theories in the fundamental-adjoint coupling constant space.

M_{0++} shows a large dip around $\beta_F = 5.7$ [89]. Thus, to avoid these artifacts requires simulations for $\beta_F \gtrsim 6.0$. On a trajectory like B the influence of X is expected to be less than along W , and it is possible that contact with continuum physics can be made on coarser lattices. If so then the trajectory B corresponds to an “improved action” in accordance with the third criteria enumerated in Section 8.

This qualitative argument for taking the continuum limit along a trajectory like B is supported by Monte Carlo renormalization group (MCRG) estimates of the renormalized trajectory [87]. Such calculations of the RT yield a negative value for β_A . As explained above, the most sensitive probe of the expected improvement on including a β_A term in the action, is the behavior of M_{0++} . Until recently tests of this conjecture were limited by statistical accuracy [166]. The recent calculations on anisotropic lattices of glueball spectra by Morningstar and Peardon [139] validate this scenario. I shall discuss their results in Section 18.3.

It is clear that there are additional phase transitions in the generalized lattice theory. For example one would get a similar diagram to that in Fig. 15 if the action was defined in terms of the 1×2 loop in the funda-

mental and adjoint representations. The question, therefore, is whether there exist other second order transition points at which all correlation lengths (measured in lattice units) diverge *i.e.* where a non-trivial continuum limit exists. The commonly held belief, supported by numerical data, is that all such points lie on the hypersurface defined by $g_{eff} = 0$. The long distance behavior of these theories is controlled by the fixed point defined with respect to a suitable renormalization group transformation. Taking the continuum limit of the lattice theory on this hypersurface provides a non-perturbative definition of QCD. This is the standard scenario.

An alternative scenario is the presence of a non-trivial fixed point at $g_{eff} \neq 0$ [140]. If this unlikely possibility turns out to be true then the perturbative relation between g_{eff} and a , based on asymptotic freedom, would change – the coupling would no longer run to zero as the momentum scale $Q \rightarrow \infty$. This could be exposed by precise measurements of α_s as a function of Q^2 , or by a detailed search for non-trivial second order phase transitions in the lattice theory.

In either case the non-perturbative lattice analyses of quantities such as the spectrum and weak matrix elements does not change. To get to the continuum limit we would still extrapolate the lattice data in a to $a = 0$. Only the PQCD relation between a and g would be invalid, consequently extrapolations using the perturbative scaling relation could not be used. My bottom line on this issue, having mentioned the heretical view, is to proceed assuming that the standard scenario is correct.

16. Errors in Lattice Results

Numerical simulations of lattice QCD introduce both statistical and a number of systematic errors. I have collected together a brief description of the various errors here, and in the remainder of the article I will only give quantitative estimates. I will label a particular systematic uncertainty as negligible/important if it is negligible/comparable in magnitude to the statistical error for that observable as determined in a state-of-the-art calculation.

16.1. Statistical errors

The monte carlo method for doing the functional integral employs statistical sampling. The results, therefore, have statistical errors. The current understanding, based on agreement of results from ensembles generated using different update algorithms, different initial starting configuration,

and different random number generators in the Markov process, is that the functional integral for QCD is dominated by a single region. Second, we find that configurations with non-trivial topology are properly represented in an ensemble generated using a Markov chain based on small changes to link variables. Based on these observations we conclude that the energy landscape is simple, consequently, the statistical errors should fall as $1/\sqrt{N}$, where N is the number of independent measurements. Tests like binning of the data and the calculation of auto-correlation times for various observables are used to determine the monte-carlo time between independent configurations.

Estimates of decorrelation times in dynamical simulations are just becoming available and the understanding of decorrelation times for various update algorithms and different fermion formulations is not complete. A recent study of decorrelations for update of Wilson fermions [142] using the hybrid Monte Carlo algorithm [141] (the algorithm of choice) presented very encouraging results. They found that decorrelation time based on tunneling rate between different topological sectors (expected to be amongst the slowest modes) was within a factor of two of standard probes like large Wilson loops and hadron correlators, and ≈ 80 units of molecular dynamics time for $m_\pi/M_\rho \geq 0.56$. Thus trajectory lengths of 5000-10000 units, which are feasible on today's computers, provide a reasonable statistical sample. Even though details like how this decorrelation time grows with decreasing quark mass still need to be filled in, it is clear that current Monte Carlo algorithms do provide a reliable way of doing the functional integral for both QCD and full QCD.

16.2. Finite Volume

This is the simplest error to visualize as it is associated with using finite lattices to represent an infinite system. Lüscher has shown that for sufficiently large lattices of side L , the finite volume corrections in the mass M of a given state fall off as e^{-ML} [27]. This analysis assumes that the interactions with the mirror states on a periodic lattice are small. On small lattices the wave-function of the states is squeezed, resulting in an much more rapid increase of the mass with decreasing lattice size. This effect is seen clearly for $L < 1.5$ fermi, where the finite size behavior fits a power-law, $\sim 1/L^3$ [143]. The goal, therefore, is to work on lattices large enough such that the finite size effects are exponentially suppressed.

To determine the lattice size needed for observables involving quark propagation I choose the pion as the test state as it the lightest and thus has the largest correlation length. Results of quenched simulations show

that for $M_\pi L \geq 5$ the exponential suppression applies and the corrections are negligible. For physical value of M_π this translates into $L \geq 7$ fermi. Another way of stating the same criteria is that the quantum mechanical properties of hadrons, typically of size ≤ 1 fermi, are unaltered if the box size is larger than 7 fermi. Current simulations explore quark masses in the range $m_q \geq m_s/4$ or equivalently $M_\pi/M_\rho \geq 0.4$. For these "heavier masses" the criteria $M_\pi L \geq 5$ translates into $L \geq 3$ fermi. The lattices used in the most recent calculations, for example the CP-PACS runs, discussed later, satisfy this condition.

Another consequence of finite L is momentum resolution. For example, the lowest non-zero momentum, and spacing, is $2\pi/La = 393$ MeV for $L = 32$ and $1/a = 2$ GeV. One would like a much finer resolution when investigating the dispersion relation for mesons and baryons or when calculating matrix elements required to study form-factors in semi-leptonic and rare decays. This can only be done by increasing either L or a . Increasing L is limited by computer resources, while increasing a increases discretization errors which are discussed next.

16.3. Finite lattice spacing errors

The lattice is merely a technical scaffold and the physical results are obtained by taking the $a \rightarrow 0$ limit. At finite a , lattice results have discretization errors whose size depend on the degree to which the lattice action and operators have been improved (see Sections 8, 12, and 13). As discussed in Section 8, we employ two strategies for removing these errors. First, improve the lattice action and operators so that the errors at fixed a are small, and secondly, repeat the simulations at a number of values of a and extrapolate to $a = 0$. Since the functional form used to characterize the errors is usually truncated to just the leading term in a , the extrapolation has an associated systematic error. To get a handle on this uncertainty, extrapolations have been done for three different formulations (Wilson, clover, staggered) for which the leading corrections are different ($O(a)$, $O(\alpha a) - O(a^2)$, $O(a^2)$). The difference between the three results in the $a = 0$ limit is a measure of the residual uncertainty. I shall illustrate this procedure for controlling the discretization errors in Section 20 where the calculation of quark masses is discussed.

16.4. Chiral Extrapolations in the light quark masses

The physical u and d quark masses are too light to simulate on current lattices. For $1/a = 2$ GeV, realistic simulations require $L/a \gtrsim 70$ to avoid

finite volume effects, *i.e.* keeping $M_\pi L \geq 5$. Current biggest lattice sizes are $L/a = 32$ for quenched and $L/a = 24$ for unquenched. Thus, to get results for quantities involving light quarks, one typically extrapolates in $m_u = m_d$ from the range $m_s/4 - 2m_s$ using functional forms derived using chiral perturbation theory but truncated to the first couple of terms. Since the range of the chiral extrapolation is large, it is important to quantify the size of the higher order chiral corrections. This requires precise data at a large number of values of quark masses. Most of the current fits are based on keeping just the lowest order terms as the quality of the data is not good enough to quantify the various higher order corrections, and/or to identify the quenched artifacts discussed in Section 16.8. The status of the current data with respect to resolving these higher order corrections is discussed in Section 18 where an analysis of SU(3) flavor symmetry breaking, for example the mass differences between mesons and baryons within SU(3) multiplets like the baryon octet and the decuplet, is presented.

Current simulations also neglect isospin breaking effects and electromagnetic contributions as these are a few MeV in nature, and smaller than the present numerical resolution. The iso-spin symmetric mass is defined by $\bar{m} = (m_u + m_d)/2$. For an exploratory analysis of iso-spin breaking effects in the presence of a strong U(1) field see Ref. [144].

I expect that our understanding of the chiral corrections will change significantly in the next two years due to the factor of 10-100 increase in the computational power made available to LQCD. To study iso-spin breaking effects properly requires dynamical lattices with roughly physical values for m_u and m_d . Such simulations are still a few years away.

16.5. Discretization of Heavy Quarks

Simulations of heavy quarks (c and b) have large discretization errors of $O(ma)$ in addition to the $O(\Lambda_{QCD}a)$ and $O(pa)$. This is because quark masses measured in lattice units, $m_c a$ and $m_b a$, are of order unity for $2 \text{ GeV} \leq 1/a \leq 5 \text{ GeV}$. Data show that these discretization errors are large even for m_c in the case of Wilson and clover actions. (Staggered fermions do not have any advantage over Wilson-like discretizations for heavy quarks and are hence not used to study heavy quarks.) A number of alternate approaches are being investigated to address this issue (see Section 13). These include the non-relativistic QCD (NRQCD) formulation [108], lattice versions of the heavy quark effective theory [145], a variant of the Dirac discretization that interpolates between NRQCD and clover discretizations [113], and non-perturbatively $O(a)$ improved SW-clover fermions [104]. Many collaborations are testing these formulations,

however, it is still too early to judge which approach will provide the best solution to the simulations of heavy quarks.

16.6. Matching between lattice and continuum scheme (renormalization constants)

Experimental data are analyzed using some continuum renormalization scheme like \overline{MS} so, to make contact with phenomenology, results in the lattice regularization scheme have to be converted to the continuum scheme. The perturbative relation between renormalized quantities in \overline{MS} and the lattice scheme are, in almost all cases, known only to 1-loop. Data show that the $O(\alpha_s)$ corrections can be large, $\sim 10 - 50\%$ depending on the quantity at hand. However, as explained by Parisi [77] and by Lepage and Mackenzie [78], these large corrections are mostly due to lattice artifacts which, once identified as coming from tadpole diagrams, can be removed. The Lepage-Mackenzie prescription for reorganizing the lattice perturbation theory to remove these artifacts has significantly reduced the associated uncertainty. Nevertheless, the improvement is not uniform and in many cases the corrections are still large, and the residual errors in 1-loop perturbative estimates are hard to ascertain. The final step in improving the reliability of the matching factors is to develop non-perturbative methods [146–148]. Significant progress has been made in setting up these calculations, and once they are complete the uncertainties in existing results due to using 1-loop Z 's will be removed. An illustration of the effect of using perturbative versus non-perturbative Z 's is presented in Section 20.1.

16.7. Operator mixing

As mentioned in Section 3, a major application of LQCD is to calculate the matrix elements of operators that arise in the effective weak Hamiltonian between hadronic states. In addition to the question of the overall normalization of these operators, the operators can, in general, mix with operators of the same, higher, and lower dimensions. On the lattice this mixing arises due to quantum corrections and discretization errors. The set of lattice operators one has to consider is usually larger than those in the continuum theory because at finite a the symmetries of the lattice theory are smaller than those of the continuum theory, for example the hard breaking of chiral symmetry in Wilson fermions. In cases where there is mixing with lower dimensional operators, the mixing coefficients have to be known very accurately otherwise the power divergences overwhelm the signal as $a \rightarrow 0$. Similarly, in cases where there is mixing with operators of

the same dimension but with different tensor structures, as in Wilson-like actions due to the explicit breaking of chiral symmetry, the chiral behavior may again be completely overwhelmed by these lattice artifacts if the coefficients are not known precisely. Examples where such mixing has posed serious computational challenges are the matrix elements of 4-fermion operators needed in the calculation of the $\Delta I = 1/2$ rule, B_K , B_6 , and B_8 . In these cases also non-perturbative methods for calculating the mixing coefficients are essential. For a discussion of these methods see Prof. Martinelli's lectures.

16.8. Quenched Approximation (QQCD)

This approximation is computationally the hardest to shed and quantify. Since it plays a key role in today's simulations, I will discuss it in some detail. The approximation consists of ignoring the fermion contribution to the path integral, *i.e.* setting $\det M = \text{constant}$ in Eq. 4.3. The sole reason for this approximation is limitations of computer power. With known algorithms the computational requirements of full QCD simulations go up by factors of $10^3 - 10^5$ depending on the quark mass. Since QQCD is confining, asymptotically free, and shows spontaneous chiral symmetry breaking ($\langle \bar{\psi}\psi \rangle_{QQCD} \neq 0$), and differs from full QCD only in the relative weighting of the background gauge configurations, the physics analyses are identical in most cases. It is therefore considered reasonable to do detailed quenched simulations in order to understand and control all other sources of errors while waiting for better algorithms or the computer technology needed to provide the required additional factor of $10^3 - 10^5$, *i.e.* 10-1000 teraflop capability.

Physically, the quenched approximation corresponds to turning off vacuum polarization effects of quark loops. This is illustrated in Fig. 16 for the pion correlator. One important consequence of neglecting vacuum polarization effects is the behavior of the potential between a $q\bar{q}$ pair as a function of the separation. In the full theory the string breaks by the creation of a $q\bar{q}$ pair out of the vacuum, therefore at large distances there is a screened potential between two mesons. For such screened potential large Wilson loops should show a perimeter law. In QQCD the string does not break and large Wilson loops show an area law. Thus the long distance behavior of the two theories is very different and one might be led to believe that QQCD is a bad approximation. However, because of confinement, the long distance scale that is relevant to hadronic physics has a natural cutoff of a few fermi. Thus, if one can match the potential between some fine scale (~ 0.01 fermi, below which asymptotic freedom makes the calcula-

tions perturbative) and the typical size of hadrons (~ 1 fermi) then QQCD could give reasonable estimates for the hadronic spectrum. This argument is really only applicable to heavy onia, where potential models work well. For light quarks one simply has to do the simulations to estimate the size of the distortions. In any case the community has proceeded by assuming that the two theories can be roughly matched by adjusting the quenched coupling constant at some scale like 0.5 fermi, or equivalently by adjusting the overall scale Λ_{QQCD} , and hopes that the quenching corrections are small, 10 – 20%, for many cases. If this bears out then even QQCD would yield many phenomenologically useful predictions.

In spite of the optimism, the fact remains that QQCD is not a unitary theory, so it is important to understand how it fails. I would like to illustrate its shortcomings with the following examples.

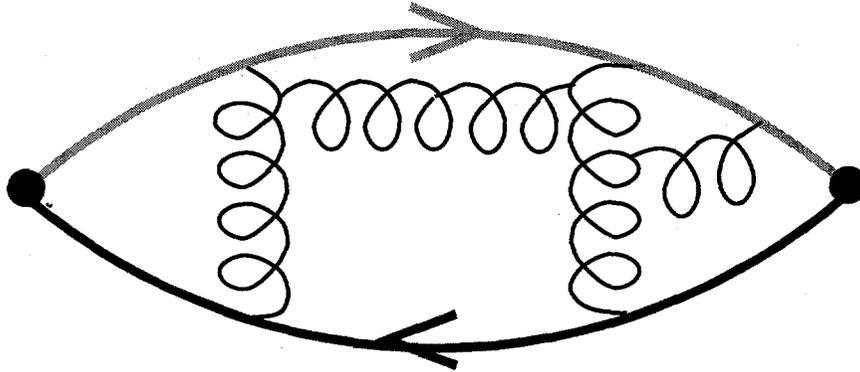
We can calculate the $\rho\pi\pi$ coupling in both QCD and QQCD by measuring the three point function $\langle T[\rho\pi\pi] \rangle$ illustrated in Fig 17A. The difference between the two values is a measure of the validity of QQCD. On the other hand the diagram shown in Fig 17B is absent in QQCD, and the ρ spectral function does not have a cut beginning at $2M_\pi$. Thus, the $\rho\pi\pi$ coupling *cannot* be extracted from an analysis of the ρ 2-point function in QQCD.

A very important difference between full QCD and QQCD is the behavior of the η' . In full QCD the singlet axial current is anomalous and the η' acquires a large mass due to the summation of vacuum polarization graphs shown in Fig 18,

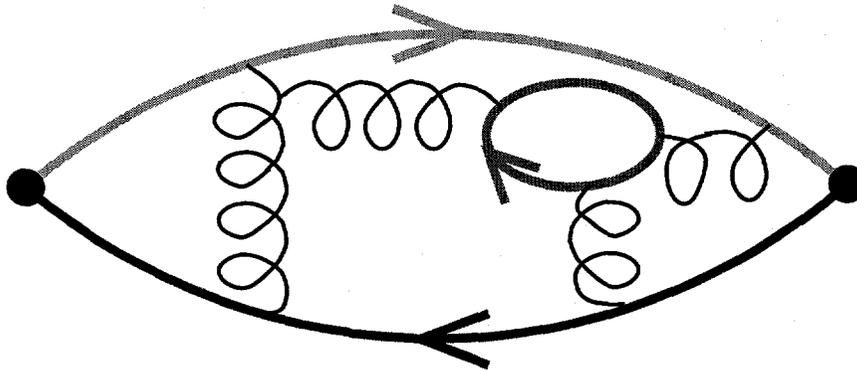
$$\frac{1}{p^2 - M^2} + \frac{1}{p^2 - M^2} M_0^2 \frac{1}{p^2 - M^2} + \dots = \frac{1}{p^2 - M^2 - M_0^2} \quad (16.1)$$

where, for simplicity I have approximated the gluonic coupling by a momentum independent constant M_0^2 . This is related to the topological susceptibility via the Witten-Veneziano relation $M_0^2 = 2n_f \chi_t / f_\pi^2 = M_{\eta'}^2 + M_\eta^2 - 2M_K^2$ [149]. In QQCD the absence of the vacuum polarization diagrams means that the η' propagator has a single and a double pole, *i.e.* $(p^2 - M^2)^{-1}$ and the “hairpin” term $(p^2 - M^2)^{-1} M_0^2 (p^2 - M^2)^{-1}$, where M^2 is the Goldstone pion mass M_π^2 . The fact that the η' , in the limit $m_q \rightarrow 0$ of QQCD, is massless has important consequences. Two groups, Sharpe and collaborators [150], and Bernard and Golterman [151], have investigated these by formulating a chiral perturbation theory for QQCD which retains η' as an additional Goldstone boson. I illustrate the differences between χ PT and $Q\chi$ PT by considering the behavior of the 2-point correlation function of the pion.

