ASPECTS OF HADRON AND INSTANTON PHYSICS IN LATTICE QUANTUM FIELD THEORIES

by

ANDREW POCHINSKY

SUBMITTED TO THE DEPARTMENT OF PHYSICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1997

Author ....
Department of Physics
December 3, 1996

Certified by ....
John William Negele
William A. Coolidge Professor of Physics
Thesis Supervisor

Accepted by ....
George Koster
Chairman, Departmental Committee on Graduate Students

Aspects of Hadron and Instanton Physics in Lattice Quantum Field Theories

by

Andrew Pochinsky

Submitted to the Department of Physics
on December 3, 1996, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

An ongoing challenge to quantum chromodynamics as the theory of strong interactions is calculating hadron masses and matrix elements from first principles. Presently lattice calculations are the most promising means to probe low energy physics of quarks and gluons. The matrix element of the polarized on-shell nucleon state $\langle ps|TJ_\mu(x)J_\nu(0)|ps\rangle$ can be reduced to a set of spin-independent longitudinal and transverse structure functions $F_i(x,Q^2)$ and spin-dependent functions $g_i(x,Q^2)$ and $h_i(x,Q^2)$. Relevant matrix elements are calculated on a large lattice in the quenched approximation. In particular, the zeroth moment of the tensor charge is calculated for light valence quarks and extrapolated to the chiral limit.

Topological excitations play an important rôle in nonperturbative quantum field theory. An introduction of the θ-term into the Lagrangian calls for special simulation techniques and sampling methods and requires a tremendous increase in statistics to get a signal. While QCD is still beyond current computational capabilities, investigations of simpler models will gain better understanding of topology related issues in lattice quantum field theories. The two dimensional $O(3)$ σ-model with the θ-term is studied in the second part of the thesis. Using the cluster update algorithms and improved estimators a numerical check of Haldane's conjecture is performed. A special updating technique has been developed to construct an improved estimator for the topological charge and other observables to overcome the sign problem.

Thesis Supervisor: John William Negele
Title: William A. Coolidge Professor of Physics

Acknowledgments

This work would not be possible without many people around the globe. My parents gave all the encouragement one can hope for and their unfailing understanding for my interests has been a great experience.

My physics teachers, Karen A. Ter-Martirosian, Yuri A. Simonov and Michail I. Polikarpov of ITEP made the first steps of my journey into the space of physical theories unforgettable. The dedication to physics and integrity they have shown throughout the years will always remain an example impossible to excel.

The interactions with my collaborators, Wolfgang Bietenholz, Richard C. Brower, Suzhou Huang, John W. Negele, Bernd Schreiber and Uwe-Jens Wiese were always stimulating and fulfilling. We shared the many joys and frustrations which make the research interesting.

I owe an especial debt of gratitude to John W. Negele for inviting me to MIT for graduate studies and being the most supportive thesis adviser.

Greg Papadopoulos, currently of Sun Microsystems Inc., provided one of their computers, which was instrumental for getting high statistics results for chapter 5.

The TEX system by Donald E. Knuth has proven to be the most demanding and exacting editor ever met, human or otherwise.

Of course, all errors and omissions are completely my own.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this thesis two major topics are described. The first deals with calculation of the hadron structure functions from the first principles. In the second part we develop a novel method of studying topology on the lattice.

Chapter 2 contains an introduction to common lattice techniques. The methodology of constructing a nucleon on the lattice is also considered there. It ends with establishing the physical scale of the lattice used for the moments calculations in the next chapter.

Twenty years of experimental high energy probes have provided detailed measurements of spin-dependent and spin-independent hadron structure functions characterizing the distribution of quarks and gluons in the nucleon. Presently, although there is no known way to solve quantum chromodynamics to calculate the structure functions directly from first principles, it is possible to calculate the moments of the structure functions using lattice QCD. The quenched calculations of moments of spin-dependent and spin-independent structure functions are considered in chapter 3. The theoretical basis and details of the lattice calculations are described and numerical results for a exploratory calculation are presented.

Since the dynamics of QCD is governed by nonperturbative effects, it is of paramount importance to understand the rôle topological objects play in formation of nucleons. An introduction of the θ-term into the Lagrangian calls for special simulation and sampling methods and requires a tremendous increase in statistics to extract a

signal. While QCD topological effects can not be fully handled by current computational methods, investigation of simpler models is important for better understanding of topology related issues in lattice quantum field theories. Difficulties of previous approaches are reviewed in chapter 4.

Studying the two-dimensional σ-model is traditionally a warm-up exercise for nonperturbative QCD. At the same time, it has another application to the behavior of one-dimensional quantum spin chains. Therefore, we develop a method to simulate the 2-d σ-model with the θ-term present on a computer in chapter 5. As an application of the technique, we study the mass gap behavior at $\theta = \pi$.

There are also three appendices in the thesis. In appendix A a portable random number generator is developed for MPP architecture. This implementation has been used in the gauge field generation for chapters 2 and 3. The jackknife procedure for error estimate is summarized in appendix B. Appendix C describes how the plateaux in the experimental data can be determined optimally.

Chapter 2

Lattice Review

A dream of understanding the properties of strong interactions from first principles is now almost fifty years old. Following the development of quantum chromodynamics in the early 1970's, QCD based calculations of the masses and other properties of hadrons were made possible by Wilson's work on lattice gauge field theory and renormalization group methods. Lattice calculations became a serious player in hadron physics around 1980 with introduction of Monte-Carlo techniques. Since that time, the lattice made its way to the particle physics community, e.g., the Particle Data Book [1] now cites lattice results for α_s and the expected glueball mass. Predictions that the 0^{++} state is the lightest are now widely believed.

In this chapter we start by reviewing basic lattice concepts and techniques. Section 2.1 gives the standard set of arguments for using lattice simulations to extract nonperturbative results for quantum field theory and introduces the lattice notation. In section 2.2 we construct both gauge and fermion actions suitable for lattice calculations. We consider implications of *quantum* aspects of the theory in section 2.3. Section 2.4 explains how the gauge conditions are implemented on the lattice. The conjugate gradient method of inverting the Dirac matrix is defined in section 2.5. After that, we show how to construct different hadron sources and discuss their merits in section 2.6. In the last section, 2.7, we establish the scale of the lattice used in the next chapter for calculation of the hadron structure function moments.

2.1 Why Lattice?

Lattice gauge theory goes back to early 70's, when Wilson [2] formulated the lattice theory for regularization purposes. Since that time, lattice calculations developed as a major player in nonperturbative field theory. Currently lattice QCD is widely used to obtain first-principles information about confinement, the hadron spectrum, electroweek decay constants, heavy quark physics etc. In fact, the range of applications is so broad that the proceedings of annual lattice conference is well above 500 pages.

The beauty of lattice field theory is that it allows to study nonperturbative effects while not imposing any *ad hoc* models. The ambition is to solve QCD from the first principles.

There is yet another appeal to study lattice theories. One can construct and study models which are difficult or impossible to implement experimentally. Besides the Ising model, the 2-d σ-model is the most common test model studied by lattice theorists. We will use it in chapter 5 to study θ-vacua.

What all lattice theories share is the departure from continuous 4-dimensional Minkowski space-time we happen to live in. The first steps are to perform a Wick rotation and to replace the flat space-time $\mathbf{R}^4$ by a manifold $\mathcal{M}$, usually a torus T^4. The next step is to substitute a discrete set of points for $\mathcal{M}$.

At this point it is convenient to introduce a notation similar to differential geometry. This notation works for both finite lattices which can be simulated on a computer and infinite lattices which are useful for analytical calculations.

One way to construct the lattice is to start by dividing a d-dimensional manifold $\mathcal{M}$ into N d-dimensional cells $c_d(n)$ without common internal points. This bisection completely defines the lattice as we shall see presently. If all cells are isomorphic then the lattice is regular, if the cells are of random form, then one has a random lattice. The lower dimensional structures are defined recursively for $k = 1 \ldots d$. A pair of k-cells intersecting over a $(k-1)$-dimensional solid define a $(k-1)$-cell: $c_{k-1}(n,m) = c_k(n) \cap c_k(m)$. If one goes on recursively, the last two steps will be 1-cells c_1 (links) and 0-cells c_0 (vertices). The collection of all cells forms the lattice

$\mathcal{L} = \{\{c_n\}, \{c_{n-1}\}, \ldots, \{c_1\}, \{c_0\}\}$ We define a *dual* d-cell $c_d^*(x)$ as a set of points in $\mathcal{M}$ which are closer to a given $c_0(x)$ than to any other c_0. Repeating the previous procedure one builds the dual lattice $\mathcal{L}^* = \left\{\{c_n^*\}, \{c_{n-1}^*\}, \ldots, \{c_1^*\}, \{c_0^*\}\right\}$. This defines a duality transformation: $(c_k)^* = c_{d-k}^*$. It is easy to see that $((c_k(x))^*)^* = c_k(x)$, so that $\mathcal{L}^{**} = \mathcal{L}$.

If the manifold $\mathcal{M}$ is orientable, one can introduce the *orientation* on $\mathcal{L}$.

The *boundary* operator d maps a k-cell into an oriented collection of $(k-1)$-cells forming the boundary of $c_k(x)$. E.g., $dc_1 = c_0(a) - c_0(b)$. Analogously, the *coboundary* operator maps a k-cell into a collection of oriented $(k+1)$-cells: $\partial c_k(x) \equiv (d((c_k(x))^*))^*$.

This language closely follows notation of differential geometry (see, e.g, [3]) and shares indeed many advantages of the latter.

One of many choices is to use a torus $T^4 = \mathbf{R}^4/\mathbf{Z}^4$ for $\mathcal{M}$ and hypercubes for all c_d. This way we get the conventional QCD lattice. Another discretization will be used in chapter 5.

The matter fields (scalars, fermions, etc.) live on c_0: $\psi_{\text{cont}}(x) \to \psi_{\text{lat}}(c_0)$. From the differential geometry point of view, the gauge field is a connection, hence it lives on c_1 defining a parallel transport of the fundamental matter fields along the c_1: $A_\mu(x) \to U(c_1)$. In addition, while the continuous gauge field A_μ was an element of the Lie algebra, the finite transport U is an element of a corresponding Lie group[1].

2.2 Lattice QCD Action

In this section we define the lattice QCD action for both quarks and gluons. From now until chapter 4 we shall work exclusively with 4-dimensional hypercube lattice. It is convenient to introduce another notation for the links and label the link $c_1(n_1, n_2, n_3)$ by the point of origin n and the direction μ: $c_1 \to (n, \mu)$. Then the gauge field depends on these two labels: $U = U(n, \mu)$. The free case $A_\mu = 0$ corresponds to $U = 1$. Occasionally we will label U by a two c_0s: $U(n, \mu) \equiv U(n, n + \hat{\mu})$.

[1] Conventionally the universal covering group is used, but other choices are possible, e.g., [4, 5, 6, 7]

The link value can be expressed through A_μ as follows:

$$U(n,\mu) = \mathrm{Pexp}\int_n^{n+\hat{\mu}} igA_\nu(x)dx^\nu \xrightarrow{a\to 0} 1 + igaA_\mu(n) + O(a^2).$$

2.2.1 Gauge Action

To construct a lattice version of the minimal gauge action F^2, we will need a lattice analog to the field strength tensor:

$$P_{\mu\nu}(n) = U(n, n+\hat{\mu})\,U(n+\hat{\mu}, n+\hat{\mu}+\hat{\nu})\,U(n+\hat{\mu}+\hat{\nu}, n+\hat{\nu})\,U(n+\hat{\nu}, n),$$

(this comes from the definition of the curvature tensor on the principal bundle, see Fig. 2-1.)

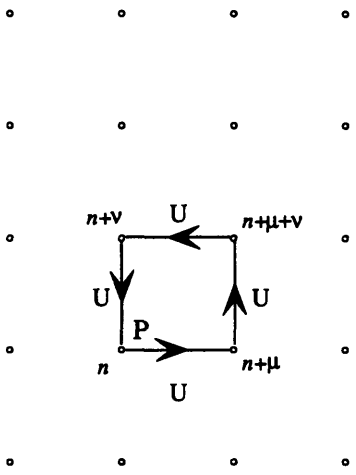

Figure 2-1: Links and a plaquette

In the continuum limit the plaquette becomes:

$$P_{\mu\nu}(n) \xrightarrow{a\to 0} 1 + iga^2F_{\mu\nu}(n) - g^2a^4F_{\mu\nu}(n)F_{\mu\nu}(n) + O(a^6), \tag{2.1}$$

(There is no summation over μ and ν on the right hand side). One immediately sees that

$$\frac{P_{\mu\nu}(n) - 1}{iga^2} \xrightarrow{a\to 0} F_{\mu\nu}(n).$$

Now we need to construct a lattice expression which will reproduce the Lagrange density $F_{\mu\nu}^2$ in the continuum. As usual on the lattice, there are many different ways

to achieve this goal. E.g., $-1/g^2 \mathrm{ReTr} \sum_{\mu\nu}(P_{\mu\nu} - 1)^2$ is a candidate. However, the definition of choice of the lattice community is

$$S_{\mathrm{glue}} = \frac{1}{N_c a^4 g^2} \sum_{\{c_2\}} \mathrm{ReTr}\left(1 - P_{\mu\nu}(n)\right). \tag{2.2}$$

One reason for that will be clear when we discuss lattice simulation techniques. This expression has $O(a^2)$ corrections in the continuum limit. Currently there is a lot of activity on improving the gauge action. (In addition to the perturbative approach of Symanzik [8, 9] there are important recent developments based on the renormalization group [10, 11, 12].)

2.2.2 Fermion Action

For the fermions the situation is more subtle. One can easily write down an expression with a correct continuum limit, e.g.,

$$S_{\mathrm{ferm}}^{\mathrm{naive}} = \frac{1}{2} \sum_{c_1} \left\{ \bar{\psi}_n \gamma_\mu U(n, n+\hat{\mu}) \psi_{n+\hat{\mu}} - \bar{\psi}_n \gamma_\mu U(n, n-\hat{\mu}) \psi_{n-\hat{\mu}} \right\} - am \sum_{c_0} \bar{\psi}_n \psi_n$$

but there is a notorious fermion doubling problem [13, 14], with the effect that this Lagrangian describes fermions with a wrong number of degrees of freedom. It has been proven that the doubling problem can not be eliminated if only interactions with a finite number of neighbors are considered [15] and both the chiral symmetry and hermiticity are kept intact. In d dimensions the number of doublers is 2^d.

Many different approaches were suggested to combat this problem [2, 16]. We will use so-called Wilson fermions:

$$S_{\mathrm{ferm}}^{W} = \sum_x \bar{\psi}_x \psi_x - \kappa \sum_{\mu,x} \left[\bar{\psi}_x \left(R - \gamma_\mu\right) U_{\mu,x} \psi_{x+\hat{\mu}} + \bar{\psi}_x \left(R + \gamma_\mu\right) U^\dagger_{\mu,x-\hat{\mu}} \psi_{x-\hat{\mu}} \right] \tag{2.3}$$

For $R \neq 0$, extra degrees of freedom acquire additional mass and are hence pushed out of the massless region. We will use $R = 1$, because this gives the theory some nice properties in the transfer matrix formalism and considerably simplifies computations. The hopping parameter κ replaces the continuum mass. In the free case, $U = 1$, the Lagrangian (2.3) describes fermions with mass

$$am_0 = \frac{1 - 2d\kappa}{2\kappa}.$$

When the gauge interaction is present, this relation is replaced by

$$m_q = \frac{Z_q}{2a}\left(\frac{1}{\kappa} - \frac{1}{\kappa_c}\right). \tag{2.4}$$

The renormalization factor Z_q depend on the form of the action. For the Wilson action at $\beta = 6.2$ its perturbative value is close to 1.12 [17, 18]. κ_c is the value of the hopping parameter where fermions becomes massless. Its value depends on the dynamics of the gluon sector. The bad news about the Wilson action (2.3) is that in the continuum limit, $a \to 0$, $R = \text{const}$, it has order a corrections. This results in the relatively strong κ-dependence of the observables and increases statistics needed for the same accuracy of the extrapolation to the chiral limit compared to the staggered fermions. Furthermore, since there is no remnant on chiral symmetry. the mass is protected against renormalization an the chiral limit requires fin tuning. However, the convenience of the action (2.3) for numerical simulations often outweights its drawbacks. It also simplifies calculating the renormalization constants.

In contrast, although staggered fermions maintain a remnant of chiral symmetry, they only partially remove the doubling problem and strictly apply to integer multiples of 4 flavors. In addition, because of thinning of degrees of freedom, the effective lattice spacing for staggered fermions is actually $2a$.

In the next chapter we will use lattice covariant derivatives for fermions:

$$(\bar{\psi}\ \overleftrightarrow{D}_\mu\ \psi)_n = (\bar{\psi}\ \vec{D}_\mu\ \psi)_n + (\bar{\psi}\ \overleftarrow{D}_\mu\ \psi)_n, \tag{2.5}$$

where

$$(\bar{\psi}\ \vec{D}\ \psi)_n = \bar{\psi}_n(U_{n,\mu}\psi_{n+\hat{\mu}} - \psi_n) \tag{2.6}$$

and

$$(\bar{\psi}\ \overleftarrow{D}\ \psi)_n = (\bar{\psi}_{n-\hat{\mu}}U^\dagger_{n-\hat{\mu},\mu} - \bar{\psi}_n)\psi_n. \tag{2.7}$$

2.3 What Makes It Tick

Now, once the action is defined in both fermion and gauge sectors, we can proceed to building the quantum theory on the lattice.

In general, the way to evaluate the functional integral $Z = \int [d\phi] e^{-S[\phi]}$ is to generate an ensemble $\mathcal{E}$ of points $\{\phi\}$ in configuration space distributed according to the weight $p(\phi) \sim e^{-S[\phi]}$ and use the estimator

$$\sum_{\phi \in \mathcal{E}} A[\phi] \xrightarrow{a \to 0} \frac{\int [d\phi] A[\phi] e^{-S[\phi]}}{Z},$$

for the observable $A[\phi]$.

In some cases, it is possible to construct an algorithm producing field configurations that are representative of a large number of points in the configuration space so that a part of the sum on the left hand side of the above equation can be done analytically. Such cluster algorithms have been constructed for spin models and generalized to the σ-model (we will use cluster updates in chapter 5), but presently there is no efficient construction known for the $SU(N)$ gauge theory.

It is very important that the ensemble $\mathcal{E}$ consists of statistically independent configurations and covers the configuration space completely. One method to generate such an ensemble is to use a Markov chain satisfying the *detailed balance principle*: If C and C' are two configurations with actions $S(C)$ and $S(C')$ respectively, then the probabilities to move from C to C' and back ($P(C', C)$ and $P(C', C)$ respectively) must satisfy the condition

$$P(C', C) e^{-S(C)} = P(C, C') e^{-S(C')}. \tag{2.8}$$

In addition, we require that $P(C, C') > 0$ for all C, C'. This ensures that any point in the configuration space can be reached by the random walk. Since the walk is defined as a discrete chain of field configurations, it is not required that the configuration space is connected in order for all topological sectors to be automatically sampled in $\mathcal{E}$ with a correct weight.

If we were able to use the whole Markov chain for the estimators, it would be the end of the story. Unfortunately, computers produce only finite sequences of configurations. Thus, the question of statistical independence of the configurations needs to be addressed. Obviously, if two configurations C and C' differ in only small number of variables, then physical observables will be strongly correlated.

So far, the Markov chain construction does not impose statistical independence on consecutive fields. If the correlation time τ of the simulation algorithm is known, then one can build an ensemble of statistically independent configurations by picking up steps from the chain separated by at least τ iterations.

One commonly used update algorithm is the heatbath [19] which can be efficiently implemented for the $SU(2)$ Yang-Mills gauge field. Its efficiency is based on the fact that the group manifold is a three dimensional sphere S^3 so that a new value of the link can be generated efficiently with the sharp peaked probability $\exp(-\beta S(U))dU$. Since for $SU(N)$, $N \geq 3$ the group manifold is not a sphere any longer and as a result, a straightforward implementation of the heatbath must deal with a sharp peak in the probability distribution. Cabibbo and Marinari [20] suggested a way around this difficulty. Their idea is to use the heatbath for updates in $SU(2)$ subgroups of $SU(N)$. The suggested algorithm automatically satisfies the detailed balance principle (2.8). While it is enough to update only two subgroups $(0,1)$ and $(1,2)$ to cover the whole $SU(3)$, the autocorrelation time is significantly reduced if the third subgroup $(1,3)$ is also updated.

Overrelaxation is another method widely used for gauge field generation [21, 22]. It can be considered as a special case of the heatbath algorithm when a new configuration has exactly the same action as the old one. The Wilson action (2.2) considered as a function of one $U_{n,\mu}$ only can be written as $C_1 + C_2 \mathrm{ReTr}(UA)$ where A is a $N_c \times N_c$ matrix and C_i are some constants. Again, the $SU(2)$ case is special: $A = kB$, $B \in SU(2)$ and, e.g., $U- > B^\dagger U^\dagger B^\dagger$ preserves the action (2.2). If $N_c > 2$ the same Cabibbo-Marinari trick could be applied as for the heatbath.

Simulating the fermion sector of the theory is a separate topic which we shall not discuss here since there are no dynamical fermions in the quenched approximation.

2.3.1 Details of the Field Generation

For calculation of the structure function moments in chapter 3 we generated $SU(3)$ quenched configurations using the Wilson action with 3-subgroup Cabibbo-Marinari interlaced with 16 overrelaxation sweeps on a $24^3 \times 32$ lattice. $SU(2)$ subgroups were

chosen in order (01), (12), (02). First the heatbath algorithm was applied to each subgroup, then 16 overrelaxation sweeps were performed. Each overrelaxation sweep consisted of sequential updates in the same subgroups. Relative numbers of heatbath and overrelaxation iteration were selected based on a tradeoff between autocorrelation time (in iterations) and run-time of the algorithm. While comprehensive studies of the autocorrelation time is prohibitively expensive, results indicate that using every 50th iteration for inverting the Dirac matrix and calculating all the observables introduces reasonably small statistical errors due to correlations between the gauge configurations. Moreover, these errors are completely overshadowed by other sources of noise.

The generation started from the cold start ($U = 1$) and first 7000 iterations were discarded to allow for system thermalization. The coupling constant was held at $\beta = 6.2$ throughout the simulation.

2.4 Gauge Fixing

In our calculations sometimes[2] we need to fix the gauge by imposing some gauge condition $G(U) = 0$. On the lattice it amounts to finding an extremum of some functional $F[U]$ with respect to gauge transformations $g : U(n, m) \to g(n)U(n, m)g^\dagger(m)$. The extremum condition for the functional F must reproduce the gauge condition G. Otherwise we are free to use any suitable functional.

2.4.1 Landau Gauge

The Landau gauge $\partial_\mu A^\mu = 0$ corresponds to the functional

$$F_L[U] = \sum_n \sum_{\mu=0}^{3} \left(U_\mu(n) + U_\mu^\dagger(n) \right)$$

[2]Currently we do the gauge fixing for distributed sources only, but it is not difficult to imagine some Dirac matrix inversion method whose convergence would from the gauge fixing also. One such example is Fourier acceleration, for which a smooth gauge like Landau gauge is desirable.

Finding an extremum of $F_L[U]$ requires an iterative procedure, e.g., one may sweep through the lattice maximizing $F_L[U]$ with respect to the local gauge transformations. Here again it is advantageous to use the Cabibbo-Marinari trick, since the maximum in $SU(2)$ case can be found by solving an algebraic equation.

2.4.2 Coulomb Gauge

The Coulomb gauge $\partial_i A^i = 0$ corresponds to the functional

$$F_G[U] = \sum_n \sum_{i=1}^{3} \left(U_i(n) + U_i^\dagger(n) \right). \tag{2.9}$$

The difference between Landau and Coulomb gauges is in the range of the internal sum over directions. In the Coulomb case it runs over spacial directions only, so that (2.9) admits gauge transformations which depend only on time. It can also be fixed using a procedure similar to that for Landau gauge. The difference from Landau gauge allows to fix the Coulomb gauge in the time slices of interest only. In our case it is enough to fix the Coulomb gauge on the source and the sink time slices instead of fixing it on the whole lattice. We did not use this miniscule optimization for two reasons. First, the time spent in gauge fixing amounts to about 2% of the full computations, and, second, it is simpler to fix the whole lattice once and for all instead of worrying about gauge transforming the propagator when building the two-point function (see section 2.7) and sequential source (section 3.6).

We use the Coulomb gauge when constructing the smeared nucleon sources in section 2.6.

2.4.3 Computation Details

Because of the nature of the functional (2.9) an iterative procedure is needed to find its minimum. A sweep through the lattice consists of changing every link $U_{n,\mu}$ in such a way that its contribution to (2.9) is minimized.

Applying a gauge transformation at the site n only, the change in (2.9) is

$$\delta F_G[U] = F_G[U^g] - F_G[U] = \sum_{i=1}^{3} \left(g_n U_i(n) + U_i^\dagger(n - \hat{i}) g_n^\dagger \right) - \sum_{i=1}^{3} \left(U_i(n) + U_i^\dagger(n - \hat{i}) \right).$$

Since δF_G is linear in g, its minimum can be easily found by cooling methods. In $SU(2)$ case the group manifold is a sphere and the minimum can be found exactly by solving a linear equation. In case $N > 2$ the Cabibbo-Marinari procedure helps again. We used 3 $SU(2)$ subgroups.

Once the gauge transformation minimizing F_G locally is found, the same procedure can be repeated on the next site until the entire lattice is swept. Because of the structure of F_G one needs to perform multiple sweeps through the lattice before the gauge can be considered fixed.

One can also combine the gauge fixing with overrelaxation methods [23, 24]. We did not do it for the present calculations though.

The overall change in F_G after one sweep determines how close to the fixed gauge the configuration is brought. Figure 2-2 shows typical behavior of F_G as a function of

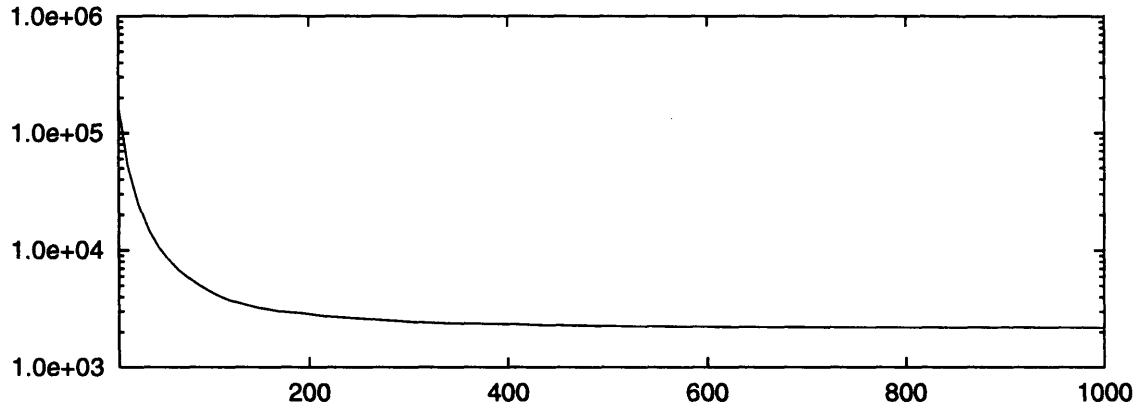

Figure 2-2: Typical gauge fixing functional vs. the iteration number

the number of iterations. In hadron structure function calculations we fix the gauge before inverting the fermion matrix. The empirical tests show that using 1000 gauge fixing sweeps bring us to the region where results do not noticeably depend on further increasing the number of iterations. Though all observables are gauge invariant, the distributed sources we are using (see section 2.6) are defined in the Coulomb gauge, so ultimately the results do depend on the gauge. One really needs to check the behavior of observables conclude that the gauge has been fixed to the acceptable level, since the functional itself changes very little from iteration to iteration. Notice the logarithmic scale on figure 2-2.

