Chapter 4

Lie Groups and Lie Algebras

In this note we'll investigate two additional notions:

1. The addition of a continuity structure on the group;
2. The addition of an algebraic structure on the group.

The former is the subject of Lie groups, and the latter is the subject of Lie algebras. These are quite different concepts. However, we put them together here because in physics we are heavily concerned with the conjunction of the two ideas.[1]

4.1 Lie Groups

Formally, we have

Def: A *Lie group* is a group, G, whose elements form an analytic manifold such that the composition $ab = c$ $(a, b, c \in G)$ is an analytic mapping of $G \times G$ into G and the inverse $a \to a^{-1}$ is an analytic mapping of G into G.

That is, a Lie group is a group with a continuity structure: derivatives may be taken. Typically, we describe Lie groups by elements that are determined differentiably by some set of continuously varying real parameters. If there are r such parameters, we have an "r-parameter Lie group".

We won't here develop the theory of Lie groups from an abstract level. Instead, we'll directly mostly think in terms of representations by matrices, where the matrices are specified by some number of continuosly varying real parameters (up to possibly discrete points of discontinuity in some situations).

[1] The reader may wish to refer back to the note on Hilbert Spaces from Ph 129a for some concepts.

As with finite groups, it is convenient when we can deal with unitary representations. This is guaranteed to be possible in the following case:

Theorem: Every finite-dimension representation of a compact Lie group is equivalent to a unitary representation, and is either irreducible or fully reducible.

By "compact" here we mean that the parameters that specify an element of the Lie group vary over a compact set (i.e., over a closed set of finite extent). The proof of this parallels the proof given for finite groups that we gave in the note on representation theory, but now using the notion of an invariant integration over the group. Compactness ensures that this integral will be finite.

The notions of compactness and invariance of the group integral are topological concepts. There is a further topological property we will sometimes assume, that the group is "connected". By this, we mean that we can get to any element of the group from the identity via a sequence of small steps.

For some examples:

- The group $O^+(3)$ (representing proper rotations in three dimensions) is a compact, connected, 3-parameter Lie group.
- The group $O(3)$ (proper and improper rotations in three dimensions) is a compact, but not a connected group. It contains two disjoint categories of elements, those with determinant $+1$, and those with determinant -1, and it is not possible to continuously go from one to the other. This may be regarded as the direct product group:
$$O(3) = O^+(3) \otimes \mathcal{I}, \tag{4.1}$$
where $\mathcal{I}$ is the inversion group.
- The Lorentz group (of proper homogeneous Lorentz transformations) is connected, but not compact. This is a little more subtle – the lack of compactness is due to the fact that there is a limit point of a sequence of group elements that is not an element (consider a sequence of velocity boosts in which $v \to 1$).
- The improper, homogeneous Lorentz group is neither connected nor compact.

We will sometimes also restrict discussion to simple compact Lie groups, recalling that a simple group is one that contains no proper invariant subgroup.

If we have a compact Lie group, then we can define the invariant integral over the group and also work with unitary representations without loss of generality. The general orthogonality relation of finite groups may be generalized to include compact Lie groups. For unitary irreducible representations $D^{(i)}$ and $D^{(j)}$ we have:
$$\int_G D^{(i)}(g)_{\mu\nu} D^{(j)*}(g)_{\alpha\beta}\mu(dg) = \frac{1}{\ell_i}\delta_{ij}\delta_{\mu\alpha}\delta_{\nu\beta}. \tag{4.2}$$

We have assumed that the invariant integral over the group is normalized to one:

$$\int_G \mu(dg) = 1. \tag{4.3}$$

Let's consider an example. In the note on representation theory, we defined the spherical harmonic functions in terms of irreducible representations of the rotation group:

$$Y_{\ell m}(\theta,\phi) \equiv \sqrt{\frac{2\ell+1}{4\pi}} D^{\ell *}_{m0}(\phi,\theta,0). \tag{4.4}$$

Suppose we wish to know the orthogonality properties of the $Y_{\ell m}$'s. We compute:

$$\begin{aligned}
&\int_{(4\pi)} Y_{\ell m}(\theta,\phi) Y^*_{\ell' m'}(\theta,\phi) d\cos\theta d\phi = & (4.5)\\
&\quad \frac{\sqrt{(2\ell+1)(2\ell'+1)}}{4\pi} \int_{(4\pi)} D^{\ell *}_{m0}(\phi,\theta,0) D^{\ell'}_{m'0}(\phi,\theta,0) d\cos\theta d\phi \\
&\quad \frac{\sqrt{(2\ell+1)(2\ell'+1)}}{4\pi} \int_{(4\pi)} D^{\ell *}_{m0}(\phi,\theta,\alpha) D^{\ell'}_{m'0}(\phi,\theta,\alpha) d\cos\theta d\phi \\
&\quad \frac{\sqrt{(2\ell+1)(2\ell'+1)}}{8\pi^2} \int_{(8\pi^2)} D^{\ell *}_{m0}(\phi,\theta,\alpha) D^{\ell'}_{m'0}(\phi,\theta,\alpha) d\cos\theta d\phi d\alpha & (4.6)
\end{aligned}$$

We have used here the invariance of the integral when adding the rotation by angle α about the x-axis, and averaging over this rotation. The result is now in the form of the general orthogonality relation:

$$\frac{1}{8\pi^2} \int_{(8\pi^2)} D^{\ell *}_{mn}(\phi,\theta,\alpha) D^{\ell'}_{m'n'}(\phi,\theta,\alpha) d\cos\theta d\phi d\alpha = \frac{1}{2\ell+1} \delta_{\ell\ell'}\delta_{mm'}\delta_{nn'}. \tag{4.7}$$

Therefore,

$$\int_{(4\pi)} Y_{\ell m}(\theta,\phi) Y^*_{\ell' m'}(\theta,\phi) d\cos\theta d\phi = \delta_{\ell\ell'}\delta_{mm'}. \tag{4.8}$$

A perhaps less-familiar but very important example may be found in classical mechanics: Consider a system with generalized coordinates $q_i, i = 1, 2, \ldots, n$ and corresponding generalized momenta $p_i = \partial_{q_i} L$, where L is the Lagrangian. Hamilton's equations are:

$$\begin{aligned}
\dot{p}_i &= -\partial_{q_i} H, & (4.9)\\
\dot{q}_i &= \partial_{p_i} H, & (4.10)
\end{aligned}$$

where H is the Hamiltonian. We may rewrite this in terms of the $2n$-dimensional vector:

$$x \equiv \begin{pmatrix} q_1 \\ \vdots \\ q_n \\ p_1 \\ \vdots \\ p_n \end{pmatrix}, \tag{4.11}$$

as:

$$\dot{x} = J\frac{\partial H}{\partial x}, \tag{4.12}$$

with

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}. \tag{4.13}$$

That is, J is a $2n \times 2n$ matrix written in terms of $n \times n$ submatrices 0 and I.

A canonical transformation is a transformation from x to y where

$$y = \begin{pmatrix} Q_1 \\ \vdots \\ Q_n \\ P_1 \\ \vdots \\ P_n \end{pmatrix}, \tag{4.14}$$

such that

$$\dot{y} = J\frac{\partial H\left[x(y)\right]}{\partial y}. \tag{4.15}$$

That is, Hamilton's equations are preserved under a canonical transformation.

