The Algebra $\mathfrak{su}(\mathrm{N})$, the Group SU(N), and Chern-Simons Theory

Charles Snider

Adviser: Antal Jevicki

1 Introduction

The Lie group $\mathbf{SU}(\mathrm{N})$ has many definitions. In a basic sense, it can be considered the group of linear transformations on N-dimensional complex vectors that preserve the scalar product: it is a group of rotations. More formally, $\mathbf{SU}(\mathrm{N})$ is the group of unitary matrices with determinant one:

$$\mathbf{SU}(\mathrm{N}) = \{A \in \mathrm{GL_N}(\mathbb{C}); \det A = 1, A^\dagger = A^{-1}\} \tag{1}$$

It gives the familiar structure of basic spin-$\frac{1}{2}$ particles from any introductory quantum mechanics course: the momentum operator J_3, and its raising and lowering operators, $J_\pm$, are built from the structure of $\mathbf{SU}(2)$ and its algebra. In quantum chromodynamics, the symmetries of color and (approximately) flavor build out of the theory of $\mathbf{SU}(3)$. In its general form, $\mathbf{SU}(\mathrm{N})$, has diverse applications to symmetries of field theory Lagrangians.

The study of $\mathbf{SU}(\mathrm{N})$ is particularly interesting because it is a Lie group: it therefore has a corresponding *Lie algebra*, denoted $\mathfrak{su}(\mathrm{N})$, such that every element of $\mathbf{SU}(\mathrm{N})$ can be written as the exponential of an element of $\mathfrak{su}(\mathrm{N})$. This allows for the direct study of the properties of $\mathbf{SU}(\mathrm{N})$ through a careful analysis of $\mathfrak{su}(\mathrm{N})$, which itself has unique properties owing to its classification as a Lie algebra.

In particular, we are interested in the applications of $\mathbf{SU}(\mathrm{N})$ and $\mathfrak{su}(\mathrm{N})$ to a peculiar field theory: Chern-Simons theory. In its non-abelian form, Chern-Simons theory takes elements of $\mathfrak{su}(\mathrm{N})$ and its representations for its gauge field and elements of $\mathbf{SU}(\mathrm{N})$ (and its representations) for its gauge transformations.

To this end, and because the study of Lie algebras is fascinating in its own right, we will establish the basic structure of $\mathfrak{su}(\mathrm{N})$ and its representations, briefly examine the theory of representations and characters of $\mathbf{SU}(\mathrm{N})$, and take this understanding and use it to study the canonical structure of Chern-Simons theory.

2 The Lie Algebra

The Lie algebra $\mathfrak{su}(\mathrm{N})$ is the set of anti-Hermitian traceless $N \times N$ matrices, with Lie bracket given by the usual matrix commutator.

$$\mathfrak{su}(\mathrm{N}) = \{A \in \mathrm{GL_N}(\mathbb{C}); \mathrm{Tr}\, A = 0, A = -A^\dagger\} \quad \text{with} \quad [A, B] = AB - BA \tag{2}$$

$\mathfrak{su}(\mathrm{N})$ is a simple Lie algebra; in particular, it is semisimple, so we can decompose it as a direct sum of a Cartan subalgebra and its corresponding root spaces (with roots α_i)

$$\mathfrak{su}(\mathrm{N}) = H \oplus L_{\alpha_1} \oplus \cdots \oplus L_{\alpha_k} \tag{3}$$

This form of the algebra is called the *root space decomposition.* From this decomposition, we will be able to derive powerful results regarding the representation theory and classification of $\mathfrak{su}(N)$.

2.1   The Defining Representation

To study the properties of $\mathfrak{su}(N)$, we must first establish its defining representation. From the defining representation, we will find a basis for the algebra, the generators, and the roots and weights; the properties of these objects will allow us to study the representations of $\mathfrak{su}(N)$.

Note: Although $\mathfrak{su}(N)$ is the set of anti-Hermitian traceless matrices, we often study its properties using a set of Hermitian matrices. When we take $\mathfrak{su}(N)$ to $\mathbf{SU}(N)$ by the exponential map, we add a factor of i, making the matrices anti-Hermitian as needed. Essentially, we redefine $\mathfrak{su}(N)$ as the set of Hermitian traceless matrices, and relate it to $\mathbf{SU}(N)$ by

$$\mathbf{SU}(N) = \{A \in \mathrm{GL}_N(\mathbb{C}); A = e^{iM} \quad \text{for some} \quad M \in \mathfrak{su}(N)\}$$

In our discussion, we will use this convention.

We first find the Cartan subalgebra H. A convenient choice for its basis is a generalization of the Paul matrices (or Gell-Mann matrices, for that matter) given by $N \times N$ matrices H_i (in block diagonal form) [3]

$$H_{j-1} = \frac{1}{\sqrt{2j(j-1)}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & \cdots & \\ 0 & 1 & 0 & \cdots & \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & -(j-1) & \\ \hline & & 0 & & 0 \end{array} \right) \tag{4}$$

where the upper left block is $j \times j$. For example, the first two H_i are

$$H_1 = \frac{1}{2} \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & -1 & \\ \hline & 0 & 0 \end{array} \right)$$

$$H_2 = \frac{1}{\sqrt{12}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & \\ \hline & 0 & & 0 \end{array} \right)$$

Note: H_1 corresponds to index $j = 2$ in Equation 4, H_2 to index $j = 3$, et cetera. There are $N - 1$ total elements H_i; the number of elements in the basis of the Cartan subalgebra is called the *rank* of the algebra.

Next, we must find the root spaces. Each root space L_{α_i} is the span of all common eigenvectors of the adjoint representation of the Cartan subalgebra, ad H, corresponding to a specific weight α_i.

Note: The weights specific to the root spaces are called *roots*; a weight, in general, is a function $\lambda : A \to F$ for some Lie subalgebra A and field F such that there is at least one

non-zero eigenvector v with $a(v) = \lambda(a)v$ for all $a \in A$. We represent the roots of $\mathfrak{su}(N)$ as vectors in $\mathbb{R}^{N-1}$.

Each L_{α_i} has a basis of matrices M, with

$$M \in \mathfrak{su}(N) \quad \text{such that} \quad \operatorname{ad} H_j(M) = [H_j, M] = (\alpha_i)_j M \tag{5}$$

where $(\alpha_i)_j$ is the jth component of the weight α_i. To find these matrices M, we define matrices S_{ij} and S_{ij}^* (with $i < j$ and $k < \ell$ to avoid multiplicity) in terms of matrix components:

$$(S_{ij})_{k\ell} = \frac{1}{2} (\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}) \tag{6}$$

$$\left(S_{ij}^*\right)_{k\ell} = \frac{i}{2} (\delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk}) \tag{7}$$

These S matrices, along with the matrices H, give a basis for $\mathfrak{su}(N)$ that is a generalization of the Pauli matrices. There are $N - 1$ matrices H, and $N^2 - N$ total S matrices: two for each matrix element above the diagonal. Furthermore, they follow an orthogonality condition that mirrors that of the Pauli matrices [1]

$$\operatorname{Tr}\left(T^a T^b\right) = \frac{1}{2}\delta_{ab} \tag{8}$$

where we use the notation T^a to represent an arbitrary one of the H, S, S^*, and the a acts as an index on them (assuming some ordering), running from $a = 1$ to $a = N^2 - 1$. This condition, also obeyed by the half Pauli matrices $\sigma_i/2$, finalizes the sense in which our basis generalizes the Pauli matrices (this condition is also obeyed by the Gell-Mann matrices, which can be found by taking the H and S matrices when $N = 3$). It is somewhat tedious to show this, but not terribly complicated.

Case (1): If both T^a and T^b are distinct elements in the Cartan subalgebra, say H_i and H_j, then, up to an overall factor, the matrix will be equal to H_k where k is the smaller of i and j. However, since the H matrices are traceless, the condition is satisfied.

Case (2): If both T^a and T^b are the same element of the Cartan subalgebra, say H_{j-1}, then we will get

$$\operatorname{Tr} H_{j-1}^2 = \frac{1}{2j(j-1)} \operatorname{Tr} \left(\begin{array}{cccc|c} 1 & 0 & 0 & \cdots & \\ 0 & 1 & 0 & \cdots & \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & (j-1)^2 & \\ \hline & & 0 & & 0 \end{array} \right) = \frac{j-1+(j-1)^2}{2j(j-1)} = \frac{1+j-1}{2j} = \frac{1}{2}$$

as needed.

Case (3): One of the elements is in S or S^* and the other is in H. Since H is diagonal, it will act on an element of S or S^* by multiplying its elements by constant factors. Thus, since S and S^* are already traceless, the condition is satisfied.

Case (4): Both elements are in S or S^*. Writing in terms of matrix components, the product of an element of S and S^* is given by

$$(S_{ij}S_{k\ell}^*)_{uw} = (S_{ij})_{uv}\left(S_{k\ell}^*\right)_{vw} = \frac{i}{4} (\delta_{iu}\delta_{jv} + \delta_{iv}\delta_{ju})(\delta_{kv}\delta_{\ell w} - \delta_{kw}\delta_{\ell v})$$

3

We are only interested in the case when $u = w$, as we will be taking a trace. This simplifies the expression to

$$
\begin{aligned}
\left(S_{ij} S_{k\ell}^*\right)_{uu} &= \left(S_{ij}^*\right)_{uv} (S_{k\ell})_{vu} \\
&= \frac{i}{4} \left(\delta_{iu}\delta_{jv} + \delta_{iv}\delta_{ju}\right) \left(\delta_{kv}\delta_{\ell u} - \delta_{ku}\delta_{\ell v}\right) \\
&= (\text{symmetric in } u, v) \, (\text{anti-symmetric in } u, v)
\end{aligned}
$$

When we sum over both u and v to get the trace, the entire term will disappear since we are summing over the product of symmetric and anti-symmetric terms.

When both terms are in S, the trace takes the form

$$
\begin{aligned}
\left(S_{ij} S_{k\ell}\right)_{uu} &= \left(S_{ij}\right)_{uv} (S_{k\ell})_{vu} \\
&= \frac{1}{4} \left(\delta_{iu}\delta_{jv} + \delta_{iv}\delta_{ju}\right) \left(\delta_{kv}\delta_{\ell u} + \delta_{ku}\delta_{\ell v}\right) \\
&= \frac{1}{2}\delta_{ik}\delta_{j\ell}
\end{aligned}
$$

Since each expression on the second line will be non-zero twice: for the first, when $i = u, j = v$ and $j = u, i = v$. For the second, when $k = v, \ell = u$ and $k = u, \ell = v$. Each time, the expression in parenthesis has value 1. These non-zero terms will only line up when $i = k, j = \ell$, which is equivalent to T^a, T^b with $a = b$, as needed.

When both terms are in S^*, we have

$$
\begin{aligned}
\left(S_{ij}^* S_{k\ell}^*\right)_{uu} &= \left(S_{ij}^*\right)_{uv} (S_{k\ell}^*)_{vu} \\
&= -\frac{1}{4} \left(\delta_{iu}\delta_{jv} - \delta_{iv}\delta_{ju}\right) \left(\delta_{kv}\delta_{\ell u} - \delta_{ku}\delta_{\ell v}\right) \\
&= \frac{1}{2}\delta_{ik}\delta_{j\ell}
\end{aligned}
$$

The same argument as when both matrices where S matrices applies, except that each the indices line up, one of the terms in parentheses will be -1 rather than 1: since $i = k, j = \ell$, when $i = u$ and $j = v$, then the k, ℓ term will pick up a negative sign to cancel the overall negative, and vice versa. This completes the proof of the trace identity.

Furthermore, since each element of this basis is itself in the Lie algebra, we know that the Lie bracket of any two of these generators will also be given by a linear combination of the other generators.

$$
\left[T^a, T^b\right] = i f^{abc} T^c \tag{9}
$$

The numbers f^{abc} are called the **structure constants**. The structure constants are difficult to define for an arbitrary N; for example, however, when $N = 2$ and we are working with $\mathfrak{su}(2)$, they are given by the Levi-Civita symbol $f^{abc} = i\varepsilon^{abc}$. In general, however, they are antisymmetric with respect to any two of the indices, and thus vanish if two indices are equal. The structure constants will prove useful when considering the quantization of non-abelian field theories.

Side Note: the adjoint representation given above, $\operatorname{ad} L$, with action

$$
\operatorname{ad} A(B) = [A, B]
$$

4

can be expressed in terms of the structure constants. We can write the form of the adjoint representation of a generator in matrix components:

$$(\operatorname{ad} T^a)_{bc} = -if^{abc}$$

With this explicit definition of the map, we see that the adjoint map of a generator takes the form of an $N^2 - 1$ dimensional matrix. If we treat the generators as states on which it acts (with T^a being the $N^2 - 1$ dimensional column vector with ith element δ_{ai}), we can consider the action of the adjoint map on the generators:

$$\operatorname{ad} T^a \left(T^b\right)_c = (\operatorname{ad} T^a)_{cd} \left(T^b\right)_d = -if^{acd}\delta_{bd} = if^{acb}$$

This corresponds to a final state which is the linear combination $-if^{acb}T^c = if^{abc}T^c$ which is precisely the value of the commutator of T^a and T^b. Thus, this definition of the adjoint representation precisely matches our original definition. This definition of the adjoint representation is often presented in physics discussions, as it offers an explicit expression for the adjoint representation, but is limited to matrix algebras (which are the main concern in physical situations).

However, despite its nice properties and applications, this choice of basis proves minimally useful for the analysis of the representation theory of $\mathfrak{su}(N)$. In particular, the S matrices are not eigenvectors of $\operatorname{ad} H$, so we have yet to find the root space decomposition of Equation 3. To find eigenvectors for $\operatorname{ad} H$, we draw inspiration from basic quantum mechanics and define "raising" and "lowering" matrices E_{ij}. In terms of matrix elements, these are given by

$$(E_{ij})_{k\ell} = \frac{1}{\sqrt{2}}\delta_{ik}\delta_{j\ell} \tag{10}$$

We can write these matrices in terms of the S matrices as we might expect

$$E_{ij} = \frac{S_{ij} - iS^*_{ij}}{\sqrt{2}} \qquad E_{ji} = \frac{S_{ij} + iS^*_{ij}}{\sqrt{2}}$$

which mirrors the structure of the raising and lowering operators from elementary quantum mechanics. If we consider the action of $\operatorname{ad} H$ on these matrices, we see

$$[H_k, E_{ij}] = (\alpha_{ij})_k E_{ij} = \left((H_k)_{ii} - (H_k)_{jj}\right) E_{ij}$$

These E_{ij} matrices are eigenvectors of $\operatorname{ad} H$, as needed. However, they themselves are not actually members of the algebra $\mathfrak{su}(N)$: they are not Hermitian. To find a root space decomposition using the E_{ij}, we define [2]

$$\bigoplus_{\alpha_i} L_{\alpha_i} \equiv L \cap \operatorname{Span}\{E_{ij} \text{ for all } i, j\} \tag{11}$$

so that we can take our root spaces to be spanned by E_{ij} without including elements not contained within the algebra. Using Equation 11 for the root spaces and the matrices from Equation 4 as the basis for the Cartan subalgebra, our root space decomposition is

$$\mathfrak{su}(N) = H \oplus \bigoplus_{\alpha_i} L_{\alpha_i} \tag{12}$$

5

We now have root spaces spanned by the E_{ij}. If we relabel our roots and root spaces to α_{ij} and $L_{\alpha_{ij}}$ to correspond to E_{ij}, we immediately find the form of our roots α_{ij}. We define $(H_k)_{jj} \equiv \nu_j^k$ for convenience, and see

$$\alpha_{ij} = \left(\nu_i^1 - \nu_j^1, \ldots, \nu_i^{N-1} - \nu_j^{N-1}\right) \tag{13}$$

where the kth component represents the eigenvalue of E_{ij} with respect to $\text{ad}\, H_k$. Since we have one E_{ij} for every non-diagonal ij, there are $N^2 - N$ elements E_{ij}. We see that we have again $N^2 - 1$ generators for the algebra.

With this definition of Cartan subalgebra and root spaces, we can now explicitly state the commutators of the generators of $\mathfrak{su}(N)$, which describes the basic structure of the defining representation

$$[H_i, H_j] = 0$$
$$[H_k, E_{ij}] = (\alpha_{ij})_k E_{ij}$$
$$[E_{ij}, E_{k\ell}] = \frac{1}{\sqrt{2}}\left(E_{i\ell}\delta_{jk} - E_{kj}\delta_{i\ell}\right)$$

However, these relations describe the algebra only in a most basic sense. A much more powerful classification comes from its root system, which we will establish later.

Note: We have introduced two sets of matrices which define the algebra. The H and S matrices, and the H and E_{ij} matrices. Any element of $\mathfrak{su}(N)$ can be written as a linear combination of matrices from either set alone; however, we note that not every linear combination of the H and E_{ij} will be Hermitian: this restriction led to the definition of Equation 11. We are free to use either set of matrices to analyze the algebra, but when using the H and E_{ij}, we will have to assume that we are careful about our choice of coefficients in linear combinations. The E_{ij}, being eigenvectors of $\text{ad}\, H$, are particularly powerful when analyzing the representation theory of $\mathfrak{su}(N)$. On the other hand, the S matrices are more useful when directly approaching non-abelian field theories. With this in mind, we will be using the E_{ij} matrices for the rest of our discussion of representation theory, and the S matrices when we return to discuss field theories directly.

We now introduce the weights for $\mathfrak{su}(N)$, which will be essential in studying the representation theory of both the algebra and the group.

We define our weights of the Cartan subalgebra to be the vectors which encode all information about the eigenvalues of the H_i acting on common eigenvectors $v \in \mathbb{C}^N$. Clearly, the common eigenvectors of the Cartan subalgebra are the vectors $v_i = (0, \ldots, 1, \ldots, 0)$ with a 1 in the ith component. We define our weights μ^j as the vector in $\mathbb{R}^{N-1}$ which has as its kth component the eigenvalue of v_j with respect to H_k. For the first few v_i, the weights are

$$\mu^1 = \left(\frac{1}{2}, \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{24}}, \ldots, \frac{1}{\sqrt{2N(N-1)}}\right)$$

$$\mu^2 = \left(-\frac{1}{2}, \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{24}}, \ldots, \frac{1}{\sqrt{2N(N-1)}}\right)$$

$$\mu^3 = \left(0, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{24}}, \ldots, \frac{1}{\sqrt{2N(N-1)}}\right)$$

The arbitrary mth weight is given by

$$\mu^m = \left(0, 0, \ldots, 0, \frac{1-m}{\sqrt{2m(m-1)}}, \frac{1}{\sqrt{2m(m+1)}}, \ldots, \frac{1}{\sqrt{2N(N-1)}} \right) \qquad (14)$$

with the first non-zero component in the $(m-1)$th spot.

