David's Research Notes

David Clarke

March 31, 2020

Math Symbols

$\exists$	There exists
$\forall$	For all
$\in$	Is a member of the set
$\leq$	Less than or equal to; is a subgroup of
$\approx$	Approximate equality
$\cong$	Isomorphic
$\equiv$	Is defined as; modular congruency
$\sim$	Equivalence relation
$\simeq$	Asymptotic equality
$\mathbb{C}$	Complex numbers
c_x	$\cos(x)$
$\log$	Natural logarithm
$\mathbb{N}$	Natural numbers
$\mathcal{O}$	Big O notation; topology on a set
$\mathbb{Q}$	Rational numbers
q	Gaussian or Student difference test; goodness-of-fit
$\mathbb{R}$	Real numbers
s_x	$\sin(x)$
$\mathrm{SU}(N)$	Special unitary group of degree N
S_V	Symmetric group on the set V
τ_{int}	Integrated autocorrelation time
$\mathbb{Z}$	Integers
$\mathbb{Z}_n$	Integers modulo n

Physics Symbols

a	Lattice spacing
β	SU(N_c) coupling constant; beta function
χ	Topological susceptibility; Polyakov loop susceptibility
c	Speed of light
G	A group; Newton's gravitational constant
g	A group element; SU(N_c) bare coupling
$\hbar$	Planck's constant
k_B	Boltzmann's constant
N_c	Number of colors
N_f	Number of fermion flavors
N_s	Lattice extension in a spatial direction
N_τ	Lattice extension in temperature/Euclidean time direction
σ	String tension
σ_i	A Pauli matrix
T_c	Deconfining phase transition temperature

High Energy Glossary

Baryogenesis	Physical processes that must have occurred during the early universe that are responsible for the imbalance between the amount of matter and antimatter.
Baryon	Three quark hadron.
Baryon number (B)	(# baryons − # anti-baryons)
Bottomness (B′)	−(# bottom quarks − # bottom anti-quarks)
Branching ratio	Ratio of decay rate of particle along a particular channel to the total decay rate of that particle.
Charm (C)	(# charm quarks − # charm anti-quarks)
Effective field theory	Field theory that is valid only at certain energy scales. The SM is therefore an effective field theory.
Field equation	Any partial differential equation describing the dynamics of some physical field.
Hadron	Color neutral particle made of multiple quarks. See e.g. meson, baryon, nucleon.
Lattice units	Units where $a = 1$ and is dimensionless.
Meson	Two quark hadron.
Naturalness	The property that free parameters or physical constants appearing in a theory should take relative values of order 1.
Nucleon	Baryons found in atomic nuclei.
Physical units	Units where a carries length.
Planck mass	Mass at which the gravitational coupling becomes of order one.

$$\alpha_G = \frac{GM_P^2}{\hbar c} \sim 1 \;\Rightarrow\; M_P \sim \sqrt{\frac{\hbar c}{G}} \approx 10^{19}\ \text{GeV}$$

Semi-leptonic decay	Weak decay of a hadron producing a lepton, corresponding neutrino, and one or more hadrons.
Strangeness (S)	$-$(# strange quarks $-$ # strange anti-quarks)
Topness (T)	(# top quarks $-$ # top anti-quarks)

Some Important Particles

Name	*Significance*	*Symbol*	*Composition*
B meson:	Light mesons with bottomness 1. Can use weak decays to extract some CKM matrix elements.	B^0 B^+ B^-	$d\bar{b}$, $b\bar{d}$ $u\bar{b}$ $b\bar{u}$
D meson:	Light mesons with charm 1. Can use weak decays to extract some CKM matrix elements.	D^0 D^+ D^-	$c\bar{u}$, $u\bar{c}$ $c\bar{d}$ $d\bar{c}$
Eta:	$\eta - \eta'$ mass difference can be explained by topology.	η η'	$u\bar{u} + d\bar{d} - 2s\bar{s}$ $u\bar{u} + d\bar{d} + s\bar{s}$
Kaon:	Light mesons with strangeness 1. Can use weak decays to extract some CKM matrix elements. Its decay constant is used as a reference in lattice calculations.	K^0 K^+ K^-	$d\bar{s}$, $s\bar{d}$ $u\bar{s}$ $s\bar{u}$
Pion:	Lightest mesons. "Would-be" Goldstone bosons of chiral symmetry breaking.	π^0 π^+ π^-	$u\bar{u}$, $d\bar{d}$ $u\bar{d}$ $d\bar{u}$
Psion:	Discovery of the charm quark.	J/ψ	$c\bar{c}$
Upsilon:	Discovery of the bottom quark.	Υ	$b\bar{b}$

Preface

This "book" is a collection of research notes that I made while researching in physics. The first chapters are dedicated to mainly mathematics, and the last chapters are dedicated to mainly physics, with an aim toward lattice field theory. Attached at the end are also some appendices. Appendix A includes some special topics in mathematics that, while sometimes useful for a physicist, don't really fit in one of the mathematics chapters at the beginning. Appendix B includes special topics in physics that I find interesting, but aren't directly relevant to any research I work on. As I continue researching, I may want to develop some of these appendices into full chapters, if they ever become directly relevant to my research.

Throughout this book, I assume the reader has some basic knowledge of mathematics, for instance what it means for operations to be *associative* or *commutative*. The reader will also find it useful to have some basic knowledge of calculus, linear algebra, differential equations, and complex analysis. (Just enough to get by in undergraduate level physics classes.)

Contents

Chapter 1

Math: Group Theory

This chapter focuses on some the basic ideas of group theory and some applications of group theory to physics. Groups appear often in physics: Boosts and rotations in special relativity generate the Lorentz group, sets of gauge transformations equipped with function composition form gauge groups, etc. Some of this presentation follows sections of Dummit and Foote [1] and Georgi [2].

1.1 Preliminaries

In order for us to understand any of the above paragraph, we must lay down a few definitions and mathematical preliminaries. A *binary operation* $\bullet$ on a set G is a function $\bullet : G \times G \to G$. A *group* , then, is a set G equipped with a binary operation $\bullet$ that satisfies the following axioms:

1. $\bullet$ is associative.

2. $\exists\, \mathbf{1} \in G$ such that $\forall\, g \in G$,

$$\mathbf{1} \bullet g = g \bullet \mathbf{1} = g. \tag{1.1}$$

   This element $\mathbf{1}$ is called the *identity*.

3. $\forall\, g \in G \quad \exists\, g^{-1} \in G$, called the *inverse* of g, such that

$$g^{-1} \bullet g = g \bullet g^{-1} = \mathbf{1}. \tag{1.2}$$

If group elements commute under $\bullet$ the group is said to be *abelian* . The *order* of a group, denoted $|G|$, is the number of unique elements in the group[1]. A *subgroup* H of G is a non-empty subset of G that itself forms a group under $\bullet$ and in this case we will write $H \leq G$. (It should be clear from context whether this symbol indicates group organization or magnitude.) Finally a group is *cyclic* if it is generated by a single element; that is, if $\exists\, g \in G$ such that $G = \{g^n : n \in \mathbb{Z}\}$.

It's actually not too common for mathematicians or physicists to write the $\bullet$ explicitly when showing the composition of two elements. So for example you will often see gh as shorthand $g \bullet h$. In general I will only refer to operations on algebraic structures explicitly when giving the definition of that structure. Therefore you can expect to see gh instead of $g \bullet h$ from here on out.

Proposition 1.1.1

A subset H of G is a subgroup of G if and only if

$$a, b \in H \Rightarrow ab^{-1} \in H$$

Proof. ($\Rightarrow$) Follows immediately from the definition of a subgroup. To show ($\Leftarrow$) let $b \in H$. Then by the above conditional, $bb^{-1} \in H$, which shows $\mathbf{1} \in H$. To show the existence of inverses in H, note $\mathbf{1}, b \in H \Rightarrow \mathbf{1}b^{-1} \in H \Rightarrow b^{-1} \in H$. Finally, associativity is inherited from G. □

Example

1. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are all groups under addition, and

$$\mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}. \tag{1.3}$$

   Each of these sets with 0 removed forms a group under multiplication. (We have to remove 0 because it has no multiplicative inverse.)

[1] This is at least true for finite groups.

2. Let $n \in \mathbb{N}$ and define an equivalence relation on $\mathbb{Z}$ by

$$a \equiv b \ (mod\ n) \ \Leftrightarrow \ n \mid (b-a). \tag{1.4}$$

(We read this as "a is congruent to b modulo n" or "a is congruent to b mod n.") Define an equivalence class by

$$\bar{a} \coloneqq \{a + mn \ : m \in \mathbb{Z}\}. \tag{1.5}$$

The set of all such equivalence classes is called *the integers modulo n* and is denoted by $\mathbb{Z}/n\mathbb{Z}$ or $\mathbb{Z}_n$. It forms a group under the "addition" operation exemplified below. As a concrete illustration, take $\mathbb{Z}_4$. It is of order 4, with elements $\bar{0}$, $\bar{1}$, $\bar{2}$, and $\bar{3}$. To see how addition works, note

$$\begin{aligned} \bar{1} + \bar{2} &= \{(1+2) + 4(m_1 + m_2) \ : m_1, m_2 \in \mathbb{Z}\} \\ &= \{3 + 4m \ : m \in \mathbb{Z}\} \\ &= \bar{3} \end{aligned} \tag{1.6}$$

and

$$\begin{aligned} \bar{3} + \bar{3} &= \{(3+3) + 4(m_3 + m_4) \ : m_3, m_4 \in \mathbb{Z}\} \\ &= \{2 + 4(1 + m_3 + m_4) \ : m_3, m_4 \in \mathbb{Z}\} \\ &= \{2 + 4m \ : m \in \mathbb{Z}\} \\ &= \bar{2} \end{aligned} \tag{1.7}$$

It should be clear from the above that $\bar{0}$ is the identity element and $\mathbb{Z}_n$ is abelian and cyclic. I would also like to emphasize that the addition defined in eqs. (1.6) and (1.7) is not the same addition as over the integers, even though I have chosen the same symbol for both cases. One should always be careful of what group operation is meant when the author is being lazy.

3. Sets of objects besides numbers also form groups. For example let V be any non-empty set of objects and let S_V be the set of all permutations of V. Then S_V forms a group under function composition called the *symmetric group* on the set V.

In many situations, one encounters two groups of the same order that behave in essentially the same ways. In a sense, one group is the same the original group, masquerading about as a unique mathematical object. We could recover the original group by merely relabeling elements of the masked group and viewing its operation differently. Let us make these ideas more precise. Let G and H be groups with operations $\bullet$ and $\circ$, respectively. A map $\phi : G \to H$ satisfying

1. $\phi(\mathbf{1}_G) = \mathbf{1}_H$ and
2. $\phi(g \bullet h) = \phi(g) \circ \phi(h)$

$\forall\, g, h \in G$ is called a *homomorphism*. If in addition to the above property we know that ϕ is a bijection, ϕ is said to be an *isomorphism* of G and H, we say that G and H are *isomorphic*, and we write $G \cong H$. Finally an *endomorphism* is a homomorphism mapping a group to itself, and an *automorphism* is a bijective endomorphism (isomorphism from a group to itself).

The above definition shows us that when $G \cong H$, H is really just G in disguise. Since ϕ is a bijection we have a way of associating each element of H with exactly one element of G and vice-versa (one could consider ϕ the relabeling), and the fact that ϕ is a homomorphism shows us that the binary operations of the groups act the same way.

1.2 Quotient groups

Let's add some structure to groups. In this section we will discuss a way of creating a new group by partitioning an old one.

An *equivalence relation* $\sim$ on a set G is a binary operation that has the following properties $\forall\, x, y, z \in G$:

1. it is *reflexive*, i.e. $x \sim x$;
2. it is *symmetric*, i.e. $x \sim y \Leftrightarrow y \sim x$; and
3. it is *transitive*, i.e. $x \sim y$ and $y \sim z \Rightarrow x \sim z$.

Let $g \in G$. The set $\bar{g} = \{x \in G : x \sim g\}$ is called an *equivalence class*.

So we see that equivalence relations are just generalizations of the $=$ sign. We actually use equivalence relations more than you probably realize. For

example when $x, y \in \mathbb{R}$ and $y = x + \epsilon$, where ϵ is small, we often just write $y = x$. Strictly speaking, these two quantities are not equal, but it's not hard to show that "equal to first order in ϵ" defines an equivalence relation.

Let $H \leq G$. The *left coset* of H with respect to G is the set

$$aH = \{ah : \ h \in H\}.$$

The *right coset* is defined similarly, but with h and a interchanged. We are now going to see that cosets define equivalence classes. To begin with let $H \leq G$; $x, y \in H$; and define a binary relation $\sim$ by

$$x \sim y \Leftrightarrow xy^{-1} \in H. \tag{1.8}$$

Proposition 1.2.1

$\sim$ is an equivalence relation.

Proof. We just need to check that it satisfies all the defining properties.

1. $xx^{-1} = \mathbf{1} \Rightarrow x \sim x$.
2. $x \sim y \Rightarrow xy^{-1} \in H \Rightarrow (xy^{-1})^{-1} \in H \Rightarrow yx^{-1} \in H$ since H is a group, so its elements have inverses. Hence $y \sim x$.
3. $x \sim y$ and $y \sim z \Rightarrow xy^{-1} \in H$ and $yz^{-1} \in H$. Then since H is closed under multiplication, $xy^{-1}yz^{-1} = xz^{-1} \in H \Rightarrow xz^{-1} \in H$ $\Rightarrow x \sim z$.

□

According to this definition, the equivalence classes of G are the sets $A \subset G$ satisfying $\forall\, x, y \in A \ xy^{-1} \in H$. To see why this is important, fix $a \in A$. Then $\forall\, x \in A$,

$$\begin{aligned} x \sim a &\Rightarrow xa^{-1} \in H \\ &\Rightarrow xa^{-1} = h, \ h \in H \\ &\Rightarrow x = ha, \ h \in H \\ &\Rightarrow A = Ha. \end{aligned} \tag{1.9}$$

In other words, the equivalence classes are the right cosets of H! If we instead defined the equivalence relation by $x \sim y \Leftrightarrow x^{-1}y \in H$, we would have found the equivalence classes to be the left cosets of H.

We can now construct our new group. The group is a collection of cosets, but it only forms a group if the cosets are special. In particular, a subgroup $N \leq G$ is said to be *normal* if $\forall\, g \in G$, $gN = Ng$. In this case we write $N \trianglelefteq G$. The *quotient group* G/N is the set of all cosets of N. We define an operation on G/N by

$$(xN)(yN) = x(Ny)N = x(yN)N = (xy)(NN) = (xy)N. \tag{1.10}$$

Hence we see why it was so crucial that the subgroup be normal: It allowed us to move the y past N in the second step, thus guaranteeing G/N is closed under this operation.

Proposition 1.2.2

G/N forms a group under the above operation.

Proof. This is pretty clearly a group because G is. The identity is $\mathbf{1}N$, and each xN has inverse $x^{-1}N$ □

Proposition 1.2.3

G/N forms a partition of G.

Proof. Let $g \in G$. We have to check that (1) g belongs to some coset of N, and further that (2) g doesn't belong to two different cosets.

1. Let $n \in N$ and define $g' = n^{-1}g$. Then $ng' = nn^{-1}g = g$, so $g \sim g'$.
2. Let $x, y \in G$. If $g \sim x$ and $g \sim y$ then $x \sim y$ since $\sim$ is transitive.

□

To summarize: If we find a normal subgroup N of G, we can make a new group G/N by partitioning G into disjoint cosets of N, which turn out to be disjoint equivalence classes. These quotient groups are useful and pop up pretty frequently; indeed the group $\mathbb{Z}/n\mathbb{Z}$, which we encountered in the first section, is a quotient group. Now you understand the notation.

Example

Take $G = (\mathbb{Z}, +)$ and $N = 4\mathbb{Z}$. Clearly $4\mathbb{Z}$ is a normal subgroup of $\mathbb{Z}$ since it's commutative. To form the quotient group $\mathbb{Z}/4\mathbb{Z}$, we find all cosets of $4\mathbb{Z}$. These are

- $4\mathbb{Z}$,
- $4\mathbb{Z} + 1 = \{4m + 1 : m \in \mathbb{Z}\}$,
- $4\mathbb{Z} + 2$, and
- $4\mathbb{Z} + 3$.

To check for instance that $4\mathbb{Z}+1$ is an equivalence class, we just need to show that $x, y \in 4\mathbb{Z} + 1 \Rightarrow xy^{-1} \in 4\mathbb{Z}$. Let $x = 4m + 1$ and $y = 4n + 1$ for some $m, n \in \mathbb{Z}$. Then

$$xy^{-1} = x - y = 4m + 1 - (4n + 1) = 4(m - n) \in 4\mathbb{Z}.$$

(This is an instance where the notation can be confusing; remember $\mathbb{Z}$ is being considered as a group under addition, so applying the inverse group operation means subtracting.) To see that there are no other cosets, note that $4\mathbb{Z} + 4 = 4\mathbb{Z}$.

1.3 Group representations

It turns out that many groups are isomorphic to sets of linear transformations on vector spaces, and similarly that many sets of matrices equipped with matrix multiplication form groups. In many cases these isomorphisms make dealing with groups less abstract and more manageable; after all, everyone is comfortable with matrices. In this section we will make these ideas precise, but we will need to use linear algebra. For a very brief review of the relevant bits, see the appendix.

The set of all automorphisms of a vector space V is called the *automorphism group* of V or the *general linear group* and is denoted $GL(V)$. The set of all $n \times n$ matrices with entries from a field F is called the *general linear group of degree* n and is denoted by $GL_n(F)$. The reader may be concerned

that the nomenclature of the previous definition is poorly chosen. However if V is a vector field over the field F, the groups $GL(V)$ and $GL_n(F)$ are actually just two ways of viewing the same thing. We shall demonstrate this fact with the following theorem.

Theorem 1.3.1

Let V be an n-dimensional vector space over F. Then

$$GL(V) \cong GL_n(F).$$

Let G be a group, F a field, and V a vector space over F. A *linear representation* of G is any homomorphism $D : G \to GL(V)$. A representation is said to be *faithful* if it is injective. The *dimension* of a representation is the dimension of V.

Example

1. Every group has a *trivial* representation, $D(g) = 1$.
2. $\mathbb{Z}_3$ has a 1D representation

$$D(\bar{0}) = 1, \quad D(\bar{1}) = e^{2\pi i/3}, \quad D(\bar{2}) = e^{4\pi i/3}. \tag{1.11}$$

In the original group addition modulo n is the binary operation, while the representation takes ordinary multiplication on the reals. An example of a 3D representation of $\mathbb{Z}_3$ is

$$D(\bar{0}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D(\bar{1}) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad D(\bar{2}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{1.12}$$

The representation of eq. (1.12) is constructed by the following general prescription. Take $V = \mathbb{R}^n$ and imagine that the group elements form an

orthonormal basis; i.e. $|e_i\rangle = |g_i\rangle$. Now define

$$D(g_1)\,|g_2\rangle = |g_1 g_2\rangle\,. \tag{1.13}$$

The dimension of this representation is clearly just the order of the group. One can recover the matrix elements via

$$[D(g)]_{ij} = \langle e_i|\,D(g)\,|e_j\rangle\,. \tag{1.14}$$

This is called the *regular* representation. Another advantage of using linear spaces to represent groups is that we can make a change of basis for our convenience. You will recall this is achieved by similarity transformations of the form

$$D(g) \to D'(g) = S^{-1}D(g)S. \tag{1.15}$$

This transformation clearly preserves the group multiplication rules because S cancels with its inverse.

Let D be a representation of G over a vector space V with subspace W. W is an *invariant subspace* of D if $\forall w \in W$

$$D(g)w \in W. \tag{1.16}$$

If D has an invariant subspace it is said to be *reducible*; otherwise it is *irreducible*. We will shorten irreducible representation as irrep. If two representations D and D' are related by eq. (1.15) they are *equivalent*. If D is equivalent to a representation whose matrix elements are in block diagonal form

$$D(g) = \begin{pmatrix} D_1(g) & 0 & \\ 0 & D_2(g) & \\ & & \ddots \end{pmatrix} \tag{1.17}$$

where the D_i are irreps, then D is called *completely reducible*. Sometimes this is written as

$$D = D_1 \oplus D_2 \oplus ... \tag{1.18}$$

and we say that D is the *direct sum* of the representations D_i.

Theorem 1.3.2

Every representation of a finite group is equivalent to a unitary transformation.

Theorem 1.3.3

Every representation of a finite group is completely reducible.

1.4 Young tableaux

Recall from Section 1.1 the permutation group on n objects, S_n. Any element in S_n can be written as a product of cycles, which are just cyclic permutations of subsets. Conventional notation writes a cycle as a list of numbers between parenthesis, indicating the set of elements are are permuted.

Example

Consider permutations of $V = \{x_1, ..., x_n\}$. Then

1. (1) takes $x_1 \to x_1$;
2. (238) takes $x_2 \to x_3 \to x_8 \to x_2$;
3. a cycle of length k is a *k-cycle*;
4. **1**=(1)(2)...(n); and
5. $(12)(3)...(n) \in S_n$ interchanges x_1 and x_2 while leaving all other elements fixed.

A simple nD representation of S_n permutes the orthonormal basis vectors of $\mathbb{R}^n$. If a permutation π takes x_i to x_j then

$$D(\pi)\left|i\right\rangle = \left|j\right\rangle, \tag{1.19}$$

which implies

$$D(\pi)_{li} = \left\langle l\right| D\left|i\right\rangle = \delta_{lj}. \tag{1.20}$$

This is called the *defining representation*.

A set is called a *conjugacy class* if $gSg^{-1} = S$. If you have a group element g_1, the *conjugation* of g_1 is gg_1g^{-1}. From the definition of normal subgroup in Section 1.2 we see that normal subgroups are conjugacy classes. The conjugacy classes of of permutations are just the cycle structure; for instance all interchanges are in the same conjugacy class. One way of seeing

this is by looking at the defining representation. If you conjugate using an interchange, all this does is switch the basis vectors $|i\rangle$ and $|j\rangle$, which clearly has no impact on the cycle structure. Then, since any permutation can be built from interchanges, it follows that an arbitrary conjugation preserves the cycle structure.

1.5 Lie algebras and Lie groups

A *Lie algebra* is a vector space L over a field F together with an operation $[\ ,\ ] : L \times L \to L$ called the *Lie bracket*that satisfies $\forall\, \alpha, \beta \in F$ and $x, y, z \in L$

1. $[\alpha x + \beta y, z] = \alpha[x, z] + \beta[y, z]$ (the Lie bracket is *bilinear*),
2. $[x, x] = 0$ (*alternating* on L), and
3. $[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$ (and satisfies the *Jacobi identity*).

The basis elements T^a of L are called the *generators*.

Proposition 1.5.1

The Lie bracket is anticommutative.

Proof. $0 = [x+y, x+y] = [x, x]+[x, y]+[y, x]+[y, y] = [x, y]+[y, x]$. □

Lie Algebras are used in physics to construct Lie groups. Lie groups are used to collect and analyze continuous symmetries of systems and structures. Let us see a way to construct Lie groups in physics. We will call the Lie group G, and the groups elements $g(\alpha) \in G$ will depend smoothly on a set of continuous parameters α. When we say "smooth", we mean that if two group elements are "close together" in G, their parameters are also close together. The identity is an important element in the group, so we will parameterize elements with respect to it. We will shorthand

$$\alpha = \left(\alpha^1,\, \alpha^2,\, ...,\, \alpha^N\right), \tag{1.21}$$

where $\alpha^a \in \mathbb{R}$. For our parameterization we set

$$g(0) = \mathbf{1}. \tag{1.22}$$

Then when we find a representation of the group, it will be parameterized in the same way, so that

$$D(0) = \mathbf{1}. \tag{1.23}$$

In a neighborhood of $\mathbf{1}$ we can expand D. We find

$$D(\epsilon\alpha) = \mathbf{1} + i\epsilon\alpha^a T^a + \mathcal{O}\big(\epsilon^2\big), \tag{1.24}$$

where

$$T^a \equiv -i\frac{\partial}{\partial\alpha^a} D(\alpha)\Big|_{\alpha=0}. \tag{1.25}$$

If we can identify the T^a here with the Lie algebra T^a, we can see how it Lie algebra generates the Lie group. The i is included so that if the representation is unitary, the T^a are Hermitian.

We can move in a fixed direction away from the identity using eq. (1.24) by simply raising it to some power. This suggests defining the representation of the group elements as

$$D(\alpha) = \lim_{k\to\infty} (1 + i\alpha^a T^a/k)^k = \exp(i\alpha^a T^a). \tag{1.26}$$

In the limit, this expression clearly goes to the representation of a group element, because k becomes large, which means the term in parentheses is a group element, which means the whole product is a group element, since the product of group elements stays in the group.

Since the exponentials are group elements, it must be that

$$\exp(i\alpha^a T^a)\exp\big(i\beta^b T^b\big) = \exp(i\delta^c T^c) \tag{1.27}$$

for some δ. Since our parameterization is smooth, we can solve for δ by Taylor expanding both sides. We write

$$i\delta^c T^c = \log\big(1 + \exp(i\alpha^a T^a)\exp\big(i\beta^b T^b\big) - 1\big). \tag{1.28}$$

Keeping terms up to only second order in α and β we find

$$i\delta^c T^c = i\alpha^a T^a + i\beta^b T^b - \frac{1}{2}[\alpha^a T^a, \beta^b T^b]. \tag{1.29}$$

We can rearrange the above to find

$$[\alpha^a T^a, \beta^b T^b] = -2i(\delta^c - \alpha^c - \beta^c)T^c \equiv i\gamma^c T^c, \tag{1.30}$$

where the bracket here denotes the ordinary commutator. From the LHS of the above, we can see that the γ^c must be some sum of products of the α^a and β^b, so we can write

$$\gamma^c = \alpha^a \beta^b f^{abc} \tag{1.31}$$

for some constants f^{abc}. Hence

$$[T^a, T^b] = i f^{abc} T^c. \tag{1.32}$$

The f^{abc} are called the *structure constants* . We have

$$f^{abc} = -f^{bac} \tag{1.33}$$

since $[A, B] = -[B, A]$. The structure constants essentially summarize the group multiplication law. Sometimes we refer to the commutator relation (1.32) as the Lie algebra, which makes sense because it essentially gives you a prescription for how the Lie bracket works.

Something worth noting is that we follow the same steps as above to prove the Campbell-Baker-Hausdorff formula. In particular if X and Y are non-commuting matrices, and ϵ is small, we can expand

$$\log\left(1 + \exp(\epsilon X)\exp(\epsilon Y) - 1\right) \tag{1.34}$$

to second order in ϵ and then do some rearranging.

Theorem 1.5.1: Campbell-Baker-Hausdorff formula

$$\exp(\epsilon X)\exp(\epsilon Y) = \exp\left(\epsilon X + \epsilon Y + \frac{1}{2}[\epsilon X, \epsilon Y] + \mathcal{O}\left(\epsilon^3\right)\right).$$

1.5.1 SU(2)

The SU(2) algebra is defined by

$$[J_i, J_j] = i\epsilon_{ij} J_k. \tag{1.35}$$

In units $\hbar = 1$, this is the algebra of the familiar angular momentum operator from QM.

Proposition 1.5.2

Let $A, B \in \mathrm{SU}(2)$ with parameterization

$$A = a_0\mathbf{1} + i\vec{a}\cdot\vec{\sigma}, \qquad B = b_0\mathbf{1} + i\vec{b}\cdot\vec{\sigma},$$

where the σ_i are the usual Pauli matrices. Then

$$AB = \mathbf{1}\left(a_0b_0 - \sum_{i=1} a_ib_i\right) + i\sum_{i=1}\sigma_i\left(a_0b_i + a_ib_0 - \sum_{j\neq k} a_jb_k\epsilon_{jki}\right).$$

Proof.

$$\begin{aligned} AB &= \left(a_0\mathbf{1} + i\sum_i a_i\sigma_i\right)\left(b_0\mathbf{1} + i\sum_i b_i\sigma_i\right) \\ &= \mathbf{1}\left(a_0b_0 - \sum_i a_ib_i\right) + i\sum_i \sigma_i\left(a_0b_i + a_ib_0\right) - \sum_{j\neq k} a_jb_k\sigma_j\sigma_k \\ &= \mathbf{1}\left(a_0b_0 - \sum_i a_ib_i\right) + i\sum_i \sigma_i\left(a_0b_i + a_ib_0\right) - i\sum_i\sum_{j\neq k} a_jb_k\epsilon_{jki}\sigma_i. \end{aligned}$$

□

Something worth noting about the parameterization of the above proposition is that it shows a bijection between SU(2) and the three-sphere S^3. In this parameterization, a_μ is a component of a 4D Euclidean vector where $a_0 \equiv a_4$. The vector, which specifies the SU(2) matrix entirely, satisfies $a_\mu a_\mu = 1$, which means it's a point on the surface of a 4D hypersphere; i.e. the point is on S^3. Additionally this parameterization is useful for carrying out matrix multiplications on the computer.

Proposition 1.5.3

Let $A, B \in \mathrm{SU}(2)$. Then

$$\frac{A+B}{\sqrt{\det(A+B)}} \in \mathrm{SU}(2).$$

Proof. This follows from $\det(cA) = c^n \det(A)$ for any $n \times n$ matrix A. □

References

[1] David Dummit and Richard Foote. *Abstract Algebra*. Wiley, United States, 2004. ISBN 978-0-471-43334-7.

[2] Howard Georgi. *Lie Algebras in Particle Physics: from Isospin to Unified Theories*. Westview, Boca Raton, 1999. ISBN 978-0-7382-0233-4.

Chapter 2

Math: Differential Geometry

2.1 Topological manifolds

Why should a physicist care about topology? Frederic Schuller gave a good motivation for learning topology in the 2015 Winter School on Gravity and Light [1]: we need to be able to speak sensibly about continuity in the most general way possible. In particular, we need to understand what it means for maps between manifolds to be continuous. Let M be a set. A *topology on* M is a subset $\mathcal{O}_M \subset 2^M$ satisfying

1. $\emptyset, M \in \mathcal{O}_M$;
2. $U \in \mathcal{O}_M$ and $V \in \mathcal{O}_M \Rightarrow U \cap V \in \mathcal{O}_M$; and
3. $\{U_\alpha\} \in \mathcal{O}_M \Rightarrow \bigcup_\alpha U_\alpha \in \mathcal{O}_M$.

A *topological space* is a pair $(M, \mathcal{O}_M)$. The sets U are said to be *open*, while a set A is said to be *closed* if and only if $M - A$ is open.

Example

1. For any set M, the *chaotic topology* is given by $\mathcal{O}_{\text{chaotic}} = \{\emptyset, M\}$. On the opposite end of the spectrum, the *discrete topology* is simply $\mathcal{O}_{\text{discrete}} = 2^M$.
2. Consider $\mathbb{R}^n$ equipped with the Euclidean metric. The *standard*

topology is

$$\mathcal{O}_{\text{standard}} = \{U \in 2^{\mathbb{R}^n} : \forall\, p \in U,\ \exists\, r > 0 \text{ with } B_r(p) \in U\}.$$

Note in this case that the topological definition of open set agrees with the definition that relies on limit points.

Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be topological spaces. A mapping $f : M \to N$ is *continuous* if $\forall V \in \mathcal{O}_N$, $f^{-1}(V) \in \mathcal{O}_M$. In other words, continuous functions are those that guarantee that you started off in an open set, if you ended up in an open set. Note that whether a function is continuous depends on the topology.

Theorem 2.1.1

Let M, N, and P be open sets with respective topologies and suppose $f : M \to N$ and $g : N \to P$ are continuous. Then $g \circ f : M \to P$ is continuous.

Proof. We need to check that the preimage of $g \circ f$ is open. Note

$$\begin{aligned}(g \circ f)^{-1}(P) &= \{m \in M : g \circ f(m) \in P\} \\ &= \{m \in M : f(m) \in g^{-1}(P)\} \\ &= f^{-1}(g^{-1}(P)).\end{aligned}$$

Since g is continuous, $g^{-1}(P)$ is open. It follows by the continuity of f that $f^{-1}(g^{-1}(P))$ is open, which completes the proof. □

Hence the composition of continuous functions is continuous, which is what you learned in calculus. Using the definition of continuity, we can know determine when two topological spaces are, in some sense, equivalent. A map between two topological spaces $f : M \to N$ is called a *homeomorphism* if

1. f is invertible;
2. f is continuous with respect to the topologies of M and N; and
3. f^{-1} is also continuous.

If such a function exists, we say M and N are *homeomorphic*. So for example when a topological space is homeomorphic to $\mathbb{R}^d$, that topological space is

essentially a warped version of $\mathbb{R}^d$. Also given a topological space, we can always make another one from a subset of that space. In fact

Proposition 2.1.1

Let $(M, \mathcal{O}_M)$ be a topological space. Let $S \subseteq M$ and $\mathcal{O}_M|_S \subseteq 2^S$ such that $\mathcal{O}_M|_S = \{U \cap S : U \in \mathcal{O}_M\}$. Then $(S, \mathcal{O}_M|_S)$ forms a topological space.

Let $d \in \mathbb{N}$. A topological space $(M, \mathcal{O}_M)$ is called a *topological manifold of dimension* d if $\forall\, x \in M$, $\exists\, U$ containing x that is homeomorphic to $\mathbb{R}^d$. Here $\mathbb{R}^d$ is taken with the standard topology. We say:

1. (U, f) is a *chart* of $(M, \mathcal{O}_M)$.
2. An *atlas* $\mathcal{A}$ is a collection of charts (U_α, f_α) such that $M = \bigcup_\alpha U_\alpha$.
3. Sometimes we call the homeomorphism for each chart a *chart map*.

Clearly mathematicians got carried away with sailing analogies. But I think they make the meaning of the terminology easy to remember. Now let's say that I have two charts (U_x, f) and (U_y, g) of points x and y in the manifold, and suppose further they overlap, i.e. $U_x \cap U_y \neq \emptyset$. Consider a point z in the intersection. Each chart gives me a perfectly reasonable representation of z. How can I easily translate from one to the other? Well, since f and g are invertible by definition the map

$$g \circ f^{-1} : f(U_x \cap U_y) \to g(U_x \cap U_y)$$

will do the trick. $g \circ f^{-1}$ is called the *chart transition map*. Informally, the chart transition map tells you how the charts of an atlas meld together to form the manifold. It's also useful in the following sense: We want to be able to do things, such as judging whether a path is continuous, by only looking at charts. If we look at a chart and find that it is continuous, can we guarantee that the path is as well? Because of the chart transition map, we can. Let's call the path γ and let f and g be two chart maps. Now suppose we find that $f \circ \gamma$ is continuous. Then it follows that $g \circ \gamma = g \circ f^{-1} \circ f \circ \gamma$ is continuous. Hence we know that we can define continuity of γ in this way, that this is a well-defined notion, because this notion does not depend on our choice of chart map.

2.2 Multilinear algebra

Let V and W be vector spaces. The set of linear maps from V to W is denoted by Hom(V,W).

Proposition 2.2.1

Hom(V,W) forms a vector space under vector addition and scalar multiplication operations of W.