We have

$$\dot{y}_i = \sum_j \frac{\partial y_i}{\partial x_j}\dot{x}_j, \tag{4.16}$$

which may be written in matrix form:

$$\dot{y} = M\dot{x}, \tag{4.17}$$

where

$$M_{ij} \equiv \frac{\partial y_i}{\partial x_j}. \tag{4.18}$$

Hence,

$$\dot{y} = MJ\frac{\partial H}{\partial x}. \tag{4.19}$$

Now

$$\frac{\partial H}{\partial x_i} = \sum_j \frac{\partial H}{\partial y_j}\frac{\partial y_j}{\partial x_i} = \sum_j \frac{\partial H}{\partial y_j}M_{ji}, \tag{4.20}$$

or,

$$\frac{\partial H}{\partial x} = M^T \frac{\partial H}{\partial y}. \tag{4.21}$$

We conclude that

$$\dot{y} = MJM^T \frac{\partial H}{\partial y}, \tag{4.22}$$

and that the transformation is canonical if

$$MJM^T = J. \tag{4.23}$$

A matrix M which satisfies the condition of Eqn. 4.23 is said to be *symplectic*. The reader is encouraged to verify that the set of $2n \times 2n$ symplectic matrices forms a group, called the *symplectic group*, denoted $Sp(2n)$.

We remark that the evolution of the system in time corresponds to a sequence of canonical transformations, and hence the time evolution corresponds to the application of successive symplectic matrices. This finds practical application in various situations, for example in accelerator physics.

We turn now to another feature of unitary representations. Let U be a unitary matrix. Write

$$U = e^{iA} \equiv \sum_{n=0}^{\infty} \frac{(iA)^n}{n!}, \tag{4.24}$$

where we leave it to the reader to investigate convergence. Now,

$$\begin{aligned}
U^{-1} &= U^\dagger = \left(e^{iA}\right)^\dagger \\
&= \left[\sum_{n=0}^{\infty} \frac{(iA)^n}{n!}\right]^\dagger \\
&= \sum_{n=0}^{\infty} \frac{\left[(-iA^*)^n\right]^T}{n!} \\
&= \sum_{n=0}^{\infty} \frac{\left[(-iA^\dagger)^n\right]}{n!} \\
&= e^{-iA^\dagger}.
\end{aligned} \tag{4.25}$$

But we also know that,

$$U^{-1} = e^{-iA}, \tag{4.26}$$

since A commutes with itself, and hence exponentials of multiples of A may be treated like ordinary numbers in products. Therefore, we may take $A = A^\dagger$. That is, A is a hermitian matrix.

Note that if we also have $\det U = 1$, then A can be taken to be traceless: The matrix A is hermitian, hence diagonalizable by a unitary transformation. Let

$$\Delta = SAS^{-1} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n), \tag{4.27}$$

be a diagonal equivalent of A, where S is unitary. Then, $A = S^{-1}\Delta S$, or

$$\begin{aligned}
1 &= \det\left(e^{iS^{-1}\Delta S}\right) \\
&= \det\sum_{k=0}^{\infty}\frac{1}{k!}\left(iS^{-1}\Delta S\right)^k \\
&= \det S^{-1}\left[\sum_{k=0}^{\infty}\frac{1}{k!}\left(i\Delta\right)^k\right]S \\
&= \det(S^{-1})\det(S)\det e^{i\Delta} \\
&= \exp\left(i\sum_{j=1}^{n}\lambda_j\right). \qquad (4.28)
\end{aligned}$$

Thus, the sum of the eigenvalues is equal to $2\pi m$, where m is an integer. Notice that if $m \neq 0$, we can define a new diagonal matrix $\Delta' = \Delta - 2\pi m\delta_{11}$, where δ_{11} is the matrix with the $i, j = 1, 1$ element equal to one, and all other elements zero. The trace of Δ' is zero. Hence $A' \equiv S^{-1}\Delta' S$ is also traceless. But $\exp(i\Delta') = \exp(i\Delta)$, and therefore

$$U = S^{-1}e^{i\Delta'}S = \exp(iS^{-1}\Delta' S) = e^{iA'}, \qquad (4.29)$$

where A' is hermitian and traceless.

Suppose D is a unitary representation of a group G. Then the elements of the group representation may be written in the form:

$$D(g) = \exp\left[i\epsilon^{\alpha}(g)X_{\alpha}\right], \qquad (4.30)$$

where the summation convention on repeated indices is used, $\{X_\alpha\}$ is a set of constant hermitian matrices, and $\{\epsilon_\alpha\}$ is a set of real parameters.

We are in particular concerned here with Lie groups (with unitary representations assumed here). In that case, if G is an r-parameter Lie group, we can find a set of r matrices $X_\alpha, \alpha = 1, 2, \ldots, r$ such that Eqn. 4.30 holds. We refer to these matrices as the *infinitesimal generators* of the group. In this case, we have the "fundamental theorem of Lie":

Theorem: The local structure of a Lie group is completely specified by the commutation relations among the generators X_α:

$$[X_\alpha, X_\beta] = C^{\gamma}_{\alpha\beta}X_\gamma, \quad \alpha, \beta = 1, 2, \ldots, r, \qquad (4.31)$$

where the coefficients $C^{\gamma}_{\alpha\beta}$ (called the *structure constants* of the group) are independent of the representation.

We investigate the proof of this, or rather of the Baker-Campbell-Hausdorff theorem, in exercise 6.

The reader is encouraged to check that the structure constants must satisfy:

$$C^{\gamma}_{\alpha\beta} = -C^{\gamma}_{\beta\alpha}, \qquad (4.32)$$

and (with summation convention over repeated indices)

$$C^{\delta}_{\alpha\beta}C^{\epsilon}_{\delta\gamma} + C^{\delta}_{\gamma\alpha}C^{\epsilon}_{\delta\beta} + C^{\delta}_{\beta\gamma}C^{\epsilon}_{\delta\alpha} = 0. \tag{4.33}$$

The matrices X_α may be regarded as operators on a vector space. If we are doing quantum mechanics, and we have a hermitian set of operators, they correspond to observables.

The commutator may be regarded as defining a kind of product, and the matrices $\{X_\alpha\}$ as generating a vector space, which is closed under this product. This brings us to the subject of Lie algrebras, in the next section.

4.2 Lie Algrebras

In the discussion of infinite groups of relevance to physics (in particular, Lie groups), it is useful to work in the context of a richer structure called an algebra. For background, we start by giving some mathematical definitions of the underlying structures:

Def: A **ring** is a triplet $\langle R, +, \circ\rangle$ consisting of a non-empty set of elements (R) with two binary operations ($+$ and $\circ$) such that:

1. $\langle R, +\rangle$ is an abelian group.
2. R is closed under $\circ$.
3. $(\circ)$ is associative.
4. Distributivity holds: for any $a, b, c \in R$

$$a \circ (b + c) = a \circ b + a \circ c \tag{4.34}$$

and

$$(b + c) \circ a = b \circ a + c \circ a \tag{4.35}$$

Conventions:
We use 0 ("zero") to denote the identity of $\langle R, +\rangle$. We speak of $(+)$ as addition and of $(\circ)$ as multiplication, typically omitting the $(\circ)$ symbol entirely (*i.e.*, $ab \equiv a \circ b$).

Def: A ring is called a **field** if the non-zero elements of R form an abelian group under $(\circ)$.

Def: An abelian group $\langle V, \oplus\rangle$ is called a **vector space** over a field $\langle F, +, \circ\rangle$ by a scalar multiplication $(*)$ if for all $a, b \in F$ and $v, w \in V$:

1. $a * (v \oplus w) = (a * v) \oplus (a * w)$ distributivity
2. $(a + b) * v = (a * v) \oplus (b * v)$ distributivity
3. $(a \circ b) * v = a * (b * v)$ associativity
4. $1 * v = v$ unit element $(1 \in F)$

Conventions:
We typically refer to elements of V as "vectors" and elements of F as "scalars." We typically use the symbol $+$ for addition both of vectors and scalars. We also generally omit the $*$ and $\circ$ multiplication symbols. Note that this definition is an abstraction of the definition of vector space given in the note on Hilbert spaces, page 6.