2.5 Solving the Dirac Equation

Another piece of machinery we need is the fermion propagator in a given gauge field background. In case of the Wilson action (2.3) this amounts to finding an inverse of the matrix:

$$M_{ab}^{\alpha\beta}(x,y) = \delta_{xy}\delta_{\alpha\beta}\delta_{ab} - \kappa\sum_{\mu}\begin{array}{l}\Big\{\delta_{x+\hat{\mu},y}\left(\delta_{\alpha\beta}-\gamma_{\alpha\beta}^{\mu}\right)U_{ab}(x,\mu)+ \\ \delta_{x-\hat{\mu},y}\left(\delta_{\alpha\beta}+\gamma_{\alpha\beta}^{\mu}\right)U_{ab}^{\dagger}(x-\mu,\mu)\Big\}\end{array} \tag{2.10}$$

To shorten the notation we will write M^{ij} instead of $M_{ab}^{\mu\nu}(x,y)$ hereafter. Though the matrix M is sparse, with $O(N)$ nonzero elements on a lattice of N sites, its inverse is not. Fortunately, as a rule we not need all $N \times N$ elements of M^{-1}, instead we are mostly interested in solving the equation

$$M^{ij}S_j = \eta^i \tag{2.11}$$

for a given right hand side η. Before trying to solve this equation, some remarks are due.

First, one notices that (2.10) builds $(M\psi)_n$ from its nearest neighbors only. Hence, if we divide the lattice into even sites ($n_1+n_2+n_3+n_4 = 0 \bmod 2$) and odd sites ($n_1+n_2+n_3+n_4 = 1 \bmod 2$), then one can rewrite (2.11) as

$$\begin{pmatrix} 1 & \kappa K_{eo} \\ \kappa K_{oe} & 1 \end{pmatrix}\begin{pmatrix} \psi_e \\ \psi_o \end{pmatrix} = \begin{pmatrix} \eta_e \\ \eta_o \end{pmatrix}.$$

This system can be immediately solved for one component of ψ. E.g., $\psi_o = \eta_o - \kappa K_{oe}\psi_e$. Then for ψ_e one has:

$$M_{ee}\psi_e = \eta'_e, \qquad M_{ee} = 1-\kappa^2 K_{eo}K_{oe}, \quad \eta'_e = \eta_e - \kappa K_{eo}\eta_o. \tag{2.12}$$

After that we apply the conjugate gradient inverter [25]. Note that since the conjugate gradient requires a Hermitian matrix, we actually solve the equation $M^{\dagger}M\psi = M^{\dagger}\eta$. Table 2.5 summarizes the standard CG algorithm for this case.

$\phi_0 = 0$, $\eta' = \eta_e + \kappa K_{eo}\eta_o$, $r_0 = \eta' - M\phi_0$, $p_0 = M^\dagger r_0$

Repeat until $|r_k|$ is small enough:

$$a_k = \frac{|M^\dagger r_k|^2}{|Mp_k|^2}$$

$$\phi_{k+1} = \phi_k + a_k p_k$$

$$r_{k+1} = r_k - a_k M p_k$$

$$b_k = \frac{|M^\dagger r_{k+1}|^2}{|M^\dagger r_k|^2}$$

$$p_{k+1} = M^\dagger r_{k+1} + b_k p_k$$

$\psi_e = \phi_{k+1}$, $\psi_o = \eta_o - \kappa K_{oe}\phi_{k+1}$

Table 2.1: Conjugate Gradient Inverter

2.6 Sources

Once we have a gauge field configuration we can inhabit it with all kinds of hadrons. Ideally, one would use a creation operator for a given particle. However, the situation is not that simple; since our building blocks are quarks, the detailed knowledge of the quark distribution is needed to create a pure hadron state. By itself this is a problem at least as complicated as, e.g., that of measuring structure functions. What we do instead is construct a source which has large enough overlap with the state of interest and propagate sufficiently far in imaginary time to project onto the ground state. If the source has some quantum numbers fixed (e.g., spin and parity) and the state N we are after is the lowest state with these quantum numbers, then in a few lattice steps all excited states will die off and one can work with N on the rest of the lattice. Of course, at large separation there are considerable fluctuations which make it difficult to pick up the signal from the noise, so in practice the usable region is somewhat limited. The larger overlap of the source with N the better, because it decreases the amplitudes of the higher excitations.

Though the true hadron wave function does not factor into a product of the valence quark wave functions, such a decomposition for the source has several advantages.

First, it is easy to construct a source with fixed spin, parity and isospin. Second, one can get considerable overlap of a simple source with the lowest state in a sector with fixed (S, P, I).

We start with creating pseudo-scalar mesons. Besides being the simplest color singlets, they are instrumental in determining the lattice scale (see section 2.7.) The following source has $I^G(J^{PC}) = 1^-(0^{-+})$:

$$\bar{J}^{(\pi)}(x) = \bar{\psi}_{a\alpha}(q,x)\gamma_5^{\alpha\beta}\psi_\beta^a(q,x).$$

Here $\bar{\psi}(q,x)$ creates a single quark. We study various choices of ψ in the next section.

For vector mesons, $I^G(J^{PC}) = 1^+(1^{--})$ one can use

$$\bar{J}_\mu^{(\rho)}(x) = \bar{\psi}_{a\alpha}(q,x)\gamma_\mu^{\alpha\beta}\psi_\beta^a(q,x).$$

In the baryon sector several sources are widely used. We will use the following lattice operator to create a nucleon

$$\bar{J}_\alpha^{(N)}(x) = \bar{\psi}_\alpha^a(q_1,x)\bar{\psi}_\beta^b(q_1,x)\,(C\gamma_5)^{\beta\delta}\,\bar{\psi}_\delta^c(q_2,x)\epsilon_{abc}.$$

One can easily check that it has $I(J^P) = 1/2(1/2^+)$.

In the above formulae, ψ can be a local quark source, $\psi_P(q,x) = q(x)$, or some kind of a smeared distribution which in general can be written as

$$\psi_\alpha^a(q,x) = \int d^3y f_{\alpha\beta}^{ab}(x,y)q_\beta^b(y). \tag{2.13}$$

Below we consider relative merits of several $f(y,x)$. Section 2.7 shows the relation of $f(x,y)$ to the right hand side of eq. (2.11).

- **Point Source** The quark fields are combined pointwise to get the hadron quantum numbers. Being a δ-function in space, this operator has extremely large overlap with higher excitations. Depending on the sector, its overlap with the lightest state could become rather small with exited states dominating most of the statistically useful region. This behavior tends to worsen as $a \to 0$.

- **Wuppertal Source** Though this source can be written in form (2.13), it is much more clear to follow the original notation [26]. Using the hopping matrix

$$H(x,x') = \sum_{i=1}^{3} \left(U_i(x)\delta_{x',x+\hat{\imath}} + U_i^{\dagger}(x-\hat{\imath})\delta_{x',x-\hat{\imath}} \right),$$

one defines $\psi(q,x) = \sum_{x'}(1+\alpha H(x,x'))^N q(x')$, where $q(x)$ is the quark creation operator. This source is manifestly gauge invariant. The smearing is controlled by two parameters (α, N).

- **Gaussian Source** In the Coloumb gauge we define the smeared source with $f^{ab}_{\alpha\beta}(x,y) = \delta^{ab}\delta_{\alpha\beta}\exp(-\rho(x-y)^2)$. Here ρ controls the spatial distribution.

- $U = 1$ **Wuppertal Source** can be built applying the technique of the Wuppertal smearing with $U = 1$ and using the resulting distribution for $f(x,y)$ in the Coloumb gauge. This allows us to compare gauge invariant and gauge fixed sources with the same spatial probability distribution for fermions.

2.6.1 Source Comparison

To determine the most suitable form of the source, we investigated the plateau in the effective mass $\ln(G(t)/G(t+1))$ for the two point functions for the pion, rho and nucleon sources. To make comparison of different sources meaningful, we used the RMS radius

$$\langle x^2 \rangle = \frac{\int d^3x x^2 (\psi^a_\mu(x))^* \psi^a_\mu(x)}{\int d^3x (\psi^a_\mu(x))^* \psi^a_\mu(x)}$$

as a quantitative measure of smearing.

Some comparisons are shown for the nucleon case in Fig. 2-3 for 7 configurations at $\kappa = 0.1519$. The Gaussian (G) and the two Wuppertal (W & U) sources were adjusted so $\sqrt{\langle x^2 \rangle} \approx 6.7a \sim 0.47\text{fm}$ for each quark field, since this smearing produced the least noisy results in all three cases. For more localized sources the excited states are more prominent, whereas for less localized sources the signal becomes noisier at large distances.

As seen in Fig. 2-3, smearing both the sources and the sink results in substantially noisier behavior than smearing only the source. On the scale of the errors in Fig. 2-3,

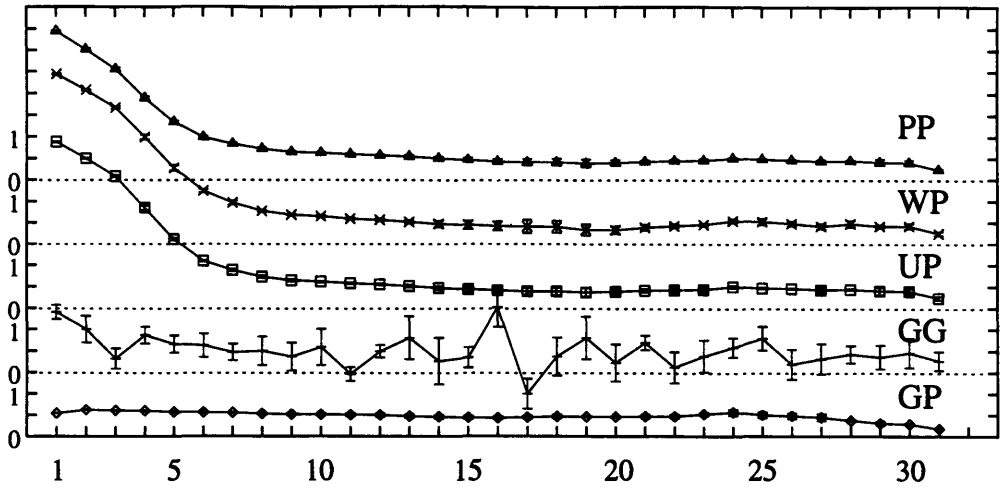

Figure 2-3: Effective mass as a function of distance for the point–point (PP), Wuppertal-point (WP), gauge fixed Wuppertal-point (UP), Gaussian–Gaussian (GG) and Gaussian–point (GP) nucleon correlators

there is no significant difference between smeared source–point sink vs. point source–smeared sink. It is interesting to note that the gauge fixed Wuppertal (U) and gauge invariant Wuppertal (W) sources are essentially equivalent.

The meson sources show similar behavior.

2.7 Lattice Scale

There are two reasons to calculate two-point functions on the lattice. First, it lets us establish the lattice scale from a hadron mass, and we will use the ρ mass. Second, we will use the two-point function on a given separation to normalize the moments of structure functions from three-point function in chapter 3.

We consider the following two-point function projected onto momentum p:

$$D_2(t,p) = \int d^3x e^{ipx} \langle T[J(t,x)\bar{J}(0)] \rangle,$$

where $\bar{J}$ is a nucleon creation operator.

If $\bar{J}$ is sufficiently far from the lattice boundary, $D_2(t)$ can be written as

$$D_2(t,p) = \sum_n |\langle 0|J(0)|n\rangle|^2 e^{-E_n(p)t},$$

after using the Euclidean translation $J(t,x) = \exp(Ht - ipx)J(0)\exp(-Hp + ipx)$ and an insertion of a complete set of states $1 = \sum_n |n\rangle\langle n|$. This relates the 2-point function to the mass spectrum in the corresponding sector.

One can substitute an explicit form for the hadron creation operator, and, after contractions one gets for the nucleon

$$\begin{aligned} D_2^{\alpha\alpha'}(t,p) = \int \; d^3x e^{ipx}(\; & \tilde{S}^{aa'}_{\alpha\alpha'}(x,t;0,t_0)\tilde{S}^{bb'}_{\beta\beta'}(x,t;0,t_0)\tilde{S}^{cc'}_{\gamma\gamma'}(x,t;0,t_0) \\ & -\tilde{S}^{ab'}_{\alpha\beta'}(x,t;0,t_0)\tilde{S}^{ba'}_{\beta\alpha'}(x,t;0,t_0)\tilde{S}^{cc'}_{\gamma\gamma'}(x,t;0,t_0)) \\ & \epsilon^{abc}\epsilon^{a'b'c'}(C\gamma_5)_{\beta\gamma}(C\gamma_5)_{\gamma\gamma'} \end{aligned}$$

and for mesons

$$D_2^{\Gamma}(t,p) = \int d^3x e^{ipx}\tilde{S}^{aa'}_{\alpha\alpha'}(x,t;0,t_0)\tilde{S}^{b'b}_{\beta'\beta}(0,t_0;x,t)\Gamma^{\alpha\beta}\Gamma^{\alpha'\beta'}\delta_{ab}\delta_{a'b'}$$

where we expressed the meson source as $j_\gamma = \bar{\psi}^a_\mu \Gamma^{\mu\nu}\psi^b_\nu\delta_{ab}$. For the pion $\Gamma = \gamma_5$, and $\Gamma = \gamma_\mu$ for the ρ meson.

Combining these two expressions for $D_2(t)$ the masses of the lowest excited states can be extracted from the t-dependence. A particularly convenient method is to use the effective mass

$$m_{\text{eff}} = \log \frac{D_2(t)}{D_2(t+1)},$$

which asymptotically approaches m_0 at large distances. Figure 2-4 shows the effective masses for N, ρ and π versus lattice time for 35 (A) and 31 (B) configurations. One can clearly see that the plateau is being reached at fairly short separations of the source and the sink. We used the same set of propagators to determine the lattice scale as for three-point function calculations. The Gaussian smeared source fixed at $x = 0$ and point momentum projected sink ($p = 0$) were used. We estimated errors using the jackknife procedure outlined in appendix B.

Figure 2-5 shows the quark mass dependence of the nucleon (N), vector (ρ) and pseudo-scalar (π) mesons. Observed data agree well with the previously published [27]

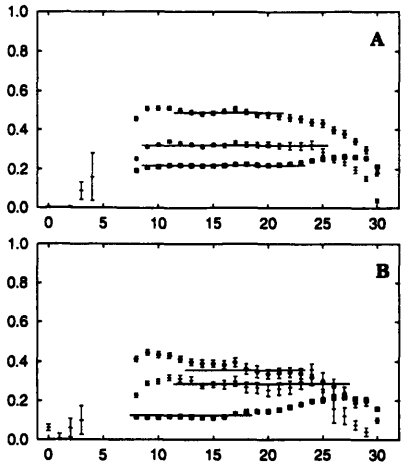

Figure 2-4: Effective masses at $\kappa = 0.15200$ (A) and $\kappa = 0.15294$ (B).

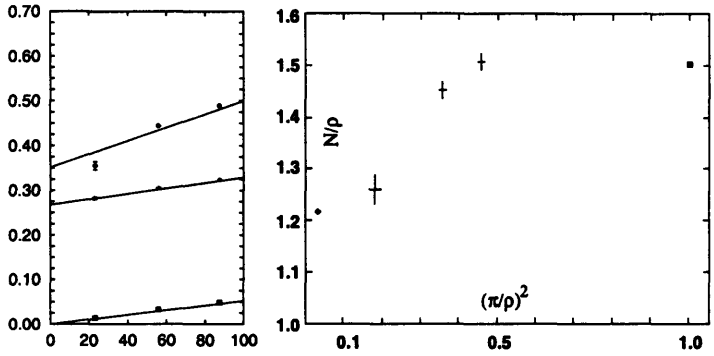

Figure 2-5: Mass κ-dependence

values of $\kappa_c = 0.15329^{+7}_{-4}$ and $a^{-1} = 2.8$GeV. Entries marked with a * indicate rescaling errors to bring χ^2 to 1. The procedure used to find plateau regions is explained in appendix C. This fixes the scale as follows:

κ		m_q, MeV	am_N	am_ρ	am_π
0.15200	35	86	0.4891(11)*	0.3245(06)	0.2197(04)
0.15246	35	55	0.4442(14)	0.3057(07)	0.1817(04)
0.15294	31	23	0.3552(29)	0.2824(16)	0.1188(08)*
0.15329		0	0.3530(110)	0.2700(80)	0

In this table we used eq. (2.4) for the quark mass with $Z = 1.12$. In the hadron

spectroscopy it is generally believed[3] that nucleon masses are very unreliable when $m_\pi < 3/La$. Our lightest quark mass is presumably too far in that region. Though both the hadron and mesons have roughly the same RMS radius, the former seems much more affected at the light quark mass. At the moment we do not have an explanation for apparently stable behavior of ρ and π.

[3]At the spectroscopy section of Lattice'95 this opinion was expressed by many participants.

Chapter 3

Hadron Structure Functions

Hadron physics has enjoyed very dynamic development in recent years. In the beginning few experimental data were available, which only allowed study of dominant effects. Now, however, with developing experimental techniques in deep inelastic scattering of leptons, e^+e^- annihilation and Drell-Yan processes, not only spin-independent functions can be measured experimentally, but the transversity distribution $h_1(x)$ and other chiral-odd distributions as well.

The parton distribution functions on the light-cone are important for understanding properties of the nucleon in high energy processes for several reasons. First, they are universal in a sense that the same distributions appear in completely different processes, so if one has measured a complete set of distribution functions from experiment or calculated them theoretically, then many hadron processes can be predicted. The distribution functions are invaluable for experiments probing physics beyond the standard model, as it allows us to consider a proton beam as a beam of quarks and gluons of *known* luminosity and thus relate observed cross sections to the fundamental vertices. Second, since the distribution functions depend on the strong interactions only, they can be calculated from the first principles, e.g., by using lattice quantum chromodynamics techniques. Alternatively, they could be used to test our understanding of non-perturbative methods and hadron structure.

In this chapter we consider calculation of the hadron structure functions on the lattice. Section 3.1 gives a review of structure functions in the continuum. In sec-

tion 3.2, we consider effects of the broken Lorenz symmetry on the various functions of interest. We move to the lattice in section 3.3. Renormalization is considered in section 3.4. One possible implementation on the lattice is described in section 3.5. We study different ways to create the proton on the lattice in section 3.6. Finally, in section 3.7 a calculation of the lowest moment of the tensor charge is presented.

3.1 Moments of Structure Functions I(Continuum)

In this section we summarize relevant continuum operators whose matrix elements will be computed numerically. Where possible we follow the notation of [28] which, in turn, is based on [29, 30, 31].

The forward virtual Compton scattering amplitude is related to the matrix element of the polarized on-shell nucleon ($p^2 = M^2$, $s^2 = -M^2$ and $s \cdot p = 0$) state

$$T_{\mu\nu}(q,p,s) \equiv i \int d^4x\, e^{iq\cdot x} \langle ps|T(J_\mu(x)J_\nu(0))|ps\rangle \tag{3.1}$$

The imaginary part of $T_{\mu\nu}$ can be written in terms of various scalar structure functions:

$$\begin{aligned}
\frac{1}{2\pi}\mathrm{Im}T_{\mu\nu} &= \frac{1}{4\pi}\int d^4x\, e^{q\cdot x} \langle ps|\left[J_\mu(x), J_\nu(0)\right]|ps\rangle \\
&= -g_{\mu\nu}\, F_1(x,Q^2) + \frac{p_\mu p_\nu}{\nu} F_2(x,Q^2) \\
&\quad + i\frac{\epsilon_{\mu\nu\lambda\sigma}q^\lambda}{\nu}\Big[s^\sigma\, g_1(x,Q^2) + (s^\sigma - p^\sigma\frac{q\cdot s}{\nu})\, g_2(x,Q^2)\Big] + \cdots
\end{aligned} \tag{3.2}$$

where $Q^2 = -q^2 > 0$, $\nu = p\cdot q \geq Q^2/2$ and $x = Q^2/2\nu$. The normalization of the nucleon state vector is chosen as $\langle ps|p's'\rangle = (2\pi)^3\, 2p^0\delta^3(\mathbf{p}-\mathbf{p}')\delta_{\mathbf{s},\mathbf{s}'}$, with $p^0 = E_{\mathbf{p}} = \sqrt{M^2+\mathbf{p}^2}$ and M being the mass of the nucleon.

In this work, we restrict our attention to the twist-2 structure functions included in eq. (3.2), which are the leading terms in the large Q^2 limit.

3.1.1 Spin-Independent Case (F_1 and F_2)

The moments of the structure functions F_1 and F_2 are related to the forward spin-averaged nucleon matrix elements through the operator product expansion (OPE).

For even values of $n \geq 2$ one has

$$\int_0^1 dx\, x^{n-1}\, F_1(x,Q^2) = \frac{1}{2}\sum_{f=u,d} c_{1,n}^{(f)}(\mu^2/Q^2, g(\mu))\, v_n^{(f)}(\mu) \tag{3.3}$$

$$\int_0^1 dx\, x^{n-2}\, F_2(x,Q^2) = \sum_{f=u,d} c_{2,n}^{(f)}(\mu^2/Q^2, g(\mu))\, v_n^{(f)}(\mu) \tag{3.4}$$

where the matrix elements $v_n^{(f)}(\mu)$ are defined by

$$\frac{1}{2}\sum_s \langle ps|\mathcal{O}^{(f)}_{\{\mu_1\cdots\mu_n\}}|ps\rangle = 2v_n^{(f)}\,[p_{\mu_1}\cdots p_{\mu_n} - \text{traces}] \tag{3.5}$$

$$\mathcal{O}^{(f)}_{\{\mu_1\mu_2\cdots\mu_n\}} = \left(\frac{i}{2}\right)^{n-1} \bar{\psi}_f(x)\gamma_{\{\mu_1}\overleftrightarrow{D}_{\mu_2}\cdots\overleftrightarrow{D}_{\mu_n\}}\psi_f(x) - \text{traces}. \tag{3.6}$$

The trace terms in the above equations is needed to construct an irreducible representation of the Lorenz group. E.g., for $n = 2$ the right hand side of eq. (3.5) is $p_{\mu_1}p_{\mu_2} - p^2/4\delta_{\mu_1\mu_2}$.

The Wilson coefficients $c_{1,n}^{(f)}(\mu^2/Q^2, g(\mu))$ and $c_{2,n}^{(f)}(\mu^2/Q^2, g(\mu))$ are known to the second order in perturbation theory; they can also be computed nonperturbatively on the lattice [32]. Although the structure functions are independent of the subtraction scale, μ, the matrix elements and the Wilson coefficients separately do depend on μ. In the parton model the moments of F_1 and F_2 can be interpreted in the following way:

$$v_n^{(f)} = \langle x^{n-1}\rangle^{(f)} \tag{3.7}$$

where x denotes the momentum fraction of a nucleon carried by quarks.

3.1.2 Spin-Dependent Case (g_1 and g_2)

The moments of the structure functions g_1 and g_2 are related to the forward spin-dependent nucleon matrix elements through the following expressions, derived from OPE.

