# Non-Abelian Gauge Invariance Notes

Physics 523, Quantum Field Theory II

Presented Monday, 5th April 2004

Jacob Bourjaily, James Degenhardt, & Matthew Ong

### Introduction and Motivation

Until now, all the theories that we have studied have been somewhat restricted to the 'canonical.' The Lagrangians of quantum electrodynamics, $\phi^4$-theory, Yukawa theory, etc. have been the starting points for our development–while we have learned to interpret them, we have not yet developed adequate methods to build Lagrangians for new theories.

On of the most important lessons of modern physics is that microscopic symmetries can be used to determine the form of physical interactions. In particular, the Standard Model of particle physics, $SU(3)\times SU(2)_L\times U(1)_Y$, has a form completely specified by its symmetries alone. Today, we will discuss some of the foundational principles from which this type of theory can be built.

We have seen already that the QED Lagrangian is invariant to a wide class of transformations. Perhaps the most striking of all is its invariance under arbitrary, position dependent phase transformations of the field $\psi$. But this should have been obvious. It is not surprising that phase differences between space-like separated points should be uncorrelated in the theory. However, a-causal regions become causal in time; therefore, there must be some way to re-correlate the arbitrary phases between points at future times. As you know, it is this argument of *local gauge invariance* in quantum electrodynamics that requires the existence of the vector field, $A_\mu$, the photon.

Clearly, the local $U(1)$ symmetry of the QED Lagrangian proves to be a powerful source of insight. Let us therefore consider promoting the concept of local gauge invariance to the highest level of physical principle. We may ask, for example, whether or not other local symmetries are observed in nature? What form must our Lagrangians take if they are to be manifestly gauge covariant?

In our presentation this afternoon, we will attempt to introduce the tools that must be used to deal with gauge symmetries. We will first review the construction of quantum electrodynamics from a simply $U(1)$ gauge principle and then extend these arguments to arbitrary $SU(n)$ symmetries. Because higher gauge groups are non-abelian, some care must be observed. We will finish our presentation with an introduction to the algebraic tools necessary to investigate gauge theories. The material of this discussion can be found in chapter 15 of Peskin and Schroeder.

### Geometry of Gauge Invariance

One of the most familiar Lagrangians should be that of quantum electrodynamics,

$$\mathscr{L}_{\text{QED}} = -\frac{1}{4}(F_{\mu\nu})^2 + \overline{\psi}(i\not{\partial} - m)\psi - e\overline{\psi}\gamma^\mu\psi A_\mu.$$

In chapter 4, we found that this Lagrangian is not only invariant under global gauge variations $\psi \to e^{i\alpha}\psi$, but to the very general, position-dependent transformation $\psi \to e^{i\alpha(x)}\psi$ if the vector field $A_\mu$ transforms simultaneously by $A_\mu \to A_\mu + \frac{1}{e}\partial_\mu\alpha(x)$. Transformations under which the phase of quantum fields are changed as arbitrary functions of position are called *local gauge transformations.* When a Lagrangian is invariant under a local gauge transformation, we say that it is *gauge covariant.*

Local gauge invariance turns out to be so powerful that the entire form of the QED Lagrangian is determined by it alone. Because quantum electrodynamics is very familiar to us, it will serve as an ideal place to begin our investigation of the consequences local gauge invariance. Because the local gauge transformation in QED is a scalar function $e^{i\alpha(x)}$, it commutes with everything. By starting with QED, we can gain a good deal of insight about the physical principles at work before rigorously exploring the symmetries of non-commutative gauge transformations.

### Derivation of $\mathscr{L}_{\text{QED}}$ from Local $U(1)$ Gauge Invariance

Let us consider the consequences of demanding that $\mathscr{L}_{\text{QED}}$ be invariant under the local gauge transformation $\psi(x) \to e^{i\alpha(x)}\psi(x)$ where $\psi$ is the standard fermion field. We see that we may already include terms like $m\overline{\psi}\psi$ because these will be manifestly covariant under any phase transformation.

What are we to make of derivative terms? Recall that, in general, the partial derivative of any field $\psi(x)$ taken in the direction of the vector $n^\mu$ can be defined by

$$n^\mu\partial_\mu\psi = \lim_{\epsilon\to 0}\frac{1}{\epsilon}[\psi(x+\epsilon n) - \psi(x)].$$

Note, however, that this derivative cannot be used for the transformed field. Because $\psi(x+\epsilon n) \to e^{i\alpha(x+\epsilon n)}\psi(x+\epsilon n)$ while $\psi(x) \to e^{i\alpha(x)}$, the subtraction–even in the limit of $\epsilon \to 0$–cannot be defined. This is because we require that our theory accommodate an *arbitrary* function $\alpha(x)$ and any subtraction technique will necessarily lose enormous generality.