At 1-loop full QCD has the diagram shown in Fig 19A which is absent in QQCD. On the other hand, in QQCD, the hairpin term gives a contribution



(A) Quenched QCD: quark loops neglected



(B) Full QCD

Fig. 16. An illustration of the difference between quenched QCD and QCD. QQCD has no vacuum polarization loops on virtual gluon lines.

via the diagram shown in Fig 19B. Such diagrams are neglected in full QCD as they iterate to produce a massive η' . The chiral expansions for M_π and f_π in full QCD have been calculated by Gasser and Leutwyler [152],

$$M_\pi^2 = 2mB \left[1 + X_\pi - \frac{1}{3} X_{\eta_8} + \frac{16}{f_0^2} (2M_K^2 + M_\pi^2)(2L_6 - L_4) + \frac{16}{f_0^2} M_\pi^2 (2L_8 - L_5) + \dots \right]$$

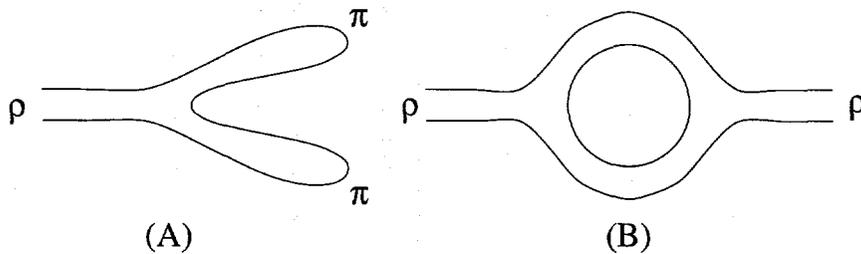


Fig. 17. The $\rho\pi\pi$ coupling from the $\rho \rightarrow \pi\pi$ decay and (B) the ρ correlator in full QCD.

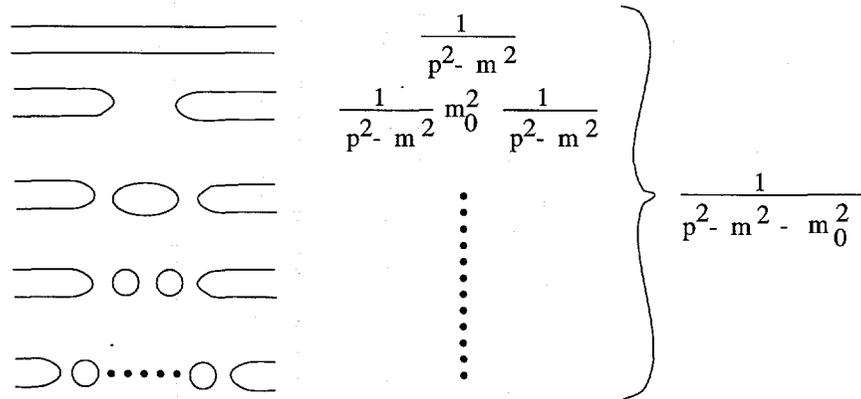


Fig. 18. The iteration of the η' propagator in full QCD. Only the first two diagrams survive in QQCD.

$$f_\pi^2 = f_0^2 \left[1 - 4X_\pi - 2X_K + \frac{16}{f_0^2} (2M_K^2 + M_\pi^2) L_4 + \frac{16}{f_0^2} M_\pi^2 L_5 + \dots \right] \quad (16.2)$$

where L_i are the $O(p^4)$ coefficients, $X_i = M_i^2 \log(\frac{M_i^2}{\Lambda_{\chi SB}^2}) / (4\pi f_0)^2$ are the chi-

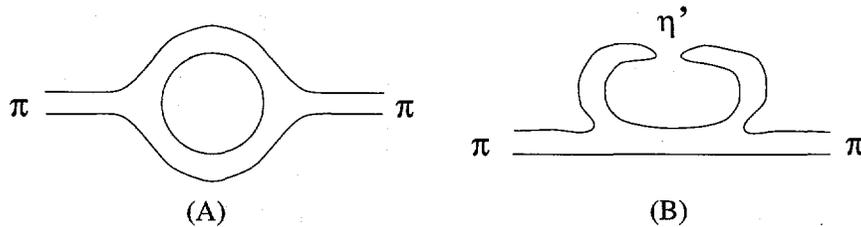


Fig. 19. The 1-loop correction to the pion propagator in (A) full QCD (B) QQCD.

ral logs, and $f_\pi = 131$ MeV. In QQCD, Sharpe and Bernard and Golterman [150,151] find that the leading terms are

$$\begin{aligned} M_\pi^2 &= 2mB_Q \left[1 - \delta \log \frac{M_\pi^2}{\Lambda_{\chi SB}^2} + \dots \right] \\ f_\pi^Q &= f_Q \left[1 - \frac{16}{f_Q^2} M_\pi^2 \tilde{L}_5 + \dots \right] \end{aligned} \quad (16.3)$$

where $\delta = M_0^2/24\pi^2 f_\pi^2$ and \tilde{L}_i are the quenched $O(p^4)$ coefficients. The term proportional to δ is a quenched chiral log and is singular in the limit $m_q = 0$.

The differences, in general, between quenched and unquenched expressions, as illustrated by the above expressions, are

- All the chiral constants are different in the two theories. For example, $B \neq B_Q$ and $f_0 \neq f_Q$.
- The normal chiral logs are missing in QQCD.
- The quenched expressions can be singular in the chiral limit due to the goldstone η' , *i.e.* through terms proportional to δ . These artifacts become dominant as $m_q \rightarrow 0$.

The first question raised by the analyses of Sharpe and Bernard and Golterman is – are the results of quenched χPT , an expansion about $m_q = 0$, meaningful if they are singular in that limit? We do not have a formal answer to this question. The approach adopted is to do the simulations and look for such chiral logs by doing fits with and without these terms. The status of this search has been reviewed in [153,154]. The data, while not conclusive, does suggest that $\delta \approx 0.15$, consistent with the value obtained by using phenomenological values for M_0 and f_π .

The second question, assuming that predictions of quenched χPT make sense, is how does one extract phenomenologically useful results for quantities that require a chiral extrapolation to \bar{m} ? One approach is to fit the quenched data using the full QCD chiral expansions in a region of quark masses where the quenched artifacts are small. Current data suggest that this region is $m_q \gtrsim m_s/4$. In this case quenching errors are, by definition, a combination of those due to the absence of quark loops and those due to fits being made at heavier quark masses. The second possibility is to fit using the QQCD expression but extrapolate keeping only the terms in the full QCD expressions. This approach has the disadvantage of increasing the number of free parameters. In either case the hope is that there

exist quantities for which the QQCD artifacts are small and the differences between QQCD and QCD coefficients are small; in such cases quenched estimates are useful phenomenological predictions. One such example is the $K^0\bar{K}^0$ mixing parameter B_K [155].

In view of these various sources of systematic error, any quantity measured on the lattice depends not only on the input parameters, *i.e.* the quark masses and the coupling α_s , but also on the lattice size L , a (discretization errors are $\propto g^m a^n$ with the powers depending on the order of improvement in the action and operators), the method used to determine the renormalization constants, and the number of dynamical quark flavors n_f . While the theoretical understanding of these errors is, in general, not complete, nevertheless, a lot is known and one can make consistency checks on the data. For example, the infinite volume continuum limit results for fixed n_f simulations should be independent of the discretization scheme, the numerical approach used, and the definition of the renormalization constants. As of this writing, results for a number of observables show stability once extrapolated to the $L \rightarrow \infty$ and $a \rightarrow 0$ limits, albeit in the quenched approximation ($n_f = 0$). We regard this consistency check as the first important step towards providing precise results. Our optimism stems from knowing that for many phenomenologically important quantities ($B_K, f_D, f_{D_s}, f_B, f_{B_s}, B_B, \alpha_s$, and the quark masses) quenched lattice QCD results are already competitive with the best non-lattice methods, and future enhancements in methodology and computer technology will systematically improve their precision.

17. Lattice Correlators

In Section 4 we discussed the reduction of the pion 2-point correlation function to a product of quark propagators and possibly gauge links, and its relation to sum over states from which M_π and f_π can be extracted. Here I enlarge the discussion to general 2-point and 3-point functions. Consider the calculation of the matrix element $\langle K^+ | V_\mu | D^0 \rangle$ which arises in the extraction of semi-leptonic form factors. The 3-point correlation function needed to calculate this ME is

$$\begin{aligned} C_\mu^{PVP}(t_x, \mathbf{p}; t_y, \mathbf{q}) &= \sum_{\mathbf{x}, \mathbf{y}} e^{-i(\mathbf{q}\cdot\mathbf{y} + \mathbf{p}\cdot\mathbf{x})} \langle T[K^+ V_\mu D^0] \rangle \\ &= \sum_{\mathbf{x}, \mathbf{y}} e^{-i(\mathbf{q}\cdot\mathbf{y} + \mathbf{p}\cdot\mathbf{x})} \langle \bar{s}(0) \gamma_5 u(0) \bar{c}(y) \gamma_\mu s(y) \bar{u}(x) \gamma_5 c(x) \rangle \end{aligned} \quad (17.1)$$

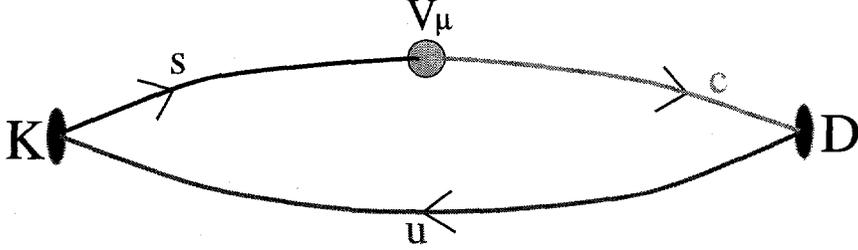


Fig. 20. Schematic of Feynman diagram for $\langle KV_\mu D \rangle$ in terms of c , s , and u quark propagators.

$$\begin{aligned}
 &= \sum_{x,y} e^{-i(q \cdot y + p \cdot x)} \langle \gamma_5 S_F^u(x, 0) \gamma_5 S_F^s(0, y) \gamma_\mu S_F^u(y, x) \rangle \\
 &\sim \frac{\exp(-E_K(\mathbf{p} + \mathbf{q})t_y - E_D(\mathbf{p})(t_x - t_y))}{(4E_K(\mathbf{p} + \mathbf{q})E_D(\mathbf{p}))} \times \\
 &\langle 0 | \bar{u} \gamma_5 s | K(\mathbf{p} + \mathbf{q}) \rangle \langle K(\mathbf{p} + \mathbf{q}) | \bar{s} \gamma_\mu c | D(\mathbf{p}) \rangle \langle D(\mathbf{p}) | \bar{c} \gamma_5 u | 0 \rangle .
 \end{aligned}$$

I have used (i) Wick contractions to express the correlation function in terms of quark propagators S_F (see Fig. 20); (ii) local bilinears for the source and sink operators and the vector current (to improve the signal one usually uses smeared operators and an improved current); and (3) the interpolating operator for the D meson and the current insertion are at definite momentum. Momentum conservation then fixes \mathbf{p}_k . The final form in Eq. 17.2 shows the behavior for large Euclidean time separations ($t_x \ll t_y \ll t_0$). Of the many terms in the right hand side, we are only interested in the matrix element $\langle K(\mathbf{p} + \mathbf{q}) | \bar{c} \gamma_\mu s | D(\mathbf{p}) \rangle$. The remaining factors can all be extracted from 2-point correlation functions. Thus, one can make individual fits to all the required correlators, or design a ratio in which the maximum number of unwanted factors cancel. In practice, to improve the numerical signal, a single fit is made to a ratio of 3-point to a combination of 2-point functions, $\langle T[KV_\mu D] \rangle / \langle T[KK] \rangle \langle T[DD] \rangle$, that gets rid of the exponential in time behavior.

To make fits to such correlation functions projected on to definite spatial momenta, we have to address three questions:

- Is the correlator even or odd in its time variables?
- Is the correlator even or odd in its momenta?
- Is the correlator real or imaginary? As mentioned before only the real part of Euclidean correlation functions has a signal. In practice we usually work in terms of bilinear operators given in Table 8 where factors of i have been dropped for brevity. In this basis of operators one needs to know

Table 7

The signs for bilinear currents under \mathcal{T} , \mathcal{P} and $CP\mathcal{H}$ transformations.

	1	γ_i	γ_4	$\gamma_i\gamma_j$	$\gamma_i\gamma_4$	$\gamma_i\gamma_5$	$\gamma_4\gamma_5$	γ_5
τ_A	+	+	-	+	-	-	+	-
π_A	+	-	+	+	-	+	-	-
χ_A	+	-	+	+	-	+	-	-

whether a given n-point function has a signal in the real or imaginary part. To answer these questions we use the discrete symmetries, \mathcal{T} , \mathcal{P} , and $CP\mathcal{H}$ discussed in Table 2 and in Eq. 10.11. For each element of the Dirac algebra, Γ_A define three signs, τ_A , π_A , and χ_A , associated with \mathcal{T} , \mathcal{P} , and $CP\mathcal{H}$:

$$\begin{aligned}
\gamma_4\gamma_5\Gamma_A\gamma_5\gamma_4 &= \tau_A\Gamma_A; \\
\gamma_4\Gamma_A\gamma_4 &= \pi_A\Gamma_A; \\
C\gamma_4\gamma_5\Gamma_A\gamma_5\gamma_4C^{-1} &= \chi_A\Gamma_A^*.
\end{aligned} \tag{17.2}$$

These signs are given in Table 7 and the application to 2-point and 3-point correlators is illustrated in the next two sub-sections.

17.1. Two-point meson correlators

Consider the general two-point correlator

$$\begin{aligned}
C^{BA}(\mathbf{p}, t) &= \sum_{\mathbf{x}} e^{-i\mathbf{p}\cdot\mathbf{x}} \langle \bar{\psi}_2(\mathbf{x}) \Gamma_B \psi_1(\mathbf{x}) \bar{\psi}_1(0) \Gamma_A \psi_2(0) \rangle, \\
&= - \sum_{\mathbf{x}} e^{-i\mathbf{p}\cdot\mathbf{x}} \langle \text{Tr} (S_2(0, \mathbf{x}) \Gamma_B S_1(\mathbf{x}, 0) \Gamma_A) \rangle,
\end{aligned} \tag{17.3}$$

where Tr is trace over spin and color indices and $t \equiv x_4$. \mathcal{T} , \mathcal{P} and $CP\mathcal{H}$ then imply the following:

- If $\tau_A\tau_B = \pm 1$, $C^{BA}(\mathbf{p}, t)$ is even (odd) in t .
- If $\pi_A\pi_B = \pm 1$, $C^{BA}(\mathbf{p}, t)$ is even (odd) in \mathbf{p} .
- If $\chi_A\chi_B = \pm 1$, $C^{BA}(\mathbf{p}, t)$ is real (imaginary).

Note that in Eq. 17.3 the second quark propagator is from $\mathbf{x} \rightarrow 0$. Using the hermiticity property $S(\mathbf{x}, 0) = \gamma_5 S(0, \mathbf{x})^\dagger \gamma_5$ we can write this in terms of $S(\mathbf{x}, 0)$. This simple property, the relation between the quark propagator from a given point and the anti-quark propagator to that point, leads to a huge computational saving. For example, if the momentum projection is done by summing over the L^3 sink points \mathbf{x} with weight $\exp(i\mathbf{p}\cdot\mathbf{x})$, then, because of the hermiticity property we need only one propagator inversion, instead of $L^3 + 1$.

Table 8

The local interpolating field operators for mesons and baryons in Wilson-like theories. Projection to zero momentum states is obtained by summing over the points \mathbf{x} on a time slice. The C-parity is only relevant for flavor degenerate meson states. The tr in baryon operators denote spin trace. $C = \gamma_2 \gamma_4$ and the symmetry properties of flavor indices for nucleons are discussed in the text. The decuplet baryon operator is completely symmetric in flavor index

State	$I^G(J^{PC})$	Operator
Scalar (σ)	$1^-(0^{++})$	$\bar{u}(x)d(x)$
	$1^-(0^{++})$	$\bar{u}(x)\gamma_4 d(x)$
Pseudoscalar	$1^-(0^{-+})$	$\bar{u}(x)\gamma_5 d(x)$
	$1^-(0^{-+})$	$\bar{u}(x)\gamma_4\gamma_5 d(x)$
Vector	$1^+(1^{--})$	$\bar{u}(x)\gamma_i d(x)$
	$1^+(1^{--})$	$\bar{u}(x)\gamma_i\gamma_4 d(x)$
Axial (a_1)	$1^-(1^{++})$	$\bar{u}(x)\gamma_i\gamma_5 d(x)$
Tensor (b_1)	$1^+(1^{+-})$	$\bar{u}(x)\gamma_i\gamma_j d(x)$
Nucleon octet	$\frac{1}{2}(\frac{1}{2}^-)$	$(u_a^T C d_b)\gamma_5 s_c \epsilon^{abc}$
	$\frac{1}{2}(\frac{1}{2}^-)$	$(u_a^T C \gamma_5 d_b) s_c \epsilon^{abc}$
	$\frac{3}{2}(\frac{3}{2}^+)$	$(u_a^T C \gamma_i d_b) s_c \epsilon^{abc}$

17.2. Interpolating operators

For Wilson-like fermions the interpolating field operators for mesons and baryons at zero-momentum are given in Table 8. The three flavors are labeled u, d, s for brevity. The baryon operators need further clarification as discussed in [156] and reproduced below.

The spin-1/2 baryons are created by the interpolating operators

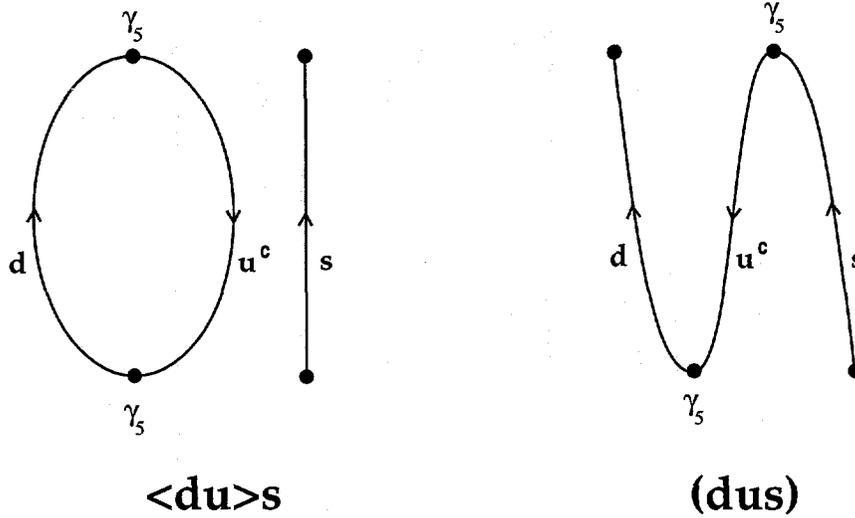
$$\mathcal{O}_{(ij)k} = (\psi_{a,i}^T C \gamma_5 \psi_{b,j}) \psi_{c,k} \epsilon^{abc}, \quad (17.4)$$

where a, b and c label color, while i, j and k label flavor. It is simple to show that $\mathcal{O}_{(ij)k} = -\mathcal{O}_{(ji)k}$, so that there are only nine independent operators—eight $SU(3)$ octets and the singlet $\sum_{ijk} \epsilon^{ijk} \mathcal{O}_{ijk}$. One way to project against the singlet is to form

$$\mathcal{B}_{ijk} = \mathcal{O}_{(ij)k} + \mathcal{O}_{(ik)j} = \mathcal{B}_{ikj}. \quad (17.5)$$

There are eight independent \mathcal{B}_{ijk} 's, the relation to the usual states being exemplified by

$$\sqrt{2}p = -\mathcal{B}_{duu} = 2\mathcal{B}_{uud} = 2\mathcal{B}_{udu},$$



The two different types of contractions for the baryon states.

$$\begin{aligned}
 \sqrt{2}\Sigma^+ &= -\mathcal{B}_{suu} = 2\mathcal{B}_{uus} = 2\mathcal{B}_{usu}, \\
 \Sigma^0 &= \mathcal{B}_{uds} + \mathcal{B}_{dsu} = -\mathcal{B}_{sud}, \\
 \sqrt{3}\Lambda^0 &= \mathcal{B}_{uds} - \mathcal{B}_{dsu}.
 \end{aligned}
 \tag{17.6}$$

The overall factor in these equations is arbitrary, while the relative normalization is fixed by $SU(3)$ symmetry.