For even values of $n \geq 0$ for g_1 and $n \geq 2$ for g_2

$$\int_0^1 dx\, x^n\, g_1(x,Q^2) = \frac{1}{4}\sum_{f=u,d} e_{1,n}^{(f)}(\mu^2/Q^2, g(\mu))\, a_n^{(f)}(\mu) \tag{3.8}$$

$$\int_0^1 dx\, x^n\, g_2(x,Q^2) = \frac{1}{4}\frac{n}{n+1}\sum_{f=u,d} \left[e_{2,n}^{(f)}(\mu^2/Q^2, g(\mu))\, d_n^{(f)}(\mu) - e_{1,n}^{(f)}(\mu^2/Q^2, g(\mu))\, a_n^{(f)}(\mu)\right] \tag{3.9}$$

where the matrix elements $a_n^{(f)}(\mu)$ and $d_n^{(f)}(\mu)$ are defined by

$$\langle ps|\mathcal{O}^{(5,f)}_{\{\sigma\mu_1\cdots\mu_n\}}|ps\rangle = \frac{1}{n+1}\, a_n^{(f)}\, [\mathcal{S}_{n+1} s_\sigma p_{\mu_1}\cdots p_{\mu_n} \;-\; \text{traces}] \tag{3.10}$$

$$\mathcal{O}^{(5,f)}_{\{\sigma\mu_1\mu_2\cdots\mu_n\}} = \left(\frac{i}{2}\right)^n \bar{\psi}_f(x)\gamma_5\gamma_{\{\sigma}\overleftrightarrow{D}_{\mu_1}\overleftrightarrow{D}_{\mu_2}\cdots\overleftrightarrow{D}_{\mu_n\}}\,\psi_f(x) - \text{traces} \tag{3.11}$$

$$\langle ps|\mathcal{O}^{(5,f)}_{[\sigma\{\mu_1]\mu_2\cdots\mu_n\}}|ps\rangle = \frac{1}{n+1}\, d_n^{(f)}\, [\mathcal{S}_n(s_\sigma p_{\mu_1} - s_{\mu_1}p_\sigma)p_{\mu_2}\cdots p_{\mu_n} \;-\; \text{traces}] \tag{3.12}$$

$$\mathcal{O}^{(5,f)}_{[\sigma\{\mu_1]\mu_2\cdots\mu_n\}} = \left(\frac{i}{2}\right)^n \bar{\psi}_f(x)\gamma_5\gamma_{[\sigma}\overleftrightarrow{D}_{\{\mu_1]}\overleftrightarrow{D}_{\mu_2}\cdots\overleftrightarrow{D}_{\mu_n\}}\,\psi_f(x) - \text{traces} \tag{3.13}$$

where $\mathcal{S}_n$ symmetrizes indices $\mu_1,\cdots,\mu_n$ only, $\mathcal{S}_{n+1}$ symmetrizes indices $\sigma,\mu_1,\cdots,\mu_n$ and $[\sigma\{\mu_1]\mu_2\cdots\mu_n\}$ is defined as

$$\mathcal{O}_{[\sigma\{\mu_1]\mu_2\cdots\mu_n\}} = \frac{1}{n+1}\Big[\mathcal{O}_{\sigma\{\mu_1\mu_2\cdots\mu_n\}} - \mathcal{O}_{\mu_1\{\sigma\mu_2\cdots\mu_n\}} + \mathcal{O}_{\sigma\{\mu_1\mu_2\cdots\mu_n\}} - \mathcal{O}_{\mu_2\{\mu_1\sigma\cdots\mu_n\}} + \cdots\Big] \tag{3.14}$$

such that the following decomposition is true

$$\mathcal{O}_{\sigma\{\mu_1\mu_2\cdots\mu_n\}} = \mathcal{O}_{\{\sigma\mu_1\mu_2\cdots\mu_n\}} + \mathcal{O}_{[\sigma\{\mu_1]\mu_2\cdots\mu_n\}} \tag{3.15}$$

As in the spin-independent case, the Wilson coefficients $c_{i,n}^{(f)}(\mu^2/Q^2, g(\mu))$ $(i = 1,2)$ are known to some order in perturbation theory. The lowest moments in the parton model are interpreted as the fraction of the nucleon spin carried by the respective quark flavors:

$$a_0^{(u)} = 2\Delta u\,, \qquad a_0^{(d)} = 2\Delta d. \tag{3.16}$$

3.1.3 Tensor Charge (h_1)

The moments of the quark distribution function (rather than the structure function) h_1 are related to the forward spin-dependent nucleon matrix elements through the expression, as defined in [33].

$$\int_0^1 dx\; x^{n-1}\,[h_1(x) - (-1)^{n-1}\bar{h}_1(x)] = t_n^{(f)}(\mu) \tag{3.17}$$

where the matrix elements $t_n^{(f)}(\mu)$ are defined by

$$\langle ps|\mathcal{O}^{(5',f)}_{\sigma\{\mu_1\cdots\mu_n\}}|ps\rangle = 2t_n^{(f)}(\mu)\; [\mathcal{S}_n(s_\sigma p_{\mu_1} - s_{\mu_1}p_\sigma)p_{\mu_2}\cdots p_{\mu_n}/M \;-\; \text{traces}] \tag{3.18}$$

$$O^{(5',f)}_{\sigma\{\mu_1\cdots\mu_n\}} = \left(\frac{i}{2}\right)^{n-2} \bar{\psi}_f(x)\gamma_5\sigma_{\sigma\{\mu_1} \overleftrightarrow{D}_{\mu_2} \cdots \overleftrightarrow{D}_{\mu_n\}} \psi_f(x) \tag{3.19}$$

Here again, $t_n^{(f)}(\mu)$ explicitly depends on the subtraction scale μ. $t_1^{(f)}(\mu)$ is also called the tensor charge of the nucleon for flavor f. It is important to note that h_1 in eq. (3.17) is only meaningful when a subtraction is implied. To relate h_1 to a physical process (the Drell-Yan process in this particular case), one needs to perform an OPE type of analysis with a given subtraction scheme. In this regard h_1 is not on the same footing as the structure functions, such as F_1, F_2, g_1 and g_2 considered earlier.

3.2 Breaking Lorenz Symmetry

As was mentioned in chapter 2, the Poincaré symmetry is explicitly broken on the lattice. In this section we consider the remnants of the rotational subgroup $SO(4)$ that survive on the lattice and the impact of this symmetry breaking on calculations. We start with a condensed review of relevant group theory. For a full treatment of the representation theory of Lie groups and reduction to subgroups, one can use any textbook on the subject, e.g., [34]. We build several low dimensional representations of the hypercubic group H_4 and decompose spin 2, 3 and 4 representations of $SO(4)$ into irreducible representations of H_4. The results will be used in section 3.3 to construct lattice operators.

H_4 is the group of symmetries of the four dimensional hypercube. It has 192 proper rotations belonging to 13 conjugacy classes. Every element of H_4 can be regarded as a product of a permutation of axes and a reflection, with combined parity even. The alternating group S_4 acting on the base vectors in $\mathbf{R}^4$ is a subgroup of H_4. On the other hand, H_4 is a subgroup of $SO(4)$ (the proper rotation group in four Euclidean dimensions), which is locally isomorphic to $SU(2) \otimes SU(2)$. Further details and the representation of H_4 can be found in [35].

Let $[n_1n_2n_3n_4]$ denote a Young diagram of S_4 with n_1 boxes in the first row, n_2 boxes in the second row and so on. For example, [4000] and [1111] are the symmetric and antisymmetric representations of S_4 respectively. (i, j) denotes the spin-i and spin-j representation of $SO(4)$ in the $SU(2) \otimes SU(2)$ format. In addition, $\overline{(i,j)} =$

$[1111] \otimes (i,j)$. We follow the simplified notation used in [36]. The correspondence with the notation used in[35] is given in Table 3.1.

For our purpose, we need to consider the decomposition of the direct product of n factors of $\mathbf{4}_1$, according to which γ_μ and D_μ transform under H_4, into a direct sum over the thirteen irreducible representations

$$\bigotimes_{i=1}^{n} \mathbf{4}_1 = \bigoplus_{\alpha=1}^{C(H_4)} m_\alpha(n)\, R^{(\alpha)} \tag{3.20}$$

where $R^{(\alpha)}$ denotes the α-th irreducible representation in Table 3.1 and $C(H_4) = 13$ is the number of classes in H_4. The integer $m_\alpha(n)$ is calculated according to

$[4000]$	$\mathbf{1}_1$	$\overline{(1,0)}$	$\mathbf{3}_5$
$[1111]$	$\mathbf{1}_2$	$\overline{(0,1)}$	$\mathbf{3}_6$
$[2200]$	$\mathbf{2}$	$(\frac{1}{2},\frac{1}{2})$	$\mathbf{4}_1$
$[3100]$	$\mathbf{3}_1$	$\overline{(\frac{1}{2},\frac{1}{2})}$	$\mathbf{4}_2$
$[2110]$	$\mathbf{3}_2$	$\mathbf{6}$	$\mathbf{6}$
$(1,0)$	$\mathbf{3}_3$	$\mathbf{8}$	$\mathbf{8}$
$(0,1)$	$\mathbf{3}_4$		

Table 3.1: The correspondence between notations in references [36] and [35].

$$m_\alpha(n) = \frac{1}{D(H_4)} \sum_{c=1}^{C(H_4)} w_c\, \chi^{(\alpha)}(c) \left(\chi^{(\mathbf{4}_1)}(c)\right)^n \tag{3.21}$$

where $D(H_4) = 192$ is the number of elements in H_4, w_c is the number of elements in class c and $\chi^{(\alpha)}(c)$ is the character of the elements of class c in the irreducible representation α.

To explicitly construct the basis vectors which transform as a given irreducible representation, let us introduce the notation $|ij\ldots\rangle$, where $i,j,\ldots = 1,2,3,4$. The inner product in this notation is $\langle i'j'\ldots|ij\ldots\rangle = \delta_{ii'}\delta_{jj'}\ldots$. The group generators act on this vector space in the following way. A reflection along the ith axis, P_i, results in a factor of -1 for each index i in $|\ldots\rangle$. For example,

$$P_1|11\rangle = |11\rangle, \qquad P_2|21\rangle = -|21\rangle, \qquad P_4|23\rangle = |23\rangle \tag{3.22}$$

A permutation (ij) interchanges indices with values i and j. For example,

$$(12)|12\rangle = |21\rangle, \qquad (14)|13\rangle = |43\rangle, \qquad (12)|34\rangle = |34\rangle \tag{3.23}$$

Since reflections P_i and permutations (ij) generate the whole group H_4, every group element can be represented as a finite composition of suitable permutations and reflections.

3.2.1 Reduction of $SO(4)$ to H_4

In general, irreducible representations of $SO(4)$ become reducible when the symmetry is restricted to H_4. One can use character orthogonality to find which irreducible representations of a small group compose a representation of the large group. The relevant formula is

$$m_\alpha(\beta) = \frac{1}{D(H_4)} \sum_{c=1}^{C(H_4)} w_c \chi^{(\alpha)}(c) \cdot \bar{\chi}^{(\beta)}(c) \tag{3.24}$$

where $\bar{\chi}^{(\beta)}(c)$ is the character of the $SO(4)$ irreducible representation labeled by β. Some of the $\bar{\chi}^{(\beta)}(c)$'s can be found in Table 3.3 of [36]. Others can be calculated by using the $SO(4)$ decomposition rules. For example,

$$\bar{\chi}^{(\frac{3}{2},\frac{3}{2})} = \bar{\chi}^{(1,1)} \cdot \chi^{(\frac{1}{2},\frac{1}{2})} - \bar{\chi}^{(\frac{3}{2},\frac{1}{2})} - \bar{\chi}^{(\frac{1}{2},\frac{3}{2})} - \chi^{(\frac{1}{2},\frac{1}{2})}. \tag{3.25}$$

This equation follows from the $SO(4)$ decomposition $(1,1) \otimes (\frac{1}{2},\frac{1}{2}) = (\frac{3}{2},\frac{3}{2}) \oplus (\frac{3}{2},\frac{1}{2}) \oplus (\frac{1}{2},\frac{3}{2}) \oplus (\frac{1}{2},\frac{1}{2})$, and

$$\bar{\chi}^{(\frac{3}{2},\frac{1}{2})} + \bar{\chi}^{(\frac{1}{2},\frac{3}{2})} = \left(\chi^{(1,0)} + \chi^{(0,1)} - 2\right) \cdot \chi^{(\frac{1}{2},\frac{1}{2})} \tag{3.26}$$

from the $SO(4)$ decompositions $(1,0) \otimes (\frac{1}{2},\frac{1}{2}) = (\frac{3}{2},\frac{1}{2}) \oplus (\frac{1}{2},\frac{1}{2})$ and $(0,1) \otimes (\frac{1}{2},\frac{1}{2}) = (\frac{1}{2},\frac{3}{2}) \oplus (\frac{1}{2},\frac{1}{2})$. In the above calculation, we have used the facts that the $SO(4)$ representations $(0,1)$, $(1,0)$ and $(\frac{1}{2},\frac{1}{2})$ are irreducible both in $SO(4)$ and H_4, so that $\bar{\chi}(c)$ in these representations are the same. We also used the identity $\bar{\chi}^{(i,j)} = \bar{\chi}^{(i,0)} \cdot \bar{\chi}^{(0,j)}$.

We are mostly interested in reduction of symmetric $SO(4)$ tensors to H_4 irreducible representations. The results up to rank 4 together with some other useful representations are given in Table 3.2.

Rank	$SO(4)$ representation	H_4 reduction
1	$(\frac{1}{2},\frac{1}{2})$	$\mathbf{4}_1$
2	$(1,1)$	$\mathbf{3}_1 \oplus \mathbf{6}$
3	$(\frac{3}{2},\frac{1}{2})$ or $(\frac{1}{2},\frac{3}{2})$	$\mathbf{8}$
3	$(\frac{3}{2},\frac{3}{2})$	$\mathbf{4}_1 \oplus \mathbf{4}_2 \oplus \mathbf{8}$
4	$(2,2)$	$\mathbf{1}_1 \oplus \mathbf{1}_2 \oplus \mathbf{2} \oplus \mathbf{3}_1 \oplus \mathbf{3}_3 \oplus \mathbf{3}_4 \oplus \mathbf{3}_5 \oplus \mathbf{3}_6 \oplus \mathbf{6}$

Table 3.2: Reduction of $SO(4)$ representations to H_4

While $SO(4)$-vectors remain irreducible under H_4, the higher rank tensors require more careful considerations.

3.2.2 Rank 2

There are 16 independent vectors $|ij\rangle$, with $i,j = 1,2,3,4$ in $\mathbf{4}_1 \otimes \mathbf{4}_1$. The decomposition

$$\mathbf{4}_1 \otimes \mathbf{4}_1 = \mathbf{1}_1 \oplus \mathbf{3}_1 \oplus \mathbf{3}_3 \oplus \mathbf{3}_4 \oplus \mathbf{6} \tag{3.27}$$

implies that the rank 2 tensors form five irreducible representations. The explicit orthonormal bases for these representations are as follows:

$$\hat{e}^{(\mathbf{1}_1)}_{A_2} = \left\{\frac{1}{2}\big(|11\rangle + |22\rangle + |33\rangle + |44\rangle\big)\right\} \tag{3.28}$$

$$\hat{e}^{(\mathbf{3}_1)}_{A_2} = \left\{\frac{1}{\sqrt{12}}\big(3|44\rangle - |11\rangle - |22\rangle - |33\rangle\big), \frac{1}{\sqrt{6}}\big(2|33\rangle - |11\rangle - |22\rangle\big), \frac{1}{\sqrt{2}}\big(|22\rangle - |11\rangle\big)\right\} \tag{3.29}$$

$$\hat{e}^{(\mathbf{3}_3)}_{B_2} = \left\{\frac{1}{\sqrt{2}}\big(|[14]\rangle + |[23]\rangle\big), \frac{1}{\sqrt{2}}\big(|[24]\rangle - |[13]\rangle\big), \frac{1}{\sqrt{2}}\big(|[34]\rangle + |[12]\rangle\big)\right\} \tag{3.30}$$

$$\hat{e}^{(\mathbf{3}_4)}_{B_2} = \left\{\frac{1}{\sqrt{2}}\big(|[14]\rangle - |[23]\rangle\big), \frac{1}{\sqrt{2}}\big(|[24]\rangle + |[13]\rangle\big), \frac{1}{\sqrt{2}}\big(|[34]\rangle - |[12]\rangle\big)\right\} \tag{3.31}$$

$$\hat{e}^{(\mathbf{6})}_{B_2} = \Big\{|\{12\}\rangle, |\{13\}\rangle, |\{14\}\rangle, |\{23\}\rangle, |\{24\}\rangle, |\{34\}\rangle\Big\} \tag{3.32}$$

where $|[ij]\rangle = (|ij\rangle - |ji\rangle)/\sqrt{2}$ and $|\{ij\}\rangle = (|ij\rangle + |ji\rangle)/\sqrt{2}$ for $i \neq j$. The basis vector sets are labeled by their irreducible representations and index patterns according to

table 3.3. Though in this case the representation index of the base vectors are unique and pattern labels do not carry any additional information, they will become necessary for higher ranks. To keep notation uniform, we retain pattern labels for all ranks.

label	index pattern	dimension	multiplets
A_2	ii	4	$\mathbf{1}_1 \oplus \mathbf{3}_1$
B_2	ij	12	$\mathbf{3}_3 \oplus \mathbf{3}_4 \oplus \mathbf{6}$

Table 3.3: Index patterns for $\mathbf{4}_1 \otimes \mathbf{4}_1 = \mathbf{1}_1 \oplus \mathbf{3}_1 \oplus \mathbf{3}_3 \oplus \mathbf{3}_4 \oplus \mathbf{6}$.

From

$$(\frac{1}{2},\frac{1}{2}) \otimes (\frac{1}{2},\frac{1}{2}) = (0,0) \oplus (0,1) \oplus (1,0) \oplus (1,1) \tag{3.33}$$

one can easily see that spin 2 representation $(1,1)$ of $SO(4)$ consists of $\{\hat{e}_{A_2}^{(\mathbf{3}_1)}\} \oplus \{\hat{e}_{B_2}^{(\mathbf{6})}\}$. Accordingly, instead of one spin 2 state, on the lattice there are 2 separate states $(\mathbf{3}_1; A_2)$ and $(\mathbf{6}; B_2)$ with different energies. In the continuum limit they become degenerate and form spin 2 states as Lorenz symmetry is restored.

3.2.3 Rank 3

There are 64 independent vectors $|ijk\rangle$, $(i,j,k = 1,2,3,4)$. The decomposition

$$\mathbf{4}_1 \otimes \mathbf{4}_1 \otimes \mathbf{4}_1 = 5 \cdot \mathbf{4}_1 \oplus \mathbf{4}_2 \oplus 5 \cdot \mathbf{8} \tag{3.34}$$

divides rank 3 tensors into eleven irreducible representations. $\mathbf{4}_2$ appears only once in the reduction and from table 3.2 it is clear that it comes from $(\frac{3}{2},\frac{3}{2})$ representation of $SO(4)$. Orthonormal bases of all representations are given below

$$\hat{e}_{A_3}^{(\mathbf{4}_1)} = \Big\{|111\rangle, |222\rangle, |333\rangle, |444\rangle\Big\} \tag{3.35}$$

$$\begin{aligned}\hat{e}_{B_3}^{(\mathbf{4}_1,1)} = \Big\{&\frac{1}{\sqrt{3}}\big(|122\rangle + |133\rangle + |144\rangle\big), \frac{1}{\sqrt{3}}\big(|211\rangle + |233\rangle + |244\rangle\big), \\ &\frac{1}{\sqrt{3}}\big(|311\rangle + |322\rangle + |344\rangle\big), \frac{1}{\sqrt{3}}\big(|411\rangle + |422\rangle + |433\rangle\big)\Big\}\end{aligned} \tag{3.36}$$

$$\hat{e}_{B_3}^{(\mathbf{8},1)} = \Big\{\frac{1}{\sqrt{6}}\big(2|122\rangle - |133\rangle - |144\rangle\big), \frac{1}{\sqrt{2}}\big(|133\rangle - |144\rangle\big),$$

$$\frac{1}{\sqrt{6}}\big(2|211\rangle - |244\rangle - |233\rangle\big), \frac{1}{\sqrt{2}}\big(|244\rangle - |233\rangle\big),$$
$$\frac{1}{\sqrt{6}}\big(2|311\rangle - |322\rangle - |344\rangle\big), \frac{1}{\sqrt{2}}\big(|322\rangle - |344\rangle\big),$$
$$\frac{1}{\sqrt{6}}\big(2|411\rangle - |422\rangle - |433\rangle\big), \frac{1}{\sqrt{2}}\big(|422\rangle - |433\rangle\big)\Big\} \tag{3.37}$$

$$\hat{e}^{(\mathbf{4}_1)}_{C_3} = \Big\{|[234]\rangle, |[134]\rangle, |[124]\rangle, |[123]\rangle\Big\} \tag{3.38}$$

$$\hat{e}^{(\mathbf{4}_2)}_{C_3} = \Big\{|\{234\}\rangle, |\{134\}\rangle, |\{124\}\rangle, |\{123\}\rangle\Big\} \tag{3.39}$$

$$\begin{aligned}\hat{e}^{(\mathbf{8},1)}_{C_3} = \Big\{&\frac{1}{\sqrt{6}}\big(2|2\{34\}\rangle - |3\{24\}\rangle - |4\{32\}\rangle\big), \frac{1}{\sqrt{2}}\big(|3\{24\}\rangle - |4\{32\}\rangle\big),\\ &\frac{1}{\sqrt{6}}\big(2|1\{34\}\rangle - |3\{14\}\rangle - |4\{31\}\rangle\big), \frac{1}{\sqrt{2}}\big(|3\{14\}\rangle - |4\{31\}\rangle\big),\\ &\frac{1}{\sqrt{6}}\big(2|1\{24\}\rangle - |2\{14\}\rangle - |4\{21\}\rangle\big), \frac{1}{\sqrt{2}}\big(|2\{14\}\rangle - |4\{21\}\rangle\big),\\ &\frac{1}{\sqrt{6}}\big(2|1\{23\}\rangle - |2\{13\}\rangle - |3\{21\}\rangle\big), \frac{1}{\sqrt{2}}\big(|2\{13\}\rangle - |3\{21\}\rangle\big)\Big\}\end{aligned} \tag{3.40}$$

$$\begin{aligned}\hat{e}^{(\mathbf{8},2)}_{C_3} = \Big\{&\frac{1}{\sqrt{6}}\big(2|2[34]\rangle + |3[24]\rangle + |4[32]\rangle\big), \frac{1}{\sqrt{2}}\big(|3[24]\rangle - |4[32]\rangle\big),\\ &\frac{1}{\sqrt{6}}\big(2|1[34]\rangle + |3[14]\rangle + |4[31]\rangle\big), \frac{1}{\sqrt{2}}\big(|3[14]\rangle - |4[31]\rangle\big),\\ &\frac{1}{\sqrt{6}}\big(2|1[24]\rangle + |2[14]\rangle + |4[21]\rangle\big), \frac{1}{\sqrt{2}}\big(|2[14]\rangle - |4[21]\rangle\big),\\ &\frac{1}{\sqrt{6}}\big(2|1[23]\rangle + |2[13]\rangle + |3[21]\rangle\big), \frac{1}{\sqrt{2}}\big(|2[13]\rangle - |3[21]\rangle\big)\Big\}\end{aligned} \tag{3.41}$$

The basis vectors for $(\mathbf{4}_1, i)_{B_3}$ and $(\mathbf{8}, i)_{B_3}$, with $i = 2, 3$, can be obtained from the basis vectors of $(\mathbf{4}_1, 1)_{B_3}$ and $(\mathbf{8}, 1)_{B_3}$ by putting the non-repeated indices at the second and third positions respectively.

Representations $\mathbf{4}_1$ and $\mathbf{8}$ appear more than once and we need an additional label to distinguish these multiple copies. It is always possible to make all the basis orthogonal with respect to this additional label. It is convenient to use a labeling scheme based on the symmetry pattern of the indices. Table 3.4 schematically summarizes the relevant index patterns.

As a result, $(\mathbf{8}, 1)_{B_3}$ and $(\mathbf{8}, 1)_{C_3}$ are mutually orthogonal. In this particular case $(\mathbf{8}, 1)_{B_3}$ and $(\mathbf{8}, 1)_{C_3}$ do not mix because of different index pattern structure. However, in general this additional symmetry does not protect copies of the same representation from mixing under renormalization.

label	index pattern(s)	dimension	multiplets
A_3	iii	4	$\mathbf{4}_1$
B_3	iij, iji, jii	36	$3 \cdot (\mathbf{4}_1 \oplus \mathbf{8})$
C_3	ijk	24	$\mathbf{4}_1 \oplus \mathbf{4}_2 \oplus 2 \cdot \mathbf{8}$

Table 3.4: Index patterns for $\mathbf{4}_1 \otimes \mathbf{4}_1 \otimes \mathbf{4}_1 = 5 \cdot \mathbf{4}_1 \oplus \mathbf{4}_2 \oplus 5 \cdot \mathbf{8}$.

3.2.4 Rank 4

There are 256 independent vectors $|ijkl\rangle$, with $i, j, k, l = 1, 2, 3, 4$. According to the decomposition

$$\begin{aligned} \mathbf{4}_1 \otimes \mathbf{4}_1 \otimes \mathbf{4}_1 \otimes \mathbf{4}_1 \;=\; & 5 \cdot \mathbf{1}_1 \oplus \mathbf{1}_2 \oplus 5 \cdot \mathbf{2} \oplus 10 \cdot \mathbf{3}_1 \oplus 6 \cdot \mathbf{3}_2 \\ & \oplus 10 \cdot \mathbf{3}_3 \oplus 10 \cdot \mathbf{3}_4 \oplus 6 \cdot \mathbf{3}_5 \oplus 6 \cdot \mathbf{3}_6 \oplus 16 \cdot \mathbf{6} \end{aligned} \tag{3.42}$$

all irreducible representations of H_4 appear in the decomposition except $\mathbf{4}_1$, $\mathbf{4}_2$ and $\mathbf{8}$. Only $\mathbf{1}_2$ shows up once. The explicit orthonormal bases are given below

$$\hat{e}_{A_4}^{(\mathbf{1}_1)} = \left\{ \frac{1}{2}\big(|1111\rangle + |2222\rangle + |3333\rangle + |4444\rangle\big) \right\} \tag{3.43}$$

$$\begin{aligned} \hat{e}_{A_4}^{(\mathbf{3}_1)} = \Big\{ & \frac{1}{2}\big(|1111\rangle + |2222\rangle - |3333\rangle - |4444\rangle\big), \\ & \frac{1}{\sqrt{2}}\big(|1111\rangle - |2222\rangle\big), \frac{1}{\sqrt{2}}\big(|3333\rangle - |4444\rangle\big) \Big\} \end{aligned} \tag{3.44}$$

$$\begin{aligned} \hat{e}_{B_4}^{(\mathbf{3}_3,1)} = \Big\{ & \frac{1}{2}\big(|4111\rangle - |1444\rangle + |3222\rangle - |2333\rangle\big), \\ & \frac{1}{2}\big(|4222\rangle - |2444\rangle - |3111\rangle + |1333\rangle\big), \\ & \frac{1}{2}\big(|4333\rangle - |3444\rangle + |2111\rangle - |1222\rangle\big) \Big\} \end{aligned} \tag{3.45}$$

$$\begin{aligned} \hat{e}_{B_4}^{(\mathbf{3}_4,1)} = \Big\{ & \frac{1}{2}\big(|4111\rangle - |1444\rangle - |3222\rangle + |2333\rangle\big), \\ & \frac{1}{2}\big(|4222\rangle - |2444\rangle + |3111\rangle - |1333\rangle\big), \\ & \frac{1}{2}\big(|4333\rangle - |3444\rangle - |2111\rangle + |1222\rangle\big) \Big\} \end{aligned} \tag{3.46}$$

$$\begin{aligned} \hat{e}_{B_4}^{(\mathbf{6},1)} = \Big\{ & \frac{1}{\sqrt{2}}\big(|2111\rangle + |1222\rangle\big), \frac{1}{\sqrt{2}}\big(|3111\rangle + |1333\rangle\big), \\ & \frac{1}{\sqrt{2}}\big(|4111\rangle + |1444\rangle\big), \frac{1}{\sqrt{2}}\big(|3222\rangle + |2333\rangle\big), \end{aligned}$$

$$\frac{1}{\sqrt{2}}\big(|4222\rangle + |2444\rangle\big), \frac{1}{\sqrt{2}}\big(|4333\rangle + |3444\rangle\big)\Big\} \tag{3.47}$$

