Lie Groups in Modern Physics

A. W. Stetz

September 21, 2011

Contents

Chapter 1

Introduction

New developments in physics are often based on recent developments in mathematics. For example, general relativity is an application of non-Euclidean differential geometry. Quantum mechanics is based on the formalism of abstract vector spaces, and analytic mechanics is based on earlier develpments in the theory of differential equations. The dates are significant. It takes about a generation for physicists to assimilate a new body of mathematics and fit it to an application. There can be no question that the modern theory of elementary particles rests heavily on the formalism of Lie groups. I am reminded of this every time someone refers to the "Standard Model" as $SU(3) \times SU(2) \times U(1)$. Apparently the two terms are interchangable! Lie groups were invented by the Swedish mathematician, Sophus Lie in the late nineteenth century, but most of the theory that we use was developed by E. Cartan, H. Weyl, and others in the 1920's. Apparently their work has still not been completely assimilated; books on group theory written for mathematicians are unreadable for most physicists, and books written for physicists deal mostly with a few familiar examples, such as the rotation group and $SU(3)$, and ignore the underlying theory. Those books that do treat the theory make frequent use of the incantation, "It can be proved that$\cdots$," followed perhaps by a reference to the Proceedings of the Royal Swedish Academy of Mathematics, *circa* 1880.

It has been my intention to write a book that would bridge the gap between the mathematical literature and the examples that can be found in standard physics texts such as Hammermesh, Georgi, and Joshi. I have tried to make the book self contained so that the reader can follow an unbroken trail of logic that begins with the definition of a group and concludes with important examples from modern physics. In order to do this I have or-

ganized the book in terms of definitions, lemmas, theorems and corollaries. I realize that this seems quaint, like an old-fashioned geometry text, but the trail is a long one, and it is necessary to have sign posts. Incidentally, the purpose of proofs, to my way of thinking, is not to defend the truth of a theorem before some court of law, but rather to reveal the inner workings of the formalism. My goal in presenting them is primarily to provide insight rather than conciseness or mathematical rigor. I have no compunctions about skipping a proof if the proof is obvious, or lavishing pages on a proof if I myself found it difficult to understand.

1.1 The Definition of a Group and Some Examples

In this chapter we introduce the formal definition of a group, and, after some examples, specialize to Lie groups. Rather than giving a formal definition of Lie groups at this point we present an intuitive survey of the rotation group as it is usually presented in quantum mechanics texts. This provides an example of all the basic ingredients of Lie group theory imbedded in a familiar context. In later chapters we will develop the mathematical formalism required to treat these topics rigorously and illustrate them with other examples.

Definition 1.1 *Group $\mathcal{G}$*

A set $\mathcal{G}$ of elements $a, b, c, \ldots$ is called a group *if the following four axioms are satisfied:*

(a) Any two elements, a and b, can be combined to make a third element c. There are exactly two ways of doing this, which we write

$$ab = c$$

$$ba = c'$$

in general $ab \neq ba$.

(b) For any three elements a, b, and c of $\mathcal{G}$

$$(ab)c = a(bc)$$

The left side of this equation is interpreted as follows: a and b are first combined, and the resulting element is then combined with c. On the right side, b and c are first combined and then subsequently combined with a.

(c) The group $\mathcal{G}$ contains an identity element e such that

$$ae = ea = a$$

for every element a in $\mathcal{G}$.

(d) For each element a of $\mathcal{G}$ there exists an inverse element a^{-1} which is also contained in $\mathcal{G}$ such that

$$aa^{-1} = a^{-1}a = e$$

This definition is very general and abstract. It does not specify what kinds of things the elements are or what sort of operation combines them. This operation is usually called "group multiplication" (or simply "multiplication"), but ordinary multiplication of two real numbers is only one instance of group multiplication. A few other instances are provided by the following examples:

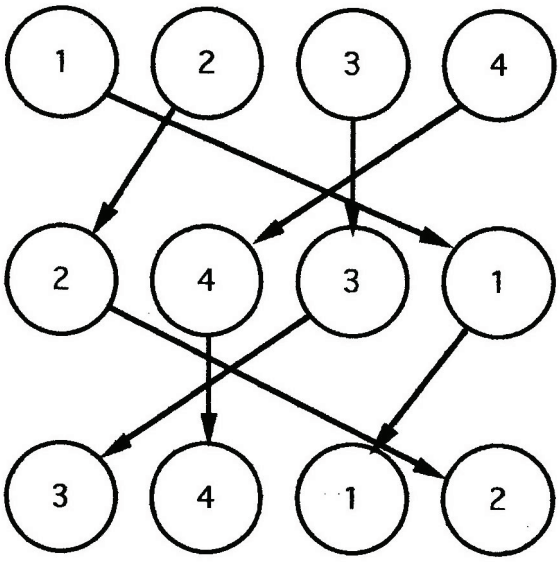


Figure 1.1: The positions of four numbered balls after two successive rearrangements.

Example 1.1 *The permutation group S_n.*

This group consists of the n! permutations of n distinguishable objects. To make this more concrete, think of four numbered balls, which could be rearranged in 4! distinct ways. Two of these permutations are illustrated in Fig. 1.1. It should be stressed that the group elements are not the configurations of these balls, but rather the operations of interchanging them (*ie.* the arrows in the Fig 1.1). Group multiplication ba means that one performs the operation a on the balls first, then rearranges them according to the instructions that constitute element b. The same end result can always be obtained with a single permutation, which is the group element $c = ba$ shown in Fig. 1.2.

Each group element can be designated by listing n integers. In the example of the four balls, element a corresponds to (4,1,3,2); in other words, the ball currently in the first position is moved to the fourth position, the second ball is moved to the first position, *etc.*. In this notation $ba = c$ becomes

$$(4, 2, 1, 3)(4, 1, 3, 2) = (3, 4, 1, 2)$$

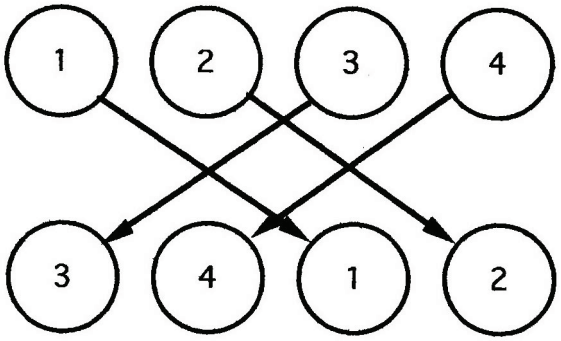


Figure 1.2: A single permutation yielding the same configuration as the two permutations of Fig. 2.1.

Since there are $n!$ elements and $(n!)^2$ ways of combining them, we could construct a table using this notation with $n!$ rows and $n!$ columns listing the outcome of all possible combinations. The first group axiom insures that such a table, called the *group multiplication table*, can always be completed (in principle). Since we will usually be dealing with groups with an infinite number of elements, it will be necessary to replace the multiplication table with a mathematical rule or function that computes the outcome of any arbitrary product of elements.

Example 1.2 *The addition group of integers.*

Take $\mathcal{G}$ to be the set of all integers N (positive, negative, and zero). Define the combination rule to be ordinary addition. In this case the identity is 0 (as $N + 0 = 0 + N = N$), and the inverse of an integer is its negative, $-N$ (as $N + (-N) = (-N) + N = 0$).

The group in Example 1.1 had a finite ($n!$) number of elements. In the second example the elements were infinite in number but still countable. Classical group theory deals mainly with such groups (called *finite groups* and *infinite discrete groups* respectively). Lie groups, on the other hand, have a non-countable infinity of elements. The group of Example 1.2 could be redefined as a Lie group by taking $\mathcal{G}$ to be the set of all real numbers (rather than integers). The elements can no longer be counted, but we can represent them with a single real continuous variable $\alpha = N$. We might say that there is a functional relationship between α and the group elements, since each numerical value of α corresponds to an element of the group.

The "function," of course, is just the identity in this case; but the fact that the group elements are specified by one (or more) continuous variable(s) through some algebraically defined function is the defining characteristic of a *Lie group*. The additive group of real numbers is thus the simplest possible example of a *one-parameter Lie group*. In general, a Lie group will have more than one parameter, *ie.* the elements will depend on several variables through some complicated functional relationship. Another ingredient in the definition is that this function is "smooth" and differentiable, so that the techniques of calculus can be applied to it. Lie group theory is thus the product of the marriage of group theory with algebra and calculus. A more formal definition is given in Chapter 2. In the meantime we consider some less trivial examples.

Example 1.3 *The matrix multiplication group.*

Take $\mathcal{G}$ to be the set of all non-singular $n \times n$ matrices. The group multiplication rule is ordinary matrix multiplication.. The product of two $n \times n$ matrices is itself of dimension $n \times n$, and the product of two non-singular matrices is non-singular. Matrix multiplication automatically satisfies the associative law, axiom (b). The identity element is the $n \times n$ unit matrix I. Finally, every non-singular matrix M has an inverse M^{-1} with $M^{-1}M = MM^{-1} = I$, so the group axioms are satisfied. This set of matrices is called the *general linear group of dimension n* abbreviated $Gl(n)$.

The group $Gl(n)$ can be regarded as a Lie group if the matrix elements are parameterized appropriately. For example, each matrix element could be a separate parameter or some function of several parameters. Thus n^2 parameters are required to specify the group completely.

Example 1.4 *The rotation group.*

Consider the set of rotations of a rigid object. Clearly these rotations constitute a group, but what are the group elements? There are at least four alternatives of consider.

(a) Let the group elements correspond to the actual rotations of the object. As in Example 1.1, the elements are operations. Group multiplication consists of making one rotation followed by another.

(b) The rotations can be parameterized with a set of three angles. The conventional choice is the set of Eulerian angles ψ, θ, and ϕ illustrated in Figure 1.3. We use two coordinate systems, a x, y, z system fixed in space and a x', y', z' system attached to the rigid object. Before any rotations have

been performed the two systems coincide. The first rotation is made around the z-axis, so that the x'-axis is displaced from the x-axis, and the y'-axis is displaced from the y-axis by an angle ϕ. The object is then rotated about its y'-axis, so that the z and z'-axes are displaced by an angle θ. Finally, the object is rotated through an angle ψ about its z'-axis. Each set of angles $0 \leq \phi \leq 2\pi$, $0 \leq \theta \leq \pi$, and $0 \leq \psi \leq 2\pi$ specifies a rotation. We can thus associate each triplet of numbers (ψ, θ, ϕ) satisfying the above inequalities with a group element. The identity element is $(0, 0, 0)$, and the inverse of (ψ, θ, ϕ) is $(-\phi, -\theta, -\psi)$.

The first axiom, $ba = c$, is interpreted as follows. Associate the element a with the triplet $(\psi_1, \theta_1, \phi_1)$ and the element b with a second rotation $(\psi_2, \theta_2, \phi_2)$. The final position of the object after these two operations have been performed could have been achieved with a single rotation $(\psi_3, \theta_3, \phi_3)$, which we associate with the element c. The group closure property (axiom a) implies the existence of a set of three functions

$$\begin{aligned} \phi_3 &= \phi_3(\psi_2, \theta_2, \phi_2; \psi_1, \theta_1, \phi_1) \\ \theta_3 &= \theta_3(\psi_2, \theta_2, \phi_2; \psi_1, \theta_1, \phi_1) \\ \psi_3 &= \psi_3(\psi_2, \theta_2, \phi_2; \psi_1, \theta_1, \phi_1) \end{aligned} \tag{1.1}$$

from which $(\psi_3, \theta_3, \phi_3)$ could be calculated given $(\psi_1, \theta_1, \phi_1)$ and $(\psi_2, \theta_2, \phi_2)$. These functions are the analog of the group multiplication table discussed in Example 1.1.

(c) The group elements can also be represented by matrices. For example, imagine a rigid object rotating about a fixed turning point, which is used as the origin of a coordinate system as shown in Figure 1.3. As a result of the rotation, any point P in the object will move to a new position P'. We define a *position vector* $\mathbf{x}_P$, which can be thought of as an arrow drawn from the origin to P. We will write

$$\mathbf{x}_P = \sum_{i=1}^{3} x_{P_i} \mathbf{e}_i \tag{1.2}$$

where $\mathbf{e}_i$, $i = 1, 2, 3$ are a set of three mutually orthogonal unit vectors parallel to the axes of the coordinate system, and the x_{P_i} are the coordinates of the point P.

A rotation of the object will change the components of the position vector, and leave the unit vectors unchanged.

$$\mathbf{x}_{P'} = \sum_{i=1}^{3} x_{P'_i} \mathbf{e}_i \tag{1.3}$$

A pure rotation must leave invariant the length of this vector, *ie.*

$$\mid \mathbf{x}_P \mid = \mid \mathbf{x}_{P'} \mid$$

as well as the angle between $\mathbf{x}_P$ and any other coordinate vector, say $\mathbf{x}_Q$. Both conditions will be met if the scalar product

$$\mathbf{x}_P \cdot \mathbf{x}_Q = \sum_{i,j=1}^{3} x_{P_i} x_{Q_j} \mathbf{e}_i \cdot \mathbf{e}_j = \sum_{i=1}^{3} x_{P_i} x_{Q_i} \tag{1.4}$$

is invariant under rotation. Define a set of 3×3 matrices M_{ij} so that

$$x'_i = \sum_{j=1}^{3} M_{ij} x_j \tag{1.5}$$

The invariance of the scalar product requires that

$$\begin{aligned} \mathbf{x}_{P'} \cdot \mathbf{x}_{Q'} = \sum_{i=1}^{3} x'_{P_i} x'_{Q_i} &= \sum_{i,j,k=1}^{3} M_{ij} M_{ik} x_{P_j} x_{Q_k} \\ &= \mathbf{x}_P \cdot \mathbf{x}_Q = \sum_{j,k=1}^{3} x_{P_j} x_{Q_k} \delta_{jk} \end{aligned} \tag{1.6}$$

consequently

$$\sum_{i=1}^{3} M_{ij} M_{ik} = \delta_{jk} \tag{1.7}$$

or, in matrix notation, $M^T M = I$, where $(M_{ij})^T = M_{ji}$ and I is the 3×3 unit matrix. Matrices that satisfy this condition are said to be *orthogonal* matrices. We are thus tempted to associate the group elements with the set of all 3×3 orthogonal matrices. There is still a subtlety to consider, however. Taking the determinant of Equation (1.7) we get

$$det(M^T M) = (detM)^2 = detI = 1$$

Thus Equation (1.7) is sufficient to ensure that M is non-singular, but not sufficient to eliminate those M's with determinant $= -1$. Such transformations can only be achieved by inverting one (or all three) of the coordinate axes, *ie.* by replacing the object by its mirror image. If we restrict the group to proper rotations, *ie.* no reflections, then the elements correspond to the set of all 3×3 orthogonal unimodular matrices. This is the matrix group

SO(3); the S stands for "special" indicating the the determinant is equal to +1. Without this restriction, the group is O(3), the orthogonal group of 3×3 matrices.

The product of two rotation matrices corresponds to two successive rotations. Thus if R_i stands for the triplet of Euler angles $(\psi_i, \theta_i, \phi_i)$, and $M(R_i)$ is the matrix that rotates an object through these angles, then

$$M(R_2)M(R_1) = M(R_3), \tag{1.8}$$

where R_3 is the set of angles $(\psi_3, \theta_3, \phi_3)$ given by (1.1). Later in this section we will return to the problem of constructing matrices with this property.

(d) We could also use the group elements to represent *operators*. Let $\psi(\mathbf{x})$ denote a scalar field, some scalar-valued function of the position vector $\mathbf{x}$. This field associates a number with each point in space, and so, in a sense, is independent of any coordinate system. In order to write it as an explicit function, however, some coordinate system is required. Once the choice is made ψ becomes a function of the *components* of $\mathbf{x}$. For example, if ψ represents the electrostatic potential associated with a uniform electric field $\mathbf{E}$, we could write

$$\psi(\mathbf{x}) = -Ex \tag{1.9}$$

We have obviously chosen a coordinate system whose x-axis is parallel to $\mathbf{E}$, and the definition of ψ is now locked into this frame of reference.

Suppose we were to evaluate ψ in a different coordinate system. Such a transformation would leave the numerical value of ψ invariant at each point in space, but it would change its functional form. This change can be thought of as the effect of an operator, conventionally called the *coordinate transformation operator*. In this example we limit ourselves to pure rotations, but the extension to more general coordinate transformations is quite straightforward.