All spin-1/2 baryon correlators are built out of the two contractions shown in Fig. 17.2. The notation $\langle DU \rangle S = \langle UD \rangle S$ corresponds to quarks of flavors U and D contracted into a closed loop, while the propagator for S carries the spin quantum numbers of the baryon. The notation (DUS) corresponds to a single ordered contraction of the three quarks. We consider two types of correlator, “ Σ -like” and “ Λ -like”. The former is exemplified by that of the Σ^0

$$\begin{aligned}
 S\{UD\} &= S\{DU\} \equiv \langle \mathcal{B}_{sud}(x) \overline{\mathcal{B}_{sud}(0)} \rangle \\
 &= \langle US \rangle D + \langle DS \rangle U + (USD) + (DSU).
 \end{aligned}
 \tag{17.7}$$

This equation defines the sign conventions for the contractions. The proton, neutron, Σ^+ , Σ^- , Ξ^0 and Ξ^- correlators are also of this type: they are, respectively, $D\{UU\}$, $U\{DD\}$, $S\{UU\}$, $S\{DD\}$, $U\{SS\}$ and $D\{SS\}$. The

second type of correlator is that of the Λ^0

$$S[UD] = S[DU] \equiv \frac{1}{3} \left[\langle US \rangle D + \langle DS \rangle U + 4 \langle UD \rangle S \right. \\ \left. - (\langle USD \rangle) - (\langle DSU \rangle) + 2 \langle SUD \rangle \right. \\ \left. + 2 \langle SDU \rangle + 2 \langle UDS \rangle + 2 \langle DUS \rangle \right]. \quad (17.8)$$

When $m_u \neq m_d$, there is also a non-vanishing $\Lambda^0 - \Sigma^0$ cross correlator, for which no useful results have been found [156].

Correlators of the form $A[AB]$ and $A\{AB\}$ are not independent—they are related by

$$A[AB] = \frac{1}{4} (3B[AA] + B\{AA\}), \\ A\{AB\} = \frac{1}{4} (B[AA] + 3B\{AA\}). \quad (17.9)$$

One can think of the results for the $A\{BB\}$ and $A[BB]$ masses as being those for the Σ and Λ , respectively, with $m_s = m_A$ and $m_u = m_d = m_B$. Unlike in the real world, there is nothing to stop $m_A = m_B$. Note, however, that in this case the Σ and Λ are also degenerate, i.e. $M(A\{AA\}) = M(A[AA])$. Indeed, the contractions in the two cases are identical.

The interpretation of the results for the completely non-degenerate correlators, $A[BC]$ and $A\{BC\}$, is more complicated. Because isospin is broken, the Σ^0 - and Λ -like states mix, with both correlators containing contributions from both physical states. Let M_+ and M_- be the masses of the heavier and lighter states, respectively, and δM the mass difference. At long times, the effective mass for both correlators will asymptote to M_- . However, at times short compared to the inverse mass difference, i.e. $\delta M t_{\max} \ll 1$, there will be an approximate plateau at a value which is a weighted average of the two masses. The surprising result, derived at lowest order in the chiral expansion in [156], is that the masses extracted from these short time plateau are insensitive to the isospin breaking and can be interpreted as those for $A[DD]$ and $A\{DD\}$ where $D = (B + C)/2$ is the mean mass for the two quarks.

17.3. Three-point correlators

A general three-point correlation function of meson operators and bilinear currents has the form

$$C^{CBA}(q, t_y; p, t_x)$$

$$\begin{aligned}
&= \sum_{x,y} e^{-i(q \cdot y + p \cdot x)} \langle \bar{\psi}_3(y) \Gamma_C \psi_2(y) \bar{\psi}_2(x) \Gamma_B \psi_1(x) \bar{\psi}_1(0) \Gamma_A \psi_3(0) \rangle, \\
&= - \sum_{x,y} e^{-i(q \cdot y + p \cdot x)} \langle \text{Tr} (S_3(0,y) \Gamma_C S_2(y,x) \Gamma_B S_1(x,0) \Gamma_A) \rangle, \quad (17.10)
\end{aligned}$$

where $t_x \equiv x_4$ and $t_y \equiv y_4$. \mathcal{T} , \mathcal{P} and \mathcal{CPH} imply

– If $\tau_A \tau_B \tau_C = \pm 1$, $C^{CBA}(\mathbf{q}, t_y; \mathbf{p}, t_x)$ is even (odd) under the simultaneous inversion of t_x and t_y .

– If $\pi_A \pi_B \pi_C = \pm 1$, $C^{CBA}(\mathbf{q}, t_y; \mathbf{p}, t_x)$ is even (odd) under the simultaneous inversion of \mathbf{p} and \mathbf{q} .

– If $\chi_A \chi_B \chi_C = \pm 1$, $C^{CBA}(\mathbf{q}, t_y; \mathbf{p}, t_x)$ is real (imaginary).

17.4. \mathcal{CPS} Symmetry

The theory has an additional discrete symmetry, \mathcal{CPS} , where \mathcal{S} is the symmetry between switching s and d quarks [157]. This symmetry holds both on the lattice and in the continuum for $m_d = m_s$. For $m_d \neq m_s$ the symmetry is softly broken, *i.e.*, violations are proportional to $m_s - m_d$. This symmetry plays a very useful role in the calculation of weak matrix elements for it (i) restricts the set of possible Wick contractions, and (ii) restricts the set of operators that mix with each other.

For a simple illustration consider $(\bar{s} \gamma_\mu L u)(\bar{u} \gamma_\mu L d)$, a local 4-quark operator which is even under \mathcal{CPS}

$$\begin{aligned}
(\bar{s} \gamma_\mu L u)(\bar{u} \gamma_\mu L d) &\xrightarrow{\mathcal{P}} (\bar{s} \gamma_4 \gamma_\mu L \gamma_4 u)(\bar{u} \gamma_4 \gamma_\mu L \gamma_4 d) \\
&\xrightarrow{\mathcal{C}} (s^T C^{-1} \gamma_4 \gamma_\mu L \gamma_4 C \bar{u}^T)(u C^{-1} \gamma_4 \gamma_\mu L \gamma_4 C \bar{d}^T) \\
&= (s^T \gamma_4 \gamma_\mu^T L \gamma_4 \bar{u}^T)(u^T \gamma_4 \gamma_\mu^T L \gamma_4 \bar{d}^T) \\
&= (s^T \gamma_\mu^T R^T \bar{u}^T)(u \gamma_\mu^T R^T \bar{d}^T) \\
&= (\bar{u} R \gamma_\mu s)(\bar{d} R \gamma_\mu u) \\
&= (\bar{u} \gamma_\mu L s)(\bar{d} \gamma_\mu L u) \\
&\xrightarrow{\mathcal{S}} (\bar{u} \gamma_\mu L d)(\bar{s} \gamma_\mu L u) \quad (17.11)
\end{aligned}$$

The brackets denote trace over spin and color and the final reordering of the two traced terms is allowed as time-ordering is implicit when calculating expectation values.

This symmetry has a further generalization as discussed by Bernard and Soni in [2,6]. For 4-fermion operators of the form $(\bar{\psi}_1 \Gamma_A \psi_2)(\bar{\psi}_3 \Gamma_B \psi_4)$, there are two general switchings: (i) \mathcal{CPS}' corresponding to $1 \leftrightarrow 2$ and $3 \leftrightarrow 4$,

and (ii) \mathcal{CPS}'' corresponding to $1 \leftrightarrow 4$ and $2 \leftrightarrow 3$. Bernard and Soni show in [2,6] how these switching symmetries are used to restrict (i) the possible contractions in the decay $K \rightarrow \pi\pi$ via the operator $(\bar{s}\gamma_\mu Lu)(\bar{u}\gamma_\mu Ld)$, and (ii) the mixing of $\mathcal{O}_\pm \equiv (\bar{s}\gamma_\mu Ld)(\bar{u}\gamma_\mu Lu) \pm (\bar{s}\gamma_\mu Lu)(\bar{u}\gamma_\mu Ld) - (u \rightarrow c)$ with other operators. I leave the details of these calculations as an exercise for the reader.

18. Lattice Calculations of the Hadron Spectrum

Successful calculations of the spectrum will test QCD. The basic steps in the lattice analysis are discussed in Sections 4 and 17. For a given set of input parameters, the mass of a given state is determined from the rate of exponential fall-off of the connected 2-point correlation function

$$\Gamma(\tau) \equiv \langle O_f(\tau) O_i(0) \rangle - \langle O_f \rangle \langle O_i \rangle = \sum_n \frac{\langle 0 | O_f | n \rangle \langle n | O_i | 0 \rangle}{2M_n} e^{-M_n \tau}. \quad (18.1)$$

We shall assume that the sum over n is restricted to the desired state by (i) optimizing O_i and O_j to get a large overlap with the wave function, *i.e.* make $\langle \text{desired state} | O | 0 \rangle$ large and the rest small, (ii) increasing the statistics so that the signal extends to large enough τ at which any remaining contamination from higher states is negligible; and (3) making a zero-momentum projection on either O_i or O_f . I shall further assume that the spectrum of possible states is discrete and has a mass-gap. Then, in order to extract M it is useful to define an effective mass

$$M_{eff}(\tau) = \log \left(\frac{\Gamma(\tau-1)}{\Gamma(\tau)} \right) \quad (18.2)$$

which, in the limit $\tau \rightarrow \infty$, converges to the desired value. The anatomy of the behavior of $M_{eff}(\tau)$ for two different states (pion and nucleon) and for two different choices of O_i and O_f are shown in Figs. 21 and 22. The points to note are:

- The convergence to the asymptotic value M can be from above or below depending on the choice of O . Only for $O_f = O_i^\dagger$ is the correlation function positive definite and the convergence is monotonic and from above.
- For $O_f \neq O_i^\dagger$, the convergence depends on subtle cancellations between contributions of various states. The relative signs are given by the product of the two amplitudes $\langle n | O_f | 0 \rangle \langle n | O_i | 0 \rangle$. Since this sign for state $|n\rangle$ can be plus or minus, the convergence is neither guaranteed to be monotonic nor from above.

- The contribution of excited states is most obvious at small τ , where M_{eff} varies rapidly before flattening out. For τ in this flat (called plateau) region one assumes that, within the numerical precision, only the lowest state contribution is significant.
- The location in τ of the onset and the length of the plateau region depends on the interpolating operators.
- For a finite statistical sample the plateau region is not flat but the data usually exhibit correlated fluctuations [158]. Therefore, unless the plateau is long enough that the fit can average over a few cycles, the extracted M_{eff} can be biased by these statistical fluctuations.
- The statistical errors grow with τ , except for the case of the pion. The reason for this was elucidated by Lepage in [159] and a synopsis of his argument is as follows. For a state with N valence quark lines, where $N = 2(3)$ for mesons (baryons), the errors are controlled by the square of the correlator which has $2N$ lines. The state of lowest energy in the squared correlator is, therefore, two pions for mesons and three pions for baryons. Thus, while the signal decreases as $\exp(-M\tau)$, the noise only decreases as $\exp(-M_\pi\tau \times N/2)$. Consequently, all states for which $M > NM_\pi/2$ the errors will increase with τ . The only exception to this condition is the pion for which the signal and noise have the same τ dependence.
- In cases where there is no clear plateau (the signal does not extend far enough in τ), one can make a fit to a larger range of τ by including the contribution of higher (radial) states in the sum in Eq. 18.1. This is illustrated in Fig. 23 for a two state fit which involves four free parameters, two amplitudes and two masses. The amplitude and mass of the excited state are less well constrained as they are trying to mimic the effect of a number of states, consequently such fits may show significant dependence on the range of τ selected. One can refine the excited state parameters by iteratively improving single state fits to $\Gamma(\tau) - \Gamma_0(\tau)$ at short τ and to the ground state $\Gamma_0(\tau)$ at large τ until the whole range is covered. A better approach to gain confidence in ground state parameters is to make simultaneous fits to a matrix of correlation functions created with different source and sink operators with the same quantum numbers. For a matrix with m (n) distinct sources (sinks) the simultaneous fit is more constraining as it involves $m + n$ amplitudes but a single common mass.

To summarize, we mostly extract M from fits to an intermediate range of τ , the "plateau" region, where M_{eff} is essentially independent of τ . At small τ the influence of excited states is large, while at large τ the statistical noise overwhelms the signal. Accuracy of the result depends on the presence of a plateau extending over a large range of τ that averages a few cycles of correlated fluctuations. The data in Figs. 21, 22, and 23 show

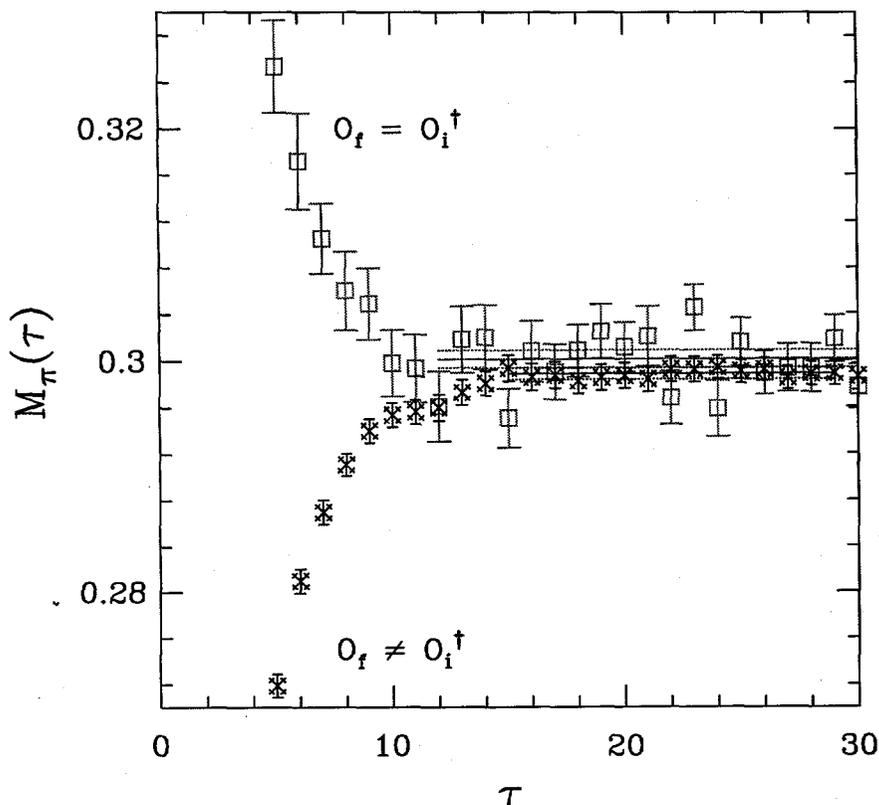


Fig. 21. Effective mass plots for M_π versus τ . The horizontal lines denote the estimate of the asymptotic value of M and the range of the plateau used in the fits, while the dotted lines show the error. The squares denote data for smeared-smeared interpolating operators ($O_f = O_i^\dagger$), while fancy crosses show data with wall-local operators.

that, with 170 lattices of size $32^3 \times 64$ at $1/a = 2.3$ GeV, this is true for the pion but marginal for other states like the vector mesons and baryons. Our present estimate is that roughly 1000 independent lattices are necessary to get a reliable signal that leads to 0.1 – 0.4% accuracy in baryon masses with quark masses in the range $m_s - m_s/4$ respectively.

Assuming that the above steps have been carried out and we have reliable estimates for hadron masses as a function of quark masses, a , L , and the number n_f of dynamical flavors, we are finally in a position to compare them to experimental data.

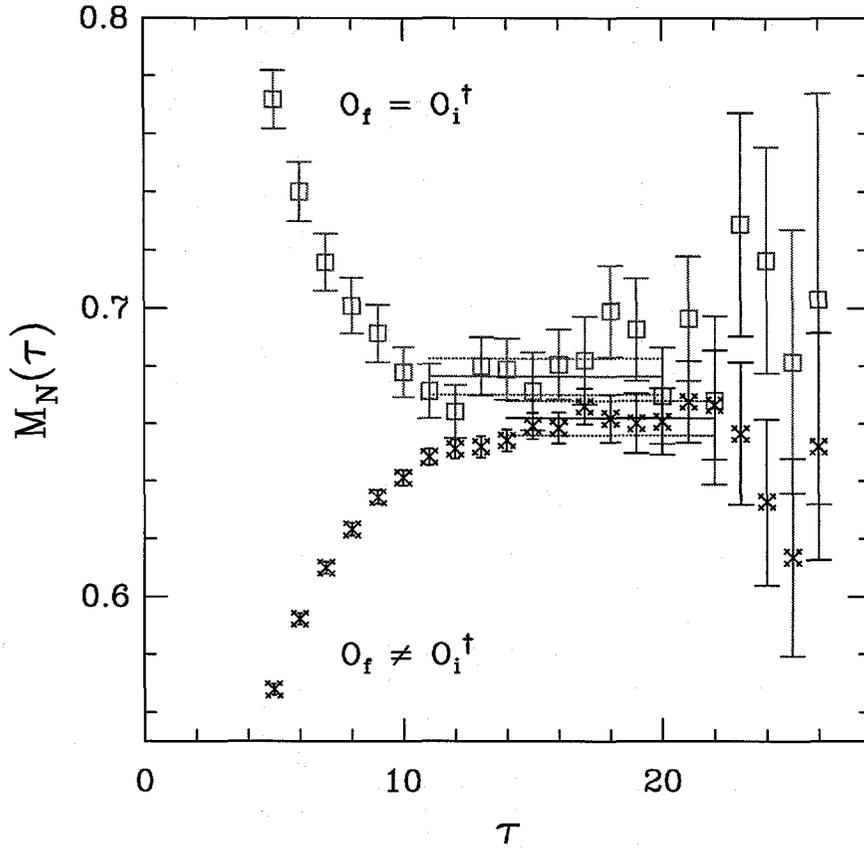


Fig. 22. Effective mass plots for M_N versus τ . The rest is same as in Fig. 21.