$$\begin{aligned}
\hat{e}_{C_4}^{(\mathbf{1}_1,1)} &= \Big\{\frac{1}{\sqrt{6}}\big(\ |\{1122\}'\rangle + |\{1133\}'\rangle + |\{1144\}'\rangle + \\
&\qquad |\{2233\}'\rangle + |\{2244\}'\rangle + |\{3344\}'\rangle\big)\Big\} \\
\hat{e}_{C_4}^{(\mathbf{2},1)} &= \Big\{\frac{1}{\sqrt{12}}\big(\ 2|\{1122\}'\rangle + 2|\{3344\}'\rangle - |\{1144\}'\rangle\ \big), \\
&\qquad -|\{2233\}'\rangle - |\{1133\}'\rangle - |\{2244\}'\rangle \\
&\quad \frac{1}{2}\big(|\{1144\}'\rangle + |\{2233\}'\rangle - |\{1133\}'\rangle - |\{2244\}'\rangle\big)\Big\} \qquad (3.48) \\
\hat{e}_{C_4}^{(\mathbf{3}_1,1)} &= \Big\{\frac{1}{\sqrt{2}}\big(|\{1144\}'\rangle - |\{2233\}'\rangle\big), \frac{1}{\sqrt{2}}\big(|\{2244\}'\rangle - |\{1133\}'\rangle\big), \\
&\quad \frac{1}{\sqrt{2}}\big(|\{3344\}'\rangle - |\{1122\}'\rangle\big)\Big\} \qquad (3.49) \\
\hat{e}_{C_4}^{(\mathbf{3}_1,4)} &= \Big\{\frac{1}{2}\big(|[1144]'\rangle + |[2244]'\rangle + |[1133]'\rangle + |[2233]'\rangle\big), \\
&\quad \frac{1}{2}\big(|[2244]'\rangle + |[3344]'\rangle + |[2211]'\rangle + |[3311]'\rangle\big), \\
&\quad \frac{1}{2}\big(|[1144]'\rangle + |[3344]'\rangle + |[1122]'\rangle + |[3322]'\rangle\big)\Big\} \qquad (3.50) \\
\hat{e}_{C_4}^{(\mathbf{3}_2,1)} &= \Big\{\frac{1}{2}\big(|[1144]'\rangle - |[2244]'\rangle - |[1133]'\rangle + |[2233]'\rangle\big), \\
&\quad \frac{1}{2}\big(|[2244]'\rangle - |[3344]'\rangle - |[2211]'\rangle + |[3311]'\rangle\big), \\
&\quad \frac{1}{2}\big(|[1144]'\rangle - |[3344]'\rangle - |[1122]'\rangle + |[3322]'\rangle\big)\Big\} \qquad (3.51) \\
\hat{e}_{D_4}^{(\mathbf{3}_3,1)} &= \Big\{\frac{1}{2}\big(|22[14]\rangle + |33[14]\rangle + |11[23]\rangle + |44[23]\rangle\big), \\
&\quad \frac{1}{2}\big(|11[24]\rangle + |33[24]\rangle - |22[13]\rangle - |44[13]\rangle\big), \\
&\quad \frac{1}{2}\big(|11[34]\rangle + |22[34]\rangle + |33[12]\rangle + |44[12]\rangle\big)\Big\} \qquad (3.52) \\
\hat{e}_{D_4}^{(\mathbf{3}_4,1)} &= \Big\{\frac{1}{2}\big(|22[14]\rangle + |33[14]\rangle - |11[23]\rangle - |44[23]\rangle\big), \\
&\quad \frac{1}{2}\big(|11[24]\rangle + |33[24]\rangle + |22[13]\rangle + |44[13]\rangle\big), \\
&\quad \frac{1}{2}\big(|11[34]\rangle + |22[34]\rangle - |33[12]\rangle - |44[12]\rangle\big)\Big\} \qquad (3.53) \\
\hat{e}_{D_4}^{(\mathbf{3}_5,1)} &= \Big\{\frac{1}{2}\big(|22\{14\}\rangle - |33\{14\}\rangle + |11\{23\}\rangle - |44\{23\}\rangle\big), \\
&\quad \frac{1}{2}\big(|11\{24\}\rangle - |33\{24\}\rangle - |22\{13\}\rangle + |44\{13\}\rangle\big), \\
&\quad \frac{1}{2}\big(|11\{34\}\rangle - |22\{34\}\rangle + |33\{12\}\rangle - |44\{12\}\rangle\big)\Big\} \qquad (3.54)
\end{aligned}$$

$$\hat{e}_{D_4}^{(\mathbf{3}_6,1)} = \Big\{\frac{1}{2}\big(|22\{14\}\rangle - |33\{14\}\rangle - |11\{23\}\rangle + |44\{23\}\rangle\big), \frac{1}{2}\big(|11\{24\}\rangle - |33\{24\}\rangle + |22\{13\}\rangle - |44\{13\}\rangle\big), \frac{1}{2}\big(|11\{34\}\rangle - |22\{34\}\rangle - |33\{12\}\rangle + |44\{12\}\rangle\big)\Big\} \quad (3.55)$$

$$\hat{e}_{D_4}^{(\mathbf{6},1)} = \Big\{\frac{1}{\sqrt{2}}\big(|33\{12\}\rangle + |44\{12\}\rangle\big), \frac{1}{\sqrt{2}}\big(|22\{13\}\rangle + |44\{13\}\rangle\big), \frac{1}{\sqrt{2}}\big(|22\{14\}\rangle + |33\{14\}\rangle\big), \frac{1}{\sqrt{2}}\big(|11\{23\}\rangle + |44\{23\}\rangle\big), \frac{1}{\sqrt{2}}\big(|11\{24\}\rangle + |33\{24\}\rangle\big), \frac{1}{\sqrt{2}}\big(|11\{34\}\rangle + |22\{34\}\rangle\big)\Big\} \quad (3.56)$$

$$\hat{e}_{D_4}^{(\mathbf{6},7)} = \Big\{\frac{1}{\sqrt{2}}\big(|33[12]\rangle - |44[12]\rangle\big), \frac{1}{\sqrt{2}}\big(|22[13]\rangle - |44[13]\rangle\big), \frac{1}{\sqrt{2}}\big(|22[14]\rangle - |33[14]\rangle\big), \frac{1}{\sqrt{2}}\big(|11[23]\rangle - |44[23]\rangle\big), \frac{1}{\sqrt{2}}\big(|11[24]\rangle - |33[24]\rangle\big), \frac{1}{\sqrt{2}}\big(|11[34]\rangle - |22[34]\rangle\big)\Big\} \quad (3.57)$$

$$\hat{e}_{E_4}^{(\mathbf{1}_1)} = \Big\{\hat{P}_{[1111]}|1234\rangle\Big\} = \Big\{|[1234]\rangle\Big\} \quad (3.58)$$

$$\hat{e}_{E_4}^{(\mathbf{1}_2)} = \Big\{\hat{P}_{[4000]}|1234\rangle\Big\} = \Big\{|\{1234\}\rangle\Big\} \quad (3.59)$$

$$\hat{e}_{E_4}^{(\mathbf{2},1)} \oplus \hat{e}_{E_4}^{(\mathbf{2},2)} = \Big\{\hat{P}_{[2200]_i}p|1234\rangle;\ p \in P_4;\ i = 1,2\Big\} \quad (3.60)$$

$$\hat{e}_{E_4}^{(\mathbf{3}_1,1)} \oplus \hat{e}_{E_4}^{(\mathbf{3}_1,2)} \oplus \hat{e}_{E_4}^{(\mathbf{3}_1,3)} = \Big\{\hat{P}_{[3100]_i}p|1234\rangle;\ p \in P_4;\ i = 1,2,3\Big\} \quad (3.61)$$

$$\hat{e}_{E_4}^{(\mathbf{3}_2,1)} \oplus \hat{e}_{E_4}^{(\mathbf{3}_2,2)} \oplus \hat{e}_{E_4}^{(\mathbf{3}_2,3)} = \Big\{\hat{P}_{[2110]_i}p|1234\rangle;\ p \in P_4;\ i = 1,2,3\Big\} \quad (3.62)$$

where $|\{iijj\}'\rangle \equiv (|iijj\rangle + |jjii\rangle)/\sqrt{2}$ and $|[iijj]'\rangle \equiv (|iijj\rangle - |jjii\rangle)/\sqrt{2}$ for $i \neq j$; $\hat{P}_{[n_1n_2n_3n_4]_i}$ stands for the irreducible symmetrizer corresponding to the standard Young tableau labeled by $[n_1n_2n_3n_4]_i$ of the permutation group P_4. The explicit expressions of $\hat{P}$ are given here

$$\hat{P}_{[2200]_1} = [24][13]\{34\}\{12\}, \qquad \hat{P}_{[2200]_2} = [14][23]\{24\}\{13\},$$
$$\hat{P}_{[3100]_1} = [14]\{123\}, \qquad \hat{P}_{[2110]_1} = [123]\{14\},$$
$$\hat{P}_{[3100]_2} = [13]\{124\}, \qquad \hat{P}_{[2110]_2} = [124]\{13\},$$
$$\hat{P}_{[2110]_3} = [134]\{12\}, \qquad \hat{P}_{[3100]_3} = [12]\{134\}.$$

It is important to emphasize that $\hat{P}$ and $p \in P_4$ in the above equations both act on the position of the index, *not* on the value. In addition, orthonormalization is also needed in the last three subspaces.

The basis vectors for $(\mathbf{3}_3, i)_{B_4}$, $(\mathbf{3}_4, i)_{B_4}$ and $(\mathbf{6}, i)_{B_4}$, with $i = 2, 3, 4$, can be obtained from the basis vectors of $(\mathbf{3}_3, 1)_{B_4}$, $(\mathbf{3}_4, 1)_{B_4}$ and $(\mathbf{6}, 1)_{B_4}$ by putting the non-repeated indices at the second, third and fourth positions respectively. Basis vectors for the other two copies of C_4 patterns: $(\mathbf{1}_1, i)_{C_4}$, $(\mathbf{2}, i)_{C_4}$, $(\mathbf{3}_1, i)_{C_4}$, $(\mathbf{3}_1, 3+i)_{C_4}$ and $(\mathbf{3}_2, i)_{C_4}$, with $i = 2, 3$, can be obtained from the corresponding basis by replacing (1122) like index patterns by (1221) and (1212) like index patterns respectively. Analogously, basis vectors for the other five copies of D_4 patterns: $(\mathbf{3}_3, i)$, $(\mathbf{3}_4, i)$, $(\mathbf{3}_5, i)$, $(\mathbf{3}_6, i)$, $(\mathbf{6}, i)$ and $(\mathbf{6}, 6+i)$, with $i = 2, 3, 4, 5, 6$, can be obtained from the corresponding basis by replacing (1123) like index patterns by (2113), (2311), (1213), (1231) and (2131) like patterns respectively.

Table 3.5 schematically summarizes the relevant index patterns.

label	index pattern	dimension	multiplets
A_4	$iiii$	4	$\mathbf{1}_1 \oplus \mathbf{3}_1$
B_4	$iiij$	48	$4 \cdot (\mathbf{3}_3 \oplus \mathbf{3}_4 \oplus \mathbf{6})$
C_4	$iijj$	36	$3 \cdot (\mathbf{1}_1 \oplus \mathbf{2} \oplus 2 \cdot \mathbf{3}_1 \oplus \mathbf{3}_2)$
D_4	$iijk$	144	$6 \cdot (\mathbf{3}_3 \oplus \mathbf{3}_4 \oplus \mathbf{3}_5 \oplus \mathbf{3}_6 \oplus 2 \cdot \mathbf{6})$
E_4	$ijkl$	24	$\mathbf{1}_1 \oplus \mathbf{1}_2 \oplus 2 \cdot \mathbf{2} \oplus 3 \cdot \mathbf{3}_1 \oplus 3 \cdot \mathbf{3}_2$

Table 3.5: Index patterns for $\mathbf{4}_1 \otimes \mathbf{4}_1 \otimes \mathbf{4}_1 \otimes \mathbf{4}_1 = 5 \cdot \mathbf{1}_1 \oplus \mathbf{1}_2 \oplus 5 \cdot \mathbf{2} \oplus 10 \cdot \mathbf{3}_1 \oplus 6 \cdot \mathbf{3}_2 \oplus 10 \cdot \mathbf{3}_3 \oplus 10 \cdot \mathbf{3}_4 \oplus 6 \cdot \mathbf{3}_5 \oplus 6 \cdot \mathbf{3}_6 \oplus 16 \cdot \mathbf{6}$

3.2.5 Permutation Symmetries

One can notice that we divided basis vectors in multiple copies of the same irreducible representation of H_4 in rank 3 and 4 tensors according to representations of the permutation groups P_3 and P_4 respectively. This is not by accident, but rather reflects the fact that permutation operations on the index positions commute with H_4 operations. The same behavior occurs for decompositions of tensor products involving Lie groups, such as $SU(2)$ and $SO(4)$. The commutativity guarantees the

simultaneous block-diagonalization of representations of the permutation group P_k and H_4 group, where k is the number of the copies of $\mathbf{4}_1$ in the product $(\mathbf{4}_1)^k$.

When $m_\alpha(n) > 1$ it is also possible to make the permutation symmetry explicit, though this is not done in the basis vectors given earlier. For example, the three copies of $(\mathbf{4}_1;\ \mathbf{8})_{B_3}$ can be put into irreducible representations of P_3 characterized by Young diagrams [300] (of dimension 1) and [210] (of dimension 2). Similarly, the four copies of $(\mathbf{3}_3;\ \mathbf{3}_4;\ \mathbf{6})_{B_4}$, the three copies of $(\mathbf{1}_1;\ \mathbf{2};\ 2\cdot\mathbf{3}_1;\ \mathbf{3}_2)_{C_4}$ and the six copies of $(\mathbf{3}_3;\ \mathbf{3}_4;\ \mathbf{3}_5;\ \mathbf{3}_6;\ 2\cdot\mathbf{6})_{D_4}$ can be symmetrized or antisymmetrized into irreducible representations of P_4.

In addition, the repetitions not common to the same index pattern at a given rank, such as $(\mathbf{8};\ i=1,2)_{C_3}$, $(\mathbf{3}_1;\ i=1,4)_{C_4}$ and $(\mathbf{6};\ i=1,7)_{D_4}$ can be further distinguished by some subgroups of P_3 and P_4. Other similar replicas, such as those appearing in index pattern E_4, are labeled by various standard Young tableaux of P_4.

Symmetry properties under permutations will turn out to be useful for determining whether logarithmic mixing occurs for those representations which appear more than once in H_4 tensor decompositions. This is because some of the operators of rank 3 or higher vanish identically upon antisymmetrization of the repeated factors of D_μ. When we constructed the basis vectors earlier, we have assumed that factors of $\mathbf{4}_1$ are distinguishable through index positions. If some of the $\mathbf{4}_1$ factors are identical, certain irreducible representations may not be realized, and hence some of the $m_\alpha(n)$'s are effectively reduced.

3.3 Moments of Structure Functions II(Lattice)

In this section we rewrite operators defined in section 3.1 on the lattice. The definitions are the same except for using symmetric lattice derivatives 2.5 in place of covariant derivatives. In the previous section, it was shown that symmetrizing over Lorenz indices is not sufficient to separate different representations in the moments. Generally, irreducible representations of $SO(4)$ used in section 3.1 are direct sums of a number of irreducible representations of H_4, since the latter is a subgroup of $SO(4)$.

For this reason, operator mixing requires more careful consideration on the lattice.

In subsections 3.3.1 and 3.3.2, we consider moments of the spin-dependent and spin-independent structure functions on the lattice. In section 3.3.3 we summarize possible operators for various moments.

3.3.1 Spin-Independent Case

In this section we consider the spin-independent case.

Any rank n tensor on the lattice can be written in a basis of irreducible representations of H_4 as

$$\mathcal{O}_{\mu_1\mu_2\cdots\mu_n} = \sum_{\alpha=1}^{C_{H_4}} \sum_{k=1}^{m_\alpha(n)} \sum_{i=1}^{d_\alpha} \bar{c}(\mu_1\mu_2\cdots\mu_n)_i^{(\alpha,k)}\, \mathcal{O}_i^{(\alpha,k)} \tag{3.63}$$

where $\mathcal{O}_i^{(\alpha,k)}$ transforms as i-th basis vector of k-th copy of the α-th irreducible representation of H_4. d_α is the dimension of the α-th irreducible representation. Label k takes care of the situation when $m_\alpha(n) > 1$. $\bar{c}(\mu_1\mu_2\cdots\mu_n)_i^{(\alpha,k)}$ are the vector components in the basis $\{\mathcal{O}_i^{(\alpha,k)}\}$.

The above equation can be inverted, giving

$$\mathcal{O}_i^{(\alpha,k)} = \sum_{\mu_1,\cdots,\mu_n=1}^{4} c(\mu_1\mu_2\cdots\mu_n)_i^{(\alpha,k)}\, \mathcal{O}_{\mu_1\mu_2\cdots\mu_n}, \tag{3.64}$$

where the $c(\mu_1\mu_2\cdots\mu_n)_i^{(\alpha,k)}$ are the Clebsch-Gordan coefficients.

Let us consider an operator $\mathcal{O}_i^{(\alpha)}$ with all $m_\alpha(n)$ either 0 or 1. Then the spin averaged matrix element $\langle p|\mathcal{O}_i^{(\alpha)}|p\rangle$ can only depend on p_μ, which also transforms as $\mathbf{4}_1$, and $\delta_{\mu\nu}$, as far as the Lorenz structure is concerned. (Hereafter we use a notation $\langle p|\mathcal{O}|p\rangle \equiv \frac{1}{2}\sum_s \langle ps|\mathcal{O}|ps\rangle$.) Furthermore, since $\mathcal{O}_i^{(\alpha)}$ transforms as the i-th unit vector in the α-th irreducible representation, $\langle p|\mathcal{O}_i^{(\alpha)}|p\rangle$ has to be proportional to a tensor made by p_μ that transforms in the same way. The proportionality constant, which is like a reduced matrix element, is the quantity we are after, i.e.,

$$\langle p|\mathcal{O}_i^{(\alpha)}|p\rangle \;=\; M(p^2,\alpha)\,\hat{e}_i^{(\alpha)}(p) \tag{3.65}$$

For example, some of the rank 2 operators have the following matrix elements:

$$\langle p|\mathcal{O}_{i=1}^{(\mathbf{3}_1)}|p\rangle \;\equiv\; \langle p|\frac{1}{\sqrt{12}}\big(3\mathcal{O}_{44} - \mathcal{O}_{11} - \mathcal{O}_{22} - \mathcal{O}_{33}\big)|p\rangle$$

$$= M(p^2, \mathbf{3}_1)\frac{1}{\sqrt{12}}\Big(3p_4p_4 - p_1p_1 - p_2p_2 - p_3p_3\Big) \quad (3.66)$$

$$\langle p|\mathcal{O}^{(\mathbf{6})}_{i=3}|p\rangle \equiv \langle p|\mathcal{O}_{\{14\}}|p\rangle = M(p^2, \mathbf{6})\, p_{\{1}p_{4\}} \quad (3.67)$$

$$\langle p|\mathcal{O}^{(\mathbf{3}_3)}_{i=1}|p\rangle \equiv \langle p|\frac{1}{\sqrt{2}}\Big(\mathcal{O}_{[14]} + \mathcal{O}_{[23]}\Big)|p\rangle = M(p^2, \mathbf{3}_3)\frac{1}{\sqrt{2}}\Big(p_{[1}p_{4]} + p_{[2}p_{3]}\Big) \quad (3.68)$$

The definitions of the basis vectors can be found in the section 3.2. There is no reason for $M(p^2, \mathbf{3}_1)$, $M(p^2, \mathbf{6})$ and $M(p^2, \mathbf{3}_3)$ to be the same on the lattice. However, in the continuum limit when $SO(4)$ symmetry is restored, one expects

$$M(p^2, \mathbf{6}) \xrightarrow{a\to 0} M(p^2, (1,1)))$$
$$M(p^2, \mathbf{3}_1) \xrightarrow{a\to 0} M(p^2, (1,1))),$$

since both $\mathbf{3}_1$ and $\mathbf{6}$ are contained in $(1,1)$ of $SO(4)$. On the other hand, $M(p^2, \mathbf{3}_3)$ could remain different even in the continuum limit, since $\mathbf{3}_3$ is a part of $(1,0)$ in $SO(4)$. Moreover, $\langle p|\mathcal{O}^{(\mathbf{3}_3)}_{i=1}|p\rangle$ will average to zero on the lattice, since $p_{[1}p_{4]} = p_{[2}p_{3]} = 0$.

The situation when $m_\alpha(n) > 1$ is more subtle. The matrix element $\langle p|\mathcal{O}^{(\alpha,k)}_i|p\rangle$ is still proportional to $\hat{e}^{(\alpha,k)}_i(p)$ as required by H_4 symmetry. However, now there are several values of k available and there is nothing to prevent mixing between them. This mixing is only logarithmical, but it has dynamical origin, so label k can not be determined by some arbitrary orthogonalization procedure. It is better to avoid such representations if possible. In the last subsection we will see that in cases we are interested in it is always possible to find H_4 representation with $m_\alpha(n) = 1$.

3.3.2 Spin-Dependent Case

When considering spin-dependent structure functions, there are two vectors, the momentum p_μ and the polarization vector s_μ in addition to the metric tensor $\delta_{\mu\nu}$ from which to construct the tensor structure of $\hat{e}^{(\alpha)}_i = \hat{e}^{(\alpha)}_i(p,s)$. For spin $1/2$ particles like the proton, polarized states can be obtained by applying the polarization operator $P(s) = (1 + \gamma_5 s)/2$ to an unpolarized matrix element

$$\langle ps|\mathcal{O}|ps\rangle = \frac{1}{2}\sum_{s'}\langle ps'|P(s)\mathcal{O}|ps'\rangle \equiv \langle p|P(s)\mathcal{O}|p\rangle, \quad (3.69)$$

which shows that $\hat{e}_i^{(\alpha)}(p,s)$ only contains terms linear in s and independent of s. The s-linear term corresponds to the piece proportional to γ_5 in $\mathcal{O}$.

For example, the rank 3 tensor has following matrix elements

$$\langle ps|\mathcal{O}_{i=3}^{(\mathbf{4}_2)}|ps\rangle = \langle p|P(s)\,\mathcal{O}_{i=3}^{(\mathbf{4}_2)}|p\rangle = \langle p|P(s)\,\mathcal{O}_{\{124\}}|p\rangle = M(p^2,\mathbf{4}_2)\,s_{\{1}p_2p_{4\}} \qquad (3.70)$$

Since the representation $\mathbf{4}_2$ appears only once, $\mathcal{O}^{(\mathbf{4}_2)}$ can not mix with any other rank 3 operators. Whether the matrix element of a rank 3 tensor $\mathcal{O}^{(\mathbf{8},2)}$ can be defined using the equation

$$\langle ps|\mathcal{O}_{i=5}^{(\mathbf{8},2)}|ps\rangle = \langle p|P(s)\,\mathcal{O}_{i=5}^{(\mathbf{8},2)}|p\rangle = \langle p|P(s)\,\mathcal{O}_{[1\{2]4\}}|p\rangle = M(p^2,(\mathbf{8},2))\,s_{[1}p_{\{2]}p_{4\}} \qquad (3.71)$$

depends on whether $\mathcal{O}^{(\mathbf{8},2)}$ mixes with the other **8** representation with the same index pattern ijk (all three indices are distinct from one another). The other three **8**'s with index pattern iij (two indices are equal and the third is different) can not be built by permuting indices in patters ijk, and hence can not mix with $\mathcal{O}^{(\mathbf{8},2)}$.

Constructing tensor structures for other tensors is a simple exercise based on section 3.2.

3.3.3 Possible Lattice Operators

In this subsection we consider criteria for selection of H_4 representations and give a list of possible candidates for lattice calculations of the moments of structure functions.

- First, we want ultimately to connect the lattice Monte-Carlo results with experimental data. Hence, the operators we chose must survive in the continuum limit. This excludes some patterns containing antisymmetrization over positions, e.g., in the spin-independent case none of the representations based of $[ij]\ldots$-like patterns can be used to extract reduced matrix elements, since any tensor structure built out of p_μ is automatically zero and any sensible lattice statistics will average to this dull number. However, such operators could be useful as consistency checks on algorithms and provide a handle on error estimates.

- Second, it is desirable to avoid operators that mix with operators of lower dimension. Usually simple symmetry considerations are enough to determine mixing properties of an operator. When mixing does occur, it is possible sometimes to correct the problem and separate contributions from lower dimensional operators, but normally this is a major undertaking, because it requires subtractions between terms with different powers of inverse lattice scale.

- Operators can also mix within the same dimension. This could happens when there is more than one copy of particular H_4 irreducible representation in $\otimes \mathbf{4}_1$. Such a mixing reflects a bad choice of basis vectors for the representation. Since this mixing has dynamical origin, it is extremely difficult to separate mixed operators on the lattice. We will avoid such operators.

- Statistical noise increases dramatically with momentum of the final state. For this reason it is better to have as few nonzero momentum components as possible.

- When two or more H_4 irreducible representations come from the same $SO(4)$ irreducible representation, they produce in general different results on the lattice. However, these states become degenerate in the continuum limit, so the gap between two values is a measure of lattice effects.

All possible lattice representations that satisfy the first two criteria are listed in Table 3.6, based on Tables 3.3, 3.4 and 3.5, and explicit basis vectors of section 3.2.

Table 3.7 lists various choices of operators for the structure function moments we consider. When more than one choice is possible for the same moment, we use an additional latin index $a, \ldots$. We assume that the nucleon spin is $s = (0, 0, 0, 1)$ where it is relevant and the p_i denote nonzero components of the 4-momentum at the sink. In our case one can choose the smallest nonzero momentum of the lattice $p_i = 2\pi/24$. It is easy to see that different representations within a given rank and a given index pattern involve similar computational complexity.

Logarithmic mixing can occur when the operator contains more than one copy of the same H_4 irreducible representation. To determine if it really occurs for operators

Moment	H_4 representations
a_0	$\mathbf{4}_1$ only
a_2	$(\mathbf{8}, i = 1,2,3)_{B_3}$
d_2	$(\mathbf{8}, i = 1,2)_{C_3}$
t_1	$(\mathbf{3}_3;\ \mathbf{3}_4)_{B_2}$
v_2	$(\mathbf{3}_1)_{A_2}$, $(\mathbf{6})_{B_2}$
v_3	$(\mathbf{8}, i = 1,2,3)_{B_3}$, $(\mathbf{4}_1;\ \mathbf{4}_2;\ \mathbf{8}, i = 1,2)_{C_3}$
v_4	$(\mathbf{2}, i = 1,2,3)_{C_4}$, $(\mathbf{3}_5, i = 1,\cdots,6;\ \mathbf{3}_6, i = 1,\cdots,6)_{D_4}$, $(\mathbf{1}_1;\ \mathbf{1}_2;\ \mathbf{2}, i = 1,2;\ \mathbf{3}_1, i = 1,2,3;\ \mathbf{3}_2, i = 1,2,3)_{E_4}$

Table 3.6: Representations that satisfy the first two criteria.

in table 3.7 let us consider permutations of the indices.