The functional form of ψ is originally defined in terms of a coordinate system given by three orthogonal unit vectors. The position vector $\mathbf{x}$ is given by

$$\mathbf{x} = \sum_{i=1}^{3} x_i \mathbf{e}_i \tag{1.10}$$

and ψ is an explicit function of the components x_i. The *coordinate system* is now rotated through a set of Eulerian angles designated by $R = (\psi, \theta, \phi)$. The new unit vectors are

$$\mathbf{e'}_j = \sum_{i=1}^{3} M(R)_{ji} \mathbf{e}_i. \tag{1.11}$$

The coordinate vector $\mathbf{x}$ remains invariant (since it points to a fixed point in space),

$$\mathbf{x} = \sum_{i=1}^{3} x_i \mathbf{e}_i = \sum_{j=1}^{3} x'_j \mathbf{e}'_j = \sum_{i,j=1}^{3} x'_j M(R)_{ji} \mathbf{e}_i,$$

so

$$x_i = \sum_{i=1}^{3} x'_j M(R)_{ji} = \sum_{j=1}^{3} [M^{-1}(R)]_{ij} x'_j \tag{1.12}$$

or

$$x'_j = \sum_{i=1}^{3} M(R)_{ji} x_i. \tag{1.13}$$

Let us regard ψ as a function of the components of $\mathbf{x}$ and emphasize this with the notation $\psi(x_i)$. We seek a new function ψ' of the transformed coordinates such that

$$\psi'(x'_i) = \psi(x_i).$$

This means that the numerical value of the new function ψ' evaluated with the new coordinates x'_i is equal to the value of the original function evaluated with the original coordinates. So

$$\psi'(x'_i) = \psi(\sum_{j=1}^{3} [M^{-1}(R)]_{ij} x'_j).$$

It is customary to drop the primes on the components with the warning that the resulting equation refers only to the new (rotated) coordinate system. Then

$$\psi'(x_i) = \psi(\sum_{j=1}^{3} [M^{-1}(R)]_{ij} x_j). \tag{1.14}$$

This can be illustrated with the example of the uniform electric field, (1.9). If the coordinate axes are rotated through an angle θ around the z-axis, the components of $\mathbf{x}$ are transformed by

$$M(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then

$$\psi'(x'_i) = -E(x'_1 \cos\theta - x'_2 \sin\theta),$$

or, if we agree to abandon the original coordinate system

$$\psi'(x_i) = -E(x_1 \cos\theta - x_2 \sin\theta).$$

Now define the coordinate transformation operator $P(R)$.

$$P(R)\psi(x_i) = \psi'(x_i) = \psi(\sum_{j=1}^{3}[M^{-1}(R)]_{ij}x_j)$$

This equation establishes a one-to-one relationship between the rotation matrices $M(R)$ and the operator $P(R)$. Since the matrices form a group, the operators must themselves form a group. The product of two operators is

$$P(R_2)P(R_1)\psi(x) = P(R_2)\psi(M^{-1}(R_1)x)$$
$$= \psi[M^{-1}(R_1)(M^{-1}(R_2)x)] = \psi(M^{-1}(R_3)x).$$

The last step uses (1.8) to introduce the matrix $M(R_3)$. The product of two operators can thus be written

$$P(R_2)P(R_1) = P(R_3). \tag{1.15}$$

A digression on conventions and notation is appropriate at this point. In Example 1.3c the vector $\mathbf{x}_P$ is rotated to form a new vector $\mathbf{x}_{P'}$. The vector components change, but the unit vectors remain fixed. In Example 1.3d the coordinate system is rotated so that the unit vectors (1.11) and vector components (1.13) transform leaving the vector $\mathbf{x}$ invariant. These two transformations are called *active* and *passive* transformations respectively. Both conventions are used in the literature, often without clearly identifying which is intended.[1] In most cases we will use the passive viewpoint. Exceptions will be clearly noted.

There is another convention that tends to confuse the matter further. Many textbooks use the symbol $\mathbf{x}$ to indicate the *components* of the vector, *ie.* $\mathbf{x} = (x_1, x_2, x_3)$. With this notation the "vector," $\mathbf{x}$, changes $\mathbf{x} \neq \mathbf{x}'$, even in the case of a passive rotation.

In this text, the boldface symbol $\mathbf{x}$ applied to a position vector will always imply a linear combination of unit vectors as in (1.10). (We will use the boldface notation with Greek letters to indicate parameter arrays in subsequent chapters. These arrays are *not* position vectors.) Vector

[1] There is actually a third convention in which the vectors and unit vectors transform in opposite directions leaving the components invariant. [Messiah, 1962, Vol. II] This might be called a "hyperactive" transformation.

components will usually be written with a subscript or superscript, *eg.* x_i or x^μ. Occasionally we will eliminate the sub- and superscripts when no confusion is likely, *eg.* $x' = Mx$ is shorthand for (1.13).

It remains to show how the rotation matrices defined in (1.5) are related to the Euler angles, $R = (\psi, \theta, \phi)$. These angles are defined in terms of a sequence of three rotations. In the passive view, each rotation can be represented by a matrix operating on the unit vectors as in (1.11).

$$M(R) = M_{z''}(\psi) M_{y'}(\theta) M_z(\phi) \tag{1.16}$$

The matrix on the right rotates the coordinate system by an angle ϕ around the z-axis yielding a new set of unit vectors, $\mathbf{e'}_i \ i = 1, 2, 3$. The second matrix rotates this coordinate system around the $\mathbf{e'}_2$ axis, and the resulting unit vectors are $\mathbf{e''}_i$. The final rotation is taken around $\mathbf{e''}_3$. Thus $M(R)$ is the product of three simple matrices.

$$M_z(\phi) = \begin{vmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{vmatrix} \quad M_{y'}(\theta) = \begin{vmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{vmatrix}$$

$$M_{z''}(\psi) = \begin{vmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

Multiplying these three matrices in the order of (1.16) yields

$$\begin{aligned}
M_{11} &= \cos\psi\cos\theta\cos\phi - \sin\psi\sin\phi \\
M_{12} &= \cos\psi\cos\theta\sin\phi + \sin\psi\cos\phi \\
M_{13} &= -\cos\psi\sin\theta \\
M_{21} &= -\sin\psi\cos\theta\cos\phi - \cos\psi\sin\phi \\
M_{22} &= -\sin\psi\cos\theta\sin\phi + \cos\psi\cos\phi \\
M_{23} &= \sin\psi\sin\theta \\
M_{31} &= \sin\theta\cos\phi \\
M_{32} &= \sin\theta\sin\phi \\
M_{33} &= \cos\theta
\end{aligned} \tag{1.17}$$

The angles are positive when drawn counterclockwise around the axes of a right-handed coordinate system. Passive rotation is assumed. To convert (1.17) to active rotations, simply change the sign of all angles.

It will be useful later on to decompose $M(R)$ into three separate rotations referred to the *same* coordinate axes. To this end we can write

$$M_{y'}(\theta) = M_z(\phi)M_y(\theta)M_z^{-1}(\phi) \tag{1.18}$$

The matrix on the right restores the coordinate axes $\mathbf{e}'_i$ to their original position. The next matrix rotates the coordinate system about the original y-axis. The coordinate system is then rotated back around the original z-axis. Using the same argument with the other rotations gives

$$M_{z''}(\psi) = M_{y'}(\theta)M_{z'}(\psi)M_{y'}^{-1}(\theta) \tag{1.19}$$

and

$$M_{z'}(\psi) = M_z(\phi)M_z(\psi)M_z^{-1}(\phi) \tag{1.20}$$

Substituting equations (1.18), (1.19), and (1.20) into (1.16) yields

$$M(R) = M_z(\phi)M_y(\theta)M_z(\psi) \tag{1.21}$$

Each rotation is referred to the same coordinate axes, but the rotations are performed in opposite order from (1.16).

Many alternative definitions of the Euler angles appear in the literature; some authors use the x-axis for the second rotation, some apparently use a left-handed coordinate system, and there is no agreement regarding the names of the angles. Quantum mechanics texts universally use the z-axis for the first and third rotations, however. The reason is that the z-axis is always chosen as the axis of quantization, so that the first and third rotations only change the wave function by a phase factor, and the resulting rotation matrices are greatly simplified. This definition has a pathology, however, that must be avoided in some applications. We will return to this point in Section 1.3.

1.2 Isomorphism and Homomorphism

In Example 1.4 we considered four rather different classes of objects as group elements: physical rotations, sets of Eulerian angles, matrices, and coordinate transformation operators. Each set constituted a group, and in some sense, it was always the same group. The phrase "in some sense" can be made more precise with the notion of isomorphism and homomorphism, which we now define.

Definition 1.2 *Mapping*

Let $\mathcal{G}$ and $\mathcal{G}'$ be two groups. A mapping ϕ of $\mathcal{G}$ onto $\mathcal{G}'$ is a rule that associates each element a of $\mathcal{G}$ with some element a' of $\mathcal{G}'$. Symbolically

$$a' = \phi(a).$$

The element a' is called the image of a. If every $a \in \mathcal{G}$ has an image $a' \in \mathcal{G}'$, we say that ϕ maps $\mathcal{G}$ **into** *$\mathcal{G}'$. If every $a' \in \mathcal{G}'$ is the image of some $a \in \mathcal{G}$, we say that ϕ maps $\mathcal{G}$* **onto** *$\mathcal{G}'$. In general a mapping from $\mathcal{G}$ onto or into $\mathcal{G}'$ does not imply a mapping from $\mathcal{G}'$ to $\mathcal{G}$. However if the mapping is one-to-one, the inverse exists and is unique, $a = \phi^{-1}(a')$. Such a mapping is said to be* **faithful**.

Definition 1.3 *Homomorphic mapping of a group $\mathcal{G}$ onto a group $\mathcal{G}'$.*

Let a and b be any two elements of $\mathcal{G}$. If ϕ is a mapping as defined above and if $\phi(b)\phi(a) = \phi(ba)$, then ϕ is said to be a homomorphic mapping. In other words, a homomorphic mapping preserves the group multiplicaton table.

Definition 1.4 *Isomorphic mapping of a group $\mathcal{G}$ onto a group $\mathcal{G}'$.*

If ϕ is a one-to-one mapping and $\phi(b)\phi(a) = \phi(ba)$, it is said to be an isomorphic mapping.

Clearly the four groups presented in Example 1.4 are related to one another through homomorphic mapping. To each physical rotation element there corresponds at least one set of Eulerian angles (θ, ψ, ϕ), at least one rotation matrix $M(R)$, and at least one coordinate transformation operator $P(R)$. Equations (1.1), (1.8), and (1.15) guarantee that the various mappings preserve the group multiplicaton table. The groups $O(3)$ and $SO(3)$ provide a simple example of a homomorphic mapping, since $O(3)$ contains every element of $SO(3)$ as well as every element that can be obtained by multiplying an element of $SO(3)$ by $-I$. Each element of $SO(3)$ is thus an

image of exactly two elements of $O(3)$; the mapping is homomorphic from $O(3)$ onto and into $SO(3)$.

Group theory provides a universal formalism for exploiting the properties of a group without regard to the explicit mathematical form of its elements. This is part of the power and usefulness of the theory; but there is a deeper aspect, which was discovered by Lie in the late nineteenth century. The continuous groups (subject to a few general restrictions) are manifestations of a much simpler mathematical structure called the Lie algebra. In its most general form the Lie algebra consists of a set of commutation relations among elements of an abstract algebra. In quantum mechanics the Lie algebra appears as a set of commutation relations among Hermitian operators defined on a Hilbert space. These operators are the familiar observables like momentum and angular momentum; but their existence indicates the presence of an underlying group, which in the case of non-relativistic quantum mechanics is the group of translations and rotations in three-dimensional Euclidian space. In any case, these simple commutation relations contain all the information required to completely reconstruct the group from which they were derived, as well as a host of other groups related to this group by homomorphic and isomorphic mapping. In this sense the Lie algebra is not unlike a strand of the DNA molecule, which contains, in principle, the information neccessary to reconstruct the entire organism from which it was taken.

In the remainder of this chapter we will explore these ideas in the context of the rotation group. In subsequent chapters we will develop the general formalism for any Lie group.

1.3 The Lie Algebra of the Rotation Group

The analysis of a Lie group in terms of its Lie algebra requires a careful study of the properties of the group in the neighborhood of the identity. For this purpose the parameterization of the rotation group in terms of the usual Euler angles as discussed in the previous section is unsatisfactory. The reason is that when $\psi = 0$, the z- and z''- axes coincide, so that ϕ and θ are no longer independent rotations. For example, $M(R) = I$ for all R of the form $R = [\phi, 0, -\phi]$, where ϕ is any arbitrary angle. Consequently the mapping from the group of rotation matrices to the group of Euler angles is infinitely multiple valued in the neighborhood of the identity.

Any parameterization of the rotation matrices will suffer from this pathology at some point. For our purposes it is sufficient to insure that these points are safely removed from the identity. The trick is obviously to choose the rotations around three distinct axes. We therefore define the angle α_1 to be a rotation about the x (or $\boldsymbol{e}_1$) axis. The angle α_2 is then defined around the new y', axis; and α_3 corresponds to a rotation about the z'' axis. This moves the singular point to $\alpha_2 = \pm\pi/2$, whereupon the x- and z''-axes coincide. This choice leads to the following functions for the matrix elements:

$$\begin{aligned}
M(R)_{11} &= \cos\alpha_2 \cos\alpha_3 \\
M(R)_{12} &= \sin\alpha_1 \sin\alpha_2 \cos\alpha_3 + \cos\alpha_1 \sin\alpha_3 \\
M(R)_{13} &= -\cos\alpha_1 \sin\alpha_2 \cos\alpha_3 + \sin\alpha_1 \sin\alpha_3 \\
M(R)_{21} &= -\cos\alpha_2 \sin\alpha_3 \\
M(R)_{22} &= \cos\alpha_1 \cos\alpha_3 - \sin\alpha_1 \sin\alpha_2 \sin\alpha_3 \\
M(R)_{23} &= \cos\alpha_1 \sin\alpha_2 \sin\alpha_3 + \sin\alpha_1 \cos\alpha_3 \\
M(R)_{31} &= \sin\alpha_2 \\
M(R)_{32} &= -\sin\alpha_1 \cos\alpha_2 \\
M(R)_{33} &= \cos\alpha_1 \cos\alpha_2
\end{aligned} \tag{1.22}$$

We have assumed a right-handed coordinate system. The angles are positive when measured in the direction a right-handed screw would turn if it were moving along the rotation axis in the positive direction.

For notational convenience we regard the three angles, $\alpha_1, \alpha_2, \alpha_3$, as forming the three components of a vector $\boldsymbol{\alpha}$.

$$M(R) = M(\alpha_1, \alpha_2, \alpha_3) = M(\boldsymbol{\alpha})$$

The functions listed in (1.22) have three properties that are essential for what we are about to do.

(1) $M(\mathbf{0}) = I$, the identity element.

(2) The elements of M are continuous, infinitely differentiable functions of the components of $\boldsymbol{\alpha}$.

(3) $M(\boldsymbol{\alpha})$ is an isomorphic mapping (at least in the neighborhood of the identity) of the group of the $\boldsymbol{\alpha}$'s onto the group of the M's.