We must therefore develop a consistent way to define what is meant to take the derivative of a field. We can do this by introducing a scalar function, $U(y,x)$, that compares two points in space. The scalar function $U(y,x)$ can be used to 'make up' for the ambiguity of $\alpha(x)$ by requiring that it transform according to

$$U(y,x) \to e^{i\alpha(y)}U(y,x)e^{-i\alpha(x)}.$$

Now the subtraction between two different points can be made consistently. This is because the combined entity $U(y,x)\psi(x)$ transforms identically to $\psi(y)$:

$$\psi(y) - \psi(x) \to e^{i\alpha(y)}\psi(y) - e^{i\alpha(y)}U(y,x)\psi(x).$$

Therefore, we may define the covariant derivative in the limiting procedure:

$$n^\mu D_\mu \psi = \lim_{\epsilon\to 0}\frac{1}{\epsilon}[\psi(x+\epsilon n) - U(x+\epsilon n, x)\psi(x)].$$

To get an explicit definition of $D_\mu$ we expand $U(y,x)$ at infinitesimally separated points.

$$U(x+\epsilon n, x) = 1 - ie\epsilon n^\mu A_\mu(x) + \mathcal{O}(\epsilon^2).$$

Here we have extracted $A_\mu$ from the Taylor expansion of the comparator function. This new vector field is called the connection. Putting this expansion into the rule for the transformation of the comparator, we find the transformation of $A_\mu$. Note that we have used the fact that $\epsilon n^\mu \partial_\mu \alpha(x) \sim \alpha(x+\epsilon n) - \alpha(x)$.

$$\begin{aligned}(1 - ie\epsilon n^\mu A_\mu(x)) &\to e^{i\alpha(x+\epsilon n)}(1 - ie\epsilon n^\mu A_\mu(x))e^{-i\alpha(x)},\\ &= e^{i\epsilon n^\mu\partial_\mu\alpha(x)+i\alpha(x)}(1 - ie\epsilon n^\mu A_\mu)e^{-i\alpha(x)},\\ &= 1 + i\epsilon n^\mu \partial_\mu\alpha(x) - ie\epsilon n^\mu A_\mu,\\ \therefore A_\mu(x) &\to A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x).\end{aligned}$$

Note that the transformation properties we demanded of our Lagrangian determined the form of the interaction field and the covariant derivative and required the existence of a non-trivial vector field $A_\mu$ .

**Gauge Invariant Terms in the Lagrangian**

We now have some of the basic building blocks of our Lagrangian. We may use any combination of $\psi$ and its covariant derivative to get locally invariant terms. Now we wish to find the kinetic term of the interaction field. To construct this term we first derive its invariance from $U(y,x)$. We expand $U(y,x)$ to $\mathcal{O}(\epsilon^3)$ and assume that it is a pure phase with the restriction $(U(x,y))^\dagger = U(y,x)$, thus

$$U(x+\epsilon n, x) = \exp[-ie\epsilon n^\mu A_\mu(x+\frac{\epsilon}{2}n) + \mathcal{O}(\epsilon^3)].$$

Now we link together comparisons f the phase direction around a small square in spacetime. This square lies in the (1,2)-plane and is defined by the vectors $\hat{1}$, $\hat{2}$ (see Fig. 15.1 in **PS** ) . We define $\mathbf{U}(x)$ to be this quantity.

$$\mathbf{U}(x) \equiv U(x, x+\epsilon\hat{2})U(x+\epsilon\hat{2}, x+\epsilon\hat{1}+\epsilon\hat{2}) \times U(x+\epsilon\hat{1}+\epsilon\hat{2}, x+\epsilon\hat{1})U(x+\epsilon\hat{1}, x).$$

$\mathbf{U}(x)$ is locally invariant and by substituting in the expansion of $U(x+\epsilon n, x)$ into $\mathbf{U}(x)$ we get a locally invariant function of the interaction field.

$$\mathbf{U}(x) = \exp\{-i\epsilon e[-A_2(x+\frac{\epsilon}{2}\hat{2}) - A_1(x+\frac{\epsilon}{2}\hat{1}+\epsilon\hat{2}) + A_2(x+\epsilon\hat{1}+\frac{\epsilon}{2}\hat{2}) + A_1(x+\frac{\epsilon}{2}\hat{1})] + \mathcal{O}(\epsilon^3)\}$$

This reduces when it is expanded in powers of $\epsilon$.

$$\mathbf{U}(x) = 1 - i\epsilon^2 e[\partial_1 A_2(x) - \partial_2 A_1(x)] + \mathcal{O}(\epsilon^3).$$

We can now see a hint of why $F_{\mu\nu}$ is gauge covariant. It is also very important to note that $A_\mu A^\mu$ is not locally invariant. Luckily however, $A_\mu$ commutes with $A_\mu$ with each other and so we do not have to worry much about this. In the next section though we won't assume that these vector fields commute. This also shows that there is no mass term for the photon.

Another way to define the field strength tensor $F_{\mu\nu}$ and to show its covariance in terms of the commutator of the covariant derivative. The commutator of the covariant derivative will be invariant in QED.

$$\begin{aligned}
[D_\mu, D_\nu]\psi(x) &= [\partial_\mu, \partial_\nu]\psi + ie([\partial_\mu, A_\nu] - [\partial_\nu, A_\mu]\psi) - e^2[A_\mu, A_\nu]\psi \\
[D_\mu, D_\nu]\psi(x) &= ie(\partial_\mu A_\nu - \partial_\nu A_\mu)\psi \\
[D_\mu, D_\nu] &= ieF_{\mu\nu}
\end{aligned}$$

One can visualize the commutator of covariant derivatives as the comparison of comparisons across a small square like we computed above. In reality, the arguments are identical.