In the case of d_2, the potential candidate of mixing is $(\mathbf{8}, 2)_{C_3}$. However, we see from the explicit basis vectors that the last two indices are antisymmetric and hence the corresponding operator vanishes identically. In the case of v_3, there are three copies of $(\mathbf{8}, i = 1,2,3)_{B_3}$, whose basis vectors can be symmetrized by symmetrizing each term as ($a, b = 1,2,3$ and $a \neq b$),

$$\hat{e}^{(\mathbf{8})}_{B_3} = \Big\{\big(|aab\rangle + |aba\rangle + |baa\rangle\big), \big(2|aab\rangle - |aba\rangle - |baa\rangle\big), \big(|aba\rangle - |baa\rangle\big)\Big\} \quad (3.72)$$

It is easy to check that $\hat{e}^{(\mathbf{8})}_{B_3,2}$ and $\hat{e}^{(\mathbf{8})}_{B_3,3}$ give the same operator. Therefore, the logarithmic mixing in this case is only two-fold. In the case of v_4, there are also three copies of $(\mathbf{2}, i = 1,2,3)_{C_4}$, whose basis vectors can be symmetrized by symmetrizing each term as

$$\begin{aligned}\hat{e}^{(\mathbf{2})}_{C_4} \;=\; &\Big\{\big(|\{aabb\}'\rangle + |\{abba\}'\rangle + |\{abab\}'\rangle\big), \big(2|\{aabb\}'\rangle - |\{abba\}'\rangle - |\{abab\}'\rangle\big), \\ &\big(|\{abba\}'\rangle - |\{abab\}'\rangle\big)\Big\} \end{aligned} \quad (3.73)$$

One can easily check again that $\hat{e}^{(\mathbf{2})}_{C_4,2}$ and $\hat{e}^{(\mathbf{2})}_{C_4,3}$ give the operator zero, so that no mixing occurs in this case. Similarly, $v_{4,a}$ does not have logarithmic mixing either.

Moment	Operator	Basis vector	Mixing
$\mathbf{p} = (0,0,0)$			
$v_{2,b}$	$\mathcal{O}_{44} - \frac{1}{3}\big(\mathcal{O}_{11} + \mathcal{O}_{22} + \mathcal{O}_{33}\big)$	$\hat{e}^{(\mathbf{3}_1)}_{A_2,1}$	–
a_0	$\mathcal{O}^5_1$	$\hat{e}^{(\mathbf{4}_1)}_{2}$	–
$a_{2,a}$	$\mathcal{O}^5_{\{334\}} - \frac{1}{2}\big(\mathcal{O}^5_{\{114\}} + \mathcal{O}^5_{\{224\}}\big)$	$\hat{e}^{(\mathbf{8},\{123\})}_{B_3,7}$	•
t_1	$\mathcal{O}^{5'}_{[24]} - \mathcal{O}^{5'}_{[13]}$	$\hat{e}^{(\mathbf{3}_3)}_{B_2,2}$	–
$\mathbf{p} = (p_1,0,0)$			
$v_{2,a}$	$\mathcal{O}_{\{14\}}$	$\hat{e}^{(\mathbf{6})}_{B_2,3}$	–
v_3	$\mathcal{O}_{\{114\}} - \frac{1}{2}\big(\mathcal{O}_{\{224\}} + \mathcal{O}_{\{334\}}\big)$	$\hat{e}^{(\mathbf{8},\{123\})}_{B_3,7}$	•
v_4	$\mathcal{O}_{\{1144\}} + \mathcal{O}_{\{2233\}} - \mathcal{O}_{\{1133\}} - \mathcal{O}_{\{2244\}}$	$\hat{e}^{(\mathbf{2},\{123\})}_{C_4,2}$	–
a_2	$\mathcal{O}^5_{\{234\}}$	$\hat{e}^{(\mathbf{4}_2)}_{C_3,3}$	–
d_2	$\mathcal{O}^5_{[2\{3]4\}}$	$\hat{e}^{(\mathbf{8},1)}_{C_3,5}$	–
$\mathbf{p} = (p_1,p_2,0)$			
$v_{3,a}$	$\mathcal{O}_{\{124\}}$	$\hat{e}^{(\mathbf{4}_2)}_{C_3,3}$	–
$v_{4,a}$	$\mathcal{O}_{\{2214\}} - \mathcal{O}_{\{3314\}} + \mathcal{O}_{\{1123\}} - \mathcal{O}_{\{4423\}}$	$\hat{e}^{(\mathbf{3}_5,\{123456\})}_{D_4,1}$	–
$\mathbf{p} = (p_1,p_2,p_3)$			
$v_{4,b}$	$\mathcal{O}_{\{1234\}}$	$\hat{e}^{(\mathbf{1}_2)}_{E_4}$	–

Table 3.7: Moments on the lattice

3.4 Renormalization Factors

In this section we review the procedures involved in calculating renormalization constants relating matrix elements in the lattice scheme and the $\overline{\text{MS}}$ scheme. The calculations are done explicitly in the Feynman gauge and in the chiral limit, both in the continuum and on the lattice, for simplicity.

We consider three operators (3.6), (3.11) and (3.10). All three operators are twist-2. $\mathcal{O}^{(f)}$ contributes to the spin-independent structure functions F_1 and F_2. $\mathcal{O}^{(5,f)}$ contributes to the spin-dependent structure functions g_1 and g_2 (corresponding to the longitudinally polarized proton target). The operator $\mathcal{O}^{(5',f)}$, related to the so-called

tensor charge, probes the spin-dependent structure function h_1 (corresponding to the transversely polarized proton target). With a special combination of the first two indices and multiplied by a factor of the current quark mass $\mathcal{O}^{(5'f)}$ also contributes to g_2 at the twist 3 level, though with negligible effect in the chiral limit.

3.4.1 Continuum Calculation

The renormalization scheme used here is the standard $\overline{\text{MS}}$ scheme. There is no operator mixing occuring in the continuum calculation for flavor non-singlet currents.

Feynman Rules

The Feynman rules for the fundamental part are given by Itzykson and Zuber [37]. The Feynman rules (at tree-level) for composite operators are given by

$$\mathcal{O}^{(1)} = \mathcal{O}^{(f)}_{\mu_1,\mu_2,\cdots,\mu_n} \quad : \qquad \not{\Delta}(\Delta\cdot p)^{n-1}, \tag{3.74}$$

$$\mathcal{O}^{(2)} = \mathcal{O}^{(5,f)}_{\mu_1,\mu_2,\cdots,\mu_n} \quad : \qquad \gamma_5\not{\Delta}(\Delta\cdot p)^{n-1}, \tag{3.75}$$

$$\mathcal{O}^{(3)} = \mathcal{O}^{(5',f)}_{\mu,\mu_1,\cdots,\mu_n} \quad : \qquad \gamma_\mu\not{\Delta}(\Delta\cdot p)^{n-1}, \tag{3.76}$$

where Δ is a null vector ($\Delta^2 = 0$) which enforces the symmetrization and removes traces. The Feynman rules for operators involving gluons can be found in the book by Muta [38].

In the chiral limit, γ_5 can be freely commuted, with at most an overall sign difference. Therefore, the calculations for (3.74) and (3.75) are practically identical.

Procedure

The steps involved are the following:

(a) draw all possible Feynman graphs up to an appropriate order in the number of loops (one-loop in our case), including symmetry factors;

(b) for each graph, first simplify the color and Dirac structures;

(c) carry out the loop integral and then remove the term with $(\epsilon^{-1} - \gamma_E + \ln 4\pi)$;

(d) add up all graphs and then factor out the tree-level contribution explicitly.

Results

The results (by B. Schreiber) are summarized in Table 3.8, graph by graph including symmetry factors. The notation is defined as follows:

$$Z_G = 1 + \frac{g^2(\mu)}{16\pi^2} C_F \left[\gamma_G \ln(p^2/\mu^2) + B_G\right] . \tag{3.77}$$

The sails and vertex graphs for the second moment are also checked by S. Huang.

Notice that there are also finite terms with structures different from that of the tree-level contribution. These finite terms should also be reproduced on the lattice and hence do not require additional renormalization for the composite operators.

Moment	G	γ_G	B_G
	sails	2	-4
first	vertex	-1/3	5/9
	self-energy	1	-1
	total	8/3	-40/9
	sails	10/3	-62/9
second	vertex	-1/6	4/9
	self-energy	1	-1
	total	25/6	-67/9
	sails	13/3	-83/9
third	vertex	-1/10	28/75
	self-energy	1	-1
	total	157/30	-2216/225

Table 3.8: Results in the $\overline{\text{MS}}$ scheme for the first three moments for operators $\mathcal{O}^{(1)}$ and $\mathcal{O}^{(2)}$. Spin-independent and spin-dependent results are the same in the chiral limit.

3.4.2 Lattice Calculation

The lattice operators are given by the same expressions as in the continuum, with the derivatives replaced by the lattice expressions (2.7) and (2.6). Operator mixing can be completely avoided when the indices $(\mu_1, \cdots, \mu_n)$ are chosen to be distinct from each other (possible up to the fourth moment). However, other choices are also available, with a proper combination of indices according to representations of the H_4 group.

Feynman Rules

The Feynman rules for the fundamental part are given by Kawai et al in [39]. The Feynman rules for composite operators are given by Kronfeld and Photiadis in [40]. and checked by B. Schreiber. Those rules relevant to our case are listed here (Wilson fermions at $r = 1$ in Feynman gauge $\alpha = 1$):

quark propagator:

$$S_F(k) \equiv \delta_{ab} \left(\frac{1}{a} \sum_\mu i\gamma_\mu \sin(k_\mu a) + M(k) \right)^{-1} \quad \text{where} \quad M(k) = \frac{1}{a} \sum_\mu \Big[1 - \cos(k_\mu a) \Big] ; \tag{3.78}$$

gluon propagator:

$$S_B(k) \equiv \delta_{AB}\, \delta_{\mu\nu} \left(\frac{2}{a^2} \sum_\rho \Big[1 - \cos(k_\rho a) \Big] \right)^{-1} ; \tag{3.79}$$

quark-antiquark-gluon vertex:

$$-ig_0 \, (T^A)_{bc} \left(\gamma_\mu \cos[(p+q)_\mu a/2] - i \sin[(p+q)_\mu a/2] \right) ; \tag{3.80}$$

quark-antiquark-gluon-gluon vertex:

$$-\frac{a}{2} g_0^2 \delta_{\mu\nu} \, \{T^A, T^B\}_{ab} \left(\cos[(p+q)_\mu a/2] - i\gamma_\mu \sin[(p+q)_\mu a/2] \right) ; \tag{3.81}$$

operator with two-quark:

$$\delta_{ab}\, \gamma_{\mu_1} \prod_{l=2}^{n} \frac{\sin(p_{\mu_l} a)}{a} ; \tag{3.82}$$

two quarks, one gluon vertex:

$$ag_0(T^A)_{ab}\,\gamma_{\mu_1}\sum_{r=2}^{n}\delta_{\mu\mu_r}\cos[(p+q)_{\mu_r}a/2]\prod_{l=2}^{r-1}\frac{\sin(p_{\mu_l}a)}{a}\prod_{l=r+1}^{n}\frac{\sin(q_{\mu_l}a)}{a}\,; \tag{3.83}$$

two quarks, two gluons vertex:

$$-\frac{1}{2}a^2g_0^2\{T^A,T^B\}_{ab}\,\gamma_{\mu_1}\sum_{r=2}^{n}\delta_{\mu\mu_r}\delta_{\nu\mu_r}\frac{\sin[(p+q)_{\mu_r}a/2]}{a}\prod_{l=2}^{r-1}\frac{\sin(p_{\mu_l}a)}{a}\prod_{l=r+1}^{n}\frac{\sin(q_{\mu_l}a)}{a}\,. \tag{3.84}$$

The composite operators are assumed to carry zero momentum here. Rules for $\mathcal{O}^{(5,f)}$ and $\mathcal{O}^{(5',f)}$ are similar to that of $\mathcal{O}^{(f)}$, with additional Dirac matrices.

Procedure

The steps involved are essentially the same as in the continuum case (specialties on the lattice are noted):

(a) draw all possible Feynman graphs up to an appropriate order in the number of loops (one-loop in our case), including symmetry factors; (there are more graphs on the lattice due to the compact formulation of gauge fields.)

(b) for each graph, first simplify the color and Dirac structures; (the color part is similar to the continuum case whereas the Dirac structure is rather different due to the explicit violation of the chiral symmetry by the Wilson action.)

(c) carry out the lattice integrals; (There are two ways of doing these integrals, one is simpler but involves fitting, used by B. Schreiber, and the other is slightly more complicated but involves no fitting, used by S. Huang.)

(d) add up all graphs and then factor out the tree-level contribution explicitly at low external momenta in lattice units.

Results

The results (by B. Schreiber) are summarized in Tables 3.9 and 3.10, graph by graph including symmetry factors. The notation are defined as follows:

$$Z_G = 1+\frac{g_0^2}{16\pi^2}C_F\left[\gamma_G\ln(p^2a^2)+B_G\right]. \tag{3.85}$$

The sails graph of the second spin-independent moment has also been checked by S. Huang. In addition, the results for the first and second spin-independent moments agree with Capitani and Rossi [41]

The result for the second spin-dependent moment agrees with Schierholz et al [28].

Moment	G	γ_G	B_G
	sails	2	-5.070
	vertex	-1/3	2.292
first	self-energy	1	-0.382
	tadpole(quark)	0	12.233
	tadpole(operator)	0	-12.233
	total	8/3	-3.160[1]
	sails	10/3	-7.608
	vertex	-1/6	1.211
second	self-energy	1	-0.382
	tadpole(quark)	0	12.233
	tadpole(operator)	0	-24.466
	total	25/6	-19.012[2]
	sails	13/3	-9.198
	vertex	-1/10	0.840
third	self-energy	1	-0.382
	tadpole(quark)	0	12.233
	tadpole(operator)	0	-36.699
	total	157/30	-33.206

Table 3.9: Results in the lattice scheme, first three spin-independent moments for the operators $\mathcal{O}^{(1)}$.

The results for the zeroth moment (or tensor charge) from $\mathcal{O}^{(3)}_{\mu\nu}$, for both $\overline{\text{MS}}$ and lattice schemes, are given in Table 3.11.

Notice that, in additional to the finite terms also occuring in the continuum, there

Moment	G	γ_G	B_G
	sails	2	-5.070
	vertex	-1/3	1.358
first	self-energy	1	-0.382
	tadpole(quark)	0	12.233
	tadpole(operator)	0	-12.233
	total	8/3	4.094
	sails	10/3	-7.608
	vertex	-1/6	0.661
second	self-energy	1	-0.382
	tadpole(quark)	0	12.233
	tadpole(operator)	0	-24.466
	total	25/6	-19.562[3]
	sails	13/3	-9.198
	vertex	-1/10	0.464
third	self-energy	1	-0.382
	tadpole(quark)	0	12.233
	tadpole(operator)	0	-36.699
	total	157/30	-33.582

Table 3.10: Results in the lattice scheme, first three spin-dependent moments for operators $\mathcal{O}^{(2)}$.

is a linearly divergent piece (a^{-1}) in the quark self-energy graph. This divergence should be taken care of by the quark mass renormalization (through fine-tuning of the hopping parameter in the Wilson fermion) in the fundamental part of the theory and hence should not be regarded as part of the composite operator renormalization.

Scheme	G	γ_G	B_G
	vertex	0	0
$\overline{\text{MS}}$	self-energy	1	-1
	total	1	-1
	vertex	0	4.386
	self-energy	1	-0.382
lattice	tadpole(quark)	0	12.233
	total	1	16.237

Table 3.11: Results in the $\overline{\text{MS}}$ scheme and the lattice scheme for the zeroth moment (or tensor charge) $\mathcal{O}^{(3)}_{\mu\nu}$.

3.5 Extracting Data from the Lattice

The expectation value of $\langle T[J(t_1)\mathcal{O}(t_2)\bar{J}(t_3)]\rangle$ at large separations $|t_1 - t_2| \gg 1$, $|t_1 - t_3| \gg 1$, $|t_2 - t_3| \gg 1$ is:

$$\langle T[J(t_1)\mathcal{O}(t_2)\bar{J}(t_3)]\rangle = |<0|J|p>|^2 e^{-E_0(p)|t_1-t_3|}\langle\mathcal{O}\rangle.$$

The dependences of the nucleon source and the mass can be factored out by dividing the above expression by $D_2(t_1, t_3)$:

$$\frac{\langle T[J(t_1)\mathcal{O}(t_2)\bar{J}(t_3)]\rangle}{\langle T[J(t_1)\bar{J}(t_3)]\rangle} = \langle\mathcal{O}\rangle. \tag{3.86}$$

Below we show how the numerator in eq. (3.86) can be computed on the lattice.

The hadron current is $J_\alpha(x) = u^a_\alpha(x)u^b_\beta(x)d^c_\gamma(x)\epsilon_{abc}\Gamma^{\beta\gamma}$. In our calculations we use $\Gamma = \Gamma^C$. In the smeared sources, the quark operators are integrated with a Gaussian smearing function.

We want to calculate the matrix element $\langle 0|J_\alpha(x)\mathcal{O}(y,y')\bar{J}_{\alpha'}(x')|0\rangle$, where $\mathcal{O}$ is bilinear in quarks: $\mathcal{O}(y,y') = \overline{\psi}^n_\nu(y)O^{\nu\nu'}_{nn'}\psi^{n'}_{\nu'}(y')$ and $O^{\nu\nu'}_{nn'}$ in general carries other Lorenz indices as well. If $\mathcal{O}(y,y')$ is not flavor changing, then it can be represented as $\mathcal{O}(y,y') = \overline{u}^n_\nu(y)^u O^{\nu\nu'}_{nn'}u^{n'}_{\nu'}(y') + \overline{d}^n_\nu(y)^d O^{\nu\nu'}_{nn'}d^{n'}_{\nu'}(y')$. Otherwise let us leave O unspecified.

This way we will be able to calculate several matrix elements at very low incremental cost.

Our goal is to calculate the matrix element

$$\langle J_\alpha(x)\mathcal{O}(y,y')\bar{J}_{\alpha'}(x')\rangle = \langle 0|u^a_\alpha(x)u^b_\beta(x)d^c_\gamma(x)\Gamma^{\beta\gamma}\epsilon_{abc}\overline{\psi}^{d'}_{\nu'}(y')O^{\nu'\nu}_{d'd}(y',y)\psi^d_\nu(y)$$
$$\overline{d}^{c'}_{\gamma'}(x')\overline{u}^{b'}_{\beta'}(x')\overline{u}^{a'}_{\alpha'}(x')\overline{\Gamma}^{\beta'\gamma'}\epsilon_{a'b'c'}|0\rangle.$$

Hereafter integration over y, y' and x is assumed. There is a momentum projection in $\int d^3x$.

Let us consider ${}^u\mathcal{O}$ first. There are four diagrams contributing to this matrix element. After contractions we have

$$\begin{aligned}\langle {}^u\mathcal{O}\rangle = \; & S^{ad'}_{\alpha\nu'}(x,y')\tilde{S}^{da'}_{\nu\alpha'}(y,x')\tilde{S}^{bb'}_{\beta\beta'}(x,x')\tilde{S}^{cc'}_{\gamma\gamma'}(x,x')O^{\nu'\nu}_{d'd}\Gamma^{\beta\gamma}\overline{\Gamma}^{\beta'\gamma'}\epsilon_{abc}\epsilon_{a'b'c'} \\ & -S^{ad'}_{\alpha\nu'}(x,y')\tilde{S}^{db'}_{\nu\beta'}(y,x')\tilde{S}^{ba'}_{\beta\alpha'}(x,x')\tilde{S}^{cc'}_{\gamma\gamma'}(x,x')O^{\nu'\nu}_{d'd}\Gamma^{\beta\gamma}\overline{\Gamma}^{\beta'\gamma'}\epsilon_{abc}\epsilon_{a'b'c'} \\ & +S^{bd'}_{\beta\nu'}(x,y')\tilde{S}^{db'}_{\nu\beta'}(y,x')\tilde{S}^{aa'}_{\alpha\alpha'}(x,x')\tilde{S}^{cc'}_{\gamma\gamma'}(x,x')O^{\nu'\nu}_{d'd}\Gamma^{\beta\gamma}\overline{\Gamma}^{\beta'\gamma'}\epsilon_{abc}\epsilon_{a'b'c'} \\ & -S^{bd'}_{\beta\nu'}(x,y')\tilde{S}^{da'}_{\nu\alpha'}(y,x')\tilde{S}^{ab'}_{\alpha\beta'}(x,x')\tilde{S}^{cc'}_{\gamma\gamma'}(x,x')O^{\nu'\nu}_{d'd}\Gamma^{\beta\gamma}\overline{\Gamma}^{\beta'\gamma'}\epsilon_{abc}\epsilon_{a'b'c'}\end{aligned}$$

One can rewrite this expression as

$$\langle {}^u\mathcal{O}\rangle = S^{ad'}_{\mu\nu'}(x,y')\tilde{S}^{da'}_{\nu\mu'}(y,x')O^{\nu\nu'}_{dd'}(y,y')\,{}^uM^{\mu\mu'\alpha\alpha'}_{aa'}(x,x'),$$

where

$${}^uM^{\mu\mu'\alpha\alpha'}_{aa'}(x,x') = \tilde{S}^{cc'}_{\gamma\gamma'}(x,x')[\;\delta^{\mu\alpha}\delta^{\mu'\alpha'}\tilde{S}^{bb'}_{\beta\beta'}(x,x')\Gamma^{\beta\gamma}\bar{\Gamma}^{\beta'\gamma'}$$
$$+\delta^{\mu\alpha}\tilde{S}^{bb'}_{\beta\alpha'}(x,x')\Gamma^{\beta\gamma}\bar{\Gamma}^{\mu'\gamma'}$$
$$+\tilde{S}^{bb'}_{\alpha\alpha'}(x,x')\Gamma^{\mu\gamma}\bar{\Gamma}^{\mu'\gamma'}$$
$$+\delta^{\alpha'\mu'}\tilde{S}^{bb'}_{\alpha\beta'}(x,x')\Gamma^{\mu\gamma}\bar{\Gamma}^{\beta'\gamma'}]\epsilon_{abc}\epsilon_{a'b'c'}.$$

Now let us calculate ${}^d\mathcal{O}$. There are only two diagrams which give:

$$\begin{aligned}\langle {}^d\mathcal{O}\rangle \; & = S^{cd'}_{\gamma\nu'}(x,y')\tilde{S}^{dc'}_{\nu\gamma'}(y,x')\tilde{S}^{aa'}_{\alpha\alpha'}(x,x')\tilde{S}^{bb'}_{\beta\beta'}(x,x')O^{\nu'\nu}_{d'd}\Gamma^{\beta\gamma}\bar{\Gamma}^{\beta'\gamma'}\epsilon_{abc}\epsilon_{a'b'c'} \\ & -S^{cd'}_{\gamma\nu'}(x,y')\tilde{S}^{dc'}_{\nu\gamma'}(y,x')\tilde{S}^{ab'}_{\alpha\beta'}(x,x')\tilde{S}^{ba'}_{\beta\alpha'}(x,x')O^{\nu'\nu}_{d'd}\Gamma^{\beta\gamma}\bar{\Gamma}^{\beta'\gamma'}\epsilon_{abc}\epsilon_{a'b'c'}\end{aligned}$$

or,

$$\langle {}^d\mathcal{O}\rangle = S^{cd'}_{\mu\nu'}(x,y')\tilde{S}^{dc'}_{\nu\mu'}(y,x')O^{\nu\nu'}_{dd'}(y,y')\,{}^dM^{\mu\mu'\alpha\alpha'}_{cc'}(x,x'),$$

where

$$
\begin{aligned}
{}^{d}M_{cc'}^{\mu\mu'\alpha\alpha'}(x,x') = [\quad & \tilde{S}_{\alpha\alpha'}^{aa'}(x,x')\tilde{S}_{\beta\beta'}^{bb'}(x,x') \\
& +\tilde{S}_{\alpha\beta'}^{aa'}(x,x')\tilde{S}_{\beta\alpha'}^{bb'}(x,x')]\Gamma^{\beta\mu}\bar{\Gamma}^{\beta'\mu'}\epsilon_{abc}\epsilon_{a'b'c'}
\end{aligned}
$$

One can easily check that for $\Gamma = C\gamma_5$ both ${}^{u}M$ and ${}^{d}M$ have nonzero entries only for $\alpha, \alpha' = 1, 2$.

3.6 Sequential Source

To calculate $S_{\mu\nu}^{ab}(x,y')$ using the existing CG code, one needs to use the following trick. First, it's easy to check that $\gamma_5^{\alpha\beta}S_{\beta\delta}^{ab}(x,y)\gamma_5^{\delta\eta} = \left(S_{\eta\alpha}^{ba}\right)^*(y,x)$. Now, the sequential source can be computed using the following identity:

$$
\begin{aligned}
S_{\mu\nu'}^{cd'}(x,y')M_{cc'}^{\mu\mu'\alpha\alpha'}(x,x') &= \gamma_5^{\mu\sigma}\left(S_{\sigma'\sigma}^{d'c}\right)^*(y',x)\gamma_5^{\sigma'\nu'}M_{cc'}^{\mu\mu'\alpha\alpha}(x,x') \\
&= \gamma_5^{\nu'\sigma'}\left(S_{\sigma'\sigma}^{d'c}(y',x)\left(\gamma_5^{\sigma\mu}M_{cc'}^{\mu\mu'\alpha\alpha}(x,x')\right)^*\right)^*
\end{aligned}
$$

The product $S_{\sigma'\sigma}^{d'c}(y',x)\left(\gamma_5^{\sigma\mu}M_{cc'}^{\mu\mu'\alpha\alpha}(x,x')\right)^*$ is just a solution of the Dirac equation

$$
\gamma_\mu^{\alpha\beta}D_{ab}^{\mu}(x)S_{\beta\gamma[\mu\mu']}^{bc}(x,y) = \Psi_{\alpha\gamma[\mu\mu']}^{ac}(x,y) \quad \Psi_{\alpha\beta[\nu\nu']}^{ab} = \left(\gamma_5^{\alpha\mu}M_{ab}^{\mu\beta\nu\nu'}\right)^*.
$$

If we denote

$$
{}^{x}S_{\alpha\beta[\nu\nu']}^{ab}(y',x') = S_{\alpha\gamma}^{ac}(y',x)\left(\gamma_5^{\gamma\delta}\,{}^{x}M_{cb}^{\delta\beta\nu\nu'}(x,x')\right)^*, \quad x = u,d
$$

then for the three point functions we have:

$$
\begin{aligned}
\langle {}^{u}O\rangle &= \tilde{S}_{\nu\mu'}^{da'}(y,x')\,{}^{u}O_{dd'}^{\nu\nu'}(y,y')\gamma_5^{\nu'\sigma'}\left({}^{u}S_{\sigma'\mu'[\alpha\alpha']}^{d'a'}(y',x')\right)^* \\
\left\langle {}^{d}O\right\rangle &= \tilde{S}_{\nu\mu'}^{dc'}(y,x')\,{}^{d}O_{dd'}^{\nu\nu'}(y,y')\gamma_5^{\nu'\sigma'}\left({}^{d}S_{\sigma'\mu'[\alpha\alpha']}^{d'c'}(y',x')\right)^*
\end{aligned}
$$

These results are obtained for smeared source—point sink three-point functions. The case of the smeared sink can be worked out along the same lines simply by adding smearing to all quark propagators reaching the point x.