These three conditions guarantee that $M(\boldsymbol{\alpha})$ can be expanded in a power series in the neighborhood of the identity.[2]

$$M(\boldsymbol{\alpha}) = I + \sum_{i=1}^{3} \alpha_i X_i + \frac{1}{2} \sum_{i,j=1}^{3} \alpha_i \alpha_j X_i X_j + O(\alpha^3) \tag{1.23}$$

The X's are a set of three constant matrices defined by

$$X_i = \left(\frac{\partial M}{\partial \alpha_i} \right)_{\boldsymbol{\alpha}=\mathbf{0}} \tag{1.24}$$

If $\boldsymbol{\alpha}$ is an infinitesimal quantity, M will have an inverse given by

$$M^{-1}(\boldsymbol{\alpha}) = I - \sum_{i=1}^{3} \alpha_i X_i + \frac{1}{2} \sum_{i,j=1}^{3} \alpha_i \alpha_j X_i X_j + O(\alpha^3)$$

Now let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ represent two sets of Euler angles, both infinitesimally close to zero. The following sequence of transformations, (1) $\boldsymbol{\beta}^{-1}$, (2) $\boldsymbol{\alpha}^{-1}$, (3) $\boldsymbol{\beta}$, (4) $\boldsymbol{\alpha}$, is called the *commutator* of the two group elements $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. In terms of matrices

$$M(\boldsymbol{\alpha})M(\boldsymbol{\beta})M^{-1}(\boldsymbol{\alpha})M^{-1}(\boldsymbol{\beta}) = I + \sum_{i,j=1}^{3} \alpha_i \beta_j (X_i X_j - X_j X_i) + O(\alpha^3) \tag{1.25}$$

Thus the group theoretical commutator, when applied to infinitesimal transformations, leads naturally to the usual quantum mechanics definition of the commutator of two linear operators

$$[X_i, X_j] = X_i X_j - X_j X_i \tag{1.26}$$

The commutator of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ must itself be a group element, which we call $\boldsymbol{\gamma}$.

$$M(\boldsymbol{\alpha})M(\boldsymbol{\beta})M^{-1}(\boldsymbol{\alpha})M^{-1}(\boldsymbol{\beta}) = M(\boldsymbol{\gamma}) \tag{1.27}$$

[2]See Section 1.5 for a derivation of this formula.

It seems natural to assume that $\boldsymbol{\gamma}$ is also infinitesimally close to the origin, so that $M(\boldsymbol{\gamma})$ has an expansion like equation (1.23). Comparing leading terms in equations (1.23) and (1.25) yields

$$[X_i, X_j] = \sum_{k=1}^{3} C_{ij}^k X_k. \tag{1.28}$$

The factors labeled C_{ij}^k are a set of n^3 numbers called the *structure constants*. It will be proved in Section 3.4 that the X_i's are linearly independent. Anticipating this result we substitute (1.28) into (1.25) and obtain an equation for C_{ij}^k,

$$\gamma_k = \sum_{i,j=1}^{3} C_{ij}^k \alpha_i \beta_j. \tag{1.29}$$

The X_i are called the *infinitesimal group generators*. They can be calculated once and for all from the parameterization of the matrices. For example, the generators of the rotation group can be found by differentiating (1.22) and setting $\boldsymbol{\alpha} = 0$.

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad X_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad X_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.30}$$

These matrices satisfy the commutation relationship

$$[X_i, X_j] = -\sum_{k=1}^{3} \epsilon_{ijk} X_k \tag{1.31}$$

so that $C_{ij}^k = \epsilon_{ijk}$. it is no accident that C_{ij}^k is a constant independent of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$. We will prove in Chapter 3 a fundamental theorem that the commutator of any two group generators can be written as a linear combination of generators with constant coefficients. The structure constants C_{ij}^k are determined ultimately by the parameterization and the multiplication table of the group. Thus $-\epsilon_{ijk}$ contains for the rotation group all the information of (1.22) in a (very) compact form.

The angular momentum commutation relations of quantum mechanics

$$[J_i, J_j] = i\hbar \sum_{k=1}^{3} \epsilon_{ijk} J_k \tag{1.32}$$

are obtained from (1.32) with the trivial replacement $X_i = \imath J_i/\hbar$. The $\imath$ is included to make the operators Hermitian, and $\hbar$ provides the appropriate dimensionality and scale.

This is the point at which disaster would have overtaken if we had used the usual Euler angles with a singular point at $\boldsymbol{\alpha} = 0$. The Taylor series expansion (1.23) would be invalid, and the resulting commutation relations would be meaningless.

There are, of course, many other satisfactory ways of parameterizing the rotation matrices; and each parameterization will produce a different set of X_i's. These sets differ from one another in a trivial way, however. It can be shown that each parameterization yields a set of three linearly independent matrices, and each reparameterization produces X_i's that are linear combinations of the original matrices. The X_i's can be regarded as the basis vectors of a linear vector space whose dimension equals the number of independent parameters in the group. Reparameterizing the group corresponds to choosing a new basis in the same vector space.

1.4 Formal Definitions

Starting with a definition of rotations in terms of a multiplication table (1.1), we are led finally to an abstract vector space governed by the commutation relations, equation (1.28). This vector space is the inner sanctum of the theory, the Lie algebra. We first give a formal definition.

Definition 1.5 *Real Lie Algebra* $\mathcal{L}$

A real Lie algebra of dimension n is a real vector space of dimension n on which is defined a commutator $[X, Y]$ *for every* X *and* $Y \in \mathcal{L}$ *such that*

(a) $[X, Y] \in \mathcal{L}$ *for all* $X, Y \in \mathcal{L}$

(b) $[c_1X + c_2Y, Z] = c_1[X, Y] + c_2[Y, Z]$ *for all* $X, Y \in \mathcal{L}$ *and all real numbers* c_1 *and* c_2*.*

(c) $[X, Y] = -[Y, X]$ *for all* $X, Y \in \mathcal{L}$*.*

(d) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ *for all* $X, Y, Z \in \mathcal{L}$*. (The Jacobi identity.)*

The definition does not specify what sort of object X,Y, and Z are except that they satisfy the usual postulates of a real vector space. Any pair of these objects can be combined in a special way called the "commutator," which is also not specified except that it must satisfy postulates (b) through (d). If the original group is a matrix group, then the elements of the Lie algebra are matrices defined by (1.24), the commutator is defined by (1.26),

and postulates (b) through (d) are satisfied automatically. There is an advantage, however, in regarding X,Y, and Z as members of an *abstract Lie algebra* for which the commutator is left undefined. The point is that any interesting group, such as the rotation group, is in fact a family of groups related by homomorphisms. Thus the rotation group can be considered as a group of abstract elements, a group of matrices (with various dimensions), or a group of operators on some vector space (there are many possibilities). In each case it is possible to define the elements X,Y,Z, and the commutator, so that all these groups that are homomorphic to one another have the same abstract Lie algebra.[3] Thus the abstract Lie algebra is a manifestation of the basic group structure and independent of the specific form in which the group elements are clothed.

Example 1.5 *The Angular Momentum Operators.*

These ideas can be illustrated by considering the Lie algebra of the rotation group of operators given in Example 1.4(d). As usual we take an infinitesimal transformation in the neighborhood of the identity. If the angles α_i in (1.22) are replaced by infinitesimal angles $\delta\alpha_i$, then the components x_i' given by (1.5) become

$$x_i' = x_i + \delta x_i$$

where

$$\delta x_i = -\sum_{j,k=1}^{3} \epsilon_{ijk}\delta\alpha_j x_k$$

Using (1.14)

$$P(\delta\alpha)\psi(x) = \psi(x - \delta x),$$

but

$$\psi(x - \delta x) = \psi(x) - \sum_{i=1}^{3} \delta x_i \frac{\partial}{\partial x_i}\psi(x)$$

$$= [1 + \sum_{i,j,k=1}^{3} \epsilon_{ijk}\delta\alpha_j x_k \frac{\partial}{\partial x_i}]\psi(x) = [1 + \sum_{i=1}^{3} \delta\alpha_i L_i]\psi(x),$$

where

$$L_i = \sum_{j,k=1}^{3} \epsilon_{ijk} x_j \frac{\partial}{\partial x_k}. \tag{1.33}$$

[3]Strictly speaking we should say that their algebras are related by Lie algebra isomorphisms. This concept is introduced in Section 3.3.

These operators are called the *generators of infinitesimal rotations.* If the coordinate system is rotated through an infinitesimal angle $\delta\alpha_i$, then ψ is changed by an amount

$$\delta\psi(x) = [P(\delta\alpha_i) - 1]\psi(x) = \delta\alpha_i L_i \psi(x)$$

The commutation relations can easily be computed from (1.33).

$$[L_i, L_j] = -\sum_{k=1}^{3} \epsilon_{ijk} L_k \tag{1.34}$$

They are identical with (1.31) in spite of the fact that the X_i's are constant matrices and the L_i's are differential operators. Thus the group of rotation matrices, Example 1.4(c), and the group of coordinate rotation operators, Example 1.4(d), have the same abstract algebra, (1.31) and (1.34).

The quantum mechanical version of (1.33) is obtained by introducing the momentum operator

$$p_k = \frac{\hbar}{\imath} \frac{\partial}{\partial x_k}.$$

Then

$$\boldsymbol{L} = \frac{\imath}{\hbar} \boldsymbol{x} \times \boldsymbol{p}.$$

The term $\boldsymbol{x} \times \boldsymbol{p}$ has the same commutation relations as $\boldsymbol{J}$ in (1.32).

1.5 Reconstructing the Group

We have clained (so far without proof) that every group has associated with it a unique Lie algebra and that all groups related to it by an isomorphism have the same Lie algebra. The converse is not quite true. It is possible for different (*ie.* not related by an isomorphism) groups to have the same Lie algebra, but in this case the groups are identical in the vicinity of the identity. Every Lie algebra thus defines a unique group. If a matrix representation of the Lie algebra is available, then it is possible to compute all the group elements using the following simple procedure.

Suppose we wish to calculate the rotation matrix $M(\boldsymbol{\alpha})$ corresponding to a fixed vector $\boldsymbol{\alpha}$. Using the dimensionless variable t, $-\infty < t < +\infty$, define $M(t) = M(\boldsymbol{\alpha}t)$, so that $M(0) = I$ and $M(1) = M(\boldsymbol{\alpha})$. It is possible to find a matrix function of t with the simple property that

$$M(s)M(t) = M(s + t) \tag{1.35}$$

where s is another scalar variable like t. Differentiating this expression with respect to s and setting $s = 0$, we obtain the matrix equation

$$\frac{dM(t)}{dt} = M(t)X \tag{1.36}$$

where

$$X = \frac{d}{dt}M(t)\Big|_{t=0} = \sum_{i=1}^{3} \alpha_i X_i \tag{1.37}$$

This equation can be satisfied with the usual power series expansion

$$M(t) = I + \sum_{m=1}^{\infty} (tX)^m/m! \tag{1.38}$$

So long as the series converges it can be used to define a matrix function called the *exponential of a matrix*,

$$M(\boldsymbol{\alpha}) = \exp(X) = I + \sum_{m=1}^{\infty} (\sum_{i=1}^{3} \alpha_i X_i)^m/m! \tag{1.39}$$

where we have set $t = 1$ to obtain the final form of the matrix.

This equation justifies the power series expansion 1.22 used in section 1.3. The formula for the inverse of M is also a trivial consequence of 1.37. Since $I = M(0) = M(t)M(t)^{-1} = M(t-t) = M(t)M(-t)$, it follows that $M(\boldsymbol{\alpha})^{-1} = M(-\boldsymbol{\alpha})$.

A set of group elements like $M(t)$ that have the property 1.34 clearly satisfies the requirements of a group. Such a set of elements is called a *one-parameter subgroup*. For a certain class of groups (the compact groups), every group element is a member of a one-parameter subgroup, which can be obtained by exponentiating an element of the Lie algebra as in 1.37.

Example 1.6 *Rotations about a single axis.*

In the case of the rotation group it is easy to compute X_i^n, and equation 1.37 can be summed to give the matrices for rotation about a single axis in terms of a power series expansion. For example, to obtain a rotation about the x-axis note that

$$X_1^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$X_1^3 = -X_1$$

so that

$$M_{22} = M_{33} = \sum_{m\ even}^{\infty} \alpha_1^m (-1)^{m/2} = \cos\alpha_1$$

$$M_{23} = -M_{32} = \sum_{m\ odd}^{\infty} \alpha_1^m (-1)^{(m+1)/2} = -\sin\alpha_1$$

and finally

$$M(\alpha_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_1 & -\sin\alpha_1 \\ 0 & \sin\alpha_1 & \cos\alpha_1 \end{pmatrix}.$$

Unfortunately, (1.38) will not reproduce the rotation matrix (1.22) for an arbitrary $\boldsymbol{\alpha}$. The problem is that (1.22) was derived by assuming a sequence of three rotations about different axes. Equation (1.38), on the other hand, assumes that all three rotations are made simultaneously as t increases from 0 to 1. Since rotations about different axes do not commute, these two procedures give different rotation matrices. Whenever a formula like (1.38) is used to parameterize a group by exponentiating a member of the Lie algebra, the resulting formula is called a *canonical parameterization.* There are often good physical reasons for choosing a non-canonical parameterization, however. See for example, the comments regarding (1.16) at the end of Section1.1.

Chapter 2

Representations

We have seen in the previous chapter how a group of operations gives rise to a family of isomorphic or homomorphic groups consisting of different kinds of mathematical objects but sharing a common multiplication table. Within this family there is an especially important sub-family, the matrix groups. These groups occupy a prominent place in group theory, in part because they are familiar and tractable objects, and in part because of a set of deep relationships among Lie groups, Lie algebras and matrices.

Consider an abstract group $\mathcal{G}$ of elements $a, b, c, ...$ satisfying the axioms in Definition 1.1. If to each element a can be assigned a non-singular $n \times n$ matrix $M(a)$ such that

$$M(ab) = M(a)M(b) \tag{2.1}$$

for every pair of elements a and $b \in \mathcal{G}$, then this set of matrices is said to provide a n-dimension *representation* M of $\mathcal{G}$. In (2.1) a and b are combined using the group multiplication law. $M(a)$ and $M(b)$ are combined with ordinary matrix multiplication.

This definition can be made more precise using the definition of mapping.

Definition 2.1 *Representation of a Group $\mathcal{G}$*

If there exists a homomorphic mapping of a group $\mathcal{G}$ onto a group of non-singular $n \times n$ matrices $M(a)$, then this group of matrices forms a n-dimensional representation M of $\mathcal{G}$.

If there is a one-to-one mapping between group elements and matrices, *ie.* if the mapping is isomorphic, then the representation is said to be *faithful*.

Example 2.1 *The rotation matrices.*

The set of 3×3 unimodular orthogonal matrices constitute a representaiton of the rotation group since

$$M(R_1 R_2) = M(R_1)M(R_2)$$

Within the appropriate range of angles the mapping is one-to-one, so the representation is faithful.

It is easy to see why the matrix groups have their privileged position. It is, for example, much easier to multiply two rotation matrices together than it is to compute the results of two successive rotations directly using the multiplication table, equation (1.1). Moreover, the rotation *operators*, Example 4(d), Chapter 1, require the rotation matrices as part of their definition. There is an important theorem due to Ado that every (finite dimensional) Lie algebra has a faithful matrix representation. We have already seen how the corresponding group can be obtained by the process of matrix exponentiation. Consequently we can concentrate on matrices without restricting any of the possibilities inherent in the abstract definition of the group.

In the next section we discuss the classification scheme that defines the classical matrix groups. In the following section we derive the basic theorems of matrix exponentiation. Following this we return to the theme of matrices as group representations.

2.1 The Classical Matrix Groups

The set of non-singular $n \times n$ matrices constitutes a group. If there are no other restrictions placed on the matrices the group is call the *general linear group*, abbreviated Gl(n,R) or Gl(n,C) depending on whether the matrix elements are real or complex numbers. In the case of Gl(n,R) there are n^2 numbers that must be specified to completely determine the matrix. Gl(n,C) requires $2n^2$ real numbers to complete the specification.

Physical applications often place some restrictions on the matrix parameters. In order to formulate these constraints it is necessary to generalize the idea of coordinate transformations given in Section 1.1. Consider a n-dimensional coordinate system with n linearly independent basis vectors $[\mathbf{e}_1, \mathbf{e}_2, ...\mathbf{e}_n]$. An arbitrary vector can be expressed as a linear combination of these basis vectors.

$$\mathbf{v} = \sum_{i=1}^{n} v^i \mathbf{e}_i \tag{2.2}$$

(Note the placement of the indices.) The n numbers, v^i, are called the *coordinates* of $\mathbf{v}$ with respect to the coordinate system of the $\mathbf{e}$'s. If we choose a different coordinate system with basis vectors related to the original basis by

$$\mathbf{e}'_j = \sum_{i=1}^{n} M_j{}^i \mathbf{e}_i \tag{2.3}$$

Then in this new coordinate system

$$\mathbf{v} = \sum_{j=1}^{n} v'^j \mathbf{e}'_j = \sum_{i,j=1}^{n} v'^j M_j{}^i \mathbf{e}_i = \sum_{i=1}^{n} v^i \mathbf{e}_i$$

Since the $\mathbf{e}_i$'s are independent

$$\sum_{j=1}^{n} v'^j M_j{}^i = v^i \tag{2.4}$$

In matrix notation

$$\mathbf{e}' = M\mathbf{e} \tag{2.5}$$
$$v^T = v'^T M \tag{2.6}$$
$$v'^T = v^T M^{-1} \tag{2.7}$$

We often need to combine two vectors to make a single number called the *scalar product* of the vectors. Symbolically

$$(\mathbf{u}, \mathbf{v}) = f \tag{2.8}$$

where f is a real or complex number depending on the definition of the vector space. The properties of the scalar product must be added as an additional axiom. It is common to require that

$$(\mathbf{u}, \alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha(\mathbf{u}, \mathbf{v}_1) + \beta(\mathbf{u}, \mathbf{v}_2) \tag{2.9}$$

and

$$(\alpha\mathbf{u}_1 + \beta\mathbf{u}_2, \mathbf{v}) = \alpha(\mathbf{u}_1, \mathbf{v}) + \beta(\mathbf{u}_2, \mathbf{v}) \tag{2.10}$$

or

$$(\alpha\mathbf{u}_1 + \beta\mathbf{u}_2, \mathbf{v}) = \alpha^*(\mathbf{u}_1, \mathbf{v}) + \beta^*(\mathbf{u}_2, \mathbf{v}) \tag{2.11}$$

The first choice, (2.10), is said to be *bilinear*, the second choice, (2.11), *sesquilinear*. Of course, if the vector space is defined on the real number field, there is no distinction.