18.1. Status of the Light Meson and Baryon Spectrum

The first test of QCD is to reproduce the masses of pseudoscalar mesons ($\mathcal{PS} = \{\pi, K, \eta\}$), vector mesons ($\mathcal{V} = \{\rho, K^*, \omega, \phi\}$), octet baryons ($\mathcal{B}_8 = \{N, \Sigma, \Xi, \Lambda\}$), and decuplet baryons ($\mathcal{B}_{10} = \{\Delta, \Sigma^*, \Xi^*, \Omega\}$) from simulations with four input parameters m_u, m_d, m_s , and α_s . I will assume that the lattices are large enough that finite size effects can be neglected. The remaining important issues are

- Simulations have mainly been done in the quenched approximation. Rather than testing QCD, these calculations quantify the accuracy of QQCD.

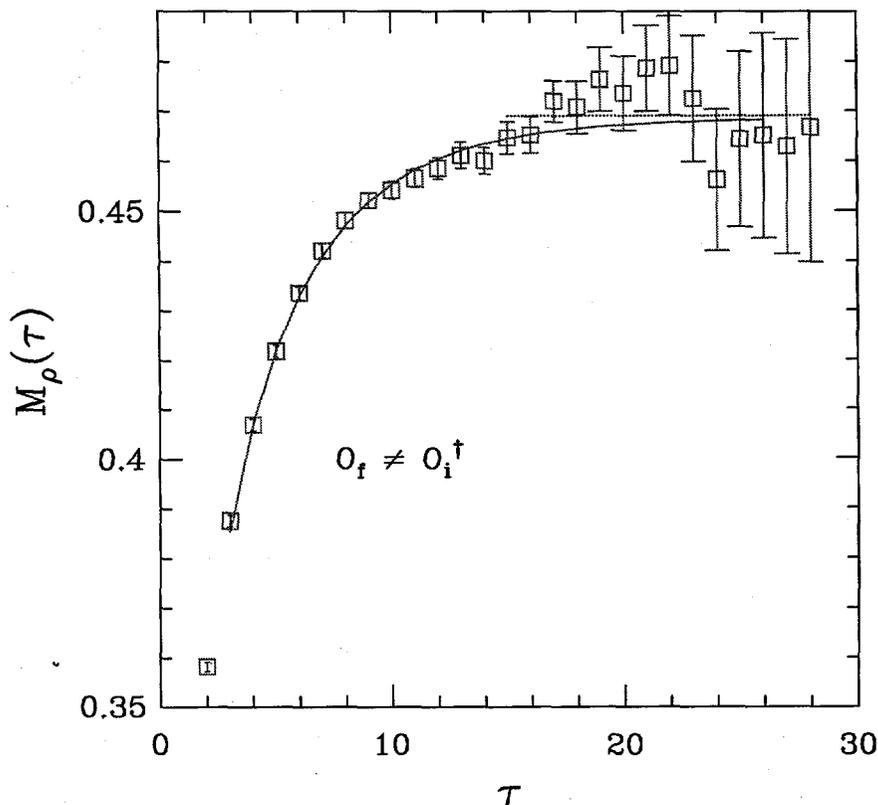


Fig. 23. Effective mass plots for M_ρ versus τ for the wall-local source. Since there is no clear plateau, a fit including two states is made to $3 \leq \tau \leq 26$. From such fits the masses of ρ and ρ^* are obtained. The mass of the ρ from the 2-state fit is consistent with that from a single state fit to data at $\tau > 15$ (blue dotted line).

- Isospin breaking effects and electromagnetic contributions are neglected.
- Simulations have been done with unphysical values of quark masses, typically in the range $[2m_s - m_s/4]$. To get physical results the data are extrapolated to $\bar{m} = (m_u + m_d)/2$ using functional forms predicted by (quenched) chiral perturbation theory.

The first step in the analyses is to extrapolate the data to physical values of \bar{m} and m_s . The simplest *Ansätze* are the lowest order chiral expressions

$$M_{\rho S}^2 a^2 = B_{\rho S} a^2 (m_1 + m_2)/2 + \dots,$$

$$M_\gamma a = A_\gamma + B_\gamma a (m_1 + m_2)/2 + \dots,$$

$$\begin{aligned}
M_{\Sigma}a &= A_N + 4F\bar{m}a + 2(F-D)m_s a + \dots, \\
M_{\Lambda}a &= A_N + 4\left(F - \frac{2D}{3}\right)\bar{m}a + 2\left(F + \frac{D}{3}\right)m_s a + \dots, \\
M_{10}a &= A_{10} + B_{10}a(m_1 + m_2 + m_3)/3 + \dots.
\end{aligned} \tag{18.3}$$

From these fits we determine the constants A_V, A_N, A_{10} and $B_{PS}, B_V, F, D, B_{10}$. Then using a suitable set, for example A_V, B_{PS}, B_V or A_{10}, B_{PS}, B_{10} , we fix the three input parameters $1/a, \bar{m}, m_s$. Unless otherwise stated the following discussion will assume that \bar{m}, a , and m_s have been fixed using M_{π}, M_{ρ} , and M_{ϕ} (or equivalently M_{K^*}) respectively. The question then is – how well do the masses of the remaining states compare with experimental numbers?

This simple analysis to test QCD (or to quantify the quenched approximation) can fall short in two ways. First, the data could show significant deviations from the linear relations Eq. 18.3. In that case the analysis becomes much more elaborate due to the increase in the number of free parameters; the possible leading non-linear terms include $O(m^2)$ corrections and chiral logarithms. To disentangle these requires very precise data at a number of values of m_q in the range $m_s/4 - 3m_s$, however, such a possibility of physics richer than that predicted by the “naive” quark model is exciting. The second possible failing is that different sets of states used to fix a, \bar{m}, m_s give different results. There could be three reasons for this failure: (i) The coefficients extracted from simulations at the significantly heavier values of m_q change as the quark masses are decreased; (ii) the differences are artifacts of discretization errors and vanish in the limit $a = 0$; (3) quenching errors. The current data show evidence for non-linearities in fits in m_q [156,160]. Also, the a, \bar{m}, m_s extracted from different sets are significantly different, and all three causes listed above contribute. As a result the focus of current analyses has been to first disentangle and quantify the contributions of the various artifacts. I will briefly illustrate these effects.

Let me start with evidence for non-linear terms in the chiral fits. Fig 24 is an example of the behavior of the vector meson mass as a function of the quark mass [156]. The plot makes it clear that in order to quantify the deviations one needs very precise data over a large range of m_q . With a statistical accuracy of $\sim 0.5\%$, a linear fit works for $0.02 \leq \bar{m}a \leq 0.04$ corresponding to $50 \leq \bar{m} \leq 100$ MeV, but not for the full range. The leading chiral corrections are proportional to $\bar{m}^{3/2}$ from chiral loops, and \bar{m}^2 . Including either term allows fits over the full range, and the data are not precise enough to distinguish between them. The effect of these non-linearities seems to be small, for example the change in M_{ρ} , after

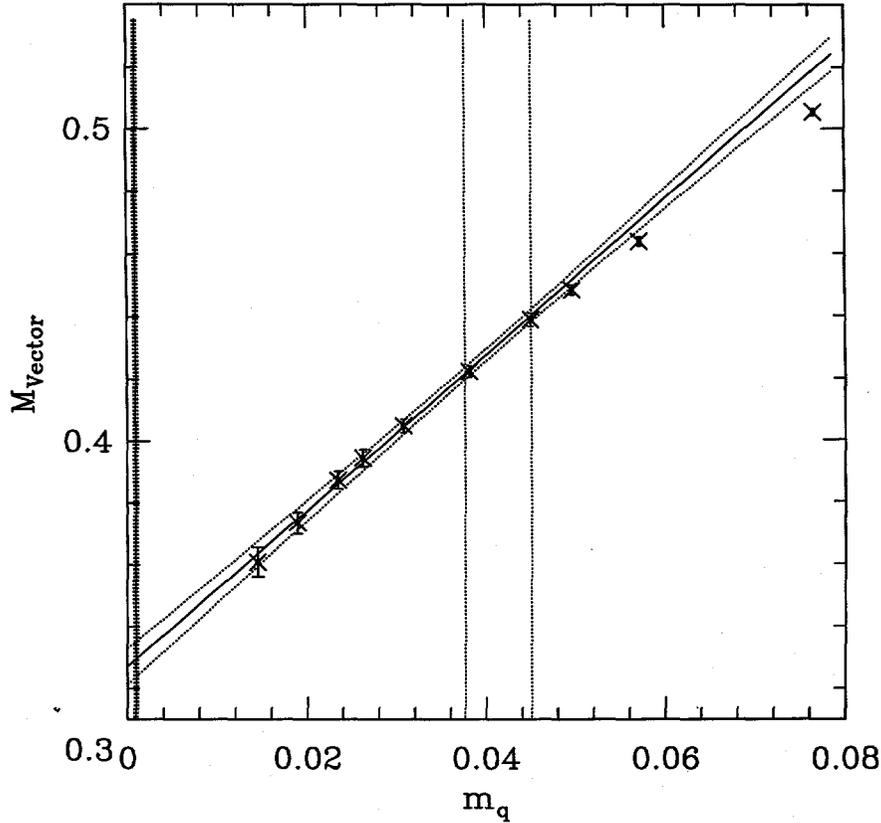


Fig. 24. Plot of M_{vector} versus the quark mass m . The data, reproduced from [156], were obtained on $32^3 \times 64$ quenched lattices at $\beta \equiv 6/g^2 = 6$ corresponding to $1/a = 2.33(4)$ GeV. The linear fit is to the lightest six points. The vertical lines show \bar{m} and the range of estimates of m_s .

extrapolation to \bar{m} , is $\lesssim 1\%$ on including either of the higher order terms in the fit.

The first test of the quenched approximation is whether it can reproduce the hyperfine splittings in the meson sector. Figure 25 shows some current data. The only input quantity is the scale $1/a$ which is used to convert lattice masses to physical units. If M_π and M_ρ are used to set \bar{m} and $1/a$ respectively, then the curves, by construction, have to pass through the experimental point at $M_{\rho S} = M_\pi$. Different choices for the scale setting quantity will change $1/a$ and shift the curves as a whole. On the other

hand, the slope at any given point is independent of the lattice scale, and is therefore the more robust lattice prediction. For this reason the quantity $J \equiv M_V \partial M_V / \partial M_{PS}^2$ is commonly used as a test of the hyperfine splittings. Lattice data (for a review see [161]) gives values for J that are $\sim 20\%$ too small. Another manifestation of this discrepancy will appear in Section 20 where I show that the estimates of m_s fixed using M_K versus M_{K^*} (or M_ϕ) differ by $\sim 15\%$.

The data shown in Fig. 25 allow two comparisons. First, the octagons (blue points) and crosses (red) are LANL data from the same 170 quenched $32^3 \times 64$ lattices at $\beta = 6.0$ with the Wilson and tadpole improved clover Dirac actions [156,162]. One does not find any significant improvement with clover fermions. Second, the fancy crosses, pluses, diamonds, and squares (green points) are Wilson fermion data from CP-PACS collaboration at $\beta = 5.9, 6.1, 6.25, 6.47$ [161]. The data at the four β values are consistent. Both comparisons imply that the smaller splittings are not due to discretization errors.

The data in Fig. 25 also show a difference between the CP-PACS and LANL Wilson fermion results at quark masses heavier than m_s . This is surprising since in this regime both collaborations have reliable measurements of the vector and pseudoscalar masses. The culprit is the lattice scale. It turns out that the estimate of $1/a$ from LANL data is 3% larger compared to the value obtained by an interpolation of the CP-PACS data. Adjusting the LANL scale shifts their data as a whole and makes all the results overlap.

The conclusion, from this and other quenched data, is that the splittings come out too small. Even though the size of discretization errors is, in my opinion, not fully resolved, the deviations are thought to be mostly due to the use of the quenched approximation.

The most extensive discussion of the mass-splittings in the baryon octet and decuplet using the Wilson action is given in [156]. Note that since this calculation was done at only one $1/a \approx 2.3$ GeV, therefore discretization errors are not resolved. There are two important observations. (i) The mass splittings in the baryon octet show significant non-linearities, for example the dependence of $M_\Sigma - M_\Lambda$ on the quark mass is shown in Fig. 26. The non-linear terms significantly increase the splittings, and extrapolations to \bar{m} give results roughly consistent with experimental values. (ii) The decuplet baryon masses show a behavior linear in $m_1 + m_2 + m_3$, and the splittings come out too small for either choice $m_s(M_K)$ or $m_s(M_\phi)$.

The state-of-the-art Wilson fermion data have been obtained by the CP-PACS collaboration. Data have been obtained at four values of the lattice spacing so a reliable extrapolation to $a = 0$ can be made. Unfortunately,

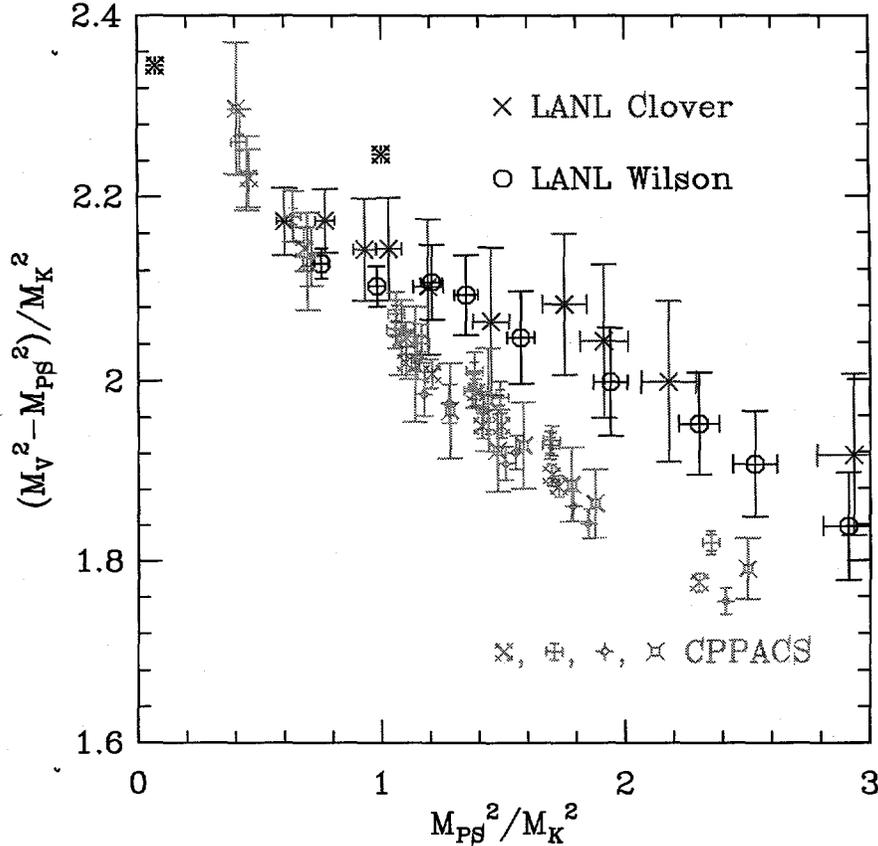


Fig. 25. Hyperfine splittings in the meson sector. Quenched data from (i) LANL collaboration with Wilson fermions (blue octagons) [156], (ii) LANL collaboration with tadpole improved clover fermions (red crosses) [162], and (iii) CP-PACS collaboration Wilson runs at $\beta = 5.9, 6.1, 6.25, 6.47$ (green fancy crosses, pluses, diamonds, and squares) [161]. The experimental points are shown by symbol burst.

the data, in particular the baryon mass-splittings have not yet been fully analyzed so some of the observations presented here should be considered preliminary. Fig. 27, taken from Yoshie's review at LATTICE 97 [161], is CP-PACS's version of the chiral fits. They find that in the range $m_s/4 - m_s$ only the nucleon and the Λ particles show any significant curvature. Even though data at more values of the quark mass are required to resolve the form of the leading higher order correction, the conclusion of their analyses is that the negative curvature found in their fits to nucleon data is suffi-

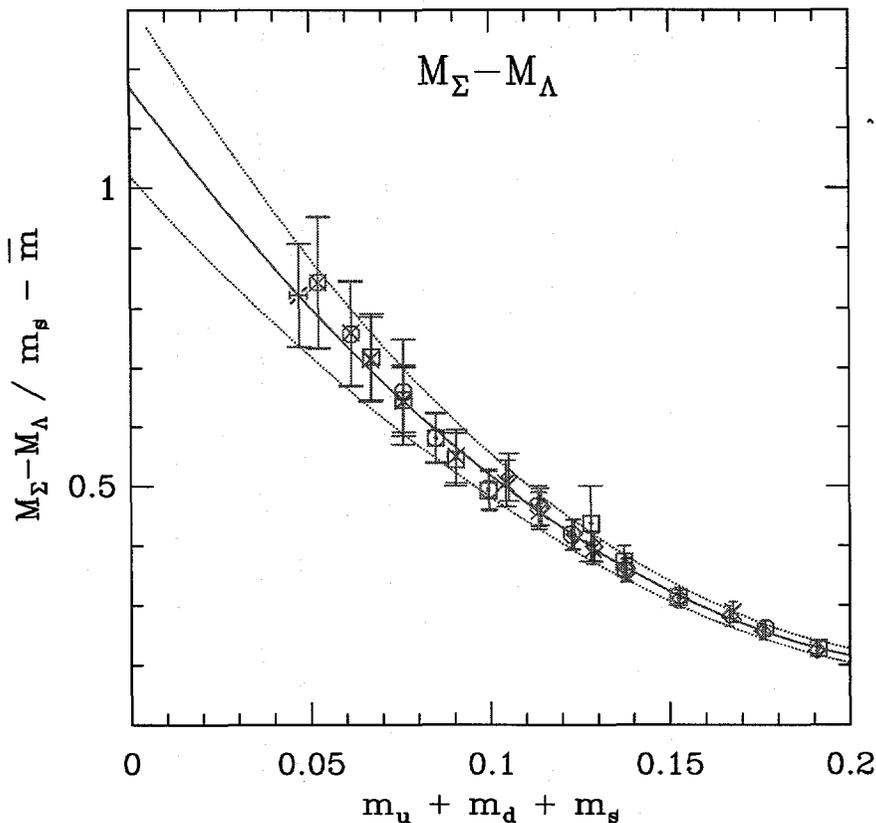


Fig. 26. Quadratic fit to $M_\Sigma - M_\Lambda / (\bar{m} - m_s)$, including baryons composed of completely non-degenerate quarks [156]. The value extrapolated to the physical point is shown by the burst symbol at the extreme left. The linear relations shown in Eq. 18.3 would predict a constant value.

ciently large to change the quenched estimate of M_N/M_ρ to almost perfect agreement with the experimental value ~ 1.22 . (This negative curvature was not seen/resolved in calculations reported before 1997, consequently a higher value in the range 1.3–1.4 was usually reported.) Their extrapolation to $a = 0$ of results with different chiral fits is shown in Fig. 28.

Second, they confirm that the decuplet masses are linear in the quark masses, and the splittings, even after extrapolation to $a = 0$, come out too small with either $m_s(M_K)$ or $m_s(M_\phi)$. Their full analyses on splittings in the baryon octet will appear soon.