Def: An **algebra** is a vector space V over a field F on which a multiplication $(\times)$ between vectors has been defined (yielding a vector in V) such that for all $u, v, w \in V$ and $a \in F$:

1. $(au) \times v = a(u \times v) = u \times (av)$
2. $(u+v) \times w = (u \times w) + (v \times w)$ and $w \times (u+v) = (w \times u) + (w \times v)$

(Once again, we often omit the multiplication sign, and hope that it is clear from context which quantities are scalars and which are vectors.)

We are sometimes interested in the following types of algebras:

Def: An algebra is called **associative** if the multiplication of vectors is associative.

We may construct the idea of a "group algebra": Let G be a group, and V be a vector space over a field F, of dimension equal to the order of G (possibly ∞). Denote a basis for V by the group elements. We can now define the multiplication of two vectors in V by using the group multiplication table as "structure constants": Thus, if the elements of G are denoted by g_i, a vector $u \in V$ may be written:

$$u = \sum a_i g_i$$

We require that, at most, a finite number of coefficients a_i are non-zero. The multiplication of two vectors is then given by:

$$\left(\sum a_i g_i\right)\left(\sum b_j g_j\right) = \sum \left(\sum_{g_i g_j = g_k} a_i b_j \right) g_k$$

[Since only a finite number of the $a_i b_j$ can be non-zero, the sum $\sum_{g_i g_j = g_k} a_i b_j$ presents no problem, and furthermore, we will have closure under multiplication.]

Since group multiplication is associative, our group algebra, as we have constructed it, is an associative algebra.

We note that an associative algebra is, in fact, a ring. Note also that the multiplication of vectors is not necessarily commutative. An important non-associative algebra is:

Def: A **Lie algebra** is an algebra in which the multiplication of vectors obeys the further properties (letting u, v, w be any vectors in V):

1. Anticommutivity: $u \times v = -v \times u$.
2. Jacobi Identity: $u \times (v \times w) + w \times (u \times v) + v \times (w \times u) = 0$.

We concentrate on Lie algebras henceforth in this note, in particular on Lie algebras associated with a Lie group. The generators, $\{X_\alpha\}$, of a Lie group generate a Lie algebra, where multiplication of vectors is defined as the commutator. Just as for groups, we have the notion of a *regular representation* (or also "adjoint representation") of the Lie algebra. We may rewrite the identity for the structure constants:

$$C^\delta_{\alpha\beta} C^\epsilon_{\delta\gamma} + C^\delta_{\gamma\alpha} C^\epsilon_{\delta\beta} + C^\delta_{\beta\gamma} C^\epsilon_{\delta\alpha} = 0. \tag{4.36}$$

in the suggestive form:

$$C^\delta_{\alpha\beta} \left(C_\delta\right)^\epsilon_\gamma + \left(-C_\beta\right)^\epsilon_\delta \left(-C_\alpha\right)^\delta_\gamma + \left(-C_\alpha\right)^\epsilon_\delta \left(C_\beta\right)^\delta_\gamma = 0. \tag{4.37}$$

Interpreting, e.g., C_α as a matrix with elements $(C_\alpha)^\epsilon_\delta$, where δ is the column index, we find:

$$[C_\alpha, C_\beta] = C^\delta_{\alpha\beta} C_\delta. \tag{4.38}$$

The matrices C_α formed from the structure constants have the same commutation relations as the generators X_α of the Lie group, and hence form a *representation of the Lie algebra*, called the regular or adjoint representation.

The problem of classifying Lie groups is essentially the problem of finding the numbers $\{C\}$ satisfying the requirements of Eqns. 4.32 and 4.33 above, and then finding the r constant matrices which satisfy the commutation relations. This problem was solved by Cartan in 1913. We list the simple Lie groups here:

The "classical Lie groups" are (except as noted, $\ell = 1, 2, \ldots$):

1. The group of unitary unimodular (i.e., determinant equal to one) $(\ell+1)\times(\ell+1)$ matrices, denoted A_ℓ or $SU(\ell+1)$. This is an $\ell(\ell+2)$-parameter Lie group, as the reader is encouraged to demonstrate.

2. The group of orthogonal unimodular $(2\ell+1)\times(2\ell+1)$ matrices, denoted B_ℓ or $SO(2\ell+1)$ or $O^+(2\ell+1)$. This is an $\ell(2\ell+1)$-parameter Lie group, as the reader is encouraged to demonstrate.

3. The group of orthogonal unimodular $(2\ell)\times(2\ell)$ matrices, for $\ell > 2$, denoted D_ℓ or $SO(2\ell)$ or $O^+(2\ell)$. This is an $\ell(2\ell-1)$-parameter Lie group, as the reader is encouraged to demonstrate. It may be noted that for $\ell \le 2$ the group is not simple.

4. The group of symplectic $(2\ell)\times(2\ell)$ matrices, denoted C_ℓ or $Sp(2\ell)$. This is an $\ell(2\ell+1)$-parameter Lie group, as the reader is encouraged to demonstrate.

In addition, there are five "exceptional groups": G_4 with 14 parameters, F_4 with 52 parameters, E_6 with 78 parameters, E_7 with 133 parameters, and E_8 with 248 parameters.

Consider briefly the example of the rotation group and associated Lie algebra in quantum mechanics.[2] In three dimensions, a rotation about the $\hat{\alpha}$ unit axis by angle ϕ can be expressed in the form:

$$R_{\hat{\alpha}}(\phi) = e^{-i\beta\cdot T}, \tag{4.39}$$

where $\beta \cdot T \equiv \beta_x T_x + \beta_y T_y + \beta_z T_z$, $\beta = \beta(\hat{\alpha}, \phi)$, and $T_{x,y,z}$ are the infinitesimal generators of rotations in three dimensions:

$$T_x \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, T_y \equiv \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, T_z \equiv \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.40}$$

We may consider the application of successive rotations (which must be a rotation):

$$\begin{aligned} e^{-i\alpha\cdot T} e^{-i\beta\cdot T} &= e^{-i\gamma\cdot T} \\ &= \sum_{m=0}^{\infty} \frac{(-i\alpha\cdot T)^m}{m!} \sum_{n=0}^{\infty} \frac{(-i\beta\cdot T)^n}{n!} \\ &= 1 - i(\alpha+\beta)\cdot T + \frac{(-i\alpha\cdot T)^2}{2!} + \frac{(-i\beta\cdot T)^2}{2!} + (-i\alpha\cdot T)(-i\beta\cdot T) + O\left[(\alpha,\beta)^3\right] \\ &= 1 - i(\alpha+\beta)\cdot T + \frac{\left[-i(\alpha+\beta)\cdot T\right]^2}{2!} - \frac{[\alpha\cdot T, \beta\cdot T]}{2!} + O\left[(\alpha,\beta)^3\right] \\ &= \exp\left\{-i(\alpha+\beta)\cdot T - \frac{[\alpha\cdot T, \beta\cdot T]}{2!} + O\left[(\alpha,\beta)^3\right]\right\}. \end{aligned} \tag{4.41}$$

Thus, to this order in the expansion, we need to have the values of commutators such as $[T_x, T_y]$, but not of products $T_x T_y$. This statement is true to all orders, as stated in the celebrated Campbell-Baker-Hausdorff theorem. Hence, every order is linear in the T's, and therefore γ exists. This is also why we can learn most of what we need to know about Lie groups by studying the commutation relations of the generators, as indicated in the general "fundamental theorem of Lie".