The scalar product of the basis vectors

$$(\mathbf{e}_i, \mathbf{e}_j) = g_{ij} \tag{2.12}$$

is called the *metric tensor*. It is a $n \times n$ matrix whose properties depend on the coordinate system. Once it is known, an arbitrary scalar product can be computed in terms of the components of the vectors.

$$(\mathbf{u}, \mathbf{v}) = \sum_{i,j=1}^{n} (u^i \mathbf{e}_i, v^j \mathbf{e}_j) = \sum_{i,j=1}^{n} u^{i(*)} g_{ij} v^j \tag{2.13}$$

The last summation is often called a *bilinear form* (without the asterisk) or a *sesquilinear form* (with the asterisk) depending on the definition of the original scalar product.

The transformation of g_{ij} under a change of basis is

$$g'_{ij} = (\mathbf{e'}_i, \mathbf{e'}_j) = \sum_{k,l} (M_i{}^k \mathbf{e}_k, M_j{}^l \mathbf{e}_l) = \sum_{k,l} M_i^{(*)\,k} g_{kl} M_j{}^l \tag{2.14}$$

The vectors $\mathbf{v}$ and $\mathbf{u}$ are not changed by this transformation, so the scalar product is also invariant. If in addition

$$g_{ij} = g'_{ij} \tag{2.15}$$

the matrix M is said to induce a *metric preserving transformation*. It is easy to prove that the set of all $n \times n$ non-singular matrices that preserve a specific metric tensor forms a group. Such groups are called *metric preserving groups*; and there is a rich "taxonomy" of these groups based on the metrics they preserve.

Most of the scalar products encountered in physics have some specific symmetry with respect to the interchange of the two vectors. The natural symmetry for a bilinear scalar product is

$$(\mathbf{u}, \mathbf{v}) = \pm(\mathbf{v}, \mathbf{u}) \quad g_{ij} = \pm g_{ji} \tag{2.16}$$

and for a sesquilinear scalar product the corresponding symmetry is

$$(\mathbf{u}, \mathbf{v}) = \pm(\mathbf{v}, \mathbf{u})^* \quad g_{ij} = \pm g^*_{ji} \tag{2.17}$$

Metrics that have one of these four symmetry properties can often be put in *cannonical form* in which the corresponding bilinear form has a simple structure. This is accomplished by a second transformation like (2.14).

$$\tilde{g}_{ij} = \sum_{k,l} S_i^{(*)\,k} g_{kl} S_j{}^l \tag{2.18}$$

Suppose the S leaves $\tilde{g}$ in canonical form. Then we may define a new set of matrices $\tilde{M}$

$$\tilde{M}_i{}^j = \sum_{k,l} S_i{}^k M_k{}^l (S^{-1})_l{}^j \tag{2.19}$$

So that

$$\tilde{g}'_{ij} = \sum_{k,l} \tilde{M}_i^{(*)\,k} \tilde{g}_{kl} \tilde{M}_j{}^l$$

$$= \sum_{klmnpqrt} [S_i^{(*)\,m} M_m^{(*)\,n} (S^{-1})_n^{(*)\,k}] [S_k^{(*)\,p} g_{pq} S_l{}^q] [S_j{}^r M_r{}^t (S^{-1})_t{}^l]$$

$$= \sum_{mrnt} S_i^{(*)\,m} (M_m^{(*)\,n} g_{nt} M_r{}^t) S_j{}^r \tag{2.20}$$

This shows that if M belongs to the group of matrices that preserve the metric, so that $g_{ij} = g'_{ij}$ as defined by (2.14), then $\tilde{M}$ preserves the metric $\tilde{g}$, *ie.*, $\tilde{g}'_{ij} = \tilde{g}_{ij}$. Furthermore, if the M's are members of a metric-preserving group, then the $\tilde{M}$'s form a metric-preserving group with the same multiplication table. (See the discussion of equivalent representations Section 2.4).

$$M(a)M(b) = M(ab)$$
$$\tilde{M}(a)\tilde{M}(b) = \tilde{M}(ab)$$

Thus a great simplification is achieved and nothing is lost by putting g_{ij} in canonical form. Notice that no special assumptions have been made about the matrix S that accomplishes this transformation except that it is non-singular. Whether or not a canonical form exists depends on the symmetry of g. We will consider the four symmetry classes individually.

(1) $g_{ij} = g^*_{ji}$ (g_{ij} could be either real or complex.)

This is by far the most common case. The metric tensor g_{ij} is a real symmetric or Hermitian matrix. It is well known that such a matrix can be diagonalized and that its eigenvalues are all real. It is always possible to find a transformation

$$\mathbf{e}'_j = \sum_i S_j{}^i \mathbf{e}_i$$

that brings g_{ij} into diagonal form.

$$\tilde{g}_{ij} = \sum_{k,l} S_i^{(*)\,k} g_{kl} S_j{}^l = \sum_i \lambda_i \delta_{ij}$$

Since the λ_i's are real it is possible to renormalize the basis vectors

$$\mathbf{e}''_i = \frac{\mathbf{e}'_i}{\sqrt{|\lambda_i|}}$$

so that

$$g''_{ij} = (\mathbf{e}''_i, \mathbf{e}''_j)$$

is a diagonal matrix whose elements are $+1, -1$, or 0. If the metric is non-singular (the usual case) zero is not allowed. This brings us to the canonical form of a Hermitian metric.

$$g_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ +1 & \text{if } 1 \leq i \leq N_+ \\ -1 & \text{if } N_+ < i \leq N_+ + N_- \end{cases} \tag{2.21}$$

where N_+ and N_- are the number of $+1$'s and -1's along the diagonal. Since no similarity transformation can change N_+ and N_-, there are $n+1$ "inequivalent" $n \times n$ Hermitian metrics and $n+1$ inequivalent matrix groups that preserve them.

Metrics of the form (2.21) are often used to "raise and lower indices." For example, the quantity

$$v_i = \sum_j g_{ij} v^j \tag{2.22}$$

transforms like the $\mathbf{e}_i$ in (2.3).

$$\begin{aligned} v'_i = \sum_j g_{ij} v'^j = \sum_{jklm} M_i^{(*)\,k} g_{kl} M_j{}^l v^m (M^{-1})_m^{\ j} \\ = \sum_{km} M_i{}^k \, (g_{km} v^m) \end{aligned}$$

The component arrays like v_i with lower indices are said to transform *covariantly*. Those like v^i with upper indices transform *contravariantly*. One advantage of the notation is that the scalar product can be writtem entirely in terms of the components of the vectors

$$(\mathbf{u}, \mathbf{v}) = \sum_i u_i v^i = \sum_j u^j v_j \tag{2.23}$$

without any reference to the metric. Another advantage is that the derivative with respect to a covariant quantity transforms contravariantly (and vice versa) so that derivatives of the form

$$\frac{\partial}{\partial x^\mu} = \partial_\mu \quad \tfrac{\partial}{\partial x_\nu} = \partial^\nu$$

act like ordinary arrays of vector components so far as their transformation properties are concerned.

Example 2.2 *Space-time coordinates in special relativity.*

Define a contravariant space-time vector x^μ with $\mu = 0, 1, 2, 3$ as follows:

$$x^0 = t \;\; x^1 = x \;\; x^2 = y \;\; x^3 = z \tag{2.24}$$

(The velocity of light, $c = 1$.) The metric tensor is

$$g_{\mu\nu} = \begin{cases} +1 \text{ for } \mu = \nu = 0 \\ -1 \text{ for } \mu = \nu = 1, 2, 3 \\ 0 \text{ for } \mu \neq \nu \end{cases} \tag{2.25}$$

The scalar product of two such vectors is

$$(\mathbf{x}, \mathbf{y}) = \sum_{\mu,\nu=0}^{3} x^\mu g_{\mu\nu} y^\nu = \sum_{\mu}^{3} x^\mu y_\mu$$

$$= x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3$$

Lorentz transformations leave this bilinear form invariant. This is equivalent to the statement that Lorentz transformations are the group of 4×4 matrices that preserve the metric given by (2.25).

For groups that preserve diagonal metrics, (2.14) is a generalization of the familiar unitarity condition $M^T = M^{-1}$. To see this connection we define $g^{ij} = g_{ij}$, so that $\sum_j g^{ij} g_{jk} = \delta^i_k$, the unit matrix. Then

$$\begin{aligned} \delta^m_i = \sum_j g_{ij} g^{jm} = \sum_{k,l,j} (M_i^{(*)\,k} g_{kl} M_j{}^l) g^{jm} \\ \sum_k M^m{}_k M_i^{(*)\,k} = \delta^m_i \\ \sum_k M^m{}_k (M^\dagger)^k{}_i = \delta^m_i \end{aligned} \tag{2.26}$$

For this reason, the group of complex $n \times n$ matrices that preserve the metric of (2.25) is called the *unitary group*, $U(N_+, N_-; C)$. If the matrices are real, the group could be called $U(N_+, N_-; R)$, or more commonly, $O(N_+, N_-)$, the *orthogonal group* in n dimensions.

Example 2.3 *Cartesian coordinate systems.*

The usual Cartesian coordinate system is defined in terms of a set of orthogonal unit vectors, $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. The metric tensor is just the unit

matrix, and the corresponding groups of metric preserving matrices are $U(n)$ and $O(n)$. In these spaces there is no difference between a tensor with upper and lower indices, so this distinction is usually dropped.

(2) $g_{ik} = -g_{ki}$

Taking the determinant of g,

$$\det(g) = \det(g^t) = \det(-g) = (-1)^n \det(g)$$

If n is odd, $\det(g) = 0$, and the metric is singular. A scalar product with this symmetry property can thus only be defined for even-dimensional spaces ($n = 2\nu$, ν integral).

The simplest anti-symmetric metric is

$$I_A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Every non-singular anti-symmetric metric can be transformed into canonical form with I_A's along the diagonal and zeros elsewhere.

$$g' = \begin{pmatrix} I_A & & & \\ & I_A & & \\ & & \cdots & \\ & & & I_A \end{pmatrix} \tag{2.27}$$

The group of $n \times n$ matrices that preserve anti-symmetric metrics is called the *symplectic group* in n dimensions. The notation is Sp(n;C) and Sp(n;R) for complex and real matrices respectively.

The coordinate system in which g_{ij} has the form of (2.27) is defined by the new symplectic coordinate basis

$$\mathbf{e}_1, \ldots, \mathbf{e}_\nu$$
$$\mathbf{e'}_1, \ldots, \mathbf{e'}_\nu \tag{2.28}$$
$$(\mathbf{e}_i, \mathbf{e}_j) = (\mathbf{e'}_i, \mathbf{e'}_j) = 0 \tag{2.29}$$
$$(\mathbf{e}_i, \mathbf{e'}_j) = -(\mathbf{e'}_i, \mathbf{e}_j) = \delta_{ij}$$

In these definitions $\nu = n/2$. The scalar product is

$$(\mathbf{u}, \mathbf{v}) = (u_1 v'_1 - u'_1 v_1) + \ldots + (u_n v'_n - u'_n v_n) \tag{2.30}$$

Example 2.4 *$Sp(2, r)$ and the Area of a Parallelogram.*

Consider two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ defined on a two-dimensional Cartesian coordinate system. The "cross" product $A_xB_y - A_yB_x$ gives the area of the parallelogram shown in Figure (). This area can be thought of as a scalar product with an antisymmetric metric.

$$(\boldsymbol{A}, \boldsymbol{B}) = \begin{bmatrix} A_x & A_y \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} B_x \\ B_y \end{bmatrix} = A^T I_A B$$

The set of real, non-singular, 2×2 matrices that preserve I_A will leave the area of the parallelogram invariant. These matrices constitute the group $Sp(2, R)$. The metric preserving condition, (2.14) and (2.15), in matrix notation is

$$I_A = M^T I_A M$$

This contains only one independent constraint,

$$M_{11}M_{22} - M_{12}M_{21} = 1$$

so that $Sp(2, R)$ is a three-parameter group.

There are many ways to choose the parameters, of course; but the following choice has a simple geometrical interpretation: Define

$$R(\theta) = \begin{bmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{bmatrix}$$

$$S(\lambda) = \begin{bmatrix} e^{\lambda/2} & 0 \\ 0 & e^{-\lambda/2} \end{bmatrix}$$

$$T(\eta) = \begin{bmatrix} \cosh(\eta/2) & \sinh(\eta/2) \\ \sinh(\eta/2) & \cosh(\eta/2) \end{bmatrix}$$

The matrices R, S, and T are linearly independent and symplectic. (The significance of the ubiquitous $1/2$ will become apparent when we discuss the Lie algebra of $Sp(2, R)$.) Clearly $R(\theta)$ is a rotation, $S(\lambda)$ is an expansion/contraction of scale along the x and y axes, and $T(\eta)$ represents an expansion/contraction along a set of axes rotated 45^o with respect to x and y. Any member of $Sp(2, r)$ can be written as a product of these three matrices.

The symplectic transformation can be applied to the coordinates of a point in the x-y plane.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = M \begin{pmatrix} x \\ y \end{pmatrix}$$

This can be thought of as a function the maps the point (x, y) into the point (x', y'). The Jacobian of this transformation is

$$\frac{\partial(x', y')}{\partial(x, y)} = \det(M) = 1$$

Consequently, the area enclosed by *any* closed curve drawn on a flat plane is preserved by this transformation, since

$$\oint dx\, dy = \oint dx'\, dy'$$

by the usual rules for changing variables within an integration.

2.2 The Exponential of a Matrix

The usual way of deriving a matrix group from a matrix representation of a Lie algebra involves exponentiation of a matrix. This technique was used to calculate the $O(3)$ rotation matrices in Section (1.5), and the method is quite general. In this section we first examine the question of convergence and then derive some useful theorems for analyzing matrix exponentials.

Definition 2.2 *The Exponential of a Matrix*
The exponential e^M of a $n \times n$ matrix M is defined by the sum

$$e^M = \sum_{m=0}^{\infty} \frac{M^m}{m!} \tag{2.31}$$

where $M^0 = I$, the $n \times n$ identity matrix.

Each term in this sum is a $n \times n$ matrix, as is the partial sum

$$S_k = \sum_{m=0}^{k} \frac{M^m}{m!} \tag{2.32}$$

Consider the element of S_k corresponding to the ith row and the jth column. We denote the element $S_{k,ij}$. The natural definition of convergence of (2.31) is that the sequence of partial sums converges separately for each pair of i and j. That is,

$$\overset{limit}{k\to\infty} S_{k,ij} = S_{ij} = [e^M]_{ij} \quad \text{for all } i,j = 1,\ldots,n \tag{2.33}$$

Theorem 2.1 *If n is finite and the elements of M are bounded, then the matrix exponential series converges uniformly.*

Proof: Let $N(M)$ denote the largest value of $|M_{ij}|$ for $i,j = 1,\ldots,n$. We assert that $N(M^m) \leq [nN(M)]^m$. This is certainly true for $M = 1$. If it holds for m, then it must hold for $m+1$:

$$N(M^{m+1}) \leq nN(M)N(M^m) \leq [nN(M)]^{m+1}$$

Since $nN(M) = c$ is finite, $\sum_{m=0}^{\infty} c^m/m! = e^c$ certainly converges. Therefore

$$\sum_{m=0}^{\infty} \frac{N(M^m)}{m!} \leq \sum_{m=0}^{\infty} \frac{c^m}{m!} = e^c$$

So the norm of each term of the series is less than the corresponding term of a convergent comparison series. Consequently the series in (2.31) converges uniformly by the Weierstrass test.