3.7 Results

We illustrate results of this chapter with measurements of the zeroth moment of the tensor charge. We analyzed 40, 40 and 36 configurations for three κ's we used in

determining the lattice scale in section 2.7 with the final momentum $\mathbf{p} = (0,0,0)$. We also analyzed from 15 to 20 of the same configurations with $\mathbf{p} = (1,1,0)$. In all cases the Gaussian smeared source has been placed at $t_1 = 8$ with the momentum projected point sink at $t = 24$. Figure 3-1 shows results for $\mathbf{p} = (0,0,0)$. For all κs the plateaux are clearly visible.

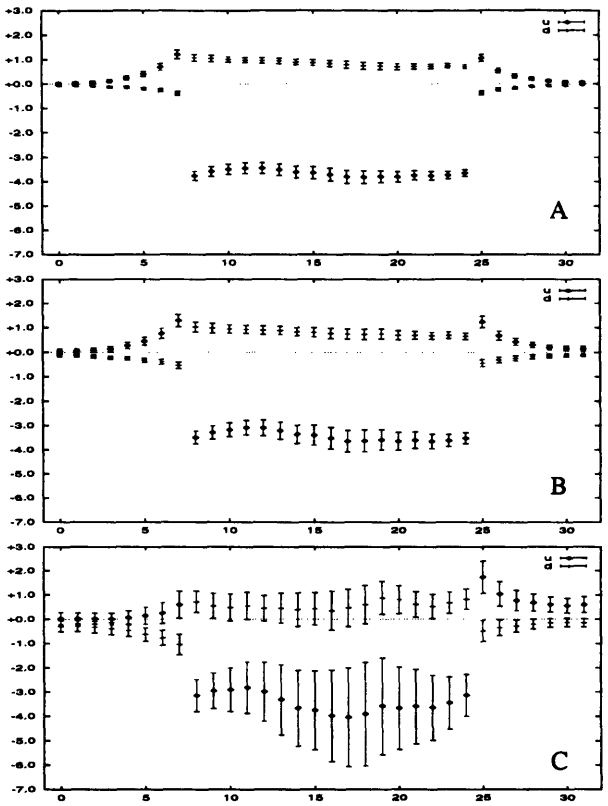

Figure 3-1: u and d contributions to t_1 for $\mathbf{p} = (0,0,0)$, $\kappa = 0.15200$(A),0.15246 (B), 0.15294 (C)

Since other moments require non-zero momentum at the sink, we examined the possibility of using a single set of propagators with non-zero momenta to measure all the observables. A comparison of results using $\mathbf{p} = (0,0,0)$ and $\mathbf{p} = (1,1,0)$ is shown in fig. 3-2 for $\kappa = 0.15200$, where we used 15 configurations for both $\mathbf{p}$'s. The result shows that one must put as many components of $\mathbf{p}$ to zero as possible, even at the expense of having to calculate additional sets of propagators.

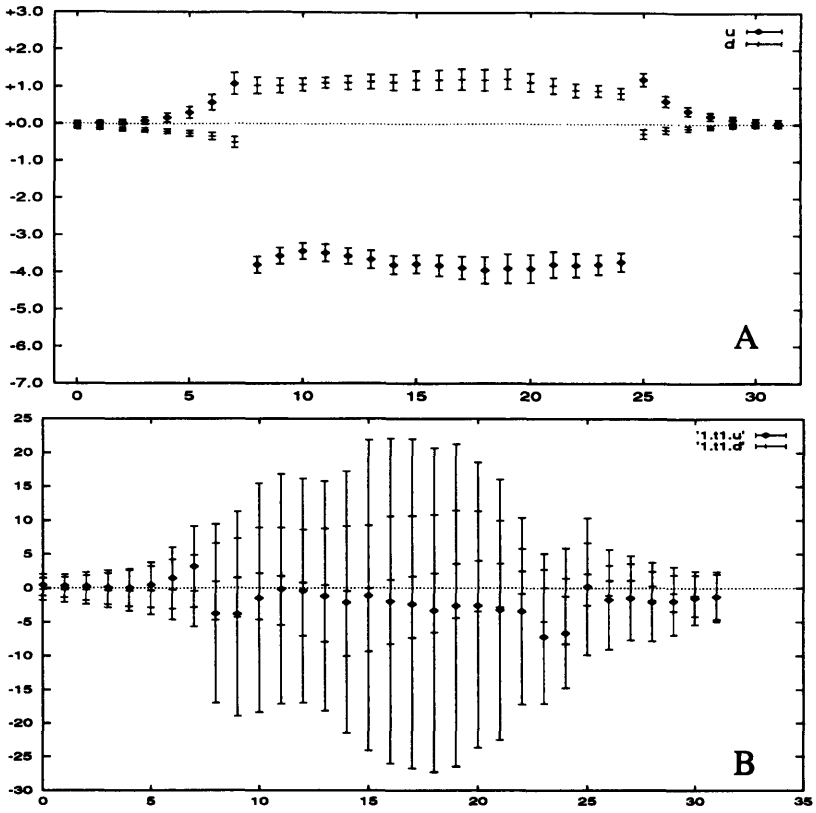

Figure 3-2: Comparison of t_1 calculated with $\mathbf{p} = (0,0,0)$ (A) and $\mathbf{p} = (1,1,0)$ (B) at $\kappa = 0.15200$. Note the different vertical scales in A and B.

Here we used the jackknife method (see appendix B) to estimate statistical errors on the ratio of three- and two-point functions and the procedure from appendix C to find the plateaux. Systematic errors due to finite lattice size and the quenched approximation are unknown at this point.

Finding the plateaux in the plots on fig. 3-1 with the χ^2 procedure one gets obtains m_q dependence of t_1 depicted on fig. 3-3. Extrapolating to the chiral limit, the contribution of the u-quark is $-3.28(2)$ before renormalization, while the d-quark gives 0.514(3). Note that here we have tacitly assumed that the nucleon in question is a proton. More properly one should call the two valence quarks of the same flavor the *like* quarks and the third quark of different flavor the *unlike* quark.)

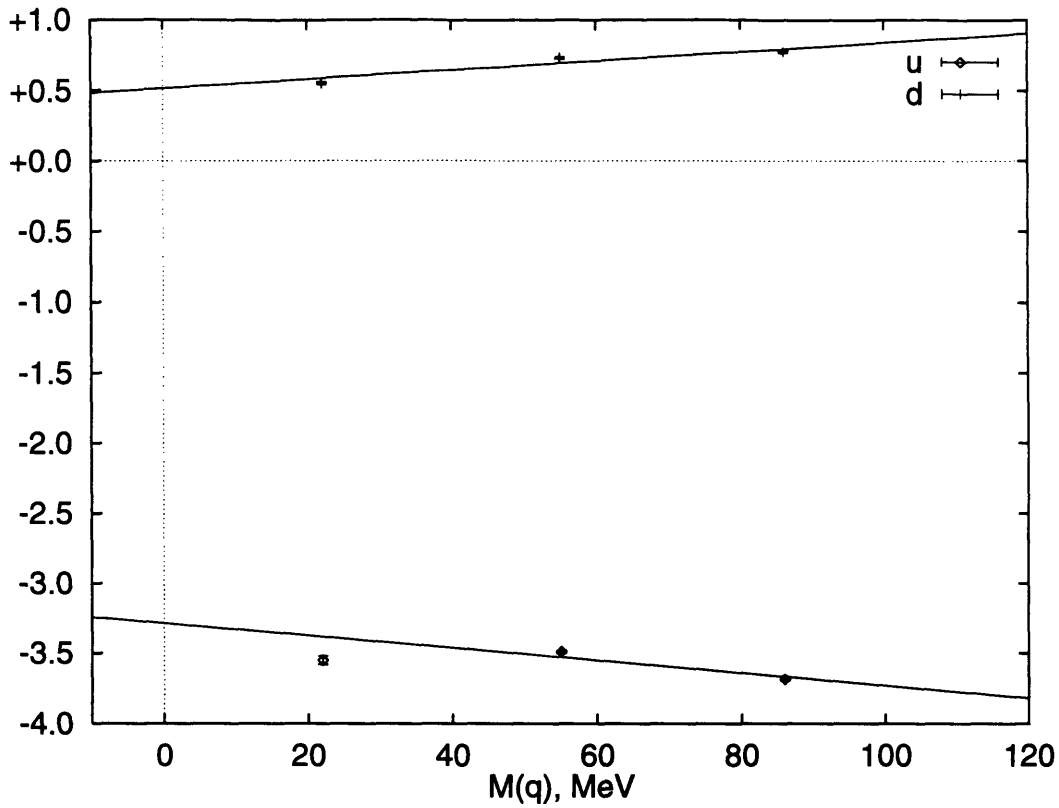

Figure 3-3: Extrapolating t_1 to the chiral limit

3.8 Conclusion

In conclusion, we would like to outline possible future developments:

- Going to higher moments could be done in two ways: *i)* using a spatially asymmetric lattice to put a very small momentum in the final state or *ii)* building the lattice derivative operators designed to decrease the noise level in the higher moments. Of course, these two approaches could be combined as well as used in conjunction with improved actions (see below.)

- Using improved (or perfect) actions for the gauge sector and the Dirac operator is an attractive way to improve approximation to the continuum.

- Understanding effects of dynamical quarks on the structure functions requires calculating the moments in full QCD. While the body of results in the quenched

approximation is growing, currently there is no results in the unquenched theory.

- Calculating renormalization constants nonperturbatively could be an attractive method, especially if a perfect action is used.
- Calculating moments of other structure functions would help in looking for chiral and infinite lattice limits by comparing with result of other groups [42, 43].

Chapter 4

Instantons on the Lattice

Topological properties are among the most fascinating aspects of the non Abelian gauge theories. Their mere nature calls for nonperturbative methods. It seems that the lattice is an ideal tool to study instantons and related phenomena, but the discreteness of lattice makes it impossible to rigorously introduce topology into the lattice formulation. Instead one has to resort to various tricks and approximations. In this chapter, we review the problems of topology studies on the lattice in section 4.1. In section 4.2, several definitions of the topological charge which are widely used in QCD are reviewed. In order to better understand topology, we turn to a simpler model, the $2d$ $O(3)$ σ-model, and show in section 4.3 that it shares the same set of problems. In the last section, we describe a definition of the topological charge which will be used with the cluster algorithm in the next chapter to study θ-vacua.

4.1 θ Terms

Most of the studies of topology start with introducing the θ-term into the action

$$S[\phi, \theta] = S[\phi] + i\theta Q[\phi], \tag{4.1}$$

and calculating the θ dependence of observables. Here Q is the topological charge. It has integer values on any field configuration $[\phi]$ and the configuration space of the continuum theory consists of disjointed sectors with fixed Q. By disjointness we mean

that if $[\phi_1]$ and $[\phi_2]$ have $Q_1 \neq Q_2$ then there is no path $[\phi]_t$ connecting them such that the action is finite for all t.

Introduction of the θ-term allows one to measure topology–related quantities. E.g., the topological susceptibility $\chi_t = \langle Q^2 \rangle / V$ is

$$\chi_t = \frac{\partial^2 Z}{Z V \partial \theta^2}.$$

From the simulation point of view, the action (4.1) is a nightmare. The major problem is that now $\exp(-S[\phi, \theta])$ is not positive definite and one can not interpret it as a probability. If one tries to solve this problem by grouping the factor $\exp(i\theta Q)$ with an observable A then there is a rapidly oscillating integrand $A \exp(i\theta Q)$ and the signal must be extracted from cancelation of negative and positive contributions. To this end a tremendous increase in statistics is required.

Several additional complications are common when dealing with the topological charge on the lattice. Among them

- Strictly speaking, there is no topological structure in the lattice field theory, since any configuration is finite action on the lattice and two field configurations can always be smoothly transformed to each other. The situation here is analogous to that with the Lorenz symmetry—it is explicitly broken on the lattice but is restored in the continuum limit.

- Q takes integer values in the continuum only. Naïve lattice transcriptions easily loose this property.

- Mixing with other operators, notably unity, and large renormalization constants are common for most lattice definitions of Q [44, 45].

There are several ways to avoid these problems. We consider them in the following section.

4.2 Topological Charge in LQCD

The continuum charge density

$$q(x) = \frac{g^2}{64\pi^2} F^a_{\mu\nu}(x) F^a_{\rho\sigma}(x) \epsilon^{\mu\nu\rho\sigma} \tag{4.2}$$

gives the second Chern number of the principle bundle

$$C_2[A] = \int d^4x q(x). \tag{4.3}$$

There are three approaches to transfer of Eq. (4.3) to the lattice. First, one can construct a lattice expression which in the continuum limit goes to (4.2). Second, one can try to build a lattice analog for (4.3) without bothering with local expressions. Third, if one is able to reconstruct the principle bundle from its lattice footprint, then one can have a well-behaved C_2 and, with a bit of luck, the local density.

4.2.1 Quick and Dirty Approach

Basically, there are two expressions in the unit hypercube for $q(x)$ which approach eq. (4.2) in the continuum, corresponding to two paths on fig. 4-1. Alternatively, one

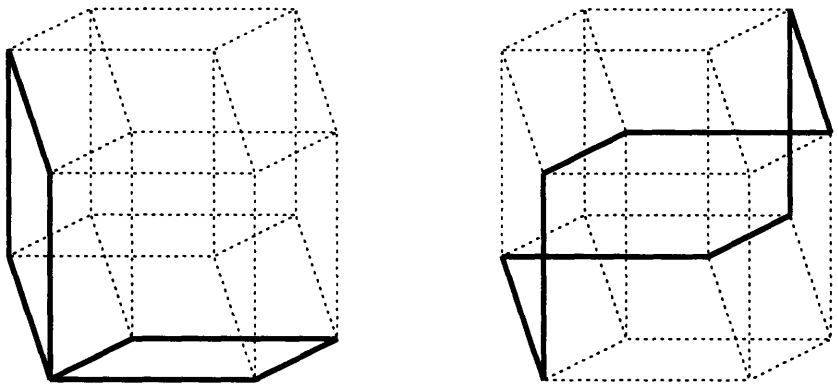

Figure 4-1: Topological charge loops

can combine these two expression to either cancel the next order in a corrections, or to suppress mixing. Being simplest, this approach suffers greatly both from mixing and renormalization.

However, these definitions are computationally cheap and are often used with cooling when their drawbacks become less prominent.

4.2.2 Topologically Correct LQCD

Here we only briefly mention two methods that give integer values for (4.3). The first was suggested by M. Lüscher [46]. It produces an integer charge for small action density and Q reduces to the Chern number in the classical continuum limit. Besides it gives a correct continuum charge density (4.2). The only drawback of this definition is its computational cost. One must reconstruct a parallel transport and calculate several 2-d integrals on every lattice site. Often double precision is required.

Another method [47] goes after the total charge (4.3) only. After extensive analytical construction it amounts to counting intersection numbers on the group manifold. The original formulation works with a simplicial lattice.

Depending on the form of the gauge action, geometrical definitions can have severe problems. For example, in the case of the Wilson action they suffer from dislocations [48]. Besides, for some configuration the geometrical charge is ill-defined, but this is a minor price for having it integer otherwise. However, these exceptional configurations could be made to have measure zero.

4.3 Topology in the σ-Model

In general the topological structure of gauge theories is very complicated and difficult to study on the lattice. To better understand the rôle topology plays in the dynamics of a field theory, we now turn to the two-dimensional $O(3)$ σ-model.

Two-dimensional σ-models are traditionally used as a warm-up exercise for non Abelian gauge theories [49]. In the continuum, the topological charge is defined as

$$Q = \int d^2x q(x), \quad q(x) = \frac{1}{8\pi}\epsilon^{abc}\epsilon^{\mu\nu}n^a\partial^\mu n^b\partial^\nu n^c. \tag{4.4}$$

Q has clear geometrical meaning. It counts the winding number of the mapping $n(x) : \mathcal{M} \to S^2$. Eq. (4.4) looks deceptively simple and easy for lattice calculations,

however, one can find in it all the pathologies of $Q_{\rm QCD}$.

A naïve discretization of (4.4) does not produce integers for all configurations and needs considerable renormalization which must be calculated nonperturbatively. In the next two subsections we consider two definitions which always (modulo a set of configurations of measure zero) give integer values for Q on the lattice.

It is more convenient to work with a triangular lattice here. For more details on the lattice layout see the next chapter.

4.3.1 Area Counting Definition

Since $q(x)$ in (4.4) is nothing but the area of an infinitesimal triangle, one can use the area covered by triangular plaquettes $\Delta_x = (n_1, n_2, n_3)$ on the lattice to calculate Q. The surface of a spherical triangle is $S = \varphi_1 + \varphi_2 + \varphi_3 - \pi$. Here φ_i is the i-th dihedral angle. When $\varphi_i < \pi/2$ for all i, one can easily rewrite S in terms of scalar products (n_i, n_j):

$$S = -\arccos\left(\frac{w_3 + w_1 + w_2 + w_3^2 + w_1^2 + w_2^2 + w_3 w_1 + w_1 w_2 + w_2 w_3 - w_3 w_1 w_2}{1 + w_3 + w_1 + w_2 + w_3 w_1 + w_1 w_2 + w_2 w_3 + w_3 w_1 w_2}\right),$$

where $w_1 = (n_2, n_3)$, $w_2 = (n_3, n_1)$ and $w_3 = (n_1, n_2)$. For large triangles (when some of $\varphi_i > \pi/2$) one should be more careful with signs. There is also an alternative set of formulae for ϕ_i:

$$\begin{aligned}
\phi_1 &= \arccos\frac{w_1 - w_2 w_3}{\sqrt{(1 - w_2^2)(1 - w_3^2)}} \\
\phi_2 &= \arccos\frac{w_2 - w_1 w_2}{\sqrt{(1 - w_1^2)(1 - w_3^2)}} \\
\phi_3 &= \arccos\frac{w_3 - w_1 w_2}{\sqrt{(1 - w_1^2)(1 - w_2^2)}} \\
S &= \phi_1 + \phi_2 + \phi_3 - \pi.
\end{aligned}$$

This also shows where the deficiency of the naïve definition comes from. The area of Δ_x is proportional to the volume of intersection of the unit solid sphere and a cone subtended by Δ_x while $n_1 \cdot (n_2 \times n_3)$ is proportional to the volume of the *tetrahedron* defined by points $0, n_1, n_2, n_3$. While both have the same continuum limit they differ at finite a.

While the final value of Q is an exact multiple of 1, it is obtained as a sum of all areas and as such is subject to rounding errors. The same is true for Lüscher's definition of Q in QCD. In addition, like QCD, this topologically correct definition of Q is much more computationally expensive.

4.3.2 Counting Triangles

It is possible to design yet another method to calculate the same geometrical definition of the topological charge which does not involve global floating point sums. This could be done if the local charge density were of no interest. The method is based on the following observation.

The winding number of the map $n(x) : \mathcal{M} \to S^2$ can be calculated by counting how may points of $\mathcal{M}$ are mapped to the same point w in S^2. In the case of a triangular plaquette, it amounts to comparing signs of a few 3×3 determinants. If w is outside of the image of the triangle $\Delta_{\mathcal{M}}$, Δ, then the associated number is 0. If w is inside, the number is $+1$ if $n : \Delta_{\mathcal{M}} \to \Delta$ preserves the orientation and -1 if it does not, see fig. 4-2

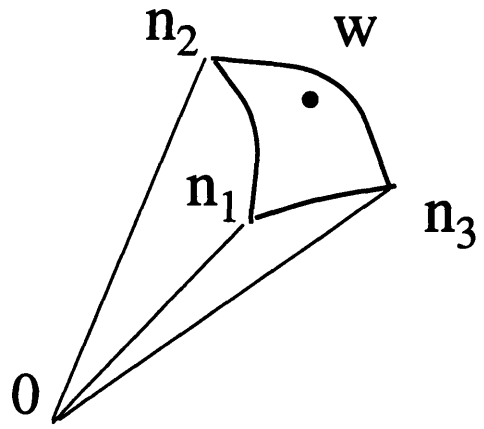

Figure 4-2: The topological charge by intersection counting

Based on this idea, the algorithm is as follows:

I Choose a random point $\vec{w}$ on a sphere. In the next steps we will find intersection numbers of triangles forming the lattice with the ray defined by $\vec{w}$. In principle

one could use any fixed point instead of a *random* one, in fact the result does not depend on the choice. However, there are certain advantages of this freedom which will be used in the next chapter.

II For every plaquette triangle Δ formed by 1-links of the lattice, find if the image of the triangle $\vec{e}(\Delta)$ on the sphere S^2 has the point $\vec{w}$ in the interior. If $\vec{w}$ is outside, assign $q_\Delta = 0$. Otherwise find if the map $\vec{e} : \mathcal{M} \to S^2$ preserves the orientation of Δ. Since $\mathcal{M}$ is orientable, orientation of all Δ is unambiguously defined by parallel transport. If $\vec{e}$ preserves the orientation, assign $q_\Delta = 1$, otherwise, set $q_\Delta = -1$.

III Sum q_Δ over the lattice. The result $Q = \sum q_\Delta$ is the topological charge of the configuration.

In the next chapter we shall see that a smart choice of w allows us to construct a Q-sensitive cluster algorithm.

Chapter 5

θ Vacua in the 2–d $O(3)$ σ–Model

As we have seen in the previous chapter, the topological charge is extremely difficult to put on the lattice.

A traditional warm-up exercise is studying the 2-dimensional $O(3)$ σ-model. This class of theories shows many important aspects of QCD and at the same time is much more tractable. Many exact results exist for the σ-model [49] which makes it even more attractive for lattice simulations, since results can be compared with exact solutions. Here we develop a method for working with the 2-d $O(3)$ σ-model with a θ-term included. We construct an update algorithm that allows dealing with arbitrary θ and apply it to show that the theory becomes massless at $\theta = \pi$.

In addition to being a prelude for the study of nonperturbative QCD, the 2-d $O(3)$ σ-model has yet another interesting application. In 1983 Haldane conjectured [50, 51] that integer and half-odd-integer 1-d antiferromagnetic quantum spin chains behave qualitatively differently; namely, while integer spin chains should have a mass-gap, half-odd-integer chains should be gap-less. This has been confirmed numerically for finite chains of spin 1 and 2 [52, 53] and analytically for half-odd-integer spins and for spin 1 [54]. The long-range physics of 1-d quantum spin chains is described by an effective 2-d classical $O(3)$-model. Haldane argued that the effective action for a chain of spins S contains a topological term $2\pi i S Q$. Here Q is the topological charge. Because of periodicity, all integer vales of S are mapped into $\theta = 0$ in field-theoretical language, while half-integer values of S correspond to $\theta = \pi$. The standard $O(3)$-

model has a mass-gap in agreement with Haldane's conjecture. On the other hand, the conjecture implies that the mass-gap of the $O(3)$-model disappears at $\theta = \pi$. A previous numerical study that was limited to $|\theta| < 0.8\ \pi$ found no phase transition in that region [55]. In fact, prior to the present work, Haldane's conjecture had not yet been verified in the context of the $O(3)$-model.

5.1 σ Model in 2 Dimensions

In this section, we review properties of the σ-model in two dimensions that will be used in the following discussion. More complete treatment of the subject can be found in an excellent review [49] and references therein.

The $O(N)$ σ-model in $(1+1)$ dimensions is a theory of N real fields $\sigma^a(x)$ $(a = 1, \ldots, N)$ which are defined on a unit sphere:

$$\sigma^a(x)\sigma^a(x) = 1. \tag{5.1}$$

σ^a transform as the vector representation of the group $O(N)$ (hence the name of the model.) The Lagrangian is usually written in the form

$$\mathcal{L} = \frac{N}{2f}\left(\partial_\mu \sigma^a(x)\right)\left(\partial^\mu \sigma^a(x)\right) \tag{5.2}$$

Although it looks like an uninteresting free field theory, solving constraint (5.1) with respect to one of the component adds a highly nontrivial interaction between the remaining components of σ^a. In addition, in the case $N = 3$ the classical field equations have solutions with finite energy—the instantons [56].

In the continuum, it is more convenient to incorporate constraint (5.1) into action (5.2) and rescale σ so that $\sigma^a\sigma^a = N/f$. The Euclidean action can be written then as

$$S[\sigma, \alpha] = \frac{1}{2}\int d^2x \left[\partial_\mu \sigma^a(x)\partial^\mu \sigma^a(x) - \frac{\alpha(x)}{\sqrt{N}}\left(\sigma^a(x)\sigma^a(x) - \frac{N}{f}\right)\right] \tag{5.3}$$

As can be easily shown, the theory undergoes a dimensional transmutation which gives a vacuum expectation value to the field $\alpha(x)$

$$\alpha(x) = \sqrt{N}m^2 + \alpha_{\rm qu}(x) \tag{5.4}$$

where

$$m^2 = M_0^2 \exp(-4\pi/f) \tag{5.5}$$

and M_0 is an ultraviolet cutoff momentum.

The topological charge in the $O(3)$ σ-model is defined as follows [56]:

$$Q = \frac{1}{8\pi} \int d^2x \epsilon^{abc} \epsilon_{\mu\nu} \sigma^a \partial^\mu \sigma^b \partial^\nu \sigma^c. \tag{5.6}$$

It is well known that the last expression is a second Pontryagin class of the map $\sigma : \mathcal{M} \to S^2$ from the coordinate space $\mathcal{M}$ to the unit sphere. The normalization factor is chosen such that $Q \in \mathbf{Z}$. The instanton realizes a local minimum of the action (5.2) in the $Q = 1$ sector. An important point is that the instanton action is strictly positive $S_{\text{inst}} > 0$.

Though instantons are absent in the σ-models for $N > 3$, there is a $SU(N)$ generalization of the sigma models, namely $CP(N-1)$ models which have instantons for all N[49].

5.2 The Lattice Version

For lattice simulation, it is more convenient to work on two-dimensional torus T^2 as the configuration space. As long as the lattice size L is large enough and the lattice spacing is small enough; $a \ll \xi \ll L$, there is not much difference between a sphere and a torus for our purposes. In particular, the instantons also exist on the torus.

We use a triangular discretization, so that every site has six neighbors and the lattice is covered with equilateral triangles.

The most symmetric fundamental domain on such a lattice is a honeycomb shown on figure 5-1; for technical reasons it is more convenient to work with an equivalent torus (see fig. 5-2.)