It may be remarked that for a general, abstract Lie algebra, we should not even think of the product $[A, B]$ as $AB - BA$, since the product AB may not be defined, while the "Lie product" denoted $[A, B]$ may be. Of course, if we have a matrix representation for the generators, then AB is defined. In physics we typically deal with matrix representations, so referring to the Lie product as a commutator is justified.

For our three-dimensional rotation generators, the Lie products are found by evaluating the commutation relations of the matrices, with the result:

$$[T_\alpha, T_\beta] = i\epsilon_{\alpha\beta\gamma} T_\gamma, \tag{4.42}$$

[2]Again, this example is considerably expanded upon in the note on the rotation group in quantum mechanics, linked to the Ph 129 web page.

where $\epsilon_{\alpha\beta\gamma}$ is the "antisymmetric tensor" (in three dimensions), or "Levi-Civita antisymmetric symbol", defined by:

$$\epsilon_{\alpha\beta\gamma} \equiv \begin{cases} +1 & \alpha,\beta,\gamma \text{ an even permutation of } 1,2,3, \\ -1 & \alpha,\beta,\gamma \text{ an odd permutation of } 1,2,3, \\ 0 & \text{otherwise.} \end{cases} \tag{4.43}$$

With these commutation relations, we may define an abstract Lie algrebra, with generators (basis vectors) t_1, t_2, t_3 satisfying the Lie products:

$$[t_\alpha, t_\beta] = i\epsilon_{\alpha\beta\gamma} t_\gamma, \tag{4.44}$$

We complete the Lie algebra by considering linear combinations of the t's, requiring:

$$[a \cdot t + b \cdot t, c \cdot t] = [a \cdot t, c \cdot t] + [b \cdot t, c \cdot t] \tag{4.45}$$

and

$$[a \cdot t, b \cdot t] = -[b \cdot t, a \cdot t]. \tag{4.46}$$

Our Lie algrebra satisfies the Jacobi indentity:

$$[a \cdot t, [b \cdot t, c \cdot t]] + [b \cdot t, [c \cdot t, a \cdot t]] + [c \cdot t, [a \cdot t, b \cdot t]] = 0. \tag{4.47}$$

The matrices T_x, T_y, T_z generate a *representation* of this Lie algebra with *dimension* three, since the matrices are 3×3 and hence operators on a 3-dimensional vector space. We note that the vector space of the Lie algebra itself is also three-dimensional, but this is not required, and the two vector spaces should not be confused.

Recalling quantum mechanics, we know that it is useful to define

$$t_+ \equiv t_1 + it_2 \tag{4.48}$$
$$t_- \equiv t_1 - it_2. \tag{4.49}$$

We may obtain the commutation relations

$$[t_3, t_+] = t_+ \tag{4.50}$$
$$[t_3, t_-] = -t_- \tag{4.51}$$
$$[t_+, t_-] = 2t_3. \tag{4.52}$$

We suppose that the t's are represented by linear transformations, J, acting on some vector space V, where V is of finite dimension, but not necessarily three dimensions. We make the correspondence $t_\pm \to J_\pm$, $t_3 \to J_3$. Since none of these generators commute, only one of $J_\pm, J_3$ can be diagonalized at a time. We have the definition:

Def: The number of generators of a Lie algebra that can simultaneously be "diagonalized" is called the *rank* of the Lie group.

Thus, the rotation group is of rank 1.

We pick J_3 to be in diagonal form with respect to some basis $\{v\}$. We label the basis vectors by the diagonal element (eigenvalue) k:

$$J_3 v_k = k v_k. \tag{4.53}$$

By repeated action of $J_\pm$ on v_k it may be demonstrated that k is either integer or $\frac{1}{2}$-integer, with some maximal value j, and the eigenvalues of J_3 are $-j, -j+1, \ldots, j$. This demonstration is commonly performed in quantum mechanics courses. There are $2j+1$ distinct eigenvalues, so the dimension of our representation is $\geq 2j+1$. If we define our space to be the space spanned by $\{v_k, k = -j, -j+1, \ldots, j\}$ then our space is said to be *irreducible* – there is no proper subspace of V which is mapped onto itself by the various J's.

As remarked earlier, for a compact Lie group we may find a unitary representation, and hence we may represent the generators of the associated Lie algebra by hermitian matrices. Assuming we have done so, we find

$$\begin{aligned} [X_\alpha, X_\beta]^\dagger &= (X_\alpha X_\beta - X_\beta X_\alpha)^\dagger \\ &= X_\beta^\dagger X_\alpha^\dagger - X_\alpha^\dagger X_\beta^\dagger \\ &= X_\beta X_\alpha - X_\alpha X_\beta \\ &= -[X_\alpha, X_\beta]. \end{aligned} \tag{4.54}$$

We thus have

$$\begin{aligned} C_{\alpha\beta}^{\delta *} X_\delta^\dagger &= C_{\alpha\beta}^{\delta *} X_\delta \\ &= [X_\alpha, X_\beta]^\dagger \\ &= -[X_\alpha, X_\beta] \\ &= -C_{\alpha\beta}^{\delta} X_\delta. \end{aligned} \tag{4.55}$$

That is, $C_{\alpha\beta}^{\delta *} = -C_{\alpha\beta}^{\delta}$, and the structure constants are thus pure imaginary for a unitary representation.

We may introduce the concept of an operator for "raising and lowering indices" or a "metric tensor", by defining:

$$g_{\mu\nu} = g_{\nu\mu} \equiv C_{\mu\alpha}^{\beta} C_{\nu\beta}^{\alpha}. \tag{4.56}$$

It may be shown that for a semi-simple Lie group $\det g \neq 0$, where g is the matrix formed by the elements $g_{\mu\nu}$. Thus, in this case, g has an inverse, which we define by:

$$g^{\mu\nu} g_{\nu\rho} = \delta_\rho^\mu, \tag{4.57}$$

where we have written the Kronecker function with one index raised.

The metric tensor may be used for raising or lowering indices, for example:

$$g^{\alpha\beta} g^{\mu\nu} g_{\nu\beta} = g^{\alpha\beta} \delta_\beta^\mu = g^{\alpha\mu}. \tag{4.58}$$

We have here "raised" the indices on $g_{\nu\beta}$. In general, given a quantity with lower indices, we may define a corresponding quantity with upper indices according to:

$$A^\alpha \equiv g^{\alpha\beta} A_\beta. \tag{4.59}$$

Or, given a quantity with raised indices, we may define a corresponding quantity with lower indices:

$$A_\alpha \equiv g_{\alpha\beta} A^\beta. \tag{4.60}$$

In particular, we may define structure constants with all lower indices:

$$C_{\alpha\beta\gamma} = C^\delta_{\alpha\beta} g_{\delta\gamma}. \tag{4.61}$$

The $C_{\alpha\beta\gamma}$ so defined is antisymmetric under interchange of any pair of indices. Note that, if $C^\delta_{\alpha\beta}$ is pure imaginary, then g is real, and $C_{\alpha\beta\gamma}$ is pure imaginary.