The following three theorems are trivial to prove using the definition of the matrix exponential:

Theorem 2.2 *If A and B are commuting matrices, then $e^{A+B} = e^A e^B$.*

Theorem 2.3 *If B is non-singular, then $Be^A B^{-1} = e^{BAB^{-1}}$.*

Theorem 2.4

$$e^{A^*} = (e^A)^* \quad e^{A^T} = (e^A)^T$$
$$e^{A^\dagger} = (e^A)^\dagger \quad e^{-A} = (e^A)^{-1}$$

Suppose the $n \times n$ matrix A has a set of n eigenvalues $\lambda_1, \ldots, \lambda_n$, each eigenvalue repeated a number of times equal to its multiplicity. The λ_j are solutions of the characteristic equation

$$\det(A - \lambda I) = 0$$

It is easy to see that the similarity transformation, $A' = BAB^{-1}$, does not change the eigenvalues.

$$\det(A' - \lambda I) = \det[B(A - \lambda I)B^{-1}]$$

$$= (\det B)\det(A - \lambda I)(\det B^{-1}) = \det(A - \lambda I) = 0$$

This allows us to prove the following theorem.

Theorem 2.5 $\det(e^A) = e^{tr\, A}$

Proof: First assume that A is diagonalizable. Let B be the matrix that diagonalizes A. Then

$$\det(e^A) = \det(Be^A B^{-1}) = \det(e^{BAB^{-1}}) = \det(e^{A'})$$

where A' is a diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$ along the diagonal. Since

$$A'^m = \begin{pmatrix} \lambda_1^m & & & & \\ & \lambda_2^m & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \lambda_n^m \end{pmatrix} \tag{2.34}$$

$$e^{A'} = \begin{pmatrix} e^{\lambda_1} & & & & \\ & e^{\lambda_2} & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & e^{\lambda_n} \end{pmatrix} \tag{2.35}$$

and

$$\det(e^A) = \det(e^{A'}) = \exp(\sum_{i=1}^{n} \lambda_i) = e^{tr\, A'} = e^{tr\, A}$$

The last equality follows because the trace of a matrix is always equal to the sum of its eigenvalues.

Of course not all matrices are diagonalizable; but all matrices can be transformed to *upper triangular form* (Miller, 1972), with the eigenvalues along the main diagonal and zeros below the diagonal. An elementary calculation shows that if A' is in upper diagonal form, A'^m is also in upper diagonal form with the eigenvalues raised to the mth power as in (2.34). The rest of the proof goes through as before.

A by-product of the proof is the following corollary.

Corollary 2.6 *If A is a $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, then the eigenvalues of e^A are $e^{\lambda_1}, \ldots, e^{\lambda_n}$.*

This is clear from (2.35) and the fact that the similarity transformation does not alter the eigenvalues.

If A and B are non-commuting matrices it is no longer true that $e^A e^B = e^{A+B}$. It is still possible to write $e^A e^B = e^C$, however, where C is another $n \times n$ matrix. There is a closed-form expression for C called the Campbell-Baker-Hausdorff formula, which we will state without proof along with an important lemma used in its derivation. The reader is referred to (Miller, 1972) for the complete proof.

Definition 2.3 *The linear transformation Ad A is defined by*

$$Ad\ A(B) = [A, B] \tag{2.36}$$

This is a linear operator that transforms B into another $n \times n$ matrix. By $(Ad\ A)^m$ we mean the operator

$$(Ad\ A)^m(B) = [A, [A, \cdots [A, B] \cdots]$$

The expression on the right contains m nested commutators.

Lemma 2.7

$$e^A B e^{-A} = (\exp(Ad\ A))B = \sum_{j=0}^{\infty} (Ad\ A)^j(B)/j! \tag{2.37}$$

Theorem 2.8 *For A, B in a sufficiently small neighborhood of the identity, $e^C = e^A e^B$, and*

$$C = B + \int_0^1 g[\exp(t\ Ad\ A)\exp(Ad\ B)](A)dt, \tag{2.38}$$

where

$$g(z) = \frac{\ln z}{z-1} = \sum_{j=0}^{\infty} \frac{(1-z)^j}{j+1}$$

The lowest order terms are

$$C = A + B + \frac{1}{2}[A,B] + \frac{1}{12}[A,[A,B]] - \frac{1}{12}[B,[B,A]] + \cdots, \tag{2.39}$$

and the matrix elements of C are analytic functions of the matrix elements of A and B.

The phrase "sufficiently small neighborhood of the identity," implies some definition of the distance between two matrices (cf. section 3.1). One possibility is the Euclidean distance function,

$$d(A,B) = \left[\sum_{i,j=1}^{n} (A_{ij} - B_{ij})^2\right]^{1/2}. \tag{2.40}$$

The theorem can be restated in terms of this function as follows: there exists an $\epsilon > 0$ such that for all matrices A, B with $d(A,I) < \epsilon$ and $d(B,I) < \epsilon$ the series implied by 2.38 converges to C and defines an analytic matrix valued function of A and C.

Equation (2.38) is truly a formidable formula: an integral over an infinite series, each term of which is an infinite series of nested commutators. It is not of much use in practical calculations except in special cases where the series (2.39) terminates after a few terms. The Campbell-Baker-Hausdorff theorem is of great theoretical significance, however, for at least two reasons. First, the Lie algebra is determined by the structure of the group in the immediate vicinity of the identity, and in this region the series expansion 2.39 converges rapidly. Second, the theorem shows that analyticity, which is part of the definition of Lie groups, survives exponentiation and multiplication.

2.3 Representations in Quantum Mechanics

The matrices in Section (2.1) were defined in terms of their action on the components of a vector. The vector $\boldsymbol{v}$ in (2.2) acts like a position vector in n-dimensional space. The matrix M has dimension $n \times n$; and (2.3) represents a homogeneous coordinate transformation. A group of matrices defined in this way constitute the *defining representation* of the group. For example, the set of all real, orthogonal, 3×3 matrices is the defining representation of the group $O(3)$. The defining representation is a special case, however. The definition of a representation leaves open the possibility that a group might have many different representations in terms of matrices of different dimensionality. This may seem artificial, but such representations arise naturally in quantum mechanics, Consider an operator $\mathcal{O}(a)$ corresponding to a group element a operating on a state vector $|A>$.

$$\mathcal{O}(a)|A> = |A'> \tag{2.41}$$

In order to evaluate this expression it is customary to expand the state in terms of a complete set of basis functions $|\psi_i>$.

$$|A> = \sum_{i=1}^{n} a_i|\psi_i> \tag{2.42}$$

By hypothesis the transformed state can also be so expanded,

$$|A'> = \sum_{i=1}^{n} a'_i|\psi_i> \tag{2.43}$$

so that the effect of the operator on the basis states is that of a matrix

$$\mathcal{O}(a)|\psi_i> = \sum_{j=1}^{n} |\psi_j> M(a)_{ji} \tag{2.44}$$

with the implicit definition that

$$\mathcal{O}(a) \sum_{i=1}^{n} a_i|\psi_i> = \sum_{i=1}^{n} a_i \mathcal{O}(a)|\psi_i> \tag{2.45}$$

In other words, $\mathcal{O}$ is a linear operator. If $\mathcal{O}$ satisfies the group postulate, $\mathcal{O}(ba) = \mathcal{O}(b)\mathcal{O}(a)$, then

$$\mathcal{O}(b)\mathcal{O}(a)|\psi_i> = \mathcal{O}(b) \sum_{j} |\psi_j> M(a)_{ji}$$

$$= \sum_{jk} |\psi_k > M(b)_{kj} M(a)_{ji} = \sum_k |\psi_k > M(ba)_{ki}$$

So that M also obeys the group law,

$$M(ba) = M(b)M(a), \tag{2.46}$$

and M is a n-dimensional representation of the group containing a and b. The odd arrangement of rows and columns in the definition (2.44) is required to make the matrices come out in the right order in (2.46). Unlike the rotation matrices in Section (1.1) which operate on the components of a position vector, these matrices operate on the elements of a vector space spanned by the basis vectors $|\psi_i >$. We call this space the *representation space* or *carrier space*. Its dimensionality, *ie.* the number of independent vectors required to span the space, is equal to the size of the matrices M.

Usually the basis states are chosen to be orthonormal so that

$$M(a)_{ji} = < \psi_j | \mathcal{O}(a) | \psi_i >. \tag{2.47}$$

Often the states are eigenstates of some other Hermitian operator, say $\mathcal{P}$, corresponding to a physical observable.

$$\mathcal{P}|\psi_i > = p_i |\psi_i >, \tag{2.48}$$

where p_i is the eigenvalue of the operator $\mathcal{P}$ corresponding to the ith eigenvector. In the simplest case $\mathcal{P}$ and $\mathcal{O}$ commute, $[\mathcal{P}, \mathcal{O}] = 0$, so that

$$\begin{aligned} \mathcal{P}\mathcal{O}(a)|\psi_i > &= \sum_{j=1}^{n} p_j |\psi_j > M(a)_{ji} \\ &= \mathcal{O}\mathcal{P}|\psi_i > = p_i \sum_{j=1}^{n} |\psi_j > M(a)_{ji} \end{aligned} \tag{2.49}$$

This can be satisfied if $p_i = p_j$. In this case the basis vectors used to define the representation all have the same eigenvalue of $\mathcal{P}$, which can be used to label the representation. (Equation (2.49) can also be satisfied if some of the elements of M_{ij} are permanently zero. This makes M a reducible representation. See Section (2.5)). The physical significance of (2.49) is that the observable quantity $\mathcal{P}$ is invariant under the operation of $\mathcal{O}$; so that $\mathcal{O}$ does not mix states corresponding to different values of p; p is a "good quantum number."

In the usual theory of angular momentum, the operator $\boldsymbol{L}^2 = L_x^2 + L_y^2 + L_z^2$ commutes with L_x, L_y, L_z and hence with the rotation operator. Thus $\boldsymbol{L}^2$ plays the role of the operator $\mathcal{P}$ in (2.48). Its eigenvalues are defined by

$$\boldsymbol{L}^2|\psi_l> = l(l+1)|\psi_l>$$

with $l = 0, 1, 2, \ldots$. There are $2l + 1$ linearly independent eigenvectors for each allowed value of l. These eigenvectors can then be used as basis vectors for a $(2l + 1)$-dimensional representation of the rotation operator. There are an infinite number of these representations, each one labeled with the integer l. If a physical state is an eigenstate of $\boldsymbol{L}^2$ corresponding to angular momentum l, then it can be expanded in terms of the appropriate $2l + 1$ eigenfunctions. The effect of a rotation can then be computed using (2.44). Of course it is possible for the original state to be a mixture of states of different l. This situation is discussed in Section (2.5).

This example is far from trivial. It was necessary to know the appropriate commuting operator (or operators) to play the role of $\mathcal{P}$, and it was necessary to construct the set of states $|\psi_l>$ corresponding to a given eigenvalue. We know how to do these things for the rotation group, because thay are done in every quantum mechanics textbook. Group theory provides a general algorithm for carrying out this program for a large set of physically interesting groups (the compact, semi-simple Lie groups). This algorithn depends on a careful study of the Lie algebra, which is the subject of Chapters () and ().

2.4 Equivalent Representations

The representation M defined in (2.44) depends on the choice of the basis functions $|\psi_i>$. There are many equivalent ways of choosing this set, however. An alternative choice is given by

$$|\psi_i'> = \sum_{j=1}^{n} S_{ji}|\psi_j> \tag{2.50}$$

where S represents any $n \times n$ non-singular matrix. The new set of basis vectors $|\psi_i'>$ spans the same vector space as the original set. The new representation M' is found as follows:

$$\mathcal{O}(a)|\psi_i'> = \sum_{m=1}^{n} M'(a)_{mi}|\psi_m'> = \sum_{k=1}^{n} S_{ki}\mathcal{O}(a)|\psi_k>$$

$$= \sum_{k,l=1}^{n} S_{ki}M(a)_{lk}|\psi_l> = \sum_{k,l,m=1}^{n} S_{ki}M(a)_{lk}(S^{-1})_{ml}|\psi_m'>.$$

In matrix notation

$$M'(a) = S^{-1}M(a)S \tag{2.51}$$

These new matrices follow the same multiplication rules as the original matrices,

$$M'(b)M'(a) = S^{-1}M(b)SS^{-1}M(a)S = S^{-1}M(ba)S = M'(ba) \tag{2.52}$$

so the M''s are also a representation. The transformed matrices are said to form an *equivalent representation*, and the transformation in (2.51) is called a *similarity transformation.*

There is an important class of operators that preserve the scalar product. Using the notation of (2.41)

$$< A'|A' >= < A|A >$$

or

$$\mathcal{O}(a)^\dagger = \mathcal{O}(a)^{-1}.$$

Such operators are said to be *unitary.* Representations of unitary operators are not necessarily unitary themselves because of the general nature of the similarity transform matrix S; however, under some circumstances it is possible to find a representation of a group such that

$$M(a)^\dagger = M^{-1}(a)$$

for all elements in the group. Such a representation is termed a *unitary representation*.

The question of whether a group posseses unitary representations can be answered on the basis of some general group properties that will be discussed in Chapter (). It is easy to show, however, that if the $|\psi_i>$'s in (2.24) are a finite set of orthonormal basis vectors, then the representation defined in (2.44) is unitary if and only if the operator $\mathcal{O}(a)$ is unitary.

$$< \psi_j|\mathcal{O}(a)^\dagger \mathcal{O}(a)|\psi_i > = < \psi_j|\psi_i >= \delta_{ji}$$
$$= \sum_{l,k=1}^{n} < \psi_l|M^*_{lj}M_{ki}|\psi_k > = \sum_{k=1}^{n} M^*_{kj}M_{ki}$$

So that in matrix notation, $M(a)^\dagger M(a) = I$.

Even if the basis vectors $|\psi_j>$ were not orthonormal they could still be used to construct an orthonormal set using the Gram-Schmidt orthogonalization procedure. The matrix that transforms the original basis to the orthonormal basis is exactly the S in (2.50) that induces the similarity transform to a unitary representation.

These arguments assume that the action of the group can be represented in terms of a finite set of basis vectors. This is not always possible as will be seen in Chapter ().

2.5 Reducible and Irreducible Representation

Suppose that the n-dimensional representation M of a group can be partitioned so that the lower left corner is a zero matrix for all elements of the group.

$$M(a) = \begin{bmatrix} M(a)_{11} & M(a)_{12} \\ 0 & M(a)_{22} \end{bmatrix} \tag{2.53}$$

This property replicates itself in matrix multiplication.

$$M(b)M(a) = \begin{bmatrix} M(b)_{11}M(a)_{11} & M(b)_{11}M(a)_{12} + M(b)_{12}M(a)_{22} \\ 0 & M(b)_{22}M(a)_{22} \end{bmatrix} \tag{2.54}$$

So that

$$M(ba)_{11} = M(b)_{11}M(a)_{11}$$

$$M(ba)_{22} = M(b)_{22}M(a)_{22}$$

This means that M_{11} and M_{22} are both representations of the same group. The partitioning has decomposed the n-dimensional representation into two representations of smaller dimensionality. This partitioning is not preserved by the similarity transformations, but it is pointless to make a fundamental distinction between equivalent representations. This leads to the following definition.

Definition 2.4 *Reducible Representation*

A representation is reducible *if it is equivalent to a representation that can be partitioned in the form of (2.53). A representation that is not reducible is said to be* irreducible.

The representations M_{11} and M_{22} might also be reducible. In this case M can be similarity transformed to a representation in which M_{11} and/or M_{22} have the form of (2.53). This sequence of operations can be continued until the transformed representation has the form

$$M' = S^{-1}MS = \begin{bmatrix} M'_{11} & M'_{12} & M'_{13} & \dots & M'_{1n} \\ 0 & M'_{22} & M'_{23} & \dots & M'_{2n} \\ 0 & 0 & M'_{33} & \dots & M'_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & M'_{nn} \end{bmatrix} \tag{2.55}$$

where all the sub-matrices along the diagonal are irreducible.