From this definition of roots and weights, it becomes clear that the roots, as defined above in Equation 13, can be alternately defined as the difference of weights. For a given root α_{ij}, a simple calculation shows we can instead define it

$$\alpha_{ij} = \mu^i - \mu^j \qquad (15)$$

We now introduce a few definitions. A *positive* weight or root is one whose last non-zero component is positive (this is actually a backwards convention, but it is convenient for our definition of the simple roots later [3]). Also, a weight or root is called *higher* than another if the difference of them is positive. Note that all the μ^i are positive except μ^N. For example, the root $\alpha_{12} = \mu^1 - \mu^2$ is positive (in fact, the positive roots are those α_{ij} such that $i < j$. The *highest* weight is, naturally, the one which is higher than all other weights. The highest weight of an irreducible representation is unique. Finally, a root is called *simple* if it is positive and cannot be written as a sum of other positive roots.

2.1.1 An Identity on the Generators

In later sections, we will need a further relation between the generators [1]

$$\sum_{a=1}^{N^2-1} (T^a)_{\alpha\beta} (T^a)_{\alpha'\beta'} = \frac{1}{2} \delta_{\alpha\beta'} \delta_{\alpha'\beta} - \frac{1}{2N} \delta_{\alpha\beta} \delta_{\alpha'\beta'}$$

Where T^a represents one of the generators, either one of the H or E_{ij}. We now set out to prove this. Clearly, for the E_{ij}, this is only non-zero when $\alpha' = \beta$ and $\alpha = \beta'$, which gives us a term

$$\sum_{a=1}^{N^2-1} (T^a)_{\alpha\beta} (T^a)_{\alpha'\beta'} = \frac{1}{2} \delta_{\alpha\beta'} \delta_{\alpha'\beta}$$

The H_i only contribute when $\alpha = \beta$ and $\alpha' = \beta'$. In the case $\alpha = \beta = \alpha' = \beta'$, we can show by induction that

$$\sum_{a=1}^{N^2-1} (T^a)_{\alpha\alpha} (T^a)_{\alpha\alpha} = \frac{1}{2} - \frac{1}{2N} \qquad (16)$$

We associate $T^a \equiv H_a$ here for convenience. To prove this, we will show that for any two possible choices of α (α and $\alpha + 1$), the two sums are equal. Consider the base case, $\alpha = 1$. Every H_i has a 1 in the position 11, so we get

$$\sum_{a=1}^{N^2-1} (T^a)_{11} (T^a)_{11} = \sum_{a=2}^{N} \frac{1}{2a(a-1)}$$

For the case when $\alpha = 2$, we have

$$\sum_{a=1}^{N^2-1} (T^a)_{22} (T^a)_{22} = \frac{1}{4} + \sum_{a=3}^{N} \frac{1}{2a(a-1)}$$

since only H_1 has a term not equal to 1 in the position 22. These two expressions are clearly equal, which proves the base case. Assume that up to some choice $\alpha \geq 3$, the inductive hypothesis holds. We will now show that the sum for this α is equal to the sum when we use $\alpha + 1$.

There are two possible terms in the sum. Using the H_i with $i = j - 1$ as used in Equation 4, when $j = \alpha$, we have a term
$$\frac{(j-1)^2}{2j(j-1)}$$
and when $j > \alpha$, we have a term
$$\frac{1}{2j(j-1)}$$
and when $j < \alpha$, we get a zero. Writing out the sum for the specified α above, we get

$$\frac{(\alpha-1)^2}{2\alpha(\alpha-1)} + \sum_{a=\alpha+1}^{N} \frac{1}{2a(a-1)} = \frac{\alpha-1}{2\alpha} + \sum_{a=\alpha+1}^{N} \frac{1}{2a(a-1)} \tag{17}$$

The first term in Equation 17 comes from H_{j-1} with $j = \alpha$, and the sum from all H_{j-1} with $j > \alpha$. Using this expression (for a fixed $j = \alpha$), let us now consider the sum for for the choice $\alpha + 1$; our first term comes when $j = \alpha + 1$ and the others for $j \geq \alpha + 2$:

$$\frac{\alpha}{2(\alpha+1)} + \sum_{a=\alpha+2}^{N} \frac{1}{2a(a-1)} \tag{18}$$

Looking back at Equation 17, we can pull out the $a = \alpha + 1$ term from the sum, to get

$$\frac{\alpha-1}{2\alpha} + \frac{1}{2\alpha(\alpha+1)} + \sum_{a=\alpha+2}^{N} \frac{1}{2a(a-1)}$$

Combining terms we see

$$\frac{(\alpha-1)(\alpha+1)}{2\alpha(\alpha+1)} + \frac{1}{2\alpha(\alpha+1)} + \sum_{a=\alpha+2}^{N} \frac{1}{2a(a-1)} = \frac{\alpha}{2(\alpha+1)} + \sum_{\alpha+2}^{N} \frac{1}{2a(a-1)}$$

By the inductive hypothesis, then, the sum is equal for any choice of α. We can therefore choose α to make the sum easiest to compute: $\alpha = N$, for which all H_i have a zero except H_{N-1}. In this case, we only have one term:

$$\sum_{a=1}^{N^2-1} (T^a)_{NN} (T^a)_{NN} = (H_{N-1})_{NN}^2 = \frac{(N-1)^2}{2N(N-1)} = \frac{N-1}{2N} = \frac{1}{2} - \frac{1}{2N}$$

as needed. There is one other non-trivial case to consider.

$$\sum_{a=1}^{N^2-1} (T^a)_{\alpha\alpha} (T^a)_{\beta\beta}$$

In this case, we will only have contributions from the H_i. Our sum can be simplified, to

$$\sum_{a=1}^{N-1} (H_a)_{\alpha\alpha} (H_a)_{\beta\beta} = \sum_{j=2}^{N} (H_{j-1})_{\alpha\alpha} (H_{j-1})_{\beta\beta}$$

We again associate $a = j - 1$ as in Equation 4. When $j > \alpha$, we get a we get a contribution

$$\frac{1}{\sqrt{2j(j-1)}}$$

When $j = \alpha$, we get a term

$$\frac{1-j}{\sqrt{2j(j-1)}}$$

and when $j < \alpha$, we get no contribution. The same applies to the β terms. Assuming the case $\alpha \neq \beta$ (since we have already covered the equality case), we have three possibilities:

$$\frac{1}{2j(j-1)} \quad \text{when} \quad j > \alpha \quad \text{and} \quad j > \beta$$
$$-\frac{1}{2j} \quad \text{when} \quad \beta < j = \alpha \quad \text{or} \quad \alpha < j = \beta$$
$$0 \quad \text{when} \quad j < \alpha \quad \text{or} \quad j < \beta$$

For a given choice of α and β, we will get a sum involving all three types of terms. Assume $\alpha > \beta$. The sum will take the form

$$0 + \cdots + 0 - \frac{1}{2\alpha} + \sum_{j=\alpha+1}^{N} \frac{1}{2j(j-1)}$$

Consider the case where, instead of α, we have $\alpha + 1$. This sum will be

$$-\frac{1}{2(\alpha+1)} + \sum_{j=\alpha+2}^{N} \frac{1}{2j(j-1)}$$

We can rearrange this sum, adding and subtracting terms, to get

$$-\frac{1}{2(\alpha+1)} - \frac{1}{2\alpha(\alpha+1)} + \sum_{j=\alpha+1}^{N} \frac{1}{2j(j-1)} = \frac{-(\alpha+1)}{2\alpha(\alpha+1)} + \sum_{j=\alpha+1}^{N} \frac{1}{2j(j-1)}$$

It is clear that this sum takes on a constant value regardless of the choice of α, just like in the previously studied case. The full inductive proof follows immediately. We can therefore evaluate it using the most convenient choice of α, $\alpha = N$. In this case, we end up with a single term

$$\sum_{j=2}^{N} (H_{j-1})_{\alpha\alpha} (H_{j-1})_{\beta\beta} = -\frac{1}{2N}$$

We now must combine the three cases we studied, which are the only non-zero cases. (1) $\alpha = \beta', \alpha' = \beta$, we get a 1/2. (2) $\alpha = \beta, \alpha' = \beta'$, we get $-1/(2N)$. (3) $\alpha = \beta = \alpha' = \beta'$, we get a $1/2 - 1/(2N)$. Since the (3) is a sub-case of (2), we will have a contribution

$$-\frac{1}{2N}\delta_{\alpha\beta}\delta_{\alpha'\beta'}$$

in the final, general expression. We also note that (2) is a sub-case of (1). In fact, the 1/2 from (2) precisely covers the only cases of (1) where (1) disappears. We can combine these two terms, then, to add a consistent contribution

$$-\frac{1}{2}\delta_{\alpha\beta'}\delta_{\alpha'\beta}$$

regardless of if $\alpha = \alpha'$. This gives us the final expression

$$\sum_{a=1}^{N^2-1} \left(T^a\right)_{\alpha\beta} \left(T^a\right)_{\alpha'\beta'} = \frac{1}{2}\delta_{\alpha\beta'}\delta_{\alpha'\beta} - \frac{1}{2N}\delta_{\alpha\beta}\delta_{\alpha'\beta'} \tag{19}$$

This condition will be needed for proofs in later sections.

2.2 The Root System

Our interest in the root system lies in two aspects. Somewhat generally, the root system classifies the algebra entirely, and any two algebras with the same root system are isomorphic. In particular, we are interested in finding explicit forms of the simple roots; with these roots, we will be able to find the fundamental weights and classify the irreducible representations of the algebra.

For the base of the root system, we use the simple roots defined above (in fact, the simple roots are defined as being the elements of the base of the root system, but we introduced them instead by their properties). Finding the simple roots directly from the definition of the roots above is tedious but straight forward; given our generators, the simple roots are the α^i given by [3]

$$\alpha^i = \mu^i - \mu^{i+1}$$

The first few simple roots, then, are

$$\alpha^1 = \mu^1 - \mu^2 = (1, 0, 0, \ldots, 0)$$

$$\alpha^2 = \mu^2 - \mu^3 = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0, \ldots, 0\right)$$

$$\alpha^3 = \mu^3 - \mu^4 = \left(0, -\frac{1}{\sqrt{3}}, \frac{2}{\sqrt{6}}, 0, \ldots, 0\right)$$

The arbitrary mth simple root (for $m \neq 1$) is given by

$$\alpha^m = \mu^m - \mu^{m+1} = \left(0, \ldots, 0, \frac{1-m}{\sqrt{2m(m-1)}}, \frac{1+m}{\sqrt{2m(m+1)}}, 0, \ldots, 0\right) \tag{20}$$

10

where the first non-zero entry is in the $(m-1)$th component. For any simple root

$$\alpha^m \cdot \alpha^m = \frac{(1-m)^2}{2m(m-1)} + \frac{(1+m)^2}{2m(m+1)} = 1$$

From the simple roots, we can reconstruct the rest of the roots by acting on the simple roots with elements of the Weyl group W. Any root β can be found by [2]

$$\beta = w_\gamma(\alpha) \qquad w \in W$$

where α and γ are simple roots, and the Weyl group is the set transformations w_γ for all γ in the base that reflect a vector v about the hyperplane normal to the simple root γ (this reflection is done by treating the simple roots as vectors in Euclidean space). We can write this transformation explicitly, with

$$w_\gamma(v) = v - \frac{2(v,\gamma)}{(\gamma,\gamma)}\gamma \tag{21}$$

where (α, β) is the inner product induced by the Killing form

$$(\alpha, \beta) = \kappa(t_\alpha, t_\beta)$$

where t_α, t_β are defined as the unique elements in the Cartan subalgebra H such that

$$\kappa(t_\alpha, k) = \alpha(k) \quad \text{for all } k \in H$$

Remember that the simple root α maps elements of the Cartan subalgebra to scalars: $\alpha(k)$ is one of the components of the root α (see Equation 13). Now that we have the simple roots as a base for our root system and a method to derive the rest of the roots, we can find the Dynkin diagram and Cartan matrix, which specify our Lie algebra entirely. We define the quantity $\langle \alpha, \beta \rangle$ by

$$\langle \alpha, \beta \rangle \equiv \frac{2(\alpha, \beta)}{(\beta, \beta)} \tag{22}$$

Since our simple roots are Euclidean vectors in $\mathbb{R}^{N-1}$, the form $(-,-)$ in Equations 21 and 22 becomes the usual inner product. To draw our Dynkin diagram, we need to check the value of

$$\langle \alpha^i, \alpha^j \rangle \langle \alpha^j, \alpha^i \rangle$$

Since the simple roots all have magnitude 1 and our form $(-,-)$ is just the dot product, this becomes

$$\langle \alpha^i, \alpha^j \rangle \langle \alpha^j, \alpha^i \rangle = 4(\alpha^i \cdot \alpha^j)^2$$

This will only be non-zero when $j = i + 1$. In that case, we get

$$4(\alpha^i \cdot \alpha^{i+1})^2 = 4\left(\frac{1+i}{\sqrt{2i(i+1)}} \frac{-i}{\sqrt{2i(i+1)}}\right)^2 = 1$$

Since this is constant for any i, between each root α^i and α^{i+1}, we draw a single line. This gives us the Dynkin diagram below.

A root system with this diagram is called type A_n. We can further find the Cartan matrix, with ij element $\langle \alpha^i, \alpha^j \rangle$. However, we already calculated this. For $i = j$, we get 2, for
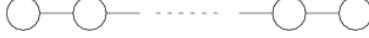
11

Figure 1: Dynkin Diagram for $\mathfrak{su}(N)$, of type A_n

$j = i + 1$ or $i = j + 1$, we get a -1, and for all other i, j we get 0. Our Cartan matrix is thus

$$
\begin{pmatrix}
2 & -1 & 0 & \dots & 0 & 0 \\
-1 & 2 & -1 & \dots & 0 & 0 \\
0 & -1 & 2 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & 2 & -1 \\
0 & 0 & 0 & \dots & -1 & 2
\end{pmatrix}
\tag{23}
$$

With this description of the root system, we have essentially classified $\mathfrak{su}(N)$ as an algebra in its entirety (the full statement as to why this gives a classification is beyond the level of this discussion). Furthermore, we now have all the necessary quantities to construct a classification of the representations of $\mathfrak{su}(N)$.

2.3   The Fundamental Weights

Irreducible representations of a Lie algebra are specified by their (unique) highest weight. With this highest weight and the simple roots, the entire algebra can be reconstructed in a lengthy but straightforward manner. In particular, if μ is the highest weight of an irreducible representation, then, for any simple root α_j

$$
\frac{2(\alpha^j, \mu)}{(\alpha^j, \alpha^j)} = \ell^j
\tag{24}
$$

The ℓ^j, for any given weight μ, are known as the Dynkin coefficients [3]. There is a unique set of highest weights μ_i that satisfy

$$
\frac{2(\alpha^j, \mu_i)}{(\alpha^j, \alpha^j)} = \delta_{ij}
\tag{25}
$$

These weights μ^i are the fundamental weights. Any highest weight of an irreducible representation μ can be written as

$$
\mu = \sum_{i=1}^{N-1} \ell^j \mu_j
$$

where the ℓ^j are the weight's Dynkin coefficients, as given in Equation 24. Looking at the form defined in Equation 24 with α^i, μ^i, we see

$$
\frac{2(\alpha^i, \mu_i)}{(\alpha^i, \alpha^i)} = 2 \left(\frac{1-i}{\sqrt{2i(i-1)}} (\mu_i)_{i-1} + \frac{1+i}{\sqrt{2i(i+1)}} (\mu_i)_i \right) = 1
$$

However, for α^{i-1}, we have

$$
\frac{2(\alpha^{i-1}, \mu_i)}{(\alpha^{i-1}, \alpha^{i-1})} = 2 \left(\frac{2-i}{\sqrt{2(i-2)(i-1)}} (\mu_i)_{i-2} + \frac{i}{\sqrt{2i(i-1)}} (\mu_i)_{i-1} \right) = 0
$$

12

The easiest way to satisfy this condition is to set μ_i to have the kth element non-zero only for $k \geq i$. This gives us

$$(\mu_i)_i = \sqrt{\frac{i}{2(i+1)}}$$

We then have

$$\frac{2(\alpha^{i+1}, \mu_i)}{(\alpha^{i+1}, \alpha^{i+1})} = \left(-\frac{i}{\sqrt{2i(i+1)}} \sqrt{\frac{i}{2(i+1)}} + \frac{2+i}{\sqrt{2(i+1)(i+2)}} (\mu_i)_{i+1} \right) = 0$$

We immediately see that

$$(\mu_i)_{i+1} = \frac{i}{\sqrt{2(i+1)(i+2)}}$$

We can repeat this process to find the full weight. However, we note that the expression [3]

$$\mu_i = \sum_{j=1}^{i} \mu^j \tag{26}$$

for weights μ^j of the defining representation (given in Equation 14) matches the properties we need. It is straightforward to check that the weights given in Equation 26 do indeed satisfy the condition of Equation 25 and are thus the set of fundamental weights.

2.3.1 Tensor Product Representations

Since any given irreducible representation is uniquely specified by its highest weight, we can associate to each fundamental weight μ_i a fundamental representation R_i. Let R be the irreducible representation specified by the highest weight μ. We can build the representation R by taking tensor products of ℓ_1 of R_1, ℓ_2 of R_2, et cetera, where ℓ_i are the Dynkin coefficients of the highest weight μ. Since we have the fundamental weights explicitly, from Equation 26, and can reconstruct their representations R_i using the root system, we now can build every irreducible representation of the algebra by finding a set of Dynkin coefficients $(\ell_1, \ldots, \ell_{N-1})$ such that the corresponding tensor product of ℓ_1 copies of R_1, ℓ_2 of R_2 cannot be decomposed into a direct sum (a process we will discuss momentarily):

$$R = R_1 \otimes R_1 \otimes \cdots \otimes R_{N-1} \tag{27}$$

Now that we have classified all irreducible representations of the algebra, we have, albeit somewhat briefly, classified every representation of the algebra.

2.4 Useful Tools

2.4.1 Tensor Methods

Given the direct relationship between tensor products of the fundamental representations and the irreducible representations of the algebra, we will briefly mention some of the properties of one of the classification tools used later.

One method to analyze representations of a Lie algebra is tensors. We can construct a Hilbert space on which representations of the algebra act, associate states of the Hilbert space to weights of the representations, and analyze the transformations of these states under representation elements.

2.4.2 A Brief Example

Because we are not going to explain in full detail the structure behind the tensor method, we will give a short example in $\mathfrak{su}(3)$ [1]. The first fundamental representation of $\mathfrak{su}(3)$ is called **3**, and is in fact the defining representation. For $\mathfrak{su}(3)$, it is convention to specify representations by their dimension, thus the name **3** since it acts on a three dimensional space. The weights (states) of **3** are

$$\left(\frac{1}{2}, \frac{\sqrt{3}}{6}\right) \equiv v_1$$

$$\left(-\frac{1}{2}, \frac{\sqrt{3}}{6}\right) \equiv v_2$$

$$\left(0, -\frac{1}{\sqrt{3}}\right) \equiv v_3$$

The Dynkin coefficients are (1,0).