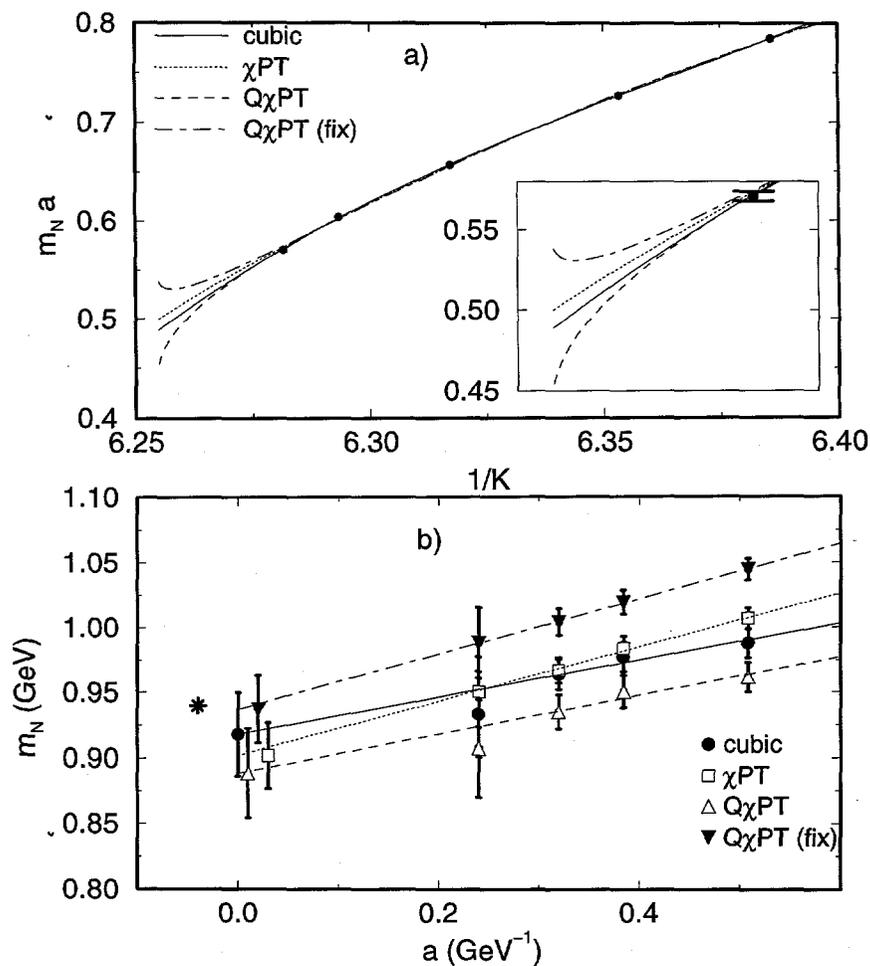


Fig. 27. (A) Chiral fit to $M_{nucleon}$ data by the CP-PACS collaboration using the Wilson gauge and fermion actions. The data show negative curvature, though more points are needed to resolve between the cubic fit $M_N = c_0 + c_1 M_\pi + c_2 M_\pi^2 + c_3 M_\pi^3$, and χ PT ($c_1 = 0$), Q χ PT ($c_3 = 0$), χ PT (fix) ($c_1 = -0.53$) fits. (B) Continuum extrapolation of the nucleon mass obtained using the different chiral expressions are shown in (A). The figures are reproduced from [161].

A final summary of the quenched results taken from [161] using the Wilson action is given in Fig. 29. This figure compares the recent CP-PACS results extrapolated to $a = 0$ with the previous best calculation by

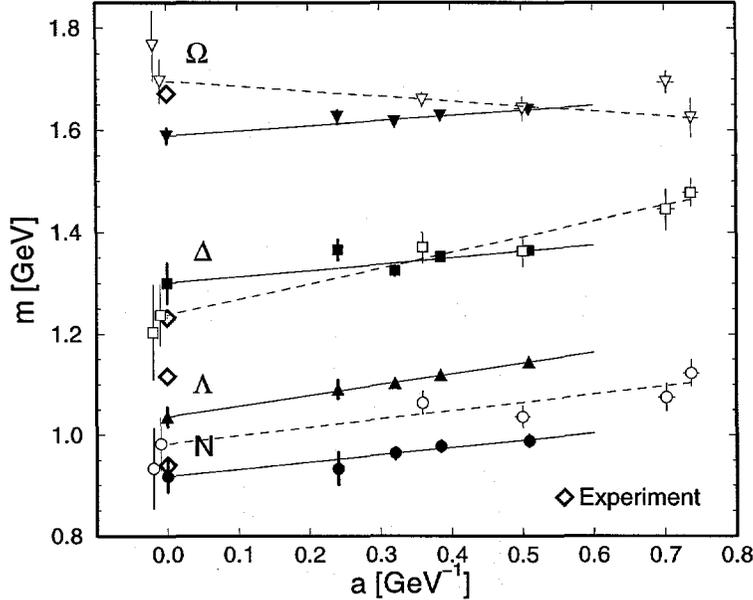


Fig. 28. Continuum extrapolation of the nucleon, Λ , Δ , and Ω masses obtained by the CP-PACS collaboration. The scale a is set by M_ρ . The previous best results by the GF11 collaboration (unfilled symbols) are also shown. Experimental values are shown by the symbol diamond. Figure reproduced from [161].

the GF11 collaboration. The main lesson to be learned from this plot is that there does not seem to be a value of the strange quark mass that allows a matching of all the meson and baryon states. I shall come back to this point when discussing the extraction of quark masses. The obvious question is – is this deviation an artifact of the quenched approximation? My view is that we still do not have full control over the quenched data, and therefore cannot point to quenching as the sole culprit. For example, the definition of m_q itself has corrections of $O(m_q a)$, incorporating which would change the nature of the chiral fits. To fully understand the discretization errors we need to wait for similarly good quality data with the staggered and $O(a)$ improved clover formulations before claiming definitive quenched values.

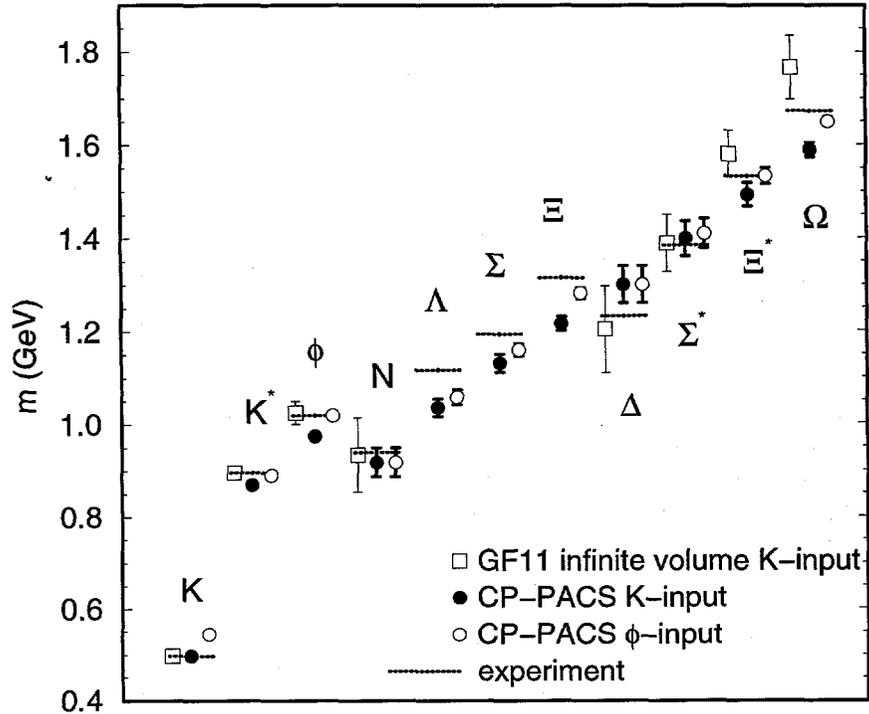


Fig. 29. State-of-the-art results from the CP-PACS collaboration for the meson and baryon masses after $a = 0$ extrapolation and using Wilson fermions. The scale a is set using M_ρ and \bar{m} using M_π . Results are shown for two ways of setting $m_s = M_K$ and M_ϕ and it is clear that neither give correct splittings in the baryon octet and decuplet. The data are reproduced from [161].

As of this writing dynamical simulations have begun in earnest, but precise quantitative results at a number of values of a and with different actions are lacking. Thus, I feel it is too early to draw any conclusions from them. For a status report on $n_f = 2$ results see the recent reviews by Yoshie and Gottlieb [161,160].

18.2. Decay Constants

In addition to the extraction of masses from the rate of exponential decay of the 2-point correlation function, the amplitudes $\langle n|O|0\rangle$ give decay constants. For example, the axial current matrix element $\langle \pi|A_\mu(x)|0\rangle = if_\pi p_\mu \exp(ip \cdot x)$ gives the pion decay constant. Some of the recent results

for decay constants have been reviewed by Prof. Martinelli at this school, and other recent reviews can be found in [163,164]. The technical points I would like to explain here are the reasons why the errors in decay constants are much larger than those in the determination of the masses. First, from Eq. 18.1 one notes that for a 1% error in the determination of the mass, the fractional error in the decay amplitude is $\approx \sqrt{\exp(0.01m\tau)} - 1 \approx 0.005m\tau$. Since fits typically have $m\tau \approx 5$ for $1/a \sim 2$ GeV lattices, the error in the decay constant is $\gtrsim 3\%$. Second, for any given configuration, the correlation function is highly correlated in τ . Plotting $\log(\Gamma(\tau))$ for each individual configuration one notes that the dominant fluctuation is in the overall normalization, *i.e.*, the amplitude. A very rapid growth in these fluctuations with decreasing m_q is, in fact, the signal for exceptional configuration as discussed in Section 10. These two sources of error combine to produce a statistical error in decay constants that is much larger than in the extraction of M_{eff} . To this one has to add a systematic uncertainty coming from the renormalization constants for the local currents.

18.3. Lattice Calculations of the Glueball Spectrum

The existence of glueballs, hadronic states composed mainly of gluons, is a major untested prediction of QCD. Assuming QCD is the correct theory, we can determine the quantum numbers of a vast number of possible glueball states, but cannot yet calculate their masses, or their mixing with $q\bar{q}$, $q\bar{q}q\bar{q}$, ... states, or understand in detail the production and decay mechanisms. Various models like bag models, flux tube model, sum rules *etc.* give estimates of masses that differ significantly from each other and the reliability of these calculations is poor. A somewhat dated summary of the theoretical expectations, compiled by Burnett and Sharpe, can be found in [165]. The lack of knowledge of properties of glueballs does not help the experimentalists who have to isolate these states from the myriad of meson states in the 1 – 2.5 GeV region. Clearly, LQCD calculations can play a vital role by calculating the spectrum and the mixing with $q\bar{q}$ mesons.

The first goal of LQCD has been to calculate the quenched spectrum in a world in which the mixing with quark states is turned off. I shall concentrate on just the 0^{++} and 2^{++} states as these have been studied most extensively. The glueball operators \mathcal{O} are constructed out of Wilson loops. Consider a $n \times m$ Wilson loop $W_{xy}^{n,m}$ lying in the xy plane. Then the rotationally invariant sum $W_{xy}^{n,m} + W_{yz}^{n,m} + W_{zx}^{n,m}$ projects onto 0^{++} state, while $W_{xy}^{n,m} - W_{yz}^{n,m}$ and its permutations project onto 2^{++} . The signal in the 2-point correlator can be improved by taking a suitable linear combination of different sized loops that maximize the overlap with physical

glueball states. One such strategy is discussed in [166]. Unfortunately, the statistical signal in glueball correlation functions is much weaker than that for meson and baryon correlation functions. Even with a sample of $O(10^5)$ independent configurations, the signal does not extend beyond $\tau \sim 10$ or $M\tau \sim 6$ and one barely has a “plateau” in the effective mass plot. The reason for this poor signal in glueball correlation functions is that they measure the fluctuations in the vacuum corresponding to the creation and propagation of a virtual glueball, whereas in meson and baryon correlators the quarks in the propagators are added as external sources.

The more recent, high statistics results for $M_{0^{++}}$ and $M_{2^{++}}$ states have been taken from Refs. [166–169]. The data and the extrapolation to $a = 0$ is shown in Fig. 30. In this plot I have chosen to set the lattice scale using $R_0 = 1.18/\sqrt{\sigma}$, where σ is the string tension [170]. The figure makes it clear that the lattice data from the various collaborations are completely consistent, though the statistical quality of the GF11 data [169] is much higher. A fit in a^2 , the leading discretization error, gives $M_{0^{++}}R_0 = 4.36(9)$ and $M_{2^{++}}R_0 = 6.13(16)$ for the continuum limit of the quenched theory. To obtain masses in physical units one needs to choose a value for R_0 or equivalently σ . This brings us back to the overall scale uncertainty inherent in all quenched calculations. There are two questions relevant to this analysis (i) what value should one choose for σ , and (ii) by how much would the results change if one had used a different quantity like M_ρ or Λ_{QCD} to set the scale? The two most quoted results are the following. The UKQCD-WUPPERTAL collaboration use $1/R_0 = 373$ MeV based on choosing $\sqrt{\sigma} = 440$ MeV and get $M_{0^{++}} = 1626(34)$ and $M_{2^{++}} = 2286(60)$ MeV [170]. The updated GF11 collaboration estimates are $M_{0^{++}} = 1648(58)$ and $M_{2^{++}} = 2273(116)$ GeV using $\Lambda_{QCD}^{(0)} = 235(6)$ MeV which is based on M_ρ to set the scale [171]. Thus, the estimates of the two groups are consistent and agree with the fit shown in Fig. 30. Therefore, in my view, based on these fits, the current combined lattice estimate is $M_{0^{++}} = 1640(30)(160)$ MeV and $M_{2^{++}} = 2280(60)(230)$ MeV, where I have added a second systematic error of 10% to take into account the scale uncertainty inherent to quenched simulations.

18.4. Improved action Results on Anisotropic lattices

To alleviate the problem that the signal in glueball 2-point correlation functions falls rapidly, Peardon and Morningstar have carried out simulations on anisotropic hypercubic lattices, $a_t \ll a_s$ [172]. The advantage of this trick is the following. Assume that there exists an approximate plateau in $M_{eff}(\tau)$ between $\tau = 2 - 4$ in an isotropic lattice simulation before the sig-

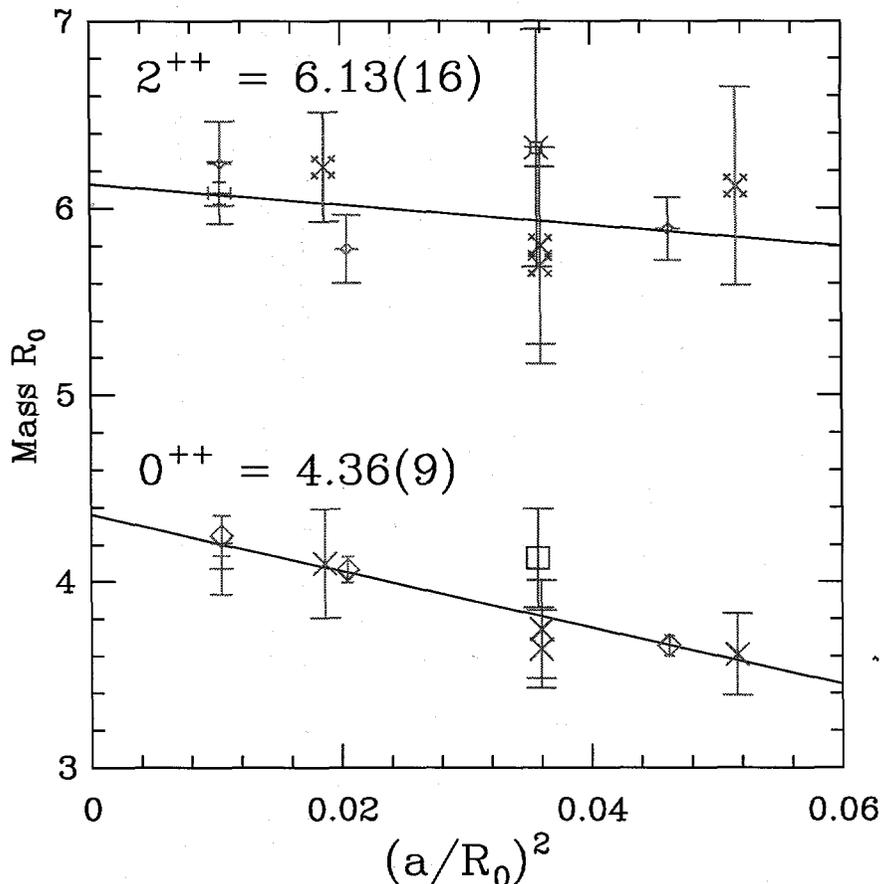


Fig. 30. Extrapolation of 0^{++} and 2^{++} glueball mass data to $a = 0$. The data are from UKQCD-WUPPERTAL collaboration (crosses), GF11 collaboration (diamond), and Los Alamos collaboration (squares). The lattice scale is set using R_0 which is related to the string tension by $1/R_0 = \sqrt{\sigma}/1.18$.

nal degrades to noise. Such a plateau is hard to infer from three time-slices. On the other hand for $a_s/a_t = 5$ lattices, one has roughly 11 time-slices to play with, thereby greatly increasing the confidence in the estimate. Using anisotropic lattices does introduce an additional parameter into the analyses, a_s/a_t or the velocity of light. This they fix non-perturbatively using the measured potential in spatial and temporal directions.

The second improvement they make is to tune the action. They tran-

scribe the tadpole improved Lüscher-Weisz glue action, Eq. 12.14, on to anisotropic lattices and include the plaquette in both the fundamental and adjoint coupling space. The motivation for the latter is to avoid the phase transition in the fundamental-adjoint space as discussed in Section 15. Their new data [139] confirms that including a negative adjoint coupling does improve the scaling in $M_{0^{++}}$.

Their results, prior to those including a negative adjoint coupling, are shown in Fig. 31. The advantage of improved action and anisotropic lattices is clearly visible. First, the calculations can be performed on much coarser lattices. Second, the discretization errors are much smaller. Lastly, the values after extrapolation to $a = 0$ are consistent with conventional methods. Their estimates $M_{0^{++}}R_0 = 3.98(15)$ and $M_{2^{++}}R_0 = 5.85(2)$, when combined with their estimate $R_0 = 410(20)$ MeV, are consistent with the values from conventional simulations given in the previous section. Thus, their results provide another confirmation of the quenched estimates.

Let me end this discussion of gluey states with pointers to other interesting calculations. Results for other spin-parity glueball states can be found in [168,172,173]; and for hybrid states in [173–176].