In many cases a representation that is reducible to the form of (2.55) can be further transformed so that all the submatrices above the diagonal are also zero. Then M is equivalent to a representation of the form

$$M'' = \begin{bmatrix} M''_{11} & 0 & 0 & \dots & 0 \\ 0 & M''_{22} & 0 & \dots & 0 \\ 0 & 0 & M''_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & M''_{rr} \end{bmatrix} \tag{2.56}$$

where the M''_{ii} are irreducible. A representation that can be brought into this form by a similarity transformation is said to be *completely reducible.*

It can be shown in most cases that a reducible representation is also completely reducible. (See Cornwell for a discussion of this point.) This is true, for example, of all unitary representations. In fact, representations for which this is *not* true tend to be mathematical curiosities. For this reason there is a tendency in physics to use the term "reducible" where we would use the term "completely reducible." Nonetheless, there are situations where the distinction is important, and we will adhere to it.

The phenomenon of reducibility is easy to understand in terms of the basis vectors introduced in Section (2.3). Let $|\psi_i>,\ i = 1 \dots n$, be a set of n linearly independent basis states in terms of which the operators $\mathcal{O}(a)$ can be represented by $M(a)$

$$\mathcal{O}(a)|\psi_i> = \sum_{j=1}^{n} M(a)_{ji}|\psi_j> \tag{2.57}$$

If $M(a)$ has the form of (2.53), then for $i = 1 \dots s_1$,

$$\mathcal{O}(a)|\psi_i> = \sum_{j=1}^{s_1} M_{11}(a)_{ji}|\psi_j> \tag{2.58}$$

for all a in the group. Thus the first s_1 basis vectors constitute an invariant subspace in the sense that any operator acting on this space creates another vector in the same space. So long as $M_{12}(a)$ is non-zero, however, the remaining basis states, $|\psi_i>,\ i = s_1 + 1, \dots, n$, are not invariant. When the representation is transformed to its completely reduced form, (2.56), the representation space breaks up into a sum of invariant subspaces, each subspace corresponding to one of the submatrices on the diagonal. We say that the representation space has been *decomposed* into a *direct sum* of

invariant subspaces. If the space spanned by the n basis vectors is denoted by V, then the decomposition is represented as follows

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_r \tag{2.59}$$

where V_1 is the space spanned by the first s_1 basis vectors, s_1 being the dimension of M''_{11} in (2.56). V_2 is the subspace on which M''_{22} operates, *etc.* The direct sum symbol $\oplus$ can also be used to denote the decomposition of the representation. For example, the equation

$$M'' = M''_{11} \oplus M''_{22} \oplus \ldots \oplus M''_{rr} \tag{2.60}$$

is short for (2.56).

The irreducible representations are the basic building blocks of group theory; they also play an important role in physics. For example, the state vectors that describe single-particle states in quantum mechanics usually span invariant subspaces V_i corresponding to simple irreducible representation of various symmetry groups.

Example 2.5 *Particles as irreducible representations of the rotation group.*

It is universally assumed that the laws of physics are invariant under ordinary rotations. The rotation group is thus a *symmetry group.* In the language of quantum mechanics, the rotation generators J_x, J_y, J_z commute with the Hamiltonian, $[\mathcal{H}, J_i] = 0$. Each single-particle state is an eigenstate of the total angular momentum, $\boldsymbol{J}^2$ with eigenvalue $j(j+1)$. To each allowed value of j there corresponds a $(2j+1)$-dimensional irreducible representation of the rotation group. Each particle with spin j can be described by a $(2j+1)$-component wave function $|\psi_m>$, where m ranges from $-j$ to $+j$ corresponding to the eigenstates of J_z. The space spanned by $|\psi_m>$ corresponding to a single value of j is the invariant subspace V_j on which the $(2j+1)$-dimensional irreducible representation on the rotation group operates.

It is useful to have some criterion for determining whether or not a representaion is irreducible. One test is provided by the following famous theorem, usually called "Schur's Lemma."

Theorem 2.9 *If $M(a)$ is a n-dimensional irreducible represenation of a group and A is a matrix that commutes with $M(a)$ for all elements a in the group, then A is a multiple of the unit matrix.*

Proof: Let V be the carrier space of M. This means that any M operating on any element of V produces another element of V. In an obvious

shorthand notation, $MV = V$. We will say that A operating on V produces an element in the subspace V', $AV = V'$. If $[M, A] = 0$, then

$$(MA)V = (AM)V$$

so

$$MV' = AV = V'.$$

Thus M operating on V' produces another vector in V', *ie.* V' is an invariant subspace of V, which is inconsistent with the assumption that M is irreducible. This contradiction would be avoided if $A = 0$ or if $\det(A) \neq 0$, because then V would be identical to V'. In this case we could still write $A = B + \lambda I$, where $\det(B) = 0$ and $\det(A) = \lambda$. Then $[M, A] = 0$ implies $[M, B] = 0$; and the only remaining way to avoid the contradiction is to admit that $B = 0$, and consequently A is a multiple of the unit matrix.

Chapter 3

Formal Properties of Lie Groups

In Chapter 1 we presented a casual definition of a Lie group as a set of group elements depending on a finite number of parameters through some smooth, differentiable functions. The most precise definition is rather subtle and requires some understanding of topology. In fact, in the realm of pure mathematics, it is often the topological aspects of Lie groups that are of primary interest. We will avoid some of these complications, however, by defining Lie groups in terms of coordinate system "patches." This carries a small disadvantage: it is possible to define Lie groups with a weaker set of assumptions than the one we will use. So far as physical applications are concerned, we give up nothing by making this simplification.

3.1 The Composition Function

In Example 1.4b, we introduced the notion of a set of functions that were analogous to the group multiplication table for finite groups. In the case of the rotation group these functions were the θ_3, ψ_3, and ϕ_3 defined by equation (1.1). In general we consider a group $\mathcal{G}$ to be an infinite set of elements depending on n parameters. We can regard these parameters as forming a vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ defined on the n-dimensional vector space R_n consisting of all real n-tuples. We will call $\alpha_1, \ldots, \alpha_n$ the *components* of $\boldsymbol{\alpha}$. The terminology is useful because the same element a can be parameterized in different ways leading to different components. This is analogous to an ordinary vector, which can have different components when referred to different coordinate systems.

It is tempting to require an isomorphic mapping from $\mathcal{G}$ into R_n, *ie.* to demand that for each element in $\mathcal{G}$ there is a unique vector in R_n. In this way we could replace all the abstract group operations with the more familiar and concrete formalism of vector spaces.[1] Unfortunately, this pushes the connection too far; but it is instructive to proceed in this direction and see where the idea breaks down.

Let us assume then that for each element in $\mathcal{G}$ there is a well defined vector in R_n. This set of vectors is itself a group, which we call G_n. Because of the group multiplication rule $ba = c$, there must exist a set of n functions $\boldsymbol{f} = (f_1, \ldots, f_n)$ of the corresponding parameter vectors such that $\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \boldsymbol{\gamma}$. Here $\boldsymbol{\alpha}$ stands for that set of n parameters that specifies the element a, $\boldsymbol{\beta}$ specifies b, and $\boldsymbol{\gamma}$ specifies c. Thus the group operation $ba = c$ has its "image" on G_n in the form of a purely algebraic operation involving functions of $2n$ variables. Again referring to the rotation group, $\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \boldsymbol{\gamma}$ is a condensed form of equation (1.1) with

$$\begin{aligned} (\theta_1, \psi_1, \phi_1) &\to \boldsymbol{\alpha} \\ (\theta_2, \psi_2, \phi_2) &\to \boldsymbol{\beta} \\ (\theta_3, \psi_3, \phi_3) &\to \boldsymbol{\gamma} = \boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha}) \end{aligned}$$

$\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha})$ is called the *composition function.*

The group postulates impose a set of general requirements on $\boldsymbol{f}$.

Definition 3.1 *The Composition Function*

(1) The composition function $\boldsymbol{f}$ is an n-tuple defined on R_n. If a,b,c are elements in $\mathcal{G}$ with $ba = c$ then the corresponding vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ in R_n must satisfy $\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \boldsymbol{\gamma}$.

(2) If $\boldsymbol{\alpha}$,$\boldsymbol{\beta}$,$\boldsymbol{\delta}$ are elements in R_n corresponding to elements in $\mathcal{G}$, then $\boldsymbol{f}(\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\beta}), \boldsymbol{\delta}) = \boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\delta}))$.

(3) There exists an identity element $\boldsymbol{e}$ in R_n such that $\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{e}) = \boldsymbol{f}(\boldsymbol{e}, \boldsymbol{\alpha}) = \boldsymbol{\alpha}$

(4) For every $a \in \mathcal{G}$ there is a corresponding $\boldsymbol{\alpha} \in G_n$ with an inverse $\boldsymbol{\omega}$ such that $\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\omega}) = \boldsymbol{f}(\boldsymbol{\omega}, \boldsymbol{\alpha}) = \boldsymbol{e}$. Moreover, $\boldsymbol{\omega}$ is the homomorphic image of a^{-1}.

These requirements are easily satisfied in the following example.

Example 3.1 *The group $Gl(2, R)$*

[1] A mapping into an analytic structure like R_n that can be written down concretely and described analytically is called a *realization.*

The group $Gl(2,R)$ consists of all 2×2 non-singular matrices with real matrix elements. Use the parameterization

$$M(\boldsymbol{\alpha}) = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix}$$

The group G_n (in this case G_4) is the set of all 4-tuples such that $\alpha_1\alpha_4 \neq \alpha_2\alpha_3$. Since

$$M(\boldsymbol{\beta})M(\boldsymbol{\alpha}) = \begin{bmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{bmatrix} \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix}$$

$$= \begin{bmatrix} \beta_1\alpha_1 + \beta_2\alpha_3 & \beta_1\alpha_2 + \beta_2\alpha_4 \\ \beta_3\alpha_1 + \beta_4\alpha_3 & \beta_3\alpha_2 + \beta_4\alpha_4 \end{bmatrix}$$

the components of the composition function $\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ are

$$\begin{aligned} f_1 &= \beta_1\alpha_1 + \beta_2\alpha_3 \\ f_2 &= \beta_1\alpha_2 + \beta_2\alpha_4 \\ f_3 &= \beta_3\alpha_1 + \beta_4\alpha_3 \\ f_4 &= \beta_3\alpha_2 + \beta_4\alpha_4 \end{aligned}$$

The identity element $\boldsymbol{e} = (1, 0, 0, 1)$. The remaining postulates are met automatically because of the rules of matrix multiplication and inversion.

In this example there is an isomorphic mapping of the group $\mathcal{G}$ (the $Gl(2,R)$ matrices) into the space R_4 of real 4-tuples. The choice of parameters together with the rules of matrix multiplication produce a composition function that satisfies the requirements of Definition 3.1 for all the elements of G_4. (G_4 is the space R_4 excluding those points for which $\alpha_1\alpha_4 = \alpha_2\alpha_3$.) This is what we had hoped for in posing this definition. Unfortunately, this example illustrates the exception rather than the rule. The next example illustrates some common difficulties.

Example 3.2 *The Group $SU(2)$*

$SU(2)$ is the group of 2×2 unitary matrices u with $\det(u) = 1$. If p and q are complex numbers with $|p|^2 + |q|^2 = 1$, then u must have the form

$$u = \begin{bmatrix} p & q \\ -q^* & p^* \end{bmatrix} \tag{3.1}$$

The identity element is $p = 1$, $q = 0$. It requires three real parameters to specify two complex numbers satisfying the constraint. A natural parameterization is given by

$$\begin{aligned} \operatorname{Im} q &= \alpha_1/2 \\ \operatorname{Re} q &= \alpha_2/2 \\ \operatorname{Im} p &= \alpha_3/2 \\ \operatorname{Re} p &= +\sqrt{1 - (\alpha_1^2 + \alpha_2^2 + \alpha_3^2)/4} \end{aligned}$$

but this is restricted to $\operatorname{Re} p \geq 0$. This spoils the homomorphism and also violates the fourth axiom, because a matrix with $\operatorname{Re} p \geq 0$ could easily have an inverse with $\operatorname{Re} p < 0$.

Here is another choice.

$$p = \cos(\alpha_1/2) \exp[i(\alpha_2 + \alpha_3)/2]$$

$$q = \sin(\alpha_1/2) \exp[i(\alpha_2 - \alpha_3)/2]$$

All allowed values of p and q are generated with this parameterization, but a more subtle problem has appeared. When $\alpha_1 = 0$ all α's with $\alpha_2 + \alpha_3 = c$ (where c is an arbitrary constant) yield the same p and q. Thus the mapping from the corresponding group elements to R_3 is infinitely multiple-valued. Furthermore, an element of the form $\boldsymbol{\alpha} = (0, \alpha_2, \alpha_3)$ does not have a unique inverse, since any $\boldsymbol{\omega} = (0, \omega_2, \omega_3)$ with $\omega_2 + \omega_3 + \alpha_2 + \alpha_3 = 0$ will satisfy $\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\omega}) = \boldsymbol{e}$. Such elements are called *singular points* because the Jacobian of the transformation that connects the parameters with the group elements vanishes at these special values of the parameters.

This example illustrates the general rule that most groups cannot be parameterized in such a way that their composition functions satisfy Definition 3.1 for all elements of the group. (We have already encountered this problem in connection with the rotation group in Section 1.3.) Nevertheless the parameterizations in Example 3.2 do yield composition functions that are valid within limited regions of G_3; and by using several alternate parameterizations we can cover the entire group. This leads to the concept of *local subgroups*, which are limited regions within a group that can be covered with a single parameterization.

The phrase "limited region" in turn implies some concept of distance. We will use the usual Euclidean distance function

$$d(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sqrt{\sum_{i=1}^{n} |\alpha_i - \beta_i|^2} \tag{3.2}$$

Where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are any two elements of R_n. The function $d(\boldsymbol{\alpha}, \boldsymbol{\beta})$ posseses all the properties normally associated with distance:

(1) $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) \geq 0$ and $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0$ if and only if $\boldsymbol{\alpha} = \boldsymbol{\beta}$.

(2) $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) = d(\boldsymbol{\beta}, \boldsymbol{\alpha})$

(3) $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) + d(\boldsymbol{\beta}, \boldsymbol{\gamma}) \geq d(\boldsymbol{\alpha}, \boldsymbol{\gamma})$

The statement "$\boldsymbol{\alpha}$ is in a region or neighborhood U of $\boldsymbol{\beta}$" means that there is a positive number ϵ such that $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) < \epsilon$.

Now suppose we are able to find a parameterization that satisfies the axioms of Definition 3.1 for some subset of $\mathcal{G}$ that includes the identity. The corresponding parameters are said to form a *local subgroup* according to the following definition.

Definition 3.2 *Local Subgroup*

Let U and L be two subsets of R_n such that $e \in U \subseteq L$. L is choosen small enough that there is an isomorphic mapping of L into $\mathcal{G}$, i.e. *every element in L is an image of one and only one element of $\mathcal{G}$. U is choosen so that $\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha}) \in L$ for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in U$ and so that the inverse of each element of U is contained in L. Within this limited region the composition function satisfies the requirements of Definition 3.1. Then L is called a* **local subgroup** *of $\mathcal{G}$, and U is called the* **germ** *of the local subgroup.*

We will also use the terms "germ" and "local subgroup" to refer to the corresponding subsets of the abstract group $\mathcal{G}$; thus $\mathcal{U} \subset \mathcal{G}$ is the isomorphic image of U, and $\mathcal{L} \subset \mathcal{G}$ is the image of L. This mapping between U and $\mathcal{G}$ is very useful, because it allows us to apply the concept of distance to at least some elements of $\mathcal{G}$. For example, if a and b are two elements of $\mathcal{G}$ with counterparts $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in U, we will say that the "distance" between a and b is $d(\boldsymbol{\alpha}, \boldsymbol{\beta})$. In this way we can apply quantatative ideas such as limit, continuity, differentiation, *etc.* to the abstract group where they would otherwise not be applicable. At this stage this definition of distance only works in that limited region of $\mathcal{G}$ that can be mapped isomorphically to R_n. We will eventually overcome this limitation (at least partially) by using "coordinate system patches." In the meantime we can define continuity and analyticity of the local subgroup.