We will now learn about a very special vector space. Given V along with this new vector space, we will be able to construct many other vector spaces. The *dual space* is given by $V^* \equiv \text{Hom}(V, \mathbb{R})$. An element of the dual space is called a *one-form* or *covector*. Let V be n-dimensional with basis $\{e^1, ..., e^n\}$. If a basis $\{\epsilon_1, ..., \epsilon_n\}$ of V^* satisfies

$$\epsilon^a(e_b) = \delta^a_b$$

then it is called the *dual basis* of V^*.

Theorem 2.2.1

If dim(V) $< \infty$, then $\left(V^*\right)^* = V$.

Example

Consider the 4D vector space of polynomials of degree 3 with real coefficients. A basis for this space is $\{1, x, x^2, x^3\}$. The dual space is the set of mixed differential operators of order 3 or smaller. It's not hard to see that the dual basis is

$$\left\{1, \frac{\partial}{\partial x}, \frac{1}{2}\frac{\partial^2}{\partial x^2}, \frac{1}{6}\frac{\partial^3}{\partial x^3}\right\}.$$

Perhaps you encountered tensors in physics and were told something useless like "a tensor is an object that transforms like a tensor." You will now get to see a more illuminating definition. Let $r, s \in \mathbb{N}$. A *tensor of rank (r,s)* is a map

$$T : \underbrace{V^* \times ... \times V^*}_{r \text{ times}} \times \underbrace{V \times ... \times V}_{s \text{ times}} \to \mathbb{R}$$

that is linear in all its arguments. In other words, a tensor is a multilinear map that eats vectors and covectors and spits out a real number. Now, you may be used to (1,1) tensors being objects that take vectors to vectors. Given a tensor T, we can actually recover such a construction.

Example

Given a tensor $T : V^* \times V \to \mathbb{R}$, where V is a vector space of finite dimension, define the map $\phi : V \to \left(V^*\right)^* = V$ by

$$\phi(v) = T(\cdot, v),$$

where $v \in V$ and the $\cdot$ represents an as-yet-unspecified covector. We see that the RHS of the above equation is an element of $(V^*)^*$ since it eats covectors and spits out real numbers. So the map ϕ takes a vector to a vector. Similarly if we define the map $\phi : V^* \to V^*$ by

$$\phi(\omega) = T(\omega, \cdot),$$

where $\omega \in V^*$ and $\cdot$ is an as-yet-unspecified vector, we see ϕ takes a covector to a covector.

Example

Here are some more familiar tensors:

1. Clearly rank (0,1) tensors are covectors, which is why they are also called one-forms. Similarly rank (1,0) tensors are vectors.
2. An inner product is a rank (0,2) tensor.
3. The gradient is a rank (0,1) tensor.

Again let V be finite dimensional. Given a basis of V and the corresponding dual basis of V^*, we can write $\forall\, v \in V$ and $\forall\, \omega \in V^*$

$$v = v^1 e_1 + ... + v^n e_n, \quad \omega = \omega_1 \epsilon^1 + ... + \omega_n \epsilon^n. \tag{2.1}$$

That is, we can write v in terms of its components v^i and ω in terms of its components ω_i. We can do the same thing for tensors. For a rank (r, s)

tensor, each component is given by

$$T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} = T(\omega^{i_1}, ..., \omega^{i_r}, e_{j_1}, ..., e_{j_s}), \tag{2.2}$$

where $1 \leq i, j \leq n$.

References

[1] Frederic Schuller. The W. E. Hereaus international winter school on gravity and light, 2014. URL https://gravity-and-light.herokuapp.com/tutorials. [Online; accessed 6-Nov-2019].

Chapter 3

Math: Probability and Statistics

This chapter is an introduction to the statistical tools needed to analyze data, especially as generated by Markov Chain Monte Carlo simulations. Parts of this presentation follow selected parts of Chapters 1 and 2 of Berg [1].

3.1 Preliminaries

The set of all possible outcomes Ω of an experiment is called the *sample space*. *Events* A are aggregates of points in the sample space. If Ω is a measurable set, then a *random variable* is a function $X : \Omega \to \mathbb{R}$. In this book we will try to always denote random variables with capital letters.

Next we need a way to define probabilities. In the case that $|\Omega| < \infty$, we might say that the probability of event A is $\mathrm{P}(A) = |A|/|\Omega|$. In this instance, we say that the probability is *uniform* since every point in the sample space is equally likely. However this need not always be the case; one outcome might be more likely than the other ones. In general for some integrable function $f : \mathbb{R} \to \mathbb{R}$, we assign a probability that X lies in the interval $[a, b]$ by

$$\mathrm{P}(X \in [a, b]) = \int_a^b \mathrm{d}x\, f(x). \tag{3.1}$$

Often statisticians say X is a *continuous* random variable, which is to be contrasted with *discrete* random variables. In this book we will only be concerned with continuous random variables, and we will simply call them

random variables. The function f is called the *probability distribution function* (PDF), and it must satisfy

$$1 = \int_{-\infty}^{\infty} \mathrm{d}x\, f(x). \tag{3.2}$$

Meanwhile the *cumulative distribution function* (CDF) is the function $F(x)$ given by

$$F(x) \equiv \mathrm{P}\left(X < x\right) = \int_{-\infty}^{x} \mathrm{d}t\, f(t). \tag{3.3}$$

Example

Two examples of important probability distributions include the *Gaussian* or *normal* distribution,

$$\mathrm{gau}(x, \hat{x}, \sigma) \equiv \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\hat{x})^2}{2\sigma^2}\right) \tag{3.4}$$

where σ is the standard deviation of the distribution and $\hat{x}$ is the mean, and the *Cauchy* distribution,

$$\mathrm{cau}(x, \alpha) \equiv \frac{\alpha}{\pi\left(\alpha^2 + x^2\right)}. \tag{3.5}$$

I will refer to these special PDFs later, particularly the normal distribution. I'll call their CDFs Gau and Cau, respectively.

Now that we know what probabilities and PDFs are, we can start thinking about ways to characterize them. For example we can think about typical values taken by a random variable from some distribution. We can get some information from the mean and variance of a distribution. These are both special cases of a more general concept. In particular, let $n \in \mathbb{N}$. The n^{th} *moment* of the distribution $f(x)$ is

$$\langle X^n \rangle = \int_{-\infty}^{\infty} \mathrm{d}x\, x^n f(x). \tag{3.6}$$

The mean and variance are the special cases $\hat{x} = \langle X \rangle$ and $\sigma^2 = \langle (X - \hat{x})^2 \rangle$. Sometimes we call the mean the *expected value* and sometimes we denote the

variance var. Note that not all probability distributions have well-defined moments. The Cauchy distribution is very ill-behaved in this regard, since its n^{th} moment diverges $\forall\, n \in \mathbb{N}$. Generally in the lab, one draws random variables from distributions about which one has no a priori knowledge. Therefore in principle one doesn't know the true moments these distributions. The definition (3.6) suggests a way to estimate them. Suppose you draw a sample $X_1, ..., X_N$: An *estimator* of the n^{th} moment is

$$\bar{X}^n \equiv \frac{1}{N} \sum_{i=1}^{N} X_i^n. \tag{3.7}$$

In the case $n = 1$ we obtain the ordinary arithmetic average. We use the hat to distinguish true values from estimators, which will generally be denoted with a bar. For estimators of moments besides the mean, we must be more careful; this is discussed in Section 3.4.

Consider two intervals $[a, b]$ and $[c, d]$ and two random variables X and Y drawn from PDFs f and g, respectively. Then X and Y are said to be *independent* if

$$\mathrm{P}\left(X \in [a, b] \text{ and } Y \in [c, d]\right) = \int_a^b \int_c^d \mathrm{d}x\, \mathrm{d}y\, f(x)\, g(y) \tag{3.8}$$

Hence we see that the *joint PDF* of X and Y is $f(x)g(y)$. On the other hand, we say X and Y are *uncorrelated* if

$$\langle XY \rangle = \langle X \rangle \langle Y \rangle, \tag{3.9}$$

The *covariance*

$$\mathrm{cov}[X, Y] \equiv \langle XY \rangle - \langle X \rangle \langle Y \rangle \tag{3.10}$$

can be used to give a measure of how correlated X and Y are, or one can use the *correlation*

$$\rho(X, Y) = \frac{\mathrm{cov}[X, Y]}{\sqrt{\sigma_X^2 \sigma_Y^2}}. \tag{3.11}$$

So equivalently we say X and Y are correlated if $\rho(X, Y) = 0$. It's worth emphasizing that if X and Y are independent, it follows that they are uncorrelated. This can be seen by applying definition (3.6) to the random variable XY, then using definition (3.8). However if X and Y are uncorrelated, *they can still be dependent.*

Example

Here's an extreme example by Cosma Shalizi [2]. Let X be uniformly distributed on [-1,1] and let $Y = |X|$. Then clearly Y depends on X. However it is easy to see that Y is uniform on [0,1] and $\langle XY \rangle = 0 = \langle X \rangle \langle Y \rangle$. Hence X and Y are not correlated.

The next two propositions show us how to add expectation values and random variables. Let X and Y be independent random variables drawn from PDFs f and g, respectively.

Proposition 3.1.1

Let $a, b \in \mathbb{R}$ be constants. Then

$$\langle aX + bY \rangle = a \langle X \rangle + b \langle Y \rangle .$$

Proof. Since X and Y are independent, their joint PDF is fg. Then

$$\begin{aligned} \langle aX + bY \rangle &= \int \mathrm{d}x \, \mathrm{d}y \, (ax + by) f(x) g(y) \\ &= a \int \mathrm{d}x \, \mathrm{d}y \, x \, f(x) g(y) + b \int \mathrm{d}x \, \mathrm{d}y \, y \, f(x) g(y) \\ &= a \int \mathrm{d}x \, x \, f(x) + b \int \mathrm{d}y \, y \, g(y) \\ &= a \langle X \rangle + b \langle Y \rangle . \end{aligned}$$

□

Proposition 3.1.2

The PDF of the random variable $Z = X + Y$ is given by the convolution

$$h(z) = \int_{-\infty}^{\infty} \mathrm{d}x \, f(x) g(z - x)$$

Proof. The CDF of Y is, according to eq. (3.8),

$$H(y) = \int_{x+y \leq z} \mathrm{d}x \, \mathrm{d}y \, f(x) g(y) = \int_{-\infty}^{\infty} \mathrm{d}x \, f(x) \int_{-\infty}^{z-x} \mathrm{d}y \, g(y).$$

The PDF h follows from the Fundamental Theorem of Calculus:

$$h(z) = \frac{\mathrm{d}H}{\mathrm{d}z} = \frac{\mathrm{d}H}{\mathrm{d}(z-x)} = \int_{-\infty}^{\infty} \mathrm{d}x\, f(x)g(z-x).$$

□

A sequence $\{X_N\}$ of random variables *converges in probability* toward the random variable X if $\forall \epsilon > 0$

$$\lim_{N\to\infty} \mathrm{P}\left(|X_N - X| > \epsilon\right) = 0, \tag{3.12}$$

and we write

$$X_N \xrightarrow{\mathrm{P}} X. \tag{3.13}$$

The sequence converges to X *almost surely* if

$$\lim_{N\to\infty} \mathrm{P}\left(X_N = X\right) = 1, \tag{3.14}$$

and in this case we write

$$X_N \xrightarrow{\mathrm{AS}} X. \tag{3.15}$$

Theorem 3.1.1: Chebyshev's inequality

Let X be drawn from a PDF with mean $\hat{x}$ and variance σ^2 and let $a > 0$. Then

$$\mathrm{P}\left(|X - \hat{x}| > a\sigma\right) < a^{-2}.$$

Proof. Let $T = (X - \hat{x})^2$ be a new random variable with PDF g. Then

$$\mathrm{P}\left(|X - \hat{x}| > a\sigma\right) = \mathrm{P}\left(T > a^2\sigma^2\right) = \int_{a^2\sigma^2}^{\infty} \mathrm{d}t\, g(t)$$

But

$$\begin{aligned}\sigma^2 &= \int_0^{\infty} \mathrm{d}t\, t\, g(t) = \left(\int_0^{a^2\sigma^2} + \int_{a^2\sigma^2}^{\infty}\right) \mathrm{d}t\, t\, g(t) \\ &\geq \int_{a^2\sigma^2}^{\infty} \mathrm{d}t\, t\, g(t) > a^2\sigma^2 \int_{a^2\sigma^2}^{\infty} \mathrm{d}t\, g(t) = a^2\sigma^2\, \mathrm{P}\left(T > a^2\sigma^2\right).\end{aligned}$$

Dividing through by $a^2\sigma^2$ completes the proof. □

Chebyshev's inequality tells you that large deviations from the mean are unlikely. Intuitively you expect that as the number of measurements increases, the sample average tends toward the true mean. This is called the *Law of Large Numbers* (LLN). To prove it, we set up as follows: Let $X_1, ..., X_N$ be a sequence of random variables drawn from a PDF with mean $\hat{x}$ and variance σ^2.

Theorem 3.1.2: Weak LLN

$$\bar{X} \xrightarrow{\mathrm{P}} \hat{x}.$$

Proof. Our proof will rely on Chebyshev's inequality, so we will first need to compute the mean and variance of the distribution of $\bar{X}$. All the X_i are drawn from the same PDF, so

$$\left\langle \bar{X} \right\rangle = \frac{1}{N} \sum_{i=1}^{N} \langle X_i \rangle = \frac{N\hat{x}}{N} = \hat{x}.$$

Meanwhile the variance of the distribution of $\bar{X}$ is

$$\sigma_{\bar{X}}^2 = \mathrm{var} \sum_{i=1}^{N} \frac{X_i}{N} = \sum_{i=1}^{N} \frac{\sigma^2}{N^2} = \frac{\sigma^2}{N}.$$

Now let $\epsilon > 0$. Then $\exists\, a > 0$ with $\epsilon = a\,\sigma_{\bar{X}}$. Hence by Chebyshev's inequality we have

$$\lim_{N\to\infty} \mathrm{P}\left(|\bar{X} - \hat{x}| > \epsilon\right) \leq \lim_{N\to\infty} \frac{\sigma_{\bar{X}}^2}{\epsilon^2} = \lim_{N\to\infty} \frac{\sigma^2}{N\epsilon^2} = 0.$$

The probability can't be less than 0, so we're done. □

The above proof relies on the PDF having a finite variance. As it turns out, the Weak LLN is true even when the variance is infinite! This can be proved using characteristic functions. But since we don't introduce characteristic functions until Section 3.3, and since we assume in practice that our data are drawn from PDFs with finite variance anyway, we direct the reader elsewhere. For example, a proof can be found on Wikipedia [3].

For completeness we also list the Strong LLN, but without proof. Like the Weak LLN, the Strong LLN is true even when the PDF variance is infinite.

Theorem 3.1.3: Strong LLN

$$\bar{X} \xrightarrow{\text{AS}} \hat{x}.$$

3.2 The normal distribution

Now we're going to focus on results about the normal distribution specifically. This first proposition will aid us in some of the calculations.

Proposition 3.2.1

Let $\alpha > 0$. Then

$$\int_{-\infty}^{\infty} \mathrm{d}x\, e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}}.$$

Proof. The trick is to just square the LHS:

$$\begin{aligned}\left(\int_{-\infty}^{\infty} \mathrm{d}x\, e^{-\alpha x^2}\right)^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{d}y\, e^{-\alpha(x^2+y^2)} \\ &= \int_0^{\infty} \mathrm{d}r\, r \int_0^{2\pi} \mathrm{d}\theta\, e^{-\alpha r^2} \\ &= \frac{\pi}{\alpha}.\end{aligned}$$

□

For the next result let X_1 and X_2 be two independent random variables drawn from normal distributions with respective means $\hat{x}_1$ and $\hat{x}_2$ and standard deviations σ_1 and σ_2.

Proposition 3.2.2

The random variable $Y = X_1 + X_2$ is normally distributed with mean $\hat{x}_1 + \hat{x}_2$ and variance $\sigma_1^2 + \sigma_2^2$.

Proof. By Proposition 3.1.2, the sum Y has the distribution

$$g(y) = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \mathrm{d}x \exp\left[-\frac{(x-\hat{x}_1)^2}{2\sigma_1^2} - \frac{(y-x-\hat{x}_2)^2}{2\sigma_2^2}\right].$$

Pull everything out of the integral that doesn't depend on x, then complete the square with what's left over. One obtains for $g(y)$

$$\frac{1}{2\pi\sigma_1\sigma_2}\exp\left[-\frac{(y-\hat{x}_1-\hat{x}_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right]\int_{-\infty}^{\infty}\mathrm{d}x\exp\left[-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}(x+C)^2\right],$$

where C is just a bunch of stuff that doesn't depend on x. Therefore you can make the substitution $u = x + C$ with $\mathrm{d}u = \mathrm{d}x$ and carry out the new integral using Proposition 3.2.1. The result is

$$g(y) = \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}\exp\left[-\frac{(y-\hat{x}_1-\hat{x}_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right].$$

□

Since the normal distribution is so important, so must be its CDF. Unfortunately the integral of the normal PDF is *non-elementary*; that is, it can't be expressed in terms of polynomials or standard functions like sin, cos, or exp. Therefore we give a name to this special function. The *error function* is

$$\mathrm{erf}(x) \equiv \frac{2}{\sqrt{\pi}}\int_0^x \mathrm{d}t\, e^{-t^2}. \tag{3.16}$$

Then we can write the Gaussian CDF with mean 0 as

$$\mathrm{Gau}(x,0,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}\mathrm{d}t\, e^{-t^2/2\sigma^2} = \frac{1}{2}+\frac{1}{2}\,\mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma}\right). \tag{3.17}$$

Now we can list some pretty powerful applications of the normal distribution. For instance one often must compare two empirical estimates of some mean. Usually these estimates are different, and one might wonder whether this disparity is real or just plain unlucky. More precisely:

Theorem 3.2.1: Gaussian difference test

Suppose $\bar{X}$ and $\bar{Y}$ are correct estimates of some expectation value, i.e. they are normally distributed with the same mean, and call their respective standard deviations σ_X and σ_Y. Then the probability that

$\bar{X}$ and $\bar{Y}$ differ by at least D is

$$\mathrm{P}\left(|\bar{X}-\bar{Y}|>D\right)=1-\mathrm{erf}\left(\frac{D}{\sqrt{2(\sigma_X^2+\sigma_Y^2)}}\right).$$

Proof. From Proposition 3.2.2, the random variable $\bar{X}-\bar{Y}$ is normally distributed with mean 0 and variance $\sigma_D^2=\sigma_X^2+\sigma_Y^2$. Therefore by eq. (3.17), the probability that $\bar{X}$ and $\bar{Y}$ are at most D apart is

$$\begin{aligned}\mathrm{P}\left(|\bar{X}-\bar{Y}|<D\right)&=\mathrm{P}\left(-D<\bar{X}-\bar{Y}<D\right)\\&=\mathrm{Gau}(D,0,\sigma_D)-\mathrm{Gau}(-D,0,\sigma_D)\\&=1-2\,\mathrm{Gau}(-D,0,\sigma_D)\\&=\mathrm{erf}\left(\frac{D}{\sqrt{2}\sigma_D}\right).\end{aligned}$$

And of course, $\mathrm{P}\left(|\bar{X}-\bar{Y}|>D\right)=1-\mathrm{P}\left(|\bar{X}-\bar{Y}|<D\right)$. □

In other words, the above theorem gives the probability that the observed difference $|\bar{X}-\bar{Y}|$ is due to chance. This probability is called the *q-value*. In practice one sets some threshold on q below which one investigates further whether the underlying distributions of the estimates are different. Often one takes the threshold as 0.05.

3.3 The central limit theorem

Let X and Y be real random variables. Then we can construct a complex random variable $F=X+iY$, and its expectation value will be

$$\langle F\rangle=\langle X\rangle+i\,\langle Y\rangle\,. \tag{3.18}$$

This allows us to speak sensibly about Fourier transformations of PDFs. In particular let X be drawn from the PDF f. The *characteristic function* of X is

$$\phi(t)\equiv\left\langle e^{itX}\right\rangle=\int_{-\infty}^{\infty}\mathrm{d}x\,e^{itx}f(x). \tag{3.19}$$

Knowing the characteristic function X is equivalent to knowing its PDF; this is because we can take the inverse Fourier transformation

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}t\, e^{-itx}\phi(t), \tag{3.20}$$

which follows from the Dirac δ-function. The derivatives of the characteristic function are easily calculated to be

$$\phi^{(n)}(t) = i^n \int_{-\infty}^{\infty} \mathrm{d}x\, x^n e^{itx} f(x); \tag{3.21}$$

therefore

$$\phi^{(n)}(0) = i^n \left\langle X^n \right\rangle. \tag{3.22}$$

If $|f(x)|$ falls off faster than x^m for any $m \in \mathbb{Z}$, then it follows from the above equation that all moments exist, and the characteristic function is analytic in t about $t = 0$.

These are some neat properties of characteristic functions; however our main use for them is summarized in the next proposition.

Proposition 3.3.1

The characteristic function of a sum of independent random variables equals the product of their characteristic functions.

Proof. Let X_1,...,X_N be drawn from PDFs f_1,...,f_N with corresponding characteristic functions $\phi_1, ..., \phi_N$, and let $Y = \sum_j X_j$. Then using the definition of the characteristic function we obtain

$$\phi_Y(t) = \left\langle e^{it\sum_j X_j} \right\rangle = \left\langle \prod_{j=1}^{N} e^{itX_j} \right\rangle = \prod_{j=1}^{N} \left\langle e^{itX_j} \right\rangle = \prod_{j=1}^{N} \phi_j(t),$$

where we used independence for the third equality. □

Now suppose you're an experimenter taking independent measurements of some observable. Furthermore suppose you don't know anything about the observable, except that it comes from some distribution with finite variance. The central limit theorem (CLT) says that armed with this information alone, you know that the sample mean will be normally distributed about the true mean. Here is the precise statement.

Theorem 3.3.1: Central limit theorem

Let $X_1, ..., X_N$ be N independent random variables drawn from PDF f. Suppose further that f has mean $\hat{x}$ and variance σ^2. Then the PDF of the estimator $\bar{X}$ converges to $\text{gau}(\bar{x}, \hat{x}, \sigma/\sqrt{N})$.

Proof. What we're going to do is look at the characteristic function ϕ_S of the random variable

$$S \equiv \bar{X} - \hat{x} = \frac{X_1 + ... + X_N - N\hat{x}}{N}.$$

If we can show that ϕ_S converges to the characteristic function corresponding to $\text{gau}(s, 0, \sigma/\sqrt{N})$, then we are finished. In order to show this, we first need the characteristic function for the distribution $\text{gau}(s, 0, \sigma/\sqrt{N})$. By completing the square and using Proposition 3.2.1, we find

$$\begin{aligned}
\phi_{\text{gau}} &= \frac{1}{\sigma}\sqrt{\frac{N}{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}s\, e^{its} \exp\left[-\frac{s^2 N}{2\sigma^2}\right] \\
&= \frac{1}{\sigma}\sqrt{\frac{N}{2\pi}} \exp\left[-\frac{\sigma^2 t^2}{2N}\right] \int_{-\infty}^{\infty} \mathrm{d}s \exp\left[-\frac{N}{2\sigma^2}(s - C)^2\right] \\
&= \exp\left[-\frac{\sigma^2 t^2}{2N}\right],
\end{aligned}$$

where C is a number that doesn't depend on s. It remains to show $\phi_S = \phi_{\text{gau}}$. By Proposition 3.3.1 we have

$$\phi_S(t) = \phi_{\frac{1}{N}\sum X_i - \hat{x}}(t) = \left[\phi_{X-\hat{x}}\left(\frac{t}{N}\right)\right]^N,$$

where $\phi_{X-\hat{x}}$ is the characteristic function corresponding to the random variable $X - \hat{x}$. Call its PDF g. From the properties of f, we know that g has mean 0 and variance σ^2. Therefore by expanding ϕ_S about $t = 0$ and using the definition (3.6), we find

$$\phi_S(t) = \left[1 - \frac{\sigma^2 t^2}{2N^2} + \mathcal{O}\left(\frac{t^3}{N^3}\right)\right]^N = \exp\left[-\frac{\sigma^2 t^2}{2N}\right] + \mathcal{O}\left(\frac{t^3}{N^2}\right),$$

as desired. □

Since the variance of the estimator $\bar{X}$ tends to 0 for large N, it follows that the sample mean converges to the true mean $\hat{x}$. In particular for large N, we expect the true mean to be within $\sigma/\sqrt{N}$ of the estimator roughly 68% of the time. Table 3.1 gives the area under a Gaussian curve for different numbers of standard deviations away from the mean.

3.4 Bias

For this section consider independent random variables $X_1, ..., X_N$ drawn from a distribution with mean $\hat{x}$ and variance σ^2. Earlier we recovered the familiar estimator for the mean, which was just the ordinary arithmetic average. But what about an estimator for the variance? Naively one might write

$$\bar{\sigma}^2_{\text{biased}} = \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})^2; \tag{3.23}$$

While we expect this estimator to converge to the exact result in the limit $N \to \infty$, it disagrees with σ^2 for small N. Most glaringly when $N = 1$, the estimator is zero, regardless of the exact result. An estimator is said to be *biased* when its expectation value does not agree with the exact result. The difference between the expectation value of the estimator and the exact result is correspondingly called the *bias*. When they agree, we say the estimator is *unbiased*.

Table 3.1: Table of areas under the curve for the normal distribution. The last column gives the probability that a random variable drawn from the distribution falls at least the given number of error bars away from the mean.

Number of σ from $\hat{x}$	Area under curve	About 1 in ...
1	0.682 689 49	3
2	0.954 499 74	22
3	0.997 300 20	370
4	0.999 936 66	15 787
5	0.999 999 43	1 744 278

Proposition 3.4.1

An unbiased estimator of the variance is

$$\bar{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2.$$

Proof. To construct an unbiased estimator of the variance, we'll determine the bias of the estimator, then remove it. Note

$$\left\langle \bar{\sigma}^2_{\text{biased}} \right\rangle = \frac{1}{N}\sum_{i=1}^{N}\left(\left\langle X_i^2 \right\rangle - 2\left\langle X_i \bar{X} \right\rangle + \left\langle \bar{X}^2 \right\rangle\right).$$

Let us analyze the above equation term by term. Since the random variables X_i are drawn from the same distribution, the first term is an unbiased estimator of $\langle X^2 \rangle$ for each i. Next the second term can be rewritten as

$$\begin{aligned}
\left\langle X_i \bar{X} \right\rangle &= \frac{1}{N}\left(\left\langle X_i^2 \right\rangle + \sum_{j\,|\,j\neq i}\left\langle X_i X_j \right\rangle\right) \\
&= \frac{1}{N}\left(\left\langle X^2 \right\rangle + (N-1)\left\langle X \right\rangle^2\right) \\
&= \frac{1}{N}\left(\left\langle X^2 \right\rangle - \left\langle X \right\rangle^2\right) + \left\langle X \right\rangle^2 \\
&= \frac{\sigma^2}{N} + \hat{x}^2,
\end{aligned}$$

where in the second line we used the independence of the X_i. Finally for the last term we have

$$\left\langle \bar{X}^2 \right\rangle = \left\langle \frac{1}{N^2}\sum_{i,j} X_i X_j \right\rangle = \frac{1}{N^2}\left(N\left\langle X^2 \right\rangle + \sum_{i,j\,|\,i\neq j}\hat{x}^2\right) = \frac{\sigma^2}{N} + \hat{x}^2,$$

where we again used independence in the second equality. Plugging everything into $\langle \bar{\sigma}^2_{\text{biased}} \rangle$ gives

$$\left\langle \bar{\sigma}^2_{\text{biased}} \right\rangle = \frac{1}{N}\sum_{i=1}^{N}\left(\left\langle X^2 \right\rangle - \frac{\sigma^2}{N} - \hat{x}^2\right) = \left(\frac{N-1}{N}\right)\sigma^2.$$

This equation shows us the bias is $-\sigma^2/N$. Therefore according to this equation, an unbiased estimator of the variance is

$$\bar{\sigma}^2 = \left(\frac{N}{N-1}\right)\bar{\sigma}^2_{\text{biased}} = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$$

as we wished to show. □

We saw that the bias of the naive variance estimator goes like $1/N$. So one might wonder: How much bias does one typically expect to encounter? Bias problems often appear whenever one wants to estimate some function of the mean $\hat{f} = f(\hat{x})$ that is not necessarily linear near the mean. One might be tempted to take the estimator

$$\bar{f}_{\text{bad}} = \frac{1}{N}\sum_{i=1}^{N} f_i, \tag{3.24}$$

where $f_i \equiv f(X_i)$. However it turns out that

$$\lim_{N\to\infty} \bar{f}_{\text{bad}} \neq \hat{f}. \tag{3.25}$$

Example 3.4.1: C

nsider N measurements of a random variable X that are drawn from a PDF with mean 0, and suppose we are interested in estimating the random variable X^2. If we try the bad estimator, we find

$$\bar{f}_{\text{bad}} \tag{3.26}$$

An estimator that never converges to its true value is called *inconsistent*; otherwise it is *consistent*. So this bad estimator is not a consistent estimator. Note that a biased estimator is not necessarily inconsistent; for instance the biased estimator of the variance eq. (3.23) is consistent. A consistent estimator of $\hat{f}$ is

$$\bar{f} = f(\bar{X}). \tag{3.27}$$

We can prove the consistency of $\bar{f}$ for a wide class of functions.

Proposition 3.4.2

Suppose $f : \mathbb{R} \to \mathbb{R}$ has a convergent Taylor series in a region about $\hat{x}$. If $\bar{X}$ maps to this region, then $\bar{f}$ has bias of order $1/N$.

Proof. If we consider f as a function of the ordinary variable x, we can expand it about $\hat{x}$ as

$$f(x) = f(\hat{x}) + f'(\hat{x})(x - \hat{x}) + \frac{1}{2}f''(\hat{x})(x - \hat{x})^2 + \mathcal{O}\big((x - \hat{x})^3\big).$$

Since $\bar{X}$ maps to the region in which this expansion is valid, we can plug it into the above formula and find its expected value. This gives

$$\big\langle \bar{f} \big\rangle - \hat{f} = f'(\hat{x}) \big\langle \bar{X} - \hat{x} \big\rangle + \frac{1}{2}f''(\hat{x}) \big\langle (\bar{X} - \hat{x})^2 \big\rangle + \mathcal{O}\big((\bar{X} - \hat{x})^3\big).$$

The LHS of this equation is the bias of $\bar{f}$. To simplify the RHS, note that by the CLT $\big\langle \bar{X} - \hat{x} \big\rangle = 0$ and $\big\langle (\bar{X} - \hat{x})^2 \big\rangle = \sigma^2/N$. Therefore

$$\big\langle \bar{f} \big\rangle - \hat{f} = \frac{1}{2}f''(\hat{x})\frac{\sigma^2}{N} + \mathcal{O}\left(\frac{1}{N^2}\right).$$

□

According to the above proposition, the bias vanishes as $N \to \infty$, which shows that $\bar{f}$ is consistent. For large N, $\bar{X}$ is very likely to be close to $\hat{x}$ by the CLT, so Proposition 3.4.2 will essentially hold whenever N is large and f is a nice enough function. There is another important consequence to this proposition: the bias decreases faster than the statistical error bar, which you will recall goes like $1/\sqrt{N}$. Hence when N becomes large enough, the bias can be ignored.

3.5 Error propagation and covariance

We will now reproduce the commonly used error propagation formula. We know that if we have a smooth function of N variables $f : \mathbb{R}^N \to \mathbb{R}$ we can

Taylor expand

$$f(x) = f(\hat{x}) + \sum_{j=1}^{N} \left.\frac{\partial f}{\partial x_j}\right|_{x=\hat{x}} (x_j - \hat{x}_j) + \mathcal{O}\big(x^2\big) \tag{3.28}$$

where $x = (x_1, ..., x_N)$ and $\hat{x} = (\hat{x}_1, ..., \hat{x}_N)$. So if $|x - \hat{x}|$ is small enough, f is a linear function of the components of x. This motivates the following situation: Suppose we have a set of M random variables Y_i, each of which is a linear function of N random variables X_j; i.e.

$$Y_i = a_{i0} + \sum_{j=1}^{N} a_{ij} X_j. \tag{3.29}$$

Then the mean is given by

$$\hat{y}_i = \langle Y_i \rangle = a_{i\,0} + \sum_{j=1}^{N} a_{ij} \langle x_j \rangle = a_{i\,0} + \sum_{j=1}^{N} a_{ij} \hat{x}_j. \tag{3.30}$$

Meanwhile the variance of Y_i is given by

$$\begin{aligned}
\sigma^2_{Y_i} = \big\langle (Y_i - \hat{y}_i)^2 \big\rangle &= \left\langle \sum_{j=1}^{N} a_{ij}(X_j - \hat{x}_j) \sum_{k=1}^{N} a_{ik}(X_k - \hat{x}_k) \right\rangle \\
&= \sum_{j=1}^{N} a_{ij}^2 \big\langle (X_j - \hat{x}_j)^2 \big\rangle + \sum_{j \neq k} a_{ij} a_{ik} \langle (X_j - \hat{x}_j)(X_k - \hat{x}_k) \rangle \\
&= \sum_{j=1}^{N} a_{ij}^2 \sigma^2_{X_j} + \sum_{j \neq k} a_{ij} a_{ik} \operatorname{cov}(X_j, X_k).
\end{aligned} \tag{3.31}$$

In the case that the X_j and X_k are independent, $\operatorname{cov}(X_j, X_k) = 0$. Furthermore if Y_i is a linear function of the X_j, we can associate the a_{ij} with partial derivatives. Then eq. (3.31) becomes

$$\sigma^2_{Y_i} = \sum_{j=1}^{N} \frac{\partial\, Y_i}{\partial X_j} \sigma^2_{X_j}. \tag{3.32}$$

This is the *error propagation formula* as it's usually stated. Berg [1] emphasizes that this relation is mnemonic because it doesn't make sense to take derivatives with respect to random variables. In practice we can apply eq. (3.32) when we

1. know how some function f depends on some variables x_i (then taking derivatives with respect to these new variables is well-defined);
2. take measurements of the variables;
3. the measurements are independent; and
4. either the function is exactly linear in the x_i or f is approximately linear in a region close to the mean.

Oftentimes in physics you'll find yourself in a situation where you want to calculate several functions of the same random variables X_j. If the X_j are close enough to their means, or if the Y_i are linear functions of the X_j, this is a situation in which (3.29) applies. Intuitively one might expect the Y_i to be correlated; this turns out to be the case. In particular

$$\begin{aligned}\operatorname{cov}(Y_i, Y_j) &= \langle (Y_i - \hat{y}_i)(Y_j - \hat{y}_j) \rangle \\ &= \sum_{k,l=1}^{N} a_{ik} a_{jl} \left\langle (X_k - \hat{x}_k)(X_l - \hat{X}_l) \right\rangle \\ &= \sum_{k=1}^{N} a_{ik} a_{jk} \sigma_{X_k}^2 + \sum_{k \neq l} a_{ik} a_{jl} \operatorname{cov}(X_k, X_l),\end{aligned} \tag{3.33}$$

which shows the Y_i are correlated even if the X_j are not.