18.5. Mixing between scalar quarkonia and glueballs

The phenomenologically important question is whether the lattice estimates can help pin down which of the two experimental candidates, $f_0(1500)$ or $f_0(1710)$, is the scalar glueball. The answer, using just the above quenched estimates, is no. The reason, in addition to the fact that the above estimates overlap with both candidates, is that effects of mixing with scalar quarkonia have not been included. Lee and Weingarten have initiated a study of mixing by calculating the two-point functions $\langle \mathcal{S}(\tau)\mathcal{S}(0) \rangle$ and $\langle G(\tau)\mathcal{S}(0) \rangle$ where \mathcal{S} is the scalar density $(\bar{u}u + \bar{d}d)/\sqrt{2}$ or $\bar{s}s$ and G is the 0^{++} glueball operator [171]. The statistical quality of these very hard calculations is not yet good. Lee and Weingarten, therefore, diagonalize an approximate 3×3 matrix that is obtained mainly from a phenomenological analyses. They then find that mixing raises $M_{0^{++}}$ by ~ 64 MeV, *i.e.* $M_{0^{++}}$ changes from $1640 \rightarrow 1704$ MeV. Also, the associated wavefunctions of the mixed states explain some of the unexpected features of decays, like why the $K\bar{K}$ decay mode of $f_0(1500)$ is suppressed and why $\Gamma(J/\psi \rightarrow \gamma f_0(1710)) > \Gamma(J/\psi \rightarrow \gamma f_0(1390)) \gg \Gamma(J/\psi \rightarrow \gamma f_0(1500))$. Consequently, Lee and Weingarten predict that $f_0(1710)$ is dominantly ($\sim 75\%$) a glueball.

In obtaining these predictions, they neglected mixing with other (multi-particle) states and decay widths in the mixing matrix and quenching errors

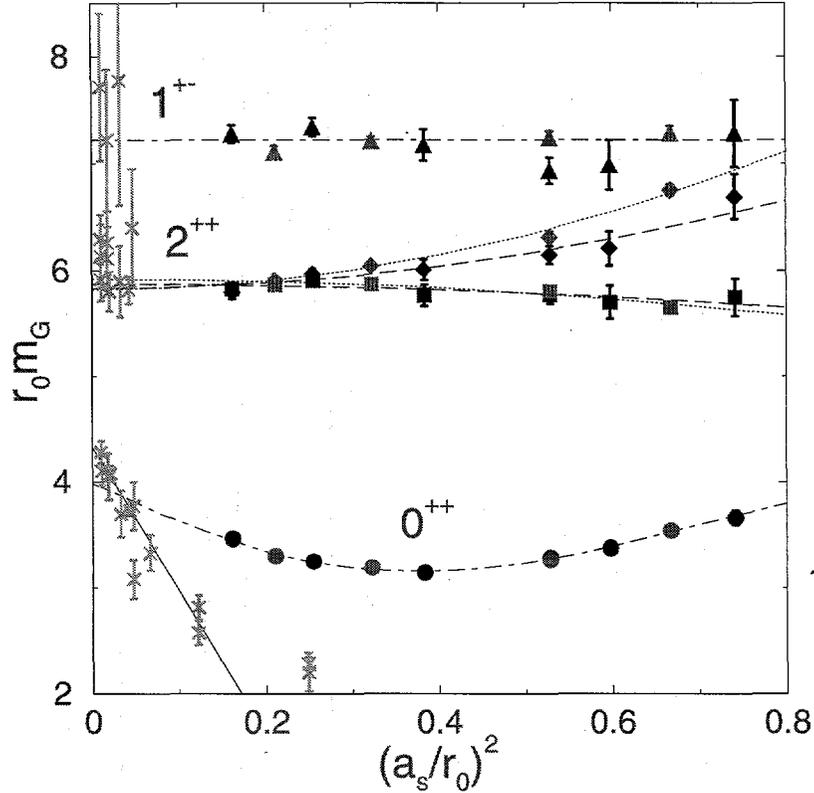


Fig. 31. Glueball mass estimates in terms of R_0 against the lattice spacing $(a_s/R_0)^2$. Results from the $\xi = 5$ simulations for the lattice irreps A_1^{++} (0^{++}), E^{++} , T_2^{++} (2^{++}), and T_1^{+-} (1^{+-}) are labeled \circ , \square , \diamond , and \triangle (blue), respectively. The corresponding symbols in red indicate the results from the $\xi = 3$ simulations. Data from Wilson action simulations taken from Refs. [167–169] are shown using green crosses. The dashed, dotted, and dash-dotted curves indicate extrapolations to the continuum limit obtained by fitting to the $\xi = 3$ data, the $\xi = 5$ data, and all data, respectively. The solid line indicates the extrapolation of the Wilson action data to the continuum limit.

in the lattice calculations. These, Lee and Weingarten claim on the basis of phenomenological arguments, are small and do not change their central conclusion. Hopefully, future LQCD calculations will provide a much better quantitative validation of their prediction.

18.6. Mass Inequalities

In addition to getting hard numbers, LQCD (actually Euclidean field theory) has also prompted the derivation of rigorous mass-inequalities. The first such was derived by Weingarten showing that the pion is the lightest meson state of non-zero isospin in QCD [177]. Consider the 2-point correlation function for mesons, *i.e.* $O = \bar{u}\Gamma d$ where Γ can be any one of the 16 elements of the Clifford algebra. Then

$$\begin{aligned}\Gamma_{\Gamma}(\tau) &= \langle 0 | \bar{d}\Gamma u(\tau, \mathbf{y}) \bar{u}\Gamma d(0, \mathbf{x}) | 0 \rangle \\ &= \langle \text{Tr} S_F^d(0, \mathbf{x}; \tau, \mathbf{y}) \Gamma S_F^u(\tau, \mathbf{y}; 0, \mathbf{x}) \Gamma \rangle \\ &= \langle \text{Tr} \Gamma \gamma_5 S_F^{d\dagger}(\tau, \mathbf{y}; 0, \mathbf{x}) \gamma_5 \Gamma S_F^u(\tau, \mathbf{y}; 0, \mathbf{x}) \rangle.\end{aligned}\quad (18.4)$$

For the special case of the pion, $\Gamma\gamma_5 = 1$, consequently $\Gamma_{\pi}(\tau)$ has the maximum value on each configuration. Furthermore, for $m_u = m_d$ it is the absolute square of the propagator. The other spin-parity states involve γ matrices. Since these have both plus and minus signs, the various spinor terms in the trace have cancellations. The second property needed to translate this property of correlation functions into an inequality on masses is that the integration measure for QCD is positive, *i.e.* each configuration contributes with the same sign. Thus, $\Gamma_{\pi}(\tau) \geq \Gamma_{\Gamma}(\tau)$ for all τ . Since, $\Gamma(\tau) \sim e^{-m\tau}$, the inequality implies that the pion is the lightest meson.

Concurrently, Witten and Nussinov derived further inequalities. Examples of these are

- The electromagnetic mass shift in the pion is positive [179]. Specifically what is shown is that $\langle V_{\mu}^3(k) V_{\mu}^3(-k) - A_{\mu}^3(k) A_{\mu}^3(-k) \rangle \geq 0$ for any Euclidean k .
- For mesons with valence flavors A and B the inequality $2M_{A\bar{B}} \geq M_{A\bar{A}} + M_{B\bar{B}}$ holds in the limit that annihilation into gluonic states can be neglected in the flavor singlet correlators [179,178].

Recently West has applied/extended these methods to glueballs [181]. He shows that 0^{++} is the lightest glueball state.

19. The strong coupling constant α_s

Spectroscopy studies, in addition to testing QCD, allow the determination of some of the fundamental parameters of the SM, *i.e.*, the coupling α_s and the quark masses. To determine α_s , one needs to know not only its strength but also the energy scale Q at which the coupling is measured. The calculation of α_s , in general, proceeds as follows. One chooses a quantity $\langle \mathcal{O} \rangle$

that can be measured very accurately, either in high energy experiments or via non-perturbative methods on the lattice, and for which a perturbative expansion is known, *i.e.*,

$$\langle \mathcal{O} \rangle = \sum_n A_n \alpha_s^n(Q), \quad (19.1)$$

where the A_n are the perturbative coefficients. Inverting this relation gives $\alpha_s(Q)$.

One promising lattice approach for calculating α_s has been pioneered by the NRQCD collaboration [110]. They examine several short-distance $\langle \mathcal{O} \rangle$, like the expectation value of small Wilson loops, that can be calculated very accurately on the lattice and for which perturbative expansions are known to at least $O(\alpha_s^2)$. The scale Q characterizing these lattice observable is typically of the form $Q = C/a$ where a is the lattice spacing and C is a constant that needs to be fixed. They determine C relevant to $\langle \mathcal{O} \rangle$ by evaluating the mean momentum flow in the perturbative expressions for these quantities. The scale $1/a$ is fixed using level splittings in quarkonia. In particular they use the spin-averaged S-P and the 1S-2S splittings in the Upsilon system. The advantage of using these splittings is that they show negligible sensitivity to the precise tuning of the bottom quark mass, and the corrections due to light quark loops are expected to be small.

Finally, to convert from α_s in the lattice scheme to $\alpha_{\overline{MS}}$ they use the two loop relation, and include reasonable estimates of the α_s^3 term in the error analysis. Their final estimate

$$\alpha_{\overline{MS}}(M_Z) = 0.1174(24) \quad (19.2)$$

includes quenching uncertainties. This remarkably precise result is consistent with the world average of experimental determinations, $\alpha_s = 0.118(3)$, as shown in Fig. 32.

An alternate method for measuring α_s has been carried out by the ALPHA collaboration [182]. In their method, called the Schrödinger functional approach, α_s is defined by measuring the response of the lattice system to an external field. They fix the lattice scale a using the force at some fixed distance ($R_0 = 0.5$ fermi). They evolve their coupling to a very high scale using a non-perturbative renormalization group stepping procedure. At this high scale they define $\Lambda_{\overline{MS}}^{(0)}$ using the two loop relation and include the three loop uncertainty, which is small, in error estimates. They have completed the calculation for the quenched theory with the result $\Lambda_{\overline{MS}}^{(0)} = 251(21)$ MeV, and are now doing the $n_f = 2$ simulations [182]. The corresponding value $\alpha_{\overline{MS}}(90 \text{ GeV}) = 0.0818(10)$ is lower than in nature. This difference is attributed to having neglected sea quark effects.

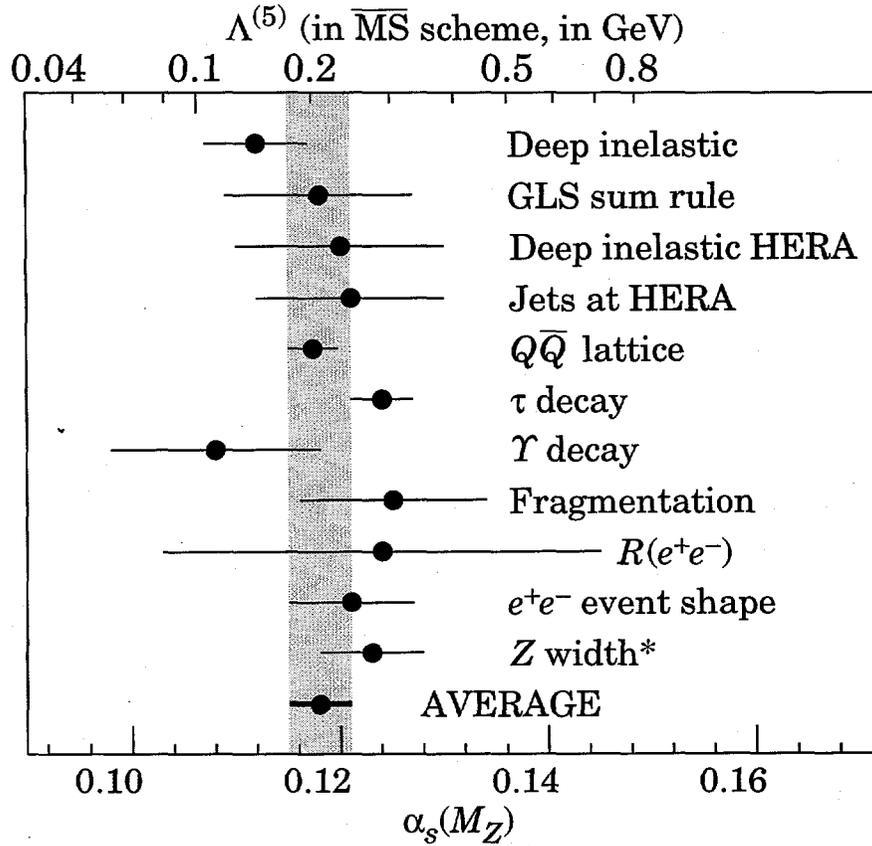


Fig. 32. Summary of the values of $\alpha_s(M_Z)$ and $\Lambda^{(5)}$ (the QCD scale for 5 active quark flavors) extracted from various processes. These are ordered from top to bottom by increasing energy scale of measurements. This figure is reproduced from the 1996 Review of Particle Physics, updated to show the latest lattice result, $\alpha_{\overline{MS}}(M_Z) = 0.1174 \pm 0.0024$. It is a remarkable confirmation of QCD that the α_s extracted from processes at different scales has the same value.

A similar low value is obtained in the NRQCD analyses if no corrections are made for the quenched approximation. For details of the Schrödinger functional method see the lectures by Prof. Lüscher at this school.

20. Quark Masses

Quarks are not asymptotic states of nature, so their masses cannot be measured in experiments. They are generally defined in some perturbative scheme like \overline{MS} and at some scale μ . Their values remain amongst the least well known of the Standard Model parameters. In this section I will discuss their extraction from lattice data.

The extraction of the up, down, and strange quark masses from the hadron spectrum is done by inverting Eqs. 18.3 once the coefficients A_π etc. are determined from fits to the lattice data. As I have mentioned before, current lattice simulations neglect iso-spin breaking effects, so we can only determine $\overline{m} = (m_u + m_d)/2$ and m_s . I will discuss these first and then briefly mention the extraction of the heavy quark masses m_c and m_b .

In the quenched approximation, the most straightforward way to estimate \overline{m} is to extrapolate M_π^2/X^2 to the physical value. Here X is a dimensionful quantity like f_π or M_ρ or M_N or M_Δ that does not depend on m_s , and de facto sets the lattice scale $1/a$. Of the above quantities, using M_ρ is the most controversial as it decays via strong interactions and has a large width, 170 MeV. On the other hand the lattice data for M_ρ is most readily available and has good signal. Also, in the quenched approximation the ρ does not decay. Using M_ρ to set a in QCD versus some other quantity can therefore give an overall scale uncertainty. The spread in $1/a$ extracted from different quantities is found to be $\approx 10\%$, and this uncertainty, inherent to the quenched approximation, will be treated as an overall systematic error. Similarly, one can set m_s by extrapolating the ratio \mathcal{M}_s/X , where \mathcal{M}_s is any state that depends on m_s like M_K , or M_{K^*} , or M_ϕ or any of the baryons containing a strange quark. The world data show that different choices of \mathcal{M}_s give rise to a $\approx 20\%$ spread in m_s . It is important to note that if one assumes the relation $M_{\mathcal{P}\mathcal{S}}^2 = B_{\mathcal{P}\mathcal{S}} m_q$ (which LQCD has not validated as true for the whole range $[\overline{m}, m_s]$), and uses M_K to fix m_s , then the ratio m_s/\overline{m} has to be 25.9 if \overline{m} is set using M_π^2 .

In the full theory X depends on m_s , and similarly \mathcal{M}_s depends on \overline{m} , through sea quark effects. Thus, \overline{m} , a , and m_s have to be fixed by solving the coupled set of equations.

Simulations of the hadron spectrum have been done by a number of collaborations around the world, and there exists a large amount of data for the masses of hadrons as a function of the quark masses and lattice spacing. (Almost all lattice collaborations, for example the APE, Columbia University, CP-PACS, Fermilab, GF11, HEMCGC, JLQCD, Los Alamos, SCRI, staggered, UKQCD, and the Wuppertal University collaborations, have contributed to these calculations [183].) In early analyses the estimates of

quark masses varied by almost a factor of two between different formulations of fermions, and the pattern of the discretization and quenching errors was not obvious from the work of any one collaboration. Clarification of this issue has been obtained by putting together all the world data for Wilson, clover, and staggered fermions [183,184]. The final picture for \bar{m} , based on the most recent, high statistics, and large lattices data is shown in Fig. 33 for the three fermion formulations. This figure shows clearly that the factor of ~ 2 difference at $1/a \sim 2$ GeV shrinks to $\sim 15\%$ after extrapolation to $a = 0$, *i.e.* most of the difference can be explained as due to discretization errors.

The behavior of m_s versus a is very similar. A summary of the quenched results in MeV at scale $\mu = 2$ GeV, based on an analysis of the 1997 world data [185], is

	Wilson	TI Clover	Staggered
$\bar{m}(M_\pi)$	4.1(1)	3.8(1)	3.5(1)
$m_s(M_{K^*})$	107(2)	99(3)	91(2)
$m_s(M_\phi)$	139(11)	117(8)	109(5)

These estimates were obtained using the following *Ansatz*.

- Chiral fits were made keeping only the lowest order terms shown in Eqs. 18.3.
- M_π^2/M_ρ is used to fix \bar{m} .
- Results for m_s are given for two choices, $M_{K^*}^2/M_\rho$ and M_ϕ/M_ρ . (Using M_{K^*}/M_ρ gives values almost identical to those with M_ϕ/M_ρ . This implies that for the choice $m_s(M_\phi)$ one could equivalently have set the scale using M_{K^*} instead of M_ρ .)
- Since linear chiral fits have been used, $m_s(M_{K^*}) \equiv 25.9\bar{m}$.
- The matching factor, Z_{m_s} , between lattice and \overline{MS} schemes is the tadpole improved 1-loop result.
- The extrapolation to $a = 0$ is done using the lowest order discretization error, *i.e.* $O(a)$ and $O(\alpha_s a)$ for Wilson and clover, and a quadratic in a^2 for staggered.

The final difference in estimates between Wilson, tadpole improved clover (TI Clover), and staggered results is $\approx 15\%$ as shown in Fig. 33. This could be due to the neglected higher order discretization errors and/or due to the difference between non-perturbative and 1-loop estimates of Z 's. Similarly, the $\approx 20\%$ variation in m_s with the state used to extract it, M_{K^*} versus M_ϕ (or equivalently M_{K^*}) could be due to the quenched approximation or partly an artifact of keeping only the lowest order discretization correction in the extrapolations. To disentangle these discretization and

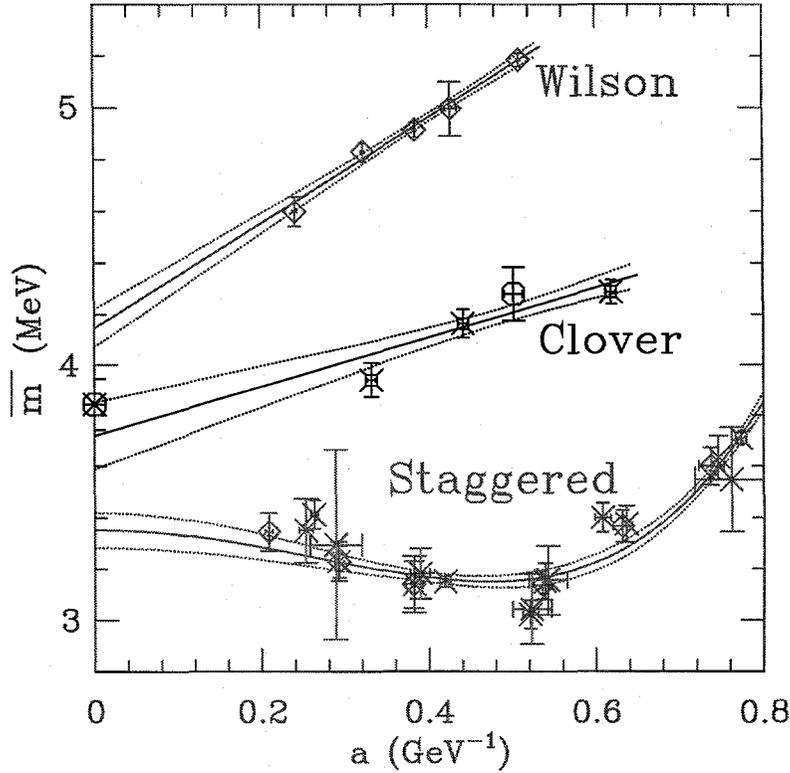


Fig. 33. The behavior of $m_u + m_d$ as a function of the lattice spacing a for three discretization schemes – Wilson, clover, and staggered. Linear extrapolations to $a = 0$ are shown for Wilson and clover formulations, while for staggered both $O(a^2)$ and $O(a^4)$ terms are incorporated in the fit. The theoretically expected form for discretization errors in the clover formulation is $O(\alpha_s a)$. The extrapolated value using this form is shown by the symbol burst.

quenching errors requires precise unquenched data with roughly physical values for light quarks.