Definition 3.3 *Continuity of the Composition Function*

The composition function $\boldsymbol{\gamma} = \boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha})$ where $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ are elements of R_n is continuous at the identity if for every neighborhood V of $\boldsymbol{e}$ there exists a neighborhood U of $\boldsymbol{e}$ such that if $\boldsymbol{\alpha} \in U$ and $\boldsymbol{\beta} \in U$ then $\boldsymbol{\gamma} = \boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha}) \in V$.

Definition 3.4 *Analyticity of the Composition Function*

The composition function is analytic in U if the following two requirements are satisfied:

(1) The function as well as all its partial derivatives to all orders with respect to all $2n$ of its arguments are continuous in U. In the language of the calculus of several variables it is a C^∞ function.

(2) For each $\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0 \in U$ the function $\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha})$ has a Taylor's series expansion that converges absolutely to $\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha})$ for all $\boldsymbol{\alpha}, \boldsymbol{\beta}$ in a neighborhood $V \subset U$ centered at $(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0)$.

It is possible to construct functions that are C^∞ but which fail to converge to the right value everywhere in V as required by (2). These are mostly mathematical curiosities, however.

It is worth noting that an analytic composition function guarantees the existence of the inverse satisfying (4) of Definition 3.2. This can be proved by reference to the implicit function theorem from the calculus of several variables (Fleming, 1977).

Theorem 3.1 *If $\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha})$ is analytic at $\boldsymbol{e}$ then there exists a neighborhood U of $\boldsymbol{e}$ such that $\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\omega}) = \boldsymbol{f}(\boldsymbol{\omega}, \boldsymbol{\alpha}) = \boldsymbol{e}$ has a unique solution for all $\boldsymbol{\alpha} \in U$. Furthermore $\boldsymbol{\omega}$ is also contained in a neighborhood U^{-1} of $\boldsymbol{e}$.*

Proof: Let $\boldsymbol{H}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha}) - \boldsymbol{e}$. Then $\boldsymbol{H}(\boldsymbol{e}, \boldsymbol{e}) = \boldsymbol{0}$. The Jacobian at this point is non-zero since

$$\left.\frac{\partial H_j(\boldsymbol{e}, \boldsymbol{\omega})}{\partial \omega_i}\right|_{\boldsymbol{\omega}=\boldsymbol{e}} = \frac{\partial \omega_j}{\partial \omega_i} = \delta_{ij}$$

The implicit function theorem guarantees that the equation $\boldsymbol{H}(\boldsymbol{\alpha}, \boldsymbol{\omega}) = \boldsymbol{0}$ has a solution with the properties stated above. This establishes the right inverse. The existence of the left inverse follows from

$$\left.\frac{\partial H_j(\boldsymbol{\omega}, \boldsymbol{e})}{\partial \omega_i}\right|_{\boldsymbol{\omega}=\boldsymbol{e}} = \frac{\partial \omega_j}{\partial \omega_i} = \delta_{ij}$$

Now let $\boldsymbol{\omega}_r$ be the solution of the equation $\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\omega}_r) = \boldsymbol{e}$ and $\boldsymbol{\omega}_l$ be the solution to $\boldsymbol{f}(\boldsymbol{\omega}_l, \boldsymbol{\alpha}) = \boldsymbol{e}$. Then

$$\boldsymbol{\omega}_l = \boldsymbol{f}(\boldsymbol{\omega}_l, \boldsymbol{e}) = \boldsymbol{f}(\boldsymbol{\omega}_l, \boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\omega}_r) = \boldsymbol{f}(\boldsymbol{f}(\boldsymbol{\omega}_l, \boldsymbol{\alpha}), \boldsymbol{\omega}_r) = \boldsymbol{f}(\boldsymbol{e}, \boldsymbol{\omega}_r) = \boldsymbol{\omega}_r$$

so the left and right inverses are identical.

Definition 3.5 *Local Lie Subgroup*

A local subgroup is also a local Lie subgroup in the neighborhood U of the identity if its composition function is analytic there.

The neighborhoods U and L in Definition 3.2 are not necessarily identical, so the local Lie subgroup is not a *group* in the usual sense of the word. (The author is not responsible for this terminology.) Starting with the germ $\mathcal{U}$, however, we can build up a group $\mathcal{G}_0$ consisting of finite products $a_n \cdots a_1$ of elements $a_i \in \mathcal{U}$. The elements in $\mathcal{G}_0$ are said to be *connected to the identity*, and $\mathcal{G}_0$ is called the *connected component* of the group $\mathcal{G}$. Clearly the connected component is itself a group. In the usual terminology, the germ $\mathcal{U}$ *generates* $\mathcal{G}_0$.

A formal definition of connectedness is given in the next section. Intuitively, however, a group or set is connected to the identity if every element can be reached by a series of small steps starting at the identity, each step lying within the set. The groups $SO(3)$ and $O(3)$ discussed in Example 1.4c illustrate this concept. The group $SO(3)$ consisting of pure rotations can be built up starting with the identity from a finite product of arbitrarily small rotations. Each small rotation by itself is a member of a local Lie group. All the elements of $SO(3)$, therefore, are connected to the identity. The group $O(3)$ contains all these elements plus another component that can only be reached by multiplying the elements of $SO(3)$ by a matrix with determinant$= -1$. These elements cannot be reached by a succession of small steps, and thus are not connected to the identity.

Now let d_i and d_{i+1} be elements in $\mathcal{G}_0$ such that

$$d_i = a_i a_{i-1} \cdots a_1 e$$

and

$$d_{i+1} = a_{i+1} a_i \cdots a_1 e$$

Take another element c that lies "between" d_i and d_{i+1} in the sense that

$$c = b d_i$$

and

$$c = b' d_{i+1}$$

where $b, b' \in \mathcal{U}$ with corresponding parameter vectors $\boldsymbol{\beta}$, $\boldsymbol{\beta}' \in U$. These parameters depend not only on c but also on the i or $i+1$ elements comprising d_i or d_{i+1}. It is useful to think of d_i as defining the origin of a local coordinate system with respect to which the coordinates of c are $\boldsymbol{\beta}$. Then d_{i+1} defines the origin of a new coordinate system in which the coordinates of c are $\boldsymbol{\beta}'$. This is illustrated in Figure 3.1. The region labeled $\mathcal{U}_i$ contains all those elements that can be obtained by multiplying $a_i \cdots a_1 e$ by an element of $\mathcal{U}$. The region labeled $\mathcal{U}_{i+1}$ is similarly obtained from $a_{i+1} a_i \cdots a_1 e$. Any point

c in the overlap region has two sets of parameters, $\boldsymbol{\beta}$ and $\boldsymbol{\beta}'$; consequently there is a mapping or function that expresses $\boldsymbol{\beta}$ in terms of $\boldsymbol{\beta}'$ and vice versa.

A set of elements like $\mathcal{U}_i$ that can all be specified with a single system of parameters (*ie.* without reparameterizing) is called a *coordinate system patch.* By chosing different elements to play the role of c and different chains of a_i's we can cover the entire group $\mathcal{G}_0$ with interlocking coordinate patches. This would be useless unless the different coordinate systems fitted together in a consistent way. This is guaranteed by the next assumption:

Definition 3.6 *Global Connected Lie Group*

A global connected Lie group is a group that can be covered with local Lie groups using the procedure described above. When two or more local groups overlap their parameter sets can be related to one another by analytic functions.

As we have seen, each coordinate system patch obtains its coordinates from an element in the germ of the local subgroup. It is customary to choose these parameters so that the identity element $\boldsymbol{e} = \mathbf{0} = (0, \ldots, 0)$. It is clear that this is always possible, since if the coordinates of the identity were $\boldsymbol{\alpha}_0$, we could define a new set of parameters $\boldsymbol{\alpha}' = \boldsymbol{\alpha} - \boldsymbol{\alpha}_0$. This linear translation is itself an analytic function, so none of the analytic properties implied by Definition 3.6 would be affected. A parameter set obtained in this way and applied to a local group is called a *local parameterization.* These parameters are formally equivalent to an n-dimensional Euclidean coordinate system. For example, the distance of any element from the identity is $d(\boldsymbol{\alpha}, \mathbf{0}) = \sqrt{\sum_{i=1}^{n} |\alpha_i|^2}$, which is just the Pythagorean theorem. In effect, we have covered the abstract group $\mathcal{G}_0$ with a patchwork of Euclidean coordinate systems.

This process of building up the global Lie group using a patchwork of local subgroups is closely analogous to the process of analytic continuation in complex function theory. We start with the neighborhood $\mathcal{U}$ of the identity. The neighborhood can be made arbitrarily large so long as the local Lie subgroup postulates hold; but in general its size will be limited by $\boldsymbol{\alpha}$'s and $\boldsymbol{\beta}$'s for which $\boldsymbol{f}$ is not continuous and/or the inverse is not unique. Now choose an element a_1 corresponding to a parameter vector $\boldsymbol{\alpha_1} \in U$ and reparameterize so that a_1 becomes the origin of a new local Lie subgroup called $\mathcal{U}_1$. We then choose an element $a_2 a_1 \in \mathcal{U}_1$ and use it as the origin of the next local subgroup. Proceeding in this way we can eventually parameterize any element of $\mathcal{G}_0$.

In the analogous process of analytic continuation, we expand a function in a power series centered at the origin. This series converges in a neighbor-

hood limited in size by the nearest singularity. We then choose a new point within the neighborhood and use it as the origin of a new power series expansion. The coefficients of the new series are determined by the coefficients of the original series and the point (analogous to $\boldsymbol{\alpha}_1$) is used as the new origin. In this way most singularities can be "outflanked," and the function is determined everywhere in the complex plane in terms of its derivatives at the origin.

It is useful to think of Lie groups as the group-theoretical analog of analytic functions. The definition of a Lie group, however it may be stated, has as its motivation the requirement that the group is globally determined from its properties in a neighborhood of the origin. The properties of this local group are in turn determined by the Lie algebra, which is a statement about the group *at* the identity.

3.2 Global Structure

We have seen how the germ of a local group generates a global connected Lie group. There is an even larger structure to consider, however. A Lie group can have several "pieces," each piece consisting of elements that are connected to one another but not to elements in other pieces. These separate pieces are usually called *components* of the group, and the component containing the identity is called the *connected component*. Within each component there can be further structure arising from the phenomenon of multiple connectedness.

Definition 3.7 *Connected Spaces*
A space M is **connected** *if for every $p, q \in M$ there exists a continuous curve $f(t)$ that maps the interval $[a, b]$ of the real axis into M so that $p = f(a)$ and $q = f(b)$.*

In other words, any two points can be connected by a smooth curve lying entirely within the set.

Definition 3.8 *Simply Connected Spaces*
The space M is **simply connected** *if every curve connecting any two points can be continuously deformed into every other such curve. If this is not possible the space is said to be* multiply connected.

The usual example of a multiply connected space is the surface or interior of a torus. Two points on the surface, for example, can be connected by paths running through the "donut hole" and by paths running along the outside of the torus, and these paths cannot be deformed into one another.

Example 3.3 *$SU(2)$ as a simply connected group.*

The group $SU(2)$ was discussed in Example 3.2. Refering to Equation 3.1 we set $p = \alpha_1 + i\alpha_2$ and $q = \alpha_3 + i\alpha_4$. The condition $|p|^2 + |q|^2 = 1$ becomes

$$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2 = 1$$

This equation describes the surface of a sphere in four-dimensional Euclidean space. It is intuitively clear in three dimensions that any curve lying on the surface of a sphere and connecting two points can be continuously deformed into any other curve connecting the same points. The same thing is true in four dimensions: $SU(2)$ is simply connected.

Example 3.4 *$SO(2)$ as a multiply connnected group.*

The set of all 2×2 special orthogonal matrices is a one-parameter group that can be parameterized as follows.

$$M(\alpha_1) = \begin{pmatrix} \cos\alpha_1 & \sin\alpha_1 \\ -\sin\alpha_1 & \cos\alpha_1 \end{pmatrix}$$

The curve generated by $\alpha_1 = 2\pi t$ with $0 \leq t \leq 1$ starts at $M(0) = I$ and ends at the same point. In this case the ends points are identical, and the curve is a loop. Another path between I and I is generated by $\alpha_1 = 0t$. These paths cannot be continuously deformed into one another, since only $\alpha_1 = 2\pi n$ with n equal to an integer can produce $M(\alpha_1) = I$. Therefore, $SO(2)$ is not simply connected. It is plausible (and true) that all the $SO(N)$ groups with $N \leq 2$ are multiply connected because of the cyclic property of sines and cosines.

Definition 3.9 *Locally Connected Spaces*
A space is **locally connected** *if every point is contained in an open neighborhood that is itself a connected space.*

For example, a space consisting of two disjoint open intervals on the real axis is locally connected but not connected.

It is a standard result for topological spaces, of which Lie groups are a special case, (*eg.* Singer and Thorpe, 1967), that the definition given above for a connected space is equivalent to any of the following conditions:

(1) M is not the union of two nonempty disjoint closed subsets.

(2) M is not the union of two nonempty disjoint open subsets.

(3) The only subsets of M that are both open and closed are M and the empty set.

The terms "open" and "closed" are used in the usual sense. A subset $A \subset M$ is said to be open if each element in it is contained within a neighborhood that itself lies entirely within the subset. The subset is closed if it contains all its limit points, *i.e.* there is no element $a \in M$ with the property that $a \notin A$ and yet every neighborhood of a intersects A.

Theorem 3.2 *$\mathcal{G}_0$ is open, connected, and closed.*

Proof: $\mathcal{G}_0$ consists of all finite products of the form $a_n \ldots a_1$ where $a_i \in \mathcal{U}$, a germ of the local subgroup. Let $a_0 = a_n \ldots a_1 \in \mathcal{G}_0$. We use the notation $a_0\mathcal{U}$ to indicate the set of all elements obtained by left multiplying an element

of $\mathcal{U}$ by a_0. Then $a_0 = a_0 e \in a_0\mathcal{U} \subset \mathcal{G}_0$. Thus $a_0\mathcal{U}$ is a set containing a_0 that is itself completely contained in $\mathcal{G}_0$, which proves that $\mathcal{G}_0$ is open.

To prove that $\mathcal{G}_0$ is connected assume the opposite, that there are two sets M and N with $\mathcal{U} \subset M$ such that $\mathcal{G}_0 = M \cup N$ and $M \cap N$ is empty. This implies that there is an $a \in M$ and a $b \in N$ such that $b = a_0 a$ with $a_0 \in \mathcal{U}$. If M and N are disjoint, however, Ma^{-1} and Na^{-1} are also disjoint. Then $a_0 = ba^{-1} \in Na^{-1}$, but a_0 is an element of the local subgroup defined in a neighborhood of the identity $e = aa^{-1} \in Ma^{-1}$, and we arrive at a contradiction.

To show that $\mathcal{G}_0$ is closed, suppose that the element $b \notin \mathcal{G}_\circ$ has the property that every neighborhood of b intersects $\mathcal{G}_\circ$. One such neighborhood is $b\mathcal{U}$. Let a_0 be an element in this intersection. Then $a_0 = bb_1 = a_1 \dots a_n$, where $b_1, a_1, \dots, a_n \in \mathcal{U}$. It follows that $b = a_1 \cdots a_n b_1^{-1} \in \mathcal{G}_0$ The contradiction proves that $\mathcal{G}_0$ is closed.

The second part of the proof shows that all elements in $\mathcal{G}_0$ are connected. The converse is also true; if an element is connected to the identity, then it is a member of $\mathcal{G}_0$. A set like this that contains all elements that can be smoothly connected to one another is called a *maximal connected set.*

Theorem 3.3 *$\mathcal{G}_0$ is a maximal connected set.*

Proof: Suppose there were an element $a \notin \mathcal{G}_0$ and a continuous curve $a(t)$ so that $a(0) = e$ and $a(1) = a$. Since the curve starts off in $\mathcal{G}_0$ and since $\mathcal{G}_0$ is closed, there would have to be some t_0 such that $a(t_0) \in \mathcal{G}_0$ and $a(t) \notin \mathcal{G}_0$ for $t_0 < t$. The element $a^{-1}(t_0)a(t_0 + \epsilon)$ exhibits the contradiction. When $\epsilon = 0$ it is equal to $e \in \mathcal{G}_0$, but for arbitrarily small positive ϵ it is no longer contained in $\mathcal{G}_0$. By definition, however, $\mathcal{G}_0$ is made up of products of elements from the neighborhood of the identity. These last two statements cannot be reconciled, so our hypothesis is wrong, and the theorem is proved.