When $\xi \gg a$ there are a number of lattice actions with the same continuum limit $S_L \to 1/(\partial_\mu \sigma^a)$ for $a/\xi \to 0$ and a smooth enough configuration $(\sigma^a(x)\sigma^a(x+\hat{a}) \approx 1)$. Details of the topological charge definition and its relation to the cluster algorithm

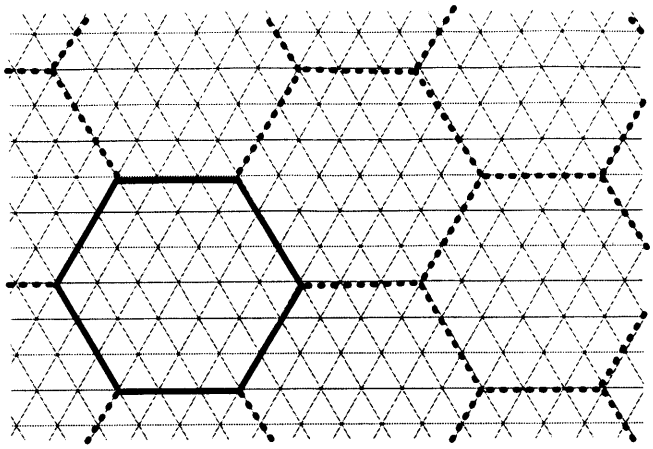

Figure 5-1: Honeycomb domains on the triangular lattice

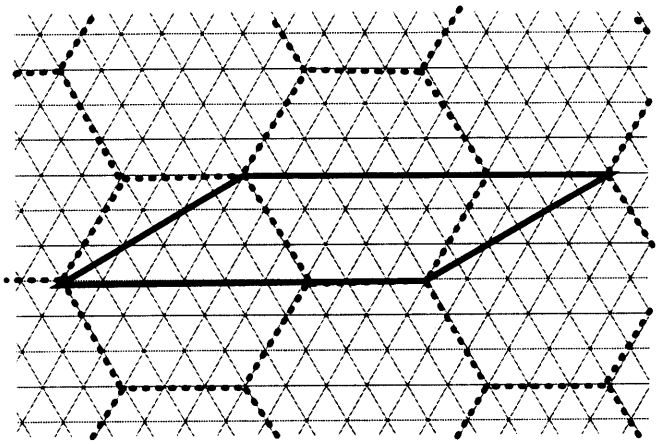

Figure 5-2: Parallelogram domains on the triangular lattice

prompt using the so-called truncated action

$$S[\vec{e}] = \sum_{\langle xy \rangle} s(\vec{e}_x, \vec{e}_y) \ , \quad s(\vec{e}_x, \vec{e}_y) = \begin{cases} \beta(1 - \vec{e}_x \cdot \vec{e}_y) & \text{if } \vec{e}_x \cdot \vec{e}_y > -1/2 \\ \infty & \text{otherwise} \end{cases} \tag{5.7}$$

Hereafter we denote the σ–field on the lattice by $\vec{e}$. This action differs from the standard one

$$S_{\mathrm{st}}[\vec{e}] = \sum_{\langle xy \rangle} \beta_{\mathrm{st}}(1 - \vec{e}_x \cdot \vec{e}_y) \tag{5.8}$$

only by effectively prohibiting very fast change of the orientation of $\vec{e}$. This cut-off is relevant only on the ultraviolet scale where other lattice effects are significant anyway. It is worth noticing that (5.7) belongs to the same universality class as the untruncated action (5.8), so the continuous observables are not affected by limiting the angle between neighbors to $]-2\pi/3, 2\pi/3[$.

Of course, one can not hope to obtain the same results for both (5.7) and (5.8) unless $\beta_{\rm st}$ is converted to β: $\langle \mathcal{O}(\beta) \rangle = \langle \mathcal{O}(\beta_{\rm st}(\beta, \mathcal{O})) \rangle_{\rm st}$ (strictly speaking, this is only an asymptotic expression when $\beta \to \infty$.) But, since both actions are in the same universality class, $\beta_{\rm st}(\beta, \mathcal{O})$ does not really depend on $\mathcal{O}$ and it suffices to identify only one observable, say ξ, in both lattice realizations to map $\beta_{\rm st} \to \beta$.

5.3 Cluster Update Algorithm

On lattices large compared to the correlation length ($L \gg \xi$) special care should be taken to generate statistically independent configurations because of critical slowing down [57, 58, 59, 60]. In addition, measuring topological quantities requires an extremely large sampling to see the signal.

Fortunately, there exist cluster algorithms [61, 62, 63] which deal with the critical slowing down by updating collective modes on the lattice instead of individual spins. In addition, the cluster construction allows to sum over large number of configuration analytically by building improved estimators for the observables.

In this section we describe a multi-cluster version of Wolff's cluster algorithm and corresponding improved estimators for the action and magnetic susceptibility. An improved estimator for the topological charge will be constructed in the next section.

For the $O(3)$ σ-model the Wolff algorithm works as follows.

- First, one chooses an arbitrary point $\vec{n}$ on S^3. Then a $\mathbf{Z}_2$ is embedded into the theory in the following way. Every spin $\vec{e}$ can be flipped with respect to $\vec{n}$:

$$\vec{e}' = \vec{e} - 2\vec{n}(\vec{n} \cdot \vec{e}). \tag{5.9}$$

  Since $|\vec{n}| = 1$, $\vec{e}'$ also belongs to S^3. The flip operation (5.9) is a reflection on a sphere with respect to the equator defined by the north pole $\vec{n}$.

  Moreover, $\vec{e}'' = \vec{e}$ and if $\vec{a}'$ and $\vec{b}'$ are flips of $\vec{a}$ and $\vec{b}$ respectively, then $s(\vec{a}, \vec{b}) = s(\vec{a}', \vec{b}')$ and $s(\vec{a}, \vec{b}') = s(\vec{a}', \vec{b})$. These properties of the $\mathbf{Z}_2$ embedding are important for the proof of detailed balance.

- Second, every 1-link on the lattice is made either "active" or "passive": if flipping one of the spins on the link decreases the action, the link is passive, otherwise the link is active with probability $p = 1 - \exp(s(\vec{x},\vec{y}) - s(\vec{x},\vec{y}'))$. Alternatively, one can think about making every link active with probability

$$p = 1 - \min(1, \exp(s(\vec{x},\vec{y}) - s(\vec{x},\vec{y}'))). \tag{5.10}$$

- Next, clusters of spins are formed by defining all sites on the lattice connected by paths of active links to belong to the same cluster. As a result, each spin belongs to one and only one cluster.

- The final step is the update. In the multi-cluster version we are using, every cluster is flipped with 50% probability.

On average, $1/2$ of all the spins will be updated on a single iteration. For the next generation, another point is picked up for $\vec{n}$.

Because the critical behavior of the percolation of the bonds updates spins on all length scales, this algorithm eliminates critical slowing down.

5.4 Improved Estimators

Finding an efficient update strategy is only one part of problem. A second issue is to improve the statistics of the measured observables. Here again, the cluster algorithm is of great help. For certain quantities it is possible to construct so-called improved estimators which take into account the cluster structure built during the update step. Here we consider two quantities, namely the energy $\langle E \rangle$ and magnetic susceptibility χ_m.

5.4.1 Action

For the average energy, $\langle E \rangle = \frac{-\partial Z}{Z \partial \beta}$, it is easy to work out an improved estimator. One can write

$$\langle E \rangle = \langle \sum_{x \sim y} (1 - (\vec{x} \cdot \vec{y})) \rangle + \langle \sum_{x \not\sim y} (1 - (\vec{x} \cdot \vec{y}) \rangle, \tag{5.11}$$

where $x \sim y$ iff x and y belong to the same cluster. The second sum could be written as $\langle \sum_{x \not\sim y} (\vec{x} + \vec{x}' \cdot \vec{y}) \rangle / 2$, which in turn is $2/3 \langle \sum_{x \not\sim y} (\vec{x} \cdot \vec{y}) \rangle$. Hence, the estimator (5.11) is nothing but

$$\langle E \rangle = \langle \sum_{x \sim y} (1 - (\vec{x} \cdot \vec{y})) \rangle. \tag{5.12}$$

5.4.2 Magnetic Susceptibility

The magnetic susceptibility $\chi_m = \frac{\partial^2 Z}{Z \partial H^2}$ also has an improved estimator. By the same argument as for $\langle E \rangle$ one finds

$$\chi_m = \langle \sum_i \left(\sum_{x \in C_i} \vec{x} \right)^2 \rangle. \tag{5.13}$$

In both cases $\langle E \rangle$ and χ_m we explicitly summed over flips of all clusters. If there are N_c clusters on average per configuration, then this trick improves statistics by a factor of 2^{N_c}. In our simulation, there is about a hundred clusters per configuration, so the number of configurations required to achieve the same accuracy with the standard estimators would be a factor of 10^{30} larger.

5.5 θ-term in the Action

The theory becomes more complicated when a θ-term is added into the partition function. Now all observables can in principle depend on the vacuum angle θ. In what follows we will concern ourselves mostly with magnetic susceptibility $\chi_m(\theta)$ and so-called topological susceptibility. The latter is defined as

$$\begin{aligned} \chi_t(\theta) &= \frac{\langle (Q - \bar{Q})^2 \rangle}{V} \\ &= -\frac{1}{V} \left(\frac{\partial^2 Z}{Z \partial \theta^2} - \left(\frac{\partial Z}{Z \partial \theta} \right)^2 \right). \end{aligned} \tag{5.14}$$

5.6 Clusters and Topological Charge

In this section we illustrate how clusters and the triangle counting definition of the topological charge of section 4.3.2 can be used together.

Let us consider what happens to the topological charge when clusters are flipped. As a first step in building an improved topological estimator, notice that there are three types of triangles:

A All three vertices belong to the same cluster.

B Two of the vertices belong to one cluster, and the third—to another.

C All vertices are in different clusters.

Second, suppose we define a charge Q_i for every cluster, so that $Q = \sum Q_i$, with the following properties: *i)* The Q_i do not depend on orientation of other clusters: after flipping cluster i one obtains the topological charge Q': $Q - Q' = Q_i - Q_i'$. *ii)* In addition, flipping the cluster reverses its charge, $Q_i' = -Q_i$, then $Q_i = (Q - Q')/2$ takes half-integer values.

Now let us try to construct Q_i which would satisfy these properties. If the direction of $\vec{w}$ is orthogonal to $\vec{n}$ (in this case $\vec{w}$ is the *west pole* first discovered by Winnie-the-Pooh [64] in a slightly different context), then the flip reverses q_Δ on the A-triangles: $(-1, 0, +1) \to (+1, 0, -1)$.

For B-triangles there are two clusters and four possible orientations.

I	J	q_Δ
$\uparrow$	$\uparrow$	$q^{(\uparrow\uparrow)}$
$\uparrow$	$\downarrow$	$q^{(\uparrow\downarrow)}$
$\downarrow$	$\uparrow$	$q^{(\downarrow\uparrow)}$
$\downarrow$	$\downarrow$	$q^{(\downarrow\downarrow)}$

It seems impossible to combine two cluster charges Q_i and Q_j to reproduce four values of $q^{()}$. However, the choice of the west pole for $\vec{w}$ guarantees that $q^{(\downarrow\downarrow)} = -q^{(\uparrow\uparrow)}$ and $q^{(\uparrow\downarrow)} = -q^{(\downarrow\uparrow)}$. Hence, B-triangles contribute $(q^{(\uparrow\uparrow)} + q^{(\uparrow\downarrow)})/2$ to Q_i and $(q^{(\uparrow\uparrow)} + q^{(\downarrow\uparrow)})/2$ to Q_j.

The situation with C-triangles is more complicated. There are now three clusters I, J and K and four orientations after taking into account the simultaneous flip of all clusters. As a result, we have a system of four equations for three variables

$$\begin{aligned} Q_i + Q_j + Q_k &= q^{(\uparrow\uparrow\uparrow)} \\ Q_i + Q_j - Q_k &= q^{(\uparrow\uparrow\downarrow)} \\ Q_i - Q_j + Q_k &= q^{(\uparrow\downarrow\uparrow)} \\ Q_i - Q_j - Q_k &= q^{(\uparrow\downarrow\downarrow)} \end{aligned}$$

which are in general inconsistent. However, if action (5.7) is used, then C-triangles could exist only if $q^{(\uparrow\uparrow\uparrow)} + q^{(\uparrow\uparrow\downarrow)} - q^{(\uparrow\downarrow\uparrow)} - q^{(\uparrow\downarrow\downarrow)} = 0$, which is exactly the consistency condition for the above system[1]. Moreover, Q_i gets $(q^{(\uparrow\uparrow\uparrow)} + q^{(\uparrow\downarrow\downarrow)})/2$ from this triangle and other charges likewise get half-integer contributions.

The final expression for Q_i is

$$Q_i = \sum_{\Delta \in A(C_i)} q_\Delta + \frac{1}{2} \sum_{\Delta \in B(C_i)} \left(q_\Delta^{(\uparrow\uparrow)} + q_\Delta^{(\uparrow\uparrow)} \right) + \frac{1}{2} \sum_{\Delta \in C(C_i)} \left(q_\Delta^{(\uparrow\uparrow\uparrow)} + q_\Delta^{(\uparrow\downarrow\downarrow)} \right). \qquad (5.15)$$

This definition has the following properties that will be used in the next section to construct an improved estimator:

- All Q_i are half-integers and the above algorithm gives *exact* result for them without any rounding.
- Flipping one cluster charges the sign of its charge and does not affect charges of other clusters.
- The total topological charge is a sum of cluster charges: $Q = \sum Q_i$.

5.7 Improving Topological Estimators

As was mentioned before, a configuration with N clusters represents an ensemble of 2^N field configurations which differ only by the cluster orientation. In our simulation,

[1] It is trivial but long to check that all triangles which can violate the consistency condition have at least one side longer than $2\pi/3$. The cluster algorithm will put an active line on such a side and the triangle in question will be either class A or B.

typically there are $\sim$ 100 clusters per configuration. The idea behind all improved estimators is to sum analytically over all cluster orientations, getting by a factor of 2^N better statistics.

Quantities we are interested in have the form:

$$\langle \mathcal{O} \rangle_\theta = \frac{1}{Z(\theta)} \int \mathcal{D}\vec{e}\ \mathcal{O}[\vec{e}]\ \exp\{-S[\vec{e}] + i\theta Q[\vec{e}]\}\,.$$

This can be rewritten as

$$\begin{aligned} \langle \mathcal{O} \rangle_\theta &= \frac{1}{Z(\theta)} \sum_q e^{i\theta q} \int \mathcal{D}\vec{e}\ \delta(Q[\vec{e}] - q)\ \mathcal{O}[\vec{e}]\ \exp\{-S[\vec{e}]\} \\ &= \frac{Z}{Z(\theta)} \sum_q p(q) e^{i\theta q} \mathcal{O}_q, \end{aligned}$$

where $\mathcal{O}_q = \int \mathcal{D}\vec{e}\ \mathcal{O}[\vec{e}]\ \delta(Q[\vec{e}] - q)\ \exp(-S[\vec{e}])/Z(q)$ is the expectation value of $\mathcal{O}$ in a given topological sector, $Z(q) = \int \mathcal{D}\vec{e}\ \delta(Q[\vec{e}] - q) \exp(-S[\vec{e}])$ is the partition function for that sector, and $p(q) = Z(q)/Z$ is the weight of the topological sector with charge q.

Hence, measuring of $\mathcal{O}$ for arbitrary θ consists of finding $\mathcal{O}_q$ and $p(q)$. It is easy to construct $\mathcal{O}_q$ along the lines of section 5.4, e.g., one can glue all charged clusters together for the calculating the improved estimators. In our simulations, usually about a half of clusters have charge zero, so the removal of charged clusters from the improved estimators still results in a factor of 2^{50} of statistical improvement.

5.7.1 Measuring $p(q)$

To calculate $p(q)$ the following trick can be used. Consider a system of clusters in a given configuration. It represents an ensemble of configurations contributing to the partition function with the same weight. However, these configurations have different topological charges. Since there is a considerable number of charged clusters, we can speak about a *distribution* of the topological charge in the ensemble. The overall distribution $p(q)$ is an average of these configuration distributions. The problem now is to build the sample distribution from the cluster set.

The following pseudo-code does the job:

```
construct the clusters;
put p_old[0] = 1, p_old[i] = 0 otherwise;
for each cluster i do {
  let charge be the charge of the cluster i;
  if ( charge == 0 )
     continue;
  put p_new[i] = 0 for all i;
  for all i sequentially do {
     p_new[i+charge] += p_old[i] / 2;
     p_new[i-charge] += p_old[i] / 2;
  }
}
return p_new;
```

This algorithm takes $O(N^2)$ steps to sum over 2^N orientations. It has only rounding errors and can be made exact. A typical distribution of $p(q)$ is shown on fig. 5-3.

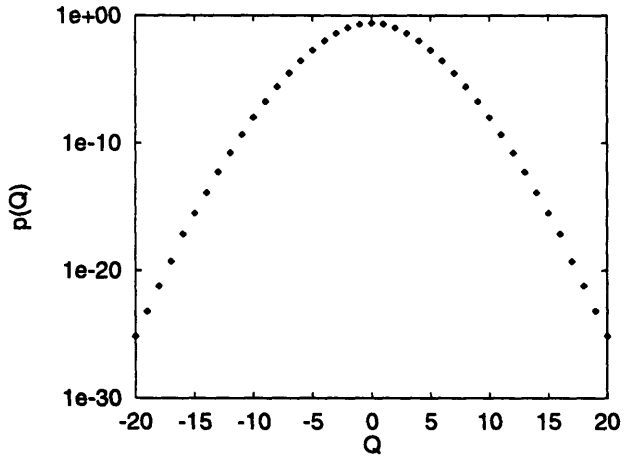

Figure 5-3: The topological charge distribution $p(Q)$ on the 36×12 lattice

5.7.2 Reweighting Technique

The distribution $p(Q)$ varies over many orders of magnitude. In particular, as one can see from fig. 5-3, large charges are suppressed exponentially. Still, at θ close to π the

contributions from the large charges are important and need to be determined with high accuracy. In this respect, a reweighting technique [65] using a trial distribution $p_t(Q)$ has been useful. The trial distribution should be as close as possible to the real distribution $p(Q)$. In general this is difficult to achieve because one has not yet determined $p(Q)$. In our case, the improved estimator described above provides a good trial distribution $p_t(Q)$. Then one works with an effective action

$$S_{\text{eff}}[\vec{e}] = S[\vec{e}] + \ln p_t(Q[\vec{e}]) \tag{5.16}$$

and determines $p(Q) = p_t(Q) \int \mathcal{D}\vec{e}\, \delta_{Q,Q[\vec{e}]}\ \exp(-S_{\text{eff}}[\vec{e}])$.

The combination of using the cluster algorithm and the reweighting technique provides enough statistics for measuring both χ_m and χ_t.

5.8 Spin Chains

The cluster formulation of the $O(3)$-model provides more than just an efficient numerical algorithm. The clusters can be interpreted as physical objects that play a crucial role in the dynamics. We identify the clusters of charge $\pm 1/2$ with merons. This definition of merons is consistent with the semiclassical picture. In fact, an instanton is built out of two Wolff clusters, both with topological charge $1/2$. At $\theta = 0$ the clusters are independent and hence the merons form an ideal gas. The resulting disorder is responsible for the mass-gap of the theory. At $\theta = \pi$ the Boltzmann factor $\exp(i\theta Q)$ is 1 for even Q and -1 for odd Q. When a configuration contains a half-odd-integer charged cluster, flipping it changes Q by an odd integer. The resulting configurations then have opposite Boltzmann weights and their contributions cancel out in the partition function. Therefore, at $\theta = \pi$ only those configurations for which all clusters carry integer charges contribute to $Z(\theta)$. This means that the merons are now bound in pairs and can no longer disorder the system. Hence it is natural to expect that the mass-gap vanishes at $\theta = \pi$. This agrees with Haldane's conjecture and it provides an exact formulation of Affleck's dynamical picture.

5.9 Numerical Results

We have performed numerical simulations with the meron-cluster algorithm at $g = \infty$ for volumes $V = 3L \times L = 18 \times 6,\ 24 \times 8,\ 30 \times 10$ and 36×12. For each lattice we have performed 10^7 sweeps for measurements. The correlation length at $\theta = 0$ is about 2.8 lattice spacings. Fig.5-3 shows the topological charge distribution $p(Q)$ on the 36×12 lattice. Due to our reweighting technique we can easily generate a distribution that covers 25 orders of magnitude.

Up to logarithmic corrections discussed below, we find $\chi_t(\pi) \propto L$ and $\chi_m(\pi) \propto L$ which indicates a *second* order phase transition. For a first order transition the susceptibilities would grow proportional to L^2. In fact, at very strong coupling one expects a first order phase transition [66, 67]. Due to the constraint in our action of eq. (5.7) the strongest bare coupling we can consider ($\beta = 0$) turns out to be too weak to reach this regime. We have not run our algorithm at $\beta < 0$ although this may be feasible. It would be interesting if one could then reach the first order domain. Affleck et al. have conjectured that the critical theory at the second order phase transition is up to additional marginal operators a conformal field theory in the universality class of the $k = 1$ Wess-Zumino-Novikov-Witten model [68, 67]. Close to the phase transition, the mass-gap of the infinite volume theory should behave as

$$m(\theta) = |\theta - \pi|^{2/3} |\ln(|\theta - \pi|)|^{-1/2}. \tag{5.17}$$

We consider Fisher's finite size scaling variable $z = m(\theta)L$ which is a renormalization group invariant measure of the physical volume. Using the critical exponents of the Wess-Zumino-Novikov-Witten model with logarithmic corrections to scaling due to the additional marginal operators [68], one expects

$$\chi_t(\theta, L) = L\ (\ln L)^{-1/2}\ g_t(z), \quad \chi_m(\theta, L) = L\ (\ln L)^{1/2}\ g_m(z), \tag{5.18}$$

close to the phase transition. Here $g_t(z)$ and $g_m(z)$ are *universal* functions. In figs. 5-4 and 5-5, $\chi_t(\theta, L)L^{-1}(\ln L)^{1/2}$ and $\chi_m(\theta, L)L^{-1}(\ln L)^{-1/2}$ are shown as functions of z. In both cases the data seem to be close to a universal curve. The logarithmic corrections to scaling are important for this. Note that no fitting or free parameters

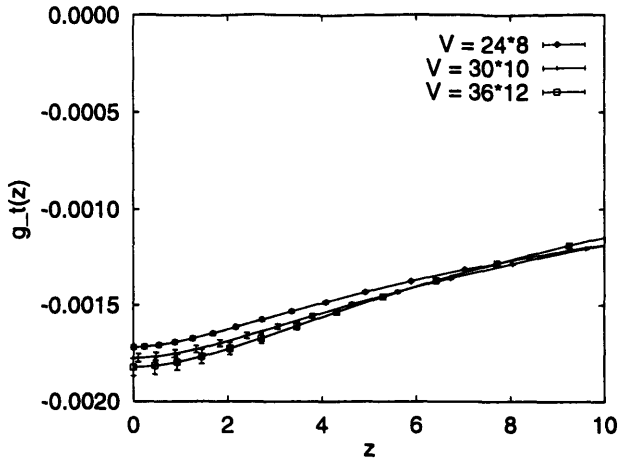

Figure 5-4: Data for the universal function $g_t(z)$

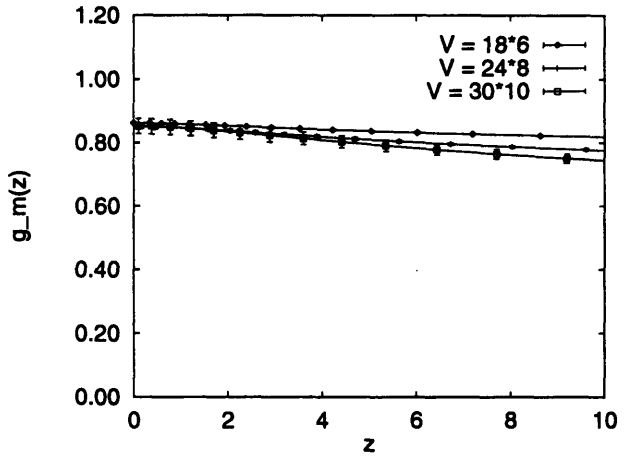

Figure 5-5: Data for the universal function $g_m(z)$

are involved. This confirms the described scenario. Cluster diagnostics on the 36×12 lattice at $g = \infty$ show that most clusters are neutral, 4 percent have charge $\pm 1/2$, 1 in 10^6 have charge ± 1 and very few have larger charges. The average sizes of the 0, $\pm 1/2$ and ± 1 charged clusters are 1.5, 10 and 32 lattice sites, respectively.

5.10 Conclusion

In conclusion, we have identified a second order phase transition at $\theta = \pi$ with the critical exponents of the $k = 1$ Wess-Zumino-Novikov-Witten model. This confirms Haldane's conjecture in the framework of the $O(3)$-model. Identifying Wolff clusters

of charge $\pm 1/2$ with merons, we have provided a precise formulation of the meron picture. We have confirmed that the merons form integer charged pairs at the phase transition. The $O(3)$-model is equivalent to the $CP(1)$-model. All $CP(N)$-models have instanton solutions and hence a θ-vacuum. In the large N limit, the $CP(N)$-model has a first order phase transition at $\theta = \pi$ [66] where CP is spontaneously broken. One expects that this persists down to $N = 2$ [66, 67, 69] although a numerical study in the $CP(3)$-model (motivated by the strong CP-problem) seems to contradict this [70]. Unfortunately, cluster algorithms do not work well for $CP(N)$-models with $N \geq 2$ [71].

The θ-vacua in gauge theories cannot yet be simulated reliably because there is no known cluster or comparably efficient algorithm.

Appendix A

Portable Random Number Generator

In a broad class of applications, different types of random data are required. Quite often the data in question live not in a simple domain like an interval or a square, but are homogeneously distributed over some complex manifold. An important example is a random point on three dimensional sphere S_3, which is widely used in the lattice QCD code.

The homogeneous distribution in this case is defined by the probability distribution $\rho(\phi,\theta) = 1$, where (ϕ,θ) are 'natural' coordinates on S_3. The measure $\mathrm{d}\Omega = \frac{1}{4\pi}\mathrm{d}\phi\sin\theta\mathrm{d}\theta$ imposes severe problems on using general methods for continuous distributions [72] because the map $(\phi,\theta) = [0,2\pi)\times[0,\pi) \to S_3$ is singular at $\theta = 0,\pi$. Moreover, any map $D_2 \to S_3$ has at least one singular point.

To circumvent this problem, the following approach is commonly used in the sequential code. First, map the manifold $\mathcal{M}$ in question into an $N+1$–dimensional cube $C_{N+1} = [-1,1]^{N+1}$ with Euclidean metric, where $\dim\mathcal{M} = N$ in such a way that there is a region $D \in C_{N+1}$, $D \cong (0,1]\times\mathcal{M}$ and the metric induced on $\mathcal{M}$ reproduces the required measure. Now, apply the following algorithm:

A1. Repeat generating points p_i in C_{N+1} until $p_k \in D$.

A2. Project p_k to $\mathcal{M}$.

We assume that there is a simple way to generate a random point in C_{N+1}.