3.6 Jackknife resampling

Let us consider a sample of independent measurements $X_1, ..., X_N$ from some distribution with mean $\hat{x}$ and variance σ^2 and a function f that has a Taylor series expansion near $\hat{x}$, but isn't necessarily linear. From Section 3.4 we know that $\bar{f} = f(\bar{X})$ is a consistent estimator of $\hat{f} = f(\hat{x})$.

The discussion of Section 3.5 gives us a way to propagate uncertainty from the random variables to f, but it is effectively unusable whenever f becomes sufficiently complicated. Even when f is simple, if the original data are correlated, the error propagation formula eq. (3.32) is still unwieldy. These are some motivations for using the jackknife. The jackknife method is pretty simple to implement, and jackknife error bars agree with usual error bars when there is no bias. Therefore it makes sense to use the jackknife method generally.

Here's how the jackknife method works: We throw away the first measurement from our sample, leaving a data set of $N-1$ resampled values. Statistical analysis is done on this smaller sample. Then we resample again, this time throwing out the second point, and so on. The *jackknife bins* are defined by

$$X_{J,i} \equiv \frac{1}{N-1}\sum_{j\neq i} X_j. \tag{3.34}$$

They allow us to construct a *jackknife estimator* for the mean $\bar{f}_J$ by

$$\bar{f}_J \equiv \frac{1}{N}\sum_{i=1}^{N} f_{J,i}, \tag{3.35}$$

where $f_{J,i} \equiv f(X_{J,i})$. The jackknife estimator for the variance of $\bar{f}_J$ is

$$\bar{\sigma}^2_{f_J} = \frac{N-1}{N}\sum_{i=1}^{N}(f_{J,i} - \bar{f}_J)^2. \tag{3.36}$$

Example

Consider the common problem of calculating the mean of the data and the variance of the mean. Using the unbiased estimator for the variance along with the CLT yields

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \quad \text{and} \quad \bar{\sigma}^2_{\bar{X}} = \frac{1}{N(N-1)}\sum_{i=1}^{N}(X_i - \bar{X})^2. \tag{3.37}$$

Meanwhile the jackknife estimator for the variance of $\bar{X}$ gives

$$\bar{\sigma}^2_{\bar{X}_J} = \frac{N-1}{N}\sum_{i=1}^{N}(X_{J,i} - \bar{X}_J)^2. \tag{3.38}$$

Some simple algebra shows that $(N-1)(X_{J,i}-\bar{X}_J) = \bar{X}-X_i$. Therefore

$$\bar{\sigma}^2_{\bar{X}_J} = \bar{\sigma}^2_{\bar{X}}. \tag{3.39}$$

Next let's see how the jackknife lets us estimate bias. From Proposition

3.4.2 we know the bias of the estimator $\bar{f}$ is of order $1/N$, which we will write

$$\text{bias } \bar{f} = \frac{A}{N} + \mathcal{O}\left(\frac{1}{N^2}\right) \tag{3.40}$$

for some constant A. Let us determine the bias of $\bar{f}_J$.

Proposition 3.6.1

If the measurements X_i are distributed relatively close to $\hat{x}$, then $\bar{f}_J$ has a bias of order $1/(N-1)$.

Proof. The assumption on the measurements is that they roughly fall within the series' radius of convergence. We rewrite

$$X_{J,i} = \hat{x} + \frac{1}{N-1}\sum_{j\neq i}(X_j - \hat{x}).$$

Then our strategy is the same as before: We expand f in the same sense as before, and take the average value of $f_{J,i}$. We obtain

$$\begin{aligned}
\langle f_{J,i}\rangle &= \langle f(X_{J,i})\rangle \\
&= \left\langle f\left(\hat{x} + \frac{1}{N-1}\sum_{j\neq i}(X_j - \hat{x})\right)\right\rangle \\
&= \hat{f} + \frac{1}{2}f''(\hat{x})\frac{1}{(N-1)^2}\sum_{\substack{j\neq i\\ k\neq i}}\langle (X_j-\hat{x})(X_k-\hat{x})\rangle + \mathcal{O}\left(\frac{1}{N^2}\right) \\
&= \hat{f} + \frac{1}{2}f''(\hat{x})\frac{1}{(N-1)^2}\left(\sum_{j\neq i}\sigma^2 + \sum_{j\neq k}\operatorname{cov}(X_j, X_k)\right) + \mathcal{O}\left(\frac{1}{N^2}\right) \\
&= \hat{f} + \frac{1}{2}f''(\hat{x})\frac{1}{N-1}\sigma^2 + \mathcal{O}\left(\frac{1}{N^2}\right),
\end{aligned}$$

where in third equality we used $\langle X_j - \hat{x}\rangle = 0$ and in the last equality we used the independence of the measurements. Since the RHS is independent of i, it follows that

$$\langle \bar{f}_J\rangle - \hat{f} = \frac{1}{2}f''(\hat{x})\frac{\sigma^2}{N-1} + \mathcal{O}\left(\frac{1}{N^2}\right).$$

□

Comparing the final steps of Propositions 3.4.2 and 3.6.1, we see that they have the same lowest order contribution, except that N is replaced by $N-1$. Therefore we can write

$$\text{bias } \bar{f}_J = \frac{A}{N-1} + \mathcal{O}\left(\frac{1}{N^2}\right) \tag{3.41}$$

with the same constant A as with eq. (3.40). Combining both of these equations, we conclude

$$A = N(N-1)\left(\langle \bar{f} \rangle - \langle \bar{f}_J \rangle\right) + \mathcal{O}\left(\frac{1}{N}\right), \tag{3.42}$$

which means that

$$\overline{\text{bias}} = (N-1)(\bar{f} - \bar{f}_J) \tag{3.43}$$

gives an estimator for the bias of $\bar{f}$, at least up to $\mathcal{O}(1/N^2)$.

Equation (3.40) shows that the bias decreases faster than the error bar. However it can happen that if A is large, and if N is relatively small, the bias is non-negligible. In practice one should be concerned if the bias is of the same order as the error bar. In such a case one should attempt to correct for this bias. Using eq. (3.43) it follows that

$$\bar{f}_C \equiv \bar{f} - \overline{\text{bias}} \tag{3.44}$$

is bias-corrected, in that it has $\mathcal{O}(1/N^2)$ disagreement with the true mean. To build an estimator for the variance of the bias corrected mean, we must do another level of jackknifing. Let $i \neq j$. The *second-level jackknife bins* are

$$X_{J,ij} \equiv \frac{1}{N-2} \sum_{k \neq i,j} X_k. \tag{3.45}$$

They allow us to construct *second-level jackknife estimators* by

$$f_{J,ij} \equiv f(X_{J,ij}). \tag{3.46}$$

We can use these second-level estimators to create jackknife samples of bias estimators and bias-corrected estimators with

$$\text{bias}_{J,i} = \frac{1}{N-1} \sum_{k \neq i} (f_{J,i} - f_{J,ik}) \quad \text{and} \quad f_{CJ,i} = f_{J,i} - \text{bias}_{J,i}. \tag{3.47}$$

With these definitions, we can then calculate estimators for the mean and variance of the mean using the same machinery as the original jackknife, i.e. from eq. (3.35) and (3.36). We get

$$\bar{f}_{CJ} = \frac{1}{N}\sum_{i=1}^{N} f_{CJ,i} \quad \text{and} \quad \bar{\sigma}^2_{f_{CJ}} = \frac{N-1}{N}\sum_{i=1}^{N}(f_{CJ,i} - \bar{f}_{CJ})^2, \tag{3.48}$$

which takes correlations between $f_{J,i}$ and $\text{bias}_{J,i}$ automatically into account. This completes the tool set needed to estimate average values of functions of data, correcting for bias and correlation.

3.7 The χ^2 distribution and fitting data

Consider a sample of N Gaussian, independent data points (X_i, Y_i), where the Y_i have standard deviations σ_i. For now we will assume the X_i have no error. We will consider a situation where we believe the Y_i are measurements of some real function y of x. Abstractly we model these data with a fit that depends on some set of M parameters

$$y = y(x; a), \tag{3.49}$$

where $a = (a_1, ..., a_M)$ is the vector of these parameters. Our goal is to estimate the a_j and their error bars, and then determine whether this fit is consistent with the data.

Assuming that $y(x, a)$ is the exact law for the data, the joint PDF of the measurements Y_i is given by eq. (3.8) to be

$$f(y_1, ..., y_N) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[\frac{-(y_i - y(x_i; a))^2}{2\sigma_i^2}\right]. \tag{3.50}$$

The PDF given by eq. (3.50) is an example of the *non-central* χ^2 *distribution.* Generally this distribution has random variable

$$X^2 = \sum_{i=1}^{N} \frac{(Y_i - \hat{y}_i)^2}{\sigma_i^2}, \tag{3.51}$$

where the random variables Y_i are drawn from $\text{gau}(y, \hat{y}_i, \sigma_i)$. In the special case that the Y_i are drawn from $\text{gau}(y, 0, 1)$ we obtain the random variable

$$X^2 = \sum_{i=1}^{N} Y_i^2. \tag{3.52}$$

In this case the PDF of X^2 is called the χ^2 *distribution*. It simplifies to

$$f(y_1, ..., y_N) = \frac{1}{(2\pi)^{N/2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{N} y_i^2\right]. \tag{3.53}$$

Let's think about a general, non-central χ^2 PDF. The likelihood that the data fall within a region near what was observed is

$$\mathrm{P} = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[\frac{-(y_i - y(x_i; a))^2}{2\sigma_i^2}\right] \mathrm{d}y_i \,. \tag{3.54}$$

Our strategy for determining the correct fit will be to find the vector a that maximizes the above probability. The happens when the argument of the exponential is closest to zero; i.e. when

$$\chi^2 \equiv \sum_{i=1}^{N} \frac{(y_i - y(x_i; a))^2}{2\sigma_i^2} \tag{3.55}$$

is minimized. This is an example of a *maximum likelihood method*. Once the parameters are found, one can then ask: What is the probability that the discrepancy between the data and the fit is due to chance?

To answer this question, let us begin with the simpler case using the χ^2 CDF (3.53). It is given by

$$F(\chi^2) = \mathrm{P}\left(X^2 \leq \chi^2\right) = \frac{1}{(2\pi)^{N/2}} \int_{\sum y_i^2 \leq \chi^2} \prod \mathrm{d}y_i \, e^{-y_i^2/2}. \tag{3.56}$$

Switching to hyperspherical coordinates, this becomes

$$F(\chi^2) = \frac{1}{(2\pi)^{N/2}} \int \mathrm{d}\Omega \int_0^{\chi} \mathrm{d}r \, r^{N-1} e^{-r^2/2}. \tag{3.57}$$

The RHS looks similar to the gamma function. With this in mind, we can make the substitution $t = r^2/2$ and use Proposition A.1.1 to obtain

$$F(\chi^2) = \frac{1}{\Gamma(N/2)} \int_0^{\chi^2/2} \mathrm{d}r \, t^{N/2-1} e^{-t}. \tag{3.58}$$

The integral

$$\Gamma(s, z) \equiv \frac{1}{\Gamma(s)} \int_0^{z} \mathrm{d}t \, t^{s-1} e^{-t} \tag{3.59}$$

with $\mathrm{Re}\, s > 0$ is called the *incomplete gamma function*. The CDF in the form (3.58) is well-suited for numerical calculation because it is straightforward to compute the incomplete gamma function.

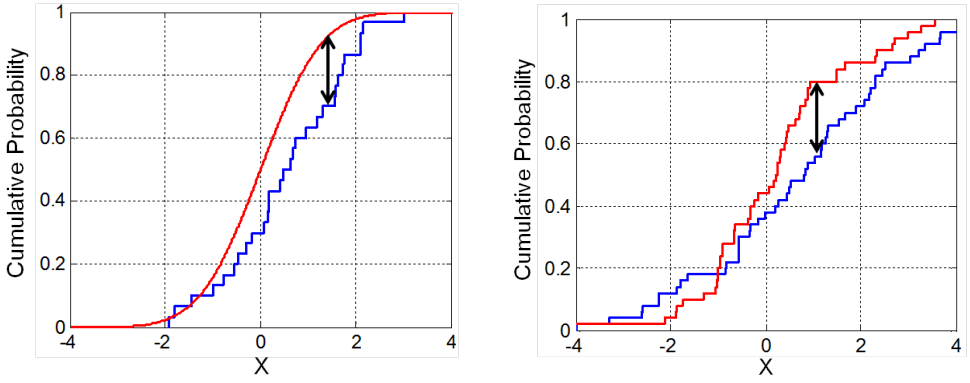

Figure 3.1: An example Kolmogorov statistic comparing an empirical CDF with a known, exact CDF (left) and a Kolmogorov statistic comparing two empirical CDFs (right). The statistic is indicated by the black, double-sided arrow. Images taken from Wikipedia [4].

3.8 The Kolmogorov test

Let's say we perform an experiment and extract a CDF $\bar{F}$ from the data. How can we tell whether the data are consistent with some true CDF F? Or if we extract another CDF $\bar{G}$, how can we tell whether $\bar{F}$ and $\bar{G}$ are consistent with each other? These questions can be answered using the *Kolmogorov test*. The statistic we will use to determine consistency is the largest difference between the CDFs, shown in Figure 3.1.

We start with some preliminary definitions. The *indicator function* of a subset A of a set B is the function $\mathbf{1}_A : B \to \{0, 1\}$ given by

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{otherwise.} \end{cases} \tag{3.60}$$

Given some measurements $X_1, ..., X_N$, we construct the *empirical CDF* as

$$\bar{F}(x) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{[X_i, \infty]}(x), \tag{3.61}$$

where each term counts the number of data less than or equal to x. The

measurements have the same CDF, so at each x we have by the LLN

$$\begin{aligned}\bar{F}(x) \xrightarrow{\mathrm{P}} \left\langle \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{[X_i,\infty]}(x) \right\rangle &= \frac{1}{N}\sum_{i=1}^{N} \left\langle \mathbf{1}_{[X_i,\infty]}(x) \right\rangle \\ &= \frac{1}{N}\sum_{i=1}^{N} \mathrm{P}\left(X_i \leq x\right) \\ &= \frac{1}{N}\sum_{i=1}^{N} F(x) \\ &= F(x),\end{aligned} \tag{3.62}$$

i.e. $\bar{F}$ is an unbiased, consistent estimator.

The *Kolmogorov statistic* is

$$\Delta \equiv \max_{x\in\mathbb{R}} \left|\bar{F}(x) - F(x)\right|. \tag{3.63}$$

Since $\bar{F}(x) \xrightarrow{\mathrm{P}} F(x)$ for all x, it follows that $\Delta \xrightarrow{\mathrm{P}} 0$. An important but surprising fact about Δ is that it is *distribution free*. The proof is relatively straightforward and given in Theorem 3.8.1. Theorem 3.8.2 gives the probability that the difference between the empirical and true CDFs is at least as extreme as what we calculate. This proof is tedious and not particularly enlightening, so it has been omitted. If you want to see a proof you can read Berg [1], who attributes it to Birnbaum and Tingey [5] and Smirnov [6].

Theorem 3.8.1

All continuous F have the same Δ.

Proof. We will start with the slightly easier case where F is monotonically increasing. In this case, F^{-1} exists and is also monotonically increasing. Then by making the variable change $y = F(x)$ we find

$$\begin{aligned}\Delta &= \max_{x\in\mathbb{R}} \left|\bar{F}(x) - F(x)\right| \\ &= \max_{y\in[0,1]} \left|\bar{F}\left(F^{-1}(y)\right) - F\left(F^{-1}(y)\right)\right| \\ &= \max_{y\in[0,1]} \left|\bar{F}\left(F^{-1}(y)\right) - y\right|.\end{aligned}$$

Let's focus on the $\bar{F}(F^{-1}(y))$ term. We can recast this as

$$\bar{F}\left(F^{-1}(y)\right) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{[X_i,\infty]}\left(F^{-1}(y)\right) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{[F(X_i),\infty]}(y),$$

where in the second step we used $F(X) \leq y \Leftrightarrow X \leq F^{-1}(y)$. This shows that the empirical CDF of $F^{-1}(y)$ is none other than the empirical CDF of the sample $F(X_1), ..., F(X_N)$. But

$$\mathrm{P}\left(F(X) \leq t\right) = \mathrm{P}\left(X \leq F^{-1}(t)\right) = F\left(F^{-1}(t)\right) = t,$$

i.e. the sample is drawn from the uniform distribution in the interval $[0, 1]$, regardless of what F is. It follows that $\bar{F}\left(F(y)\right)$ is independent of F, and hence so is Δ.
In the case F is not monotonic, the inverse is not guaranteed. We define

$$F^{+}(y) \equiv \min\{x \,|\, F(x) \geq y\}.$$

The important thing is that $F^{+}(y) \leq z \Leftrightarrow y \leq F(z)$. To see this note that

$$F^{+}(y) \leq z \Rightarrow y \leq F\left(F^{+}(y)\right) \leq F(z)$$

and

$$y \leq F(z) \Rightarrow F^{+}(y) \leq F^{+}\left(F(z)\right) \leq z.$$

The theorem follows by replacing F^{-1} with F^{+} in the first paragraph. □

Theorem 3.8.2

Let $D > 0$. Then

$$\mathrm{P}\left(\Delta > D\right) = \sum_{k=0}^{K} \binom{N}{k} D \left(D + \frac{k}{N}\right)^{k-1} \left(1 - D - \frac{k}{N}\right)^{N-k},$$

where K is determined by the condition that $1 - D - k/N$ cannot be negative. If N is large enough, this probability can be approximated as

$$\mathrm{P}\left(\Delta > D\right) \approx e^{-2ND^2}.$$

We now have all the ingredients we need to carry out a Kolmogorov test, and by Theorem 3.8.1 we are guaranteed it will work, regardless of the underlying distribution, under the modest assumption that its CDF is continuous. In practice one can proceed as follows:

1. Sort the measurements, then place them into an array $\{X_i\}$.
2. The X_i are indexed by i, so the corresponding empirical CDF is just the array of fractions $\{i/n\}$. For example $X_1 \leq X_1$, so $\bar{F}(X_1) = 1/N$. Since X_1, $X_2 \leq X_2$ we have $\bar{F}(X_2) = 2/N$.
3. Compute the exact PDF. For instance if we think the data come from $\mathrm{Gau}(x, 0, \sigma)$, we can compute the exact CDF using eq. (3.17); or if we think the data come from the uniform distribution on $[0, 1]$, we can just heapsort them.
4. Determine Δ.
5. Calculate $\mathrm{P}(\Delta > D)$ using Theorem 3.8.2. If the calculation of the exact probability is slow, you can use the approximation, but only if N is big enough, say larger than 100 or so.
6. The underlying assumption is that empirical CDF is an estimator for true CDF, so if the probability is below some threshold, say 0.05, then this assumption becomes suspect.

3.9 Statistical analysis of Markov chains

Suppose we have computed using MCMC a time series of N measurements $\{X_1, ..., X_N\}$. In principle each element of this sample is drawn from a PDF with mean $\langle X_i \rangle = \langle X \rangle = \hat{x}$ and variance $\sigma^2 = \langle (X_i - \hat{x})^2 \rangle$, i.e. they all have the same mean and variance. Unbiased estimators for the mean and variance are

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i \quad \text{and} \quad \bar{\sigma}^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2. \tag{3.64}$$

The variance of the random variable $\bar{X}$ is

$$\sigma_{\bar{X}}^2 = \left\langle (\bar{X} - \hat{x})^2 \right\rangle = \frac{1}{N^2} \left(\sum_{i \neq j} \langle X_i X_j \rangle + N \left\langle X^2 \right\rangle \right) - \hat{x}^2. \tag{3.65}$$

In the case that the measurements are uncorrelated, the expected values factorize, and we obtain

$$\sigma_{\bar{X}}^2 = \sigma^2/N \tag{3.66}$$

in agreement with the CLT. But in practice measurement $i+1$ is often correlated with measurement $i+t$ because they are from the same time series. To measure this we draw inspiration from definition (3.11). The *autocovariance* between measurements X_i and X_{i+t} is

$$c(X_i, X_{i+t}) \equiv \langle (X_i - \hat{x})(X_{i+t} - \hat{x}) \rangle = \langle X_i X_{i+t} \rangle - \langle X_i \rangle \langle X_{i+t} \rangle , \tag{3.67}$$

For a Markov process in equilibrium, the autocovariance depends only on the separation t, so we define $c(t) \equiv c(X_i, X_{i+t})$. Finally note that $c(0) = \sigma^2$, which motivates the definition of the *autocorrelation*

$$\gamma(t) \equiv \frac{c(t)}{\sigma^2}. \tag{3.68}$$

The autocorrelation decays in t as a sum of exponentials. I don't know why this is true, and I couldn't find a reference, but this is what everybody says. Assuming this is the case we can write

$$\gamma(t) = A_{\exp} e^{-t/\tau_{\exp}} + \sum_{i=1}^{\infty} A_i e^{-t/\tau_i}, \tag{3.69}$$

where the As are constants and we have picked out the leading exponential behavior; i.e. for all i

$$\tau_{\exp} > \tau_i. \tag{3.70}$$

$\tau_{\exp}$ is called the *exponential autocorrelation time*.

Plugging definition (3.67) into eq. (3.65) we have

$$\sigma_{\bar{X}}^2 = \frac{1}{N^2} \sum_{i,j} c(X_i, X_j). \tag{3.71}$$

In the last sum, $|i-j| = 0$ occurs N times, and $|i-j| = t$ occurs $2(N-t)$ times. Note $1 \leq t \leq N-1$. Therefore

$$\sigma_{\bar{X}}^2 = \frac{1}{N^2} \left(N\, c(0) + 2 \sum_{t=1}^{N-1} (N-t) c(t) \right). \tag{3.72}$$

Finally we use $c(0) = \sigma^2$ to find

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{N}\left(1 + 2\sum_{t=1}^{N-1}\left(1 - \frac{t}{N}\right)\gamma(t)\right) \equiv \frac{\sigma^2}{N}\tau_{\text{int}}. \tag{3.73}$$

The quantity

$$\tau_{\text{int}} = \left(1 + 2\sum_{t=1}^{N-1}\left(1 - \frac{t}{N}\right)\gamma(t)\right) \tag{3.74}$$

is called the *integrated autocorrelation time*.

From eq. (3.73) we see that τ_{int} is just the ratio between the estimated variance of the sample mean and what this variance would have been if the data were uncorrelated.

In practice, we often don't know the true mean $\hat{x}$ of the time series. Therefore along the lines of eq. (3.64), we construct an unbiased estimator of the autocovariance

$$\bar{c}(t) = \frac{N}{(N-1)(N-t)}\sum_{i=1}^{N-t}(X_i - \bar{X})(X_{i+t} - \bar{X}), \tag{3.75}$$

where it is the factor $N/(N-1)$ that removes the bias, just as before. Also in most situations we work in the limit where N is large. In this limit, we can construct an estimator for τ_{int} by

$$\bar{\tau}_{\text{int}}(n) = 1 + 2\sum_{t=1}^{n}\bar{\gamma}(t), \tag{3.76}$$

where $n < N$. To understand the above estimator look at definition (3.74). When t is small, $1 - t/N \approx 1$. Large t terms are doubly suppressed by the exponential decay of $\gamma(t)$ and by $1 - t/N \approx 0$. If the estimator still makes you uncomfortable, note that in the overly simplistic case where $\gamma(t)$ has only one exponential term, one can prove

$$\lim_{N\to\infty}\tau_{\text{int}} = 1 + 2\sum_{t=1}^{\infty}\gamma(t), \tag{3.77}$$

which parallels eq. (3.76) more closely. To construct a final estimator for τ_{int}, one looks for a window in n for which eq. (3.76) becomes roughly independent of n. This serves as the final $\bar{\tau}_{\text{int}}$.

References

[1] B. A. Berg. *Markov Chain Monte Carlo Simulations and Their Statistical Analysis*. World Scientific, Singapore, 2004.

[2] C. Shalizi. Reminder no. 1: Uncorrelated vs. independent, 2013. URL http://www.stat.cmu.edu/~cshalizi/uADA/13/reminders/uncorrelated-vs-independent.pdf. [Online; accessed 25-May-2017].

[3] Wikipedia contributors. Law of large numbers — Wikipedia, the free encyclopedia, 2017. URL https://en.wikipedia.org/w/index.php?title=Law_of_large_numbers&oldid=817455983. [Online; accessed 17-February-2018].

[4] Wikipedia contributors. Kolmogorov Smirnov test — Wikipedia, the free encyclopedia, 2017. URL https://en.wikipedia.org/w/index.php?title=Kolmogorov%E2%80%93Smirnov_test&oldid=816607137. [Online; accessed 17-February-2018].

[5] Z. W. Birnbaum and F. H. Tingey. One-sided confidence contours for probability distribution functions. *Ann. Math. Statist.*, 22(4):592–596, 1951.

[6] N. Smirnoff. Sur les écarts de la courbe de distribution empirique. *Mat. Sbornik*, 6:3–26, 1939.

Chapter 4

Physics: Statistical Physics

4.1 Equations of state

This section follows parts of chapter 8 of Ref. [1]. More details can be found in this book. You can also look at the references therein.

Probably the equation of state that you are most familiar with is the ideal gas law,

$$PV = NkT. \tag{4.1}$$

The volume V and the number of particles N both scale linearly with the size of the system; we call such variables *extensive*. Meanwhile the pressure P and the temperature T are *intensive* , i.e. they do not scale with the system.

There is an equation of state for each intensive variable required for the description of thermodynamic states. For example since

$$\mathrm{d}U = T\,\mathrm{d}S - P\,\mathrm{d}V + \mu\,\mathrm{d}N\,, \tag{4.2}$$

we know that

$$T = \frac{\partial U}{\partial S}. \tag{4.3}$$

This intensive variable T depends on only the extensive variables; generally we could write

$$T = T(U, S, V, N). \tag{4.4}$$

This is what we call an *equation of state*. Knowing every equation of state is enough to reconstruct the fundamental equation, and therefore enough to determine the physics of the system.

4.2 Legendre transforms

Equation (4.2) tells us that we can think of the internal energy U of a system in equilibrium at (T, P, μ) as a function of S, V, and N. However one of these extensive variables, such as S, may be difficult or impossible to measure, and therefore we would rather think in terms of the more accessible quantity T, which is the derivative of U with respect to S. Hence we want

1. to look at U in terms of a derivative with respect to S rather than S itself; moreover

2. we don't want to lose any information we had before, i.e. we want this process to be invertible.

These are the purposes of a Legendre transformation. The second point may seem too obvious to state, but it's worth emphasizing here because what makes thermodynamic potentials such as U special is that you're supposed to be able to determine state variables like T from them. Since no information is lost, Legendre transforms guarantee that thermodynamic potentials get transformed to other thermodynamic potentials.

Before we define a Legendre transformation, let's look at an example due to Markus Deserno [2] where a naive transformation can go wrong and information can be lost.

Example

We consider a function $y(x)$ and define a new variable

$$p \equiv y'(x). \tag{4.5}$$

In order to accomplish goal (1) above, one might naively solve eq. (4.5) for x, obtaining the function $x(p)$ and then plug this back into $y(x)$ to obtain

$$Y(p) = y\left(y'^{-1}(p)\right). \tag{4.6}$$

To see that this procedure destroys information, consider the example

$$y(x) = \frac{1}{2}(x - x_0)^2. \tag{4.7}$$

The derivative is $p = y'(x) = x - x_0$, and hence $x = p + x_0$. Plugging

this into eq. (4.7) we get the function

$$Y(p) = y(x(p)) = \frac{1}{2}\left(x(p) - x_0\right)^2 = \frac{1}{2}p^2. \tag{4.8}$$

Evidently all functions of the form (4.7) get transformed to the same function regardless of the value of x_0; therefore there is no way starting from $Y(p)$ to figure out what x_0 was. Hence information was destroyed.

Now on to the definition. Recall that a function f is *convex* in a region if the graph of the function lies below the line segment connecting any two points in that region. It is *concave* if $-f$ is convex. Consider a function $y : \mathbb{R} \to \mathbb{R}$. Then the *Legendre transform* is defined by

$$Y(p) = \begin{cases} \min_x[y(x) - xp] & \text{if } y \text{ is convex} \\ \max_x[y(x) - xp] & \text{if } y \text{ is concave.} \end{cases} \tag{4.9}$$

This definition makes sense in view of goal (1) at the beginning of this section, assuming y is differentiable. To see this, note that the maximum or the minimum corresponds to a critical point, and so

$$0 = \frac{d}{dx}\left[y(x) - xp\right] = y'(x) - p. \tag{4.10}$$

We now state two useful facts about Legendre transforms. These can be relatively straightforwardly proven. We emphasize that Legendre transformations are only defined for concave or convex functions, and that the concavity or convexity is important to prove the second point, because it guarantees that the derivative is monotonic.

Proposition 4.2.1

1. The Legendre transformation of a convex function is concave and vice versa.
2. The Legendre transformation is its own inverse.

4.3 Helmholtz free energy

Let us now return to our issue of dropping S in favor of $T = \partial U/\partial S$. According to the above section if we Legendre transform $U(S)$ out of the variable S, then the minimization will guarantee that the new independent variable is T. Calling this new function F, we obtain

$$F = U - ST. \tag{4.11}$$

This new function is guaranteed to be a thermodynamic potential, and we give it a special name: the *Helmholtz free energy*.

4.4 Cumulants

References

[1] Raza Tahir-Kheli. *General and statistical thermodynamics*. Graduate texts in physics. Springer, Berlin ; New York, 2012. ISBN 978-3-642-21480-6 978-3-642-21481-3. OCLC: ocn731919885.

[2] Markus Deserno. Legendre transforms, 2012. URL https://www.andrew.cmu.edu/course/33-765/pdf/Legendre.pdf. [Online; accessed 26-Nov-2019].

Chapter 5

Physics: Quantum Field Theories

The Standard Model (SM) of particle physics classifies all known *elementary particles*, i.e. particles with no known substructure, and describes three fundamental forces: the electromagnetic, weak, and strong forces. Elementary particles can be divided into *matter particles* (quarks and leptons); *gauge bosons*, which mediate the three aforementioned forces; and a *scalar boson*, the Higgs boson, whose field interacts directly with elementary particles that thereby acquire their mass. For each particle there exists a corresponding antiparticle; sometimes a particle is its own antiparticle. Figure 5.1 gives a schematic overview of the SM. The SM has a long history of experimental confirmations culminating in the 2012 discovery of the Higgs boson by the ATLAS and CMS experiments [1; 2].

The theoretical framework underlying the SM is an example of a Quantum Field Theory (QFT). QFTs are consistent with both quantum mechanics and relativity. Lattice gauge theories are a kind of QFT; therefore it is important for the reader to know a little bit about them.

5.1 The principle of stationary action

This section follows a fairly well known and delightful lecture by Feynman [4].

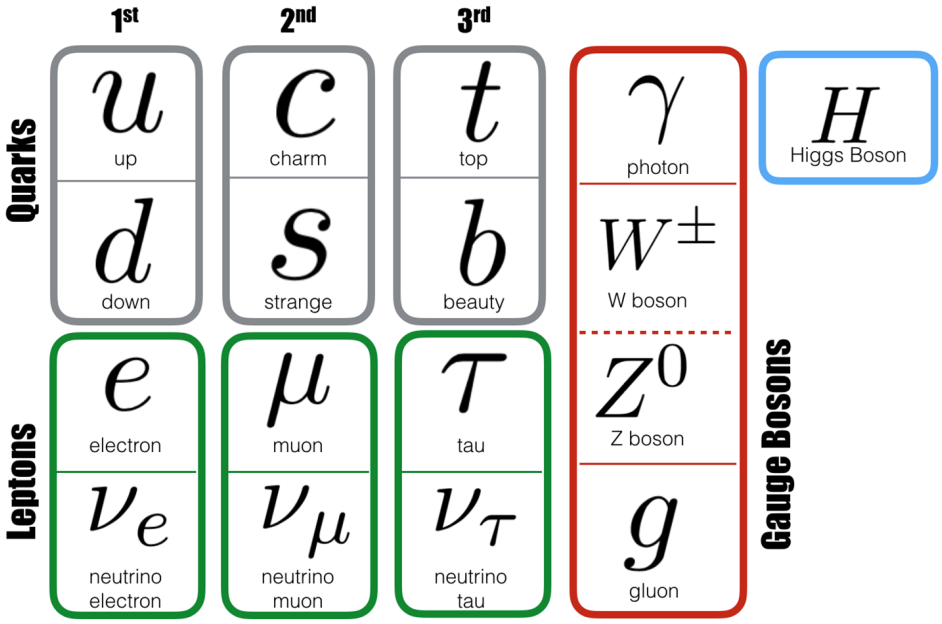

Figure 5.1: Summary of elementary SM particles. The first three columns give the three generations of matter particles. Image taken from the Physics Institute at University of Zurich [3].

5.2 The path integral

References

[1] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716(1):1–29, 2012. ISSN 03702693. doi: 10.1016/j.physletb.2012.08.020. URL http://linkinghub.elsevier.com/retrieve/pii/S037026931200857X.

[2] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716(1):30–61, 2012. ISSN 03702693. doi: 10.1016/j.physletb.2012.08.021. URL http://linkinghub.elsevier.com/retrieve/pii/S0370269312008581.

[3] University of Zurich Physik-Institut. Standard model, 2018. URL https://www.physik.uzh.ch/en/researcharea/lhcb/outreach/StandardModel.html. [Online; accessed 19-September-2018].

[4] Gottlieb, M. A. and Pfeiffer, R. The Feynman lectures on physics vol. II ch. 19: The principle of least action, 2013. URL http://www.feynmanlectures.caltech.edu/II_19.html.

Chapter 6

LGT: The Basics

LGT is introduced in Section 6.1 by reviewing local gauge symmetries in QFT, regularizing a pure gauge theory on the lattice, discussing the true continuum limit, and introducing finite temperature. In Section 6.2, reference scales are defined, and systematic error within the context of scale setting is explored. Topological observables are introduced in Section 6.3, and effects of topology barriers are considered.

6.1 Lattice gauge theory

Physically, QFT is defined on a 4D Minkowskian space-time. In LGT the 4D space-time is instead equipped with a Euclidean metric, which is related to the original metric via a *Wick rotation*

$$t \to i\tau. \tag{6.1}$$

Therefore we will work with a Euclidean metric and use downstairs summation indices. We will also use *natural units* $\hbar = c = k_B = 1$. In natural units, every physical quantity has units of some power of length. For example time has units of length, while energy, mass, and momentum have units of inverse length. We first work in the continuum, then discretize the theory by defining the lattice.

6.1.1 Local gauge symmetries

Local gauge symmetries play a central role in the SM. Starting from a Lagrangian that depends on the derivatives of some field, the requirement of

local gauge invariance suggests that we introduce a *gauge field*. This gauge field allows one to define a *covariant derivative* whose transformation law will respect the local gauge symmetry. Excitations of the gauge field are gauge bosons, which are the force-carrying particles of the SM.