**Building a Lagrangian**

We now have all of the ingredients that we can use to construct a general local gauge invariant Lagrangian. The most general Lagrangian with operators of dimension 4 is:

$$\mathscr{L}_4 = \overline{\psi}(i\partial_\mu\gamma^\mu + ieA_\mu)\psi - \frac{1}{4}(F_{\mu\nu})^2 - c\epsilon^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu} - m\overline{\psi}\psi.$$

If we want a theory that is $P$, $T$ invariant then we must set $c$ to zero since it would violate these discrete symmetries. This Lagrangian contains only two free parameters, $e$, and $m$. Higher dimensional Lagrangians are also generally allowed. Such as:

$$\mathscr{L}_6 = ic_1\overline{\psi}\sigma_{\mu\nu}F^{\mu\nu}\psi + c_2(\overline{\psi}\psi)^2 + c_3(\overline{\psi}\gamma^5\psi)^2 + \cdots .$$

While allowed by gauge symmetry, these terms are irrelevant to low energy physics and can therefore be ignored.

**The Yang-Mills Lagrangian**

We may generalize our work on local gauge invariance to consider theories which are invariant under a much wider class of local symmetries. Let us consider a theory of $\psi$, where $\psi$ is an $n$-tuple of fields given by

$$\psi \equiv \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix}.$$

We will investigate the consequences of demanding this theory be invariant under a local $SU(n)$ transformation $\psi \to V(x)\psi$ where

$$V(x) = e^{it_a\alpha^a(x)}.$$

Here, $\alpha^a(x)$ are arbitrary, differentiable functions of $x$ and the $t_a$'s are the generators of $su_n$.

It is important to note that in all but the most trivial case, the generators of the algebra do not commute $[t_a, t_b] \neq 0$. In particular, as will be explained later, $[t_a, t_b] = if_{abc}t_c$ where $f_{abc}$ are the structure constants of the theory.

As before, we must redefine the derivative so that it is consistent with arbitrary local phase transformations. To do this, we will introduce a comparator function $U(y, x)$ that transforms as

$$U(y, x) \to V(y)U(y, x)V^\dagger(x)$$

so that $U(x, y)\psi(y)$ transforms identically to $\psi(x)$. Therefore, we may define the covariant derivative of $\psi$ in the direction of $n^\mu$ by

$$n^\mu D_\mu\psi(x) \equiv \lim_{\epsilon\to 0} \frac{1}{\epsilon}\left[\psi(x + \epsilon n) - U(x + \epsilon n, x)\psi(x)\right].$$

We will set $U(x, x) = 1$ and we may restrict $U(y, x)$ to be a unitary matrix. Therefore, because $U(x + \epsilon n, x) \sim \mathcal{O}(1)$, we may expand $U$ in terms of the infinitesimal generators of $su_n$:

$$U(x + \epsilon n, x) = 1 + ig\epsilon n^\mu A^a_\mu t_a + \mathcal{O}(\epsilon^2)$$

where $g$ is a constant extracted for convenience. Therefore, we see that the covariant derivative is given by

$$\begin{aligned} n^\mu D_\mu\psi(x) &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\psi(x+\epsilon n) - \left(1 + ig\epsilon n^\mu A_\mu^a t_a + \mathcal{O}(\epsilon^2)\right)\psi(x)\right], \\ &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\psi(x+\epsilon n) - \psi(x)\right] - \frac{1}{\epsilon}\left(ig\epsilon n^\mu A_\mu^a t_a + \mathcal{O}(\epsilon^2)\right)\psi(x), \\ &= n^\mu\partial_\mu\psi(x) - \lim_{\epsilon\to 0}\left(ign^\mu A_\mu^a t_a + \mathcal{O}(\epsilon)\right)\psi(x), \\ &= n^\mu\left(\partial_\mu - igA_\mu^a t_a\right)\psi(x), \\ &\therefore D_\mu = \partial_\mu - igA_\mu^a t_a. \end{aligned}$$

It is clear that the covariant derivative *requires* a vector field $A_\mu^a$ for each of the generators $t_a$.

Let us compute how the gauge transformation acts on $A_\mu^a$. By a direct application of the definition of the transformation of $U$ and our expansion of $U(x+\epsilon n, x)$ near $\epsilon\to 0$, we see that