There is a tremendous difficulty in making an efficient port of this algorithm to SIMD environment. The problem is that step A1 in the data–parallel edition reads

A1'. Repeat generating points p_i in C_{N+1} until $p_k \in D$ on all positions of the data set.

In this work an alternative approach is investigated. The idea is based on the observation that a random number calculated on the position k is perfectly suitable for any other position j. So, instead of waiting for every position to produce a random point, we employ the following algorithm:

B1. *Initialization.* Set $k \leftarrow 0$. Result will be stored in r.

B2. Generate a point in C_{N+1} on every position.

B3. Mark the positions containing points inside D. Let them be positions $j_1, j_2, ..., j_m$.

B4. Copy $p_{j_1}, p_{j_2}, \ldots$ to $r_k, r_{k+1}, r_{k+2}, \ldots$. (Stop copying when the last position is filled.)

B5. Set $k \leftarrow k + m$.

B6. Repeat steps B2–B5 while $k <$ total number of positions.

We assume that a generic scan operation is available for step **B4**. An implementation of the algorithm B on C* 7.1 [73] uses the `scan()` procedure which is quite time consuming on CM–5 when data leave the physical processing node.

The table below shows that even on the CM-5 with it's expensive communication [74] there is considerable performance improvement.

As the scan operation is quite general in the SIMD world, algorithm **B** does not use any CM–5 specific features and may be implemented easily on other platforms.

# of positions	Alg. A time	Alg. B time
2	55.34	27.07
8	94.35	29.96
32	135.74	30.42
128	172.53	30.34
512	211.03	27.96
2048	482.79	48.01
4096	1044.24	90.67
8192	2174.27	170.11

Table A.1: Time is in milliseconds for S_3 on 32 node VU CM-5 in dedicated mode

Appendix B

Jackknife

In this appendix we outline the jackknife method used to estimate errors of effective masses and three-point functions in chapters 2 and 3 respectively. A more complete treatment of the jackknife and bootstrap methods can be found in refs. [75, 76]. We start by reviewing the basic idea behind the jackknife method. After that an outline of its application to the effective mass and three-point function calculations is given.

Consider a statistical ensemble $\{x_i, i = 1 \ldots N\}$ with the average $\langle x \rangle = \sum_i x_i/N$ and the dispersion $D(x) = \frac{\sum_i (x_i - \langle x \rangle)^2}{N-1}$. The standard way to estimate a statistics $f(x)$ is the following.

$$\langle f \rangle = \frac{\sum_i f(x_i)}{N} \tag{B.1}$$

$$\sigma(f) = \sqrt{\frac{\sum_i (f(x_i) - \langle f \rangle)^2}{N(N-1)}}. \tag{B.2}$$

This simple approach, however, does not work well in many situations. E.g., if $D(x)$ and $|f''(x)|$ are large enough, the error of the $\langle f \rangle$ is overestimated by (B.2). One can redeem the situation somewhat by the following trick.

Consider a derivative ensemble $\{x^{(i)}, i = 1 \ldots N\}$, where

$$x^{(i)} = \frac{N\langle x \rangle - x_i}{N-1}, \tag{B.3}$$

and use the following formulae for statistics $f(x)$:

$$\langle f \rangle_J = \frac{\sum_i f(x^{(i)})}{N} \tag{B.4}$$

$$\sigma_J(f) = \sqrt{\frac{(N-1)\sum_i (f(x^{(i)}) - \langle f\rangle_J)^2}{N}}. \tag{B.5}$$

One can easily convince oneself that $\langle x\rangle = \langle x\rangle_J$ and $\sigma(x) = \sigma_J(x)$. Moreover, $\langle f\rangle_J$ is an unbiased estimator for $\bar{f}$ if all x's are independent random variables, $\sigma(x)$ is finite and $f(x)$ and $p(x)$ satisfy some mild conditions.

Notice, that (B.3) is nothing but an average of all xs except for x_i and (B.5) differs from the standard definition by a scaling factor chosen to reproduce expected results in case $f(x) = x$.

An interested reader will find a more complete treatment of the procedure in the literature cited above.

To apply this scheme to the effective mass calculation, we use a set of propagators $\{D_i(t), i = 1 \ldots N\}$ each calculated at a given gauge field configuration U_i. The derived ensemble is

$$D^{(i)}(t) = \frac{\sum_{k=1}^{N} D_k(t) - D_i(t)}{N-1}$$

and the statistics

$$m_{\text{eff}}(D,t) = \log \frac{D(t)}{D(t+1)}.$$

After that formula (B.5) is applied to get an estimator for $\sigma(m)$. However, further analysis shows that $\sigma(m)$ is overestimated by this method. One possible explanation is that first, there is a strong correlation between $D_i(t)$ and $D_i(t+1)$ and, second, the normal distribution is not a very good model statistics for $D_i(t)$.

In case of three-point functions, we seek to estimate an average and the standard deviation of $O(t_1,t_2,t_3)/D(t_1,t_2)$. On the lattice each gauge configuration U_i yields $O_i(t_1,t_2,t_3)$ and $D_i(t_1,t_2)$. Again, a derivative ensemble $\{O^{(i)}, D^{(i)}\}$ is formed as

$$\begin{aligned} D^{(i)} &= \frac{\sum_{k=1}^{N} D_k - D_i}{N-1} \\ O^{(i)} &= \frac{\sum_{k=1}^{N} O_k - D_i}{N-1}, \end{aligned}$$

and the ratio and its variance are estimated using (B.4) and (B.5). Probably for the same reasons as for m_{eff}, the resulting σ seems to be too large.

Appendix C

Search for Plateaux

If statistics itself is more an art than a science, then error analysis is black magic with all its smoke and mirrors. We certainly do not pretend to have a bulletproof procedure for searching the plateaux in the effective masses or three-point functions. While it is sometimes possible to find methodical flows in some approaches, it is much more difficult if indeed possible at all to suggest a "correct" method. Hence, the following should be considered but a recipe and taken with a grain of salt.

In general the problem can be stated as follows. *Given experimental results for* $x(t) \pm \delta(t), t = 1 \ldots N$ *and assuming that for* t *large enough* $x(t)$ *reaches a plateau* $x(t) \to x_0$, $t \gg 1$, *derive an estimate for* x_0 *and find its error* $\sigma(x_0)$ *under the condition that random noise* $\delta(t)$ *increases at large* t.

We illustrate the procedure used with an example. Table C.1 shows a set of three-point function "measurements" cooked-up to illustrate how the technique works. As can be seen in fig. C-1, there is a source at $t = 8$, and a sink at $t = 24$. The data set shows noticeable contributions of excited states at small separations and an increase in noise at large t.

We begin with defining a time window $w(S, M)$ starting at $t = S$ and including M consecutive points. Using these M points we calculate an estimator for χ^2 statistics:

$$\frac{\chi^2(S,M)}{M} = \sum_{i=S}^{S+M-1} \frac{(x_i - \bar{x})^2}{M\sigma_i},$$

t	x	t	x	t	x	t	x
0	0.03(5)	8	7.23(18)	16	5.18(17)	24	0.41(54)
1	0.05(1)	9	6.34(23)	17	4.99(11)	25	0.01(21)
2	0.09(12)	10	5.73(7)	18	5.14(9)	26	0.19(6)
3	0.00(3)	11	5.45(10)	19	4.97(14)	27	0.14(19)
4	0.01(4)	12	5.21(7)	20	5.02(19)	28	0.07(20)
5	0.03(6)	13	5.21(8)	21	5.09(30)	29	0.04(9)
6	0.01(2)	14	5.11(6)	22	4.90(25)	30	0.02(7)
7	0.04(5)	15	5.16(14)	23	4.53(15)		

Table C.1: Result of a gedanken experiment

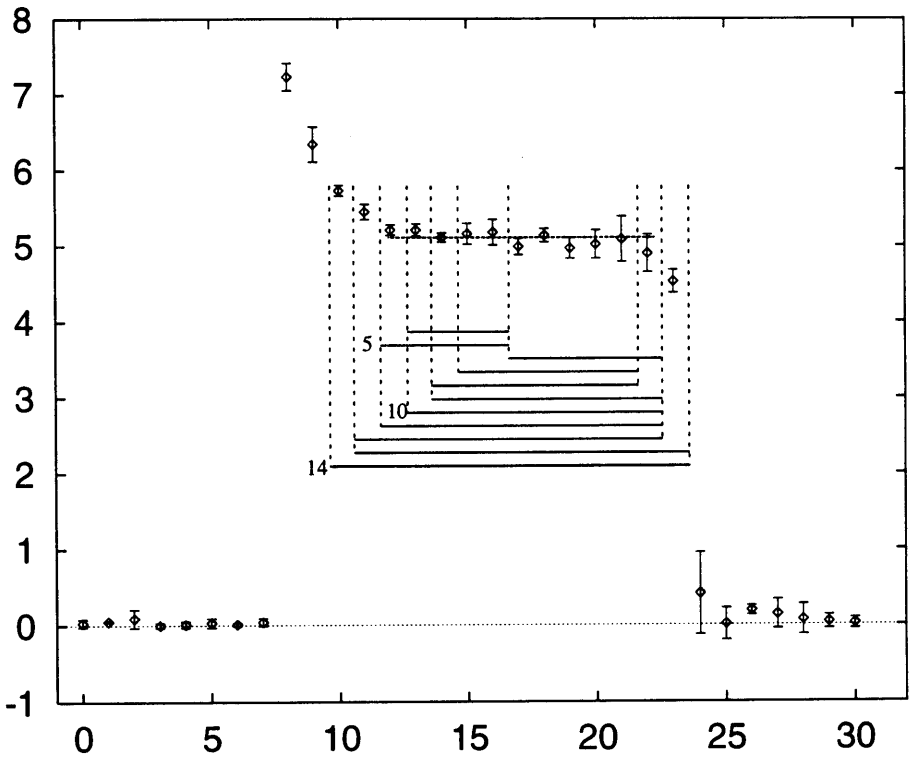

Figure C-1: An example of plateau search results

where

$$w = \sum_{i=S}^{S+M-1} 1/\sigma_i^2, \quad \bar{x} = \frac{1}{wM} \sum_{i=S}^{S+M-1} \frac{x_i}{\sigma_i^2}.$$

After that we search for the global minimum $\min_S \chi^2(S,M)/M$. As an example, fig. C-2 shows $\chi^2(S,6)/6$. One sees that in the region of interest, the minimum is

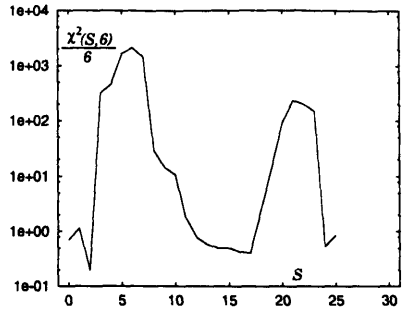

Figure C-2: $\chi^2(S,M)/M$ for $M = 6$

reached at $S = 17$. The windows corresponding to $\min \chi^2(S,M)/M$ are plotted as solid horizontal lines on fig. C-1. Table C.2 shows the minimal values of $\chi^2(S,M)/M$ for $M = 4 \ldots 14$. The following observations can be made:

S	M	$\min \chi^2(S,M)/M$	S	M	$\min \chi^2(S,M)/M$
4	13	0.35	10	13	0.57
5	12	0.39	11	12	0.67
6	17	0.41	12	11	1.45
7	15	0.42	13	11	2.74
8	14	0.38	14	10	7.26
9	14	0.40			

Table C.2: $\min \chi^2(S,M)/M$ and corresponding S for various M

1. If the time window is too small, there is no single prominent minimum of chi^2. As the window size grows, a single deep minimum is formed while other minima either disappear or are pushed up. In addition these unstable minima tend to change their position rapidly as the size of the window changes. Windows of sizes 4 and 5 in the example show this behavior.

2. Once a stable minimum is formed, changing the window size produces no significant changes in the window position ($M >= 6$ in our case.)

3. Plotted as a function of M, $\min_S \chi^2(S,M)/M$ reaches a plateau in the stable window régime with a sudden jump up when the window becomes too large (see fig. C-3.

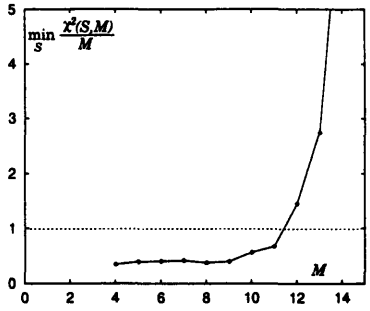

Figure C-3: $\min_S \chi^2(S,M)/M$ vs. M

Finally, we choose S and M corresponding to the end stable χ^2 region. In our case $S = 12$, $M = 11$. The corresponding average and its standard deviation is plotted on fig. C-1 as a dotted horizontal bar. We see that, as expected, approximately 1/3 of data points deviate from the average by more than one σ. The final result of the analysis is

$$\bar{x} = 5.13 \qquad \sigma(x) = 0.01.$$

The fine print: The expected value of χ^2/M is close to 1. If the χ^2-plateau is reached at a value significantly different from 1, it could indicate that σ_i have not been correctly estimated. As a quick remedy, one can rescale the error to bring χ^2/M to 1, however, an extreme cation is required here, because more often than not too large or too small χ^2 is a result of a systematic bias or a simple error in the analysis.

Bibliography

[1] M. Aguilar-Benitez, C. Amsler, B. Armstrong, R. M. Barnett, P. R. Burchat, C. D. Carone, C. Caso, G. Conforto, R. L. Crawford, S. Eidelman, J. L. Feng, P. S. Gee, C. Grab, D. E. Groom, A. Gurtu, K. Hagiwara, K. G. Hayes, J. J. Hernandez, K. Hikasa, K. Honscheid, F. James, M. Mangano, A. Manohar, K. Monig, L. Montanet, H. Murayama, K. Nakamura, K. A. Olive, M. Roos, R. H. Schindler, R. E. Shrock, M. Tanabashi, T. G. Trippe, G. S. Wagman, C. G. Wohl, and R. L. Workman. Review of particle physics. *Phys. Rev.*, D54:1, 1996.

[2] K. Wilson. Confinement of quarks. *Phys. Rev.*, D10:2445, 1974.

[3] S. Kobayashi and K. Nomizu. *Foundations of differential geometry.* Interscience Publishers, New York, 1969.

[4] Tamiaki Yoneya. Monopole condensation and quark confinement in a weak coupling $SU(N)$ lattice gauge model. *Nuclear Physics*, B153:431, 1979.

[5] G. Mack and V. B. Petkova. Sufficient condition for confinement of static quarks by a vortex condensation mechanism. *Ann. of Phys.*, 125:117, 1980.

[6] Yu. M. Makeenko and S. B. Khokhlachev. Symmetry of gauge group center and the problem of quark confinement in quantum chromodynamics. *Sov. Phys. JETP*, 51:448, 1981.

[7] Yu. M. Makeenko and M. I. Polikarpov. Phase diagram of mixed lattice gauge theory from viewpoint of large N. *Nuclear Physics*, B205:386, 1982.

[8] K. Symanzik. Continuum limit and improved action in lattice theories. 1. Principles and ϕ^4 theory. *Nuclear Physics*, B226, 1983.

[9] K. Symanzik. Continuum limit and improved action in lattice theories. 2. $O(N)$ nonlinear σ model in perturbation theory. *Nuclear Physics*, B226, 1983.

[10] G. P. Lepage. QCD for coarse lattices. *Nucl. Phys. B (Proc. Suppl.) 47*, page 3, 1996.

[11] T. DeGrand, A. Hasenfratz, P. Hasenfratz, and F. Niedermayer. Fixed point actions for $SU(3)$ gauge theory. *Physics Letters*, B365:233, 1996.

[12] W. Bietenholz, R. Brower, S. Chandrasekharan, and U.J. Wiese. Progress on perfect lattice actions for QCD. In *Lattice 96: 14th International Symposium on Lattice Field Theory*, St. Louis, MO, 1996.

[13] H. B. Nielsen and M. Ninomiya. Absence of neutrinos on a lattice. 1. Proof by homotopy theory. *Nuclear Physics*, B185:541, 1981.

[14] H. B. Nielsen and M. Ninomiya. Absence of neutrinos on a lattice. 2. Intuitive topological proof. *Nuclear Physics*, B193:173, 1981.

[15] A. K. Trivedi. The Nielsen-Ninomiya "no go" theorem is false. hep-lat/9309012.

[16] L. Susskind. Lattice fermions. *Phys. Rev.*, D16:3031, 1977.

[17] A. Gonzalez-Arroyo, F.J. Yndurain, and G. Martinelli. computation of the relation between the quark masses in lattice gauge theories and on the continuum. *Physics Letters*, 117B:437, 1982.

[18] M. F. L. Golterman and J. Smit. selfenergy and flavor interpretation of staggered fermions. *Nuclear Physics*, B245:61, 1984.

[19] M. Creutz. Monte carlo study of quantized $SU(2)$ gauge theory. *Phys. Rev.*, D21:2308, 1980.

[20] N. Cabibbo and E. Marinari. A new method for updating $SU(N)$ matrices in computer simulations of gauge theories. *Physics Letters*, 119B:387, 1982.

[21] S. L. Adler and Tsvi Piran. Relaxation methods for gauge field equilibrium equations. *Rev. Mod. Phys.*, 56:1, 1984.

[22] S. L. Adler. Overrelaxation algorithms for lattice field theories. *Phys. Rev.*, D37:458, 1988.

[23] J. E. Mandula and M. Ogilvie. Efficient gauge fixing via overrelaxation. *Physics Letters*, B248:156, 1990.

[24] M. L. Paciello, C. Parrinello, S. Petrarca, B. Taglienti, and A. Vladikas. $SU(3)$ lattice gauge fixing with overrelaxation and Gribov copies. *Physics Letters*, B276:163, 1992.

[25] M. R. Hestenes. *Conjugate direction methods in optimization.* Springer-Verlag, New York, 1980.

[26] S. Güsken, U. Low, K. H. Mutter, R. Sommer, A. Patel, and K. Schilling. Nonsinglet axial vector couplings of the baryon octet in lattice QCD. *Physics Letters*, B227:266, 1989.

[27] C. R. Allton, R. M. Baxter, S. P. Booth, K. C. Bowler, S. Collins, H. Duong, D. S. Henty, A. Hulsebos, A. C. Irving, R. D. Kenway, G. Martinelli, C. McNeile, A. McKerrell, C. Michael, B. J. Pendleton, M. Prisznyak, C. T. Sachrajda, D. G. Richards, J. N. Simone, A. D. Simpson, P. W. Stephenson, and B. E. Wilkes. Quenched hadrons using Wilson and $O(a)$ improved fermion actions at $\beta = 6.2$. *Physics Letters*, B284:377, 1992.

[28] M. Göckeler, R. Horsley, E. M. Ilgenfritz, H. Perlt, G. Schierholz, and A. Schiller. Polarized and unpolarized nucleon structure functions from lattice QCD. *Phys. Rev.*, D53:2317, 1996.

[29] R. L. Jaffe. G_2: the nucleon's other spin dependent structure function. *Comments Nucl. Part. Phys*, 14:239, 1990.

[30] R. L. Jaffe and X.-D. Ji. Chiral odd parton distributions and polarized Drell-Yan. *Physics Letters*, B266:458, 1991.

[31] R. L. Jaffe. Spin, twist and hadron structure in deep inelastic processes. In *Ettore Majorana International School of Nucleon Structure*, Erice, Italy, 1995.

[32] G. Martinelli, C. Pittori, C. T. Sachrajda, M. Testa, and A. Vladikas. A general method for nonperturbative renormalization of lattice operators. *Nuclear Physics*, B445:81, 1995.

[33] R. L. Jaffe and X.-D. Ji. Chiral odd parton distributions and Drell-Yan processes. *Nuclear Physics*, D375:527, 1992.

[34] J. E. Humphreys. *Introduction to Lie Algebras and Representation Theory.* Springer-Verlag, New York, 1972.

[35] J. E. Mandula. Representations of the rotation reflection symmetry group of the four-dimensional cubic lattice. *Nuclear Physics*, B228:91, 1983.

[36] G. Beccarini, M. Bianchi, S. Capitani, and G. Rossi. Deep inelastic scattering in improved lattice QCD. 2. the second moment of structure functions. *Nuclear Physics*, B456:271, 1995.

[37] C. Itzykson and J.-B. Zuber. *Quantum Field Theory.* McGraw–Hill, New York, 1980.

[38] T. Muta. *Foundations of quantum chromodynamics : an introduction to perturbative methods in gauge theories.* World Scientific, Singapore, 1987.

[39] Hikaru Kawai, Ryuichi Nakayama, and Koichi Seo. Comparison of the lattice lambda parameter with the continuum lambda parameter in massless QCD. *Nuclear Physics*, B189:40, 1981.

[40] S. Kronfeld and D. M. Photiadis. Phenomenology on the lattice: composite operators in lattice gauge theory. *Phys. Rev.*, D31:2939, 1985.

[41] S. Capitani and G. Rossi. Deep inelastic scattering in improved lattice QCD. 1. the first moment of structure functions. *Nuclear Physics*, B433:351, 1995.

[42] S. Aoki, M. Doui, T. Hatsuda, and Y. Kuramashi. Tensor charge of the nucleon in lattice QCD.

[43] M. Göckeler, R. Horsley, E. M. Ilgenfritz, H. Perlt, P. Rakow, G. Schierholz, and A. Schiller. Lattice computation of structure functions. In *Lattice 96: 14th International Symposium on Lattice Field Theory*, 1996.

[44] M. Yamada and Hiroshi Koibuchi. Topology collapse of nontrivial $U(1)$ bundle on latticized sphere. *Modern Physics Letters*, A3:1489, 1988.

[45] B. Alles, A. Di Giacomo, and M. Giannetti. Topological susceptibility from different definitions of topological charge in lattice QCD. *Physics Letters*, B249:490, 1990.

[46] M. Lüscher. Topology of lattice gauge fields. *Comm. in Math. Phys.*, 85:39, 1982.

[47] A. V. Phillips and D. A. Stone. The computation of characteristic classes of lattice gauge fields. *Comm. in Math. Phys.*, 131:255, 1990.

[48] B. Berg and M. Lüscher. Definition and statistical distributions of a topological number in the lattice $O(3)$ σ model. *Nuclear Physics*, B190:412, 1981.

[49] V. A. Novikov, M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov. Two-dimensional σ models: modeling nonperturbative effects of quantum chromodynamics. *Phys. Rev.*, 116:103, 1984.

[50] F. D. M. Haldane. Continuum dynamics of the 1-D Heisenberg antiferromagnet: identification with the $O(3)$ σ model. *Physics Letters*, 93A:464, 1983.

[51] F. D. M. Haldane. Nonlinear field theory of large-spin Heisenberg antiferromagnets: semiclassically quantized solutions of the one-dimensional easy-axis Néel state. *Phys. Rev. Letters*, 50:1153, 1983.

[52] U. Glaus and T. Schneider. Critical properties of spin-1 Heisenberg chain with uniaxial anisotropy. *Phys. Rev.*, B30:215, 1984.

[53] I. Affleck. Quantum spin chains and the Haldane gap. *J. Phys.: Condens. Matter*, 1:3047, 1989.

[54] E. H. Lieb, T. Schultz, and D. J. Mattis. ?? *Ann. Phys.*, 16:407, 1961.

[55] G. Bhanot, R. Dashen, N. Seiberg, and H. Levine. Scaling and θ dependence in the $O(3)$ σ-model. *Phys. Rev. Letters*, 53:519, 1984.

[56] A. A. Belavin, A. M. Polyakov, A. S. Shvarts, and Yu. S. Tyupkin. Pseudoparticle solutions of the Yang-Mills equations. *Physics Letters*, 59B:85, 1975.

[57] R. G. Edwards and A. D. Sokal. Dynamic critical behavior of wolff's collective mode Monte Carlo algorithm for the two-dimensional $O(N)$ nonlinear σ model. *Phys. Rev.*, page 1374, 1989.

[58] B. Mehlig, D. W. Heermann, and A. L. C. Ferreira. Critical dynamics of the hybrid Monte Carlo algorithm. *Physics Letters*, B291:151, 1992.

[59] J. Sloan, D. Kusnezov, and A. Bulgac. Critical behavior of global demons and hybrid Monte Carlo in the XY model. *Physics Letters*, B296:379, 1992.

[60] G. Mana, T. Mendes, A. Pelissetto, and A. D. Sokal. Dynamic critical behavior of multigrid Monte Carlo for two-dimensional nonlinear σ models. *Nucl. Phys. B (Proc. Suppl.) 47*, page 796, 1996.

[61] R. H. Swendsen and J.-S. Wang. Nonuniversal critical dynamics an Monte Carlo simulations. *Phys. Rev. Letters*, 58:86, 1987.

[62] U. Wolff. Comparison between cluster Monte Carlo algorithms in the Ising model. *Physics Letters*, B228:379, 1989.

[63] U. Wolff. Cluster algorithms for nonlinear σ model. In *Workshop on Fermion Algorithms*, page 213, Julich, Germany, 1991.

[64] A. A. Milne. *Winnie-the-Pooh.* Methuen & Co., Ltd, 1926.

[65] U.-J. Wiese. Numerical simulation of lattice θ vacua: the $2-D$ $U(1)$ gauge theory as a test case. *Nuclear Physics*, B318:153, 1989.

[66] N. Seiberg. θ physics at strong coupling. *Phys. Rev. Letters*, 53:637, 1984.

[67] I. Affleck. Field theory methods and quantum critical phenomena. In E. Brezin and J. Zinn-Justin, editors, *Fields, Strings and Critical Phenomena*, page 563, Les Houches, 1988.

[68] I. Affleck, D. Gepner, H. J. Schulz, and T. Ziman. Critical behaviour os spin-s Heisenberg antiferromagnetic chains: analitic and numerical results. *J. Phys.*, A22:511, 1989.

[69] I. Affleck. Nonlinear σ model at $\theta = \pi$: euclidean lattice formulation and solid-on-solid models. *Phys. Rev. Letters*, 66:2429, 1991.

[70] S. Olejnik and G. Schierholz. On the existence of a first order phase transition at small vacuum angle theta in the CP^3 model. In *Nucl. Phys. B (Proc. Suppl.) 34*, page 709, 1994.

[71] K. Jansen and U.-J. Wiese. Cluster algorithms and scaling in CP^3 and CP^4 models. *Nuclear Physics*, B370:762, 1992.

[72] Donald E. Knuth. *The Art of Computer Programming.* Addison–Wesley, Reading, MA, 1981. v. 2.

[73] Thinking Machines Corporation. *C* Programming Guide*, 1993.

[74] Thinking Machines Corporation. *The Connection Machine CM–5 Technical Summary*, 1992.

[75] B. Efron and R. Tibshirani. *An introduction to the bootstrap.* Chapman & Hall, New York, 1993.

[76] J. Shao and D. Tu. *The jackknife and bootstrap.* Springer Verlag, New York, 1995.