As an example consider N_c complex scalar fields $\phi_i(x)$ equipped with a global $\mathrm{SU}(N_c)$ symmetry. The Lagrangian is

$$\mathcal{L}_M = -\partial_\mu \phi^\dagger(x) \partial_\mu \phi(x) + m^2 \phi^\dagger(x)\phi(x), \tag{6.2}$$

where $\phi(x)$ is the N_c-dimensional vector formed by these fields. $\mathcal{L}_M$ becomes invariant under local $\mathrm{SU}(N_c)$ transformations, i.e. transformations of the form

$$\phi(x) \to U(x)\phi(x), \tag{6.3}$$

where $U(x) \in \mathrm{SU}(N_c)$, when one replaces the partial derivative by the covariant derivative D_μ, which transforms as

$$D_\mu(x) \to U(x) D_\mu(x) U^\dagger(x). \tag{6.4}$$

We define

$$D_\mu(x) \equiv \partial_\mu + A_\mu(x), \qquad A_\mu(x) \equiv -igA^a_\mu(x)T^a, \tag{6.5}$$

where g is the bare coupling constant, $A_\mu(x)$ is the gauge field, and T^a, $a = 1, \ldots, N^2 - 1$, are the generators of the $\mathrm{SU}(N_c)$ Lie algebra $\mathfrak{su}(N_c)$. For notational convenience we now suppress dependence on x.

Proposition 6.1.1

Using this definition of D_μ, the gauge fields must change according to

$$A_\mu \to U A_\mu U^\dagger - (\partial_\mu U) U^\dagger.$$

Proof. The transformed D can be written $U D_\mu U^\dagger = \partial'_\mu + A'_\mu$. Solving for A'_μ gives

$$\begin{aligned} A'_\mu &= U(\partial_\mu + A_\mu)U^\dagger - \partial_\mu \\ &= U(\partial_\mu U^\dagger) + U A_\mu U^\dagger - \partial_\mu \\ &= \partial_\mu - (\partial_\mu U)U^\dagger + U A_\mu U^\dagger - \partial_\mu \\ &= U A_\mu U^\dagger - (\partial_\mu U) U^\dagger. \end{aligned}$$

□

The gauge field becomes dynamic by adding the kinetic part

$$\mathcal{L}_G = \frac{1}{4} F^a_{\mu\nu} F^a_{\mu\nu} = -\frac{1}{2g^2} \operatorname{tr} F_{\mu\nu} F_{\mu\nu}, \tag{6.6}$$

where

$$F^a_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu, \qquad F_{\mu\nu} \equiv -igF^a_{\mu\nu} T^a, \tag{6.7}$$

and f^{abc} are the structure constants of $\mathrm{SU}(N_c)$. $\mathcal{L}_G$ is also invariant under the transformation of (6.4). One way to see this is to use the following fact.

Proposition 6.1.2

$$F_{\mu\nu} = [D_\mu, D_\nu]\,.$$

Proof. Use the definition of D_μ and apply the above commutator to some field ψ. We get

$$\begin{aligned}
[D_\mu, D_\nu]\,\psi &= (\partial_\mu + A_\mu)(\partial_\nu \psi + A_\nu \psi) - (\mu \leftrightarrow \nu) \\
&= \partial_{\mu\nu}\psi + \partial_\mu A_\nu \psi + A_\nu \partial_\mu \psi + A_\mu \partial_\nu \psi + A_\mu A_\nu \psi - (\mu \leftrightarrow \nu) \\
&= \partial_\mu A_\nu \psi - \partial_\nu A_\mu \psi + [A_\mu, A_\nu]\,\psi \\
&= -ig\left(\partial_\mu A^a_\mu - \partial_\nu A^a_\mu\right) T^a \psi - ig^2 A^b_\mu A^c_\nu f^{bca} T^a \psi \\
&= -ig\left(\partial_\mu A^a_\mu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu\right) T^a \psi \\
&= F_{\mu\nu}\psi.
\end{aligned}$$

□

Taken altogether, the gauge-invariant, dynamical, scalar theory is described by the Lagrangian

$$\mathcal{L} = -(D_\mu \phi)^\dagger D_\mu \phi + m^2 \phi^\dagger \phi - \frac{1}{2g^2} \operatorname{tr} F_{\mu\nu} F_{\mu\nu}. \tag{6.8}$$

We would like to point out that the definitions (6.5) and (6.7) are somewhat different than the convention of many QFT books such as Srednicki [1]. An advantage of the convention we have taken, which is also used in, for instance, Montvay and Münster [2], is that one can explicitly see the dependence of the Lagrangian (6.6) on the coupling.

In this chapter we will be primarily interested in a theory with $\mathcal{L}_G$ only and gauge group $\mathrm{SU}(N_c)$; such a theory is referred to as *pure* $\mathrm{SU}(N_c)$. Often the gauge particles of pure $\mathrm{SU}(N_c)$ theories are referred to as "gluons," even when $N_c \neq 3$. Because it is non-Abelian, it has nonzero structure constants, which means it contains self-interactions of the form AAA and $AAAA$. These self-interactions are responsible for confinement, which we discuss further in the next section. For the purpose of a lattice study, it is helpful to look at a non-Abelian theory, which has a well-defined continuum limit.

6.1.2 Lattice regularization

We now define QFT on a lattice. Let $N_1, N_2, N_3, N_4 \in \mathbb{N}$. The *lattice* Ł is defined by

$$Ł \equiv \{x \,|\, x_\mu = a\, n_\mu,\, n_\mu \leq N_\mu,\, \mu = 1,2,3,4\}. \tag{6.9}$$

Here a is called the *lattice spacing*. After our Wick rotation, we identify N_1, N_2, and N_3 as the extensions of the lattice in the spatial directions, and N_4 is taken to be the extension in the Euclidean time direction. Matter fields and gauge transformations are defined on the *sites* $x \in Ł$. We shall take the lattice to have periodic boundary conditions (BCs), i.e.

$$x + aN_\mu \hat{\mu} = x, \tag{6.10}$$

where $\hat{\mu}$ is the unit vector in the direction indicated by μ. Since the lattice is discrete, one must replace partial derivatives by finite differences,

$$\partial_\mu f(x) \to \Delta_\mu f(x) \equiv \frac{f(x + a\hat{\mu}) - f(x - a\hat{\mu})}{2a}, \tag{6.11}$$

and similarly replace integrals with sums,

$$\int d^4x \to a^4 \sum_x . \tag{6.12}$$

Moreover the BCs (6.10) imply for every direction that the momentum is discretized as

$$p_\mu = \frac{2\pi}{a} \frac{n_\mu}{N_\mu}, \tag{6.13}$$

which means that momentum space integrals must also be replaced by sums

$$\int \frac{d^4p}{(2\pi)^4} \to \frac{1}{a^4 N_1 N_2 N_3 N_4} \sum_p . \tag{6.14}$$

Putting QFT on a lattice regularizes the theory. To see this, consider a field ϕ defined on the lattice. Its Fourier transform

$$\widetilde{\phi}(p) = a^4 \sum_x e^{-ipx} \phi(x) \tag{6.15}$$

is periodic in momentum space, which gives us the correspondence $p_\mu \leftrightarrow p_\mu + 2\pi/a$. Hence we can restrict momenta to the *first Brillouin zone*,

$$-\frac{\pi}{a} < p_\mu \leq \frac{\pi}{a} \tag{6.16}$$

and one obtains a UV cutoff $|p_\mu| \leq \pi/a$.

Now we define the building blocks necessary to construct paths on the lattice. The directed *link* connects x with the neighboring point $x + a\hat{\mu}$, and its corresponding *link variable* $U_\mu(x) \in \mathrm{SU}(N_c)$ is defined by

$$U_\mu(x) = e^{-aA_\mu(x)}, \tag{6.17}$$

where $A_\mu(x) \in \mathfrak{su}(N_c)$. A link variable is depicted in Fig. 6.1 (left). We associate to any path $\mathcal{C}$ the ordered product of its link variables $U(\mathcal{C})$. If we follow a path and then reverse our steps, we should end up back where we started; hence

$$U_{-\mu}(x + a\hat{\mu}) U_\mu(x) = \mathbf{1}. \tag{6.18}$$

Furthermore $U^\dagger(x)U(x) = \mathbf{1}$, so we can see the effect of the dagger on link variables:

$$U_\mu^\dagger(x) = U_{-\mu}(x + a\hat{\mu}). \tag{6.19}$$

Let $\mathcal{C}_x$ be a path on the lattice that originates and terminates at the point x. The corresponding *Wilson loop* is defined by $\mathrm{tr}\, U(\mathcal{C}_x)$. Under local gauge transformations, link variables transform as

$$U_\mu(x) \to U(x) U_\mu(x) U^\dagger(x + a\hat{\mu}), \qquad U(x) \in \mathrm{SU}(N_c), \tag{6.20}$$

which ensures the gauge invariance of Wilson loops. Maybe at this point it is worth pointing out that link variables will have a Lorentz index attached to them (they live on the links) but gauge transformations do not (they live on the sites). A *plaquette*, shown in Figure 6.1 (middle), is the smallest Wilson loop, an oriented square of side length a with corresponding link variable

$$U_{\mu\nu}^{\square}(x) = U_\mu(x) U_\nu(x + a\hat{\mu}) U_\mu^\dagger(x + a\hat{\nu}) U_\nu^\dagger(x). \tag{6.21}$$

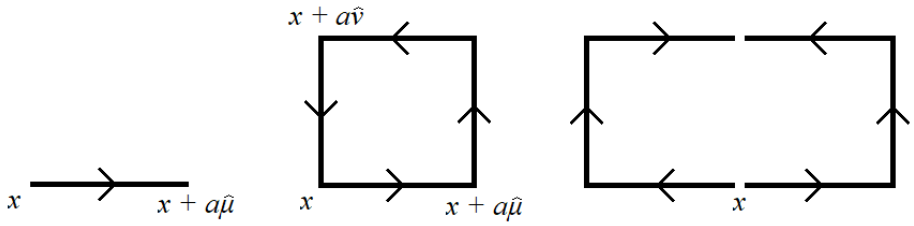

Figure 6.1: Left: A link variable. Middle: A plaquette. Right: A staple matrix in 2D.

Every link variable in 4D LGT is part of six plaquettes. The remaining three edges of any particular plaquette are shaped like a staple; therefore we call the combination

$$U_\mu^{\sqcup}(x) = \sum_{\nu\neq\mu} \Big[U_\nu(x) U_\mu(x+a\hat{\nu}) U_\nu^\dagger(x+a\hat{\mu}) + U_\nu^\dagger(x-a\hat{\nu}) U_\mu(x-a\hat{\nu}) U_\nu(x-a\hat{\nu}+a\hat{\mu}) \Big] \tag{6.22}$$

the *staple matrix*. A 2D staple matrix is shown in Fig. 6.1 (right); alternatively one can view it as one of the three terms in the sum (6.22).

Plaquettes are used to construct the gauge invariant SU(N_c) *Wilson action* [3], given by

$$S_W \equiv \beta \sum_{x,\mu<\nu} \left(1 - \frac{1}{N_c} \operatorname{Re}\operatorname{tr} U_{\mu\nu}^{\square}(x) \right). \tag{6.23}$$

The factor β is given this name in analogy to the inverse temperature in statistical mechanics. The Wilson action becomes identical with the pure SU(N_c) action in the continuum limit because the plaquette is directly related to the field tensor, which is proven in the following proposition.

Proposition 6.1.3

$$U^{\square}_{\mu\nu}(x) = \exp\left[-a^2 F_{\mu\nu}(x) + \mathcal{O}\big(a^3\big)\right].$$

Proof. Starting with the definition of the plaquette variable, we have

$$\begin{aligned}
U^{\square}_{\mu\nu}(x) &= U(x, x+a\hat{\nu})U(x+a\hat{\nu}, x+a\hat{\nu}+a\hat{\mu}) \\
&\quad \times U(x+a\hat{\mu}+a\hat{\nu}, x+a\hat{\mu})U(x+a\hat{\mu}, x) \\
&= \exp\left[aA_\nu(x)\right]\exp\left[aA_\mu(x+a\hat{\nu})\right] \\
&\quad \times \exp\left[-aA_\nu(x+a\hat{\mu})\right]\exp\left[-aA_\mu(x)\right] \\
&= \exp\left[aA_\nu(x)\right]\exp\left[a\left(A_\mu(x) + a\Delta_\nu A_\mu(x)\right) + \mathcal{O}\big(a^3\big)\right] \\
&\quad \times \exp\left[-a\left(A_\nu(x) + a\Delta_\mu A_\nu(x)\right) + \mathcal{O}\big(a^3\big)\right]\exp\left[-aA_\mu(x)\right] \\
&= \exp\left[aA_\nu + aA_\mu + a^2\Delta_\nu A_\mu + \frac{1}{2}\left[aA_\nu, aA_\mu\right] + \mathcal{O}\big(a^3\big)\right] \\
&\quad \times \exp\left[-aA_\nu - a^2\Delta_\mu A_\nu - aA_\mu + \frac{1}{2}\left[-aA_\nu, -aA_\mu\right] + \mathcal{O}\big(a^3\big)\right] \\
&= \exp\left[a^2\Delta_\nu A_\mu + a^2\left[A_\mu, A_\nu\right] - a^2\Delta_\mu A_\nu + \mathcal{O}\big(a^3\big)\right] \\
&= \exp\left[-a^2 F_{\mu\nu} + \mathcal{O}\big(a^3\big)\right].
\end{aligned}$$

In the fourth step we applied the Campbell-Baker-Hausdorff formula and dropped the x dependence for notational convenience, since at this step all the gauge fields depend on the same space-time point anyway. The fifth step uses another application of the Campbell-Baker-Hausdorff formula. □

After some algebra, the connection between the Wilson action and the action corresponding to eq. (6.6) becomes clear.

Proposition 6.1.4

$$S_W = -\frac{\beta}{4N_c}\sum_x a^4 \operatorname{tr} F_{\mu\nu}(x)F_{\mu\nu}(x) + \mathcal{O}\big(a^5\big).$$

Proof. Using the definition (6.23) and Proposition 6.1.3 we have

$$\begin{aligned} S_W &= \beta \sum_{x,\mu<\nu} \left(1 - \frac{1}{N_c} \operatorname{Re} \operatorname{tr} U^{\square}_{\mu\nu}(x)\right) \\ &= \beta \sum_{x,\mu<\nu} \left(1 - \frac{1}{2N_c} \operatorname{tr} \left[U^{\square}_{\mu\nu}(x) + U^{\square}_{\mu\nu}(x)^\dagger\right]\right) \\ &= \beta \sum_{x,\mu<\nu} \left(1 - \frac{1}{2N_c} \operatorname{tr} \left[2\mathbf{1} + \frac{a^4}{2} F_{\mu\nu}(x)^2 + \mathcal{O}\big(a^5\big)\right]\right) \\ &= \beta \sum_{x,\mu<\nu} \left(-\frac{a^4}{2N_c} \operatorname{tr} F_{\mu\nu}(x)^2 + \mathcal{O}\big(a^5\big)\right) \\ &= -\frac{\beta}{4N} \sum_x a^4 \operatorname{tr} F_{\mu\nu}(x) F_{\mu\nu}(x) + \mathcal{O}\big(a^5\big). \end{aligned}$$

The cancellation of the $\mathcal{O}(a^2)$ term can be seen as follows: The role of the $\dagger$ in $\mathrm{SU}(N_c)$ is to take the inverse. For a path of link variables, this is the same as following the path in reverse. Following a plaquette in reverse just interchanges μ and ν, which flips the sign of the leading term in the exponential of Proposition 6.1.3 because $F_{\mu\nu}$ is antisymmetric. □

In the limit $a \to 0$, the Wilson action coincides with the action $S_G = \int d^4x \, \mathcal{L}_G$ when one identifies

$$\beta = \frac{2N_c}{g^2}. \tag{6.24}$$

Because of this identification, β is also (besides g) sometimes referred to as the *coupling constant*.

We close this subsection with a remark about confinement. Let $\mathcal{C}_{RT}$ be a rectangular loop on the lattice of side lengths R and T and let $W(\mathcal{C}_{RT})$ be the corresponding Wilson loop. Then the *static quark potential* $V(R)$ is defined by

$$V(R) \equiv -\lim_{T\to\infty} \frac{1}{T} \log W(\mathcal{C}_{RT}) \tag{6.25}$$

and gives the energy of the gauge field due to two color sources separated by

a distance R. The *string tension* σ is defined by

$$\sigma \equiv \lim_{R\to\infty} \frac{1}{R} V(R). \tag{6.26}$$

If the string tension is non-vanishing, then the potential scales linearly with R in the large R limit; this phenomenon has been observed in LGT simulations [2]. Thus we see one of the major successes of LGT: it proffers an explanation of confinement.

6.1.3 Renormalization group and the continuum limit

In the limit $a \to 0$, physical quantities P should agree with experimental results, which means they should become independent of a, "forgetting" about the lattice structure. Since P depends in general also on g, this means that changes in a have to be compensated by changes in g to keep the physics constant. More precisely, it must be that

$$\lim_{a\to 0} P\Big(g(a), a\Big) = P_0 \tag{6.27}$$

where P_0 is the physical quantity's experimental value. Both Callan [4] and Symanzik [5; 6] independently formulated the requirement of constant physics as a differential equation

$$\left(\frac{\partial}{\partial \log a} + \frac{\partial g}{\partial \log a}\frac{\partial}{\partial g}\right) P = 0. \tag{6.28}$$

(The RHS of this equation is more precisely $\mathcal{O}((a/\xi)^2 \log(a/\xi))$ for a lattice system with correlation length ξ [2].) Equation (6.28) relates to a semi-group of scale changing transformations called the *renormalization group* (RG). The coefficient of the second term is called the *beta function*,

$$\beta \equiv -\frac{\partial g}{\partial \log a}, \tag{6.29}$$

and it measures how the bare coupling g must change when a changes. The use of the symbol β here is unfortunately a convention; it is not to be confused with the coupling constant. It is usually clear from context what is meant. In practice β can be determined from perturbation theory. An explicit dependence of g on a is then determined by solving the differential equation (6.29).

For example the pure SU(N_c) lattice beta function has been calculated up to 3-loop order in perturbation theory. It is given by

$$\beta_L(g) = -b_0 g^3 - b_1 g^5 - b_2^L g^7 + \mathcal{O}\big(g^9\big) \tag{6.30}$$

where

$$\begin{aligned} b_0 &= \frac{11}{3}\frac{N_c}{16\pi^2}, \qquad b_1 = \frac{34}{3}\left(\frac{N_c}{16\pi^2}\right)^2, \\ b_2^L &= \left(-366.2 + \frac{1433.8}{N_c^2} - \frac{2143.0}{N_c^4}\right)\left(\frac{N_c}{16\pi^2}\right)^3 \end{aligned} \tag{6.31}$$

have been calculated at one-loop [7; 8], two-loop [9; 10; 11], and three-loop order [12], respectively. The constants b_0 and b_1 are universal in the sense that they do not depend on the regularization scheme; however b_2 does depend on the regularization scheme, with b_2^L being the value using lattice regularization. The RG equation on the lattice is

$$\beta_L(g) = -a\frac{\mathrm{d}g}{\mathrm{d}a}, \tag{6.32}$$

and its solution is given by

$$a\Lambda_L = \exp\left(\int^g \frac{\mathrm{d}g'}{\beta_L(g')}\right) = f_{as}\big(g^2\big) \equiv f_{as}^0\big(g^2\big)\sum_{i=0}^{\infty} q_i\, g^{2i}, \tag{6.33}$$

where $q_0 = 1$, the other q_i are coefficients that can be, in principle, calculated perturbatively, and

$$f_{as}^0\big(g^2\big) \equiv \exp\left(-\frac{1}{2b_0 g^2}\right)(b_0 g^2)^{-b_1/2b_0^2}. \tag{6.34}$$

In fact from eq. (6.30) and (6.31), one obtains

$$q_1 = \frac{b_1^2 - b_2^L b_0}{2b_0^3} = \begin{cases} 0.08324 & \text{for SU(2)} \\ 0.18960 & \text{for SU(3).} \end{cases} \tag{6.35}$$

The integration constant Λ_L has units of mass and is called the *lattice* Λ-*parameter*. From eq. (6.33) one sees that

$$\Lambda_L = \lim_{g\to 0}\frac{1}{a} f_{as}^0\big(g^2\big). \tag{6.36}$$

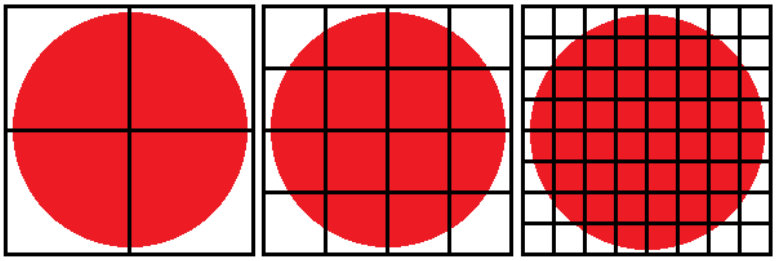

Figure 6.2: A schematic representation of the continuum limit. The red object represents some physical quantity. As the images progress to the right, the lattice spacing decreases relative to the physical length, and the bare coupling becomes weaker.

The fact that pure SU(N_c) theory has a negative beta function (6.30) has a profound physical implication. In particular when we invert eq. (6.33) keeping only universal terms, we find

$$g(a)^{-2} = b_0 \log\left(a^{-2}\Lambda_L^{-2}\right) + \frac{b_1}{b_0}\log\log\left(a^{-2}\Lambda_L^{-2}\right) + \mathcal{O}\big(1/\log\left(a^2\Lambda_L^2\right)\big). \quad (6.37)$$

Two consequences are that the coupling $g(a)$ is driven to zero as a approaches zero (UV cutoff), which is known as *asymptotic freedom*, while at low energies, $g(a)$ becomes too large for reliable perturbative analysis.

From eq. (6.33) we see that taking $g \to 0$ drives $a \to 0$. However the limit $g \to 0$ is not enough to ensure a well-defined continuum limit. The physical size of the lattice is proportional to a^4, and hence collapses to zero unless we also increase the number of sites. Therefore we extrapolate to the continuum limit by calculating our observable of interest at different values of the coupling constant, with the extensions N_1, N_2, N_3, and N_4 chosen so that the physical size of the lattice is large enough for a reliable calculation of the observable of interest. A schematic representation is shown in Figure 6.2. We note that two kinds of systematic uncertainty arise in this context. Namely, to what extent do finite lattice spacing (which limits the smallest wavelength) and finite lattice size (which limits the largest wavelength) affect our results? These questions are discussed in detail in Section 6.2.

6.1.4 Finite temperature

We now restrict our attention to lattices that have extension $N_1 = N_2 = N_3 \equiv N_s$ and $N_4 \equiv N_\tau$. Expectation values of physical observables X are given in 4D, Euclidean, pure SU(2) LGT at zero temperature by

$$\langle X \rangle = \frac{1}{Z} \int \mathcal{D}U \, e^{-S(U)} X(U), \tag{6.38}$$

where the action is related to the Lagrangian by

$$S = \int \mathrm{d}^4 x \, \mathcal{L}, \tag{6.39}$$

Z is the *partition function*

$$Z \equiv \int \mathcal{D}U \, e^{-S(U)}, \tag{6.40}$$

and the integration measure, called the *Haar* is

$$\int \mathcal{D}U \equiv \int \prod_{x,\mu} \mathrm{d}U_\mu(x) \, . \tag{6.41}$$

The quantities X and S appearing in the integral (6.38) are functionals of the configuration U, and this integral is called a *functional integral*. The Haar measure is a product of measures, one measure per link, each running over all possible values of the link; in other words, the Haar measure runs over all possible configurations. The functional integral is therefore a weighted average of the observable X over all possible configurations, each configuration receiving a weighting factor $\int \mathcal{D}U \, e^{-S}/Z$.

The functional integral for a 3D, pure SU(2) LGT system in contact with a thermal reservoir at temperature T has the same structure, except that the corresponding action is

$$S(T) = \int_0^{1/T} \mathrm{d}x_4 \int \mathrm{d}^3 x \, \mathcal{L}, \tag{6.42}$$

and the Haar measure runs over fields that are periodic in the x_4 direction. Because the functional integral for both systems is formally the same, we interpret a 4D system with $N_s \gg N_\tau$ as a 3D system at finite temperature,

with x_4 running along a temperature direction rather than a time direction. The continuum limit of the finite temperature system corresponds to $a \to 0$ with aN_s and aN_τ fixed. The *physical temperature* is seen to be

$$T = \frac{1}{aN_\tau}. \tag{6.43}$$

6.2 Reference scales

Lattice computations deliver dimensionless quantities $L = \ell/a$, where ℓ is some physical length. The requirement that the theory has a well-defined continuum limit means that for two length scales ℓ_i and ℓ_j

$$r_{ij} \equiv \frac{\ell_i}{\ell_j} = \lim_{a\to 0} \frac{L_i}{L_j} \equiv \lim_{a\to 0} R_{ij}, \tag{6.44}$$

i.e. in the continuum limit, length ratios attain their physical values. Continuum limit extrapolations of a particular length ℓ_i therefore depend on how one determines R_{ij} and on the choice of the *reference scale* or *reference length* ℓ_j. Choosing a reference scale to use for continuum limit extrapolation is called *scale setting*, and commonly one says "we set the scale with ℓ_j."

Calculation of the constants R_{ij} is prone to nontrivial statistical and systematic errors because they come from MCMC simulations performed on finite lattices with nonzero spacing. Therefore it is desirable to set the scale with a quantity that is computable with low numerical effort, has small systematic uncertainties, and good statistical precision. Controlling systematic error is discussed in Section 6.2.2, while the discussion of statistical error is postponed to Chapter 7. We begin by introducing some reference scales.

6.2.1 Defining reference scales

One choice of scale in this project is the *deconfining phase transition temperature*

$$T_c = \frac{1}{a(\beta_c)\, N_\tau}. \tag{6.45}$$

For $T < T_c$ gluons are bound into glueballs, while at higher temperatures $T > T_c$ they exist in a gluon plasma. The deconfining phase transition is a second-order phase transition for SU(2) (see Engels et al. [13] and references

therein) and a first-order transition for $\mathrm{SU}(N_c)$ when $N_c > 2$. The order parameter for this transition is the *Polyakov loop*,

$$P(\vec{x}) = \mathrm{tr} \prod_{\tau} U_4(\vec{x}, \tau), \tag{6.46}$$

which is a straight Wilson loop of length N_τ that is parallel to the Euclidean time axis and closes due to the periodic BCs. In practice, we determine β_c by looking at plots of the Polyakov loop susceptibility,

$$\chi = N_s^3 \left(\left\langle |P|^2 \right\rangle - \langle |P| \rangle^2 \right), \qquad P \equiv \sum_{\vec{x}} P(\vec{x}), \tag{6.47}$$

as a function of β and estimating (in the infinite volume limit) where it diverges. Numerical estimates of T_c are prone to systematic error because the simulations are performed at finite lattice size while T is only sharp in the infinite volume limit. It is therefore necessary to extrapolate, for fixed N_τ, the dependence of $\beta_c(N_\tau)$ on the spatial size N_s to the infinite volume limit $N_s \to \infty$. Inverting $\beta_c(N_\tau)$ gives our first length scale $N_\tau(\beta)$, which we call the *deconfinement scale*.

A reference scale due to Lüscher [14] involves using the *gradient flow*. We begin by introducing a fictitious *flow time* t and evolve the system according to the evolution equation

$$\dot{V}_\mu(x,t) = -g^2 V_\mu(x,t)\, \partial_{x,\mu} S[V(t)] \tag{6.48}$$

with initial condition

$$V_\mu(x,0) = U_\mu(x). \tag{6.49}$$

In the above, the $\mathrm{SU}(N_c)$ link derivatives are defined by

$$\begin{aligned} \partial_{x,\mu} f(V) &\equiv i \sum_a T^a \frac{d}{ds} f\left(e^{itX^a} V\right)\Big|_{t=0}, \\ X^a(x',\mu') &\equiv \begin{cases} T^a & \text{if} (x',\mu') = (x,\mu) \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{6.50}$$

Lüscher showed that the gradient flow averages the gauge field A_μ over a sphere with mean-square radius $\sqrt{8t}$ in 4D. Hence t has dimension length squared, and $\sqrt{8t}$ is interpreted as the *smoothing range* of the flow. From

eq. (6.48) we see that the gradient flow lowers the action. For pure SU(2) the link derivative of the action takes the simple form

$$g^2 \partial_{x,\mu} S(V) = \frac{1}{2} \left(V_\mu^{\square}(x) - V_\mu^{\square}(x)^\dagger \right). \tag{6.51}$$

After choosing an energy density discretization E (for example one might use the Wilson action) a scale is defined by choosing an appropriate, fixed, dimensionless *target value* y and integrating the gradient flow equation until

$$y = t^2 E(t). \tag{6.52}$$

As a function of β, a *gradient scale*

$$s(\beta) = \sqrt{t(\beta)} \tag{6.53}$$

scales like a length, provided that

1. lattice sizes are chosen so that $aN_{\text{min}} \gg \sqrt{8t}$, where $N_{\text{min}} = \min N_i$ for simulations on an $N_1 N_2 N_3 N_4$ lattice;

2. the target values are large enough so that $\sqrt{8t} \gg a$ for the smallest used flow time; and

3. the values of β are large enough to be in the SU(2) scaling region.

In contrast to the deconfinement scale, the computation of a gradient scale does not require fits or extrapolations. The only remaining ambiguity is how to choose a target value.

An alternative to the gradient flow that is similar and algorithmically simpler is known as *cooling*. Cooling was introduced as part of an investigation of topological charge in the 2D O(3) sigma model [15]. Bonati and D'Elia showed that using cooling as a smoothing technique produces similar results for topological observables as the gradient flow for pure SU(3) LGT [16]. In pure SU(2) a standard cooling step is

$$V_\mu(x, n_c) = \frac{V_\mu^{\sqcup}(x, n_c - 1)}{\sqrt{\det V_\mu^{\sqcup}(x, n_c - 1)}}, \tag{6.54}$$

where n_c is the number of cooling steps. The update (6.54) minimizes the local contribution to the action, so that the "cooling flow" decreases the

action. Like with the gradient flow, one picks a target value and iterates eq. (6.54) until

$$y = t_c^2 E(t_c), \tag{6.55}$$

and a *cooling scale* is given by

$$u(\beta) = \sqrt{t_c(\beta)}. \tag{6.56}$$

6.2.2 Continuum limit and finite size scaling

One desires to know the ratio r_{ij} of two scales in the continuum limit. In principle this could be estimated by simulating very near to the continuum limit. The continuum limit of LGT is defined in the vicinity of a second order phase transition in the bare coupling. Because the correlation length diverges near critical points, subsequent configurations become more correlated, and it requires more configurations to obtain effectively independent data. This is called *critical slowing down*. In practice, one therefore calculates R_{ij} at multiple β (hence multiple a) and extrapolates the continuum limit result based on these data. We now discuss two possible fitting forms for continuum limit extrapolation.

Using the Wilson action, ratios of observables that have units of length are known to scale as

$$R_{ij} \equiv \frac{L_i}{L_j} = \frac{\ell_i}{\ell_j}\Big(1 + \mathcal{O}\big(a^2\Lambda_L^2\big)\Big). \tag{6.57}$$

In the continuum limit, ratios of lengths approach their continuum limit values. Sometimes corrections depending on a, such as in the equation above, are referred to as *lattice artifacts*. In general the approach to the continuum limit is thought to have lattice artifacts of power p (RG considerations show that these a^p artifacts are modified by powers of logarithms [2]) where p depends on the lattice discretization. The Wilson action in particular has $p = 2$. Equation (6.57) suggests a two-parameter fit of the form

$$R_{ij} = r_{ij} + c_{ij}\left(\frac{1}{L_j}\right)^2, \tag{6.58}$$

where r_{ij} and c_{ij} are the fit parameters. We will refer to this behavior as *standard scaling*.

Another possibility for continuum limit extrapolation uses the asymptotic scaling relation (6.33)

$$a\Lambda_L = f_{\rm as}(\beta). \tag{6.59}$$

We start by noting that the scale L_i calculated on the lattice is some function of the spacing, so it can be expanded as a power series in a:

$$L_i = \frac{c_i}{a\Lambda_L}\left(1 + \sum_{k=1}^{\infty} \alpha_{i\,k} (a\Lambda_L)^k\right), \tag{6.60}$$

where the $\alpha_{i\,k}$ are expansion coefficients. Allton suggested using this equation to fit the approach to the continuum limit [17]. Inserting eq. (6.59) into the above power series yields

$$L_i = \frac{c_i}{f_{\rm as}(\beta)}\left(1 + \sum_{k=1}^{\infty} \alpha_{i\,k}\, f_{\rm as}(\beta)^{\,k}\right). \tag{6.61}$$

In practice $f_{\rm as}$ is only known up to three loops, so we must truncate it at some order m. Furthermore to have a finite number of fit parameters, we must truncate the power series at some order n. Hence, the approach of a length to the continuum limit can be fit according to

$$L_i = \frac{c_i^{mn}}{f_{\rm as}^m(\beta)}\left(1 + \sum_{k=1}^{n} \alpha_{i\,k}^{mn}\, f_{\rm as}^m(\beta)^{\,k}\right), \tag{6.62}$$

where upper indices m and n are attached to quantities that will change if m or n change. The fit parameters are c_i^{mn} and the α_{ik}^{mn}.

In general, asymptotic scaling would allow $\mathcal{O}(a)$ corrections. In order to ensure non-perturbative corrections are $\mathcal{O}(a^2)$, we improve on Allton by demanding that all scales have the same $k = 1$ term $\alpha_{i,1}^{mn}$; then terms of order a cancel in the ratio. Using eq. (6.62) along with this restriction, one obtains

$$R_{ij} = r_{ij} + \sum_{k=2}^{n} \kappa_{i\,k}^{mn}\, f_{\rm as}^m\,(L_j)^{\,k}, \tag{6.63}$$

where the fit parameters are now r_{ij} and the $\kappa_{i\,k}^{mn}$. One can switch the domain of $f_{\rm as}^m$ from β to the reference L_j using, for instance, eq. (6.62). The continuum limit estimate r_{ij} also depends on m and n, but we have suppressed these indices for clearer comparison with the standard scaling

fit (6.58). We will refer to the behavior of eq. (6.62) or (6.63) as *asymptotic scaling*.

If we carry out a naive continuum limit without changing the extension of the lattice, its physical volume collapses to zero. Ideally, calculations would be performed in the *thermodynamic limit*, where $N_s \to \infty$ and $N_\tau \to \infty$, and then take the limit $a \to 0$. In practice, the infinite volume observable is determined by simulating at fixed β on lattices of several sizes, then extrapolating to the thermodynamic limit. For some observables, the dependence on finite lattice size is known from theory. For example the critical coupling constant $\beta_c(N_\tau)$ is known [13] to depend on N_s as

$$\beta_c(N_\tau, N_s) = \beta_c(N_\tau) + a_1(N_\tau) N_s^{a_2(N_\tau)}. \tag{6.64}$$

The $N_s = \infty$ result $\beta_c(N_\tau)$ can then be extracted from a fit of the three parameters $\beta_c(N_\tau)$, $a_1(N_\tau)$, and $a_2(N_\tau)$.