We can thus consider some state of a tensor product of two copies of **3**, $\mathbf{3} \otimes \mathbf{3}$. We can take some state in the product representation, $v_i u_j$, where both v_i and u_j are one of the three states above. This state is neither symmetric nor antisymmetric in the indices. We can, however, manipulate the tensor

$$v_i u_j = \frac{1}{2}\left(v_i u_j + v_j u_i\right) + \frac{1}{2}\left(v_i u_j - v_j u_i\right)$$

$$= \frac{1}{2}\left(v_i u_j + v_j u_i\right) + \frac{1}{2}\varepsilon_{ijk}\varepsilon^{k\ell m} v_\ell u_m$$

turning the element $v_i u_j$ into a sum of a symmetric tensor with an antisymmetric tensor. Without providing excessive detail, the first term is a state of the representation **6**, and the second of the representation $\overline{\mathbf{3}}$, which is the other fundamental representation (this can be shown by explicitly examining the states written, as the product of two states is found by adding the respective weights; one will quickly see that the states found from the first term are exactly the states of the representation **6**, and similarly for the second, states of $\overline{\mathbf{3}}$). Thus, we see that the product states of $\mathbf{3} \otimes \mathbf{3}$ transform as the direct sum of something which transforms as **6** and something which transforms as $\overline{\mathbf{3}}$, which gives us:

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \overline{\mathbf{3}}$$

Note the dimensionality: the tensor product has dimension given by the product of the dimensions, and the direct sum by the sum. This relationship holds for any such decomposition.

2.4.3 Young Tableaux

Young tableaux are diagrams which are in one-to-one correspondence with the conjugacy classes of a group, and thus with the irreducible representations. We can therefore use them to classify irreducible representations.

We make a Young tableau for an irreducible representation by taking the Dynkin coefficients and constructing a diagram with ℓ_1 columns with one row each, ℓ_2 columns of two rows each, et cetera. To get an explicit state from the tableau, we place in each box an index and then construct a tensor that is symmetric with respect to indices in the same row and antisymmetric with respect to indices in the same column (we also normalize the tensors if necessary). We can product and sum tableaux just like representations. Furthermore, fundamental representations are given by tableau of a single column (since for a fundamental representation, only one Dynkin coefficient is non-zero).

For our example above, we could make the Young tableau

$$\mathbf{3} \otimes \mathbf{3} = \square \otimes \square = \boxed{} \oplus \begin{array}{c}\square\\\square\end{array}$$

We can find a state from the tableau by placing indices in the boxes.

$$\boxed{i} \otimes \boxed{j} = u_i v_j = \boxed{i \mid j} \oplus \begin{array}{c}\boxed{i}\\\boxed{j}\end{array} = (v_i u_j + v_j u_i) + (v_i u_j - v_j u_i)$$

$$\rightarrow \frac{1}{2}(v_i u_j + v_j u_i) + \frac{1}{2}(v_i u_j - v_j u_i)$$

$$= \frac{1}{2}(v_i u_j + v_j u_i) + \frac{1}{2}\varepsilon_{ijk}\varepsilon^{k\ell m}v_\ell u_m$$

as we saw before, where we normalized the tensor in the second line by multiplying by a constant. As a general example, if we had some (irreducible) representation of the algebra given by Dynkin coefficients (1,3,2,1), we would have the tableau
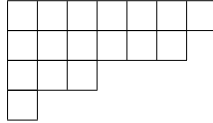
The process to decompose a product of Young tableaux into a direct sum is much more straightforward than directly manipulating tensor indices. However, we will not examine it, as we have already shown our main point: that to each irreducible representation with some highest weight (and thus some set of Dynkin coefficients), we can draw a unique Young tableau, and vice versa. Furthermore, this relationship generalizes to characters of the group: to each character of an irreducible representation, we can associate a Young tableau. From this tableau, we can find the character's corresponding representation and reconstruct its entire algebra, allowing for a complete classification.

3 Characters of SU(N)

With a complete classification scheme for representations of the algebra and group, we can now turn our attention to our main goal: finding the characters of $\mathbf{SU}(N)$. We will go

about this by two different methods: the Weyl character formula and the Casimir operator. However, via each method, we will see a strong correspondence between the wave functions of free fermions, characters of the group, and the Dynkin coefficients.

3.1   Preliminaries

Before looking at characters, we establish a few basic definitions. We remember our definition of $\mathbf{SU}(\mathrm{N})$ from the first section

$$\mathbf{SU}(\mathrm{N}) = \{A \in \mathrm{GL_N}(\mathbb{C}); \det A = 1, A^\dagger = A^{-1}\} \tag{28}$$

In particular, $U \in \mathbf{SU}(\mathrm{N})$ is a unitary matrix and thus diagonalizable. We can therefore make a similarity transformation on some arbitrary U

$$U = V^\dagger \Omega V$$

where Ω is a diagonal matrix, with elements ω_i, the eigenvalues of U, on its diagonal. We also note that the determinant of U must be 1. This gives us the condition on the ω_i

$$\prod_{i=1}^{N} \omega_i = 1 \longrightarrow \omega_k = e^{i\varphi_k} \quad \text{such that} \quad \sum_{k=1}^{N} \varphi_k = 2\pi m$$

for some integer m. Since we are working with characters, we will be looking at traces of elements of $\mathbf{SU}(\mathrm{N})$. Traces are permutation invariant, so we get

$$\mathrm{Tr}(V^\dagger \Omega V) = \mathrm{Tr}\,\Omega$$

Thus, the trace of some arbitrary representation will be a function of the ω_i only. Since we now have that the trace is a permutation invariant function of the ω_i, we also have the trace can be written as some combination of the invariant polynomials W_i

$$W_1 = \mathrm{Tr}\,U = \omega_1 + \cdots + \omega_N$$
$$W_2 = \mathrm{Tr}\,U^2 = \omega_1^2 + \cdots + \omega_N^2$$
$$\vdots$$
$$W_N = \mathrm{Tr}\,U^N = \omega_1^N + \cdots + \omega_N^N$$

3.2   The Weyl Character Formula

The Weyl Character formula, in general, specifies the characters for any representation of a semisimiple finite Lie algebra, but that is far beyond the reach of this discussion. In particular, for $\mathbf{SU}(\mathrm{N})$, it can be shown that the formula simplifies to [1]

$$\chi_\ell(U) = \frac{1}{\Delta}[\omega^{\ell_1}, \omega^{\ell_2}, \dots, \omega^{\ell_N}] \tag{29}$$

Where we define the quantity $[\omega^{\ell_1}, \omega^{\ell_2}, \ldots, \omega^{\ell_N}]$ as the totally antisymmetric function

$$[\omega^{\ell_1}, \omega^{\ell_2}, \ldots, \omega^{\ell_N}] \equiv \begin{vmatrix} \omega_1^{\ell_1} & \omega_2^{\ell_1} & \cdots & \omega_N^{\ell_1} \\ \omega_1^{\ell_2} & \omega_2^{\ell_2} & \cdots & \omega_N^{\ell_2} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_1^{\ell_N} & \omega_2^{\ell_N} & \cdots & \omega_N^{\ell_N} \end{vmatrix}$$

with the ω_i being the eigenvalues of U. The function Δ is given by the same form, but with $\ell_i \to N - i$. We require $\ell_N = 0$.

Given the rather lengthy discussion to classify representations of the algebra, it seems rather anti-climactic that we should find the complete set of group characters so quickly. However, the characters have interesting non-trivial properties; most notably, with some manipulation, we can relate them directly to fermionic wave functions. If we multiply by

$$e^{i\delta} \quad \text{with} \quad \delta = -\frac{N-1}{2}(\varphi_1 + \cdots + \varphi_N)$$

which is identically one since we have required the sum of φ_i to be an integer multiple of 2π, then we see that Δ becomes

$$\Delta = [e^{i(N-1)\varphi}, \ldots, e^{i0\varphi}] \to [e^{i(N-1)\varphi/2}, e^{i(N-3)\varphi/2}, \ldots, e^{-i(N-1)\varphi/2}]$$

which is exactly the ground state of N non-interacting fermions constrained to single dimension (in particular, constrained to a circle). Consider also the numerator of Equation 29. First, define

$$\lambda_i = \ell_i - N + i$$

In terms of these λ_i, the numerator is

$$[\omega^{N-1+\lambda_1}, \omega^{N-2+\lambda_2}, \ldots, \omega^{\lambda_N}]$$

Adding the phase and writing as exponentials, we get

$$\left[e^{i\left(\frac{N-1}{2}+\lambda_1\right)\varphi}, e^{i\left(\frac{N-3}{2}+\lambda_2\right)\varphi}, \ldots, e^{i\left(-\frac{N-1}{2}+\lambda_N\right)\varphi} \right]$$

and we see that this function corresponds to a set of fermions with excitations λ_i. A natural question, then, is to check the energies of such wave functions, which is given by the sum of the squares of the exponents. The energy of the gound state is

$$E_{GS} = \left(\frac{N-1}{2}\right)^2 + \left(\frac{N-3}{2}\right)^2 + \cdots + \left(-\frac{N-1}{2}\right)^2 = \sum_{k=1}^{N} \left(\frac{N-2k+1}{2}\right)^2$$

The energy of the excited state, then, is given by

$$E = \left(\frac{N-1}{2}+\lambda_1\right)^2 + \left(\frac{N-3}{2}+\lambda_2\right) + \cdots = \sum_{k=1}^{N} \left(\frac{N-2k+1}{2}+\lambda_k\right)^2$$

Since the character itself is given as the quotient of the excited state and the ground state, we can think of the energy of the character's "state" as the difference in energies. We can thus write

$$E_\chi = E - E_{GS} \tag{30}$$

Furthermore, it turns out that the Young tableau corresponding to the character (and its representation) can be given by a tableau with λ_1 boxes in the first row, λ_2 boxes in the second, et cetera; this correspondence allows us to associate the character with the representations of the algebra examined previously.

3.2.1 Examples

Since we did not derive carefully the above claims (notably, the Weyl character formula itself) as to do so would far surpass the reach of this discussion, we will give a few examples to demonstrate.

Consider a character with $\ell = (\ell_1, \ell_2, \ell_3) = (3, 1, 0)$. We are in $\mathbf{SU}(3)$, and we have $\lambda_1 = 1$, $\lambda_2 = 0$, and $\lambda_3 = 0$. Its Young tableau is given by

$$\square$$

The ground state is given by the determinant

$$
\begin{vmatrix} \omega_1^2 & \omega_2^2 & \omega_3^2 \\ \omega_1 & \omega_2 & \omega_3 \\ 1 & 1 & 1 \end{vmatrix} = (\omega_1 - \omega_2)(\omega_2 - \omega_3)(\omega_1 - \omega_3)
$$

$$
= \omega_1^2\omega_2 - \omega_1^2\omega_3 + \omega_1\omega_3^2 - \omega_1\omega_2^2 - \omega_2\omega_3^2 + \omega_3\omega_2^2
$$

The excited state is given by

$$
\begin{vmatrix} \omega_1^3 & \omega_2^3 & \omega_3^3 \\ \omega_1 & \omega_2 & \omega_3 \\ 1 & 1 & 1 \end{vmatrix} = \omega_1^3\omega_2 - \omega_1^3\omega_3 + \omega_1\omega_3^3 - \omega_1\omega_2^3 - \omega_2\omega_3^3 + \omega_3\omega_2^3
$$

The character is then given by

$$
\chi = \frac{\omega_1^3\omega_2 - \omega_1^3\omega_3 + \omega_1\omega_3^3 - \omega_1\omega_2^3 - \omega_2\omega_3^3 + \omega_3\omega_2^3}{\omega_1^2\omega_2 - \omega_1^2\omega_3 + \omega_1\omega_3^2 - \omega_1\omega_2^2 - \omega_2\omega_3^2 + \omega_3\omega_2^2} = \omega_1 + \omega_2 + \omega_3
$$

It is simple to check this calculation. Thus, the Weyl character gives the expected result for the defining representation.

Consider a character with $\ell = (4, 1, 0)$ and $\lambda = (2, 0, 0)$. This character has tableau

$$\square\square$$

Our excited state this time is given by

$$
\begin{vmatrix} \omega_1^4 & \omega_2^4 & \omega_3^4 \\ \omega_1 & \omega_2 & \omega_3 \\ 1 & 1 & 1 \end{vmatrix} = \omega_1^4\omega_2 - \omega_1^4\omega_3 + \omega_1\omega_3^4 - \omega_1\omega_2^4 - \omega_2\omega_3^4 + \omega_3\omega_2^4
$$

and we find the character to be

$$
\chi = \frac{\omega_1^4\omega_2 - \omega_1^4\omega_3 + \omega_1\omega_3^4 - \omega_1\omega_2^4 - \omega_2\omega_3^4 + \omega_3\omega_2^4}{\omega_1^2\omega_2 - \omega_1^2\omega_3 + \omega_1\omega_3^2 - \omega_1\omega_2^2 - \omega_2\omega_3^2 + \omega_3\omega_2^2}
$$

18

It is an exercise in algebra to show that the character is equal to

$$\chi = \omega_1^2 + \omega_2^2 + \omega_3^2 + \omega_1\omega_2 + \omega_2\omega_3 + \omega_1\omega_3 = \frac{1}{2}\left((\omega_1 + \omega_2 + \omega_3)^2 + \omega_1^2 + \omega_2^2 + \omega_3^2\right)$$

Clearly, for larger N, the characters become exceedingly difficult to explicitly calculate in this manner. Primarily, however, the Weyl method is valuable for the insight it offers into the relationship between free fermions and the characters of $\mathbf{SU}(\mathrm{N})$.

3.3 The Casimir Method

If we continue in the avenue of analyzing the characters as wavefunctions, the next natural step is to define a Hamiltonian and takes our wavefunctions as eigenfunctions and returns, as eigenvalues, their energies. Our wavefunctions are characters, so a immediate choice is the Casimir operator of the group.

The Casimir takes the characters of the group's representations as eigenvectors. The Casimir is defined [1]

$$\hat{\mathbf{C}} = 2\sum_{a=1}^{N^2-1}\hat{\mathbf{E}}^a\hat{\mathbf{E}}^a$$

$$\hat{\mathbf{E}}^a = \mathrm{Tr}\left(T^a U \frac{\partial}{\partial U}\right) = (T^a)_{\alpha\beta} U_{\beta\gamma}\frac{\partial}{\partial U_{\alpha\gamma}}$$

where the T^a are the generators of the algebra, specifically, the H and E_{ij}. We scale by a factor, for convenience, to get

$$\hat{\mathbf{E}}^a = (1 - 1/N)^{-\frac{1}{2}}\,\mathrm{Tr}\left(T^a U \frac{\partial}{\partial U}\right) = (1 - 1/N)^{-\frac{1}{2}}(T^a)_{\alpha\beta}U_{\beta\gamma}\frac{\partial}{\partial U_{\alpha\gamma}}$$

Consider the character of the defining representation of $\mathbf{SU}(\mathrm{N})$, which is just the trace of the matrix U:

$$\mathrm{Tr}\,U = U_{\tau\tau}$$

We can apply the Casimir operator to this quantity, which gives us

$$\hat{\mathbf{C}}\left(\mathrm{Tr}\,U\right) = 2\left(1 - 1/N\right)^{-1}(T^a)_{\alpha\beta}U_{\beta\gamma}\frac{\partial}{\partial U_{\alpha\gamma}}(T^a)_{\alpha'\beta'}U_{\beta'\gamma'}\frac{\partial}{\partial U_{\alpha'\gamma'}}U_{\tau\tau}$$

$$= 2\left(1 - 1/N\right)^{-1}(T^a)_{\alpha\beta}U_{\beta\gamma}\frac{\partial}{\partial U_{\alpha\gamma}}(T^a)_{\alpha'\beta'}U_{\beta'\gamma'}\delta_{\alpha'\tau}\delta_{\gamma'\tau}$$

$$= 2\left(1 - 1/N\right)^{-1}(T^a)_{\alpha\beta}U_{\beta\gamma}\frac{\partial}{\partial U_{\alpha\gamma}}(T^a)_{\tau\beta'}U_{\beta'\tau}$$

$$= 2\left(1 - 1/N\right)^{-1}(T^a)_{\alpha\beta}U_{\beta\gamma}(T^a)_{\tau\beta'}\delta_{\beta'\alpha}\delta_{\tau\gamma}$$

$$= 2\left(1 - 1/N\right)^{-1}(T^a)_{\alpha\beta}(T^a)_{\gamma\alpha}U_{\beta\gamma}$$

$$= 2\left(1 - 1/N\right)^{-1}\left(\frac{1}{2}\delta_{\alpha\alpha}\delta_{\gamma\beta} - \frac{1}{2N}\delta_{\alpha\beta}\delta_{\alpha\gamma}\right)U_{\beta\gamma}$$

$$= U_{\beta\beta} \to NU_{\beta\beta}$$

19

Where in the last step, we added a factor of N which accounts for the multiplicity in taking the derivatives. Thus, we see that a character of the defining representation is an eigenvector of the Casimir with eigenvalue N. We can use this type of analysis in general to find some very useful identities about joining and splitting the trace functions W_i which we defined previously [1].

$$\left(2\sum_a \hat{\mathbf{E}}^a \hat{\mathbf{E}}^a\right) W_n = n \sum_{m=0}^{n-1} W_m W_{n-m}$$

$$2\sum_a \left(\hat{\mathbf{E}}^a W_n\right)\left(\hat{\mathbf{E}}^a W_m\right) = nm W_{n+m}$$

We will now prove these two identities.