$$\begin{aligned} 1 + ig\epsilon n^\mu A_\mu^a t_a &\to V(x+\epsilon n)\left(1 + ig\epsilon n^\mu A_\mu^a t_a\right)V^\dagger(x), \\ &= V(x+\epsilon n)V^\dagger(x) + V(x+\epsilon n)ig\epsilon n^\mu A_\mu^a t_a V^\dagger(x), \\ &= \left[\left(1+\epsilon n^\mu\partial_\mu + \mathcal{O}(\epsilon^2)\right)V(x)\right]V^\dagger(x) + \left[\left(1+\epsilon n^\mu\partial_\mu + \mathcal{O}(\epsilon^2)\right)V(x)\right]ig\epsilon n^\mu A_\mu^a t_a V^\dagger(x), \\ &= 1 + \epsilon n^\mu\left(\partial_\mu V(x)\right)V^\dagger(x) + V(x)ig\epsilon n^\mu A_\mu^a t_a V^\dagger(x) + \mathcal{O}(\epsilon^2), \\ &= 1 + V(x)\epsilon n^\mu\left(-\partial_\mu V^\dagger(x)\right) + V(x)ig\epsilon n^\mu A_\mu^a t_a V^\dagger(x) + \mathcal{O}(\epsilon^2), \\ &= 1 + V(x)\left(ig\epsilon n^\mu A_\mu^a t_a - \epsilon n^\mu\partial_\mu\right)V^\dagger(x) + \mathcal{O}(\epsilon^2), \\ &\therefore A_\mu^a t_a \to V(x)\left(A_\mu^a t_a + \frac{i}{g}\partial_\mu\right)V^\dagger(x). \end{aligned}$$

For infinitesimal transformations, we can expand $V(x)$ in a power series of the functions $\alpha^a(x)$. Note that we should change our summation index $a$ on two of the sums below to avoid ambiguity. To first order in $\alpha$, we see

$$\begin{aligned} A_\mu^a t_a &\to V(x)\left(A_\mu^a t_a + \frac{i}{g}\partial_\mu\right)V^\dagger(x), \\ &= \left(1 + i\alpha^a(x)t_a\right)\left(A_\mu^b t_b + \frac{i}{g}\partial_\mu\right)\left(1 - i\alpha^c(x)t_c\right), \\ &= A_\mu^a t_a + i\left(\alpha^a(x)t_a A_\mu^b t_b - A_\mu^b t_b\alpha^a(x)t_a\right) + \frac{1}{g}\partial_\mu\alpha^a(x)t_a, \\ &= A_\mu^a t_a + \frac{1}{g}\partial_\mu\alpha^a(x)t_a + i\left[\alpha^a t_a, A_\mu^b t_b\right], \\ &\therefore A_\mu^a t_a \to A_\mu^a t_a + \frac{1}{g}\partial_\mu\alpha^a(x)t_a + f_{abc}A_\mu^b\alpha^c(x). \end{aligned}$$

As our work before showed, we can build more general gauge-covariant terms into our Lagrangian using the covariant derivative. As before, we find that the commutator of the covariant derivative is itself not a differential operator bet merely a multiplicative matrix acting on $\psi$. This implies that the commutator of the covariant derivative will transform by

$$[D_\mu, D_\nu]\psi(x) \to V(x)[D_\mu, D_\nu]\psi(x).$$

Let us now compute this invariant matrix directly.

$$\begin{aligned} [D_\mu, D_\nu]^a &= \left(\partial_\mu - igA_\mu^a t_a\right)\left(\partial_\nu - igA_\nu^b t_b\right) - \left(\partial_\nu - igA_\nu^b t_b\right)\left(\partial_\mu - igA_\mu^a t_a\right), \\ &= [\partial_\mu, \partial_\nu] - ig\left(\partial_\mu A_\nu^a t_a + A_\nu^a t_a\partial_\mu\right) + ig\left(\partial_\nu A_\mu^a t_a + A_\mu^a t_a\partial_\nu\right) + (-ig)^2\left(A_\mu^a t_a A_\nu^b t_b - A_\nu^b t_b A_\mu^a t_a\right), \\ &= -ig\left(\partial_\mu A_\nu^a t_a\right) + ig\left(\partial_\nu A_\mu^a t_a\right) + (-ig)^2\left[A_\mu^a t_a, A_\nu^b t_b\right], \\ &= -ig\left(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a - ig\left[t_b, t_c\right]A_\mu^b A_\nu^c\right)t_a, \\ &= -ig\left(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf_{abc}A_\mu^b A_\nu^c\right)t_a, \\ &\therefore [D_\mu, D_\nu]^a \equiv -igF_{\mu\nu}^a t_a, \quad \text{with} \quad F_{\mu\nu}^a \equiv \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf_{abc}A_\mu^b A_\nu^c. \end{aligned}$$

The transformation law for the commutator of the covariant derivative implies that

$$F_{\mu\nu}^a t_a \to V(x)F_{\mu\nu}^a t_a V^\dagger(x).$$

Therefore, the infinitesimal transformation is simply

$$F^a_{\mu\nu} t_a \to F^a_{\mu\nu} t_a + i \left[ \alpha^a t_a, F^b_{\mu\nu} t_b \right],$$
$$= F^a_{\mu\nu} t_a + f_{abc} \alpha^b F^c_{\mu\nu}.$$

It is important to note that the field strength tensors $F^a_{\mu\nu}$ are themselves not gauge invariant quantities because they transform nontrivially. However, we may easily form gauge-invariant quantities by combining all three of the field strength tensors. For example,

$$\mathscr{L} = -\frac{1}{2} \mathrm{Tr} \left[ \left( F^a_{\mu\nu} t_a \right)^2 \right] = -\frac{1}{4} \left( F^a_{\mu\nu} \right)^2$$

is a gauge-invariant kinetic energy term for $A^a_\mu$.

We now know how to make a large number of gauge invariant Lagrangians. In particular, to make an 'old' Lagrangian locally gauge invariant, we may simply promote the derivative $\partial_\mu$ to the covariant derivative $D_\mu$ and add a kinetic term for the $A^a_\mu$'s. For example, we may promote our standard Lagrangian of quantum electrodynamics to one which is locally invariant under a $SU(n)$ gauge symmetry by simply making the transformation,

$$\mathscr{L}_{\mathrm{QED}} = -\frac{1}{4} \left( F_{\mu\nu} \right)^2 + \overline{\psi} \left( i \not{D} - m \right) \psi$$
$$\to \mathscr{L}_{YM} = -\frac{1}{4} \left( F^a_{\mu\nu} \right)^2 + \overline{\psi} \left( i \not{D} - m \right) \psi.$$

**Basic Facts About Lie Algebras**

A Lie group is a group with a smooth finite-dimensional manifold structure. We usually distinguish between *real* and *complex* Lie groups, which have the structure of a real (resp. complex) manifold.

Associated to any Lie group $G$ is its Lie algebra $\mathcal{G}$, defined to be the tangent space of $G$ at the identity. If we think of $G$ as a group of linear transformations on a vector space, $\mathcal{G}$ is the space of infinitesimal transformations associated to $G$. On the other hand, if we parametrize elements in a neighborhood of the identity of $G$ in terms of arcs passing through the identity, $\mathcal{G}$ consists of the derivatives of these so-called 1-parameter subgroups at the identity. In symbols,

$$\mathcal{G} = \left\{ \frac{d}{dt} g(t) \bigg|_{t=0} ; \; g(t) \text{ is a smooth curve through the identity in } G \right\}$$

Conversely, we have a smooth map from $\mathcal{G}$ to $G$ given by the *exponential map.* If $G$ and $\mathcal{G}$ are written in terms of $n \times n$ matrices, this map is concretely defined by the power series

$$\exp(M) = 1 + M + \frac{M^2}{2!} + \cdots .$$

On a purely algebraic level, we can also speak of Lie algebras just as finite dimensional vector spaces (over $\mathbb{R}$ or $\mathbb{C}$) equipped with a commutator $[,]$. This bracket is bilinear, anti-symmetric, and satisfies the Jacobi identity:

$$[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0, \qquad \forall A, B, C \in \mathcal{G}.$$

If we pick a basis $\{T^a\}$ for $\mathcal{G}$ (often called a set of generators of the Lie group $G$), closure of $\mathcal{G}$ under the Lie bracket gives rise to a set of constants $\{f^{abc}\}$ defined by

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

The $f^{abc}$'s are called *structure constants.* We can choose a basis for $\mathcal{G}$ such that these constants are totally antisymmetric in the indices $a, b, c$. This is the same basis that diagonalizes the symmetric bilinear form

$$(t^a, t^b) \quad \mapsto \quad \mathrm{Tr} \left[ t^a t^b \right].$$

**Classification of Compact, Simple Lie Groups**

In gauge theory we are primarily interested in Lie groups that admit finite dimensional unitary representations. This requires that we consider compact groups. The classification scheme of such groups is aided by looking only at so-called *simple* Lie groups, ones which have no non-trivial normal Lie subgroups. In particular, such groups have no Lie subgroups isomorphic to $U(1)$ contained in its center (the subgroup of elements commuting with all of $G$). These simple compact Lie groups run in three infinite families:

(1) $SU(n) = \{M \in GL(n,\mathbb{C}) | MM^\dagger = I\}$
(2) $SO(n) = \{M \in GL(n,\mathbb{R}) | MM^\top = I\}$
(3) $Sp(2n) = \{M \in GL(2n,\mathbb{R}) | MJM^\top = I\}$

where

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

The Lie algebras associated to these groups are:

(1) $su_n = \{M \in gl(n,\mathbb{C}) | M + M^\dagger = 0\}$
(2) $so_n = \{M \in gl(n,\mathbb{R}) | M + M^\top = 0\}$
(3) $sp_{2n} = \{M \in gl(2n,\mathbb{R}) | MJ + JM^\top = 0\}$

where $gl(n,\mathbb{C})$ is the space of all $n \times n$ $\mathbb{C}$-valued matrices. Analogously for $gl(n,\mathbb{R})$.

The dimensions of the above algebras (as real vector spaces) are $n^2-1, n(n-1)/2$, and $2n(2n+1)/2$, respectively.

There are also five exceptional Lie algebras, denoted by $F_4, G_2, E_6, E_7, E_8$.

**Representation Theory**

The possible forms of a Lagrangian invariant under the action of a Lie group $G$ is determined by the representations of $G$. The representations of simple Lie groups and their corresponding Lie algebras have the nice property of being completely decomposable. If such a group acts on a finite dimensional vector space $V$, we can decompose $V$ as a direct sum of subspaces $V_i$, where each $V_i$ is preserved under the action of $G$ and admits no further invariant subspace. Hence to understand the representation theory of $G$ it suffices to look at its action on spaces that contain no proper non-zero invariant subspace. Such representations are called *irreducible*.

Examples of irreducible representations:

(1) Standard representation of $su_n$: Just $su_n$ acting by matrix multiplication on $\mathbb{C}^n$.
(2) Spin representations of $su_2$: For $j \in \frac{1}{2}\mathbb{Z}$, take a $2j+1$-dimensional space with basis $\langle -j|, \langle -j+1|, \ldots, \langle j|$, along with raising and lowering operators $J_\pm$, and $J_3 = [J_+, J_-]$. Over $\mathbb{C}$, $su_2 \cong \langle J_\pm, J_3\rangle$, and each basis element $\langle k|$ is an eigenvector of $J_3$ with eigenvalue $k$. These give all the irreducible representations of $su_2$.

For any Lie algebra $\mathcal{G}$ we also have the *adjoint* representation, where $\mathcal{G}$ acts on itself via the commutator $[,]$. That is,

$$t : u \quad \mapsto \quad [t,u] = t_G(u), \forall t,u \in \mathcal{G}.$$

In this case, the matrix representing the linear transformation $t_G$ is given by the structure constants.

Also, if we go back to the definition of the general covariant derivative $D_\mu$, we see that $D_\mu$ acts on a field by

$$D_\mu \vec{\phi} \quad = \quad \partial_\mu \vec{\phi} - igA_\mu^b t_G^b \vec{\phi}.$$

Likewise, the vector $\vec{A_\mu}$ transforms infinitesimally as

$$A_\mu^a \quad \to \quad A_\mu^a + \frac{1}{g}(D_\mu \alpha)^a.$$

**Some Algebraic Results**

For any irreducible representation $r$ of $\mathcal{G}$ of a vector space $V$, if we pick any basis $\{T^a\}$ for $\mathcal{G}$, we can construct the operator

$$T^2 = \sum_a (T^a)^2.$$

This element commutes with the action of $\mathcal{G}$ on $V$. To see this, we pick a basis for $\mathcal{G}$ that diagonalizes the bilinear form given above. Then expanding $[T^2, T^b]$ in terms of structure constants, using the fact that they are totally antisymmetric, gives the result. Since the representation $r$ is irreducible, $T^2$ is a constant times the identity matrix:

$$\sum_a (T^a)^2 \quad = \quad C_2(r) I_d$$

where $d = d(r)$ is the dimension of the representation. This operator $T^2$ is called the *quadratic Casimir operator* associated to the representation $r$.

Our basis of $\mathcal{G}$ is chosen so that $Tr((T^a)^2) = C(r), \forall a$, where $C(r)$ is another constant. Taking the trace then implies

$$d(r)C_2(r) = d(G)C(r).$$

The constants $C_2(r), C(r)$ will depend on the choice of basis for $\mathcal{G}$, but they are always related by the equation above.

**A Word About Tensor Products and Direct Sums of Representations**

If we have two irreducible representations of $\mathcal{G}$ on vector spaces $V_1$ and $V_2$, we get an action of $\mathcal{G}$ on the direct sum $V_1 \oplus V_2$ in a natural way:

$$t \cdot (v \oplus w) = (t \cdot v) \oplus (t \cdot w).$$

In concrete terms, the matrix representation of $t$ is given by

$$\begin{pmatrix} t_1 & 0 \\ 0 & t_2 \end{pmatrix},$$

where $t_1$ is the matrix representing the action of $t$ on $V_1$, and analogously for $t_2$.

Likewise, we get a representation of $\mathcal{G}$ on the tensor product $V_1 \otimes V_2$ by

$$t \cdot (v \times w) \quad = \quad (t \cdot v) \otimes (1 \cdot w) + (1 \cdot v) \otimes (t \cdot w).$$

Note that this is *not* the same as the action given by the direct sum.

## Solutions to Suggested Problems

### 15.1 Brute-Force Computations in $SU(3)$.

The standard basis for the fundamental representation of $SU(3)$ is given by

$$t_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad t_2 = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad t_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad t_4 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$t_5 = \frac{1}{2}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad t_6 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad t_7 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad t_8 = \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

**a)** Why are there exactly eight matrices in this basis?

When we represent this group in $\mathrm{Mat}_3(\mathbb{C})$, the basis will be given by a set of $3 \times 3$ matrices over the field $\mathbb{C}(\cong \mathbb{R} \times \mathbb{R})$ and so there can be at most $2 \cdot 3^2$ matrices corresponding to the real and imaginary parts of each entry.

Because the elements of the $SU(3)$ group satisfy $U^\dagger U = 1_{3\times 3}$ (by the definition of unitarity). This gives us 3 conditions corresponding to the diagonal entries of $1_{3\times3}$ and $2 \cdot \frac{3^2-3}{2}$ conditions for the independent off-diagonal entries. There is also one additional condition which demands that the determinant is 1.

Therefore, the complete set of $3 \times 3$ unitary matrices over $\mathbb{C}$ are spanned by no more than $2 \cdot 3^2 - 3 - 3^2 + 3 - 1 = 3^2 - 1 = 8$ basis matrices. Because we have found 8 above, there must be exactly eight.

To generalize this argument to any $SU(N)$ group, simply replace every instance of '3' with '$N$'.