6.3 Topological invariants

6.3.1 Topological winding number

This section follows Chapter 93 of Srednicki [1]; more details can be found there. We start by considering classical, pure SU(2) gauge theory

$$\mathcal{L} = -\frac{1}{2g^2} F_{\mu\nu} F_{\mu\nu} \tag{6.65}$$

at fixed x_4, focusing for the moment on U that are time-independent. Let $U \equiv U(\vec{x}) \in \mathrm{SU}(2)$, and set the BC $U(\infty) = U_0$ for some constant matrix U_0. The *topological winding number* or *Pontryagin index* of the map U is

$$n \equiv \frac{1}{24\pi^2} \int \mathrm{d}^3x \, \epsilon_{ijk} \, \mathrm{tr}\, U \, \partial_i U^\dagger U \, \partial_j U^\dagger U \, \partial_k U^\dagger. \tag{6.66}$$

The winding number is invariant under coordinate changes since the Jacobian of the measure cancels the Jacobian of the partial derivatives. Given the BC, it is also invariant under smooth deformations of U. To show this, we first establish a useful Lemma.

Lemma 6.3.1

$$\delta\left(U\partial_k U^\dagger\right) = -U\partial_k\left(U^\dagger\delta U\right)U^\dagger.$$

Proof. Note that $\delta U^\dagger = -U^\dagger\delta U\, U^\dagger$. Hence

$$\begin{aligned}\delta\left(U\partial_k U^\dagger\right) &= \delta U\partial_k U^\dagger + U\partial_k\delta U^\dagger \\ &= \quad -U\partial_k\left(U^\dagger\delta U\, U^\dagger\right) \\ &= \quad -U\left(\partial_k U^\dagger\delta U U^\dagger + U^\dagger\partial_k\delta U\, U^\dagger + U^\dagger\delta U\partial_k U^\dagger\right).\end{aligned}$$

Cancelling the first and last terms and using the product rule gives the result. □

Theorem 6.3.1

The topological winding number is invariant under smooth deformations of U.

Proof. We integrate eq. (6.66) over a time-slice of space-time, which we call Ω. Then

$$\begin{aligned}\delta n &= -\frac{1}{24\pi^2}\,\delta\int_\Omega \mathrm{d}^3x\,\epsilon_{ijk}\,\mathrm{tr}\,U\partial_i U^\dagger\,U\partial_j U^\dagger\,U\partial_k U^\dagger \\ &= -\frac{1}{8\pi^2}\int_\Omega \mathrm{d}^3x\,\epsilon_{ijk}\,\mathrm{tr}\,\delta\left(U\partial_i U^\dagger\right)U\partial_j U^\dagger\,U\partial_k U^\dagger \\ &= +\frac{1}{8\pi^2}\int_\Omega \mathrm{d}^3x\,\epsilon_{ijk}\,\mathrm{tr}\,\partial_i\left(U^\dagger\delta U\right)U^\dagger\partial_j U\,U^\dagger\partial_k U \\ &= +\frac{1}{8\pi^2}\int_{\partial\Omega} \mathrm{d}S_i\,\epsilon_{ijk}\mathtt{r}\mathtt{U}^\dagger\delta\mathtt{U}\,\mathtt{U}^\dagger\partial_{\mathtt{j}}\mathtt{U}\,\mathtt{U}^\dagger\partial_{\mathtt{k}}\mathtt{U} \\ &\quad -\frac{1}{8\pi^2}\int_\Omega \mathrm{d}^3x\,\epsilon_{ijk}\,\mathrm{tr}\,U^\dagger\delta U\partial_i\left[U^\dagger\partial_j U\,U^\dagger\partial_k U\right].\end{aligned}$$

In the second step we used that the trace is cyclic. In the third step we used Lemma 6.3.1 as well as $U\partial_\mu U^\dagger = -\partial_\mu U U^\dagger$. In the last step we integrated by parts. The first integral is over the time-slice boundary evaluated at infinity. Since $\partial_j U = \partial_j U_0 = 0$ there, this termm

vanishes. The integrand of the remaining integral is expanded as

$$\epsilon_{ijk}\,\mathrm{tr}\Big[\partial_i U^\dagger \partial_j U U^\dagger \partial_k U + \partial_j U^\dagger \partial_i U U^\dagger \partial_k U + U^\dagger \partial_{ij} U U^\dagger \partial_k U + U^\dagger \partial_j U U^\dagger \partial_{ik} U\Big].$$

Terms with double derivatives vanish, as they are symmetric with respect to exchange of indices, while ϵ is antisymmetric. The remaining terms are also shown to vanish using the antisymmetry of ϵ in addition to cyclically permuting terms under the trace. □

The quantity (6.66) is called a winding number because it counts the number of times the mapping U "winds around" or "covers" the integration region. Let us see how this works in the present case. The integration region is the 3D surface of space-time, which is homeomorphic to the 3-sphere S^3. A point $\hat{x} \in S^3$ is specified by two polar angles χ and ψ and an azimuthal angle ϕ as

$$\hat{x} = \begin{pmatrix} \mathrm{s}_\chi\,\mathrm{s}_\psi\,\mathrm{c}_\phi \\ \mathrm{s}_\chi\,\mathrm{s}_\psi\,\mathrm{s}_\phi \\ \mathrm{s}_\chi\,\mathrm{c}_\psi \\ \mathrm{c}_\chi \end{pmatrix}. \tag{6.67}$$

Then the mapping $U : S^3 \to \mathrm{SU}(2)$ given by

$$U(\hat{x}) = \begin{pmatrix} \mathrm{c}_\chi + i\,\mathrm{s}_\chi\,\mathrm{c}_\psi & i\,\mathrm{s}_\chi\,\mathrm{s}_\psi\,e^{-im\phi} \\ i\,\mathrm{s}_\chi\,\mathrm{s}_\psi\,e^{im\phi} & \mathrm{c}_\chi - i\,\mathrm{s}_\chi\,\mathrm{c}_\psi \end{pmatrix}, \tag{6.68}$$

has winding number m. Intuitively, one can see this in the following manner: Any SU(2) matrix can be written in terms of four real components as

$$U = a_4 \mathbf{1} + i\vec{a}\cdot\vec{\sigma}, \tag{6.69}$$

where $a_\mu a_\mu = 1$. The vector corresponding to the map (6.68) is

$$\hat{a} = \begin{pmatrix} \mathrm{s}_\chi\,\mathrm{s}_\psi\,\mathrm{c}_{m\phi} \\ \mathrm{s}_\chi\,\mathrm{s}_\psi\,\mathrm{s}_{m\phi} \\ \mathrm{s}_\chi\,\mathrm{c}_\psi \\ \mathrm{c}_\chi \end{pmatrix}. \tag{6.70}$$

We see that if we sweep through ϕ, $\hat{x}$ sweeps over S^3 once while $\hat{a}$ sweeps over S^3 m times. And indeed, one can see that eq. (6.66) extracts the winding

number by plugging in the mapping (6.68). This is done in the following Proposition.

Proposition 6.3.1

Consider the map $U : S^3 \to \mathrm{SU}(2)$ given by

$$U(\hat{x}) = \begin{pmatrix} \mathrm{c}_\chi + i\,\mathrm{s}_\chi\,\mathrm{c}_\psi & i\,\mathrm{s}_\chi\,\mathrm{s}_\psi\,e^{-im\phi} \\ i\,\mathrm{s}_\chi\,\mathrm{s}_\psi\,e^{im\phi} & c_\chi - i\,\mathrm{s}_\chi\,\mathrm{c}_\psi \end{pmatrix}.$$

Then U has winding number m.

Proof. Plugging this map into eq. (6.66) we find

$$\begin{aligned} n &= -\frac{1}{24\pi^2}\int_{S^3} \mathrm{d}^3x\,\epsilon_{ijk}\,\mathrm{tr}\,U\partial_i U^\dagger\,U\partial_j U^\dagger\,U\partial_k U^\dagger \\ &= -\frac{1}{24\pi^2}\int_0^\pi \mathrm{d}\chi\int_0^\pi \mathrm{d}\psi\int_0^{2\pi}\mathrm{d}\phi\,\epsilon_{\alpha\beta\gamma}\,\mathrm{tr}\,U\partial_\alpha U^\dagger\,U\partial_\beta U^\dagger\,U\partial_\gamma U^\dagger, \end{aligned}$$

where α, β, $\gamma \in \{\chi, \psi, \phi\}$ and $\epsilon_{\chi\psi\phi} \equiv +1$. Since the trace is cyclic, all even permutations of χ, ψ, ϕ give the same contribution to the integral, and similarly for all odd permutations. Hence

$$\begin{aligned} n = -\frac{1}{8\pi^2}\int_0^\pi \mathrm{d}\chi\int_0^\pi \mathrm{d}\psi\int_0^{2\pi}\mathrm{d}\phi\,\epsilon_{\chi\psi\phi}\,\mathrm{tr}\,\Big(&U\partial_\chi U^\dagger\,U\partial_\psi U^\dagger\,U\partial_\phi U^\dagger \\ &- U\partial_\chi U^\dagger\,U\partial_\phi U^\dagger\,U\partial_\psi U^\dagger\Big). \end{aligned}$$

Next we compute

$$\begin{aligned} U^\dagger &= \begin{pmatrix} \mathrm{c}_\chi - i\,\mathrm{s}_\chi\,\mathrm{c}_\psi & -i\,\mathrm{s}_\chi\,\mathrm{s}_\psi\,e^{-im\phi} \\ -i\,\mathrm{s}_\chi\,\mathrm{s}_\psi\,e^{im\phi} & c_\chi + i\,\mathrm{s}_\chi\,\mathrm{c}_\psi \end{pmatrix} \\ \partial_\chi U^\dagger &= \begin{pmatrix} -\,\mathrm{s}_\chi - i\,\mathrm{c}_\chi\,\mathrm{c}_\psi & -i\,\mathrm{c}_\chi\,\mathrm{s}_\psi\,e^{-im\phi} \\ -i\,\mathrm{c}_\chi\,\mathrm{s}_\psi\,e^{im\phi} & -s_\chi + i\,\mathrm{c}_\chi\,\mathrm{c}_\psi \end{pmatrix} \\ \partial_\psi U^\dagger &= \begin{pmatrix} +i\,\mathrm{s}_\chi\,\mathrm{s}_\psi & -i\,\mathrm{s}_\chi\,\mathrm{c}_\psi\,e^{-im\phi} \\ -i\,\mathrm{s}_\chi\,\mathrm{c}_\psi\,e^{im\phi} & -i\,\mathrm{s}_\chi\,\mathrm{s}_\psi \end{pmatrix} \\ \partial_\phi U^\dagger &= \begin{pmatrix} 0 & -m\,\mathrm{s}_\chi\,\mathrm{s}_\psi\,e^{-im\phi} \\ m\,\mathrm{s}_\chi\,\mathrm{s}_\psi\,e^{im\phi} & 0 \end{pmatrix} \end{aligned}$$

and plug into the above equation. I didn't see any further simplification, so all that remains is to multiply the matrices, take the trace, and carry out the integrations. That seemed tedious, and I'm certain I would make a mistake, so I just plugged this into Mathematica to find

$$n = m.$$

□

Finally we prove another fact about the winding number that will be useful in the following section.

Proposition 6.3.2

Let $U_n : S^3 \to S^3$ have winding number n and $U_k : S^3 \to S^3$ have winding number k. Then the map $U_n U_k$ has winding number $n + k$.

Proof. The total winding number for the map $U_n U_k$ can be written

$$w = \frac{1}{24\pi^2} \int_0^\pi \mathrm{d}\chi \int_0^\pi \mathrm{d}\psi \left(\int_0^\pi + \int_\pi^{2\pi} \right) \mathrm{d}\phi \, \epsilon_{\alpha\beta\gamma} \operatorname{tr} U_n U_k \partial_\alpha (U_n U_k)^\dagger \times (\beta \text{ term}) (\gamma \text{ term}) .$$

From Theorem 6.3.1, we know we can smoothly deform U_n to $\mathbf{1}$ for $x_3 < 0$ without changing w. Then for $0 \leq \phi \leq \pi$, we have $\partial_i U_k = 0$ and $U_n U_k = U_k$, and we can clearly identify the first contribution to the above integral as k. Similarly, we smoothly deform U_k to $\mathbf{1}$ for $x_3 > 0$ and find the second contribution to be n. Thus,

$$w = n + k.$$

□

6.3.2 Theta vacua

Now we have shown that (6.66) is invariant under smooth deformations of U, and we can believe that it really extracts the number of times U as a mapping covers the integration region. Next we want to understand how the

winding number relates to physics. Consider two maps U and U' that are gauge transformations of zero and with different winding numbers. Since the winding number is a topological invariant, the only way to deform U to U' is to pass through configurations with $F_{\mu\nu} \neq 0$; in other words, there is an energy barrier between U and U'. The corresponding quantum theory therefore has degenerate vacuum states characterized by their winding numbers.

Between two quantum states $|n\rangle$ and $|n'\rangle$, there is a transition amplitude $\langle n'| H |n\rangle$ where H is the Hamiltonian. Let us discuss its matrix elements. Since

1. the product of two maps with winding numbers n and k has winding number $n+k$, which one can see from eq. (6.66) by smoothly deforming U_n to $\mathbf{1}$ for $x_3 < 0$ and smoothly deforming U_k to be $\mathbf{1}$ for $x_3 > 0$;

2. the winding number is odd under parity, which follows from the fact that negative winding numbers reverse orientation; and

3. the Yang-Mills Hamiltonian is parity invariant,

it follows that the tunneling amplitude depends on $|n - n'|$ only. This can be seen in a few steps. From the first point, we see that a gauge transformation U_k maps a field configuration with winding number n to one with winding number $n + k$. In the corresponding quantum theory, this transformation is achieved through a unitary operator

$$\mathcal{U}_k |n\rangle = |n + k\rangle . \tag{6.71}$$

The pure SU(2) Hamiltonian is built out of gauge invariant field strengths, so we must have

$$\mathcal{U}_k H \mathcal{U}_k^\dagger = H. \tag{6.72}$$

Inserting factors of $\mathcal{U}_k^\dagger \mathcal{U}_k = \mathbf{1}$ into the matrix element $\langle n| H |n'\rangle$ and using eq. (6.71) and (6.72) yields

$$\langle n| H |n'\rangle = \langle n + k| H |n' + k\rangle . \tag{6.73}$$

Hence the matrix element depends on $n - n'$ only. From the second and third points, we see that $P |n\rangle = |-n\rangle$ and $PHP^{-1} = H$; therefore we similarly find

$$\langle n| H |n'\rangle = \langle -n| H |-n'\rangle , \tag{6.74}$$

showing that the matrix element depends on $|n - n'|$ only. One can use this fact to show that H is diagonalized by *theta vacua*, which are states of the form

$$|\theta\rangle \equiv \sum_{n=-\infty}^{\infty} e^{-in\theta} |n\rangle . \tag{6.75}$$

The parameter θ is referred to as the *vacuum angle*.

6.3.3 Topological charge and instantons

We will now discuss the topology of gauge field configurations defined on all space-time. Let $r = (x_\mu x_\mu)^{1/2}$. We require that

$$A_\mu(x) \to U(x)\partial_\mu U^\dagger(x) \tag{6.76}$$

as $r \to \infty$ to keep the action finite. (Infinite actions are exponentially suppressed in the path integral.) The 3D integration region will be the surface of space-time at infinity. In addition to the BC $U(\infty) = U_0$, we specify U at $x_4 = -\infty$ to have winding number n_- and U at $x_4 = +\infty$ to have winding number n_+. The entire boundary is homeomorphic to S^3, and the winding number of U is

$$Q \equiv n_+ - n_-, \tag{6.77}$$

where the relative minus sign is due to the surfaces at $x_4 = \pm\infty$ having opposite orientation. We indicate this winding number with a Q, and it is often called the *topological charge*.

By viewing the integrand of eq. (6.66) as the surface integral over a 4D region, defining the *Chern-Simons current*

$$J_\mu^{CS} \equiv 2\epsilon_{\mu\nu\rho\sigma} \operatorname{tr}\left(a_\nu F_{\rho\sigma} + \frac{2}{3} A_\nu A_\rho A_\sigma\right), \tag{6.78}$$

and applying Gauss's theorem, one can identify the winding number as an integral over the four-divergence of J_μ^{CS}. We find

$$Q = \frac{1}{16\pi^2} \int \mathrm{d}^4x \operatorname{tr} {}^*F_{\mu\nu} F_{\mu\nu} \equiv \int \mathrm{d}^4x\, q, \tag{6.79}$$

where

$${}^*F_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma} F_{\rho\sigma} \tag{6.80}$$

is the *dual* field strength tensor. The quantity q is called the *topological charge density*. This form of the topological charge is helpful to look find vacuum solutions to the Euclidean field equations, and it also motivates one of the possible definitions for topological charge on the lattice. We prove it now.

Proposition 6.3.3

The topological charge can be written in terms of the field strength as

$$Q = \frac{1}{16\pi^2} \int \mathrm{d}^4x \, \mathrm{tr} \, {}^*F_{\mu\nu} F_{\mu\nu}.$$

Proof. Starting from the definition of the winding number we have

$$Q = -\frac{1}{24\pi^2} \int \mathrm{d}^3x \, \epsilon_{\nu\rho\sigma} \, \mathrm{tr} \, U\partial_\nu U^\dagger \, U\partial_\rho U^\dagger \, U\partial_\sigma U^\dagger.$$

Recasting this integral as a 4D surface integral and noting that $\epsilon_{r\chi\psi\phi} = -1$, which by looking at the Jacobian for this change of variables leads to an overall minus sign, we obtain

$$\begin{aligned} Q &= \frac{1}{24\pi^2} \int \mathrm{d}S_\mu \, \epsilon_{\mu\nu\rho\sigma} \, \mathrm{tr} \, U\partial_\nu U^\dagger \, U\partial_\rho U^\dagger \, U\partial_\sigma U^\dagger \\ &= \frac{1}{24\pi^2} \int \mathrm{d}S_\mu \, \epsilon_{\mu\nu\rho\sigma} \, \mathrm{tr} \, A_\nu A_\rho A_\sigma. \end{aligned}$$

From the BCs we know that $F_{\rho\sigma} = 0$ on this surface, so we are able to replace the integrand in the winding number with J^{CS}. We get

$$Q = \frac{1}{32\pi^2} \int \mathrm{d}S_\mu \, J^{CS}_\mu = \frac{1}{32\pi^2} \int \mathrm{d}^4x \, \partial_\mu J^{CS}_\mu$$

by the divergence theorem. It remains to compute $\partial_\mu J^{CS}_\mu$. We have

$$\begin{aligned}
\partial_\mu J^{CS}_\mu &= 2\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,\Big[\partial_\mu A_\nu F_{\rho\sigma} + A_\nu \partial_\mu F_{\rho\sigma} \\
&\qquad\qquad + \frac{2}{3}\left(\partial_\mu A_\nu A_\rho A_\sigma + A_\nu \partial_\mu A_\rho A_\sigma + A_\nu A_\rho \partial_\mu A_\sigma\right)\Big] \\
&= 2\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,\Big[\partial_\mu A_\nu F_{\rho\sigma} + A_\nu \partial_\mu F_{\rho\sigma} + 2\partial_\mu A_\nu A_\rho A_\sigma\Big] \\
&= \epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,\Big[\partial_\mu A_\nu F_{\rho\sigma} - \partial_\nu A_\mu F_{\rho\sigma} + 2A_\nu \partial_\mu [A_\rho, A_\sigma] + 4\partial_\mu A_\nu A_\rho A_\sigma\Big] \\
&= \epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,\Big[\partial_\mu A_\nu F_{\rho\sigma} - \partial_\nu A_\mu F_{\rho\sigma} \\
&\qquad\qquad + [A_\mu, A_\nu]\big(\partial_\rho A_\sigma - \partial_\sigma A_\rho + [A_\rho, A_\sigma]\big)\Big] \\
&= \epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\,F_{\mu\nu}F_{\rho\sigma} \\
&= 2\,\mathrm{tr}\,{}^*F_{\mu\nu}F_{\mu\nu}.
\end{aligned}$$

To get to the second line, we used the fact that cyclic permutations of products under the trace leave the trace unchanged; the fact that ϵ is antisymmetric; and relabelled dummy indices. To get to the third line, we expanded the field strength tensor; and used the fact that terms with second-derivatives are symmetric and therefore vanish when contracted with ϵ. Finally to get to the fourth line, one can use the same tricks as with the second line. In addition, note that $\epsilon\,\mathrm{tr}\,AAAA = 0$ because cyclic permutations of four indices in ϵ flip the sign, while cyclic permutations of the $AAAA$ indices under the trace leave it unchanged; therefore we can add terms of this form inside the trace with impunity and obtain the $[A, A][A, A]$ term. Plugging this result back into our expression for the winding number completes the proof. □

With eq. (6.79) we can find vacuum solutions to the Euclidean field equations

$$D_\mu F_{\mu\nu} = 0. \tag{6.81}$$

The trick is to construct a lower bound on the action. Then if we can find a solution saturating the bound, it must solve the field equations, since it minimizes the action. This is called a *Bogomolny bound*. Using eq. (6.65), we find

$$S \geq 8\pi^2 |Q|/g^2, \tag{6.82}$$

which becomes saturated when

$$ {}^*F_{\mu\nu} = (\text{sign } n) F_{\mu\nu}. \tag{6.83} $$

Proposition 6.3.4

For configurations with topological charge Q, the action is bounded below by

$$ S \geq \frac{8\pi^2 |Q|}{g^2}. $$

Proof. Note that ${}^*F_{\mu\nu}{}^*F_{\mu\nu} = F_{\mu\nu}F_{\mu\nu}$, so

$$ \frac{1}{2} \operatorname{tr} \left({}^*F_{\mu\nu} \pm F_{\mu\nu} \right)^2 = \operatorname{tr} F_{\mu\nu}F_{\mu\nu} \pm \operatorname{tr} {}^*F_{\mu\nu}F_{\mu\nu}. $$

The LHS of the above equation is non-negative, so

$$ \int \mathrm{d}^4x \ \operatorname{tr} F_{\mu\nu}F_{\mu\nu} \geq \left| \int \mathrm{d}^4x \ \operatorname{tr} {}^*F_{\mu\nu}F_{\mu\nu} \right|. $$

The LHS of the above equation is $2g^2S$ while the RHS is, according to Proposition 6.3.3, $16\pi^2|Q|$. This completes the proof. □

We arrive at an explicit solution to the above equation using the map (6.68) with $Q = 1$ ($m = 1$).

Proposition 6.3.5

The equation

$$ {}^*F_{\mu\nu} = F_{\mu\nu} $$

is solved by

$$ A_\mu(x) = \frac{r^2}{r^2 + R^2} U(\hat{x}) \partial_\mu U^\dagger(\hat{x}), $$

where $\hat{x} = x/r$.

Proof. We start with the ansatz

$$ A_\mu(x) = f(r) U(\hat{x}) \partial_\mu U^\dagger(\hat{x}), $$

with $f(\infty) = 1$ and $f(0) = 0$. Plugging this ansatz into the field tensor, we get

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu f\, U\partial_\nu U^\dagger + f\partial_\mu U\partial_\nu U^\dagger + f^2 U\partial_\mu U^\dagger U\partial_\nu U^\dagger - (\mu \leftrightarrow \nu) \\ &= \partial_\mu f\, U\partial_\nu U^\dagger + f(1-f)\partial_\mu U\partial_\nu^\dagger - (\mu \leftrightarrow \nu). \end{aligned}$$

Terms symmetric in μ and ν vanished, and we utilized $\partial_\mu U^\dagger = -U^\dagger \partial_\mu U U^\dagger$. To proceed, we need to know the components of ∂. They are

$$\partial = e_r \frac{\partial}{\partial r} + e_\chi \frac{1}{r}\frac{\partial}{\partial \chi} + e_\psi \frac{1}{r s_\chi}\frac{\partial}{\partial \psi} + e_\phi \frac{1}{r s_\chi s_\psi}\frac{\partial}{\partial \phi},$$

where e_i is the unit vector in direction i. Since f is a function of r only and U is a function of the angles only, this implies

$$F_{r\chi} = \frac{1}{r} f' U\partial_\chi U^\dagger$$

and

$$F_{\psi\phi} = \frac{1}{r^2 s_\chi^2 s_\psi} f(1-f)\left(\partial_\psi U\partial_\phi U^\dagger - \partial_\phi U\partial_\psi U^\dagger\right).$$

From the definition of the dual tensor, we have ${}^*F_{r\chi} = -F_{\psi\phi}$, since $\epsilon_{r\chi\psi\phi} = -1$. To satisfy the instanton equation ${}^*F_{\mu\nu} = F_{\mu\nu}$ we must therefore have $F_{r\chi} = -F_{\psi\phi}$. Because the variables are separated in F, we conclude

$$kf' = kf(1-f)$$

and

$$U\partial_\chi U^\dagger = -\frac{1}{c s_\chi^2 s_\psi}\left(\partial_\psi U\partial_\phi U^\dagger - \partial_\phi U\partial_\psi U^\dagger\right)$$

for some constant k. Plugging the explicit mapping into the latter equation yields $k = 2$. The former, ordinary differential equation is then easily solved. The result is

$$f(r) = \frac{r^2}{r^2 + R^2},$$

where R is a constant of integration. □

This solution is called the *instanton* [18] and the integration constant R

is called the *instanton size*. The instanton mediates between vacuum configurations at times $+\infty$ and $-\infty$ with winding numbers n_+ and n_-. When $Q = -1$ we have an *anti-instanton*. When $|Q| > 1$, the mediating solution is constructed of multiple instantons or anti-instantons. When separations are large compared to their sizes, we call this a *dilute gas* of instantons or anti-instantons. From eq. (6.82) we see that each instanton or anti-instanton contributes $8\pi^2/g^2$ to the Bogomolny bound.

The *topological susceptibility* is defined as

$$\chi_Q \equiv \int \mathrm{d}^4x \, \langle q(x)q(0)\rangle \,, \tag{6.84}$$

where q is the topological charge density of eq. (6.79). The topological susceptibility gives evidence that the topological structure of the underlying gauge fields has phenomenological significance. In particular, by performing a calculation in the large N_c limit, Witten and Veneziano [19; 20] showed that at $N_c = \infty$ the η' mass is related to the topological susceptibility through

$$m_{\eta'}^2 + m_\eta^2 - 2m_K = \frac{4N_f\chi_Q}{f_\pi^2}, \tag{6.85}$$

where m_η is the η mass, m_K is the mass of the kaon, N_f is the number of fermion flavors, and f_π is the pion decay constant. This mechanism can be used to explain the $\eta - \eta'$ mass difference. Plugging experimental values into the above formula for $N_f = 3$, one finds

$$\chi_Q \approx (180 \text{ MeV})^4. \tag{6.86}$$

While a conventional derivation of the Witten-Veneziano formula depends on large N_c, lattice calculations for pure SU(2) and pure SU(3) land relatively close to eq. (6.86).

6.3.4 Topological charge on the lattice

It may seem counter-intuitive that a well-defined notion of topology exists at all on the lattice. By taking an open cover of the lattice as the manifold (e.g., one can associate a hypercube to each site, keeping periodic BCs) one can define a topological charge on the lattice; the non-trivial topology is contained in transition functions connecting cells of the open cover. For example Lüscher [21] showed the existence of a well-defined topological charge, which

approaches Q in the continuum limit, for configurations with a small action density, i.e., with

$$\max \operatorname{tr}\left(\mathbf{1} - U^{\square}\right) < \epsilon \tag{6.87}$$

for some small $\epsilon > 0$. The maximum is taken over all plaquettes. Configurations not satisfying this inequality are called *exceptional*. Since

$$\left\langle \operatorname{tr}\left(\mathbf{1} - U^{\square}\right)\right\rangle = \frac{3}{8}g^2 + \mathcal{O}\left(g^4\right), \tag{6.88}$$

one sees that exceptional configurations become suppressed as the lattice spacing decreases. In order for a configuration of one topological charge to tunnel to another, it must pass through such an exceptional configuration; hence as g decreases (as β increases), it becomes increasingly more difficult for a configuration to tunnel out of its topological sector. This phenomenon is sometimes called *topological freezing*.

Further definitions of topological charge on the lattice can be found in reviews such as the review by Kronfeld [22]. For our definition of topological charge, we follow the example of eq. (6.79) using the rule (6.12). It is reasonable to measure a topological charge on the lattice by

$$Q_L = a^4 \sum_x q_L(x), \tag{6.89}$$

where the sum is over all lattice sites and

$$q_L(x) = -\frac{1}{2^9\pi^2} \sum_{\mu\nu\rho\sigma=\pm 1}^{\pm 4} \tilde{\epsilon}_{\mu\nu\rho\sigma} \operatorname{tr} U^{\square}_{\mu\nu}(x) U^{\square}_{\rho\sigma}(x). \tag{6.90}$$

Here $\tilde{\epsilon} = \epsilon$ for positive indices while $\tilde{\epsilon}_{\mu\nu\rho\sigma} = -\tilde{\epsilon}_{(-\mu)\nu\rho\sigma}$ for negative indices. The summation over backwards indices along with the definition of $\tilde{\epsilon}$ ensures q_L has negative parity. The restriction of generated configurations to a subset with some fixed topological charge is what we mean by *topological sector*. The lattice expression for the topological susceptibility is

$$\chi_L = a^4 \sum_x \langle q_L(x) q_L(0)\rangle = \frac{1}{N^4}\left\langle Q_L^2\right\rangle, \tag{6.91}$$

where we have assumed a geometry $N \equiv N_1 = N_2 = N_3 = N_4$ and utilized the translational invariance due to periodic BCs.

Lattice gauge theories typically experience local fluctuations of the gauge fields, which are produced stochastically. These fluctuations blur the topological structure of the lattice, and must therefore be stripped away from the configuration before measuring Q_L. The signal is considerably improved by *smoothing*, where one replaces each link by a local average of links; Q_L is then constructed on the smoothed field.

Standard cooling minimizes the local contribution to the action, which forces a gauge field to take a more typical (smoother) value given its neighbors. As mentioned earlier, the gradient flow averages the gauge field over a neighborhood, and therefore also has a smoothing effect. Ideally, these methods work because they make local modifications, which therefore leave the global topological charge relatively intact. A delicate issue with these smoothing algorithms is that they can destroy physical instantons; in fact after protracted cooling, a lattice will eventually be brought to $Q_L = 0$. This happens because exceptional configurations or *dislocations* do not allow for a well-defined topological charge. A lattice can then change its topological charge by passing through these exceptional configurations. In practice, one cools just enough that topological observables become *quasi-stable*, i.e. just enough that they do not change after many additional cooling sweeps.

6.4 The Polyakov loop

6.4.1 The Polyakov loop and deconfinement

We begin with a brief review of some relevant properties of the Polyakov loop. For a theory with underlying $\mathrm{SU}(N_c)$ gauge group, we define the *Polyakov loop* by

$$P_{\vec{x}} \equiv \frac{1}{N_c} \operatorname{tr} \prod_{\tau} U_4(\vec{x}, \tau). \tag{6.92}$$

Because of the trace and periodic BCs, the Polyakov loop is invariant under gauge transformations. For the purpose of, for instance, data analysis, it is convenient to define the spatial average of the Polyakov loop,

$$P \equiv \frac{1}{N_\sigma^3} \sum_{\vec{x}} P_{\vec{x}}. \tag{6.93}$$

Finally, some quantities of interest to us are defined more naturally using the untraced Polyakov loop, or *thermal Wilson line*, which is just

$$L_{\vec{x}} \equiv \prod_{\tau} U_4(\vec{x}, \tau). \tag{6.94}$$

Each of the above is also sometimes referred to as a Polyakov loop, which can be confusing because the untraced Polyakov loop is not gauge invariant.

We now specialize to SU(3). The Polyakov loop relates to $F_{q\bar{q}}$, the color-averaged free energy of a static quark-antiquark pair in equilibrium at temperature T by [23; 24]

$$\exp\left[-\frac{F_{q\bar{q}}(r,T)}{T}\right] = \left\langle P_{\vec{x}} P^{\dagger}_{\vec{y}} \right\rangle, \qquad rT = |\vec{x} - \vec{y}|\,/N_\tau. \tag{6.95}$$

At low temperatures with static quarks, the Polyakov loop expectation value $\langle |P| \rangle$ is zero. This can be seen from at least two viewpoints. First note that at large r, $P_{\vec{x}}$ and $P_{\vec{y}}$ become essentially uncorrelated; therefore one finds

$$\left\langle P_{\vec{x}} P^{\dagger}_{\vec{y}} \right\rangle \approx \langle P_{\vec{x}} \rangle \left\langle P^{\dagger}_{\vec{y}} \right\rangle = \langle P \rangle \left\langle P^{\dagger} \right\rangle = \langle P \rangle \langle P \rangle^{\dagger} = |\langle P \rangle|^2, \tag{6.96}$$

and from eq. (6.95) we can (at least schematically) write something like

$$|\langle P \rangle|^2 \sim \exp\left[-\frac{F_{q\bar{q}}(\infty,T)}{T}\right]. \tag{6.97}$$

Since $F_{q\bar{q}}$ is growing linearly with separation at low temperatures due to confinement, the RHS of the above equation is zero. Another viewpoint is as follows: $\langle |P| \rangle$ is zero due to the global $\mathbb{Z}_3$ symmetry of the gauge action. In particular one can show that the plaquette is unchanged when multiplying all temporal links in a given time slice $x_4 = t_0$ with the same element $z \in \mathbb{Z}_3$, i.e. the gauge action is unchanged under

$$U_4(\vec{x}, t_0) \to z U_4(\vec{x}, t_0). \tag{6.98}$$

The Polyakov loop is in general changed by this transformation because it winds around the time direction. However since the action remains unchanged, one can write

$$\langle P \rangle = \frac{1}{3}\left\langle P + zP + z^2 P \right\rangle = \frac{1}{3}\left(1 + z + z^2\right)\langle P \rangle = 0 \tag{6.99}$$

since the sum over the center elements is zero.

At higher temperatures the first equality of eq. (6.99) no longer holds. At the deconfinement temperature T_c, this center symmetry is spontaneously broken, and $\langle |P| \rangle$ acquires a nonzero value, signalling a finite static quark-antiquark free energy and hence deconfinement. At finite N_σ, $\langle |P| \rangle$ has an inflection point at T_c, and the slope at this point diverges in the infinite volume limit. The Polyakov loop susceptibility, defined as

$$\chi_{|P|} = N_\sigma^3 \left(\left\langle |P|^2 \right\rangle - \langle |P| \rangle^2 \right). \tag{6.100}$$

exhibits a pronounced peak in a finite volume $N_\sigma^3 \times N_\tau$ at T_c. The peak height diverges in the infinite volume limit as $\chi_{|P|}^{\max} \sim N_\sigma^3$, reflecting the first order nature of the deconfinement phase transition in pure SU(3) gauge theory.