Consider the splitting operation, given by the first identity. Writing out the left hand side explicitly, we get

$$2\left(1-1/N\right)^{-1}\left(T^a\right)_{\mu\nu} U_{\nu\rho} \frac{\partial}{\partial U_{\mu\rho}}\left(T^a\right)_{\alpha\beta} U_{\beta\gamma} \frac{\partial}{\partial U_{\alpha\gamma}}(\underbrace{U_{a_1 a_2}\dots U_{a_i a_j}}_{\text{r elements}}\underbrace{U_{a_j a_k} U_{a_k a_l}\dots U_{a_n a_1}}_{\text{n-r elements}})$$

Where the last quantity in parentheses is the matrix element expansion of the trace, and the expression is summed over all repeated indices. Let there be r elements before $U_{a_j a_k}$ and $n-r$ elements after it (including $U_{a_j a_k}$ itself). We now act on $U_{a_j a_k}$ with $\partial/\partial U_{\alpha\gamma}$

$$2\left(1-1/N\right)^{-1}\left(T^a\right)_{\mu\nu} U_{\nu\rho} \frac{\partial}{\partial U_{\mu\rho}}\left(T^a\right)_{\alpha\beta} U_{\beta\gamma}(U_{a_1 a_2}\dots U_{a_i a_j} U_{a_k a_l}\dots U_{a_n a_1})\delta_{a_j\alpha}\delta_{a_k\gamma}$$

Applying the delta functions, we get

$$2\left(1-1/N\right)^{-1}\left(T^a\right)_{\mu\nu} U_{\nu\rho} \frac{\partial}{\partial U_{\mu\rho}}\left(T^a\right)_{a_j\beta} U_{\beta a_k}(U_{a_1 a_2}\dots U_{a_i a_j} U_{a_k a_l}\dots U_{a_n a_1})\delta_{a_j\alpha}\delta_{a_k\gamma}$$

We now apply the second derivative to another arbitrary element, $U_{a_r a_s}$

$$2\left(1-1/N\right)^{-1}\left(T^a\right)_{\mu\nu} U_{\nu\rho} \frac{\partial}{\partial U_{\mu\rho}}\left(T^a\right)_{a_j\beta} U_{\beta a_k} U_{a_1 a_2}\dots U_{a_i a_j}\underbrace{U_{a_k a_l}\dots U_{a_q a_r}}_{\text{b elements}} U_{a_r a_s}\underbrace{U_{a_s a_t}\dots U_{a_n a_1}}_{\text{(n-1)-r-b-1 elements}}$$

$$= 2\left(1-1/N\right)^{-1}\left(T^a\right)_{\mu\nu} U_{\nu\rho}\left(T^a\right)_{a_j\beta} U_{\beta a_k} U_{a_1 a_2}\dots U_{a_i a_j} U_{a_k a_l}\dots U_{a_q a_r} U_{a_s a_t}\dots U_{a_n a_1}\delta_{a_r\mu}\delta_{a_s\rho}$$

Applying the delta functions gives

$$2\left(1-1/N\right)^{-1}\left(T^a\right)_{a_r\nu} U_{\nu a_s}\left(T^a\right)_{a_j\beta} U_{\beta a_k} U_{a_1 a_2}\dots U_{a_i a_j} U_{a_k a_l}\dots U_{a_q a_r} U_{a_s a_t}\dots U_{a_n a_1}$$

Taking the sum over the repeated index a and applying the condition on the generators T^a, we get

$$U_{\nu a_s} U_{\beta a_k} U_{a_1 a_2}\dots U_{a_i a_j} U_{a_k a_l}\dots U_{a_q a_r} U_{a_s a_t}\dots U_{a_n a_1}\delta_{a_r\beta}\delta_{\nu a_j}$$

Applying the delta functions gives

$$U_{a_j a_s} U_{a_r a_k} U_{a_1 a_2}\dots U_{a_i a_j} U_{a_k a_l}\dots U_{a_q a_r} U_{a_s a_t}\dots U_{a_n a_1}$$

Rearranging, we get

$$(U_{a_1 a_2} \ldots U_{a_i a_j} U_{a_j a_s} U_{a_s a_t} \ldots U_{a_n a_1})(U_{a_k a_l} \ldots U_{a_q a_r} U_{a_r a_k})$$

The above equation is clearly the product of two traces (note how the indices repeat with neighboring elements and connect at the beginning and end, at a_1 and a_k). To see which W_m correspond to the above elements, we carefully count the indices. As defined before, we have r elements from $U_{a_1 a_2}$ to $U_{a_i a_j}$ inclusive.

$$(\underbrace{U_{a_1 a_2} \ldots U_{a_i a_j}}_{r \text{ elements}} U_{a_j a_s} \underbrace{U_{a_s a_t} \ldots U_{a_n a_1}}_{n\text{-}r\text{-}b\text{-}2 \text{ elements}})(\underbrace{U_{a_k a_l} \ldots U_{a_q a_r}}_{b \text{ elements}} U_{a_r a_k})$$

The first term therefore has $n - b - 1$ terms, and the second has $b + 1$ terms. Thus, the term gained from this choice of $U_{a_r a_s}$ and $U_{a_j a_k}$ is

$$W_{n-b-1} W_{b+1}$$

If we define $m = b + 1$, we get

$$W_{n-m} W_m$$

We will get a term for all possible choices of b, so we must sum over every value of b, which can range from 0 to n. However, by our method of indexing, the index $m = n + 1$ is equivalent to $m = 0$, so we must sum over m from 0 to $n - 1$. Our result now becomes

$$\sum_{m=0}^{n-1} W_{n-m} W_m$$

Which is almost the correct identity, but we have yet to account for multiplicity. Consider the arrangement of terms from before

$$2 \left(T^a\right)_{\mu\nu} U_{\nu\rho} \frac{\partial}{\partial U_{\mu\rho}} \left(T^a\right)_{\alpha\beta} U_{\beta\gamma} \frac{\partial}{\partial U_{\alpha\gamma}} (\underbrace{U_{a_1 a_2} \ldots U_{a_i a_j}}_{r \text{ elements}} \underbrace{U_{a_j a_k} U_{a_k a_l} \ldots U_{a_n a_1}}_{n\text{-}r \text{ elements}})$$

If we shift our selection regions over by one index, we get

$$2 \left(T^a\right)_{\mu\nu} U_{\nu\rho} \frac{\partial}{\partial U_{\mu\rho}} \left(T^a\right)_{\alpha\beta} U_{\beta\gamma} \frac{\partial}{\partial U_{\alpha\gamma}} (\underbrace{U_{a_2 a_3} \ldots U_{a_i a_j} U_{a_j a_k}}_{r \text{ elements}} \underbrace{U_{a_k a_l} \ldots U_{a_n a_1} U_{a_1 a_2}}_{n\text{-}r \text{ elements}})$$

If we then proceed with the exact same process, but instead deriving with respect to $U_{a_k a_l}$ first, and later with respect to $U_{a_s a_t}$ (in order to preserve the values of r and b), we get the exact same result of

$$W_{n-m} W_m$$

Thus, by shifting our indices, we get the same result. We can do n such shifts before we come back to the original choice of indices, and thus we end up with a multiplicity of n for any given term

$$W_{n-m} W_m$$

and our final result is therefore

$$n \sum_{m=0}^{n-1} W_{n-m} W_m$$

21

as needed.

For the second identity,

$$2\left(1-1/N\right)^{-1}\sum_{a}\left(\hat{\mathbf{E}}^{a}W_{n}\right)\left(\hat{\mathbf{E}}^{a}W_{m}\right)=nmW_{n+m}$$

we begin similarly, by writing the left hand side out explicitly

$$2\left(1-1/N\right)^{-1}\left(T^{a}\right)_{\mu\nu}U_{\nu\rho}\frac{\partial}{\partial U_{\mu\rho}}\underbrace{U_{a_1a_2}\dots U_{a_ia_j}U_{a_ja_k}U_{a_ka_l}\dots U_{a_na_1}}_{\text{n terms}}\times$$

$$\left(T^{a}\right)_{\alpha\beta}U_{\beta\gamma}\frac{\partial}{\partial U_{\alpha\gamma}}\underbrace{U_{a_qa_r}\dots U_{a_sa_t}U_{a_ta_u}U_{a_ua_w}\dots U_{a_{m+q}a_q}}_{\text{m terms}}$$

We then apply the derivatives to $U_{a_ja_k}$ and $U_{a_ta_u}$

$$2\left(1-1/N\right)^{-1}\left(T^{a}\right)_{\mu\nu}U_{\nu\rho}\underbrace{U_{a_1a_2}\dots U_{a_ia_j}U_{a_ka_l}\dots U_{a_na_1}}_{\text{n-1 terms}}\delta_{a_j\mu}\delta_{a_k\rho}\times$$

$$\left(T^{a}\right)_{\alpha\beta}U_{\beta\gamma}\underbrace{U_{a_qa_r}\dots U_{a_sa_t}U_{a_ua_w}\dots U_{a_{m+q}a_q}}_{\text{m-1 terms}}\delta_{a_t\alpha}\delta_{a_u\gamma}$$

Applying the delta functions, we get

$$2\left(1-1/N\right)^{-1}\left(T^{a}\right)_{a_j\nu}U_{\nu a_k}\underbrace{U_{a_1a_2}\dots U_{a_ia_j}U_{a_ka_l}\dots U_{a_na_1}}_{\text{n-1 terms}}\times$$

$$\left(T^{a}\right)_{a_t\beta}U_{\beta a_u}\underbrace{U_{a_qa_r}\dots U_{a_sa_t}U_{a_ua_w}\dots U_{a_{m+q}a_q}}_{\text{m-1 terms}}$$

And summing over a and using the condition on the generators, we get

$$U_{\nu a_k}\underbrace{U_{a_1a_2}\dots U_{a_ia_j}U_{a_ka_l}\dots U_{a_na_1}}_{\text{n-1 terms}}\times$$

$$U_{\beta a_u}\underbrace{U_{a_qa_r}\dots U_{a_sa_t}U_{a_ua_w}\dots U_{a_{m+q}a_q}}_{\text{m-1 terms}}\delta_{a_j\beta}\delta_{\nu a_t}$$

Finally, applying the delta functions, we arrive at

$$U_{a_ta_k}\underbrace{U_{a_1a_2}\dots U_{a_ia_j}U_{a_ka_l}\dots U_{a_na_1}}_{\text{n-1 terms}}\times$$

$$U_{a_ja_u}\underbrace{U_{a_qa_r}\dots U_{a_sa_t}U_{a_ua_w}\dots U_{a_{m+q}a_q}}_{\text{m-1 terms}}$$

Clearly, we have overall $n+m$ terms. Rearranging, we get

$$U_{a_qa_r}\dots U_{a_sa_t}U_{a_ta_k}U_{a_ka_l}\dots U_{a_na_1}U_{a_1a_2}\dots U_{a_ia_j}U_{a_ja_u}U_{a_ua_w}\dots U_{a_{m+q}a_q}$$

Here, once again, we have extracted the equation for a trace. In particular, we have

$$W_{m+n}$$

To find the multiplicity factor, we consider the choice of elements to derive. Clearly, had we chosen to derive any element besides $U_{a_j a_k}$ or $U_{a_t a_u}$ in the factors, we would have arrived at the same result. We therefore have n ways to get this result from W_n and m ways to get it from W_m, so overall we have multiplicity mn. Thus, the final result is the identity

$$2 \sum_a \left(\hat{\mathbf{E}}^a W_n \right) \left(\hat{\mathbf{E}}^a W_m \right) = nm W_{n+m}$$

We can use these identities to find eigenfunctions of the Casimir. A few low dimensional characters are (with corresponding Young tableau, specified as $(\lambda_1, \lambda_2, \dots)$)

$$\chi_{(1)} = \operatorname{Tr} U \quad \square$$

$$\chi_{(2)} = \frac{1}{2} \left[(\operatorname{Tr} U)^2 + \operatorname{Tr} U^2 \right] \quad \square\square$$

$$\chi_{(2,1)} = \frac{1}{3} \left[(\operatorname{Tr} U)^3 - \operatorname{Tr} U^3 \right]$$

$$\chi_{(2,2)} = \frac{1}{12} \left[(\operatorname{Tr} U)^4 - 4 \operatorname{Tr} U^3 \operatorname{Tr} U + 3 \left(\operatorname{Tr} U^2 \right)^2 \right]$$

Note that the first two cases precisely match the characters found from the Weyl character formula; our two methods of finding the characters, though differing in process, are equivalent.

Having shown a process to find the characters using the Casimir operator, and also how to find the eigenvalues of the characters, we can o relate the eigenvalues of a character specified by some general set of $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N)$ to the fermionic energies we established earlier.

The general formula for the eigenvalues of a character specified by $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N)$ with respect to the Casimir operator is [4]

$$\epsilon(\lambda_1, \dots, \lambda_N) = \lambda_1 (N - 1) + \lambda_1^2 + \lambda_2 (N - 3) + \lambda_2^2 + \cdots = \sum_{k=1}^{N} \left(\lambda_k^2 + \lambda_k (N - (2k - 1)) \right)$$

We can complete the square for each λ_k term to find

$$\epsilon = \sum_{k=1}^{N} \left(\lambda_k^2 + \lambda_k (N - (2k - 1)) \right)$$

$$= \sum_{k=1}^{N} \left(\left(\lambda_k + \frac{N - 2k + 1}{2} \right)^2 - \left(\frac{N - 2k + 1}{2} \right)^2 \right)$$

This is clearly equal exactly to the energies we found for the fermionic wave functions derived from the Weyl Character formula in Equation 30. We see now that we can fully treat the characters of $\mathbf{SU}(N)$ as the wave functions of one dimensional free fermions. We can take the Casimir operator of the group to be the Hamiltonian, and the eigenvalues of the characters as the energies. This correspondence suggests there is something inherently physical about the group $\mathbf{SU}(N)$, at the very least in application. One such application is non-abelian gauge field theory; in particular, for our discussion, Chern-Simons theory.

4 Yang-Mills Theory and Applications of SU(N)

Before we turn our attention to Chern-Simons theory, we will briefly touch on one significant and direct application of the group theory we have established. Consider first a Yang-Mills theory with Lagrangian density

$$\mathcal{L} = -\frac{1}{4} \operatorname{Tr} F_{\mu\nu} F^{\mu\nu} \tag{31}$$

where we use the non-abelian field strength (with coupling constant equal to one, for simplicity)

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$$

We take our gauge field A_μ to be elements of $\mathfrak{su}(N)$ and our gauge transformations to come from $\mathbf{SU}(N)$. We can expand our gauge field in terms of the generators of $\mathfrak{su}(N)$ as we established before, using

$$A_\mu \equiv A_\mu^a T^a$$

The generators T^a obey the trace constraint of Equation 8 and the completeness condition of Equation 19. If we specifically consider the case of $1 + 1$ dimensions, the Lagrangian reduces to

$$\mathcal{L} = -\frac{1}{2} \operatorname{Tr} F_{01} F^{01}$$

To put this in canonical form, we first note that there is only one canonical conjugate momentum, corresponding to A_1:

$$\Pi = \frac{\delta \mathcal{L}}{\delta \dot{A}_1} = -\frac{1}{2} \left(\dot{A}_1 - \partial_1 A_0 + [A_0, A_1] \right) \left(\dot{A}_1 - \partial_1 A_0 + [A_0, A_1] \right) = -F_{01}$$

where we are dropping the trace for now for convenience. We can thus write the Lagrangian as

$$\mathcal{L} = -\Pi \dot{A}_1 - \mathcal{H} + \lambda \mathcal{F} = \frac{1}{2} F_{01} F^{01} = -\frac{1}{2} \Pi^2$$

where $\mathcal{H}$ is the Hamiltonian density, and $\mathcal{F}$ is a constraint function, with λ a Lagrange multiplier. It is immediately clear that we can write the Lagrangian as

$$\mathcal{L} = -\Pi \dot{A}_1 + \frac{1}{2} \Pi^2 - \Pi \left(-\partial_1 A_0 + [A_0, A_1] \right)$$

Looking at the last term, we can rearragne to find

$$\begin{aligned} \Pi \left(-\partial_1 A_0 + [A_0, A_1] \right) &= -\Pi \partial_1 A_0 + \Pi A_0 A_1 - \Pi A_1 A_0 \\ &= A_0 \partial_1 \Pi + A_0 A_1 \Pi - A_0 \Pi A_1 \\ &= A_0 \left(\partial_1 \Pi + [A_1, \Pi] \right) \end{aligned}$$

where we remember that we are in a trace and thus free to cyclically permute independent terms. Our Lagrangian density becomes

$$\mathcal{L} = -\Pi \dot{A}_1 + \frac{1}{2} \Pi^2 - A_0 \left(\partial_1 \Pi + [A_1, \Pi] \right)$$

We recognize now our canonical variables Π and A_1, the Hamiltonian density $\frac{1}{2}\Pi^2$, and our constraint $\partial_1\Pi + [A_1, \Pi]$ with Lagrange multiplier A_0. We restore the trace and rewrite our Lagrangian in terms of the generators and coefficients to get

$$\mathcal{L} = \text{Tr}\left(-\Pi^a T^a \dot{A}_1^b T^b + \frac{1}{2}\Pi^a T^a \Pi^b T^b - A_0^a T^a \left(\partial_1 \Pi^b T^b + A_1^b \Pi^c [T^b, T^c]\right)\right)$$

$$= -\frac{1}{2}\Pi^a \dot{A}_1^a + \frac{1}{4}\left(\Pi^a\right)^2 - \frac{1}{2}A_0^a \left(\partial_1 \Pi^a + A_1^b \Pi^c f^{abc}\right)$$

Our canonical variables are thus $N^2 - 1$ fields A_1^a and Π^a. With a coupling constant set to one, our Poisson bracket will be a delta function of position, so we can define our commutator for a quantum theory by

$$[A_1^a(x), \Pi^b(y)] = i\delta_{ab}\delta(x - y)$$

We can thus associate our momenta Π^a to a functional derivative with respect to the field A_1

$$\Pi^a \equiv \frac{1}{i}\frac{\delta}{\delta A_1^a(x)}$$

Our Hamiltonian can thus be written as a second order differential operator

$$\mathcal{H} = -\frac{1}{4}\left(\frac{1}{i}\frac{\delta}{\delta A_1^a}\right)\left(\frac{1}{i}\frac{\delta}{\delta A_1^a}\right)$$

However, we can use a change of variables. If we define an arbitrary element U of $\mathbf{SU}$(N) as $\exp(iA) \rightarrow \exp(iA_1^a T^a + \text{independent of } A_1)$, we see

$$\frac{1}{i}\frac{\delta}{\delta A_1^a} = \text{Tr}\left(\frac{1}{i}\frac{\delta U}{\delta A_1^a}\frac{\delta}{\delta U}\right) = \text{Tr}\left(T^a U \frac{\delta}{\delta U}\right)$$

where we take a trace to eliminate the matrix degrees of freedom. In terms of these group variables, our Hamiltonian takes the form

$$\mathcal{H} = -\frac{1}{4}\text{Tr}\left(T^a U \frac{\delta}{\delta U}\right)\text{Tr}\left(T^a U \frac{\delta}{\delta U}\right)$$

However, we recognize this Hamiltonian as being precisely the Casimir of $\mathbf{SU}$(N), scaled by a constant. We can immediately see, therefore, that the eigenstates of the $1+1$ dimensional Yang-Mills theory will come from the characters of $\mathbf{SU}$(N). We can, however, establish this a bit more formally.