**c)** Let us check that this representation satisfies the orthogonality condition,

$$\mathrm{Tr}\left[t^a t^b\right] = C(r)\delta^{ab}.$$

We note by direct calculation that $\mathrm{Tr}\left[t^a t^b\right] = 0 \ \forall a \neq b$. Therefore, we must only verify the eight matrices themselves. Again by direct calculation, we see that

$$\mathrm{Tr}\left[t_1 \cdot t_1\right] = \mathrm{Tr}\left[t_2 \cdot t_2\right] = \mathrm{Tr}\left[t_3 \cdot t_3\right] = \mathrm{Tr}\left[t_4 \cdot t_4\right] = \mathrm{Tr}\left[t_5 \cdot t_5\right] = \mathrm{Tr}\left[t_6 \cdot t_6\right] = \mathrm{Tr}\left[t_7 \cdot t_7\right] = \mathrm{Tr}\left[t_8 \cdot t_8\right] = \frac{1}{2}.$$

Therefore, $C(r) = \frac{1}{2}$.

**d)** We are to compute the value of the quadratic Casimir operator $C_2(r)$ directly from its definition and verify its relation to $C(r)$.

By direct calculation, we find that

$$t^a t^a = \frac{4}{3} 1_{3\times3},$$

and therefore $C_2(r) = \frac{4}{3}$. This agrees with the expected value $C_2(3) = \frac{3^2-1}{2\cdot 3} = \frac{4}{3}$.

### 15.2 Adjoint Representation of $SU(2)$.

We are to write down the basis matrices of the adjoint representation of $SU(2)$ and verify its canonical properties. From the definition of the adjoint representation, we see that

$$t_1 = i\epsilon_{i1j} = i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad t_2 = i\epsilon_{i2j} = i\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad t_3 = i\epsilon_{i3j} = i\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

It is almost obvious that $\mathrm{Tr}\left[t_i t_j\right] = 0 \ \forall i \neq j$. It is interesting to note that the product of any two distinct matrices will yield a matrix with single nonzero entry of $-1$ in an off-diagonal position; hence, the trace of a product of distinct matrices will vanish.

It is easy to verify that $\mathrm{Tr}\left[t_1^2\right] = \mathrm{Tr}\left[t_2^2\right] = \mathrm{Tr}\left[t_3^2\right] = 2$, and therefore $C(G) = 2$.
Likewise, notice that

$$t_1^2 + t_2^2 + t_3^2 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = 2\delta_{ij},$$

and therefore $C_2(G) = 2$.

**15.5 Casimir Operator Computations**.

This problem develops an alternative way of deriving the Casimir operators and other constants associated with an algebra and its representation.

**a)** Let us consider a group $\mathcal{G}$ which contains a subalgebra isomorphic to $su_2$. In general, any irreducible representation $r$ of $\mathcal{G}$ may be decomposed into a sum of representations of $su_2$:

$$r \to \sum j_i,$$

where $j_i$ are the spins of $SU(2)$ representations.

As we have done before, let us denote the generators of $su_2$ by $J_\pm, J_3$, with $J_3$ the diagonal operator. Because we are given an irreducible representation of $\mathcal{G}$ on a vector space $V$, when we restrict the action to $su_2$, $V$ decomposes into a sum of irreducible subspaces $\bigoplus_i V_{j_i}$, where $V_{j_i}$ is the spin $j_i$ representation space of $su_2$.

Therefore, by the definition of the of $C(r)$, we see that

$$\mathrm{Tr}\left[J_-^2 + J_+^2 + J_3^2\right] = 3C(r).$$

However, using the decomposition stated above, we see that

$$3C(r) = \mathrm{Tr}\left[J_-^2 + J_+^2 + J_3^2\right] = \mathrm{Tr}\left[\sum_i (J_-)_{j_i}^2 + (J_+)_{j_i}^2 + (J_3)_{ji}^2\right]$$

where, for example, $(J_-)_{j_i}$ is the restriction of $J_-$ to $V_{j_i}$. Then using the identity, $d(r)C_2(r) = d(G)C(r)$, we see that

$$\begin{aligned} 3C(r) &= \mathrm{Tr}\left[\sum_i (J_-)_{j_i}^2 + (J_+)_{j_i}^2 + (J_3)_{ji}^2\right], \\ &= 3C(j_i). \end{aligned}$$

Now, noting the definition of $C(r)$ and the fact that $(J_3)_{j_i}\langle k| = k\langle k|$, we get

$$\begin{aligned} C(j_i) = \mathrm{Tr}\left[(J_3)_{j_i}^2\right] &= 2\sum_k k^2, \\ &= 2\frac{j_i(j_i+1)(2j_i+1)}{6}, \\ &= \frac{1}{3} j_i(j_i+1)(2j_i+1). \end{aligned}$$

Note that this formula holds for any set of integer or half-integer values of $k$.

Combining these results, we see that

$$\therefore 3C(r) = \sum_i j_i(j_i+1)(2j_i+1).$$

*ὅπερ ἔδει δεῖξαι*

**b)** Under an $su_2$ subgroup of $su_n$, the fundamental representation $N$ transforms as a 2-component spinor ($j = \frac{1}{2}$) and $(N-2)$ singlets. We are to use this to show that $C(N) = \frac{1}{2}$ and that the adjoint representation of $su_n$ decomposes into one spin 1, $2(N-2)$ spine-$\frac{1}{2}$, and singlets to show that $C(G) = N$.