One can argue that, as far as susceptibilities go, it is the susceptibility of $\mathrm{Re}\, P$ rather than $|P|$ that more directly contains the relevant physics. This point can be seen as follows: By doing a hopping parameter expansion of the fermion determinant, which is valid in the limit of heavy quarks, and by doing a strong coupling expansion in the kinetic part, one finds the leading order contributions

$$S = S_G + S_F \approx \sum_{\vec{x},\vec{y}} \mathrm{tr}\, L_{\vec{x}}\, \mathrm{tr}\, L_{\vec{y}}^\dagger + h \sum_{\vec{x}} \left(\mathrm{tr}\, L_{\vec{x}} + \mathrm{tr}\, L_{\vec{x}}^\dagger \right), \tag{6.101}$$

where h is an effective coupling that we will treat like an external field strength. The first term in the fermionic part is coming from loops winding around the lattice in the positive temporal direction, while the second term comes from loops oriented in the negative temporal direction. With the action in this form, one can more easily see the analogy with an Ising model solid subject to an external magnetic field. The susceptibility is then by definition

$$\chi = \frac{\partial^2 \log Z}{\partial h^2}, \tag{6.102}$$

where

$$Z = \int \mathcal{D}U\, e^{-S(h;U)} \tag{6.103}$$

is the partition function. We compute

$$\frac{\partial Z}{\partial h} = \int \mathcal{D}U \sum_{\vec{x}} \left(\mathrm{tr}\, L_{\vec{x}} + \mathrm{tr}\, L_{\vec{x}}^\dagger \right) e^{-S(h;U)} = Z \left\langle \sum_{\vec{x}} \left(\mathrm{tr}\, L_{\vec{x}} + \mathrm{tr}\, L_{\vec{x}}^\dagger \right) \right\rangle \tag{6.104}$$

and similarly

$$\frac{\partial^2 Z}{\partial h^2} = Z \left\langle \left\{ \sum_{\vec{x}} \left(\operatorname{tr} L_{\vec{x}} + \operatorname{tr} L_{\vec{x}}^\dagger \right) \right\}^2 \right\rangle . \tag{6.105}$$

Hence

$$\begin{aligned} \chi &= \frac{\partial^2 \log Z}{\partial h^2} \\ &= \frac{1}{Z} \frac{\partial^2 Z}{\partial h^2} - \frac{1}{Z^2} \left(\frac{\partial Z}{\partial h} \right)^2 \\ &= \left\langle \{\}^2 \right\rangle - \langle \{\} \rangle^2 \\ &= N_\sigma^3 \left(\left\langle \operatorname{Re} P^2 \right\rangle - \langle \operatorname{Re} P \rangle^2 \right) . \end{aligned} \tag{6.106}$$

So from this point of view, the susceptibility of $\operatorname{Re} P$ is what falls out the most naturally from the effective Polyakov loop action. According to Frithjof, looking at $\chi_{|P|}$ became popular because in quenched QCD in the confined phase, one finds two sectors with nonzero $\operatorname{Im} P$, and it is easier to just measure $|P|$ than to rotate measurements in these two sectors back to the real axis.

6.4.2 Free energy and Polyakov loop renormalization

The gauge invariant color-averaged Polyakov loop correlator can be decomposed into (in general gauge-dependent) color singlet F_1 and color octet F_8 contributions [23; 24; 25]. In particular

$$\exp\left[-\frac{F_{q\bar{q}}(r,T)}{T} \right] = \frac{1}{9} \exp\left[-\frac{F_1(r,T)}{T} \right] + \frac{8}{9} \exp\left[-\frac{F_8(r,T)}{T} \right] , \tag{6.107}$$

where

$$\begin{aligned} \exp\left[-F_1(r,T)/T \right] &= \frac{1}{3} \left\langle \operatorname{tr} L_{\vec{x}} L_{\vec{y}}^\dagger \right\rangle \\ \exp\left[-F_8(r,T)/T \right] &= \frac{9}{8} \left\langle P_{\vec{x}} P_{\vec{y}}^\dagger \right\rangle - \frac{1}{24} \left\langle \operatorname{tr} L_{\vec{x}} L_{\vec{y}}^\dagger \right\rangle , \end{aligned} \tag{6.108}$$

which clearly depend on the gauge. For small distances ($rT \ll 1$ and $r \ll 1/\Lambda_{\mathrm{QCD}}$) it can be shown within zero temperature perturbation theory [26] that

$$F_1(r, T=0) \equiv V_{T=0}(r) = -8F_8(r, T=0) + \mathcal{O}\left(g^4\right) = -\frac{g^2}{3\pi r} \left(1 + \mathcal{O}\left(g^2\right) \right) . \tag{6.109}$$

The exponent in eq. (6.107) is positive for the singlet channel but negative for the octet channel in this limit, which means F_1 dominates, and we find

$$F_{q\bar{q}} - F_1 = T \log 9. \tag{6.110}$$

Therefore,

$$\lim_{r\to 0} F_{q\bar{q}} - F_1 = T \log 9 \qquad \forall\, T. \tag{6.111}$$

The Polyakov loop requires a multiplicative renormalization. When each link is renormalized with factor $Z(\beta)$, it follows from eq. (6.92) that the Polyakov loop renormalizes as

$$P^{\text{ren}} = Z(\beta)^{N_\tau} P. \tag{6.112}$$

From eq. (6.95) it is then clear that $F_{q\bar{q}}$ requires an additive renormalization. We can write

$$a\, F^{\text{ren}}_{q\bar{q}} = a\, F_{q\bar{q}} + c(\beta), \tag{6.113}$$

and then by eq. (6.111) we can similarly write

$$a\, F^{\text{ren}}_1 = a\, F_1 + c(\beta) \tag{6.114}$$

with the same $c(\beta)$ for both. From eqs. (6.95), (6.112), and (6.113) we can write

$$c(\beta) = -2 \log Z(\beta). \tag{6.115}$$

In practice, renormalized Polyakov loops can be calculated using the $q\bar{q}$-scheme [26]. The additive renormalization is determined by matching the singlet free energy to the zero temperature potential at short distances, i.e.

$$c(\beta) = a\, V_{T=0}(r_s) - a\, F_1(r_s, T), \tag{6.116}$$

where r_s is one of the shortest distances. (Often it is better to take the third or fourth shortest available distance because the shortest distance suffers the most from lattice artifacts.) The renormalized Polyakov loop is determined as follows: By eqs. (6.95) and (6.96)

$$\lim_{r\to\infty} F^{\text{ren}}_{q\bar{q}}(r, T) = -T \log |\langle P \rangle|^2 + c(\beta) T. \tag{6.117}$$

Moreover as the distance between the pair increases, this free energy should approach that of two quarks that do not interact with each other, i.e.

$$\lim_{r\to\infty} F^{\text{ren}}_{q\bar{q}}(r, T) = 2 F^{\text{ren}}_q(T), \tag{6.118}$$

where F_q is the free energy of a heavy, static quark in the thermal medium. The renormalized Polyakov loop is then recovered from its relationship to this free energy, namely

$$P^{\text{ren}} = \exp\left[-\frac{F_q^{\text{ren}}(T)}{T}\right], \tag{6.119}$$

and from the above three equations we finally find

$$P^{\text{ren}} = \exp\left[-\frac{1}{2}\left(\log|\langle P\rangle|^2 - c(\beta)\right)\right]. \tag{6.120}$$

6.4.3 Finite size scaling

In this section we discuss some finite volume effects related to the Polyakov loop. In the confined phase in pure SU(3), there are three sectors that P clusters around in the complex plane, corresponding to the three roots of unity. At finite quark mass, the real sector is preferred, which can be seen for example in Fig. 6.3. I've been told that one can show from strong coupling arguments that the real sector is preferred at finite quark mass, but don't understand that yet. Also we expect $\langle \operatorname{Im} P\rangle = 0$ at $\mu = 0$. To see this note that for any action symmetric under $U \to U^\dagger$, we have

$$\begin{aligned}
\langle \operatorname{Im} P\rangle &= \int \mathcal{D}U\, e^{-S(U)} \operatorname{Im} P(U) \\
&= \int \mathcal{D}U^\dagger\, e^{-S(U^\dagger)} \operatorname{Im} P(U^\dagger) \\
&= -\int \mathcal{D}U\, e^{-S(U)} \operatorname{Im} P(U) \\
&= -\langle \operatorname{Im} P\rangle\,,
\end{aligned} \tag{6.121}$$

where in the third step we used the invariance of the Haar measure and the action. Finite chemical potential breaks the $U \to U^\dagger$ symmetry, so we can't anticipate that $\langle \operatorname{Im} P\rangle$ will vanish anymore.

Let us discuss the large N_σ behavior of some other quantities. To simplify the notation a little, we introduce

$$x \equiv \operatorname{Re} P, \quad y \equiv \operatorname{Im} P, \quad \text{and} \quad V \equiv N_\sigma^3. \tag{6.122}$$

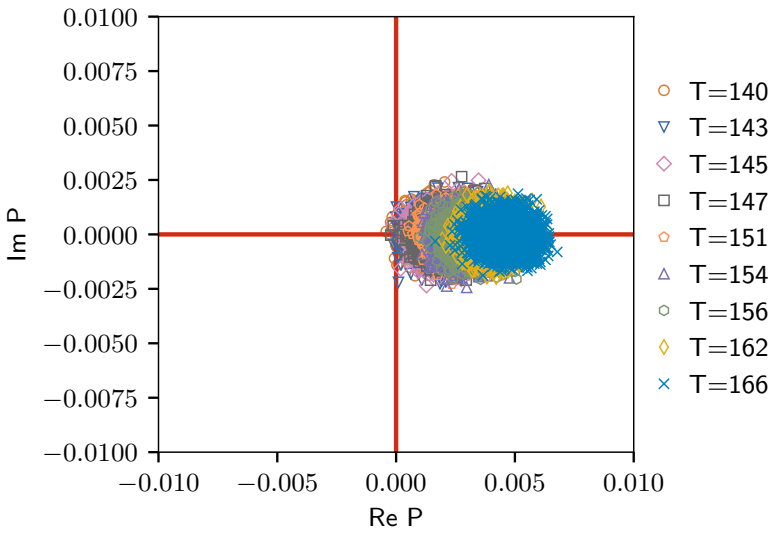

Figure 6.3: Scatter plot of P values for $56^3 \times 8$ HISQ configurations with $N_f = 2+1$ at $m_s/m_l = 80$. Temperatures are listed in MeV. The real sector is preferred at finite quark mass.

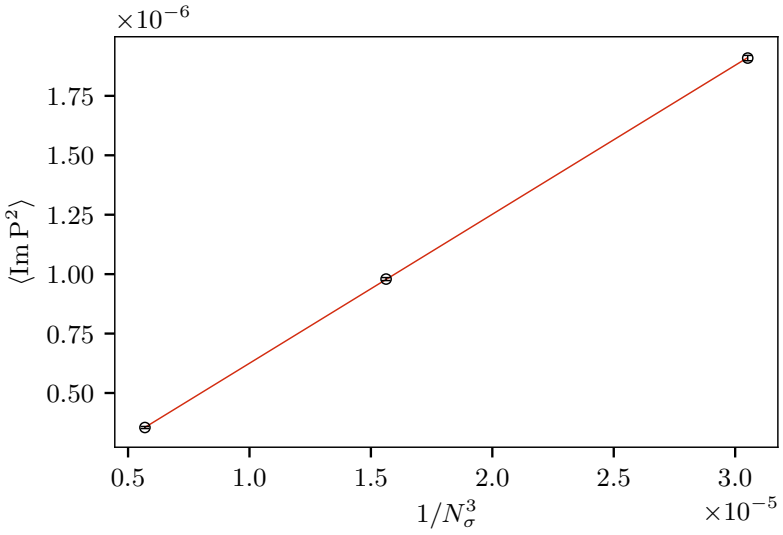

Figure 6.4: Volume dependence of the bare $\langle \mathrm{Im}\, P^2 \rangle$ on $N_\tau = 8$ lattices with $\beta = 6.445$, $m_s/m_l = 80$, and $N_\sigma = 32$, 40, and 56. The linear fit has $\chi^2/\mathrm{d.o.f.} = 0.11$.

From the above discussion we have in this notation $\langle y \rangle = 0$. Note that $\langle y^2 \rangle$ is generally not zero, although it vanishes as $1/V$ with increasing V. One can see $\langle y^2 \rangle \neq 0$ in Fig. 6.3.

In particular one expects that $\langle y^2 \rangle$ scales as $1/V$. This is because in general $\log Z$ is proportional to the free energy, so it is extensive and therefore scales as V; therefore cumulants of extensive observables, which are coefficients in the cumulant expansion, scale as V. It follows that the susceptibility of an *intensive* observable scales as $1/V$, and hence

$$\left\langle y^2 \right\rangle - \left\langle y \right\rangle^2 = \left\langle y^2 \right\rangle = \frac{A}{V} \tag{6.123}$$

for some volume-independent constant A. Indeed Fig. 6.4 shows $\langle y^2 \rangle$ scales linearly with $1/V$.

It also follows from the scaling of terms in the cumulant expansion that

$$\left\langle x^2 \right\rangle - \left\langle x \right\rangle^2 = \frac{B}{V} \quad \text{and} \quad \left\langle x \right\rangle = C \tag{6.124}$$

for some volume-independent constants B and C. By combining the two

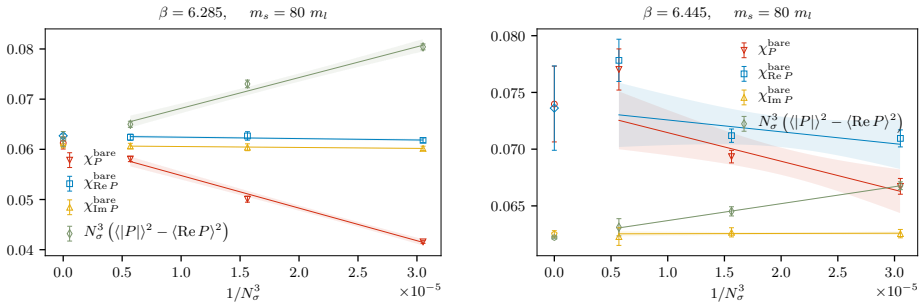

Figure 6.5: Finite size scaling of various susceptibilities for $N_\tau = 8$, $m_s/m_l = 80$ lattices for $\beta = 6.285$ (left) and $\beta = 6.445$ (right). Also shown are linear fits in $1/N_\sigma^3$ along with the infinite volume extrapolations.

equations in (6.124) one sees

$$\left\langle x^2 \right\rangle = C^2 + \frac{B}{V}. \tag{6.125}$$

In Figure 6.5 some of these relations are seen to hold on $N_f = 2 + 1$ HISQ lattices both above and below the chiral crossover. We also see in this figure that $\langle |P| \rangle$ does not scale with volume in the same way as $\langle \text{Re}\, P \rangle$ or $\langle \text{Im}\, P \rangle$. This makes sense because extensive quantities should scale linearly with the system size, but absolute values are not linear.

References

[1] M. Srednicki. *Quantum Field Theory*. Cambridge, Cambridge, 2007. ISBN 978-0-521-86449-7.

[2] I. Montvay and G. Münster. *Quantum Fields on a Lattice*. Cambridge, Cambridge, 1994.

[3] K. G. Wilson. Confinement of quarks. *Phys. Rev. D*, 10(8):2445–2459, 1974. URL https://journals.aps.org/prd/abstract/10.1103/PhysRevD.10.2445.

[4] C. G. Callan. Broken scale invariance in scalar field theory. *Phys. Rev. D*, 2(8):1541–1547, 1970. ISSN 0556-2821. doi: 10.1103/PhysRevD.2.1541. URL https://link.aps.org/doi/10.1103/PhysRevD.2.1541.

[5] K. Symanzik. Small distance behaviour in field theory and power counting. *Commun. Math. Phys.*, 18(3):227–246, 1970. ISSN 0010-3616, 1432-0916. doi: 10.1007/BF01649434. URL http://link.springer.com/10.1007/BF01649434.

[6] K. Symanzik. Small-distance-behaviour analysis and Wilson expansions. *Commun. Math. Phys.*, 23(1):49–86, 1971. ISSN 0010-3616, 1432-0916. doi: 10.1007/BF01877596. URL http://link.springer.com/10.1007/BF01877596.

[7] D. J. Gross and F. Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30(26):1343–1346, 1973. URL https://journals.aps.org/prd/abstract/10.1103/PhysRevD.8.3497.

[8] H. D. Politzer. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.*, 30(26):1346–1349, 1973. URL https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.30.1346.

[9] A. A. Belavin and A. A. Migdal. Calculation of anomalous dimensionalities in non-Abelian gauge field theories. *JETP Lett.*, 19(5):181–182, 1974.

[10] W. E. Caswell. Asymptotic behavior of non-abelian gauge theories to two-loop order. *Phys. Rev. Lett.*, 33(4):244–246, 1974. URL https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.33.244.

[11] D. R. T. Jones. Two-loop diagrams in Yang-Mills theory. *Nucl. Phys. B*, 75(3):531–538, 1974. URL http://www.sciencedirect.com/science/article/pii/0550321374900935.

[12] B. Allés, A. Feo, and H. Panagopoulos. The three-loop β function in SU(N) lattice gauge theories. *Nucl. Phys. B*, 491(1-2):498–512, 1997. URL http://www.sciencedirect.com/science/article/pii/S0550321397000928.

[13] J. Engels, S. Mashkevich, T. Scheideler, and G. Zinovjev. Critical behaviour of SU(2) lattice gauge theory. A complete analysis with the χ^2-method. *Phys. Lett. B*, 365(1-4):219–224, 1996. URL http://www.sciencedirect.com/science/article/pii/037026939501280X.

[14] M. Lüscher. Properties and uses of the Wilson flow in lattice QCD. *J. High Energy Phys.*, 2010(8):071, 2010. ISSN 1029-8479. doi: 10.1007/JHEP08(2010)071. URL http://link.springer.com/10.1007/JHEP08(2010)071.

[15] B. A. Berg. Dislocations and topological background in the lattice O(3) sigma model. *Phys. Lett. B*, 104(6):475–480, 1981. doi: 10.1016/0370-2693(81)90518-9.

[16] C. Bonati and M. D'Elia. Comparison of the gradient flow with cooling in SU(3) pure gauge theory. *Phys. Rev. D*, 89(10):105005, 2014. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.89.105005. URL http://link.aps.org/doi/10.1103/PhysRevD.89.105005.

[17] C. R. Allton. Lattice Monte Carlo data versus perturbation theory. *Nucl. Phys. B (Proc. Suppl.)*, 53(1-3):867–869, 1997.

[18] A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Y. S. Tyupkin. Pseudoparticle solutions of the Yang-Mills equations. *Phys. Lett.*, 59(1):85–87, 1975.

[19] E. Witten. Current algebra theorems for the U(1) "Goldstone boson". *Nucl. Phys. B*, 156(2):269–283, 1979. ISSN 05503213. doi: 10.1016/0550-3213(79)90031-2. URL http://linkinghub.elsevier.com/retrieve/pii/0550321379900312.

[20] G. Veneziano. U(1) without instantons. *Nucl. Phys. B*, 159(1-2):213–224, 1979. ISSN 05503213. doi: 10.1016/0550-3213(79)90332-8. URL http://linkinghub.elsevier.com/retrieve/pii/0550321379903328.

[21] M. Lüscher. Topology of lattice gauge fields. *Commun. Math. Phys.*, 85:39–48, 1982.

[22] A. S. Kronfeld. Topological aspects of lattice gauge theories. *Nucl. Phys. B (Proc. Suppl.)*, 4:329–351, 1988. ISSN 09205632. doi: 10.1016/0920-5632(88)90123-5. URL http://linkinghub.elsevier.com/retrieve/pii/0920563288901235.

[23] Larry D. McLerran and Benjamin Svetitsky. A Monte Carlo study of SU(2) Yang-Mills theory at finite temperature. *Phys. Lett. B*, 98(3):195–198, January 1981. ISSN 03702693. doi: 10.1016/0370-2693(81)90986-2.

[24] Larry D. McLerran and Benjamin Svetitsky. Quark liberation at high temperature: A Monte Carlo study of SU(2) gauge theory. *Phys. Rev. D*, 24(2):450–460, July 1981. ISSN 0556-2821. doi: 10.1103/PhysRevD.24.450.

[25] Sudhir Nadkarni. Non-Abelian Debye screening: The color-averaged potential. *Physical Review D*, 33(12):3738–3746, June 1986. ISSN 0556-2821. doi: 10.1103/PhysRevD.33.3738.

[26] O Kaczmarek, F Karsch, P Petreczky, and F Zantow. Heavy quark–antiquark free energy and the renormalized Polyakov loop. *Phys. Lett. B*, 543(1-2):41–47, September 2002. ISSN 03702693. doi: 10.1016/S0370-2693(02)02415-2.

Chapter 7

LGT: MCMC Simulations

As discussed in Section 6.1.4, expectation values of physical observables X in pure SU(2) LGT are given by functional integrals

$$\langle X \rangle = \frac{1}{Z} \int \mathcal{D}U \, e^{-S(U)} X(U). \tag{7.1}$$

Even though the integral (7.1) is well-defined on a lattice because there are finitely many sites, it is not feasible to evaluate it numerically; even relatively small lattices have 4×10^4 links. The goal of an MCMC simulation is to estimate $\langle X \rangle$ by randomly generating configurations, distributed with probability e^{-S}, and on each configuration, making a measurement X_i. The average

$$\bar{X} = \frac{1}{N_{\text{conf}}} \sum_{i=1}^{N_{\text{conf}}} X_i \tag{7.2}$$

serves as the estimator.

In Section 7.1 we introduce MCMC simulations as they are applied to the project. Section 7.2 summarizes some of the tools needed to statistically analyze the generated data. The final Section 7.3 provides details of how our simulation is implemented on the computer. Further details can be found in, for instance, Berg [1].

7.1 Markov chain Monte Carlo

To generate our configurations, we start from some arbitrary configuration C_0 and construct a stochastic sequence of configurations. Configuration C_i

is generated based on configuration C_{i-1}, which we call an *update* or *Monte Carlo step*. The result is a *Markov chain*

$$C_0 \to C_1 \to C_2 \to ... \tag{7.3}$$

of configurations.

Markov chain Monte Carlo (MCMC) is characterized by the probability $W^{CC'} \equiv \mathrm{P}\,(C'|C)$, the probability to jump to configuration C' given that the system started in configuration C. The MCMC *transition matrix*

$$W \equiv \left(W^{CC'}\right) \tag{7.4}$$

is constructed to bring the system to *equilibrium*. In equilibrium, the chain should have no sinks or sources of probability, which means that the probability of jumping into a configuration C' should be the same as jumping out of C'. This property is called *balance*

$$\sum_C W^{CC'}\,\mathrm{P}\,(C) = \sum_C W^{C'C}\,\mathrm{P}\,(C')\,, \tag{7.5}$$

with the LHS representing the total probability to end up in C' and the RHS representing the probability to transition out of C'. If W satisfies

1. *ergodicity*, i.e.

$$\mathrm{P}\,(C) > 0 \;\text{ and }\; \mathrm{P}\,(C') > 0 \;\;\Rightarrow\;\; \exists\, n \in \mathbb{N} \;\text{ s.t. }\; \left(W^n\right)^{CC'} > 0; \tag{7.6}$$

2. *normalization*, i.e.

$$\sum_{C'} W^{CC'} = 1; \tag{7.7}$$

3. and balance,

then the Markov process is guaranteed to bring the ensemble toward equilibrium. Using normalization, one finds from eq. (7.5)

$$\sum_C W^{CC'}\,\mathrm{P}\,(C) = \mathrm{P}\,(C')\,, \tag{7.8}$$

which shows that the equilibrium distribution is a fixed point of the Markov chain. The first property, ergodicity, guarantees that it is possible to transition from C to C' in a finite number of steps. In realistic simulations, it is important that the n appearing in eq. (7.6) is not too large. For example the Markov chain may have difficulty connecting different topological sectors in configuration space.

7.1.1 Update: Metropolis and heat bath

In this and the following subsection, we omit the Lorentz index and space-time point from link variables to avoid clutter. We use U to indicate the link to be updated, $U^{\sqcup}$ to indicate the staple matrix attached to U, and U' to indicate a trial link. We will use the Boltzmann distribution $\mathrm{P}(C) \propto e^{-S_C}$.

One trivial way to satisfy the balance condition (7.5) is to find an update that satisfies it term-by-term. For such an update,

$$W^{CC'}\,\mathrm{P}(C) = W^{C'C}\,\mathrm{P}(C')\,. \tag{7.9}$$

This property is known as *detailed balance*. One of the most well-known Monte Carlo updates that satisfies detailed balance is the *Metropolis algorithm* [2]. In the Metropolis algorithm, a trial configuration C' is selected with some probability distribution $\mathrm{T}(C'|C)$. Then C' is accepted with likelihood

$$\mathrm{P}(C \to C') = \min\left[1, \frac{\mathrm{T}(C|C')\,e^{-S_{C'}}}{\mathrm{T}(C'|C)\,e^{-S_C}}\right], \tag{7.10}$$

where S_C is the action corresponding to C. If C' is rejected, the unchanged configuration is counted in the Markov chain. Using the fact that the total probability to transition from C to C' is $W^{CC'} = \mathrm{T}(C'|C)\,\mathrm{P}(C \to C')$, one can show that this update satisfies detailed balance.

Another update is the *heat bath* (HB). In our simulations, a new configuration is generated from an old one by updating one link. For the SU(2) HB algorithm, the trial link distribution is

$$\mathrm{d}\,\mathrm{T}(U') \propto dU' \exp\left(\frac{\beta}{2}\,\mathrm{tr}\;U'U^{\sqcup}\right) \tag{7.11}$$

and the transition probability is

$$\mathrm{P}(C \to C') = \min\left[1, e^{-(S_{C'}-S_C)}\right]. \tag{7.12}$$

This construction also satisfies detailed balance. The new configuration is automatically accepted whenever it lowers the action, and increases in the action are exponentially suppressed. HB updates ensure local equilibrium, but they often take more CPU time. For SU(2) the guarantee of local equilibrium turns out to be more impactful, so heat bath updates are more efficient than general Metropolis updates.

Single link Metropolis or HB updates of links carried out in a systematic (as opposed to random) order fulfill balance, but do not fulfill detailed balance.

7.1.2 Update: Over-relaxation

An additional useful update for SU(2) is the *over-relaxation* (OR) update. Adler introduced OR algorithms [3] and they were further developed by Creutz [4] and others. The idea of the OR algorithm is to speed up relaxation by generating a group element "far away" from U without destroying equilibrium, which is here achieved by keeping the action constant.

More precisely let $U \in \mathrm{SU}(N_c)$ and suppose we have some method of choosing another link variable U_0 that maximizes the action for this staple. We assume that this method of selection has no dependence on U. Pick some element $V \in \mathrm{SU}(N_c)$ such that $U = VU_0$; viewed in this way, U is "on one side of U_0," and the element "on the other side" is $U' = V^{-1}U_0$. Note that

$$V = UU_0^{-1}, \tag{7.13}$$

which implies

$$U' = U_0U^{-1}U_0. \tag{7.14}$$

This manner of constructing a new link variable U', which generates a group element "far away" from U without changing the action, is what we mean by over-relaxation.

In principle an OR update should be more efficient than a Monte Carlo update. This is because we chose the new link variable to be two group elements away from the old one, thrusting us further along configuration space. However unlike Metropolis updates, OR updates only sample the subspace of constant action, and are therefore not ergodic. Hence to ensure an approach to equilibrium, they must be supplemented with, for instance, HB updates.

We implement the SU(2) OR update by

$$U \to U' = \frac{1}{\det U^{\sqcup}} \left(U^{\sqcup}UU^{\sqcup}\right)^{\dagger}. \tag{7.15}$$

Proposition 7.1.1

This update does not change the SU(2) Wilson action.

Proof. Since $\det(kA) = k^n \det(A)$ for any constant k and $n \times n$ matrix A, one can show that the sum of two SU(2) matrices is proportional

to an SU(2) matrix. Hence we can write

$$U^{\sqcup} = u^{\sqcup}\sqrt{\det U^{\sqcup}}$$

where $u^{\sqcup} \in \mathrm{SU}(2)$. After updating, the local contribution to the Wilson action becomes

$$\operatorname{tr} U'U^{\sqcup} = \frac{1}{\det U^{\sqcup}} \operatorname{tr} \left(U^{\sqcup}UU^{\sqcup}\right)^{\dagger} = \operatorname{tr} U^{\sqcup}U \left(u^{\sqcup}\right)^{\dagger} u^{\sqcup} = \operatorname{tr} U^{\sqcup}U,$$

which is what it was originally. □

Since the action is unchanged, the proposal is always accepted. This simple behavior is special to U(1) and SU(2) LGT. Its usefulness is extended to SU(N_c) when $N_c > 2$ via the method of Cabibbo and Marinari [5].

7.2 Statistical analysis

Since C_i is generated based on C_{i-1}, measurements on subsequent configurations are correlated. In our simulations, these correlations are reduced in two ways:

1. Subsequent configurations are separated by multiple updating sweeps; and then

2. configurations are grouped into N_{conf} blocks or bins.

The final measurements X_i used in data analysis are obtained by averaging within each block. To check whether the final data are effectively independent, one can use the integrated autocorrelation time. For statistically independent measurements, we expect the variance $\sigma^2_{\bar{X}}$ of $\bar{X}$ to be

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{N_{\mathrm{conf}}} \tag{7.16}$$

due to the CLT. In practice, however, one finds

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{N_{\mathrm{conf}}}\tau_{\mathrm{int}}. \tag{7.17}$$

The factor τ_{int} is the integrated autocorrelation time. It is the ratio between the estimated variance of the sample mean and what this variance would have been if the data were independent. For effectively independent data, $\tau_{\text{int}} = 1$.

So, the final measurements are drawn from some distribution with mean $\langle X \rangle$ and variance σ^2 and are effectively independent. The estimator $\bar{X}$ of the mean is the average (7.2), while the unbiased estimator $\bar{\sigma}^2$ of the variance is

$$\bar{\sigma}^2 = \frac{1}{N_{\text{conf}} - 1} \sum_{i=1}^{N_{\text{conf}}} \left(X_i - \bar{X} \right)^2 . \tag{7.18}$$

An estimator is biased if its mean for finite N_{conf} does not agree with the exact result; the bias is the difference. Generally, problems with bias emerge whenever one wishes to estimate some non-linear function f of the mean $\langle X \rangle$. Naively one might guess

$$\bar{f}_{\text{bad}} = \frac{1}{N_{\text{conf}}} \sum_{i=1}^{N_{\text{conf}}} f(X_i) \tag{7.19}$$

as an estimator; however it can be shown that the bias of $\bar{f}_{\text{bad}}$ is $\mathcal{O}(1)$, i.e. it never converges to the exact result. An estimator for $f(\langle X \rangle)$ that converges to its true value is

$$\bar{f} = f(\bar{X}); \tag{7.20}$$

in particular, the bias of this estimator is $\mathcal{O}(1/N_{\text{conf}})$. Therefore in the large N_{conf} limit, the bias vanishes faster than the statistical error bar.

We have introduced a way to estimate the mean and variance of some operator, as well as a way to estimate the mean of some function of that operator. Now we need a way to estimate the error bar of that function. We cannot use

$$\bar{\sigma}^2_{\bar{f}} = \frac{\bar{\sigma}^2_{\bar{f}}}{N_{\text{conf}}} = \frac{1}{N_{\text{conf}} \left(N_{\text{conf}} - 1 \right)} \sum_{i=1}^{N_{\text{conf}}} \left(f(X_i) - \bar{f} \right)^2 \tag{7.21}$$

because $f(X_i)$ is not a valid sample point. One could analytically produce an error bar for $\bar{f}$ using error propagation. However when the function is complicated, error propagation becomes extremely unwieldy.

Jackknifing allows one to extract a mean and error bar, and it is straightforward to implement; therefore it makes sense to use the jackknife method

generally. The idea of jackknifing is to throw away the first measurement, leaving $N_{\text{conf}}-1$ resampled values. Then we resample again, this time throwing out the second point, and so on. The resulting jackknife bins are

$$X_{J,i} = \frac{1}{N_{\text{conf}}-1}\sum_{j\neq i} X_j. \tag{7.22}$$

The jackknife estimator for $f(\langle x\rangle)$ is then

$$\bar{f}_J = \frac{1}{N_{\text{conf}}}\sum_{i=1}^{N_{\text{conf}}} f(X_{J,i}), \tag{7.23}$$

while the estimator for the variance of $\bar{f}_J$ is

$$\bar{\sigma}^2_{f_J} = \frac{N_{\text{conf}}-1}{N_{\text{conf}}}\sum_{i=1}^{N_{\text{conf}}}\left(f(X_{J,i})-\bar{f}_J\right)^2. \tag{7.24}$$

In many instances, we will need to compare two estimates of the same quantity against each other and decide whether the difference between them is significant. This can happen, for example, if we want to compare another group's results with our own. Let their result be $\bar{X}$ with uncertainty $\sigma_{\bar{X}}$ and ours be $\bar{Y}$ with uncertainty $\sigma_{\bar{Y}}$. Then the probability that these two estimates differ by at least D is

$$q = \mathrm{P}\left(|\bar{X}-\bar{Y}| > D\right) = 1 - \mathrm{erf}\left(\frac{D}{\sqrt{2\left(\sigma^2_{\bar{X}}+\sigma^2_{\bar{Y}}\right)}}\right) \tag{7.25}$$

assuming $\bar{X}$ and $\bar{Y}$ are normally distributed with the same mean. This is called a Gaussian difference test. The quantity q is called the q-value. In practice we take $q \leq 0.05$ to be an indication of a possible discrepancy between $\bar{X}$ and $\bar{Y}$, keeping in mind that $q \leq 0.05$ by chance one out of twenty times.