The constraint function $\mathcal{F}$ effectively serves to enforce gauge invariance. Since any physical states in this theory must obey the constraint, if we can construct our states in terms of manifestly gauge invariant variables, then the constraint will be automatically satisfied. To this end, we can phrase our theory in terms of *path ordered phase factors* [5]. This factor is defined by

$$U \equiv P_{xy}\exp\left(i\int_x^y A_\mu dz\right)$$

Under a gauge transformation, the field A_μ transforms as

$$A_\mu \rightarrow g^{-1}A_\mu g + g^{-1}\partial_\mu g = g^{-1}\left(A_\mu g + \partial_\mu g\right)$$

Plugging this transformation into our phase factor, we see that it transforms as

$$U \to U' = \exp\left(i \int g^{-1}\left(A_\mu g + \partial_\mu g\right) dx\right)$$

If we expand this as a power series, we see we will get terms of the form

$$\int \prod_i dx_i \left(g^{-1}\left(A_\mu g + \partial_\mu g\right)\right)^n$$

The clearly, the object in parenthesis, when raised to the nth power, will simplify to

$$\left(g^{-1}\left(A_\mu g + \partial_\mu g\right)\right)^n = g^{-1}\left(A_\mu + \partial_\mu g g^{-1}\right)^{n-1}\left(A_\mu g + \partial_\mu g\right)$$

We can make two adjustments to our phase factors that will immediately give us the gauge invariant variables we desire. First, we will consider only closed (spatial) paths. This is equivalent in the $1+1$ dimensional case to requiring periodic boundary conditions, or rather constraining the system to a circle. Secondly, we will take the trace of the phase factor. Taking this trace allows us to cyclically permute our terms, which simplifies our expression to

$$g^{-1}\left(A_\mu + \partial_\mu g g^{-1}\right)^{n-1}\left(A_\mu g + \partial_\mu g\right) = \left(A_\mu + \partial_\mu g g^{-1}\right)^n$$

And thus, our phase factor becomes

$$\operatorname{Tr} U' = \operatorname{Tr}\exp\left[i \oint dx\left(A_\mu + \partial_\mu g g^{-1}\right)\right]$$

If we consider an infinitesimal gauge transformation, with $g \approx 1 + i\varepsilon\alpha$, for $\alpha \in \mathfrak{su}(N)$ and $\varepsilon \ll 1$, we have

$$\operatorname{Tr} U' = \operatorname{Tr}\exp\left[i \oint dx\left(A_\mu + \partial_\mu(1 + i\varepsilon\alpha)(1 - i\varepsilon\alpha)\right)\right] = \operatorname{Tr}\exp\left[i \oint dx\left(A_\mu + i\varepsilon\partial_\mu\alpha\right)\right]$$

But the remaining gauge contribution is a total derivative, so when we integrate over a closed loop, it will not contribute. We see now that these "trace paths over closed loops" are manifestly gauge invariant. This gives us a set of gauge invariant variables given by the $\mathbf{SU}(N)$ group element generated by integrating the gauge field over closed paths. These quantities are called *Wilson Loops*:

$$W = \operatorname{Tr}\left[\exp\left(i \oint A dx\right)\right]$$

Since we have constrained our theory to a circle, the only closed paths will be those which wind around the circle an integer number of times. Thus, the complete set of Wilson loops will be given by

$$W_n = \operatorname{Tr}\left[\exp\left(i \oint A dx\right)\right]^n = \operatorname{Tr} U^n$$

where we note that the exponentiation of the gauge field gives an element of the gauge group. By now, the correspondence between this formulation of the $1+1$ dimensional Yang-Mills on a circle and the study of characters of $\mathbf{SU}(N)$ should be abundantly clear. To write

our Hamiltonian in terms of these variables, we note that the $\hat{\mathbf{E}}^a$ operators are differential operators in terms of U. We can make a change of variables by

$$\hat{\mathbf{E}}^a = T^a_{ij} U_{jk} \frac{\partial}{\partial U_{ik}} = T^a_{ij} U_{jk} \frac{\partial W_n}{\partial U_{ik}} \frac{\partial}{\partial W_n} = \left(\hat{\mathbf{E}}^a W_n \right) \frac{\partial}{\partial W_n}$$

This change of variables immediately gives us the Hamiltonian [5]

$$\mathcal{H} = \sum_a \left[\hat{\mathbf{E}}^a \left(\sum_n \hat{\mathbf{E}}^a W_n \frac{\partial}{\partial W_n} \right) \right]$$

$$= \sum_a \left[\sum_n \left(\hat{\mathbf{E}}^a \hat{\mathbf{E}}^a W_n \frac{\partial}{\partial W_n} \right) + \sum_{n,m} \left(\hat{\mathbf{E}}^a W_m \hat{\mathbf{E}}^a W_n \frac{\partial}{\partial W_n} \frac{\partial}{\partial W_n} \right) \right]$$

where we have dropped the factor of $-\frac{1}{4}$ for convenience, as it does not affect the states of the system, which are our primary concern. Here, we note we can explicitly use the identities we proved previously concerning the joining and splitting of the loops:

$$\left(2 \sum_a \hat{\mathbf{E}}^a \hat{\mathbf{E}}^a \right) W_n = n \sum_{m=0}^{n-1} W_m W_{n-m}$$

$$2 \sum_a \left(\hat{\mathbf{E}}^a W_n \right) \left(\hat{\mathbf{E}}^a W_m \right) = nm W_{n+m}$$

Using these identities, our Hamiltonian becomes

$$\mathcal{H} = \sum_n \sum_{m=0}^{n-1} n W_m W_{n-m} \frac{\partial}{\partial W_n} + \sum_{n,m} nm W_{n+m} \frac{\partial}{\partial W_n} \frac{\partial}{\partial W_m}$$

again, dropping unnecessary scaling constants. We can furthermore treat the loops W_n as creation operators and the derivative with respect to the loops as annihilation operators

$$W_n \to \sqrt{n} a_n^\dagger \qquad \frac{\partial}{\partial W_n} \to \frac{1}{\sqrt{n}} a_n$$

which gives us the Hamiltonian [5]

$$\mathcal{H} = N \sum_{n \neq 0} |k_n| a_n^\dagger a_n + \frac{1}{2} \sum_{\substack{n,\, m\, >\, 0 \\ n,\, m\, <\, 0}} \sqrt{nm|n+m|} \left(a_n^\dagger a_m^\dagger a_{n+m} + a_{n+m}^\dagger a_n a_m \right)$$

the physical states of which, as we already know, are the characters of $\mathbf{SU}(\text{N})$. A few of the states are, in terms of the creation operators $a^\dagger$,

$$|1\rangle = a_1^\dagger |0\rangle$$

$$|2\rangle = \left[\left(a_1^\dagger \right)^2 + a_2^\dagger \right] |0\rangle$$

$$|2,1\rangle = \left[\left(a_1^\dagger \right)^3 - a_3^\dagger \right] |0\rangle$$

27

where $|0\rangle$ gives the vacuum state, and the states are specified by $(\lambda_1, \lambda_2, \dots)$ which gives the Young tableau as before. These states, as we already showed, are fermionic states corresponding to free fermions on a circle. The $1+1$ dimensional Yang-Mills constrained to a circle is therefore equivalent to the problem of free fermions on a circle.

This method, of phrasing the problem in terms of the algebraic structure of the gauge field and the topological properties of the space on which we constrain our system (i.e. the loops) will become more challenging, but equally insightful as we turn our attention to Chern-Simons theory: first on a disc (with one type of loop) and then on a torus (with two types of loops).

5 Chern-Simons Theory

5.1 Abelian Chern-Simons Theory

Having now established (and perhaps gained some appreciation for) the classical structure of $\mathfrak{su}(N)$, as well as established a direct physical application of the theory, we can turn towards our main goal: Chern-Simons theory. Chern-Simons theory is a gauge invariant theory that is unique to $2+1$ dimensions, allowing for interactions between a physical current and gauge field that differs from standard Maxwell theory. When elevated to a non-abelian theory, it takes elements of $\mathfrak{su}(N)$ for its gauge field A_μ, connecting it directly to our previous discussion.

To study the canonical structure of non-abelian Chern-Simons theory, we will first study the simpler abelian case. Remembering our familiar Maxwell theory, we have a Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - A_\mu J^\mu \tag{32}$$

where we represent our current by J^μ and gauge field by A_μ, as is usual. The field strength tensor $F_{\mu\nu}$ is defined in the usual way

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

which corresponds to the familiar equation of motion

$$\partial_\mu F^{\mu\nu} = J^\nu$$

However, Chern-Simons theory is specified by a different Lagrangian, in terms of the same gauge field A_μ and current J^μ

$$\mathcal{L} = \frac{\kappa}{2} \varepsilon^{\mu\nu\rho} A_\mu \partial_\nu A_\rho - A_\mu J^\mu \tag{33}$$

This equation can also be written in an alternate form

$$\mathcal{L} = \frac{\kappa}{4} \varepsilon^{\mu\nu\rho} A_\mu \left(\partial_\nu A_\rho + \partial_\nu A_\rho \right) = \frac{\kappa}{4} \varepsilon^{\mu\nu\rho} A_\mu \left(\partial_\nu A_\rho - \partial_\rho A_\nu \right) = \frac{\kappa}{4} \varepsilon^{\mu\nu\rho} A_\mu F_{\nu\rho} \tag{34}$$

28

Under an infinitesimal shift in the gauge field, $A_\mu \rightarrow A_\mu + \delta A_\mu$, we see that the source free Lagrangian changes by

$$\mathcal{L} + \delta\mathcal{L} = \frac{\kappa}{2}\varepsilon^{\mu\nu\rho}\left(A_\mu + \delta A_\mu\right)\partial_\nu\left(A_\rho + \delta A_\rho\right)$$

$$= \frac{\kappa}{2}\varepsilon^{\mu\nu\rho}\left(A_\mu\partial_\nu A_\rho + \delta A_\mu\partial_\nu A_\rho + A_\mu\partial_\nu\delta A_\rho\right)$$

$$\text{integration by parts:} \quad A_\mu\partial_\nu\delta A_\rho = -\delta A_\rho\partial_\nu A_\mu$$

$$= \frac{\kappa}{2}\varepsilon^{\mu\nu\rho}\left[A_\mu\partial_\nu A_\rho + \delta A_\mu\left(\partial_\nu A_\rho - \partial_\rho A_\nu\right)\right]$$

$$\delta\mathcal{L} = \frac{\kappa}{2}\varepsilon^{\mu\nu\rho}\delta A_\mu F_{\nu\rho}$$

If we include the source, this equation takes the form

$$\delta\mathcal{L} = \frac{\kappa}{2}\varepsilon^{\mu\nu\rho}\delta A_\mu\left(F_{\nu\rho} + J^\mu\right)$$

By requiring that the variation in the action disappears for a physical path (and thus that the variation in the Lagrangian disappears as well), we recover the equation of motion

$$\frac{\kappa}{2}\varepsilon^{\mu\nu\rho}F_{\nu\rho} = J^\mu$$

We can also find the equations of motion using the Euler-Lagrange equations.

$$\frac{\partial\mathcal{L}}{\partial A_\lambda} - \partial_\sigma\frac{\partial\mathcal{L}}{\partial(\partial_\sigma A_\lambda)} = 0$$

$$\text{with} \quad \frac{\partial\mathcal{L}}{\partial A_\lambda} = \frac{\kappa}{2}\varepsilon^{\lambda\nu\rho}\partial_\nu A_\rho - J^\lambda$$

$$\text{and} \quad \frac{\partial\mathcal{L}}{\partial(\partial_\sigma A_\lambda)} = \frac{\kappa}{2}\varepsilon^{\mu\sigma\lambda}A_\mu \longrightarrow \partial_\sigma\frac{\partial\mathcal{L}}{\partial(\partial_\sigma A_\lambda)} = \frac{\kappa}{2}\varepsilon^{\mu\sigma\lambda}\partial_\sigma A_\mu = \frac{\kappa}{2}\varepsilon^{\lambda\nu\rho}\partial_\rho A_\nu$$

where in the last line, we used the antisymmetry of the Levi-Civita symbol twice, and renamed the repeated indices (as we are free to do). Combining, this gives us the final equation of motion

$$\frac{\kappa}{2}\varepsilon^{\lambda\nu\rho}\left(\partial_\nu A_\rho - \partial_\rho A_\nu\right) = \frac{\kappa}{2}\varepsilon^{\lambda\nu\rho}F_{\nu\rho} = J^\lambda \tag{35}$$

Looking at the form of the Lagrangian, it is apparent that it is not directly gauge invariant. However, under a gauge transformation $A_\mu \rightarrow A_\mu + \partial_\mu\Lambda$, we see

$$\mathcal{L} + \delta\mathcal{L} = \frac{\kappa}{4}\varepsilon^{\mu\nu\rho}\left(A_\mu + \partial_\mu\Lambda\right)F_{\nu\rho} = \mathcal{L} + \frac{\kappa}{4}\varepsilon^{\mu\nu\rho}\partial_\mu\Lambda F_{\nu\rho} = \frac{\kappa}{2}\partial_\mu\Lambda\partial_\nu A_\rho$$

But if we note that

$$\partial_\mu\left(\varepsilon^{\mu\nu\rho}\Lambda\partial_\nu A_\rho\right) = \varepsilon^{\mu\nu\rho}\partial_\mu\Lambda\partial_\nu A_\rho + \varepsilon^{\mu\nu\rho}\Lambda\partial_\mu\partial_\nu A_\rho$$

And that, because of the antisymmetry of the Levi-Civita symbol (and commutivity of partial derivatives), the second term is zero. We now see we can write the change in the Lagrangian as a total derivative [6]

$$\delta\mathcal{L} = \frac{\kappa}{2}\partial_\mu\left(\varepsilon^{\mu\nu\rho}\Lambda\partial_\nu A_\rho\right) \tag{36}$$

The Lagrangian is not gauge invariant. However, since it changes by a total derivative under a gauge transformation, we can make the action gauge invariant with a careful choice of boundary conditions; since the action is gauge invariant, all physical measurements we might make will be gauge invariant.

5.2 Canonical Quantization

To canonically quantize the theory, we want to break down the Lagrangian such that it takes the form

$$\mathcal{L} = p\dot{q} - \mathcal{H} + \lambda F$$

where F is some constraint function, and λ a Lagrange multiplier for the constraint. We first separate the terms in order to eliminate the A_0, ∂_0 freedom in the Lagrangian.

$$\mathcal{L} = \frac{1}{2}\kappa\varepsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho$$

$$= \frac{1}{2}\kappa\left(\varepsilon^{i0j}A_i\dot{A}_j + \varepsilon^{0ij}A_0\partial_i A_j + \varepsilon^{ij0}A_i\partial_j A_0\right)$$

using $\quad \varepsilon^{i0j} = -\varepsilon^{ij} = \varepsilon^{0ij} = \varepsilon^{ij0}$

$$= \frac{1}{2}\kappa\varepsilon^{ij}\left(-A_i\dot{A}_j + A_0\partial_i A_j + A_i\partial_j A_0\right)$$

integration by parts: $\quad A_i\partial_j A_0 = -A_0\partial_j A_i$

$$= \frac{1}{2}\kappa\varepsilon^{ij}\left(-A_i\dot{A}_j + A_0\partial_i A_j - A_0\partial_j A_i\right)$$

$$= \frac{1}{2}\kappa\varepsilon^{ij}\left(A_j\dot{A}_i + A_0 F_{ij}\right)$$

using $\quad \varepsilon^{ij}F_{ij} = F_{12} - F_{21} = 2F_{12}$

$$= \frac{\kappa}{2}\varepsilon^{ij}A_j\dot{A}_i + \kappa A_0 F_{12}$$

integration by parts: $\quad A_2\dot{A}_1 - A_1\dot{A}_2 = 2A_2\dot{A}_1$

$$= \kappa A_2\dot{A}_1 + \kappa A_0 F_{12}$$

where we noted the non-obvious steps in the derivation (and liberally using the rule that $A\partial_\mu B = -B\partial_\mu A$ within the Lagrangian). Clearly, we now have a Lagrangian of the canonical form mentioned above, with

$$p\dot{q} = \kappa A_2\dot{A}_1 \qquad \mathcal{H} = 0 \qquad \lambda F = \kappa A_0 F_{12}$$

We now define our p, q variables as

$$p = \sqrt{\kappa}A_2 \qquad q = \sqrt{\kappa}A_1 \tag{37}$$

We can also check the Poisson bracket

$$\{p, q\} = \{\sqrt{\kappa}A_1(x), \sqrt{\kappa}A_2(y)\} = \frac{\delta A_1(x)}{\delta A_1(w)}\frac{\delta A_2(y)}{\delta A_2(w)} - \frac{\delta A_1(x)}{\delta A_2(w)}\frac{\delta A_2(y)}{\delta A_1(w)} = \delta(x-y) \tag{38}$$

where we dropped the factors of κ, because the functional derivative removes them. In a quantum theory, we elevate the Poisson bracket into the commutator, which gives us non-commuting elements of the gauge field:

$$[A_1(x), A_2(y)] = i\{A_1(x), A_2(y)\} = \frac{i}{\kappa}\delta(x-y) \tag{39}$$

where we set $\hbar = 1$. The constraint function imposes the generalization of a Gauss Law constraint to our Chern-Simons theory. We also note that the Hamiltonian is identically zero

in a pure Chern-Simons theory. The physically allowed states, therefore, are constrained only by the Gauss Law constraint F. In this sense, the Gauss Law constraint acts as an annihilator operator on physically allowed states:

$$F_{12}|\phi\rangle = 0$$

where we have dropped the κA_0 term, as they act as a Lagrange multiplier. Drawing inspiration from basic quantum theory and our commutator in Equation 39, we treat the canonical momentum A_2 as a functional derivative:

$$A_2(x) = \frac{1}{i\kappa}\frac{\delta}{\delta A_1(x)}$$

For convenience, we are going to drop the position dependence notation $A_i(x) \to A_i$ and assume all fields are being evaluated at the same position (the functional derivatives will ensure this, otherwise). Using the functional derivative notation, we get our constraint equation

$$\left(\frac{1}{i\kappa}\partial_1\frac{\delta}{\delta A_1} - \partial_2 A_1\right)|\phi\rangle = 0$$

If we assume we have some position space wave function (using our canonical variable A_1), we can write this as

$$\left(\partial_1\frac{\delta}{\delta A_1} - i\kappa\partial_2 A_1\right)\phi(A_1) = 0$$

An easy solution to this equation is

$$\phi(A_1) = \exp\left(i\kappa A_1\frac{\partial_2}{\partial_1}A_1\right)$$

where we treat the operator $1/\partial_1$ as an inverse derivative operator. If the reader is unconvinced by the convenient definition of an inverse derivative operator, merely think of the equation in terms of matrices. We can make an infinite column vector for A_1 with ith component $A_1(x)$ and $(i+1)$th component $A_1(x+\delta x)$, in which case the derivative operator takes the form of an infinite square matrix

$$\partial = \frac{1}{\delta x}\begin{pmatrix} 1 & -1 & 0 & 0 & \cdots \\ 0 & 1 & -1 & 0 & \cdots \\ 0 & 0 & 1 & -1 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

The inverse matrix of this is simply

$$\frac{1}{\partial} = \delta x\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots \\ 0 & 1 & 1 & 1 & \cdots \\ 0 & 0 & 1 & 1 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

Although we will not explicitly use such matrix forms for the derivative and inverse derivative operators, they give some intuition to the actual structure of the operation and merit a

31

mention. Alternatively, for those versed in field theory, the inverse derivatives can be thought of in the same way simple propagators are thought of as "inverse d'Alembertian" operators.

We can repeat the above analysis in terms of a different pair of complex variables, a and $a^\dagger$. We define these variables

$$a = A_1 + iA_2 \qquad a^\dagger = A_1 - iA_2$$

These variables mirror the definitions of the raising and lowering operators from the standard quantum harmonic oscillator. The commutator of these two variables is

$$[a, a^\dagger] = [A_1 + iA_2, A_1 - iA_2] = -2i\,[A_1, A_2] = \frac{2}{\kappa}\delta(x - y)$$

where we add the delta function for clarity (as usual, it serves to enforce a measurement of the fields at the same position and will be dropped from now on). We can scale each by a factor of $\sqrt{\kappa/2}$ to normalize the commutator, giving us

$$[a, a^\dagger] = 1 \quad \text{with} \quad a = \sqrt{\frac{\kappa}{2}}\,(A_1 + iA_2)$$

We next define similar complex derivative operators:

$$D_+ = \frac{\partial_1 + i\partial_2}{i\sqrt{2}} \qquad D_- = \frac{\partial_1 - i\partial_2}{i\sqrt{2}}$$

In terms of these variables, we can rewrite our Gauss Law constraint

$$
\begin{aligned}
F_{12} &= \partial_1 A_2 - \partial_2 A_1 \\
&= \partial_1 A_2 - \partial_2 A_1 + \frac{1}{2i}\left(\partial_1 A_1 + \partial_2 A_2 - \partial_1 A_1 - \partial_2 A_2\right) \\
&= \frac{1}{2i}\left[\partial_1 A_1 + i\partial_1 A_2 - i\partial_2 A_1 + \partial_2 A_2 - \left(\partial_1 A_1 - i\partial_1 A_2 + i\partial_2 A_1 + \partial_2 A_2\right)\right] \\
&= \frac{D_- a - D_+ a^\dagger}{\sqrt{\kappa}}
\end{aligned}
$$

By analogy with our original approach, we now treat $a^\dagger$ as a functional derivative

$$a^\dagger = -\frac{\delta}{\delta a}$$

Using the same process as above, we can find our physical states by applying the Gauss Law constraint as an annihilation operator

$$\sqrt{\frac{1}{\kappa}}\left(D_- a + D_+ \frac{\delta}{\delta a}\right)\phi(a) = 0$$

Which has a similar solution to the first case

$$\phi(a) = \exp\left(-a\frac{D_-}{D_+}a\right)$$

5.3 Non-Abelian Chern-Simons

Having now examined the canonical structure of the abelian theory, we can turn to the non-abelian case, where many of the results mirror the structure of the abelian case. The non-abelian Lagrangian is defined by

$$\mathcal{L} = \kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr}\left(A_\mu \partial_\nu A_\rho + \frac{2}{3} A_\mu A_\nu A_\rho \right) \tag{40}$$

For the non-abelian gauge fields, we define

$$A_\mu \equiv A_\mu^a T^a \tag{41}$$

where T^a are the generators of $\mathfrak{su}(\mathrm{N})$. As in the abelian case, we can vary the Lagrangian by an infinitesimal shift in the gauge field, $A_\mu \to A_\mu + \delta A_\mu$. The corresponding shift in the Lagrangian is [6]

$$\mathcal{L} + \delta\mathcal{L} = \kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr}\left((A_\mu + \delta A_\mu) \partial_\nu (A_\rho + \delta A_\rho) + \frac{2}{3} (A_\mu + \delta A_\mu)(A_\nu + \delta A_\nu)(A_\rho + \delta A_\rho) \right)$$

$$\delta\mathcal{L} = \kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr}\left(\delta A_\mu \partial_\nu A_\rho + A_\mu \partial_\nu \delta A_\rho + \frac{2}{3} (A_\mu A_\nu \delta A_\rho + \text{ cyclic permutations in } \mu,\nu,\rho) \right)$$

$$= \kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr}\left(\delta A_\mu (\partial_\nu A_\rho - \partial_\rho A_\nu) + 2\, \delta A_\mu A_\nu A_\rho \right)$$

$$= \kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr}\left(\delta A_\mu (\partial_\nu A_\rho - \partial_\rho A_\nu + [A_\nu A_\rho]) \right)$$

$$= \kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr}\left(\delta A_\mu F_{\nu\rho} \right)$$

where we find the non-abelian field strength as in Yang-Mills

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$$

Using the association that $A_\mu = A_\mu^a T^a$ and that, since $F_{\mu\nu}$ is a function of elements of $\mathfrak{su}(\mathrm{N})$ and the bracket of elements of $\mathfrak{su}(\mathrm{N})$ and thus is also contained in $\mathfrak{su}(\mathrm{N})$, we can associate

$$F_{\mu\nu} \equiv F_{\mu\nu}^a T^a$$

Plugging this in, we see

$$\kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr}\left(\delta A_\mu F_{\nu\rho} \right) = \kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr}\left(\delta A_\mu^a T^a F_{\nu\rho}^b T^b \right)$$

$$= \frac{\kappa}{2} \varepsilon^{\mu\nu\rho} \delta A_\mu^a F_{\nu\rho}^a$$

Requiring this variation to vanish for non-zero shift δA_μ gives us equations of motion that are analogous to the abelian case:

$$\kappa \varepsilon^{\mu\nu\rho} F_{\nu\rho}^a = 0$$

which is equivalent to saying

$$\kappa \varepsilon^{\mu\nu\rho} F_{\nu\rho} = 0$$

The exact correspondence between the abelian and non-abelian equations of motion suggest the non-zero source equation of motion

$$\kappa \varepsilon^{\mu\nu\rho} F_{\nu\rho} = J^{\mu}$$

which suggests the form of the non-zero source Lagrangian

$$\mathcal{L} = \kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr} \left(A_{\mu} \partial_{\nu} A_{\rho} + \frac{2}{3} A_{\mu} A_{\nu} A_{\rho} - A_{\mu} J^{\mu} \right)$$

for non-abelian source $J^{\mu} = J^{\mu\,a} T^a$. However, unlike in the abelian case, the non-abelian Chern-Simons action changes by an additive factor under a gauge transformation. Our non-abelian gauge transformation is

$$A_{\mu} \to g^{-1} A_{\mu} g + g^{-1} \partial_{\mu} g \quad \text{with} \quad g \in \mathbf{SU}(N)$$

The full form of the shift in the Lagrangian from a gauge transformation is given by [6]

$$\delta\mathcal{L} = -\kappa \varepsilon^{\mu\nu\rho} \partial_{\mu} \operatorname{Tr} \left(\partial_{\nu} g g^{-1} A_{\rho} \right) - \frac{\kappa}{3} \varepsilon^{\mu\nu\rho} \operatorname{Tr} \left(g^{-1} \partial_{\mu} g g^{-1} \partial_{\nu} g g^{-1} \partial_{\rho} g \right) \tag{42}$$

The first term represents a total derivative, which we can ignore with appropriate boundary conditions, and the second term is the winding number density. Again, with appropriate boundary conditions, the integral of the winding number density will be integer valued, which leaves the overall factor $\exp(iS)$ gauge invariant, given certain restrictions on the coupling κ.

We can check this for an infinitesimal gauge transformation, $g \approx 1 + i\varepsilon\alpha$ for ε very small and $\alpha \in \mathfrak{su}(N)$. Our gauge transformation takes the form

$$A_{\mu} \to (1 - i\varepsilon\alpha) A_{\mu} (1 + i\varepsilon\alpha) + (1 - i\varepsilon\alpha) \partial_{\mu} (1 + i\varepsilon\alpha) = A_{\mu} + i\varepsilon \left([A_{\mu}, \alpha] + \partial_{\mu}\alpha \right)$$

We first note that under an infinitesimal transformation as above, the winding number density will go to zero, as it will only have terms of order three or higher in ε. Plugging our gauge transformed variable into the Lagrangian, we get

$$\delta\mathcal{L} = \kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr} \Bigg(\left(A_{\mu} + i\varepsilon \left([A_{\mu}, \alpha] + \partial_{\mu}\alpha \right) \right) \partial_{\nu} \left(A_{\rho} + i\varepsilon \left([A_{\rho}, \alpha] + \partial_{\rho}\alpha \right) \right)$$

$$+ \frac{2}{3} \left(A_{\mu} + i\varepsilon \left([A_{\mu}, \alpha] + \partial_{\mu}\alpha \right) \right) \left(A_{\nu} + i\varepsilon \left([A_{\nu}, \alpha] + \partial_{\nu}\alpha \right) \right) \left(A_{\rho} + i\varepsilon \left([A_{\rho}, \alpha] + \partial_{\rho}\alpha \right) \right) \Bigg) - \mathcal{L}$$

$$= \kappa \varepsilon^{\mu\nu\rho} \operatorname{Tr} \Bigg(i\varepsilon A_{\mu} \partial_{\nu} \left([A_{\rho}, \alpha] + \partial_{\rho}\alpha \right) + i\varepsilon \left([A_{\mu}, \alpha] + \partial_{\mu}\alpha \right) \partial_{\nu} A_{\rho}$$

$$+ i\varepsilon \frac{2}{3} \left(A_{\mu} A_{\nu} \left([A_{\rho}, \alpha] + \partial_{\rho}\alpha \right) + \left([A_{\mu}, \alpha] + \partial_{\mu}\alpha \right) A_{\nu} A_{\rho} + A_{\mu} \left([A_{\nu}, \alpha] + \partial_{\nu}\alpha \right) A_{\rho} \right) \Bigg)$$

Looking at the three terms in the last line, we note that we can expand and rewrite them

$$A_{\mu} A_{\nu} \left([A_{\rho}, \alpha] + \partial_{\rho}\alpha \right) = A_{\mu} A_{\nu} A_{\rho} \alpha - A_{\mu} A_{\nu} \alpha A_{\rho} + A_{\mu} A_{\nu} \partial_{\rho}\alpha$$

$$= A_{\mu} A_{\nu} A_{\rho} \alpha - A_{\rho} A_{\mu} A_{\nu} \alpha + A_{\mu} A_{\nu} \partial_{\rho}\alpha$$

$$= A_{\mu} A_{\nu} \partial_{\rho}\alpha$$

where the first two terms disappear by the antisymmetry of the Levi-Civita symbol. Thus, for the last three terms in the gauge variation, the only important part is the ordering of the indices. We can thus simplify the variation to

$$
\delta\mathcal{L} = \kappa\varepsilon^{\mu\nu\rho}\,\mathrm{Tr}\left(i\varepsilon A_\mu\partial_\nu\Big([A_\rho,\alpha]+\partial_\rho\alpha\Big) + i\varepsilon\Big([A_\mu,\alpha]+\partial_\mu\alpha\Big)\partial_\nu A_\rho\right.
$$

$$
\left. + i\varepsilon\frac{2}{3}\Big(A_\mu A_\nu\Big([A_\rho,\alpha]+\partial_\rho\alpha\Big) + \Big([A_\mu,\alpha]+\partial_\mu\alpha\Big)A_\nu A_\rho + A_\mu\Big([A_\nu,\alpha]+\partial_\nu\alpha\Big)A_\rho\Big)\right)
$$

$$
= \kappa\varepsilon^{\mu\nu\rho}\,\mathrm{Tr}\left(i\varepsilon A_\mu\partial_\nu\Big([A_\rho,\alpha]+\partial_\rho\alpha\Big) + i\varepsilon\Big([A_\mu,\alpha]+\partial_\mu\alpha\Big)\partial_\nu A_\rho\right.
$$

$$
\left. + i\varepsilon\frac{2}{3}\Big(A_\mu A_\nu\partial_\rho\alpha + A_\nu A_\rho\partial_\mu\alpha + A_\rho A_\mu\partial_\nu\alpha\Big)\right)
$$

But again, these terms combine by the properties of the Levi-Civita symbol, so we get

$$
\delta\mathcal{L} = \kappa\varepsilon^{\mu\nu\rho}\,\mathrm{Tr}\left(i\varepsilon A_\mu\partial_\nu\Big([A_\rho,\alpha]+\partial_\rho\alpha\Big) + i\varepsilon\Big([A_\mu,\alpha]+\partial_\mu\alpha\Big)\partial_\nu A_\rho + 2i\varepsilon A_\mu A_\nu\partial_\rho\alpha\right)
$$

We can further simplify the first two terms. Expanding the first, we see

$$
\begin{aligned}
A_\mu\partial_\nu\Big([A_\rho,\alpha]+\partial_\rho\alpha\Big) &= A_\mu\partial_\nu(A_\mu\alpha) - A_\mu\partial_\nu(\alpha A_\mu) + A_\mu\partial_\nu\partial_\rho\alpha\\
&= A_\mu\partial_\nu(A_\mu\alpha) - A_\mu\partial_\nu(\alpha A_\mu)\\
&= A_\mu\partial_\nu A_\rho\alpha + A_\mu A_\rho\partial_\nu\alpha - A_\rho A_\mu\partial_\nu\alpha - \partial_\nu A_\rho A_\mu\alpha\\
&= 2A_\mu A_\rho\partial_\nu\alpha + A_\mu\partial_\nu A_\rho\alpha - \partial_\nu A_\rho A_\mu\alpha\\
&= 2A_\mu A_\rho\partial_\nu\alpha + A_\mu\partial_\nu A_\rho\alpha + A_\rho\partial_\nu A_\mu\alpha\\
&= 2A_\mu A_\rho\partial_\nu\alpha
\end{aligned}
$$

where we integrated by parts and used the symmetry properties of the Levi-Civita symbol on several occasions. Clearly, however, by those same symmetry properties, this term precisely cancels the last term in the overall expression above. Thus, our gauge variance is given only by the second term. Expanding it, we see

$$
\begin{aligned}
([A_\mu,\alpha]+\partial_\mu\alpha)\partial_\nu A_\rho &= (A_\mu\alpha - \alpha A_\mu + \partial_\mu\alpha)\,\partial_\nu A_\rho\\
&= \partial_\nu A_\rho A_\mu\alpha - A_\mu\partial_\nu A_\rho\alpha + \partial_\mu\alpha\partial_\nu A_\rho\\
&= \partial_\nu A_\rho A_\mu\alpha + \partial_\nu A_\mu A_\rho\alpha + \partial_\mu\alpha\partial_\nu A_\rho\\
&= \partial_\mu\alpha\partial_\nu A_\rho\\
&= \partial_\mu\alpha\partial_\nu A_\rho + \partial_\nu\partial_\mu\alpha A_\rho\\
&= \partial_\nu\left(\partial_\mu\alpha A_\rho\right)
\end{aligned}
$$

Where in the penultimate line, we added a term that is identically zero because of the Levi-Civita symbol. Thus, our overall gauge variation is given by the (remarkably simple)

$$
\delta\mathcal{L} = \kappa\varepsilon^{\mu\nu\rho}\partial_\nu\,\mathrm{Tr}\left(i\varepsilon\partial_\mu\alpha A_\rho\right)
$$

35

Looking back at the original formula (and ignoring the winding term, which is zero in the small transformation limit), we have

$$\delta\mathcal{L} = -\kappa\varepsilon^{\mu\nu\rho}\partial_\mu \operatorname{Tr}\left(\partial_\nu gg^{-1}A_\rho\right) = \kappa\varepsilon^{\mu\nu\rho}\partial_\nu \operatorname{Tr}\left(\partial_\mu gg^{-1}A_\rho\right)$$

Clearly, we need only check the contents of the trace. Again using the approximate gauge transformation, $g \approx 1 + i\varepsilon\alpha$, we have

$$\partial_\mu gg^{-1}A_\rho = \partial_\mu(1 + i\varepsilon\alpha)(1 - i\varepsilon\alpha)A_\rho = i\varepsilon\partial_\mu\alpha(1 - i\varepsilon\alpha)A_\rho = i\varepsilon\partial_\mu\alpha A_\rho$$

as needed.

5.4   Canonical Quantization

Having established the essential properties of non-abelian Chern-Simons theory, we can now mimic our approach in the abelian case and quantize the theory. We first separate the 0th component terms in the Lagrangian

$$
\begin{aligned}
\mathcal{L} &= \kappa\varepsilon^{\mu\nu\rho}\operatorname{Tr}\left(A_\mu\partial_\nu A_\rho + \frac{2}{3}A_\mu A_\nu A_\rho\right) \\
&= \kappa\varepsilon^{0ij}\operatorname{Tr}\left(A_0\partial_i A_j + \frac{2}{3}A_0 A_i A_j\right) \\
&\quad + \kappa\varepsilon^{i0j}\operatorname{Tr}\left(A_i\dot{A}_j + \frac{2}{3}A_i A_0 A_j\right) \\
&\quad + \kappa\varepsilon^{ij0}\operatorname{Tr}\left(A_i\partial_j A_0 + \frac{2}{3}A_i A_j A_0\right)
\end{aligned}
$$

We can use integration by parts and other tricks from our previous abelian calculation to simplify this

$$
\begin{aligned}
\mathcal{L} &= \kappa\varepsilon^{ij}\operatorname{Tr}\left(A_0\partial_i A_j + \frac{2}{3}A_0 A_i A_j\right) \\
&\quad - \kappa\varepsilon^{ij}\operatorname{Tr}\left(A_i\dot{A}_j + \frac{2}{3}A_0 A_j A_i\right) \\
&\quad + \kappa\varepsilon^{ij}\operatorname{Tr}\left(-A_0\partial_j A_i + \frac{2}{3}A_0 A_i A_j\right) \\
&= \kappa\varepsilon^{ij}\operatorname{Tr}\left(A_0\left(\partial_i A_j - \partial_j A_i\right) + A_j\dot{A}_i + 2A_0 A_i A_j\right) \\
&= \kappa\varepsilon^{ij}\operatorname{Tr}\left(A_0\left(\partial_i A_j - \partial_j A_i + 2A_i A_j\right) + A_j\dot{A}_i\right)
\end{aligned}
$$

We can use the anti-symmetry of ε^{ij} to simplify the expression as before, giving us

$$\mathcal{L} = \kappa\varepsilon^{ij}\operatorname{Tr}\left(A_0\left(\partial_i A_j - \partial_j A_i + 2A_i A_j\right) + A_j\dot{A}_i\right) = 2\kappa\operatorname{Tr}\left(A_0 F_{12} + A_2\dot{A}_1\right) \quad (43)$$

where, by the same tricks, we recover the non-abelian field strength tensor by taking advantage of the symmetry properties of ε^{ij}. We can use the Lie algebraic structure to simplify the non-abelian field strength somewhat, if we do so in terms of components:

$$F_{12}^a = \partial_1 A_2^a T^a - \partial_2 A_1^a T^a + [A_1^b T^b, A_2^c T^c] = \partial_1 A_2^a T^a - \partial_2 A_1^a T^a + if^{bca}A_1^b A_2^c T^a$$

36

We can use the component forms of the gauge fields and field strength to eliminate the trace and the generators of $\mathfrak{su}(N)$ from our Lagrangian, giving us $N^2 - 1$ scalar field equations:

$$
\begin{aligned}
\mathcal{L} &= 2\kappa \operatorname{Tr}\left(A_0 F_{12} + A_2 \dot{A}_1\right) \\
&= 2\kappa \operatorname{Tr}\left(A_0^d T^d \left(\partial_1 A_2^a T^a - \partial_2 A_1^a T^a + i f^{bca} A_1^b A_2^c T^a\right) + A_2^a \dot{A}_1^b T^a T^b\right) \\
&= 2\kappa A_0^d \left(\partial_1 A_2^a - \partial_2 A_1^a + i f^{bca} A_1^b A_2^c\right) \operatorname{Tr} T^d T^a + 2\kappa A_2^a \dot{A}_1^b \operatorname{Tr} T^a T^b \\
&= \kappa A_0^a \left(\partial_1 A_2^a - \partial_2 A_1^a + i f^{bca} A_1^b A_2^c\right) + \kappa A_2^a \dot{A}_1^a
\end{aligned}
$$

This form of the Lagrangian mirrors precisely the form of the abelian case, but with $N^2 - 1$ fields and an additional term in our constraint. The Hamiltonian density is still zero, our canonical variables and commutators are the same

$$
\left[A_1^a(x), A_2^b(y)\right] = \frac{i}{\kappa}\delta_{ab}\delta(x - y) \tag{44}
$$

and once again, we can find wave functions that represent physically valid states by finding solutions to the Gauss Law constraint $F_{12} = 0$ with Lagrange multiplier A_0. We can also treat A_2 as a differential operator with respect to A_1 as before, giving us

$$
A_2^a = \frac{1}{i\kappa}\frac{\delta}{\delta A_1^a}
$$

5.5   A Motivating Example: On a Disc

We first call our attention back to when we checked the overall shift in the Lagrangian from a gauge transformation. In that case, we argued that one of the two terms disappeared, as it was a total derivative; since we generally assume that the fields go to zero at infinity (to keep amplitudes normalized and non-divergent), we could ignore it. However, on a bounded manifold, such as the disc, we can have non-trivial contributions from a total derivative term. To account for this, we have to add a boundary term in the action [7]:

$$
S = \kappa \int_M d^3x\, \varepsilon^{\mu\nu\rho} Tr\left(A_\mu \partial_\nu A_\rho + \frac{2}{3}A_\mu A_\nu A_\rho\right) + \kappa \int_{\delta M} d^2x\, \operatorname{Tr}(A_2 A_0)
$$

This boundary term precisely acts to cancel the non-zero boundary contributions from a variation in the action. We can now quantize the theory on a disc to build some intuition for our later discussion on the torus. On a disc, we can solve our constraint with a "pure gauge" solution, using fields represented as [7]

$$
A_\mu = -\partial_\mu U U^{-1} \qquad U \in \mathbf{SU}(N)
$$

where we denote dimensions by $0 = t, 1 = r, 2 = \varphi$. This representation of the gauge field satisfies the constraint: to see this, we can check for a "small" field A, such that $U \approx 1 + i\varepsilon\alpha$, for $\varepsilon \ll 1$ and $\alpha \in \mathfrak{su}(N)$ (note that this implies $\alpha = \alpha^a T^a$ as usual with Lie algebra valued elements). This gives us a gauge field, in our representation,

$$
A_\mu = -\partial_\mu(1 + i\varepsilon\alpha)(1 - i\varepsilon\alpha) = -i\varepsilon\partial_\mu\alpha
$$

Our constraint is given by

$$\partial_1 A_2^a - \partial_2 A_1^a + i f^{abc} A_1^b A_2^c = 0$$

Plugging in our form of the infinitesimal field, we have

$$i\varepsilon \partial_1 \partial_2 \alpha^a - i\varepsilon \partial_2 \partial_1 \alpha^a - i\varepsilon^2 f^{abc} \partial_1 \alpha^b \partial_2 \alpha^c = 0$$

as needed. Thus, for the simple case of the disc, the representation give above represents a valid choice of gauge field. With gauge fields in this form, we can rewrite the action as

$$S = \kappa \int_M d^3x \, \varepsilon^{\mu\nu\rho} Tr\left(A_\mu \partial_\nu A_\rho + \frac{2}{3} A_\mu A_\nu A_\rho\right) + \kappa \int_{\delta M} d^2x \, \mathrm{Tr}(A_2 A_0)$$

$$= \frac{\kappa}{3} \int_M d^3x \, \varepsilon^{\mu\nu\rho} \, \mathrm{Tr}\left(U^{-1}\partial_\mu U U^{-1}\partial_\nu U U^{-1}\partial_\rho U\right) + \kappa \int_{\delta M} dt \, d\varphi \, \mathrm{Tr}\left(U^{-1}\partial_\varphi U U^{-1}\partial_t U\right)$$

To see this, we first note that the second term is obvious: it comes directly from plugging in our representation of the gauge field into the boundary term in the action we recently introduced. The first term comes from the Chern-Simons Lagrangian term:

$$A_\mu \partial_\nu A_\rho + \frac{2}{3} A_\mu A_\nu A_\rho = \partial_\mu U U^{-1}\partial_\nu\left(\partial_\rho U U^{-1}\right) - \frac{2}{3}\left(\partial_\mu U U^{-1}\partial_\nu U U^{-1}\partial_\rho U U^{-1}\right)$$

$$= \partial_\mu U U^{-1}\partial_\nu\partial_\rho U U^{-1} + \partial_\mu U U^{-1}\partial_\rho U \partial_\nu U^{-1}$$

$$\quad - \frac{2}{3}\left(\partial_\mu U U^{-1}\partial_\nu U U^{-1}\partial_\rho U U^{-1}\right)$$

$$= \partial_\mu U U^{-1}\partial_\rho U U^{-1} U \partial_\nu U^{-1} - \frac{2}{3}\left(\partial_\mu U U^{-1}\partial_\nu U U^{-1}\partial_\rho U U^{-1}\right)$$

where we added a factor of the identity, $U^{-1}U$, in the middle of the first term. Noting that

$$\partial_\mu\left(UU^{-1}\right) = 0 = \partial_\mu U U^{-1} + U\partial_\mu U^{-1}$$

we can simplify this further, to get

$$= -\partial_\mu U U^{-1}\partial_\rho U U^{-1}\partial_\nu U U^{-1} - \frac{2}{3}\left(\partial_\mu U U^{-1}\partial_\nu U U^{-1}\partial_\rho U U^{-1}\right)$$

$$= \frac{1}{3}\partial_\mu U U^{-1}\partial_\nu U U^{-1}\partial_\rho U U^{-1}$$

which gives us the first term in the action, as needed. Rewriting our action in terms of our small field approximation, we have

$$S = \frac{\kappa}{3} \int_M d^3x \, \varepsilon^{\mu\nu\rho} \, \mathrm{Tr}\left(-i\varepsilon^3 \partial_\mu\alpha\partial_\nu\alpha\partial_\rho\alpha\right) + \kappa \int_{\delta M} dt \, d\varphi \, \mathrm{Tr}\left(-\varepsilon^2\partial_t\alpha\partial_\varphi\alpha\right)$$

In terms of the generators, we have

$$S = \frac{\kappa}{3} \int_M d^3x \, \varepsilon^{\mu\nu\rho}\left(-i\varepsilon^3 \partial_\mu\alpha^a\partial_\nu\alpha^b\partial_\rho\alpha^c\right) \mathrm{Tr}\, T^a T^b T^c + \frac{\kappa}{2}\int_{\delta M} dt \, d\varphi \left(-\varepsilon^2\partial_t\alpha^a\partial_\varphi\alpha^a\right)$$

where we used our trace condition on the generators for the boundary term. To simplify the triple generator trace, we need another identity. We can take a standard identity on the generators

$$T^a T^b = \frac{1}{2N}\delta_{ab} + \frac{1}{2}(if^{abc} + d^{abc})T^c$$

38

for d^{abc} a set of constants symmetric in all indices (they come from the anti-commutators of the generators). We can multiply either side by T^d to get

$$T^a T^b T^d = \frac{1}{2N} \delta_{ab} T^d + \frac{1}{2} (i f^{abc} + d^{abc}) T^c T^d$$

If we then take a trace of both sides, we have

$$\operatorname{Tr} T^a T^b T^d = \operatorname{Tr} \left(\frac{1}{2N} \delta_{ab} T^d \right) + \frac{1}{4} \left(i f^{abd} + d^{abd} \right) = \frac{1}{4} \left(i f^{abd} + d^{abd} \right)$$

since the generators T^a are inherently traceless. This gives us the needed identity. Plugging this in, we have

$$S = \frac{\kappa}{12} \int_M d^3 x \, \varepsilon^{\mu\nu\rho} \left(-i\varepsilon^3 \partial_\mu \alpha^a \partial_\nu \alpha^b \partial_\rho \alpha^c \right) \left(i f^{abc} + d^{abc} \right) + \frac{\kappa}{2} \int_{\delta M} dt \, d\varphi \left(-\varepsilon^2 \partial_t \alpha^a \partial_\varphi \alpha^a \right)$$

$$= \frac{\kappa}{12} \int_M d^3 x \, i f^{abc} \varepsilon^{\mu\nu\rho} \left(-i\varepsilon^3 \partial_\mu \alpha^a \partial_\nu \alpha^b \partial_\rho \alpha^c \right) + \frac{\kappa}{2} \int_{\delta M} dt \, d\varphi \left(-\varepsilon^2 \partial_t \alpha^a \partial_\varphi \alpha^a \right)$$

where we can drop the d^{abc} since it is contracted with the anti-symmetric $\varepsilon^{\mu\nu\rho}$.

The states in this theory on the disc correspond to elements of the loop group, similar to the $1+1$ dimensional Yang-Mills theory. However, since the states must be gauge invariant, states which can transform into one another under gauge transformations must be identified as the same. Thus, the states of the theory come from the quotient group of the loop group modulo the gauge group [7].

5.6 Another Approach to the Disc

There is another approach we can take to this problem of putting the theory on a disc, using holomorphic fields [6]. We can assign

$$A_z^a = \frac{1}{2} (A_1 + i A_2)^a \qquad A_{\bar{z}}^a = \frac{1}{2} (A_1 - i A_2)^a$$

$$z = x_1 + i x_2 \qquad \bar{z} = x_1 - i x_2$$

$$\partial_z = \frac{1}{2} (\partial_1 + i \partial_2) \qquad \partial_{\bar{z}} = \frac{1}{2} (\partial_1 - i \partial_2)$$

The commutator of our new fields is

$$[A_z^a, A_{\bar{z}}^a] = \frac{1}{4} [A_1^a + i A_2^a, A_1^a - i A_2^a] = \frac{1}{4} \left(-i[A_1^a, A_2^a] + i[A_2^a, A_1^a] \right) = \frac{1}{2\kappa}$$

Thus, as before, we can associate

$$A_{\bar{z}}^a = \frac{1}{2\kappa} \frac{\delta}{\delta A_z^a}$$

In this scheme, we can use a similar representation of the gauge field [8]

$$A_z = -\partial_z U U^{-1}$$

We can write the action similarly to before, but with a different boundary condition, in terms of these new complex variables:

$$S = \kappa \int_M d^3x \, \varepsilon^{\mu\nu\rho} Tr \left(A_\mu \partial_\nu A_\rho + \frac{2}{3} A_\mu A_\nu A_\rho \right) + \kappa \int_{\delta M} d^2x \, \mathrm{Tr}(A_z A_{\bar{z}})$$

$$= \frac{\kappa}{3} \int_M d^3x \, \varepsilon^{\mu\nu\rho} \, \mathrm{Tr} \left(U^{-1} \partial_\mu U U^{-1} \partial_\nu U U^{-1} \partial_\rho U \right) + \kappa \int_{\delta M} dz \, d\bar{z} \, \mathrm{Tr} \left(U^{-1} \partial_z U U^{-1} \partial_{\bar{z}} U \right)$$

We must also rewrite our constraint in terms of the new variables. It turns out that it is nearly identical. Basic algebra gives us immediately

$$\partial_1 A_2^a - \partial_2 A_1^a = 2i \left(\partial_z A_{\bar{z}}^a - \partial_{\bar{z}} A_z^a \right)$$

The non-abelian term is more difficult. Consider the term

$$f^{bca} A_z^b A_{\bar{z}}^c = \frac{1}{4} f^{bca} \left(A_1^b + i A_2^b \right) \left(A_1^c - i A_2^c \right)$$

Expanding this out, we get

$$\frac{1}{4} f^{bca} \left(A_1^b + i A_2^b \right) \left(A_1^c - i A_2^c \right) = \frac{1}{4} f^{bca} \left(A_1^b A_1^c + A_2^b A_2^c + -i A_1^b A_2^c + i A_1^c A_2^b \right)$$

The structure constants are anti-symmetric with respect to the swap of any two indices. Thus, the $A_1^b A_a^c$ and $A_2^b A_2^c$ terms are zero, since they are a symmetric quantity contracted with an anti-symmetric quantity. We can also switch the indices and signs on the last term, and we get

$$\frac{1}{4} f^{bca} \left(A_1^b A_1^c + A_2^b A_2^c + -i A_1^b A_2^c + i A_1^c A_2^b \right) = \frac{1}{4} f^{bca} \left(-i A_1^b A_2^c - i A_1^b A_2^c \right) = -\frac{i}{2} f^{bca} A_1^b A_2^c$$

Clearly, then, we have a similar relationship as the abelian part:

$$f^{bca} A_1^b A_2^c = 2i f^{bca} A_z^b A_{\bar{z}}^c$$

We are thus free to make the swap $1 \to z$ and $2 \to \bar{z}$ and our constraint remains exactly the same. In this new coordinate system, then, we have the constraint equation

$$F_{z\bar{z}} |\phi\rangle = 0$$

for some physical state $|\phi\rangle$. In the small field approximation, we have $A_z^a \approx -i\varepsilon \partial_z \alpha^a$ and $A_{\bar{z}}^a \approx -i\varepsilon \partial_{\bar{z}} \alpha^a$, so we see that our new representations of the gauge fields satisfy the constraint, as before

$$\partial_z \left(-i\varepsilon \partial_{\bar{z}} \alpha^a \right) - \partial_{\bar{z}} \left(-i\varepsilon \partial_z \alpha^a \right) + \mathcal{O} \left(\varepsilon^2 \right) = 0$$

as needed. In this scheme, we can write our states explicitly as [6]

$$\psi = \exp\left(4S \right)$$

where the factor of four acts essentially as a scaling factor on the action (it's needed to make the states obey the constraint, but ultimately only has the effect of changing the choice of scale from the original Lagrangian). We can check this in the small field approximation. We

note that our small field approximation of the action will be identical to the previous case, up to a change in the boundary term:

$$S = \frac{\kappa}{12} \int_M d^3x \, i f^{abc} \varepsilon^{\mu\nu\rho} \left(-i\varepsilon^3 \partial_\mu \alpha^a \partial_\nu \alpha^b \partial_\rho \alpha^c \right) + \frac{\kappa}{2} \int_{\delta M} dz \, d\bar{z} \left(-\varepsilon^2 \partial_z \alpha^a \partial_{\bar{z}} \alpha^a \right)$$

We also rewrite our constraint as a functional derivative operator

$$\partial_z A_{\bar{z}}^a - \partial_{\bar{z}} A_z^a + i f^{abc} A_z^b A_{\bar{z}}^c = \frac{1}{2\kappa} \left(\partial_z \frac{\delta}{\delta A_z^a} \right) - \partial_{\bar{z}} A_z^a + \frac{1}{2\kappa} \left(i f^{abc} A_z^b \frac{\delta}{\delta A_z^c} \right)$$

In terms of a small field approximation, we have

$$-\frac{1}{2\kappa} \frac{1}{i\varepsilon} \left(\partial_z \frac{\delta}{\delta(\partial_z \alpha^a)} \right) + i\varepsilon \partial_{\bar{z}} \partial_z \alpha^a + \frac{1}{2\kappa} \left(i f^{abc} \partial_z \alpha^b \frac{\delta}{\delta(\partial_z \alpha^c)} \right)$$

We first note that the functional derivatives will only touch the boundary term. The action of the functional derivative on the states will be

$$\frac{\delta}{\delta(\partial_z \alpha^a)} \exp(4S) = \frac{\delta}{\delta(\partial_z \alpha^a)} \exp\left(\frac{\kappa}{3} \int_M d^3x \, i f^{ijk} \varepsilon^{\mu\nu\rho} \left(-i\varepsilon^3 \partial_\mu \alpha^i \partial_\nu \alpha^j \partial_\rho \alpha^k \right) \right.$$

$$\left. + 2\kappa \int_{\delta M} dz \, d\bar{z} \left(-\varepsilon^2 \partial_z \alpha^i \partial_{\bar{z}} \alpha^i \right) \right)$$

$$= \left(-2\varepsilon^2 \kappa \right) \partial_{\bar{z}} \alpha^i \delta_{ia} \exp\left(\dots \right)$$

Thus, acting on our states with the constraint yields

$$\left[-\frac{1}{2\kappa} \frac{1}{i\varepsilon} \left(\partial_z \frac{\delta}{\delta(\partial_z \alpha^a)} \right) + i\varepsilon \partial_{\bar{z}} \partial_z \alpha^a + \frac{1}{2\kappa} \left(i f^{abc} \partial_z \alpha^b \frac{\delta}{\delta(\partial_z \alpha^b)} \right) \right] \exp(4S)$$

$$= \left[-\frac{1}{2\kappa} \frac{1}{i\varepsilon} \left(\partial_z \left(-2\varepsilon^2 \kappa \right) \partial_{\bar{z}} \alpha^i \delta_{ia} \right) + i\varepsilon \partial_{\bar{z}} \partial_z \alpha^a + \frac{1}{2\kappa} \left(i f^{abc} \partial_z \alpha^b \left(-2\varepsilon^2 \kappa \right) \partial_{\bar{z}} \alpha^i \delta_{ic} \right) \right] \exp(4S)$$

$$= \left[-i\varepsilon \partial_z \partial_{\bar{z}} \alpha^a + i\varepsilon \partial_{\bar{z}} \partial_z \alpha^a + \mathcal{O}(\varepsilon^2) \right] \exp(4S) = 0$$

as needed. With this straightforward case in mind, we can now consider what happens when we quantize our theory on a more interesting (and non-trivial) manifold: the torus.

5.7 Non-Abelian Chern-Simons on a Torus

When we considered Chern-Simons theory on a disc, we did not need to worry about the topological properties of the space. On the torus, however, we cannot be so cavalier about the space. To account for the non-contractible loops of the torus, we have to make an adjustment to our gauge field. In fact, we can use an adjusted form of the representation we used previously:

$$A_i = -\partial_i U U^{-1} + U \theta_i(t) U^{-1}$$

for the spatial components of the gauge field. The two functions θ_i, which are elements of $\mathfrak{su}(N)$, encode the information regarding the two independent loops of the torus. As before, we can check that this form of the gauge field obeys our constraint:

$$\text{Tr} \left(A_0 F_{12} \right) = \text{Tr} \left(A_0 \left(\partial_1 A_2 - \partial_2 A_1 + [A_1, A_2] \right) \right) = \partial_1 A_2^a - \partial_2 A_1^a + i f^{abc} A_1^b A_2^c = 0$$

41

We can again check this for a small field approximation, with $\partial_\mu U U^{-1} \approx i\varepsilon \partial_\mu \alpha$. However, since θ_j is also part of the field, we have to assume that θ_j is also small: $\theta_j \to \varepsilon \theta_j$. We now have

$$U\theta_j U^{-1} = (1 + i\varepsilon\alpha)\varepsilon\theta_j(1 - i\varepsilon\alpha) = \varepsilon\theta_j + \mathcal{O}(\varepsilon^2)$$

Plugging these approximate fields into our constraint, we have

$$\begin{aligned}
0 &= \partial_1 A_2^a - \partial_2 A_1^a + if^{abc} A_1^b A_2^c \\
&= \partial_1\Big(-i\varepsilon\partial_2\alpha^a + \varepsilon\theta_2^a + \mathcal{O}(\varepsilon^2)\Big) - \partial_2\Big(-i\varepsilon\partial_1\alpha^a + \varepsilon\theta_1^a + \mathcal{O}(\varepsilon^2)\Big) \\
&= -i\varepsilon\partial_1\partial_2\alpha^a + i\varepsilon\partial_2\partial_1\alpha^a
\end{aligned}$$

as needed. We can also rephrase our action in terms of this form of the gauge field. Since the gauge fields identically cancel the constraint, we can use the adjusted form of the action

$$S = 2\kappa \int d^3x \ \mathrm{Tr}\Big(A_0 F_{12} + A_2 \dot{A}_1\Big) = 2\kappa \int d^3x \ \mathrm{Tr}\Big(A_2 \dot{A}_1\Big)$$

where we no longer need the boundary term, as we are working on a surface without boundary. In fact, it is more instructive to use a slightly "older" form of the Lagrangian for our action, taking

$$S = \kappa \int d^3x \ \mathrm{Tr}\Big(\varepsilon^{ij} A_j \dot{A}_i\Big)$$

Plugging in the representation of our gauge fields, we have

$$\begin{aligned}
S &= \kappa \int d^3x \ \mathrm{Tr}\Big(\varepsilon^{ij} A_j \dot{A}_i\Big) \\
&= \kappa \int d^3x \ \mathrm{Tr}\left(\varepsilon^{ij}\Big(-\partial_j U U^{-1} + U\theta_j U^{-1}\Big)\Big(-\partial_0\big(\partial_i U U^{-1}\big) + \partial_0\big(U\theta_i U^{-1}\big)\Big)\right) \\
&= \kappa \int d^3x \ \mathrm{Tr}\left(\varepsilon^{ij}\Big(\partial_j U U^{-1}\partial_0\big(\partial_i U U^{-1}\big) - \partial_j U U^{-1}\partial_0\big(U\theta_i U^{-1}\big) - U\theta_j U^{-1}\partial_0\big(\partial_i U U^{-1}\big)\right. \\
&\qquad\qquad \left. + U\theta_j U^{-1}\dot{U}\theta_i U^{-1} + U\theta_j\dot\theta_i U^{-1} + U\theta_j\theta_i\dot{U}^{-1}\Big)\right)
\end{aligned}$$

Here, we note that because $\partial_0(U U^{-1}) = 0 = \dot{U}U^{-1} + U\dot{U}^{-1}$, we have

$$U\theta_j U^{-1}\dot{U}\theta_i U^{-1} = \theta_j U^{-1}\dot{U}\theta_i = -U\theta_i\theta_j\dot{U}^{-1}$$

Which, combined with some integrations by parts, gives us the integrand

$$\begin{aligned}
&\mathrm{Tr}\left(\varepsilon^{ij}\Big(\partial_j U U^{-1}\partial_0\big(\partial_i U U^{-1}\big) + \partial_0\big(\partial_j U U^{-1}\big)U\theta_i U^{-1} - \partial_0\big(\partial_i U U^{-1}\big)U\theta_j U^{-1} + \theta_j\dot\theta_i\Big)\right) \\
&= \mathrm{Tr}\left(\varepsilon^{ij}\Big(\partial_j U U^{-1}\partial_0\big(\partial_i U U^{-1}\big) + 2\partial_0\big(\partial_j U U^{-1}\big)U\theta_i U^{-1} + \theta_j\dot\theta_i\Big)\right)
\end{aligned}$$

Expanding the terms further, we see that the middle term becomes

$$2\partial_0\big(\partial_j U U^{-1}\big)U\theta_i U^{-1} = 2\Big(\partial_j\dot{U}U^{-1} + \partial_j U\dot{U}^{-1}\Big)U\theta_i U^{-1}$$

With more integration by parts, this becomes

$$2\left(\partial_j \dot{U}\theta_i U^{-1} + \partial_j U^{-1}\dot{U}\theta_i\right) = 2\left(U^{-1}\partial_j \dot{U}\theta_i - U^{-1}\partial_j \dot{U}\theta_i\right) = 0$$

Our action therefore becomes

$$S = \kappa \int d^3x \,\mathrm{Tr}\,\varepsilon^{ij}\left(\partial_j UU^{-1}\partial_0\left(\partial_i UU^{-1}\right) + \theta_j \dot{\theta}_i\right)$$

However, we have separated our loop variables θ_i from the gauge variables. We can integrate the remaining "pure gauge" term, leaving us with an effective action (up to a constant) [7]

$$S = \kappa \int dt \,\mathrm{Tr}\left(\varepsilon^{ij}\theta_j \dot{\theta}_i\right) = \kappa \int dt \,\mathrm{Tr}\left(\theta_2 \dot{\theta}_1 - \theta_1 \dot{\theta}_2\right) = \frac{\kappa}{2}\int dt\left(\theta_2^a \dot{\theta}_1^a - \theta_1^a \dot{\theta}_2^a\right) = \kappa \int dt\,\theta_2^a \dot{\theta}_1^a$$

In this form, we have eliminated entirely the gauge variables, and end up with a simple action that strongly resembles previous forms of the Lagrangian we've come across in our analysis. We can let $\theta_2 = p$ and $\theta_1 = q$ just as with A_2, A_1 before. This gives us a commutator

$$[\theta_1^a, \theta_2^b] = \frac{i}{\kappa}\delta_{ab}$$

And a sense of θ_2 as a functional derivative with respect to θ_1:

$$\theta_2^a = \frac{1}{i\kappa}\frac{\delta}{\delta\theta_1^a}$$

To construct the states, we first note that the two loop variables, θ_i, are compact Lie algebra (in particular, $\mathfrak{su}(N)$) valued variables: they roughly represent translation of the gauge fields about the two non-contractible loops of the torus. From basic quantum mechanics (e.g. particle in a box), we can say that having a compact position variable (i.e. θ_1) leads to quantized momentum (i.e. θ_2). However, the momentum is also compact: this tells us that the position variable must also be quantized. This presents us the problem of defining a lattice and lattice spacings for the two interconnected variables. Given the form of the momentum operator, our momentum eigenstates will be of the form

$$\psi(\theta_1) = \exp\left(i\lambda \cdot \theta_1\right)$$

for some eigenvalue λ corresponding to the momentum. Furthermore, because the space θ_1 is compact, we ought to have a set of vectors α such that the states corresponding to θ_1 and $\theta_1 + 2\pi\alpha$ are identical. This gives us a condition

$$\exp\left(i\lambda \cdot \theta_1\right) = \exp\left(i\lambda \cdot (\theta_1 + 2\pi\alpha)\right) = \exp\left(i\lambda \cdot \theta_1 + 2\pi i\lambda \cdot \alpha\right)$$

$$\text{therefore}\quad \lambda \cdot \alpha \in \mathbb{Z}$$

Given that the two spaces θ_1 and θ_2 are both compact and quantized, as well as Lie algebra valued, we can guess that the lattices and lattices spacings for these variables will be related to the weight and root lattices of the algebra. In fact, we can postulate that the lattice spacing vectors α above are on the root lattice (this guess, which also just so happens to be correct, follows from the fact that with respect to the weights, the roots act as lattice spacings already; furthermore, the roots play a role in the integration measure on the group [1], so it naturally follows that they would play a significant role in other realms) and derive the properties of the momenta λ and their corresponding eigenstates. In fact, it follows immediately from the assumption that α is a root and the requirement that $\lambda \cdot \alpha \in \mathbb{Z}$ that λ is on the weight lattice. To see this, we take a brief tangent to our old friend: group theory.

43

5.7.1 A Note on Roots

To explore the above assertion, we need a few facts about roots. We will be brief here, but for a more detailed approach see [2]. We will quickly state a few facts:

1. If α is a root, then $-\alpha$ is also a root.

2. If L_α is the root space corresponding to α and $L_{-\alpha}$ the root space for $-\alpha$, then there exist $x \in L_\alpha$ and $y \in L_{-\alpha}$ such that the subalgebra $\mathrm{Span}(x, y, [x, y])$ is isomorphic to $\mathfrak{sl}(2)$.

3. In fact, by this method, we can associate to each root α a subalgebra isomorphic to $\mathfrak{sl}(2)$, with basis $e_\alpha, f_\alpha, h_\alpha$ for $e_\alpha \in L_\alpha$, $f_\alpha \in L_{-\alpha}$, and h_α an element of the Cartan subalgebra.

This element h_α in the basis of the $\mathfrak{sl}(2)$ subalgebra is known as the *coroot of α* (it also happens to be unique). We now recall a few definitions. Notably, when we refer to a weight, we are referring to an element of the dual space of the Cartan subalgebra, i.e. the set of maps from the Cartan subalgebra to $\mathbb{C}$.

Note: Since we use the terms *weights* and *roots* frequently, this definition merits some clarification (we mentioned this before, when we introduced these notions originally): both the weights, as the eigenvalues of the elements of the Cartan subalgebra, and the roots, as the eigenvalues of the adjoint representation of the Cartan subalgebra, are technically *weights*. They are all maps from the Cartan subalgebra to $\mathbb{C}$. However, they have special properties that have resulted in these somewhat unfortunate names.

In particular, the weight lattice is defined to be the set of weights (or the set of elements of the dual space of the Cartan subalgebra, or the set of maps from the Cartan subalgebra to $\mathbb{C}$) such that when evaluated for h_α (from the $\mathfrak{sl}(2)$ subalgebra) for any root α, it is an integer [9]. Formally:

$$\text{the weight lattice} \quad \Lambda^w = \{\beta \in H^* \quad \text{such that} \quad \beta(h_\alpha) \in \mathbb{Z}\}$$

Both our weights (as the eigenvalues of the Cartan subalgebra) and the roots are elements of the weight lattice. In fact, the root lattice Λ^R is the lattice formed by only the roots of the algebra. In this manner, the root lattice $\Lambda^R \subset \Lambda^w$.

We also recall our discussion of the simple roots: a set of roots such that every root can be written as a linear combination of these roots (there are stronger classifications, but this is sufficient for our discussion). We also recall our definition of the fundamental weights: a set of weights classified by their orthogonality to the simple roots. Furthermore, we remember that all highest weights of irreducible representations could be written in terms of the fundamental weights.

In the abstract, we can perform a similar procedure for the weight lattice. We can define the fundamental weights ω_i of the weight lattice as the set of weights in the dual space of the Cartan subalgebra such that

$$\omega_i(h_{\alpha_j}) = \delta_{ij}$$

In fact, any weight on the weight lattice can be written as a linear combination of the fundamental weights. More specifically, any weight $\beta \in \Lambda^w$ can be written

$$\beta = \sum \beta(h_{\alpha_i})\omega_i$$

This becomes even more specific for the roots. Any simple root can be written

$$\alpha^j = \sum \alpha^j (h_{\alpha_i}) \omega_i = \sum A_{ji} \omega_i$$

where A_{ji} is the Cartan matrix we introduced so long ago. In the particular case of $\mathfrak{su}(N)$, these fundamental weights ω_i are equal to the set of fundamental weights μ_i we found in our discussion of the root space classification [9]. We now have enough information to show the association of the momenta and lattice spacings of the states to the weight and root lattices.

We need $\lambda \cdot \alpha \in \mathbb{Z}$ for some momenta λ and root α. We can write each in terms of a corresponding basis. If λ is on the weight lattice, we can write it in terms of the fundamental weights:

$$\lambda = \sum \lambda(h_{\alpha_i}) \mu_i$$

but we have $\lambda(h_{\alpha_i}) \in \mathbb{Z}$ by definition of the weight lattice. Thus, this simplifies to

$$\lambda = \sum c_i \mu_i \quad \text{for} \quad c_i \in \mathbb{Z}$$

Furthermore, we can write the roots in terms of the simple roots:

$$\alpha_i = \sum_j d^i_j \alpha^j \quad \text{for} \quad d^i_j \in \mathbb{Z}$$

for simple roots α^j and integer coefficients (see the discussion of the root spaces). The dot product of these two will be

$$\lambda \cdot \alpha_i = \sum_k c_k d^i_k (\mu_k \cdot \alpha^k)$$

However, we defined our fundamental weights by their orthogonality to the simple roots. Thus, this becomes simply

$$\lambda \cdot \alpha_i = \sum_k c_k d^i_k \in \mathbb{Z}$$

as needed. It is thus appropriate to associate the momenta λ with the elements of the weight lattice, given a sense of periodicity with respect to the roots. However, this is not entirely complete. We must also include a sense of compactness on the momenta. If we shift θ_2 by the same quantity $2\pi\alpha$, we have

$$\left(\theta_2 + 2\pi\alpha \right) \exp \left(i\lambda \cdot \theta_1 \right) = \left(\frac{1}{i\kappa} \frac{\delta}{\delta\theta_1} + 2\pi\alpha \right) \exp \left(i\lambda \cdot \theta_1 \right) = \left(\frac{\lambda}{\kappa} + 2\pi\alpha \right) \exp \left(i\lambda \cdot \theta_1 \right)$$

If we are to associate the momenta also with momenta shifted by this "periodic" condition, then we would require that

$$\frac{\lambda}{\kappa} + 2\pi\alpha \equiv \lambda$$

Alternatively, we get the same result by

$$\theta_2 \exp \left(i \left(\lambda + 2\pi\kappa\alpha \right) \cdot \theta_1 \right) = \frac{1}{i\kappa} \frac{\delta}{\delta\theta_1} \exp \left(i \left(\lambda + 2\pi\kappa\alpha \right) \cdot \theta_1 \right) = \left(\frac{\lambda}{\kappa} + 2\pi\alpha \right) \exp \left(\dots \right)$$

Thus, we see that the action of shifting $\theta_2 \to \theta_2 + 2\pi\alpha$ is equivalent to shifting λ to $\lambda + 2\pi\kappa\alpha$. This tells us we ought to associate our momenta λ not as elements of the weight lattice, but

rather as elements of the quotient set of the weight lattice modulo the root lattice times a constant, $2\pi\kappa\Lambda^R$. Furthermore, in order to maintain gauge invariance, we must associate states whose momenta gauge transformations of each other. This corresponds to modding out the action of the Weyl group W (see our discussion of the root spaces) [7]. We now have our final classification of the momentum eigenstates of non-abelian Chern-Simons theory on the torus. They are given by wave-functions

$$\psi(\lambda) = \exp\left(i\lambda \cdot \theta_1\right)$$

such that λ is an element of

$$\lambda \in \frac{\Lambda^w}{W \ltimes 2\pi\kappa\Lambda^R}$$

5.8 It Always Comes Back to Group Theory

By themselves, these momenta labels seem somewhat meaningless. The derivation of them is somewhat abstract, and they certainly don't seem to have much context. However, a closer look at the momenta reveals, at least, some interesting properties.

The momentum eigenstates are specified by their momentum λ. However, these momenta are in direct correspondence to elements of the weight lattice modulo the roots. This is, in fact, equivalent to saying the momenta are in direct correspondence with the highest weights. Since adding roots to the momenta doesn't change them (with respect to the states), we are free to shift up whichever weight we start with using the roots until it is the highest weight without changing the state. It is thus, in a manner of speaking, sufficient to associate the momentum eigenstates with the highest weights of $\mathfrak{su}(N)$.

However, this result is quite potent. We have already established so much of the theory of $\mathfrak{su}(N)$ that from here we can draw a few rough conclusions. For one thing, the structure of $\mathfrak{su}(N)$ lends to the states a sense of being "fermionic". The highest weights are in a one to one correspondence with the irreducible representations, as we already established so long ago. Furthermore, the irreducible representations are in one to one correspondence with their characters (from basic group theory), which we already showed in some detail were equivalent to fermionic states. In this sense, the states of Chern-Simons theory on a torus correspond (rather roughly) to fermionic particles.

Furthermore, we can associate these states to Young tableaux. Since they correspond to the highest weights, we can extend that correspondence to the Young tableaux. Not only does this reinforce the sense that the states are related to the characters of $\mathbf{SU}(N)$ (and thus fermionic states), it also grants us powerful tools with which to study and classify the states.

There is much more subtlety and nuance to this relationship, which is explored fully in [10]. Unfortunately, it is not possible to include and discuss in detail the conclusions of [10], but the curious reader ought to give it a glance.

6   Conclusion

Throughout this discussion, we explored the classical theory of $\mathfrak{su}(N)$, even some of the fundamentals of the general theory of Lie algebras (hopefully gaining some appreciation for the beauty of the theory in the process). In exploring the uniquely physical properties of the characters of its group, **SU**(N), we explored only a the beginning of the powerful applications of the special unitary group in physics. In fact, we were able to reinforce this relationship with a brief study of low dimensional Yang-Mills theory.

We then turned to Chern-Simons theory: a much more complicated, and far less obviously physical, field theory. We quantized its various forms, studied its states, and drew what conclusions we could. Perhaps most importantly, we tied it back to the group theory we so carefully established, further reinforcing the potency of the theory of $\mathfrak{su}(N)$. The applications and uses of Chern-Simons theory are varied and fascinating, and most of them far exceed the reach of this discussion. Perhaps this study will open the door to a more complete understanding.

The goal of this discussion never was to present something novel. In fact, a large part of the material we discussed could be considered "textbook". Rather, the purpose lies in the intermediate steps, in showing the reader the detailed calculations and process that go into the well known conclusions, to build a sense of understanding and completeness. And, perhaps, to impart to the reader a bit of that same understanding.

7   Bibliography

References

[1] A. Jevicki, "Lectures on Group Theory", Brown University (lecture notes)

[2] K. Erdmann and M. Wildon, *Introduction to Lie Algebras*, (Springer, London 2006)

[3] H. Georgi, *Lie Algebras in Particle Physics*, (Westview Press, 1999)

[4] A. Jevicki, "Non-perturbative collective field theory", *Nucl. Phys. B* **376** (1992) 75-98.

[5] "Yang Mills Theory and Loop Space" (author unknown, lecture notes)

[6] G. Dunne, "Aspects of Chern-Simons Theory", arXiv:hep-th/9902115v1.

[7] Junggi Yoon, "(2+1) Dimensional Non-Abelian Chern-Simon Theory"

[8] Chen Ning Yang, "Condition of Self-Duality for SU(2) Gauge Fields on Euclidean Four-Dimensional Space", *Phys. Rev. Lett.* **38** (1977) 1377.

[9] P. Woit, "Topics in Representation Theory: Fundamental Representations and Highest Weight Theory", Columbia (lecture notes)

[10] M. Douglas, "Chern-Simons-Witten Theory as a Topological Fermi Liquid", arXiv:hep-th/9403119v1