Because we have that $su_n$ decomposes into a 2-component spinor and singlets, we may use our result from part (a) to see that

$$\begin{aligned} 3C(N) &= \tfrac{1}{2}(\tfrac{1}{2}+1)(2\cdot\tfrac{1}{2}+1) + 0 = 3/2, \\ \therefore C(N) &= \frac{1}{2}. \end{aligned}$$

Let us now prove the properties of the adjoint representation. Let $\mathcal{G} = su_n$ under the restriction $\mathcal{H} = su_2$, we will embed $\mathcal{H}$ into $\mathcal{G}$ by the map

$$t \mapsto \begin{pmatrix} t & 0 \\ 0 & 0 \end{pmatrix}$$

where $t$ is a $2\times 2$ block of the matrix. The adjoint action of $\mathcal{H}$ on itself gives rise to a spin 1 representation of dimension 3, leaving a $n^2-4$ dimensional complementary invariant subspace. Those elements of $\mathcal{G}$ whose entries are in the lower right $(n-2)\times(n-2)$ block commute with $\mathcal{H}$ and hence give rise to 1 dimensional subspaces (singlets).

If we let $E_{i,j}$ denote the matrix with 1 in the $(i,j)^{\text{th}}$ position and zeros everywhere else, then the subspaces $V_j^+ = \langle E_{1,j}, E_{2,j}\rangle$ and $V_j^- = \langle E_{j,1}, E_{j,2}\rangle$ for $j = 3, \ldots, n$ are invariant under $su_2$. There are $2(n-2)$ such subspaces and so this completes the decomposition.

**c)** Symmetric and antisymmetric 2-index tensors form irreducible representations of $su_n$. We are to compute $C_2(r)$ for each of these representations. The direct sum of these representations is the product representation $N \times N$. We are to verify that our results $C_2(r)$ agree with the work in the text.

Let $\mathcal{A}$ denote the antisymmetric representation, $\mathcal{S}$ the symmetric representation of $su_n$, and let $V$ be the fundamental representation of $su_n$ with the standard basis $\{e_1, \ldots, e_n\}$. We can compute $C(\mathcal{A})$ and $C(\mathcal{S})$ by determining the action of the diagonal operator $J_3$ on $Sym_2(V)$ and $V \wedge V$.

The standard basis for $Sym^2(V)$ is given by $\{e_i \otimes e_j\}, 1 \leq i \leq j \leq n$ with dimension $n(n+1)/2$. Then $J_3$ acts on the basis by

$$\begin{aligned}
&e_1 \otimes e_j \mapsto \tfrac{1}{2}(e_1 \otimes e_j) \qquad && j > 2,\\
&e_2 \otimes e_j \mapsto -\tfrac{1}{2}(e_2 \otimes e_j) && j > 2,\\
&e_1 \otimes e_2 \mapsto 0,\\
&e_1 \otimes e_1 \mapsto e_1 \otimes e_j,\\
&e_2 \otimes e_2 \mapsto -(e_2 \otimes e_j).
\end{aligned}$$

These give the diagonal entries for the matric representation of $J_3$ and all other entries are zero. Taking the sum of the squares of the diagonal entries gives

$$C(\mathcal{S}) = \tfrac{1}{2}(n-2) + 2 = \frac{n+2}{2}.$$

We can do the same with the antisymmetric representation, which has a basis of $\{e_i \wedge e_j\}, 1 \leq i \leq j \leq n$ with dimension $n(n-1)/2$. The action of $J_3$ is then given by

$$\begin{aligned}
&e_1 \wedge e_j \mapsto \tfrac{1}{2}(e_1 \wedge e_j) \qquad && j > 2,\\
&e_2 \wedge e_j \mapsto -\tfrac{1}{2}(e_2 \wedge e_j) && j > 2,\\
&e_1 \wedge e_2 \mapsto 0.
\end{aligned}$$

Comparing this with our previous result, we see that

$$C(\mathcal{A}) = \frac{n-2}{2}.$$

We can now compute the quadratic Casimir operators $C_2(\mathcal{A})$ and $C_2(\mathcal{S})$ using the identity $d(r)C_2(r) = d(G)C(r)$, we get

$$\begin{aligned}
C_2(\mathcal{S}) &= \frac{(n^2-1)C(\mathcal{S})}{n(n+1)/2},\\
&= \frac{(n-1)(n+2)}{n}.\\
C_2(\mathcal{A}) &= \frac{(n^2-1)C(\mathcal{A})}{n(n-1)/2},\\
&= \frac{(n+1)(n-2)}{n}.
\end{aligned}$$

We may check this result with the work of Peskin and Schroeder by computing

$$\begin{aligned}
(C_2(\mathcal{A}) + C_2(\mathcal{S}))d(\mathcal{A})d(\mathcal{S}) &= C_2(\mathcal{A})d(\mathcal{A}) + C_2(\mathcal{S})d(\mathcal{S}),\\
&= (n^2-1)\left(\frac{n+2}{2} + \frac{n-2}{2}\right),\\
&= \left(n^2-1\right)n.
\end{aligned}$$

This result checks.