In practice, the true variances $\sigma_{\bar{X}}$ and $\sigma_{\bar{Y}}$ are not known. If one wishes to use the estimators $\bar{\sigma}_{\bar{X}}$ and $\bar{\sigma}_{\bar{Y}}$ instead, one can perform a *Student difference test* or *t-test* to investigate whether the discrepancy D is due to chance. Suppose the estimate $\bar{X}$ comes from M_{conf} data, while $\bar{Y}$ comes from N_{conf}

data. Assume $\sigma_{\bar{X}} = \sigma_{\bar{Y}}$, which happens when the sampling methods used are identical. We introduce the random variable

$$t = \frac{D}{\bar{\sigma}_D}, \tag{7.26}$$

where $D = \bar{X} - \bar{Y}$, and

$$\bar{\sigma}_D^2 = \left(\frac{1}{M_{\text{conf}}} + \frac{1}{N_{\text{conf}}}\right) \frac{(M_{\text{conf}} - 1)\,\bar{\sigma}_{\bar{X}}^2 + (N_{\text{conf}} - 1)\,\bar{\sigma}_{\bar{Y}}^2}{M_{\text{conf}} + N_{\text{conf}} - 2}. \tag{7.27}$$

Then the probability that these two estimates differ by at least D is

$$q = 2 \begin{cases} I\left(z, \frac{\nu}{2}, \frac{1}{2}\right) & \text{for } t \leq 0, \\ 1 - \frac{1}{2} I\left(z, \frac{\nu}{2}, \frac{1}{2}\right) & \text{otherwise,} \end{cases} \tag{7.28}$$

where I is the incomplete beta function, $\nu = M_{\text{conf}} + N_{\text{conf}} - 2$, and

$$z = \frac{\nu}{\nu + t^2}. \tag{7.29}$$

To estimate finite size corrections and carry out continuum limit extrapolations, we need a way to fit data to curves. Consider a sample of N_{sim} Gaussian, independent data points (X_i, Y_i), where the Y_i have standard deviations σ_i and the X_i have no errors. For instance, if one is interested in a continuum limit extrapolation, the X_i are β values while the Y_i are ratios of scales evaluated at that β. We model these data with a fit that depends on some set of M parameters

$$y = y(x; a), \tag{7.30}$$

where $a = (a_1, ..., a_M)$ is the vector of these parameters. Our goal is to estimate the a_j. Assuming that $y(x; a)$ is the exact law for the data, the probability distribution for the measurements Y_i is

$$f(y_1, ..., y_{N_{\text{sim}}}) = \prod_{i=1}^{N_{\text{sim}}} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[\frac{-(y_i - y(x_i; a))^2}{2\sigma_i^2}\right]. \tag{7.31}$$

The probability that the data fall within a region near what was observed is

$$\mathrm{P} = \prod_{i=1}^{N_{\text{sim}}} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[\frac{-(y_i - y(x_i; a))^2}{2\sigma_i^2}\right] dy_i. \tag{7.32}$$

Our strategy for determining the correct fit will be to find the vector a that maximizes the above probability. This happens when

$$\chi^2(a) \equiv \sum_{i=1}^{N_{\text{sim}}} \frac{(y_i - y(x_i; a))^2}{2\sigma_i^2} \tag{7.33}$$

is minimized. This strategy is an example of a *maximum likelihood method*.

We now describe an iterative method to search for the minimum of χ^2. Let a_n be the vector of parameters for the n^{th} iteration. As long as a is in a small enough neighborhood of a_n, we can safely approximate

$$\chi^2(a) \approx \chi^2(a_n) + (a - a_n) \cdot b + \frac{1}{2}(a - a_n)\, A\, (a - a_n), \tag{7.34}$$

where the coefficients of the vector b and the $M \times M$ matrix A are given by the first and second derivatives of χ^2 evaluated at a_n. In the *Newton-Raphson method*, the next iteration a_{n+1} is determined from the condition $\nabla\chi^2(a)|_{a=a_{n+1}} = 0$, which yields

$$a_{n+1} = a_n - A^{-1}b. \tag{7.35}$$

If the approximation (7.34) is not good, one can instead move a small step in the direction of the gradient by

$$a_{n+1} = a_n - c\, b, \tag{7.36}$$

where c is a constant that is small enough not to overshoot direction of steepest descent. This is an example of a *steepest descent method*. The Levenberg-Marquardt method [6; 7], which is our method of choice, varies smoothly between (7.35) and (7.36). Steepest descent is used far from the minimum, and then it switches to the Newton-Raphson method when the minimum is approached.

7.3 Computer implementation

Now that we have introduced the general idea of MCMC, along with some specific updating schemes, and complications for statistical analysis, we are ready to discuss the computer implementation.

As mentioned earlier, we design the simulation using *local updates*, which means we update the links one at a time. This is done in a systematic order,

because there is some computational advantage compared to updating in a random order [1]. An updating *sweep* updates every link on the lattice once. To maximize efficiency while maintaining ergodicity, our updating sweeps have a combination of HB and OR updating. We call this a *Monte Carlo Over-relaxation* (MCOR) sweep.

An MCMC simulation of LGT broadly consists of three essential steps:

1. *Initialization*: The first thing to do is get everything ready for the simulation. This includes initializing the random number generator, and setting up an initial configuration.

2. *Equilibration*: To avoid over-sampling rare configurations, one must perform many sweeps to bring the system to its equilibrium distribution. The structure of this section looks like

```
do from n=1 to n=nequi
  call MCOR update
end do
```

3. *Measurements*: All observables of interest are measured on the equilibrated configurations. To help reduce correlations between measurements, multiple updating sweeps are performed in between. This section is structured as

```
do from n=1 to n=nmeasurements
  do from n=1 to n=ndiscarded
    call MCOR update
  end do
  take measurement
end do
```

For simulations like what I did in grad school, it may take months (or years!) for a single-processor MCMC simulation to generate enough data to get reasonable error bars. Therefore it is advantageous to divide the lattice into smaller sublattices, updating simultaneously on each sublattice, passing relevant information between the sublattices whenever necessary. *Parallelizing* in this way offers a speed up factor somewhat less than the number of

sublattices used. A standard way to parallelize code is to use the Message Passing Interface (MPI). MPI allows for efficient exchange of information between processors and is easily included in Fortran or C programs.

The goal of some simulations is to determine phase transition points. Close to these points, on a finite lattice, the susceptibility of the relevant order parameter attains its maximum. The most straightforward strategy of estimating this maximum is to run multiple simulations in the vicinity of the transition point. Because this strategy requires multiple runs, it is inefficient. *Reweighting* (see [8] and references therein) is an efficient alternative. Consider the expectation value of an observable X calculated at β'. We have

$$\begin{aligned} \langle X \rangle_{\beta'} &= Z_{\beta'}^{-1} \int \mathrm{d}\phi \, e^{-\beta' E(\phi)} X(\phi) e^{(\beta-\beta)E(\phi)} \\ &= Z_{\beta'}^{-1} \int \mathrm{d}\phi \, e^{(\beta-\beta')E(\phi)} X(\phi) e^{-\beta E(\phi)} \\ &= Z_{\beta'}^{-1} Z_\beta \left\langle e^{(\beta-\beta')E} X \right\rangle_\beta \\ &= \left\langle \frac{Z_\beta}{Z_{\beta'}} e^{(\beta-\beta')E} X \right\rangle_\beta . \end{aligned} \tag{7.37}$$

We can calculate the expectation value in the last line using data from a time series generated at β, and this gives us an estimate for $\langle X \rangle_{\beta'}$. Reweighting is only useful when $E\Delta\beta = \mathcal{O}(1)$. Provided that the critical parameter β_c is sufficiently close to the simulation point β, it suffices to have only one simulation, estimating the maximum by reweighting to multiple nearby β'.

References

[1] B. A. Berg. *Markov Chain Monte Carlo Simulations and Their Statistical Analysis*. World Scientific, Singapore, 2004.

[2] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953. ISSN 0021-9606, 1089-7690. doi: 10.1063/1.1699114. URL http://aip.scitation.org/doi/10.1063/1.1699114.

[3] S. L. Adler. Over-relaxation method for the Monte Carlo evaluation of

the partition function for multiquadratic actions. *Phys. Rev. D*, 23(12): 2901–2904, 1981.

[4] M. Creutz. Overrelaxation and Monte Carlo simulation. *Phys. Rev. D*, 36(2):515–519, 1987.

[5] N. Cabibbo and E. Marinari. A new method for updating SU(N) matrices in computer simulations of gauge theories. *Phys. Lett. B*, 119(4-6):387–390, 1982.

[6] K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quart. Appl. Math.*, 2(2):164–168, 1944. ISSN 0033-569X, 1552-4485. doi: 10.1090/qam/10666. URL http://www.ams.org/qam/1944-02-02/S0033-569X-1944-10666-0/.

[7] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math.*, 11(2):431–441, 1963. ISSN 0368-4245, 2168-3484. doi: 10.1137/0111030. URL http://epubs.siam.org/doi/10.1137/0111030.

[8] A. M. Ferrenberg and R. H. Swendsen. New Monte Carlo technique for studying phase transitions. *Phys. Rev. Lett.*, 63:1658, 1989.

Chapter 8

LFT: Fermions

We will now introduce Fermions on the lattice. This presentation follows Chapter 5 of [1]. We will start in the continuum and introduce a naive discretization. Then we will go into detail, point out a problem with the naive discretization, and fix it. We will omit space-time dependence sometimes for the sake of brevity, but it should be clear which objects have space-time dependence anyway. Primarily we are interested in studying QCD, so $N_c = 3$. This chapter uses Dirac algebra extensively, but the algebra is different when the metric is Euclidean. For details see Appendix A.

In the continuum theory, the free fermion propagator is

$$S_F = \int \mathrm{d}^4x\, \bar{\psi}(\not{\partial} + m)\psi. \tag{8.1}$$

Using the rules (6.11) and (6.12), one naively discretizes this action on the lattice as

$$S_F = a^4 \sum_x \bar{\psi}(x) \left(\sum_{\mu=1}^{4} \gamma_\mu \frac{\psi(x + a\hat{\mu}) - \psi(x - a\hat{\mu})}{2a} + m\psi(x) \right). \tag{8.2}$$

The fermionic partition function is

$$Z_F = \int \mathcal{D}\psi\, \mathcal{D}\bar{\psi}\, e^{-S_F}, \tag{8.3}$$

with integration measure

$$\int \mathcal{D}\psi\, \mathcal{D}\bar{\psi} = \prod_{x,f,\alpha,c} \mathrm{d}\bar{\psi}^f_{\alpha c}(x)\, \mathrm{d}\psi^f_{\alpha c}(x)\,, \tag{8.4}$$

where f runs over flavors, α runs over Dirac indices, and c runs over colors.

In addition, we want the fermionic action to be gauge invariant. The gauge transformation for fermion fields

$$\psi(x) \to U(x)\psi(x), \qquad \bar{\psi}(x) \to \bar{\psi}(x)U(x)^\dagger \tag{8.5}$$

with $U \in \mathrm{SU}(3)$ reminds us of the gauge transformation for scalar fields given in Chapter 6. Working in the continuum we introduce, just as before, a gauge field $A_\mu(x)$ so that the fermionic action becomes

$$S_F = \int \mathrm{d}^4x\, \bar{\psi}\big(\gamma_\mu(\partial_\mu + A_\mu) + m\big)\psi. \tag{8.6}$$

The gauge-transformed fermionic action is

$$S'_F = \int \mathrm{d}^4x\, \bar{\psi}\, U^\dagger\big(\gamma_\mu(\partial_\mu + A'_\mu) + m\big)U\psi, \tag{8.7}$$

so that the requirement of gauge invariance leads us to conclude

$$(\partial_\mu + A_\mu)\psi = U^\dagger(\partial_\mu + A'_\mu)U\psi. \tag{8.8}$$

Solving for A'_μ we find

$$A'_\mu = (\partial_\mu U)U^\dagger - UA_\mu U^\dagger, \tag{8.9}$$

just as we did before.

On the lattice, link variables transform as

$$U_\mu(x) \to U'_\mu(x) = U(x)U_\mu(x)U^\dagger(x + a\hat{\mu}), \tag{8.10}$$

which ensures that closed loops of link variables are gauge invariant. In addition the discretized fermionic action becomes gauge invariant when it is written as

$$S_F = a^4 \sum_x \bar{\psi}(x) \left(\sum_{\mu=1}^{4} \gamma_\mu \frac{U_\mu(x)\psi(x + a\hat{\mu}) - U_{-\mu}(x)\psi(x - a\hat{\mu})}{2a} + m\psi(x) \right), \tag{8.11}$$

where

$$U_{-\mu}(x) = U_\mu(x - a\hat{\mu})^\dagger \tag{8.12}$$

transforms as

$$U_{-\mu}(x) \to U'_{-\mu}(x) = U(x)U_{-\mu}(x)U^\dagger(x - a\hat{\mu}). \tag{8.13}$$

Since eq. (8.11) is the gauge invariant action, this will be what we investigate for fermion doubling.

8.1 Grassmann numbers

Fermions are objects that by definition anticommute with each other. With this in mind, we introduce Grassmann numbers. We consider a set of numbers η_i, $1 \leq i \leq N$ obeying

$$\eta_i \eta_j = -\eta_j \eta_i \tag{8.14}$$

for all i and j. These are called *Grassmann numbers*. It follows that

$$\eta_i^2 = 0. \tag{8.15}$$

If a power series of a function f of the Grassmann numbers exists, then eq. (8.15) guarantees that the series terminates almost immediately. In general we would write

$$f(\eta) = a + \sum_i a_i \eta_i + \sum_{i<j} a_{ij} \eta_i \eta_j + ... + a_{1...N} \eta_1 ... \eta_N \tag{8.16}$$

with a, a_i, a_{ij}, ..., $a_{1...N} \in \mathbb{C}$. We refer to eq. (8.16) as a *Grassmann polynomial*. Grassmann polynomials are closed under addition and multiplication, and are thus said to form a *Grassmann algebra*. The η_i are the *generators* of the Grassmann algebra.

To learn how differentiation ought to work, we use the simple example $N = 2$. Then

$$f(\eta) = a + a_1 \eta_1 + a_{12} \eta_1 \eta_2. \tag{8.17}$$

A definition of the derivative that follows our intuition is

$$\frac{\partial f}{\partial \eta_1} = a_1 + a_{12} \eta_2. \tag{8.18}$$

However, the defining characteristic of Grassmann numbers is that they anticommute. Therefore we could also have expanded f as

$$f(\eta) = a + a_1 \eta_1 - a_{12} \eta_2 \eta_1. \tag{8.19}$$

In order for the derivative f to make sense we must therefore require

$$\frac{\partial}{\partial \eta_1} \eta_2 = -\eta_2 \frac{\partial}{\partial \eta_1}. \tag{8.20}$$

Similarly if we take another derivative, this time with respect to η_2, we find that partial derivatives with respect to different Grassmann variables must

also anticommute to maintain consistency. Altogether, the differentiation rules for Grassmann variables become

$$\begin{aligned} \frac{\partial}{\partial\eta_i}1 &= 0; \\ \frac{\partial}{\partial\eta_i}\eta_i &= 1; \\ \frac{\partial}{\partial\eta_i}\eta_j &= -\eta_j\frac{\partial}{\partial\eta_i}; \\ \frac{\partial^2}{\partial\eta_i\partial\eta_j} &= -\frac{\partial^2}{\partial\eta_j\partial\eta_i}. \end{aligned} \tag{8.21}$$

Next we move on to integration. We will construct integrals that work like integrals over subsets $\Omega \in \mathbb{R}^N$ for which the integrand vanishes at the boundary $\partial\Omega$. Thus we demand that the Grassmann integral is a complex number,

$$\int \mathrm{d}^N\eta\, f \in \mathbb{C}; \tag{8.22}$$

that with $\lambda_1, \lambda_2 \in \mathbb{C}$ it is linear,

$$\int \mathrm{d}^N\eta\,(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 \int \mathrm{d}^N\eta\, f_1 + \lambda_2 \int \mathrm{d}^N\eta\, f_2; \tag{8.23}$$

and that its integrand vanishes at the boundary,

$$\int \mathrm{d}^N\eta\, \frac{\partial}{\partial\eta_i} f = 0. \tag{8.24}$$

If some function f of $N-1$ Grassmann numbers can be written as the derivative of another function g of N Grassmann numbers, it follows from eq. (8.24) that the integrand vanishes. An integral of a Grassmann polynomial of N variables is therefore proportional to the coefficient $a_{1\ldots N}$, since $\eta_1 \ldots \eta_N$ cannot be written as a derivative of N variables. In particular if we demand a normalization

$$\int \mathrm{d}^N\eta\, \eta_1 \ldots \eta_N = 1, \tag{8.25}$$

we obtain the rule

$$\int \mathrm{d}^N\eta\, f = a_{1\ldots N}. \tag{8.26}$$

Finally to make Grassmann integration more analogous to the integration that we're used to, we define

$$\mathrm{d}^N\eta \equiv \mathrm{d}\eta_N \ldots \mathrm{d}\eta_1 \,. \tag{8.27}$$

Altogether then, the rules of Grassmann integration can be summarized by eq. (8.23) along with

$$\begin{aligned} \int \mathrm{d}\eta_i \, 1 &= 0; \\ \int \mathrm{d}\eta_i \, \eta_i &= 1; \\ \mathrm{d}\eta_i \, \mathrm{d}\eta_j &= -\,\mathrm{d}\eta_j \, \mathrm{d}\eta_i \,. \end{aligned} \tag{8.28}$$

Using these properties we can define integration over subsets of Grassmann variables. Interestingly, the measures obey the same algebraic properties as the derivatives.

Next let us discuss how Grassmann integrals behave under linear transformations. In general we can write such a change of variables as

$$\eta'_i = M_{ij}\eta_j, \tag{8.29}$$

where M is a complex $N \times N$ matrix. More succinctly, one writes $\eta' = M\,\eta$. Applying this change of variables to the normalization eq. (8.25) we obtain

$$\begin{aligned} \int \mathrm{d}^N\eta \, \eta_1 \ldots \eta_N &= 1 \\ &= \int \mathrm{d}^N\eta' \, \eta'_1 \ldots \eta'_N \\ &= \int \mathrm{d}^N\eta' \, M_{1\,i_1} \ldots M_{N\,i_N} \, \eta_{i_1} \ldots \eta_{i_N} \\ &= \int \mathrm{d}^N\eta' \, M_{1\,i_1} \ldots M_{N\,i_N} \, \epsilon_{i_1 \ldots i_N} \eta_1 \ldots \eta_N \\ &= \int \mathrm{d}^N\eta' \det M \, \eta_1 \ldots \eta_N . \end{aligned} \tag{8.30}$$

It follows that under the transformation (8.29), the measure transforms according to the rule

$$\mathrm{d}^N\eta = \det M \, \mathrm{d}^N\eta' \,, \tag{8.31}$$

which is in some sense the "opposite" of the usual transformation rule, where the $\det M$ would have been on the LHS.

The stuff between the parentheses in the naive action (8.2) can be viewed as an operator. Keeping this in mind, we're going to show that the fermionic partition function can be viewed as a determinant. The setup is as follows: We have a Grassmann algebra with 2N generators $\bar{\eta}_i$ and η_i that all anticommute with each other and an $N \times N$ linear transformation M.

Theorem 8.1.1: Matthews-Salam formula

$$\int \prod_{i=1}^{N} \mathrm{d}\eta_i \, \mathrm{d}\bar{\eta}_i \, e^{\bar{\eta} M \eta} = \det M.$$

Proof. Let $\eta' = M\eta$. Then from eq. (8.29) we have

$$\begin{aligned}
\int \prod_{i=1}^{N} \mathrm{d}\eta_i \, \mathrm{d}\bar{\eta}_i \, e^{\bar{\eta} M \eta} &= \det M \int \prod_{i=1}^{N} \mathrm{d}\eta'_i \, \mathrm{d}\bar{\eta}_i \exp\left(\sum_{j=1}^{N} \bar{\eta}_j \eta'_j \right) \\
&= \det M \int \prod_{i=1}^{N} \mathrm{d}\eta'_i \, \mathrm{d}\bar{\eta}_i \exp\left(\bar{\eta}_i \eta'_i \right) \\
&= \det M \int \prod_{i=1}^{N} \mathrm{d}\eta'_i \, \mathrm{d}\bar{\eta}_i \left(1 + \bar{\eta}_i \eta'_i \right) \\
&= \det M.
\end{aligned}$$

The second line follows since pairs $\bar{\eta}_i \eta_i$ commute with each other, the third line is a power series expansion of the second line, and the last line follows from eq. (8.28). □

If we replace M with the Dirac operator, we see that the fermionic partition function is just the determinant of the Dirac operator. By the way, the ordering of the differentials in the measure should be $\mathrm{d}\eta \, \mathrm{d}\bar{\eta}$. This is because the $\bar{\eta}$ variables in the integrand will always come on the left. By placing the $\mathrm{d}\bar{\eta}$ differentials on the right, it will always hit the $\bar{\eta}$ variables first.

A generalization of this gives us the generating functional for fermions. Here we consider $4N$ Grassmann numbers $\bar{\eta}_i$, η_i, $\bar{\theta}_i$, and θ_i. The $\bar{\theta}_i$ and the θ_i are the source terms. We define

$$W[\theta, \bar{\theta}] \equiv \int \prod_{i=1}^{N} \mathrm{d}\bar{\eta}_i \, \mathrm{d}\eta_i \, e^{\bar{\eta} M \eta + \bar{\theta} \eta + \bar{\eta} \theta} \tag{8.32}$$

Theorem 8.1.2

$$W[\theta,\bar{\theta}] = e^{-\bar{\theta}M^{-1}\theta}\det M.$$

Proof. Using the definition of the generating functional and completing the square, we have

$$\int\prod_{i=1}^{N}\mathrm{d}\bar{\eta}_i\,\mathrm{d}\eta_i\,e^{\bar{\eta}M\eta+\bar{\theta}\eta+\bar{\eta}\theta} = \int\prod_{i=1}^{N}\mathrm{d}\bar{\eta}_i\,\mathrm{d}\eta_i\,e^{(\bar{\eta}+\bar{\theta}M^{-1})\,M\,(\eta+M^{-1}\theta)-\bar{\theta}M^{-1}\theta}.$$

Now we define new variables $\bar{\eta}' = \bar{\eta}+\bar{\theta}M^{-1}$ and $\eta' = \eta + M^{-1}\theta$. From eq. (8.28) it follows that

$$\int\prod_{i=1}^{N}\mathrm{d}\bar{\eta}_i\,\mathrm{d}\eta_i = \int\prod_{i=1}^{N}\mathrm{d}\bar{\eta}'_i\,\mathrm{d}\eta'_i\,.$$

Therefore by the Matthews-Salam formula,

$$\begin{aligned}\int\prod_{i=1}^{N}\mathrm{d}\bar{\eta}_i\,\mathrm{d}\eta_i\,e^{(\bar{\eta}+\bar{\theta}M^{-1})\,M\,(\eta+M^{-1}\theta)-\bar{\theta}M^{-1}\theta} &= \int\prod_{i=1}^{N}\mathrm{d}\bar{\eta}'_i\,\mathrm{d}\eta'_i\,e^{\bar{\eta}'M\eta'-\bar{\theta}M^{-1}\theta}\\ &= e^{-\bar{\theta}M^{-1}\theta}\det M.\end{aligned}$$

□

With the knowledge we now have, we can derive Wick's theorem, which lets us calculate fermionic expectation values. First we define

$$\langle\eta_{i_1}\bar{\eta}_{j_1}...\eta_{i_n}\bar{\eta}_{j_n}\rangle_F \equiv \frac{1}{Z_F}\int\prod_{k=1}^{N}\mathrm{d}\eta_k\,\mathrm{d}\bar{\eta}_k\,\bar{\eta}_{j_1}...\eta_{i_n}\bar{\eta}_{j_n}e^{-\bar{\eta}M\eta}. \tag{8.33}$$

Theorem 8.1.3: Wick's theorem

$$\langle\eta_{i_1}\bar{\eta}_{j_1}...\eta_{i_n}\bar{\eta}_{j_n}\rangle_F = (-1)^n\sum_{P_{1,...,n}}\mathrm{sign}P\,M^{-1}_{i_1,j_{P_1}}...\,M^{-1}_{i_n,j_{P_n}}.$$

Proof. From the definitions (8.32) and (8.33) we have

$$\langle\eta_{i_1}\bar{\eta}_{j_1}...\eta_{i_n}\bar{\eta}_{j_n}\rangle_F = \frac{1}{Z_F}\frac{\partial}{\partial\theta_{j_1}}\frac{\partial}{\partial\bar{\theta}_{i_1}}...\frac{\partial}{\partial\theta_{j_n}}\frac{\partial}{\partial\bar{\theta}_{i_n}}W[\theta,\bar{\theta}]\Big|_{\theta=\bar{\theta}=0}.$$

Let us now apply Theorem 8.1.2 to the RHS and carry out the derivatives. The $\det M$ will cancel with Z_F because of the Matthews-Salam formula. □

8.2 Fermion doubling

Let us now discuss one of the important problems with the naive fermionic action. First we introduce some results about Fourier transformations on the lattice. Define

$$V \equiv N_1 N_2 N_3 N_4 \tag{8.34}$$

with N_μ even. We generalize to *toroidal BCs*, i.e.

$$f(x + aN_\mu \hat{\mu}) = e^{2\pi i \theta_\mu} f(x). \tag{8.35}$$

Periodic BCs then have $\theta_\mu = 0$ and *anti-periodic BCs* have $\theta_\mu = 1/2$. The momentum space becomes

$$p_\mu = \frac{2\pi}{aN_\mu}(k_\mu + \theta_\mu), \qquad -\frac{N_\mu}{2} < k_\mu \leq \frac{N_\mu}{2}, \tag{8.36}$$

which reduces to the first Brillouin zone when periodic BCs are employed. By including the boundary phases in the momentum definition, plane waves

$$\exp(ip_\mu x_\mu) \tag{8.37}$$

will also obey the BCs. Now we derive a basic formula for Fourier transformations on the lattice. Let N be even and ℓ be an integer $0 \leq \ell \leq N-1$.

Proposition 8.2.1

$$\frac{1}{N} \sum_{j=-N/2+1}^{N/2} \exp\left(\frac{2\pi i \ell}{N}\right)^j = \delta_{\ell 0}.$$

Proof. Since N is even there are N terms in the above sum, so we find

the LHS to be 1. For $\ell \neq 0$ let $m \equiv j + N/2 + 1$ and define

$$q \equiv \exp\left(\frac{2\pi i \ell}{N}\right).$$

Then

$$\frac{1}{N} \sum_{j=-N/2+1}^{N/2} \exp\left(\frac{2\pi i \ell}{N}\right) \propto \sum_{m=0}^{N-1} q^m = \frac{1-q^N}{1-q} = 0$$

since $q^N = 1$. □

Applying this formula in each space-time direction, we arrive at

$$\frac{1}{V} \sum_p \exp\left(i p_\mu (x-x')_\mu\right) = \delta(x-x') = \delta_{n_1 n_1'} \delta_{n_2 n_2'} \delta_{n_3 n_3'} \delta_{n_4 n_4'}, \tag{8.38}$$

where $x_\mu = a n_\mu$, and

$$\frac{1}{V} \sum_x \exp\left(i (p-p')_\mu x_\mu\right) = \delta(p-p') = \delta_{k_1 k_1'} \delta_{k_2 k_2'} \delta_{k_3 k_3'} \delta_{k_4 k_4'}. \tag{8.39}$$

We define the Fourier transform as

$$\tilde{f}(p) = \frac{1}{\sqrt{V}} \sum_x f(x) \exp(-i p_\mu x_\mu). \tag{8.40}$$

The inverse transform

$$f(x) = \frac{1}{\sqrt{V}} \sum_p \tilde{f}(p) \exp(i p_\mu x_\mu) \tag{8.41}$$

can be verified by plugging the Fourier transform into the LHS.

Using these results about Fourier transformations on the lattice, we can begin to investigate fermion doubling. We shall do this for a single flavor for notational convenience; the result clearly generalizes to a summation over flavors. The gauge-invariant, naive fermion action (8.11) is bilinear in ψ and $\bar{\psi}$, so it can be written as

$$S_F = a^4 \sum_{x,y} \sum_{\alpha,\beta,c_1,c_2} \bar{\psi}(x)_{\alpha c_1} D(x|y)_{\alpha\beta c_1 c_2} \psi(y)_{\beta c_2}, \tag{8.42}$$

where α and β are Dirac indices, c_1 and c_2 are color indices, and the Dirac operator on the lattice is

$$D(x|y)_{\alpha\beta c_1 c_2} \equiv \sum_{\mu=1}^{4} (\gamma_\mu)_{\alpha\beta} \frac{U_\mu(x)_{c_1 c_2}\delta_{x+a\hat{\mu},y} - U_{-\mu}(x)_{c_1 c_2}\delta_{x-a\hat{\mu},y}}{2a} + m\delta_{\alpha\beta}\delta_{c_1 c_2}\delta_{xy}. \tag{8.43}$$

In this form, we can apply the Matthews-Salam formula or Wick's theorem by identifying $M = a^4 D$. For simplicity, we set $U_\mu(x) = \mathbf{1}$ for all links, just so we can see clearly how the doubling arises. Since we have free fermions, we may as well suppress color indices, and we will also use vector/matrix notation in Dirac space, so that Dirac indices are also suppressed. We find

$$\begin{aligned}
\tilde{D}(p|q) &= \frac{1}{V}\sum_{x,y} e^{-ipx} D(x|y) e^{-iqy} \\
&= \frac{1}{V}\sum_{x,y} e^{-ipx}\left(\sum_\mu \gamma_\mu \frac{\delta_{x+a\hat{\mu},y} - \delta_{x-a\hat{\mu},y}}{2a} + m\delta_{x,y}\mathbf{1}\right) e^{-iqy} \\
&= \frac{1}{V}\sum_{x} e^{-i(p+q)x}\left(\sum_\mu \gamma_\mu \frac{e^{iq_\mu a} - e^{-iq_\mu a}}{2a} + m\mathbf{1}\right) \\
&= \delta(p+q)\left(\frac{i}{a}\sum_\mu \gamma_\mu \sin(q_\mu a) + m\mathbf{1}\right) \\
&\equiv \delta(p+q)\tilde{D}(q).
\end{aligned} \tag{8.44}$$

In Gattringer and Lang [1] they take the transform of the y part to have opposite sign, which they say makes the similarity transformation unitary. For the purpose of seeing fermion doubling the sign does not make a difference; it only changes the delta function from $\delta(p-q)$ to $\delta(p+q)$. Therefore I decided to stick with the convention (8.40). If we want to use Wick's theorem, we will need to calculate the inverse of the Dirac operator. We find

$$\tilde{D}(p)^{-1} = \frac{m\mathbf{1} - ia^{-1}\sum_\mu \gamma_\mu \sin(p_\mu a)}{m^2 + a^{-2}\sum_\mu \sin(p_\mu a)^2}, \tag{8.45}$$

which can be verified by multiplying both sides by $\tilde{D}(p)$. This is the propagator for free fermions in momentum space. At $m = 0$ we find as $a \to 0$

$$\tilde{D}(p)^{-1} \to \frac{-i\sum_\mu \gamma_\mu p_\mu}{p^2}. \tag{8.46}$$

This propagator has a pole at $p = 0$. Poles in propagators correspond to physical particles, so in the continuum theory, the propagator corresponds to a single particle satisfying the Dirac equation. On the lattice if the fermions are massless, eq. (8.45) has a pole anywhere $p_\mu a = \pi$. In the first Brillouin zone, this happens 15 places besides $p_\mu = 0$. Hence on the lattice at finite spacing, the propagator has 15 unphysical poles that nevertheless correspond to some fermions. This is called *fermion doubling*, and we call the 15 unwanted particles the *doublers*.

This proliferation of extra particles can cause problems in the interacting theory. The additional states can be pair produced through interactions of the fermion field [2]. For instance even if all particles on the external lines of a diagram are the real particles, the doublers can appear in virtual loops. Therefore one typically wants to remove the doublers. One way to remove the doublers is to introduce an extra term that cancels the doublers on the lattice, and still reduces to the correct continuum value. With this formulation, the lattice momentum space Dirac operator is

$$\tilde{D}(p) = m\mathbf{1} + \frac{i}{a}\sum_\mu \gamma_\mu \sin(p_\mu a) + \frac{1}{a}\sum_\mu \mathbf{1}\big(1 - \cos(p_\mu a)\big). \tag{8.47}$$

This second term is called the *Wilson term*. This formulation of fermions is called *Wilson fermions*. The complete Dirac operator using this formulation becomes

$$D(x|y)^f_{\alpha\beta c_1 c_2} = \left(m^f + \frac{4}{a}\right)\delta_{\alpha\beta}\delta_{c_1 c_2}\delta_{xy} - \frac{1}{2a}\sum_{\mu=\pm 1}^{\pm 4}(\mathbf{1} - \gamma_\mu)_{\alpha\beta}U_\mu(x)_{c_1 c_2}\delta_{x+a\hat{\mu},y}, \tag{8.48}$$

where

$$\gamma_{-\mu} \equiv -\gamma_\mu. \tag{8.49}$$

8.3 Questions

1. When can we find a power series expansion for a function of Grassmann numbers?

2. Why do we want the integrand to vanish at the boundary?

3. Why do we know γ_μ and $U(x) \in \mathrm{SU}(N)$ commute?

References

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

[2] I. Montvay and G. Münster. *Quantum Fields on a Lattice*. Cambridge, Cambridge, 1994.

Appendix A

Special Topics in Mathematics

A.1 Hyperspherical coordinates

Here we work out how to compute spherical volumes in N dimensions. We warm up with 3D. In 3D we have the radial coordinate

$$r^2 = x^2 + y^2 + z^2, \tag{A.1}$$

an *azimuthal angle* $\phi \in [0, 2\pi]$ tracking the projection on the x-y plane, and a *polar angle* $\theta \in [0, \pi]$ measuring the inclination from the $\hat{z}$ direction. You can keep from mixing up the names by thinking of θ as measuring the distance from the North pole. The Cartesian coordinates are recovered from the spherical coordinates by

$$\begin{aligned} z &= r\,\mathrm{c}_\theta \\ x &= r\,\mathrm{s}_\theta\,\mathrm{c}_\phi \\ y &= r\,\mathrm{s}_\theta\,\mathrm{s}_\phi\,. \end{aligned} \tag{A.2}$$

The strange ordering of the coordinates will make sense when we move to N dimensions. Now the Jacobian corresponding to the transformation between Cartesian and spherical coordinates is

$$\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)} = \begin{pmatrix} \mathrm{s}_\theta\,\mathrm{c}_\phi & r\,\mathrm{c}_\theta\,\mathrm{c}_\phi & -r\,\mathrm{s}_\theta\,\mathrm{s}_\phi \\ \mathrm{s}_\theta\,\mathrm{s}_\phi & r\,\mathrm{c}_\theta\,\mathrm{s}_\phi & r\,\mathrm{s}_\theta\,\mathrm{c}_\phi \\ \mathrm{c}_\theta & -r\,\mathrm{s}_\theta & 0 \end{pmatrix}. \tag{A.3}$$

Knowing the Jacobian, we compute the volume element

$$\mathrm{d}^3x = \det\frac{\partial(x,y,z)}{\partial(r,\theta,\phi)}\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi = r^2\,\mathrm{d}r\,\mathrm{s}_\theta\,\mathrm{d}\theta\,\mathrm{d}\phi = r^2\,\mathrm{d}r\,\mathrm{d}\Omega\,, \tag{A.4}$$

where we will generically use $\mathrm{d}\Omega$ to stand in for the solid angle, regardless of the dimension. Integrating over the solid angle gives

$$\mathrm{d}\Omega = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{s}_\theta \, \mathrm{d}\theta = 4\pi, \tag{A.5}$$

so that the surface area of the 3-ball becomes

$$S_3 = R^2 \int \mathrm{d}\Omega = 4\pi R^2, \tag{A.6}$$

and the corresponding volume is

$$V_3 = \int_0^R \mathrm{d}r \, r^2 \int \mathrm{d}\Omega = \frac{4}{3}\pi R^3. \tag{A.7}$$

Now let us work in $N > 3$ dimensions. We define a radial coordinate

$$r^2 = \sum_{i=1}^{N} x_i^2 \tag{A.8}$$

with one azimuthal angle $\phi \in [0, 2\pi]$ and $N-2$ polar angles $\theta_i \in [0, \pi]$. To see that this mapping covers all of $\mathbb{R}^N$, note that the azimuthal angle covers an entire 2D plane; then one only has to sweep the 2D plane from 0 to π to cover an entire 3D hyperplane; and continuing in this way fills the entire N-dimensional volume, because the entire $(N-1)$-dimensional hyperplane was filled out in the preceding step. The Cartesian coordinates generalize the pattern of the 3D case as

$$\begin{aligned}
x_1 &= r \, \mathrm{c}_{\theta_1} \\
x_2 &= r \, \mathrm{s}_{\theta_1} \mathrm{c}_{\theta_2} \\
&\vdots \\
x_{N-2} &= r \, \mathrm{s}_{\theta_1} \ldots \mathrm{s}_{\theta_{N-3}} \mathrm{c}_{\theta_{N-2}} \\
x_{N-1} &= r \, \mathrm{s}_{\theta_1} \ldots \mathrm{s}_{\theta_{N-3}} \mathrm{s}_{\theta_{N-2}} \mathrm{c}_{\phi} \\
x_N &= r \, \mathrm{s}_{\theta_1} \ldots \mathrm{s}_{\theta_{N-3}} \mathrm{s}_{\theta_{N-2}} \mathrm{s}_{\phi},
\end{aligned} \tag{A.9}$$

so that the coordinate we call "x_1" is our new "z" coordinate.