Definition 3.10 *$\mathcal{F}$ is a* **subgroup** *of a group $\mathcal{G}$ if it is a subset and if the elements of $\mathcal{F}$ satisfy the group postulates.*

Definition 3.11 *Invariant subgroup*
$\mathcal{F}$ is an **invariant** *or* **normal** *subgroup of $\mathcal{G}$ if $a\mathcal{F}a^{-1} \subseteq \mathcal{F}$ for all $a \in \mathcal{G}$.*

Definition 3.12 *Coset*
Let $\mathcal{F}$ be a subgroup of a group $\mathcal{G}$. Then for any $a \in \mathcal{G}$ the set $a\mathcal{F}$ is called the left coset of $\mathcal{F}$ with respect to a. Similarly the set $\mathcal{F}a$ is called the right coset.

Unfortunately, there seems to be no agreement in the literature whether $a\mathcal{F}$ should be called the left coset, because a is to the left of $\mathcal{F}$, or the right coset, because $\mathcal{F}$ is to the right of a.

The following theorem is stated for right cosets, but it is equally true for left cosets.

Theorem 3.4
(1) If $a \in \mathcal{F}$, then $\mathcal{F}a = \mathcal{F}$.
(2) Two right cosets of $\mathcal{F}$ are either identical or they are disjoint.

Proof:

(1) If $a \in \mathcal{F}$, then certainly $\mathcal{F}a \subseteq \mathcal{F}$. It is possible to prove the much stronger statement, however, that $\mathcal{F}a$ contains every element of $\mathcal{F}$ once and only once. Let b be an arbitrary element of $\mathcal{F}$, and define $c = ba^{-1} \in \mathcal{F}$. Then $ca \in \mathcal{F}a$ and $ca = b$, so $\mathcal{F}a$ contains every element of $\mathcal{F}$ at least once. Now suppose that it contained an element more than once. This would imply that for some $b_1, b_2 \in \mathcal{F}$ $b_1a = b_2a$ but $b_1 \neq b_2$, and these statements are obviously inconsistent.

This statement is sometimes called the "rearrangement theorem." It is equally true if we replace the subgroup $\mathcal{F}$ with the entire group $\mathcal{G}$. It implies that the operation of left (or right) multiplication merely rearranges the elements in the group.

(2) Suppose that $\mathcal{F}a$ and $\mathcal{F}a'$ contain a common element, *ie.* for some $b, b' \in \mathcal{F}$ $ba = b'a'$. Then $b'^{-1}b = a'a^{-1}$, and $a'a^{-1} \in \mathcal{F}$. From (1) $\mathcal{F}(a'a^{-1}) = \mathcal{F}$. Now $\mathcal{F}a = \mathcal{F}(a'a^{-1})a^{-1} = \mathcal{F}a'$. Consequently, if there is any common element, the two cosets are identical.

These proofs depend on the assumption that $a \in \mathcal{F}$. In general, $\mathcal{F}a \neq \mathcal{F} \neq a\mathcal{F}$, and the left and right cosets are not necessarily identical. However, if $a\mathcal{F} = \mathcal{F}a$ for all $a \in \mathcal{G}$ then $a^{-1}\mathcal{F}a = \mathcal{F}$ and $\mathcal{F}$ is an invariant subgroup. In this case we can construct a new group $\mathcal{G}/\mathcal{F}$ known as the *factor group*. The "elements" of this group are the distinct right cosets of $\mathcal{F}$; in other words, we ignore the internal structure of the cosets. If $b \in \mathcal{F}$ and $a \in \mathcal{G}$, then $\mathcal{F}a$ and $\mathcal{F}ba$ are identical. Multiplication is defined as follows.

Definition 3.13 *Product of right cosets.*
Let $\mathcal{F}$ be an invariant subgroup. The product of the cosets $\mathcal{F}a$ and $\mathcal{F}b$ is defined by

$$(\mathcal{F}a)(\mathcal{F}b) = \mathcal{F}(ab)$$

.

It is necessary to show that this definition is consistent, *ie.* if $a' \in \mathcal{F}a$ and $b' \in \mathcal{F}b$ then $a'b' \in \mathcal{F}(ab)$.

Proof: There must be some $c, c' \in \mathcal{F}$ such that $a' = ca$, $b' = c'b$, and $a'b' = cac'b$. Since $ac' \in a\mathcal{F}$, and since $\mathcal{F}$ is an invariant subgroup, there must be an element $c'' \in \mathcal{F}$ such that $ac' = c''a$. Then $a'b' = (cc'')(ab)$ and $cc'' \in \mathcal{F}$.

It is now trivial to show that the elements $\mathcal{F}a$, $\mathcal{F}b, \ldots$ form a group called $\mathcal{G}/\mathcal{F}$ or the *factor group of* $\mathcal{G}$ *by* $\mathcal{F}$.

Theorem 3.5 *Let $\mathcal{G}$ be a Lie group and let $\mathcal{G}_0$ be the connected component containing the identity e as in Theorem 3.2. Then $\mathcal{G}_0$ is an invariant Lie subgroup, and the components of $\mathcal{G}$ are the cosets of $\mathcal{G}_0$.*

Proof: $\mathcal{G}_0$ easily satisfies the group axioms. It is a Lie subgroup since it uses the same multiplication table as $\mathcal{G}$.

To show that it is an invariant subgroup, let $a \in \mathcal{G}$, and consider the set $a\mathcal{G}_0a^{-1}$. This set contains the identity $e = aea^{-1}$, and all its elements are connected to the identity, so that $a\mathcal{G}_0a^{-1} \subseteq \mathcal{G}_0$.

The operation of right (or left) multiplication is clearly a mapping in the sense of Definition 1.2. Every element in $\mathcal{G}_0a$ is an image of an element in $\mathcal{G}_0$. Since we are dealing with Lie groups the mapping is analytic, so that if two elements in $\mathcal{G}_0$ can be connected by a smooth curve, then the images in $\mathcal{G}_0a$ have the same property. Now suppose that there were an element b that was not contained in $\mathcal{G}_0a$, but which could be connected to elements in $\mathcal{G}_0a$. Then ba^{-1} would not be contained in $\mathcal{G}_0$ but yet would be connected to elements in $\mathcal{G}_0$. This is inconsistent with the fact that $\mathcal{G}_0$ is a maximal connected subset. Just as $\mathcal{G}_0$ contains all possible elements that can be connected to e, so $\mathcal{G}_0a$ contains all elements that can be connected to a. If $a \in \mathcal{G}_0$ these sets are identical, but with a suitable choice of right multipliers that cosets will comprise the entire group.

The next theorems make use of the concept of homomorphism. This was introduced in Definition 1.3 for general groups, but an extra refinement is required for Lie groups. Roughly speaking we require that the mapping function ϕ (Definition 1.2) be C^∞. In order to differentiate ϕ, however, it must be expressed in terms of parameters rather than abstract group elements. This leads to the following definition.

Definition 3.14 *Lie group homomorphism.*

A mapping ϕ of $\mathcal{G}$ into $\mathcal{G}'$ is a Lie group homomorphism if, (1) it is a homomorphism, and (2) the coordinates of a' are C^∞ functions of the

coordinate a for every $\phi(a) = a' \in \mathcal{G}'$ and for every local parameterization of a and a'.

Several other related mappings are defined in obvious ways. A *Lie group isomorphism* is an isomorphism (Definition 1.3) for which (2) is true. A *local* Lie group homomorphism (or isomorphism) is one in which (2) holds in some neighborhood U of e in $\mathcal{G}$. In what follows we will assume that (2) is true whenever the words homomorphism and isomorphism are used.

Definition 3.15 *Kernel of a homomorphic mapping.*

Let ϕ be a homomorphic mapping of a group $\mathcal{G}$ into $\mathcal{G}'$. Then the set of elements $a \in \mathcal{G}$ such that $\phi(a) = e'$, the identity on $\mathcal{G}'$ is called the **kernel**, *which we will call $\mathcal{K}$.*

Theorem 3.6 *Let ϕ be a homomorphic mapping of $\mathcal{G}$ into $\mathcal{G}'$ and let $\mathcal{K}$ be the kernel of this mapping. Then*

(1) $\mathcal{K}$ is an invariant subgroup of $\mathcal{G}$.

(2) Because of (1) we can define a factor group $\mathcal{K}a$ where a ranges over all of $\mathcal{G}$. Every element of this group can be mapped one-to-one onto $\mathcal{G}'$ by the mapping θ defined by

$$\theta(\mathcal{K}a) = \phi(a)$$

Furthermore, θ is an isomorphic mapping of $\mathcal{G}/\mathcal{K}$ onto $\mathcal{G}'$.

Proof: (1) It is trivial to show that the elements of $\mathcal{K}$ satisfy the group axioms. To prove that it is invariant let $a \in \mathcal{G}$ and $k \in \mathcal{K}$. Then $\phi(aka^{-1}) = \phi(a)\phi(k)\phi(a^{-1}) = \phi(a)e'\phi(a^{-1}) = e'$; so $aka^{-1} \in \mathcal{K}$.

(2) Let ka be an arbitrary element of $\mathcal{K}a$. $\phi(ka) = e'\phi(a) = \phi(a)$; so every element in $\mathcal{K}a$ maps to the same element in $\mathcal{G}'$. The mapping is one-to-one, since if $\theta(\mathcal{K}a_1) = \theta(\mathcal{K}a_2)$ then $\mathcal{K}a_1 = \mathcal{K}a_2$. This is proved as follows: $\theta(\mathcal{K}a_1) = \theta(\mathcal{K}a_2)$ implies $\phi(a_1) = \phi(a_2)$. There must be a third element $a_3 \in \mathcal{G}$ such that $a_1 = a_2a_3$. Then $\phi(a_3) = e'$, $a_3 \in \mathcal{K}$, and $a_1 \in \mathcal{K}a_2$. From part (2) of Theorem 3.4, $\mathcal{K}a_1 = \mathcal{K}a_2$.

The mapping is homomorphic since

$$\theta(\mathcal{K}a_1)\theta(\mathcal{K}a_2) = \phi(a_1)\phi(a_2) = \phi(a_1a_2) = \theta((\mathcal{K}a_1)(\mathcal{K}a_2)).$$

It is also one-to-one and thus isomorphic.

Definition 3.16 *The center $\mathcal{C}$ of a group $\mathcal{G}$.*

The center *of a group $\mathcal{G}$ is the subgroup consisting of all elements of $\mathcal{G}$ that commute with every element of $\mathcal{G}$.*

Since $\mathcal{C}$ is obviously invariant it is sometimes called the central invariant subgroup. A group like $\mathcal{C}$ whose elements all commute with one another is said to be *abelian*.

The following important theorems are stated without proof. They are proved in several standard references including Hausner and Schwartz (1968).

Theorem 3.7 *If two simply connected groups $\mathcal{G}$ and $\mathcal{G}'$ are locally isomorphic they are isomorphic in the large.*

Thus a simply connected group is completely determined by the germ of its local group. Two simply connected groups with that same (in the sense of isomorphic) germ are the same group.

This does not happen with multiply connected groups, but there is a unique simply connected group that is locally isomorphic to any multiply connected group. The details are contained in the following remarkable theorem.

Theorem 3.8 *Let $\mathcal{G}$ be a connected group. Then there is a unique simply connected group $\tilde{\mathcal{G}}$ that is locally isomorphic to $\mathcal{G}$. The maping ϕ of $\tilde{\mathcal{G}}$ onto $\mathcal{G}$ is homomorphic and locally isomorphic. The kernel $\mathcal{K}$ of ϕ is an invariant central subgroup of $\tilde{\mathcal{G}}$, and it is discrete. Moreover, $\mathcal{G}$ is globally isomorphic to the factor group $\tilde{\mathcal{G}}/\mathcal{K}$.*

The group $\tilde{\mathcal{G}}$ is called the *universal covering group* for $\mathcal{G}$. If $\mathcal{G}$ itself is simply connected, then $\mathcal{G}$ and $\tilde{\mathcal{G}}$ are isomorphic by the previous theorem. The following examples illustrate the universal covering group for some multiply connected groups.

Example 3.5 *$SO(2)$ and $U(1)$.*

$U(1)$ is the set of all 1×1 unitary matrices. A convenient parameterization is given by $e^{i\alpha} = M(\alpha)$. The identity $e = M(0)$, and since $M(\beta)M(\alpha) = M(\beta + \alpha)$ the composition function $f(\beta + \alpha) = \beta + \alpha$. All the group elements are reachable with this parameterization.

The group $SO(2)$ was introduced in Example 3.1. The mapping

$$\phi(e^{i\alpha}) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

is a Lie group isomorphism. $U(1)$, like $SO(2)$, is multiply connected.

The covering group is the multiplicative group of real positive numbers, often called R_+. Let a, b, c be real positive numbers, and define the group

multiplication rule to be ordinary multiplication; then the group axioms are satisfied with $a^{-1} = 1/a$. The group is simply connected in a rather trivial way, since there is only one path between any two points. A good parameterization is $a = e^{\tilde{\alpha}}$ with $-\infty < \tilde{\alpha} < +\infty$. The composition function is $\tilde{\gamma} = \tilde{\alpha} + \tilde{\beta}$.

The mapping

$$\phi(e^{\tilde{\alpha}}) = e^{i\tilde{\alpha}}$$

is a Lie homomorphism because

$$\phi(e^{\tilde{\beta}})\phi(e^{\tilde{\alpha}}) = \phi(e^{\tilde{\beta}}e^{\tilde{\alpha}})$$

and because the relation $\tilde{\alpha} = i\alpha$ is C^∞. In the neighborhood of the identity $\tilde{\alpha} = \alpha = 0$ the mapping is one-to-one and thus isomorphic; but all elements of R_+ of the form $\exp(\tilde{\alpha}+2\pi n)$ with n an integer map to the same element $e^{i\alpha}$ of $U(1)$, so the mapping is globally homomorphic. This is easy to visualize geometrically as shown in Figure 3.2. The kernel of the homomorphism consists of the elements $e^{2\pi n}$, which constitute a discrete, abelian subgroup of R_+ as required by the theorem.

The factor group $\tilde{\mathcal{G}}/\mathcal{K}$ consists of all the distinct elements of the form $\mathcal{K}a = e^{2\pi n}e^{\tilde{\alpha}}$. According to the definition of a factor group, $\mathcal{K}e^{\tilde{\alpha}}$ and $\mathcal{K}e^{\tilde{\alpha}+2\pi m}$ are the same element, since $e^{2\pi n}$ and $e^{2\pi(n+m)}$ (m an integer) are both elements of $\mathcal{K}$. The distinct element of $\tilde{\mathcal{G}}/\mathcal{K}$ can be written $\mathcal{K}e^{\tilde{\alpha}}$ with $0 \le \tilde{\alpha} < 2\pi$. This is isomorphic with $U(1)$ verifying the last statement of the theorem.

The statements made about $U(1)$ are equally true for $SO(2)$, so R_+ is the covering for both groups.

Example 3.6 *$SO(3)$ and its covering group $SU(2)$.*

$SO(3)$ is the group of real orthogonal matrices in three dimensions with determinant= $+1$. The rotation matrices, Equation 1.21, belong to this group. Like $SO(2)$ it is multiply connected. $SU(2)$ is the group of 2×2 complex unitary matrices u discussed in Examples 3.2 and 3.3. It is not obvious that they are even locally isomorphic, but this can be exhibited with the following indirect argument.

The $SO(3)$ matrices M_{ij} operate on a space of three-component column matrices $\boldsymbol{x}$ as in equation (1.5). In matrix notation $\boldsymbol{x}' = M\boldsymbol{x}$. We can construct a new representation space on which the $SU(2)$ matrices act consisting of 2×3 traceless, Hermitian matrices

$$r = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}$$

The operation $\boldsymbol{x}' = M\boldsymbol{x}$ has its counterpart on this space in the operation $r' = uru^{-1}$. The scalar product is defined by the operation

$$\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = \frac{1}{2} Tr(r_1 r_2)$$

The $SU(2)$ transformations preserve this scalar product.

$$\boldsymbol{x}_1' \cdot \boldsymbol{x}_2' = \frac{1}{2} Tr(r_1' r_2') = \frac{1}{2} Tr[(ur_1u^{-1})(ur_2u^{-1})] = \frac{1}{2} Tr[u(r_1 r_2)u^{-1}] = \frac{1}{2} