University of Rochester

Group Theory for Physicists
Professor Sarada Rajeev

Characterization of $SU(N)$

David Mayrhofer

PHY 391 Independent Study Paper
December 13th, 2019

1 Introduction

At this point in the course, we have discussed $SO(N)$ in detail. We have determined the Lie algebra associated with this group, various properties of the various reducible and irreducible representations, and dealt with the specific cases of $SO(2)$ and $SO(3)$. Now, we work to do the same for $SU(N)$. We determine how to use tensors to create different representations for $SU(N)$, what difficulties arise when moving from $SO(N)$ to $SU(N)$, and then delve into a few specific examples of useful representations.

2 Review of Orthogonal and Unitary Matrices

2.1 Orthogonal Matrices

When initially working with orthogonal matrices, we defined a matrix O as orthogonal by the following relation

$$O^T O = \mathbb{1} \tag{1}$$

This was done to ensure that the length of vectors would be preserved after a transformation. This can be seen by

$$v \to v' = Ov \implies (v')^2 = (v')^T v' = v^T O^T O v = v^2 \tag{2}$$

In this scenario, matrices then must transform as $A \to A' = OAO^T$, as then we will have

$$\begin{aligned}(Av)^2 \to (A'v')^2 = (OAO^T Ov)^2 = (OAO^T Ov)^T (OAO^T Ov) \\ = v^T O^T O A^T O^T O A O^T O v = v^T A^T A v = (Av)^2\end{aligned} \tag{3}$$

Therefore, when moving to unitary matrices, we want to ensure similar conditions are met.

2.2 Unitary Matrices

When working with quantum systems, we not longer can restrict ourselves to purely real numbers. Quite frequently, it is necessarily to extend the field we are with with to the complex numbers. For instance, when working with a wave function $|\psi\rangle = \sum_i \psi^i |i\rangle$ where $\psi^i \in \mathbb{C}$, we need $\langle\psi|\psi\rangle = \sum_i |\psi^i|^2$ to be invariant. Transformations U must then satisfy

$$|\psi\rangle \to |\psi'\rangle = U|\psi\rangle \implies \langle\psi'|\psi'\rangle = \langle U\psi|U\psi\rangle = \langle\psi|U^\dagger U|\psi\rangle = \langle\psi|\psi\rangle, \tag{4}$$

where the last equality is satisfied if and only if $U^\dagger U = \mathbb{1}$. This is the condition for a matrix to be unitary. We first confirm that these U do form a group. We can see that they are closed by taking A and B both unitary and showing their product is in fact unitary.

$$(AB)^\dagger(AB) = B^\dagger A^\dagger AB = \mathbb{1} \tag{5}$$

Since $1^\dagger 1 = (1)1 = 1$, the identity matrix is unitary. To determine if each element's inverse is in the group, we need to show that the adjoint of each unitary matrix is unitary,

$$(U^\dagger)^\dagger U^\dagger = UU^\dagger = \mathbb{1} \tag{6}$$

Since associativity of these matrices follows from the general associativity of matrix multiplication, the unitary matrices are a group.

2.3 Special Unitary Matrices

When dealing with orthogonal transformations, to exclude reflection, we required that the determinant of any matrix $\det O = +1$. This subset of $O(N)$ was designated as the special orthogonal matrices, $SO(N)$. We require a similar condition when dealing with unitary matrices. We can determine the determinant of any unitary matrix by noting

$$\begin{aligned}\det \mathbb{1} = \det U^\dagger U = \det U^\dagger \det U = (\det U)^* \det U = |\det U|^2 = 1 \\ \implies \det U = e^{i\varphi}\end{aligned} \tag{7}$$

The determinant of U therefore reduces to some phase factor. We can require that $\varphi = 0$, giving the special unitary matrices with $\det U = 1$. We can then break $U(N)$ into two parts: $SU(N)$, and the identity matrix scaled by $e^{i\varphi}$. Note that there is an overlap between these two section of $U(N)$, as if we have are dealing with the group $SU(d)$, $e^{2\pi i \frac{n}{d}}\mathbb{1}$ will both have determinant one and be a scaling of the identity matrix if $n \in \mathbb{Z}$.

3 Tensor Representations of $SU(N)$

We build higher dimensional representation of $SU(N)$ by first starting with fundamental representation of the group, consisting of the $N \times N$ special unitary matrices. We then take some tensor m indices $\varphi^{i_1 i_2 \cdots i_m}$. This tensor transforms under under a unitary operation by $\varphi^{i_1 i_2 \cdots i_m} \to \varphi'^{i_1 i_2 \cdots i_m} = U^{i_1 j_1} \ldots U^{i_m j_m} \varphi^{j_1 \cdots j_m}$. This tensor therefore will provide a representation of $SU(N)$. To make this more clear, we work through an example of a totally symmetric tensor φ^{ijk} that we will use to represent $SU(3)$. This tensor has ten independent components. We can count these components by noting there are 3 independent components that correspond to each index being equal, 1 independent component that corresponds to each index being different, and 6 independent components that correspond to two indices being the same.

This group solved a puzzle in particle physics in the 1960s. There were nine baryonic particles of similar mass, which we can expect to be associated with some symmetry group. Gell-Mann knew that the strong interaction's symmetric group is $SU(3)$, so he guessed that there must be a tenth particle that was similar to these nine known particles to correspond to this ten dimensional representation of $SU(3)$. When he was proven correct, particle physicists began to more uniformly embrace group theory.

3.1 Contractions of Tensors Used to Represent $SU(N)$

Naturally for such tensors, we may need to take a contraction over some collection of indices. However, we cannot do so in exactly the same way as we did for the tensors representing $SO(N)$. Recall for some tensor T^{ijk} that furnishes a representation of $SO(N)$, we can contract over two of the indices by taking $T^{iik} = \delta^{ij}T^{ijk}$. This quantity will then transform under a rotation as

$$\begin{aligned}\delta^{ij}T^{ijk} \to \delta^{ij}T'^{ijk} &= \delta^{ij}O^{if}O^{jg}O^{kh}T^{fgh} = (\delta^{ij}O^{if}O^{jg})(O^{kh}T^{fgh}) \\ &= (O^T)^{fi}\delta^{ij}O^{jg}(O^{kh}T^{fgh}) = (O^TO)^{fg}(O^{kh}T^{fgh}) \\ &= \delta^{fg}O^{kh}T^{fgh}\end{aligned} \tag{8}$$

However, the last inequality only holds because $O^TO = \mathbb{1}$, which is not the case for unitary matrices. Instead, we need to determine how to define the contraction so that tensor transforms as it does in the case of $SO(N)$.

To do this, we first define a contravariant and a covariant of our vector, $\psi_i = \psi^{i*}$. Then, the inner product in quantum mechanics can be defined as

$$\langle\phi|\psi\rangle = \left(\sum_j \langle j|\phi^{j*}\right)\left(\sum_i \psi^i|i\rangle\right) = \psi^i\phi_i \tag{9}$$

Now, when we do a transformation we take

$$\psi^i \to \psi'^i = U^i{}_j\,\psi^j, \tag{10}$$

where we have lowered the second index to match the upper index on the ψ^j. The covariant then transforms as

$$\psi_i \to \psi'_i = \psi_j(U^\dagger)^j{}_i \tag{11}$$

Note that the inner product is still preserved,

$$\langle\phi|\psi\rangle \to \langle\phi'|\psi'\rangle = \langle U\phi|U\psi\rangle = \phi_i(U^\dagger)^j{}_i U^i{}_k\,\psi^k = \psi_j\delta^j{}_k\psi^k = \phi_i\psi^i \tag{12}$$

Generally speaking, we can now have a tensor with m upper and and n lower indices, $\varphi^{i_1\cdots i_m}_{j_1\cdots j_n}$. Then the tensor will transform as

$$\varphi^{ij}_k \to \varphi'^{ij}_k = U^i{}_l\,U^j{}_m\,(U^\dagger)^n{}_k\varphi^{lm}_n \tag{13}$$

We can now take a contraction on two of the indices. Note that under this formalism, we must take the contraction over one upper and lower index.

$$\varphi^{ij}_j = U^i{}_j\,U^j{}_m\,(U^\dagger)^n{}_j\varphi^{lm}_n = U^i{}_l\,\delta^n{}_m\varphi^{lm}_n = U^i{}_l\,\varphi^{lm}_m \tag{14}$$

3.2 Raising and Lowering Indices

We know by definition for any $U \in SU(N)$, we must have $\det U = 1$. However, introducing contravariant and covariant vectors now mean there are two ways to take the determinant

$$\epsilon_{i_1 i_2 \cdots i_N} U_1^{i_1} U_2^{i_2} \cdots U_N^{i_N} = 1, \tag{15}$$
$$\epsilon^{i_1 i_2 \cdots i_N} U_{i_1}^{1} U_{i_2}^{2} \cdots U_{i_N}^{N} = 1 \tag{16}$$

We can also express the first equation as

$$\epsilon_{i_1 i_2 \cdots i_N} U_{j_1}^{i_1} U_{j_2}^{i_2} \cdots U_{j_N}^{i_N} = \epsilon_{j_1 j_2 \cdots j_N} \tag{17}$$

If we multiply this by $(U^\dagger)^{j_N}_{p_N}$, then we can eliminate one of the matrices on the right hand side to get

$$\epsilon_{i_1 i_2 \cdots i_{N-1} p_N} U_{j_1}^{i_1} U_{j_2}^{i_2} \cdots U_{j_{N-1}}^{i_{N-1}} = \epsilon_{j_1 j_2 \cdots j_N} (U^\dagger)^{j_N}_{p_N} \tag{18}$$

By repeating this process, we can write the exact same relationship in Eq. (17) using $U^\dagger$ instead of U.

We can then use this to exchange upper and lower indices of tensors. For instance, we can define a tensor $\varphi_{kpq} = \epsilon_{ijpq} \varphi_k^{ij}$. When transforming this tensor, we have

$$\varphi_{kpq} \to \varphi'_{kpq} = \epsilon_{ijpq} U^i{}_l U^j{}_m (U^\dagger)^n{}_k \varphi_n^{lm} = (U^\dagger)^s{}_p (U^\dagger)^t{}_p (U^\dagger)^n{}_k (\epsilon_{lmst} \varphi_n^{lm}) = (U^\dagger)^s{}_p (U^\dagger)^t{}_p (U^\dagger)^n{}_k \varphi_{nst}, \tag{19}$$

which is exactly how we expect this tensor to transform.

3.3 Symmetric and Anti-Symmetric Tensor Representations

As we have defined it here, the symmetry or anti-symmetry of a tensor is not changed by group transformations. For our purposes, we restrict our attention to strictly tensors that are strictly symmetric or anti-symmetric under exchange of upper or lower indices. Furthermore, since the contraction transforms like a vector, we can subtract it off of any tensor we choose to provide our representation. With these restrictions, we can determine the dimension of the representation provided by, for example, a symmetric tensor $S_j^{ij} = +S_k^{ji}$ to be $\frac{1}{2}N^2(N+1) - N = \frac{1}{2}N(N-1)(N+2)$. Similarly, we can determine the dimension of the representation provided by the antisymmetric tensor$A_k^{ij} = -A_k^{ji}$ to be $\frac{1}{2}N^2(N-1) - N = \frac{1}{N}(N-2)(N+1)$.

These representation have several naming conventions. One such convention labels representation by the number of indices above and below the tensor. For example, a tensor with m upper indices and n lower indices would be labeled as (m, n). Often, the symmetry of the tensor is denoted by using brackets for antisymmetric tensors and curly braces for symmetric tensors. For instance, a symmetric tensor S_k^{ij} would be labeled as {2,1}. If there are no upper or lower indices, then that term may suppressed, with an asterisk added if there are no upper indices. For instance, $[2, 0]$ can be written as $[2]$ and $[0, 2]$ as $[2]^*$.

4 Construction of the Lie Algebra

We begin the same way as in $SO(N)$, by taking the infinitesimal transformation $U \approx \mathbb{1} + iH$, where H is some matrix with complex entries. Then, using the defining relation of the unitary matrices we have

$$U^\dagger U = \mathbb{1} \approx (\mathbb{1} - iH^\dagger)(\mathbb{1} + iH^\dagger) \approx \mathbb{1} - i(H^\dagger - H) \tag{20}$$

$$\implies H^\dagger = H \tag{21}$$

As before, we can use the infinitesimal transformations to determine that for any unitary U, we can write $U = e^{iH}$, for some Hermitian H. We can check that any matrix of the form e^{iH} is in fact unitary by evaluating its Taylor series

$$(e^{iH})^\dagger = \left(\sum_{k=0}^{\infty} \frac{(iH)^k}{k!}\right)^\dagger = \sum_{k=0}^{\infty} \frac{(-iH^\dagger)^k}{k!} = \sum_{k=0}^{\infty} \frac{(-iH)^k}{k!} = e^{-iH} \tag{22}$$

$$\implies e^{iH}(e^{iH})^\dagger = e^{iH}e^{-iH} = \mathbb{1} \tag{23}$$

We now want to determine what conditions we must impose on this form of U to ensure $\det U = 1$. To do this, we first note that we can write any matrix as the product of its eigenvalues. If H has eigenvalues λ_j, then e^{iH} has eigenvalues $e^{i\lambda_j}$. We can show this by evaluating the Taylor series expansion of e^{iH} applied to an eigenvector of H, say v_j

$$e^{iH}v_j = \left(\sum_{k=0}^{\infty} \frac{(iH)^k}{k!}\right) v_j = \sum_{k=0}^{\infty} \frac{(iH)^k v_j}{k!} = \left(\sum_{k=0}^{\infty} \frac{(i\lambda_j)^k}{k!}\right) v_j = e^{i\lambda_j} v_j \tag{24}$$

We will also make use of the fact that the trace of any matrix is the sum of its eigenvalues. We can then write the determinant of U as

$$\det U = \det e^{iH} = \prod_{j=1}^{N} e^{i\lambda_j} = e^{i\sum_{j=1}^{N} \lambda_j} = e^{i \operatorname{tr} H} \tag{25}$$

For $\det U = 1$, we can then require $\operatorname{tr} H = 0$. With this, we now know that we only need to consider traceless, Hermitian matrices. For 2×2 matrices, this has general form

$$H = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}, \tag{26}$$

where the x_i are all real numbers. We can also write this matrix in terms of the Pauli matrices,

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{27}$$

Then write this matrix H can be written as

$$H = x_i \sigma^i \tag{28}$$

If we take $N = 3$, then the equivalent of the Pauli matrices are the Gell-Mann matrices. We can note in both cases, and in the general case, we have $N^2 - 1$ matrices. To determine the number of independent components needed to characterize one group, we count the number of such components on the diagonal. Since there the diagonal must be real, each component contributes has only one free real number, and the condition that the matrix is traceless give a total of $N - 1$ free parameters from the diagonal. The off diagonal terms contribute $N(N - 1)$ real parameters, but since the lower diagonal is the complex conjugate of the upper diagonal, there are only $\frac{1}{2}N(N-1)$ free terms. We then need a total of $N - 1 + \frac{1}{2}N(N-1) = N^2 - 1$ free real numbers to characterize the group $SU(N)$. Therefore, we can write any $N \times N$ Hermitian, traceless matrix in terms of $N^2 - 1$ real numbers and and generators. The elements of $SU(N)$ can be then written as

$$U = e^{i\theta^a T^a}, \tag{29}$$

where the θ^a are the real numbers and T^a are the generators.[1]

5 Determining the Structure Constants of the Lie Algebra

We can measure the commutativity of two arbitrary unitary matrices $U_1 \approx \mathbb{1} + A$ and $U_2 \approx \mathbb{1} + B$ by taking

$$U_2^{-1} U_1 U_2 \approx (\mathbb{1} - B)(\mathbb{1} + A)(\mathbb{1} + B) \approx \mathbb{1} + A + AB - BA = U_1 + [A, B] \tag{30}$$

The quantity then deviates from U_1 by $[A, B]$. We write $A = i\sum_a \theta^a T^a$ and $B = i\sum_b \theta'^b T^b$ so the commutator takes the form

$$[A, B] = -\sum_{a,b} \theta^a \theta'^b [T^a, T^b] \tag{31}$$

Since the all of the T are traceless, Hermitian matrices, the commutator is traceless and antiHermitian. Therefore, there must be some linear combination of the generators that can be used to write commutator,

$$[T^a, T^b] = i f^{abc} T^c \tag{32}$$

These f^{abc} are the fine structure constants of the Lie algebra. They are by definition of the commutator antisymmetric.