At this step one could find the Jacobian and its determinant, and thus obtain the new integration measure. Looking at the coordinates (A.9) we see the volume element will generically separate as

$$\mathrm{d}^N x = r^{N-1} \, \mathrm{d}\Omega \,, \tag{A.10}$$

where the power $N-1$ is due to the fact that one column of the Jacobian has no r dependence, because we differentiated with respect to it. Unfortunately proceeding directly in this manner would require us to explicitly calculate an $N \times N$ matrix, calculate its determinant, then integrate over the resulting solid angle, which seems tedious at best. Thankfully there is a clever way to do this calculation.

Proposition A.1.1

The solid angle in N dimensions is

$$\int \mathrm{d}\Omega = \frac{2\pi^{N/2}}{\Gamma(N/2)}.$$

Proof. Recall

$$\int_{-\infty}^{\infty} \mathrm{d}x\, e^{-x^2} = \pi^{1/2}.$$

Therefore

$$\int \prod_{i=1}^{N} \mathrm{d}x_i\, e^{-x_i^2} = \left(\int_{-\infty}^{\infty} \mathrm{d}x\, e^{-x^2}\right)^N = \pi^{N/2}.$$

Switching to hyperspherical coordinates, carrying out a substitution, and using the definition of the gamma function, we can rewrite the LHS of the above as

$$\begin{aligned}\int \prod_{i=1}^{N} \mathrm{d}x_i\, e^{-x_i^2} &= \int \mathrm{d}\Omega \quad \int_0^{\infty} \mathrm{d}r\, r^{N-1} e^{-r^2} \\ &= \frac{1}{2}\int_0^{\infty} \mathrm{d}u\, u^{N/2-1} e^{u} \\ &= \frac{1}{2}\,\Gamma(N/2).\end{aligned}$$

Solving for the solid angle completes the proof. □

With this proposition, the surface area is easily dispatched,

$$S_N = \frac{2\pi^{N/2} R^{N-1}}{\Gamma(N/2)}, \tag{A.11}$$

as is the volume of the N-ball,

$$V_N = \frac{2\pi^{N/2} R^N}{N\Gamma(N/2)}. \tag{A.12}$$

As expected, these formulas agree with eqs. (A.6) and (A.7) when $N = 3$.

The Jacobian can be used to obtain the integration measure. For $N > 3$ dimensions, however, calculating the determinant can be a little unwieldy. An equivalent way of getting the measure right is to use wedge products. For example using properties like $\mathrm{d}x_1 \wedge \mathrm{d}x_2 = -\,\mathrm{d}x_2 \wedge \mathrm{d}x_1$ and $\mathrm{d}x_1 \wedge \mathrm{d}x_1 = 0$, one finds very quickly

$$\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3 \wedge \mathrm{d}x_4 = r^3\, \mathrm{s}_{\theta_1}^2\, s_{\theta_2}\, \mathrm{d}r \wedge \mathrm{d}\theta_1 \wedge \mathrm{d}\theta_2 \wedge \mathrm{d}\phi \tag{A.13}$$

or, as it is usually written,

$$\mathrm{d}^4x = r^3\, \mathrm{s}_{\theta_1}^2\, s_{\theta_2}\, \mathrm{d}r\, \mathrm{d}\theta_1\, \mathrm{d}\theta_2\, \mathrm{d}\phi\,. \tag{A.14}$$

Once we know the measure, we can also save ourselves some work finding the hyperspherical gradient operator. Each factor of the measure goes to the denominator of one of the terms. Make sure each denominator has units of length. In the present case, the gradient operator is read off from the measure to be

$$\partial = e_r \frac{\partial}{\partial r} + e_{\theta_1} \frac{1}{r} \frac{\partial}{\partial \theta_1} + e_{\theta_2} \frac{1}{r s_{\theta_1}} \frac{\partial}{\partial \theta_2} + e_\phi \frac{1}{r s_{\theta_1} s_{\theta_2}} \frac{\partial}{\partial \phi}, \tag{A.15}$$

where e_i are unit vectors in direction i and the ∂ on the LHS is meant to represent a vector, while ∂s on the RHS represents just the partial derivative operator.

A.2 Linear algebra reminders

Let F be a field. A *vector space* over F is a set V together with a binary operation $+$ under which F is abelian and a mapping $F \times V \to V$, denoted av $\forall\, a \in F$ and $v \in V$, satisfying

1. $(a + b)v = av + bv$,
2. $(ab)v = a(bv)$,

3. $a(v+w) = av + aw$, and

4. $1v = v$

$\forall\, a, b \in F$ and $v, w \in V$. The elements v and w are called *vectors*[1] A *subspace* is a subset W of V that still forms a vector space over F. I will become immediately less formal and refer to the underlying field only if it's absolutely necessary. Let V and V' be vector spaces over a field F. A *linear transformation* of V to V' is a mapping $T : V \to V'$ such that

1. $T(v+w) = T(v) + T(w)$ and

2. $T(av) = aT(v)$

$\forall\, a \in F$ and $v, w \in V$. We will often drop the parentheses and simply write $T(v) = Tv$. A linear transformation mapping a vector space to itself is called a *linear operator*. The *null space* or *kernel* of T is

$$\mathcal{N}(T) \equiv \{v \in V : Tv = 0\}; \tag{A.16}$$

i.e. it is the set of vectors annihilated by T. Finally T is said to be *idempotent* if $T^2 = T$. To connect some of the above definitions to what we learned in Chapter 1, we see that a linear operator is an endomorphism. Furthermore we see that if a linear operator has an inverse, then it is an automorphism.

Let M be any set. A *distance function* or *metric* is a function

$$d : M \times M \to \mathbb{R}$$

such that $\forall\, x, y, z \in M$

1. $d(x,y) = 0 \Leftrightarrow x = y$,

2. $d(x,y) = d(y,x)$,

3. $d(x,z) \leq d(x,y) + d(y,z)$ (this is called the *triangle inequality*).

A *metric space*, then, is a set M equipped with a metric d. One can generalize the idea of a metric and relax some of these conditions. For now this definition is sufficient for our purposes.

[1]Note that within the contexts of high energy physics and relativity, the word "vector" has a slightly more specific meaning: A "vector" in those contexts is a vector in the mathematical sense along with a demand on how it behaves under Lorentz transformations.

Example

$\mathbb{R}^n$ forms a metric space when equipped with

$$d(x,y) = |x-y| = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}. \tag{A.17}$$

This known as the *Euclidean metric*.

Next we turn to the concepts of orthogonality and dimensionality, which essentially tell us when two vectors are independent and how many independent generators it takes to make a vector space. We will formulate these ideas in terms of inner products. Let V be a vector space over $\mathbb{C}$. An *inner product* or *scalar product* is a mapping $(\ ,\ ) : V \times V \to \mathbb{C}$ satisfying $\forall\, x, y, z \in V$ and $\alpha \in F$

1. $(x,y) = (y,x)^*$,
2. $(x, y+z) = (x,y) + (x,z)$
3. $(x, \alpha y) = \alpha(x,y)$

Note that the above properties also imply $(x+y, z) = (x,z) + (y,z)$ and $(\alpha x, y) = \alpha^*(x,y)$. The *magnitude* or *norm* of a vector x is $|x| \equiv (x,x)^{1/2}$. If $(x,y) = 0$ then x and y are said to be *orthogonal* and we write $x \perp y$. If in addition $|x| = |y| = 1$, then they are *orthonormal*.

Proposition A.2.1

Let V be a vector space with inner product and let W a subspace of V with basis $\{w_1, ..., w_N\}$. Consider the mapping $P : V \to W$ defined by

$$Pv = \sum_{i=1}^{N} \frac{(v, w_i)}{|w_i|^2} w_i.$$

Then $v - Pv \perp W$.

Proof. We have

$$\begin{aligned}(v - Pv, w_j) &= (v, w_j) - (Pv, w_j) \\ &= (v, w_j) - \sum_{i=1}^{N} \frac{(v, w_i)}{|w_i|^2}(w_i, w_j) \\ &= (v, w_j) - \frac{(v, w_j)}{|w_j|^2}|w_j|^2 \\ &= (v, w_j) - (v, w_j) \\ &= 0.\end{aligned}$$

Since $v - Pv$ is orthogonal to every basis vector of W, it is orthogonal to every vector in W. □

The map of Proposition A.2.1 is called the *orthogonal projection*. A linear transformation U is *unitary* if $U^\dagger U = \mathbf{1}$. It is *hermitian* if $U^\dagger = U$, and it is *orthogonal* if $U^T U = \mathbf{1}$. A matrix is *positive semidefinite* if its eigenvalues are nonnegative.

Theorem A.2.1

Hermitian matrices are diagonalizable.

Proposition A.2.2

Let A and B be $n \times n$ matrices with differentiable elements. Then

$$\partial_\mu(AB) = (\partial_\mu A)B + A(\partial_\mu B).$$

Proof. We use the summation convention and abuse notation slightly: in the proposition statement ∂_μ has an identity matrix implicitly attached to it, but in this proof this symbol will also be used to denote the partial derivative operator. This is confusing, but it is also very common. Now since the matrix elements obey the product rule, we have

$$\begin{aligned}\big(\partial_\mu(AB)\big)_{mk} &= \delta_{mn}\partial_\mu(A_{nl}B_{lk}) \\ &= \delta_{mn}\big(\partial_\mu(A_{nl})B_{lk} + A_{nl}(\partial_\mu B_{lk})\big) \\ &= \big((\partial_\mu A)B + A(\partial_\mu B)\big)_{mk}.\end{aligned}$$

□

The *determinant* of an $N \times N$ matrix M with entries $m_{ij} \in \mathbb{C}$ is given by

$$\det M \equiv \epsilon_{i_1 \ldots i_N} m_{1\, i_1} \ldots m_{N\, i_N}. \tag{A.18}$$

A.3 Dirac algebra

In this section we summarize Section 3.2 of [1]. *Gamma matrices* appear when one works with spin-1/2 representations of the Lorentz group. Here we remind the reader of algebraic properties of gamma matrices. We work in 4D Minkowski space and demand that

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbf{1}, \tag{A.19}$$

where $\mathbf{1} \equiv \mathbf{1}_4$. The representation of the Lorentz algebra is then

$$S^{\mu\nu} = \frac{i}{4}\left[\gamma^\mu, \gamma^\nu\right]. \tag{A.20}$$

An explicit 4D representation of the gamma matrices, called the *chiral* or *Weyl* representation, uses the Pauli matrices:

$$\gamma^0 = \begin{pmatrix} 0 & \mathbf{1}_2 \\ \mathbf{1}_2 & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \tag{A.21}$$

In this representation, generators of boosts and rotations are, respectively,

$$S^{0i} = \frac{i}{4}\left[\gamma^0, \gamma^i\right] = -\frac{i}{2}\begin{pmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{pmatrix} \tag{A.22}$$

and

$$S^{ij} = \frac{i}{4}\left[\gamma^i, \gamma^j\right] = \frac{1}{2}\epsilon^{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}. \tag{A.23}$$

A four-component field that transforms under boosts and rotations according to the above generators is called a *Dirac spinor*, which we usually denote with ψ.

Knowing a little bit about gamma matrices, we can introduce "Feynman slash" notation,

$$\not{p} \equiv \gamma_\mu p^\mu, \tag{A.24}$$

and we can then write the *Dirac equation* in natural units as

$$(i\not{\partial} - m\mathbf{1})\psi = 0. \tag{A.25}$$

Next we define

$$\bar{\psi} \equiv \psi^\dagger \gamma^0, \tag{A.26}$$

which allows us to write down the Lorentz scalar $\bar{\psi}\psi$ and the Lorentz vector $\bar{\psi}\gamma^\mu\psi$. The matrix

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{A.27}$$

allows us to define left-handed and right-handed projection operators

$$P_L \equiv \frac{1}{2}(1-\gamma_5) \quad \text{and} \quad P_R \equiv \frac{1}{2}(1+\gamma_5), \tag{A.28}$$

which project out the left-handed and right-handed components of Dirac spinors like

$$\psi_R = P_R\psi, \quad \psi_L = P_L\psi, \quad \bar{\psi}_R = \bar{\psi}P_L, \quad \text{and} \quad \bar{\psi}_L = \bar{\psi}P_R. \tag{A.29}$$

The following Propositions tell us how to manipulate gamma matrices and projection operators.

Proposition A.3.1

Gamma matrix identities in $4D$:

$$\begin{array}{lrcl}
1. & \gamma_5^2 & = & \mathbf{1}; \\
2. & \{\gamma^\mu, \gamma_5\} & = & \mathbf{0}; \\
3. & \mathrm{tr}\,[\text{odd no. of } \gamma\text{s}] & = & \mathbf{0}; \\
4. & \not{a}\not{b} & = & -\not{b}\not{a} - 2(ab)\mathbf{1}; \\
5. & \mathrm{tr}\,\gamma^\mu\gamma^\nu & = & -4g^{\mu\nu}\mathbf{1}; \\
6. & \gamma_\mu\gamma^\mu & = & -4\mathbf{1}.
\end{array}$$

Proposition A.3.2

Projection operator identities:

1. $P_R^2 = P_R$;
2. $P_L^2 = P_L$;
3. $\gamma^\mu P_L = P_R \gamma^\mu$;
4. $\gamma^\mu P_R = P_L \gamma^\mu$.

References

[1] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview, Boulder, 1995. ISBN 978-0-201-50397-5.

Appendix B

Special Topics in Physics

Section B.1 follows Chapter 9 of Thomson [1]. Section B.2 follows Section 11.1 of Peskin and Schroeder [2]. Sections B.3 and **??** follow Chapter 7 of Gattringer and Lang [3].

B.1 Isospin and hypercharge

In quantum mechanics you learn that spin is a quantum number of charged particles. For example the electron is a spin-1/2 particle. The z-component S_3 of the spin operator S commutes with the Hamiltonian, and you learn that the eigenvectors of S_3 are the $+1/2$ and -1/2 states. You also learn that the components of S are related to the Pauli matrices by

$$S_i = \frac{\hbar}{2}\sigma_i. \tag{B.1}$$

Early on in nuclear physics, scientists noticed that the proton and neutron had about the same mass, and that the nuclear force between two nucleons (i.e. protons or neutrons) was approximately charge independent. Therefore Heisenberg suggested that protons and neutrons were two states of a single particle (the nucleon) just as there are spin-up and spin-down states of a spin-1/2 particle. The quantum number corresponding to this property is called *isospin*. Using this idea, the proton and neutron form an isospin doublet with total isospin $I = 1/2$ and z-component $I_3 = \pm 1/2$. Thus the Pauli matrices

also give a suitable representation of the isospin operator, and we write

$$I^2 = \sum_{i=1}^{3} I_i^2, \qquad I_i = \frac{1}{2}\sigma_i. \tag{B.2}$$

I know this is sloppy, but I want to leave it to the reader to determine from context whether I represents the operator or the eigenvalue.

The concept of isospin can be extended in the same way to quarks. In the simplest case we have $N_f = 2$ and consider the lightest quarks u and d. The SU(2) isospin symmetry is only approximate because the u and d quarks have slightly different masses. The isospin doublet then has a u component and a d component. Generally with N_f flavors of fermion, the symmetry group is SU(N_f) and we form N_f component multiplets in flavor space, one component per flavor.

Let's introduce the s quark and do the $N_f = 3$ case. Using the Pauli matrices and the fact that SU(3) has 8 generators, you can figure out what the Gell-Mann matrices are. We will say that u, d, and s are eigenvectors of isospin and write

$$u = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad d = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad s = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{B.3}$$

Then u and d span a 2D subspace of flavor space, so from the earlier discussion we should know that the generators of SU(2) are contained in the generators of SU(3). Hence

$$\lambda_1 = \begin{pmatrix} 0 & 1 & \\ 1 & 0 & \\ & & 0 \end{pmatrix}, \qquad \lambda_2 = \begin{pmatrix} 0 & -i & \\ i & 0 & \\ & & 0 \end{pmatrix}, \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & \\ 0 & -1 & \\ & & 0 \end{pmatrix}. \tag{B.4}$$

But there's nothing special about u and d; u and s will similarly form a subspace, and so will d and s. In both cases, we will use the Pauli matrices as generators. We can similarly write

$$\lambda_4 = \begin{pmatrix} 0 & & 1 \\ & 0 & \\ 1 & & 0 \end{pmatrix}, \qquad \lambda_5 = \begin{pmatrix} 0 & & -i \\ & 0 & \\ i & & 0 \end{pmatrix}, \qquad \lambda_X = \begin{pmatrix} 1 & & 0 \\ & 0 & \\ 0 & & -1 \end{pmatrix}, \tag{B.5}$$

$$\lambda_6 = \begin{pmatrix} 0 & & \\ & 0 & 1 \\ & 1 & 0 \end{pmatrix}, \qquad \lambda_7 = \begin{pmatrix} 0 & & \\ & 0 & -i \\ & i & 0 \end{pmatrix}, \qquad \lambda_Y = \begin{pmatrix} 0 & & \\ & 1 & 0 \\ & 0 & -1 \end{pmatrix}. \tag{B.6}$$

Finally note the fact that there are only 8 linearly independent generators, so two of these matrices should be linearly dependent. Since the $u \leftrightarrow d$ isospin symmetry is the closest to being exact, we choose the last generator to be a linear combination of λ_X and λ_Y that treats the u and d quarks symmetrically. Thus

$$\lambda_8 = \frac{1}{\sqrt{3}}\lambda_X + \frac{1}{\sqrt{3}}\lambda_Y = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & & \\ & 1 & \\ & & -2 \end{pmatrix}. \tag{B.7}$$

The new isospin and total isospin operators are

$$I^2 = \frac{1}{4}\sum_{i=1}^{8} I_i^2, \qquad I_i = \frac{1}{2}\lambda_i. \tag{B.8}$$

In the case of SU(2), the operators I_i do not commute, and therefore are not simultaneously diagonalizable. For SU(3) in the Gell-Mann basis, I_3 and I_8 are both diagonal, so they correspond to compatible observables. The observable we associate with I_8 is rescaled as

$$Y = \frac{1}{\sqrt{3}}\lambda_8, \tag{B.9}$$

and Y is called the *hypercharge*.

Theorem B.1.1: Gell-Mann-Nishijima Formula

The electric charge Q of a particle is related to its isospin and hypercharge by

$$Q = I_3 + \frac{1}{2}Y$$

B.2 Spontaneous symmetry breaking

Let $\phi(x)$ denote a vector (in the mathematical sense) of N real, scalar fields $\phi^i(x)$. Then the Lagrangian

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi\right)^2 + \frac{1}{2}\mu^2\phi^2 - \frac{\lambda}{4}\phi^4 \equiv \frac{1}{2}\left(\partial_\mu \phi\right)^2 + V(\phi) \tag{B.10}$$

is invariant under $O(N)$. (Recall orthogonal transformations are the ones with $R^T R = \mathbf{1}$.) This is the Lagrangian of the *linear sigma model*. Note that is is a generalization of ϕ^4 theory, but we have replaced the positive mass parameter m^2 with a negative parameter $-\mu^2$ and rescaled λ to eliminate a factor of 6. Classically, the potential is minimized when ϕ lies on an N-dimensional sphere of radius $\sqrt{\mu^2/\lambda}$, i.e. it is minimized for vectors $\phi_{\min}$ satisfying

$$\phi_{\min}^2 = \frac{\mu^2}{\lambda}. \tag{B.11}$$

To interpret the theory, we first choose coordinates so that $\phi_{\min}$ lies entirely along the N direction

$$\phi_{\min} = (0, 0, ..., 0, v), \tag{B.12}$$

where $v = \sqrt{\mu^2/\lambda}$ is the *vacuum expectation value* or VEV. Then, we define a set of shifted fields π_k and σ relative to this point by writing

$$\phi(x) = (\pi_1(x), \pi_2(x), ..., \pi_{N-1}(x), v + \sigma(x)) \tag{B.13}$$

Written in terms of π, the $N-1$ dimensional vector with components π_k, and σ, the new Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \pi\right)^2 + \frac{1}{2}\left(\partial_\mu \sigma\right)^2 - \frac{1}{2}\left(2\mu^2\right)\sigma^2 - \sqrt{\lambda}\mu\sigma^3 - \sqrt{\mu}\pi^2\sigma - \frac{\lambda}{4}\sigma^2 - \frac{\lambda}{2}\pi^2\sigma^2 - \frac{\lambda}{4}\pi^4, \tag{B.14}$$

where we have removed constant terms, because they do not change the physics. Equation (B.14) is the Lagrangian of $N-1$ massless, dynamic fields π_k and a dynamic field σ with mass $\sqrt{2}\mu$. Written in this form, the original O(N) symmetry is now obscured. There is a remaining O($N-1$) symmetry rotating the π_k among themselves. This is an example of *spontaneous symmetry breaking* (SSB), and we say something like "the original O(N) symmetry spontaneously breaks to the subgroup O($N-1$)."

Let's try to gain some geometric intuition for this phenomenon. Looking at eq. (B.12), we see that in ϕ space, the σ field corresponds to oscillations

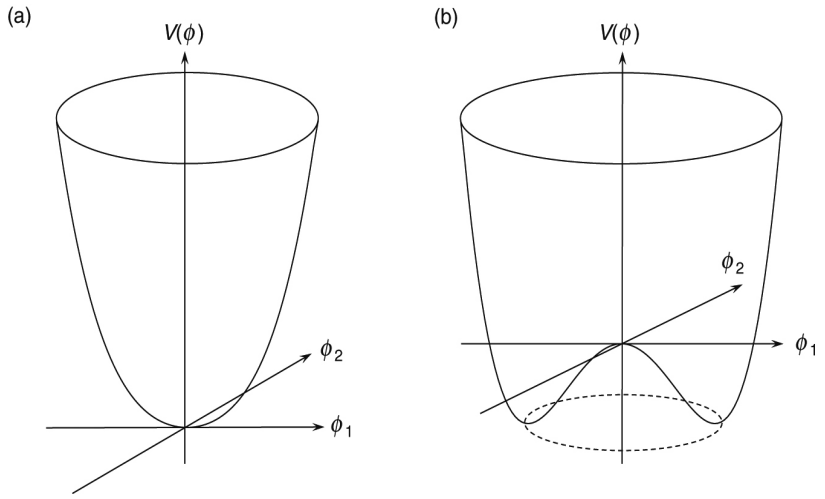

Figure B.1: Linear sigma model potential for $N = 2$. In (a) the ϕ field mass term $m^2 > 0$, while (b) gives this potential when m^2 is replaced by a negative parameter $-\mu^2$. Oscillations along the circle of minima in (b) correspond to the π field. Oscillations in the radial direction correspond to the σ field. Image taken from Thompson [1].

of ϕ orthogonal to the $N-1$ dimensional hypersurface, while the massless π_k fields corresponds to oscillations along the hypersurface. An example with $N = 2$ is shown in Figure B.1. If we take the ground state vector (B.12) and hit it with $\mathrm{O}(N)$, it will be rotated somewhere else on the hypersurface. The subgroup $\mathrm{O}(N-1)$ hits the first $N-1$ components of the ground state, which are all 0, thereby leaving it unchanged. In the original ϕ^4 theory with $m^2 > 0$, the ground state vector was 0, so the $\mathrm{O}(N)$ symmetry was also a symmetry of the ground state. After SSB, $\mathrm{O}(N)$ changes the ground state vector in general, which is why we say the symmetry is broken. Generally any symmetry respected by the Lagrangian but not by the ground state vector is a broken symmetry.

In the linear sigma model, massless π particles appeared after SSB. This is a special example of a general result known as Goldstone's theorem. The generated massless particles are referred to as *Goldstone bosons*. Many light

bosons can be interpreted as approximate Goldstone bosons; as we will see in Section B.3, the pion can be viewed in this manner.

Theorem B.2.1: Goldstone's theorem

Consider a Lagrangian of the form

$$\mathcal{L} = (\text{kinetic term for } \phi) + (\text{terms independent of } \phi) - V(\phi),$$

where ϕ is the N-dimensional vector of real, scalar fields ϕ_k, and $\mathcal{L}$ is invariant under a continuous, global transformation of ϕ with generators T^a. Then for every spontaneously broken generator there exists a corresponding Goldstone boson.

Proof. Let ϕ_{min} be a constant field minimizing V. Expanding V about this minimum we get to leading order

$$V(\phi) = V(\phi_{\text{min}}) + \frac{1}{2}(\phi - \phi_{\text{min}})_i(\phi - \phi_{\text{min}})_j \left.\frac{\partial^2 V}{\partial\phi_i \partial\phi_j}\right|_{\phi=\phi_{\text{min}}}.$$

The differences $\phi - \phi_{\text{min}}$ give the new fields of the theory after SSB; for example in the linear sigma model this difference is, from equations (B.12) and (B.13),

$$\phi - \phi_{\text{min}} = (\pi_1, ..., \pi_{N-1}, \sigma).$$

Therefore the coefficient of the quadratic term is a symmetric matrix whose eigenvalues give the masses of these fields. If we can prove that each broken generator implies a zero eigenvalue for this matrix, we are done. The kinetic term for ϕ is already invariant under the global transformation, so if $\mathcal{L}$ is invariant, it follows that V must be as well. Then we can write

$$V\left((\mathbf{1} - i\omega^a T^a)\phi\right) = V(\phi),$$

where ω is some infinitesimal parameter. Expanding to linear order yields

$$\frac{\partial V}{\partial \phi_j}(T^a\phi)_j = 0.$$

Differentiating the above with respect to ϕ_i and evaluating at $\phi_{\min}$ gives

$$\frac{\partial^2 V}{\partial \phi_i \partial \phi_j}\Big|_{\phi=\phi_{\min}} (T^a \phi_{\min})_j = 0,$$

i.e. $T^a\phi_{\min}$ is annihilated by the mass matrix. If T^a is a broken generator, we have $T^a\phi_{\min} \neq 0$, so the above equation implies $T^a\phi_{\min}$ is an eigenvector of the mass matrix with eigenvalue zero. □

B.3 Chiral symmetry in the continuum

In the continuum, the massless fermion action for a single flavor reads

$$S_F = \int \mathrm{d}^4x\, \mathcal{L}_F = \int \mathrm{d}^4x\, \bar{\psi} \not{D} \psi. \tag{B.15}$$

We refer to $\not{D}$ as the *massless Dirac operator*. A *chiral rotation* of the fermion fields is a transformation mapping

$$\psi \to e^{i\alpha\gamma_5}\psi \quad \text{and} \quad \bar{\psi} \to \bar{\psi} e^{i\alpha\gamma_5}, \tag{B.16}$$

where $\alpha \in \mathbb{R}$. This is probably called a chiral rotation because γ_5 is used to define the operators (A.28), which project out the left-handed and right-handed components of the fermion field according to eq. (A.29). Using identity 2 of Proposition A.3.1, we find that $\mathcal{L}_F$ transforms under this rotation as

$$\mathcal{L}_F \to \bar{\psi} e^{i\alpha\gamma_5} \not{D} e^{i\alpha\gamma_5}\psi = \bar{\psi} e^{i\alpha\gamma_5} e^{-i\alpha\gamma_5} \not{D}\psi = \mathcal{L}_F, \tag{B.17}$$

i.e. it is invariant under chiral rotations. Using Proposition A.3.2, one can decompose $\mathcal{L}_F$ into its left-handed and right-handed parts as

$$\mathcal{L}_F = \bar{\psi}_L \not{D} \psi_L + \bar{\psi}_R \not{D} \psi_R, \tag{B.18}$$

and we colloquially say that the chiral components of $\mathcal{L}_F$ "do not talk to each other." If one were to include a mass term in $\mathcal{L}_F$, it would decompose as

$$m\bar{\psi}\psi = m\left(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R\right), \tag{B.19}$$

which mixes the chiral components, thereby breaking chiral symmetry. This is why one refers to the limit $m \to 0$ as the *chiral limit*.

We now generalize these ideas to N_f flavors of fermion. In Gattringer and Lang eq. (7.11), they write the fermion action as

$$S_F = \int \mathrm{d}^4x\, \mathcal{L}_F = \int \mathrm{d}^4x\, \bar{\psi} \left(\not{D} + M \right) \psi, \tag{B.20}$$

where M is the *mass matrix*

$$M \text{ “}=\text{”} \ \mathrm{diag}(m_1, m_2, ..., m_{N_f}) \tag{B.21}$$

acting in flavor space. M cannot be a $N_f \times N_f$ matrix, because we have defined γ^μ as a 4×4 matrix. The only way adding these matrices makes sense is if M is a $4N_f \times 4N_f$ matrix, i.e. if

$$M = \begin{pmatrix} m_1 \mathbf{1}_4 & & & \\ & m_2 \mathbf{1}_4 & & \\ & & \ddots & \\ & & & m_{N_f} \mathbf{1}_4 \end{pmatrix}. \tag{B.22}$$

Then ψ must be a $4N_f$ vector (in the mathematical sense, not in the Lorentz transformation sense) looking something like

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{N_f} \end{pmatrix}, \tag{B.23}$$

where each ψ_i is a 4-component Dirac spinor, and presumably when we write γ^μ in eq. (B.20) what we really mean is

$$\gamma^\mu = \gamma^\mu \mathbf{1}_{N_f} = \begin{pmatrix} \gamma^\mu & & & \\ & \gamma^\mu & & \\ & & \ddots & \\ & & & \gamma^\mu \end{pmatrix}. \tag{B.24}$$

I pray the reader forgives the notational perversion in the above equation.

In the chiral limit, the action eq. (B.20) is again invariant under chiral rotations, also in this context called *axial vector rotations*, taking the form

$$\psi \to e^{i\alpha\gamma_5 T^a}\psi, \qquad \bar{\psi} \to \bar{\psi} e^{i\alpha\gamma_5 T^a}, \tag{B.25}$$

$$\psi \to e^{i\alpha\gamma_5 \mathbf{1}}\psi, \qquad \bar{\psi} \to \bar{\psi} e^{i\alpha\gamma_5 \mathbf{1}}, \tag{B.26}$$

where the T^a are the N_f^2-1 generators of $\mathrm{SU}(N_f)$, $\mathbf{1} \equiv \mathbf{1}_{4N_f}$, and again $\alpha \in \mathbb{R}$. In this limit the action is also invariant under the *vector transformations*

$$\psi \to e^{i\alpha T^a}\psi, \qquad \bar{\psi} \to \bar{\psi}e^{-i\alpha T^a}, \tag{B.27}$$

$$\psi \to e^{i\alpha \mathbf{1}}\psi, \qquad \bar{\psi} \to \bar{\psi}e^{-i\alpha \mathbf{1}}. \tag{B.28}$$

This invariance under the above two equations extends to the case of degenerate masses, when $m_1 = m_2 = ... = m_{N_f} \equiv m$, and is the familiar isospin symmetry generalized to N_f flavors. The symmetry (B.28) holds for arbitrary masses, and the conserved quantity is the *baryon number*

$$B = \frac{1}{3}(n_q - n_{\bar{q}}). \tag{B.29}$$

You can see why it's called baryon number: A baryon is made of 3 quarks, and hence has $B = 1$. Anti-quarks have $B = -1$, and mesons have $B = 0$.

Returning once again to the massless limit, the invariance of the action under eqs. (B.25), (B.26), (B.27), and (B.28) represents the global symmetry group

$$\mathrm{SU}(N_f)_L \times \mathrm{SU}(N_f)_R \times \mathrm{U}(1)_V \times \mathrm{U}(1)_A. \tag{B.30}$$

Breaking the global symmetry group (B.30) has important implications in QCD phenomonology. For example we will show that in the quantized, massless theory the fermion determinant changes under (B.26), breaking the $\mathrm{U}(1)_A$ symmetry explicitly. The remaining symmetry is

$$\mathrm{SU}(N_f)_L \times \mathrm{SU}(N_f)_R \times \mathrm{U}(1)_V. \tag{B.31}$$

If the fermion masses are degenerate, the symmetry $\mathrm{SU}(N_f)_L \times \mathrm{SU}(N_f)_R$ breaks to its subgroup $\mathrm{SU}(N_f)_V$. Now the remaining symmetry is

$$\mathrm{SU}(N_f)_V \times \mathrm{U}(1)_V. \tag{B.32}$$

Finally allowing non-degenerate masses breaks the symmetry further. There remains

$$\underbrace{\mathrm{U}(1)_V \times \mathrm{U}(1)_V \times ... \times \mathrm{U}(1)_V}_{N_f \text{ times}}. \tag{B.33}$$

The typical QCD scale is ~ 1 GeV (think protons, which have a mass of about 940 MeV) so the u, d, and s quarks have masses that are relatively small (about 5 MeV, 5 MeV, and 100 MeV, respectively). Since these masses

are so close to zero on this scale, we can say that when $N_f = 2$, and partly for $N_f = 3$, eq. (B.31) is an approximate symmetry. If the u and d quarks were massless, this symmetry would be exact, and it can be argued [3] that a nucleon and its negative parity partner should have the same mass. However one finds the negative parity nucleon to have a mass of about 1535 MeV, and this difference of about 600 MeV is too large to be explained by the slight breaking due to the u and d masses. We conclude that something else is happening. This something is called *spontaneous chiral symmetry breaking*, and it comes out of the dynamics of QCD. If quarks were massless, the pions would arise as the Goldstone bosons of chiral SSB; hence we refer to pions as "would-be" Goldstone bosons.

References

[1] M. Thomson. *Modern Particle Physics*. Cambridge, New York, 2013. ISBN 978-1-107-03426-6.

[2] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview, Boulder, 1995. ISBN 978-0-201-50397-5.

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

Index