For best estimates of quenched results I average the data and use the spread as the error. To these, I add a second systematic error of 10% due to the uncertainty in the determination of the scale $1/a$, which partly reflects the uncertainty due to the quenched approximation. The results, in the \overline{MS} scheme and evaluated at 2 GeV, are [185]

$$\overline{m} = 3.8(4)(4) \text{ MeV}$$

$$m_s = 110(20)(11) \text{ MeV}. \quad (20.1)$$

20.1. Light Quark Masses from Ward Identities

A second method for extracting quark masses is based on axial and vector current Ward identities

$$\begin{aligned} Z_A \partial_\mu A_\mu &= (m_1 + m_2)_R Z_P \mathcal{P}, \\ Z_V \partial_\mu V_\mu &= (m_1 - m_2)_R Z_S \mathcal{S}, \end{aligned} \quad (20.2)$$

where 1, 2 are the flavor indices and \mathcal{P} and \mathcal{S} are the lattice pseudoscalar and scalar densities. From these one can construct ratios of 2-point correlation functions

$$\begin{aligned} \langle \partial_\mu A_\mu(\tau) \mathcal{P} \rangle &= (m_1 + m_2)_R \frac{Z_P}{Z_A} \langle \mathcal{P}(\tau) \mathcal{P}(0) \rangle \\ \langle \partial_\mu V_\mu(\tau) \mathcal{S}(0) \rangle &= (m_1 - m_2)_R \frac{Z_S}{Z_V} \langle \mathcal{S}(\tau) \mathcal{S}(0) \rangle, \end{aligned} \quad (20.3)$$

which give the renormalized quark masses once the various renormalization constants Z_i are known. Since the Ward identities are operator identities, Eq. 20.3 holds for all τ except at very short times where there is influence of contact terms. Thus it is not necessary for the signal to persist to asymptotic τ as needed in spectroscopy method to isolate the lowest state.

The two estimates of quark masses, from spectroscopy and from the Ward identities, should agree in the continuum limit. A comparison of the two data sets for the Wilson action and using 1-loop perturbative Z 's is shown in Fig. 34. The two estimates are significantly different at finite a , however, a linear extrapolation works for both sets of data, and the extrapolated values are in surprisingly good agreement. One could therefore attribute the differences at finite a to discretization errors.

The recent non-perturbative calculations of the Z 's [186] challenge this nice picture. Fig. 34 also shows the same data analyzed using these non-perturbative Z 's. At fixed a the non-perturbative Z 's increase both the hadron spectroscopy and WI results to roughly the same value. This suggests that most of the difference is due to the $O(\alpha_s^2)$ corrections to the 1-loop Z 's and not due to the $O(a)$ discretization errors as mentioned before. However, the non-perturbative calculation of Z 's themselves have $O(a)$ errors that have not been resolved in the results of [186]. As a result we do not have a clear separation between the effects of the $O(a)$ discretization errors and the $O(\alpha_s^2)$ errors in these Z 's.

The bottom line is that, even though we have not resolved the two kinds of errors, and while both the value at fixed a and the slope in a

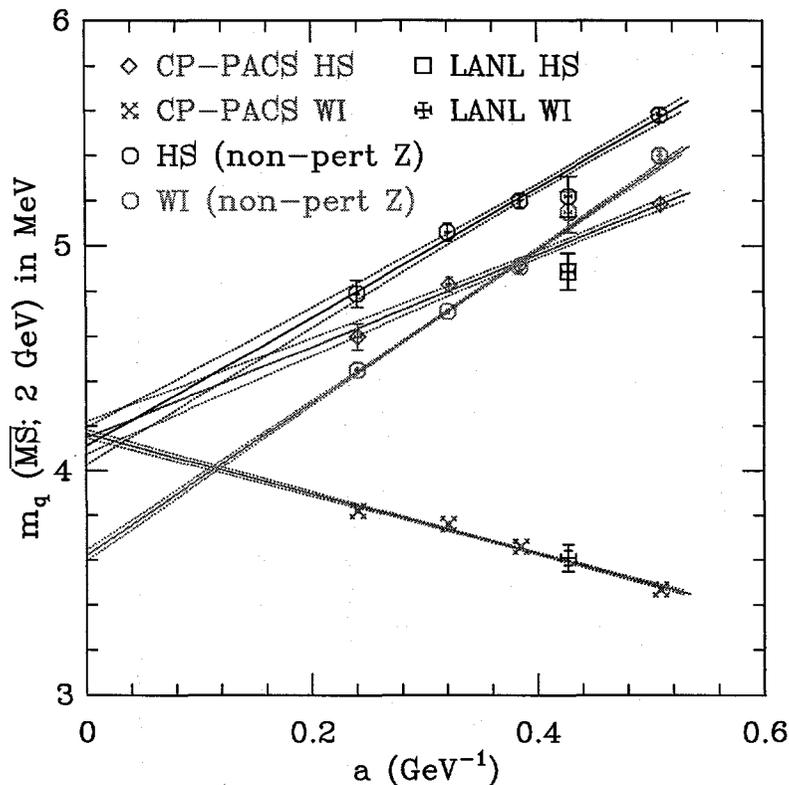


Fig. 34. The behavior of $\bar{m} = (m_u + m_d)/2$ as a function of the lattice spacing a for Wilson fermions. Data from the hadron spectroscopy and Ward Identity method are shown for two ways of estimating the renormalization constants, *i.e.* one loop perturbative and non-perturbative.

increases significantly on changing from 1-loop to non-perturbative Z 's, the extrapolated value is roughly unchanged.

To conclude, I believe that we have finally obtained values for light quark masses with reliable estimates of statistical and all systematic errors other than quenching. Calculations are already underway to significantly increase the number of data points with $n_f = 2$ and 4 flavors of dynamical quarks to reduce the quenching uncertainty. Also, new ways to calculate the renormalization constants non-perturbatively are being developed. Therefore, I expect that in the next few years there will be a significant increase in the precision of both QCD and full QCD estimates of light quark masses.

20.2. Schrodinger Functional Approach

In the above discussion the term non-perturbative Z 's referred to only the lattice scheme. In the required matching constants one still has to use perturbative expressions for the renormalization constants in the continuum scheme and make a good estimate of the matching point. The ALPHA Collaboration has proposed a method for removing this last vestige of dependence on perturbation theory [187]. The details of their method have been discussed by Prof. Lüscher at this school, so I will just summarize the key points.

- (i) Calculate the quark mass using the axial WI method at some infrared lattice scale. At the same time calculate Z_P and Z_A non-perturbatively in the Schrödinger functional scheme.
- (ii) Evolve this lattice result to some very high scale using the step scaling function calculated non-perturbatively.
- (iii) At this high scale, where α_s is small, define the renormalization group independent mass

$$\hat{m} = \lim_{a \rightarrow 0} m_R (2\beta_0 g^2)^{-\gamma_0/2\beta_0} \quad (20.4)$$

This \hat{m} is scheme independent, thus subsequent conversion to a running mass in some scheme and at some desired scale can be done in the continuum using the most accurate versions (4-loop) of the scale evolution equations [188].

To reiterate, the advantage of this method is that the lattice calculations directly predict the scale and scheme invariant quantity \hat{m} . The calculations by the ALPHA Collaboration are still in progress and at this time estimates for \hat{m} have not been released. It will be interesting to see how they compare with conventional hadron spectroscopy or WI methods discussed above. My prejudice is that they will be very similar.

20.3. Role of quark masses in CP violation

A very opportune application of the precise determination of the light quark masses is in the Standard Model prediction of direct CP violation in kaon decays characterized by the ratio ϵ'/ϵ (see lectures by Prof. Buras and

[189]). The reason is that ϵ'/ϵ is, to a good approximation, proportional to $1/(m_d + m_s)^2$ [190],

$$\epsilon'/\epsilon = A \left\{ c_0 + [c_6 B_6^{1/2} + c_8 B_8^{3/2}] \left[\frac{158 \text{ MeV}}{(m_s + m_d)} \right]^2 \right\}, \quad (20.5)$$

where all quantities are to be evaluated at the scale $m_c = 1.3 \text{ GeV}$. Eq. 20.5 highlights the dependence on the light quark masses and the bag parameters $B_6^{1/2}$ and $B_8^{3/2}$ which LQCD is also expected to provide. The coefficients c_0, c_6 , and c_8 have been calculated to next-to-next-to-leading order, and with M_{top} now known the main uncertainty in their extraction comes from the value of Λ_{QCD} . The quantity $A = |V_{ub}||V_{cb}|\sin\delta$ in the parameterization of Eq. 2.4, requires input of other SM parameters like $f_B, |V_{ub}|, |V_{cb}|, B_K, \sqrt{B_{B_d}}f_{B_d}$. Using the central values quoted by Buras at scale $m_c = 1.3 \text{ GeV}$ [190] one gets $A = 1.29 \times 10^{-4}$, $c_0 = -1.4$, $c_6 = 7.9$, $c_8 = -4.0$. Unfortunately, the errors in these quantities make the resulting theoretical prediction, $\epsilon'/\epsilon = 3.6(3.4) \times 10^{-4}$ [190], very loose.

Current experimental estimates are $7.4(5.9) \times 10^{-4}$ from Fermilab E731 [191] and $23(7) \times 10^{-4}$ from CERN NA31 [192]. The new generation of experiments, Fermilab E832, CERN NA48, and DAΦNE KLOE, will reduce the uncertainty to $\approx 1 \times 10^{-4}$. First results from these experiments (see lectures by L. Fayard) should be available in the next couple of years. Thus, it is very important to tighten the theoretical prediction. Clearly, lower estimates of quark masses suggested by LQCD would significantly increase the size of direct CP violation, and make its experimental determination and validation much easier.

20.4. Heavy quark masses m_c and m_b

The masses of the heavy quarks, m_c and m_b , have been calculated using the charmonium and upsilin spectroscopy. For heavy quarks there exist two different definitions of the quark mass -- the pole and \overline{MS} masses. The pole mass is defined to be $m^{(pole)} = (M_{onia} - E_{binding})/2$, where M_{onia} is taken from experimental data and $E_{binding}$ is calculated on the lattice. This calculation has been done using two approaches, NRQCD [193] and heavy quark effective theory (HQET) [194], and the results are consistent. The \overline{MS} mass, which has the advantage of being free of renormalon ambiguities that are inherent in $m^{(pole)}$, is obtained from $m^{(pole)}$ using perturbation theory. The present results are

$$\begin{aligned} m_b^{\overline{MS}}(m_b^{\overline{MS}}) &= 4.15(5)(20) \text{ GeV} \quad (\text{APE}) \\ &= 4.16(15) \text{ GeV} \quad (\text{NRQCD}). \end{aligned} \quad (20.6)$$

These estimates are consistent with determinations from other non-lattice methods.

The agreement is not as good for various lattice estimates of $m_c^{\overline{MS}}$. Current estimates lie between 1.2 – 1.5 GeV. For example, three different methods investigated by the APE collaboration [195,186], Fermilab collaboration [196,197], and by Bochkarev and Forcrand [198] who evaluate the correlation functions arising in QCD sum-rules on the lattice, give

$$\begin{aligned} m_c^{\overline{MS}}(m_c^{\overline{MS}}) &= 1.525(40)(125) \text{ GeV} && \text{APE} \\ &= 1.33(8) \text{ GeV} && \text{Fermilab} \\ &= 1.22(5) \text{ GeV} && \text{B\&F.} \end{aligned} \tag{20.7}$$

At this stage I do not consider the difference significant as the various systematic errors have not been resolved equally well in all the calculations.

Conclusions and Acknowledgements

I hope that these lectures succeed in providing a pedagogical introduction to Lattice QCD, and to the analyses of numerical data. Undoubtedly, some topics have been left out and many are sketchy. My purpose was to give a broad overview. I have concentrated on providing the basics and have illustrated the analyses of Monte Carlo data by the calculations of the hadron spectrum, strong coupling constant α_s , and quark masses. Prof. Martinelli will, among other things, discuss the extraction of matrix elements of weak Hamiltonian, while Prof. Lüscher will provide a panoramic view of what future calculations will look like and the improvement possible with better discretization schemes. Together, I hope these set of lectures provide sufficient details and excitement to motivate some of you to probe deeper into the field.

I express my gratitude to all the lecturers, students, and staff who made this Les Houches school fun and exciting. It, most certainly, was a memorable six weeks for me. In preparing these lectures I would like to thank G. Bali, U. Heller, C. Michael, M. Peardon, the CP-PACS collaboration, in particular S. Hashimoto and T. Yoshie, for providing unpublished data. Special thanks go to my collaborators T. Bhattacharya, D. Daniel, G. Kilcup, A. Patel, and S. Sharpe from whom I have, over the years, learned so much. Finally, I would like to thank T. Bhattacharya, D. Daniel, W. Lee, and S. Sharpe for proof reading these notes and for their suggestions.

References

- [1] M. Creutz, *Quarks, Gluons, and Lattices*, Cambridge University Press, 1983.
- [2] *Quantum Fields on the Computer*, Ed. M. Creutz, World Scientific, 1992.
- [3] I. Montvay and G. Münster, *Quantum Fields on a Lattice*, Cambridge University Press, 1994.
- [4] H. J. Rothe, *Quantum Gauge Theories: An Introduction*, World Scientific, 1997.
- [5] *Phenomenology and Lattice QCD*, Eds. G. Kilcup and S. Sharpe, World Scientific, 1995.
- [6] 1989 TASI Lectures, Eds. T. Degrاند and D. Toussaint, World Scientific, 1990.
- [7] J. Kogut, *Rev. Mod. Phys.* **51** (1979) 659.
- [8] J. Kogut, *Rev. Mod. Phys.* **55** (1983) 775.
- [9] LATTICE93 *Nucl. Phys. (Proc. Suppl.)* **B34** (1994) 1; LATTICE94 *Nucl. Phys. (Proc. Suppl.)* **B42** (1995) 1; LATTICE95 *Nucl. Phys. (Proc. Suppl.)* **B47** (1996) 1; LATTICE96 *Nucl. Phys. (Proc. Suppl.)* **B53** (1997) 1; LATTICE97 *Nucl. Phys. (Proc. Suppl.)* **B63** (1998) 1.
- [10] A. Buras, hep-ph/9711217; these proceedings.
- [11] J. Richman, these proceedings.
- [12] P. van Baal, *Nucl. Phys. (Proc. Suppl.)* **B63** (1998) 126.
- [13] C. Detar, To appear in 'Quark Gluon Plasma 2', edited by R. Hwa, World Scientific, 1998 (hep-ph/9504325); A. Ukawa, Lectures at INT, Published in Ref. [5]. J.-P. Blaizot, *J. Korean Physical Society* **25** (1992) S65-S98; J. Kapusta, *Finite Temperature Field Theory*, Cambridge University Press, 1989.
- [14] H. B. Nielsen and M. Ninomiya, *Nucl. Phys.* **B185** (1981) 20; Erratum *Nucl. Phys.* **B195** (1982) 541; *Nucl. Phys.* **B193** (1981) 173.
- [15] Y. Shamir, *Nucl. Phys. (Proc. Suppl.)* **B53** (1997) 664; *Nucl. Phys. (Proc. Suppl.)* **B47** (1996) 212.
- [16] R. Narayanan, hep-lat/9707035.
- [17] M. Testa, hep-lat/9707007.
- [18] P. Hasenfratz, hep-lat/9802007; M. Lüscher, hep-lat/9802011.
- [19] K. G. Wilson, *Phys. Rev.* **D10** (1974) 2445; in *New Phenomena in Subnuclear Physics* (Erice 1975), Ed. A. Zichichi, Plenum New York, 1977.
- [20] B. Sheikholesalami and R. Wohlert, *Nucl. Phys.* **B259** (1985) 572.
- [21] J. Kogut and L. Susskind, *Phys. Rev.* **D11** (1975) 395; T. Banks, J. Kogut, and L. Susskind, *Phys. Rev.* **D13** (1996) 1043; L. Susskind, *Phys. Rev.* **D16** (1977) 3031.
- [22] G. Parisi, *Statistical Field Theory*, Benjamin/Cummings, 1988.
- [23] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, Oxford University Press, Oxford, 1993.
- [24] K. Osterwalder and R. Schrader, *Comm. Math. Phys.* **31** (1973) 83; *Comm. Math. Phys.* **42** (1975) 281.
- [25] M. Lüscher, *Comm. Math. Phys.* **54** (1977) 283.
- [26] C. Michael, *Nucl. Phys.* **B327** (1990) 515; T. DeGrand, *Phys. Rev.* **D43** (1991) 2296.
- [27] M. Lüscher, *Selected Topics In Lattice Field Theory*, Lectures at the 1988 Les Houches Summer School, North Holland, 1990.
- [28] M. Lüscher, *Comm. Math. Phys.* **104** (1977) 177; *Comm. Math. Phys.* **105** (1986) 153; *Nucl. Phys.* **B354** (1991) 531; *Nucl. Phys.* **B364** (1991) 237.