Now consider the quantity

$$F \equiv g_{\alpha\beta} X^\alpha X^\beta = X^\alpha X_\alpha = X_\alpha X^\alpha, \tag{4.62}$$

where the X_α are the infinitesimal generators of the Lie algebra. Consider the commutator of F with any generator:

$$\begin{aligned}
[F, X_\gamma] &= g^{\alpha\beta}[X_\alpha X_\beta, X_\gamma] \\
&= g^{\alpha\beta}\left\{X_\alpha[X_\beta, X_\gamma] + [X_\alpha, X_\gamma]X_\beta\right\} \\
&= g^{\alpha\beta}\left(C^\delta_{\beta\gamma} X_\alpha X_\delta + C^\delta_{\alpha\gamma} X_\delta X_\beta\right) \\
&= g^{\alpha\beta} C^\delta_{\beta\gamma} X_\alpha X_\delta + g^{\beta\alpha} C^\delta_{\beta\gamma} X_\delta X_\alpha \\
&= g^{\alpha\beta} C^\delta_{\beta\gamma} \left(X_\alpha X_\delta + X_\delta X_\alpha\right) \\
&= g^{\alpha\beta} g^{\delta\epsilon} C_{\beta\gamma\epsilon} \left(X_\alpha X_\delta + X_\delta X_\alpha\right) \\
&= C_{\beta\gamma\epsilon}\left(X^\beta X^\epsilon + X^\epsilon X^\beta\right) \\
&= C_{\epsilon\gamma\beta}\left(X^\epsilon X^\beta + X^\beta X^\epsilon\right) \\
&= -C_{\beta\gamma\epsilon}\left(X^\epsilon X^\beta + X^\beta X^\epsilon\right) \\
&= 0,
\end{aligned} \tag{4.63}$$

since it is equal to its negative. Thus, F commutes with every generator, hence commutes with every element of the algebra. By Schur's lemma, F must be a multiple of the identity, since if F commutes with every generator, then it must commute with every element of the group in some irreducible representation. An operator which commutes with every generator is known as a *Casimir operator*.

For example, consider again the rotation group in quantum mechanics. The structure constants are

$$C^\gamma_{\alpha\beta} = i\epsilon_{\alpha\beta\gamma}. \tag{4.64}$$

The metric tensor is thus

$$\begin{aligned}
g_{\mu\nu} &= C^\beta_{\mu\alpha} C^\alpha_{\nu\beta} \\
&= -\epsilon_{\mu\alpha\beta}\epsilon_{\nu\beta\alpha} \\
&= 2\delta_{\mu\nu}.
\end{aligned} \tag{4.65}$$

Hence, $J^\alpha = 2J_\alpha$, and $J^2 = J_1^2 + J_2^2 + J_3^2$ is a Casimir operator, a multiple of the identity. To determine the multiple, we consider the action of J^2 on a basis vector. This may be accomplished by writing it in the form $J^2 = J_z^2 + \frac{1}{2}(J_+J_- + J_-J_+)$, where $J_\pm \equiv J_x \pm iJ_y$. This exercise yields the familiar result

$$J^2 v_k = j(j+1)v_k, \tag{4.66}$$

where $2j+1$ is the dimension of the representation. Thus,

$$J^2 = j(j+1)I. \tag{4.67}$$

4.3 Example: $SU(3)$

The group $SU(3)$ consists of the set of unitary unimodular 3×3 matrices. In the exercises, you show that it is an eight parameter group. Thus, we know that the associated Lie algebra must have eight linearly independent generators. That is, we wish to find a set of eight linearly independent traceless hermitian 3×3 matrices. It is readily demonstrated that the vector space of such matrices is in fact eight dimensional, that is, our generators provide a basis for the vector space of traceless hermitian 3×3 matrices.

There are many ways we could pick our basis for the Lie algebra. However, it is generally wise to make as many as possible diagonal. In this case, there are three linearly-independent 3×3 diagonal hermitian matrices, but the traceless requirement reduces these to only two. The number of simultaneously diagonalizable generators is called the *rank* of the Lie algebra, hence $SU(3)$ is rank two.

A common choice for the generators, with two diagonal generators, is the "Gell-Mann matrices":

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{4.68}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \tag{4.69}$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{4.70}$$

Notice the $SU(2)$ substructure. For example, the upper left 2×2 submatrices of λ_1, λ_2, and λ_3 are just the Pauli matrices. The group $SU(3)$ contains subgroups isomorphic with $SU(2)$ (but not invariant subgroups).

One area where $SU(3)$ plays an important role is in the "Standard Model" – $SU(3)$ is the "gauge group" of the strong interaction (Quantum Chromodynamics). In this case, the group elements describe transformations in "color" space, where color is the analog of charge in the strong interaction. Instead of the single dimension of electromagnetic charge, color space is three-dimensional. The $SU(3)$ symmetry reflects the fact that all colors couple with the same strength – there is no preferred "direction" in color space. In field theory, once the gauge symmetry is specified, the form of the interaction is determined.

There is another example in particle physics where $SU(3)$ enters. Instead of the color symmetry just discussed, there is a "flavor" symmetry. The three lightest quarks are called "up" (u), "down" (d), and "strange" (s). The quantum number that distinguishes these is called flavor. The strong interaction couples with the same strength to each flavor. Thus, we may make "rotations" in this three-dimensional flavor space without changing the interaction. These rotations are described by the elements of $SU(3)$. The symmetry is actually broken, because the u, d, and s quarks have different masses (also, the electromagnetic and weak interaction couplings depend on flavor), but it is still a useful approximation in many situations. We'll develop this application somewhat further here.

We use the Gell-Mann representation, in which λ_3 and λ_8 are the diagonal generators. According to the assumption of $SU(3)$ flavor symmetry, our operators in flavor space commute with the Hamiltonian. We'll label our quark flavor basis according to the eigenvalues of λ_3 and λ_8. It is conventional to notice the $SU(2)$ substructure of $(\lambda_1, \lambda_2, \lambda_3)$ and refer to the two-dimensional operations of these generators as operations on "isospin" (short for isotopic spin) space. This is the ordinary nuclear isospin. It really doesn't have anything to do with angular momentum, but gets its "spin" nomenclature from the analogy with angular momentum where $SU(2)$ also enters. By analogy with angular momentum, a two-dimensional representation gets "third-component" quantum numbers of $\pm 1/2$. That is, we define, in this representation,

$$I_3 = \frac{1}{2}\lambda_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.71}$$

The eigenstates with $I_3 = +1/2, -1/2$ are called the u quark and the d quark, respectively. The strange quark in this convention has $I_3 = 0$, it is an $I = 0$ state (an isospin "singlet"). Note that this three-dimensional representation of $SU(2)$ is reductible to two-dimensional and one-dimensional irreps.

For the other quantum number, we define the "hypercharge" operator, in this representation:

$$Y = \frac{1}{\sqrt{3}}\lambda_8 = \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{4.72}$$

Thus, the u and d quarks both have $Y = 1/3$, and the s quark has $Y = -2/3$. The basis for this three-dimensional representation of flavor $SU(3)$ is illustrated in Fig. 4.1.

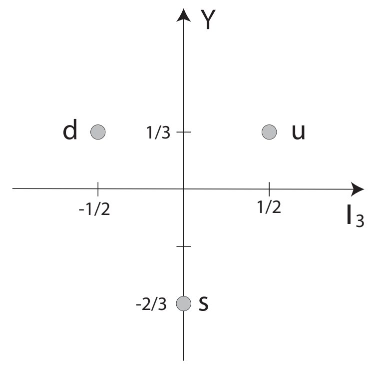

Figure 4.1: The 3 representation of $SU(3)$, in the context of quark flavors.

Now, we can also generate additional representations of $SU(3)$, and interpret in this physical context. Under complex conjugation of an element of $SU(3)$,

$$U = e^{i\epsilon^\alpha \lambda_\alpha} \to U^* = e^{-i\epsilon^\alpha \lambda_\alpha^*}. \tag{4.73}$$

This generates a new three-dimensional representation, called $\bar{3}$. The I_3 and Y quantum numbers switch signs. Thus, the diagram for $\bar{3}$ looks like the diagram for 3 reflected through the origin. We label the states $\bar{u}, \bar{d}, \bar{s}$, reflecting their interpretation as anti-quark states. Notice that the complex conjugate representation $\bar{3}$ is not equivalent to the 3 representation. This is a difference from $SU(2)$, where the two representations (2 and $\bar{2}$) are equivalent.