We can think of these T^a either as explicit matrices, or as any objects with satisfy the above commutation relationship. The advantage of the latter is that

[1] Here, the superscripts a are simply labels to denote the θ and T, and therefore have no associated notion of contravariance or covariance. I will generally use superscripts for such objects, but there is no difference between θ_a and θ^a, T_a and T^a, etc. Einstein summation notation is still assumed in these cases.

it is easier to consider higher dimensional representations of SU(N) than the fundamental $N \times N$ representation. We can easily determine the generator T^a by taking the derivative of $U = \mathbb{1} + i\theta^a T^a$ with respect to θ^a. Also note we can write the variation in a tensor as

$$\delta\varphi^p = i\theta^a (T^a)^p{}_q \varphi^q, \tag{33}$$

where p and q are the indices of the matrix representing T^a.

5.1 The Adjoint Representation

We can calculate the fine structure constants using the adjoint representation of $SU(N)$ as well. This adjoint representation is furnished by a tensor $\varphi^i{}_j$. This representation will have a dimension of $N^2 - 1$. To see this, we first note that this representation transforms as

$$\varphi^i_j \to \varphi'^i_j = U^i{}_l \varphi^l_n (U^\dagger)^n{}_j, \tag{34}$$

i.e., it transforms in the same way as a matrix. We note that if φ^i_j is traceless and Hermitian, we can then also write

$$\varphi^i_j = A^a (T^a)^i_j \tag{35}$$

We know that there must be $N^2 - 1$ total A^a, as there are $N^2 - 1$ total T^a. However, since there are $N^2 - 1$ such A^a, we can use these to provide another representation for $SU(N)$. We then want work out how these A^a transform. First, we determine how a general representation φ transforms

$$\begin{aligned}\varphi \to \varphi' &\approx (\mathbb{1} + i\theta^a T^a)\varphi(\mathbb{1} + i\theta^a T^a)^\dagger \approx \varphi + i\theta^a T^a - \varphi i\theta^a T^a \\ &= \varphi + i\theta^a [T^a, \varphi]\end{aligned} \tag{36}$$

The variational in φ is then

$$\delta\varphi = i\theta^a [T^a, \varphi] \tag{37}$$

Now, to determine how these A^a specifically transform, we take

$$(\delta A^b) T^b = \delta(A^b T^b) = i\theta^a [T^a, A^c T^c] = i\theta^a A^c [T^a, T^c] = i\theta^a A^c i f^{acb} T^b \tag{38}$$

However, from Eq. (35), we have

$$\delta A^b = i\theta^a (T^a)^b{}_c A^c \tag{39}$$

Therefore, we get

$$(T^a)^{bc} = -i f^{abc} \tag{40}$$

This means that to determine the fine structure constant, we simply need to find the adjoint representation of the generators.

6 Conclusion

We have effectively reproduced most of the results that were derived earlier for $SO(N)$, but now for $SU(N)$. We see that many of the results carry directly over. However, a few definitions such as the contraction must be adjusted in order to account for the complex nature of the matrix entries. This is remedied by the introduction of contravariant and covariant vectors which are the complex conjugates of one another. The Lie algebra is formed by taking unitary matrices infinitesimally close to the identity, and we can in turn develop its generators. These generators have the defining relation $[T^a, T^b] = if^{abc}T^c$. Lastly, we determined we can calculate the fine structure constant by determining the generators in the adjoint representation of $SU(N)$.

7 References

[1] Zee, A., *Group Theory in a Nutshell for Physicists*, Chapter IV.4, Princeton University Press, 2016.