- [29] S. Sharpe, R. Gupta, and G. Kilcup, Nucl. Phys. **B383** (1992) 309; R. Gupta, A. Patel, and S. Sharpe, Phys. Rev. **D48** (1993) 388;
- [30] M. Fukugita, Y. Kuramashi, H. Mino, M. Okawa, and A. Ukawa, Phys. Rev. **D52** (1995) 3003.
- [31] K. Wilson, Phys. Rev. **D10** (1974) 2445.
- [32] R. Celmaster, Phys. Rev. **D26** (1982) 2955.
- [33] N.H. Christ, R. Friedberg, and T.D. Lee, Nucl. Phys. **B202** (1982) 89; Nucl. Phys. **B210** (1982) 310, 337.
- [34] J. Mandula, G. Zweig, J. Goverts, Nucl. Phys. **B228** (1983) 91, 109
- [35] G. P. Lepage, hep-lat/9607076
- [36] H. Kawai, R. Nakayama, and K. Seo, Nucl. Phys. **B189** (1981) 40.
- [37] S. Elitzur, Phys. Rev. **D12** (1975) 3978.
- [38] D. Leinweber, *et.al.*, hep-lat/9803015.
- [39] C. Parrinello, *et.al.*, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 245.
- [40] R. Gupta, D. Daniel and J. Grandy, Phys. Rev. **D48** (1993) 3330.
- [41] V. Giménez, *et.al.*, hep-lat/9806006.
- [42] V. N. Gribov, Nucl. Phys. **B139** (1978) 1.
- [43] P. van Baal, hep-th/9711070.
- [44] A. Cucchieri and T. Mendes, Nucl. Phys. (Proc. Suppl.) **B53** (1997) 811.
- [45] J. Hetrick and Ph. de Forcrand, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 838.
- [46] M.C. Paciello, *et.al.*, Phys. Lett. **289B** (1992) 405; Ph. de Forcrand and J. Hetrick, Nucl. Phys. (Proc. Suppl.) **B42** (1995) 861; A. Cucchieri, Nucl. Phys. **B508** (1997) 353.
- [47] G. Martinelli, *et.al.*, Phys. Lett. **411B** (1997) 141.
- [48] L. Maiani and M. Testa, Phys. Lett. **245B** (1990) 585.
- [49] K. Symanzik, Nucl. Phys. **B226** (1983) 187 and 205.
- [50] M. Lüscher, Comm. Math. Phys. **85** (1982) 39.
- [51] J. Grandy and R. Gupta, Nucl. Phys. (Proc. Suppl.) **B42** (1995) 246; J. Grandy and G. Kilcup, Nucl. Phys. (Proc. Suppl.) **B53** (1997) 560.
- [52] T. DeGrand, A. Hasenfratz, and T. Kovacs, hep-lat/9801037.
- [53] P. Hasenfratz, hep-lat/9802007.
- [54] P. Ginsparg and K. W. Wilson, Phys. Rev. **D25** (1982) 2649.
- [55] M. Lüscher, hep-lat/9802011.
- [56] K. G. Wilson, in *New Phenomena in Subnuclear Physics*, ed. A. Zichichi (Plenum Press, 1977).
- [57] M. Bochicchio, L. Maiani, G. Martinelli, G.C. Rossi and M. Testa, Nucl. Phys. **B262** (1985) 331; Phys. Lett. **176B** (1986) 445; Phys. Lett. **178B** (1986) 265; Phys. Lett. **181B** (1986) 344; Nucl. Phys. **B289** (1987) 505; and Nucl. Phys. **B293** (1987) 420.
- [58] Karsten and J. Smit, Nucl. Phys. **B183** (1981) 103.
- [59] W. Bardeen, *et.al.*, Phys. Rev. **D57** (1998) 1633.
- [60] A. Duncan, E. Eichten, and H. Thacker, hep-lat/9806020.
- [61] M. Gell-Mann, R.J. Oakes, and B. Renner, Phys. Rev. **175** (1968) 2195.
- [62] D. Daniel, R. Gupta, G. Kilcup, A. Patel, and S. Sharpe, Phys. Rev. **D46** (1992) 3130.
- [63] G. Kilcup and S. Sharpe, Nucl. Phys. **B283** (1987) 493.
- [64] H.S. Sharatchandra, H.J. Thun, and P. Weisz, Nucl. Phys. **B192** (1981) 205; C. van den Doel and J. Smit, Nucl. Phys. **B228** (1983) 122; M. Golterman and J. Smit, Nucl. Phys. **B245** (1984) 61.

- [65] H. Kluberg-Stern, A. Morel, O. Napoly, and B. Petersson, Nucl. Phys. **B220** (1983) 447.
- [66] F. Gliozzi, Nucl. Phys. **B204** (1982) 419.
- [67] D. Daniel, and T.D. Kieu, Phys. Lett. **175B** (1986) 73.
- [68] M. Golterman, Nucl. Phys. **B273** (1986) 663.
- [69] R. Gupta, G. Guralnik, G. Kilcup, and S. Sharpe, Phys. Rev. **D43** (1991) 2003.
- [70] M. Golterman and J. Smit, Nucl. Phys. **B255** (1985) 328.
- [71] S. Sharpe *Standard Model Phenomenology and Weak Decay on the lattice*, Ed. G. Martinelli, World Scientific, 1998.
- [72] G. Kilcup, *et.al.*, Phys. Rev. Lett. **64** (1990) 25; G. Kilcup, R. Gupta, and S. Sharpe, Phys. Rev. **D57** (1998) 1654;
- [73] M. Lüscher and P. Weisz, Commun. Math. Phys. **97** (1985) 59; Erratum Commun. Math. Phys. **98** (1985) 433.
- [74] G. Curci, P. Menotti and G. Paffuti, Phys. Lett. **130B** (1983) 205.
- [75] P. Weisz and R. Wohlert, Nucl. Phys. **B236** (1984) 397.
- [76] M. Lüscher and P. Weisz, Phys. Lett. **158B** (1985) 250.
- [77] G. Parisi, *High Energy Physics - 1980, AIP Conference Proceedings No. 68*, Eds. L. Durand and L. Pondrom, New York, 1981. .
- [78] P. Lepage and P. Mackenzie, Phys. Rev. **D48** (1993) 2250.
- [79] M. Alford, T. Klassen, and P. Lepage, hep-lat/9712005.
- [80] P. Weisz and R. Wohlert, Nucl. Phys. **B236** (1984) 397.
- [81] M. Alford, *et.al.*, Nucl. Phys. (Proc. Suppl.) **B42** (1994) 787.
- [82] M. Alford, *et.al.*, Phys. Lett. **361B** (1995) 87.
- [83] K. Wilson and J. Kogut, Phys. Rev. **12** (1974) 76.
- [84] K. Wilson, in *Recent Developments in Gauge Theories*, edited by G. t'Hooft, *et.al.*, Plenum Press, N.Y. 1980.
- [85] M. Fisher, *Lecture notes in Physics, Vol. 186*, Springer Verlag, 1983.
- [86] R. Gupta, in proceedings of LATTICE 85, Wuppertal 1985, Eds. K-H Mütter and K. Schilling, Plenum Press (1986).
- [87] A. Patel and R. Gupta, Phys. Lett. **183B** (1987) 193.
- [88] R. Gupta, lectures at TASI 89 [6].
- [89] R. Gupta, in *Quantum Fields on the Computer*, Ed. M. Creutz, World Scientific, 1992.
- [90] P. Hasenfratz, hep-lat/9803027.
- [91] Y. Iwasaki, University of Tsukuba report UTHEP-118, 1983.
- [92] K. Kanaya, hep-lat/9804006.
- [93] R. Gupta, *et.al.*, Phys. Rev. **D43** (1991) 2301.
- [94] QCDTARO Collaboration, Nucl. Phys. (Proc. Suppl.) **B53** (1997) 938.
- [95] QCDTARO Collaboration, Nucl. Phys. (Proc. Suppl.) **B63** (1997) 928; hep-lat/9806008.
- [96] P. Hasenfratz and F. Niedermayer, Nucl. Phys. **B414** (1994) 785.
- [97] P. Hasenfratz, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 53.
- [98] R. Swendsen, Phys. Rev. Lett. **47** (1981) 1775.
- [99] T. DeGrand, A. Hasenfratz, P. Hasenfratz, F. Niedermayer, Phys. Lett. **365B** (1996) 233.
- [100] T. DeGrand, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 913, and hep-lat/9802012.
- [101] F. Farchioni and A. Papa, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 910.
- [102] H. Hamber and C.M.Wu, Phys. Lett. **133B** (183) 351.
- [103] R. Wohlert, DESY 87/069; M. Lüscher and P. Weisz, Nucl. Phys. **B479** (1996)

429.

- [104] M. Lüscher, ALPHA Collaboration, Nucl. Phys. **B478** (1996) 365.
- [105] R. Edwards, U. Heller, and T. Klassen, hep-lat/9711052.
- [106] G. Martinelli, C.T. Sachrajda and A. Vladikas, Nucl. Phys. **B358** (1991) 212.
- [107] M. Alford, T. Klassen, and P. Lepage, Nucl. Phys. **B496** (1997) 377.
- [108] P. Lepage and B. Thacker, Phys. Rev. **D43** (1991) 196; G. Lepage, L. Magnea, C. Nakhleh, U. Magnea and K. Hornbostel, Phys. Rev. **D46** (1992) 4052.
- [109] A. Ali Khan, *et.al.*, hep-lat/9801038.
- [110] C.T.H. Davies, *et.al.*, Phys. Rev. **D56** (1997) 2755.
- [111] C.T.H. Davies, *et.al.*, hep-lat/9802024; C.T.H. Davies, Nucl. Phys. (Proc. Suppl.) **B60A** (1998) 124; hep-lat/9705039.
- [112] W. Bietenholz, R. Brower, S. Chandrasekharan, and U.-J. Wiese, Nucl. Phys. (Proc. Suppl.) **B53** (1997) 921.
- [113] A. X. El-Khadra, A. S. Kronfeld, P. B. Mackenzie, Phys. Rev. **D55** (1997) 3933.
- [114] D. Kaplan, Phys. Lett. **288B** (1992) 342; Y. Shamir, Nucl. Phys. **B406** (1993) 90.
- [115] T. Blum and A. Soni, hep-lat/9712004.
- [116] G. 't Hooft, Comments at the 1972 Marseilles Conference, Unpublished; H. D. Politzer, Phys. Rev. Lett. **30** (1973) 1346; D. Gross and F. Wilczek, Phys. Rev. Lett. **30** (1973) 1343.
- [117] W. Celmaster and R. J. Gonsalves, Phys. Rev. **D20** (1979) 1420.
- [118] G.S. Bali and K. Schilling, Phys. Rev. **D47** (1993) 661; G.S. Bali and K. Schilling, Int. J. Mod. Phys. **C4** (1993) 1167.
- [119] G.S. Bali *et al.* (T χ L collab), Nucl. Phys. (Proc. Suppl.) **B63** (1998) 209; G.S. Bali, *et.al.*, (T χ L collab), in preparation; U. Glaässner, *et.al.*, (SESAM Collab.) PLB383 (1997) 98.
- [120] M. Creutz, Phys. Rev. **D21** (1980) 2308.
- [121] H. Wittig, *et.al.*, UKQCD collaboration, Nucl. Phys. (Proc. Suppl.) **B42** (1995) 288.
- [122] M. Guagnelli, R. Sommer, and H. Wittig, hep-lat/9806005.
- [123] A. Yu. Dublin, A. Kaidalov, and Yu. A. Simonov, Phys. Lett. **323B** (1994) 41.
- [124] W. Lucha, F. Schöberl, and D. Gromes, Phys. Rev. **200** (1991) 128.
- [125] L. Burakovsky, hep-ph/9805286.
- [126] S. Veseli and M. Olsson, Phys. Lett. **383B** (1996) 109.
- [127] R. Sommer, Nucl. Phys. **B411** (1994) 839.
- [128] B. Svetitsky and L. Yaffe, Nucl. Phys. **B210** (1982) 423; B. Svetitsky, Phys. Rev. **132** (1986) 1,
- [129] J. Engels, S. Mashkevich, T. Scheideler, and G. Zinovev, Phys. Lett. **365B** (1996) 219; G. Boyd, *et.al.*, Nucl. Phys. **B469** (1996) 419.
- [130] A. Guth, Phys. Rev. **D21** (1980) 2291; J. Frölich and T. Spencer, Comm. Math. Phys. **83** (1982) 411.
- [131] M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rev. **D20** (1979) 1915; T. Degrand and D. Toussaint, Phys. Rev. **D22** (1980) 2478.
- [132] M. Creutz, Phys. Rev. Lett. **43** (1979) 553; M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rev. Lett. **42** (1979) 1390.
- [133] G. Münster and P. Weisz, Phys. Lett. **119B** (1980) 119.
- [134] U. Heller and F. Karsch, Nucl. Phys. **B251** [FS13] (1985) 254; and private communications.
- [135] G. Bhanot and M. Creutz, Phys. Rev. **D24** (1981) 3212.
- [136] R. Gavai, Nucl. Phys. **B474** (1996) 446.

- [137] U. Heller, Phys. Lett. **362B** (1995) 123.
- [138] K. Moriarty, Phys. Lett. **106B** (1981) 130; G.G. Batrouni and B. Svetitsky, Phys. Rev. Lett. **52** (1984) 2205; Y. Makeenko and M. Polikarpov, Nucl. Phys. **B205** (1982) 386.
- [139] C. Morningstar and M. Peardon, in preparation.
- [140] A. Patrascioiu and E. Seiler, Phys. Rev. **D57** (1998) 1394.
- [141] S. Duane, A.D. Kennedy, B. Pendleton, D. Roweth, Phys. Lett. **195B** (1987) 216.
- [142] B. Allès, et.al., hep-lat/9803008.
- [143] S. Aoki, et.al., Phys. Rev. **D50** (1994) 486.
- [144] A. Duncan, E. Eichten, and E. Eichten, Phys. Rev. Lett. **76** (1996) 3894.
- [145] E. Eichten and B. Hill, Phys. Lett. **243B** (1990) 427, Phys. Lett. **240B** (1990) 193; Phys. Lett. **234B** (1990) 511. M. Neubert, Phys. Rev. **245** (1994) 259; G. Martinelli and C. Sachrajda, Nucl. Phys. **B478** (1996) 660; J. Mandula and M. Ogilvie, Phys. Rev. **D57** (1998) 1397.
- [146] M. Lüscher, et.al., ALPHA Collaboration, hep-lat/9611015.
- [147] A. Donini, et.al., Nucl. Phys. (Proc. Suppl.) **B53** (1997) 883; G. Martinelli, et.al., Phys. Lett. **411B** (1997) 141.
- [148] G. M. de Divitiis and R. Petronzio, Phys. Lett. **419B** (1998) 311.
- [149] E. Witten, Nucl. Phys. **B156** (1979) 269; G. Veneziano, Nucl. Phys. **B159** (1979) 213.
- [150] S. Sharpe, Phys. Rev. **D46** (1992) 3146; J. Labrenz and S. Sharpe, Phys. Rev. **D54** (1996) 4595; S. Sharpe and Y. Zhang Phys. Rev. **D53** (1996) 5125.
- [151] C. Bernard and M. Golterman, Phys. Rev. **D46** (1992) 853; M. Golterman, hep-lat/9405002 and hep-lat/9411005.
- [152] J. Gasser and H. Leutwyler, Nucl. Phys. **B250** (1985) 465.
- [153] R. Gupta, Nucl. Phys. (Proc. Suppl.) **B42** (1995) 85;
- [154] S. Sharpe, Nucl. Phys. (Proc. Suppl.) **B53** (1997) 181.
- [155] G. Kilcup, S. Sharpe, R. Gupta, and A. Patel Phys. Rev. Lett. **64** (1990) 25; G. Kilcup, R. Gupta, and S. Sharpe, Phys. Rev. **D57** (1998) 1654.
- [156] T. Bhattacharya, R. Gupta, G. Kilcup, and S. Sharpe, Phys. Rev. **D53** (1996) 6486.
- [157] C. Bernard, T. Draper, A. Soni, D. Politzer and M. Wise, Phys. Rev. **D32** (1985) 2343.
- [158] S. Aoki, et.al., JLQCD Collaboration, Nucl. Phys. (Proc. Suppl.) **B47** (1996) 354.
- [159] P. Lepage, 1989 TASI Lectures, Eds. T. Degrand and D. Toussaint, World Scientific, 1990.
- [160] S. Gottlieb, Nucl. Phys. (Proc. Suppl.) **B53** (1997) 155.
- [161] T. Yoshie, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 3; hep-lat/9711017.
- [162] T. Bhattacharya, et.al., in preparation.
- [163] J. Flynn, Nucl. Phys. (Proc. Suppl.) **B53**, (1997) 168.
- [164] T. Onogi, Nucl. Phys. (Proc. Suppl.) **B63A-C**, (1998) 59; A. Ali Khan, Nucl. Phys. (Proc. Suppl.) **B63A-C**, (1998) 71; Y. Yoshie, Nucl. Phys. (Proc. Suppl.) **B63A-C**, (1998) 3.
- [165] T. Barnett and S. Sharpe, *Ann. Rev. Nuc. Part. Sci.* **40** (1990) 327.
- [166] R. Gupta, A. Patel, C. Baillie, G. Kilcup, S. Sharpe, Phys. Rev. **D43** (1991) 2301.
- [167] C. Michael and M. Teper, Nucl. Phys. **B314** (1989) 347.
- [168] G. Bali and K. Schilling, Phys. Lett. **309B** (1993) 378.
- [169] H. Chen et.al., Nucl. Phys. (Proc. Suppl.) **B34** (1994) 357.
- [170] C. Michael, hep-ph/9710502.

- [171] W. Lee and D. Weingarten, hep-lat/9805029.
- [172] C. Morningstar and M. Peardon, Phys. Rev. **D56** (1997) 4043.
- [173] M. Peardon, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 22.
- [174] P. Lacock, *et.al.*, Phys. Lett. **401B** (1997) 308.
- [175] T. Manke, *etal.*, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 332.
- [176] S. Collins, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 335.
- [177] D. Weingarten, Phys. Rev. Lett. **51** (1983) 1830.
- [178] S. Nussinov, Phys. Rev. Lett. **51** (1983) 2081.
- [179] E. Witten, Phys. Rev. Lett. **51** (1983) 2351.
- [180] C. Vafa and E. Witten, Phys. Rev. Lett. **51** (1983) 2351.
- [181] G. West, Phys. Rev. Lett. **77** (1996) 2622.
- [182] S. Capitani, *et.al.*, Nucl. Phys. (Proc. Suppl.) **B63A-C** (1998) 153.
- [183] R. Gupta and T. Bhattacharya, Phys. Rev. **D55** (1997) 7203.
- [184] B. Gough, *et.al.*, Phys. Rev. Lett. **79** (1997) 1622.
- [185] T. Bhattacharya and R. Gupta, Nucl. Phys. (Proc. Suppl.) **B63A-C** (1998) 95.
- [186] V. Gimenez, *et.al.*, hep-lat/9801028.
- [187] S. Capitani *et.al.*, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 153.
- [188] J.A.M. Vermaseren *et.al.*, Phys. Lett. **405B** (1997) 327.
- [189] A. Buras, hep-ph/9610461.
- [190] A. Buras, M. Jamin, and M. E. Lautenbacher, Phys. Lett. **389B** (1996) 749.
- [191] L. K. Gibbons, *et.al.*, Phys. Rev. Lett. **70** (1993) 1203.
- [192] G. D. Barr, *et.al.*, Phys. Lett. **317B** (1993) 233.
- [193] C.T.H. Davies, *et.al.*, Phys. Rev. Lett. **73** (1994) 2654.
- [194] V. Gimenez, G. Martinelli, and C. Sachrajda, Phys. Lett. **393B** (1997) 124.
- [195] C. Allton, *et.al.*, Nucl. Phys. **B431** (1994) 667.
- [196] A. El-Khadra, hep-lat/9509381.
- [197] A. Kronfeld, Nucl. Phys. (Proc. Suppl.) **B63** (1998) 311.
- [198] A. Bochkarev, P. de Forcrand, Nucl. Phys. **B477** (1996) 489; Nucl. Phys. (Proc. Suppl.) **B53** (1997) 305.