We may also generate higher dimension representations of $SU(3)$ by forming direct product representations. Some of these have special interpretation in particle physics: Combining a quark with an anti-quark, that is, forming the $3\otimes\bar{3}$ representation, gives meson states. Combining three quarks, $3\otimes 3\otimes 3$, gives baryons. As usual, these direct product representations may be expected to be reducible. For example, we have the reduction to irreducible representations: $3 \otimes \bar{3} = 8 \oplus 1$. We will discuss the graph in Fig. 4.2 in class.

4.4 Exercises

1. Show that $SU(n)$ requires $(n-1)(n+1)$ real parameters to specify an element.

2. Show that $C_{\alpha\beta\gamma}$ is antisymmetric under interchange of any pair of indices.

3. Show that the complex conjugate representation, $\bar{2}$, of $SU(2)$ is equivalent to the original 2 representation.

4. Consider the Helmholtz equation in two dimensions:

$$\nabla^2 f + f = 0, \tag{4.74}$$

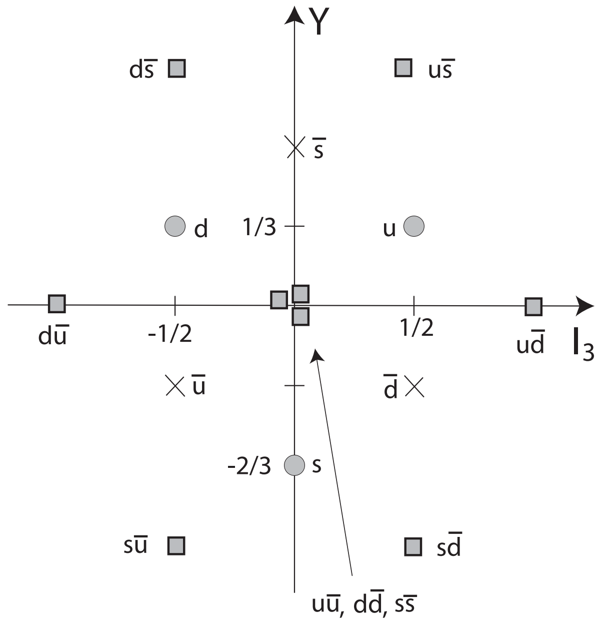

Figure 4.2: The $3 \otimes \bar{3} = 8 \oplus 1$ representation of $SU(3)$, in the context of quark flavors.

where

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \tag{4.75}$$

(a) Show that the equation is left invariant under the transformation:

$$\tau(\epsilon, \theta, \alpha, \beta) : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta + \alpha \\ x\epsilon\sin\theta + y\epsilon\cos\theta + \beta \end{pmatrix}, \tag{4.76}$$

where $\epsilon = \pm 1$, $-\pi \leq \theta < \pi$, and α and β are any real numbers (actually, α, β, and θ could be complex, but we'll restrict to real numbers here).

(b) The set of transformations $\{\tau(\epsilon, \theta, \alpha, \beta)\}$ obviously forms a Lie group, where group multiplication is defined as the application of successive transformations. Is it a compact group? Is it connected? What is the identity element? The group multiplication table can be shown to be:

$$\tau(\epsilon_1, \theta_1, \alpha_1, \beta_1)\tau(\epsilon_2, \theta_2, \alpha_2, \beta_2) = \tau(\epsilon_3, \theta_3, \alpha_3, \beta_3), \tag{4.77}$$

where

$$\epsilon_3 = \epsilon_1\epsilon_2 \tag{4.78}$$

$$\theta_3 = \epsilon_2\theta_1 + \theta_2 \quad [\mathrm{mod}(-\pi,\pi)] \tag{4.79}$$
$$\alpha_3 = \alpha_2\cos\theta_1 - \beta_2\sin\theta_1 + \alpha_1 \tag{4.80}$$
$$\beta_3 = \epsilon_1(\alpha_2\sin\theta_1 + \beta_2\cos\theta_1) + \beta_1. \tag{4.81}$$

What is the inverse $\tau^{-1}(\epsilon,\theta,\alpha,\beta)$?

5. We consider some properties of a group algebra which can be useful for obtaining characters: Let the elements of a class be denoted $\{a_1, a_2, \ldots, a_{n_a}\}$, the elements of another class be denoted $\{b_1, b_2, \ldots, b_{n_b}\}$, etc. Form element $A = \sum_{i=1}^{n_a} a_i$ of the group algebra, and similarly for B, etc.

   Suppose D is an n-dimensional irreducible representation. You showed in problem 19 that
$$D(A) \equiv \sum_{i=1}^{n_a} D(a_i) = \frac{n_a}{n}\chi(A)I, \tag{4.82}$$
   where $\chi(A)$ is the character of irrep D for class A.

   (a) Now consider the multiplication of two elements, A and B, of the group algebra. Show that the product consists of complete classes, i.e.,
$$AB = \sum_C s_C C, \tag{4.83}$$
   where s_C are non-negative integers. You may find it helpful to show that $g^{-1}ABg = AB$ for all group elements g.

   (b) Finally, prove the potentially useful relation:
$$n_a\chi(A)n_b\chi(B) = n\sum_C s_c n_c \chi(C). \tag{4.84}$$

6. We have discussed Lie algrebras (with Lie product given by the commutator) and Lie groups, in our attempt to deal with rotations. At one point, we asserted that the structure (multiplication table) of the Lie group in some neighborhood of the identity was completely determined by the structure (multiplication table) of the Lie algebra. We noted that, however intuitively pleasing this might sound, it was not actually a trivial statement, and that it followed from the "Baker-Campbell-Hausdorff" theorem. Let's try to tidy this up a bit further here.

   First, let's set up some notation: Let $\mathcal{L}$ be a Lie algebra, and $\mathcal{G}$ be the Lie group generated by this algebra. Let $X, Y \in \mathcal{L}$ be two elements of the algebra. These generate the elements $e^X, e^Y \in \mathcal{G}$ of the Lie group. We assume the notion that if X and Y are close to the zero element of the Lie algebra, then e^X and e^Y will be close to the identity element of the Lie group.

   What we want to show is that the group product $e^X e^Y$ may be expressed in the form e^Z, where $Z \in \mathcal{L}$, at least for X and Y not too "large". Note that the non-trivial aspect of this problem is that, first, X and Y may not commute, and second, objects of the form XY may not be in the Lie algebra. Elements of $\mathcal{L}$ generated by X and Y must be linear combinations of X, Y, and their repeated commutators.

(a) Suppose X and Y commute. Show explicitly that the product $e^X e^Y$ is of the form e^Z, where Z is an element of $\mathcal{L}$. (If you think this is trivial, don't worry, it is!)

(b) Now suppose that X and Y may not commute, but that they are very close to the zero element. Keeping terms to quadratic order in X and Y, show once again that the product $e^X e^Y$ is of the form e^Z, where Z is an element of $\mathcal{L}$. Give an explicit expression for Z.

(c) Finally, for more of a challenge, let's do the general theorem: Show that $e^X e^Y$ is of the form e^Z, where Z is an element of $\mathcal{L}$, as long as X and Y are sufficiently "small". We won't concern ourselves here with how "small" X and Y need to be – you may investigate that at more leisure.

Here are some hints that may help you: First, we remark that the differential equation

$$\frac{df}{du} = Xf(u) + g(u), \tag{4.85}$$

where $X \in \mathcal{L}$, and letting $f(0) = f_0$, has the solution:

$$f(u) = e^{uX} f_0 + \int_0^u e^{(u-v)X} g(v)dv. \tag{4.86}$$

This can be readily verified by back-substitution. If g is independent of u, then the integral may be performed, with the result:

$$f(u) = e^{uX} f_0 + h(u, X)g, \tag{4.87}$$

Where, formally,

$$h(u, X) = \frac{e^{uX} - 1}{X}. \tag{4.88}$$

Second, if $X, Y \in \mathcal{L}$, then

$$e^X Y e^{-X} = e^{X_c}(Y), \tag{4.89}$$

where I have introduced the notation "X_c" to mean "take the commutator". That is, $X_c(Y) \equiv [X, Y]$. This fact may be demonstrated by taking the derivative of

$$A(u; Y) \equiv e^{uX} Y e^{-uX} \tag{4.90}$$

with respect to u, and comparing with our differential equation above to obtain the desired result.

Third, assuming $X = X(u)$ is differentiable, we have

$$e^{X(u)}\frac{d}{du}e^{-X(u)} = -h(1, X(u)_c)\frac{dX}{du}. \tag{4.91}$$

This fact may be verified by considering the object:

$$B(t,u) \equiv e^{tX(u)}\frac{\partial}{\partial u}e^{-tX(u)}, \tag{4.92}$$

and differentiating (carefully!) with respect to t, using the above two facts, and finally letting $t = 1$.

One final hint: Consider the quantity

$$Z(u) = \ln\left(e^{uX}e^{Y}\right). \tag{4.93}$$

The series:

$$\ell(z) = \frac{\ln z}{z-1} = 1 - \frac{z-1}{2} + \frac{(z-1)^2}{3} - \cdots \tag{4.94}$$

plays a role in the explicit form for the result. Again, you are not asked to worry about convergence issues.

7. In the next few problems we'll pursue further the example we discussed in the notes and in class with $SU(3)$. We consider systems made from the u, d, and s quarks (for "up", "down", and "strange"). Except for the differences in masses, the strong interaction is supposed to be symmetric as far as these three different "flavors" of quarks are concerned. Thus, if we imagine our matter fields to be a triplet:

$$\psi = \begin{pmatrix} \psi_u \\ \psi_d \\ \psi_s \end{pmatrix}, \tag{4.95}$$

then we expect invariance (under the strong interaction) under the transformations

$$\psi \to \psi' = U\psi, \tag{4.96}$$

where U is any 3×3 matrix. Thus, U is any element of $SU(3)$, and the interaction possesses $SU(3)$ symmetry.

You have already shown that $SU(n)$ is an $(n^2 - 1)$ parameter group. Thus, $SU(3)$ has 8 parameters, and an arbitrary element in $SU(3)$ can be expressed in the form:

$$U = \exp\left\{\frac{i}{2}\sum_{j=i}^{8} a_j\lambda_j\right\}$$

where the $\{\lambda_j\}$ is a set of eight 3×3 traceless, hermitian matrices. One such set is the following: (Gell-Mann)

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix},$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

If the a_j are infinitesimal numbers, we have

$$\psi' = (1 + \frac{i}{2} \sum a_j \lambda_j)\psi$$

and hence, the quantities $\Lambda_j = \frac{1}{2}\lambda_J$ are called the generators of the infinitesimal transformations, or, simply, the generators of the group. These generators satisfy the commutation relations: (and we have a Lie algebra)

$$[\Lambda_i, \Lambda_j] = i f_{ijk} \Lambda_k$$

.

Evaluate the structure constants, f_{ijk}, of $SU(3)$.

8. We may find ourselves interested in "states" consisting of more than one quark, thus we must consider (infinitesimal) transformations of the form

$$\begin{aligned} \psi \to \psi' &= (1 + i\vec{\alpha} \cdot \vec{\Lambda})\psi && (4.97) \\ \vec{\alpha} \cdot \vec{\Lambda} &\equiv \sum_{j=1}^{8} a_j \Lambda_j \end{aligned}$$

where the Λ_j may be represented by matrices of dimension other than 3. Let us develop a simple graphical approach to dealing with this problem (We could also use less intuitive method of Young diagrams, as in the final problem of this problem set).

First, let us introduce the new operators ("canonical form"):

$$\begin{aligned} I_\pm &= \Lambda_1 \pm i\Lambda_2 && (4.98) \\ U_\pm &= \Lambda_6 \pm i\Lambda_7 && (4.99) \\ V_\pm &= \Lambda_4 \pm i\Lambda_5 && (4.100) \\ I_3 &= \Lambda_3 \qquad (\text{"3rd component of isotopic spin"}) && (4.101) \\ Y &= \frac{2}{\sqrt{3}}\Lambda_8 \qquad (\text{"hypercharge"}) && (4.102) \end{aligned}$$

Second, we remark that only two of the 8 generators of $SU(3)$ can be simultaneously diagonalized (e.g., see the explicit λ matrices I wrote down earlier). [Thus, $SU(3)$ is called a group of *rank* 2 – in general, $SU(n)$ has rank $n-1$.] We choose I_3 and Y to be the diagonalized generators. Thus, our states will be eigenstates of these operators, with eigenvalues which will denote by i_3 and y. With the structure constants, you may easily find, e.g.,

$$[I_3, I_\pm] = \pm I_\pm$$

Thus, if $\psi(i_s)$ is an eigenstate of I_3 with eigenvalue i_s:

$$\begin{aligned} I_3 I_+ \psi(i_s) &= I_+(1+I_3)\psi(i_s) = I_+(1+i_s)\psi(i_3) \\ &= (1+i_s)I_+\psi(i_s) \end{aligned} \tag{4.103}$$

So I_+ acts as a "raising" operator for i_3, since $I_+\psi(i_s)$ is again an eigenstate of I_3, with eigenvalue $1+i_s$. Likewise, we have other commutation relations, such as:

$$[I_3, U_\pm] = \mp\frac{1}{2}U_\pm \tag{4.104}$$

$$[I_3, V_\pm] = \pm\frac{1}{2}V_\pm \tag{4.105}$$

$$[Y, I_\pm] = 0 \tag{4.106}$$

$$[Y, U_\pm] = \pm U_\pm \tag{4.107}$$

$$[Y, V_\pm] = \pm V_\pm \tag{4.108}$$

$$[I_3, Y] = 0 \tag{4.109}$$

etc.

Thus, the action of the "raising and lowering" operators $I_\pm, U_\pm, V_\pm$ can be indicated graphically, as in Fig. 4.3.

Thus, we may generate all states of an irreducible representation starting with one state by judicious application of the raising and lowering operators. As a simplest example, and to keep the connection to quarks alive, we consider the 3-dimensional representation: Let's start at the $u-$quark; it has $i_3 = \frac{1}{2}$ and $y = \frac{1}{3}$. See Fig. 4.4.

Why did we stop after we generated d and s, starting from u? Well, clearly we can't have more components (or "occupied sites") than the dimensional-maximum allowed. In fact, since this a 3-dimensional representation, we can just look at the matrices we gave earlier and see that the eigenvalues of I_3 are going to be $\pm\frac{1}{2}$ and 0, and those of Y will be $\frac{1}{3}$, $\frac{1}{3}$, and $-\frac{2}{3}$. A little more consideration of the matrices convinces us that, e.g., $I_+u = 0, I_+s = 0, U_+d = 0, etc.$,

We have given the $i_3 - y$ graph for the "3" representation of $SU(3)$. Now give the corresponding graph for the "3^*" (or $\bar{3}$) representation, that is, the conjugate representation.

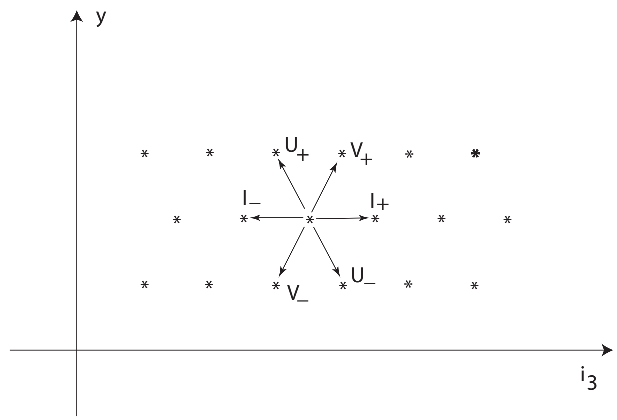

Figure 4.3: The actions of the $SU(3)$ raising and lowering operators $SU(3)$, in the $i_3 - y$ state space.

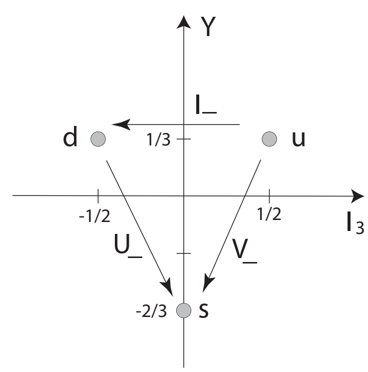

Figure 4.4: The 3 irreducible representation of $SU(3)$, in the $i_3 - y$ state space.

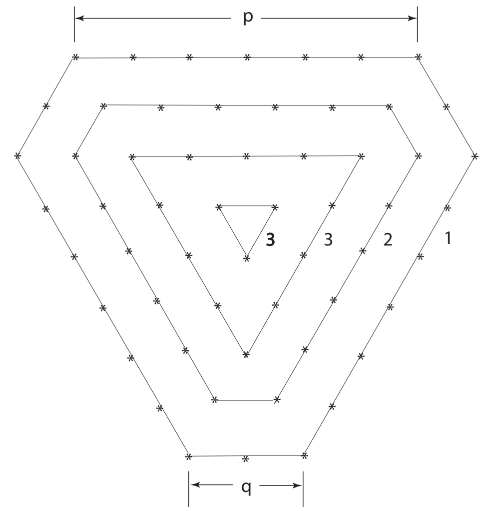

Figure 4.5: The graph of the $SU(3)$ irreducible representation $(p, q) = (6, 2)$. The numbers indicate the multiplicities at each site.

9. You are encouraged to develop the detailed arguments, using the commutation relations for the following observations: The graph for a given irreducible representation is a convex graph which is 6-sided in general (or three-sided if a side length goes to zero). A graph (of an irreducible rep.) is uniquely labelled by two numbers (p, q). An example will suffice to get the idea across. Fig. 4.5 shows the graph for $(p, q) = (6, 2)$. The origin of the $I_3 - y$ coordinate system is inside the innermost triangle. The rule giving the multiplicity of states at each site is that i) the outermost ring has multiplicity of $1, ii$) as one moves to inner rings, the multiplicity increases by one at each ring, until a triangular ring is reached, whereupon no further increases occur.

   By counting the total number of states (*i.e.*, by counting sites, weighted according to multiplicity), we arrive at the dimesionality of the representation. The result, as you may wish to convince yourselves, is

   $$\dim = N = \frac{1}{2}(p+1)(q+1)(p+q+2)$$

   For the 3 and 3^* representations, give the corresponding pairs $(p, q,)$, and check that the dimensions come out correctly.

   One more remark: If we have $p \geq q$, we denote the representation by its dimensionality N. If $p < q$. we call it a conjugate representation, and denote it by $N^*[e.g., (2, 0)$ is the representation 6, but $0, 2$ is 6^*.] An alternative notation is to use a "bar", e.g., $\bar{N}$ to denote the conjugate representa-

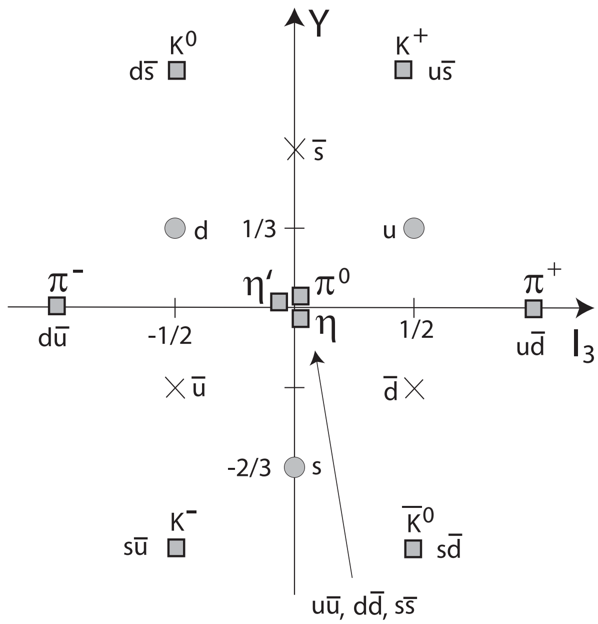

Figure 4.6: The graph of the $SU(3)$ representation for $3 \times \bar{3}$. Physical particle names for the lowest pseudoscalar mesons are indicated at each site.

tion (since for unitary representations the adjoint and complex conjugate representations are the same).

10. We know that the mesons are states of a quark and an antiquark. If you have done everything fine so far, you will see that we can thus generate the mesons by $3 \otimes 3^*$. The result is shown in Fig. 4.6 (don't worry about the particle names, unless you're interested) Using the rules given above concerning irreducible representations, we find, from this graph, the decomposition $3 \otimes 3^* = 8 \oplus 1$.

    We know baryons are made of three quarks (no antiquarks). Make sure you understand how I did the mesons, and apply the same graphical approach to the baryons, and determine the decomposition of $3 \otimes 3 \otimes 3$ into a direct sum of irreducible reps. Do not use Young diagrams (next problem) to do this problem, although you are encouraged to check you answer with Young diagrams. You may find it amusing, if you know something about particle properties, to assign some known baryon names to the points on your graphs.

11. Go to the URL: http://pdg.lbl.gov/2007/reviews/youngrpp.pdf. Study the section on "$SU(n)$ Multiplets and Young Diagrams," and use the techniques described there to answer the following question: We consider

the special unitary group $SU(4)$. This is the group of unimodular unitary 4×4 matrices. We wish to consider the product representation of the irreducible representation given by the elements of the group itself with the irreducible representation formed by the isomorphism of taking the complex conjugate of every element. This turns out to yield a representation which is not equivalent to 4. We could call this new representation 4^*, but it is perhaps more typical to use the notation $\bar{4}$. Note that, since we are dealing with unitary matrices, the complex conjugate and the adjoint representation are identical, so this notation is reasonable.

The question to be answered is: What are dimensions of the irreducible representaations obtained in the decomposition of the product representation $4 \otimes \bar{4}$?

The principal purpose of this problem (which is mechanically very simple) is to alert you to the existence of convenient graphical techniques in group theory – most notably that of Young diagrams. We make no attempt yet to understand "why it works".

A few more words are in order concerning the language on the web page: Since it is taken from the Particle Data Group's "Review of Particle Properties," it is concerned with the application to particle physics, and the language reflects this. However, it is easily understood once one realizes that the number of particles in a "multiplet" is just the dimension of a representation for the group. Effectively, the particles are labels for basis vectors in a space of dimension equal to the multiplet size. [The basic physics motivation for the application of $SU(n)$ to the classification and properties of mesons and baryons is that the strong interaction is supposed to be symmetric as far as the different flavors are concerned. The "n" in this $SU(n)$ is just the number of different flavors. Note that this (flavor) $SU(n)$ is a different application from the "color" $SU(3)$ symmetry in QCD.] Those of you who know something about particles may find it amusing to try to attach some known particle names to the $4 \otimes \bar{4}$ multiplets.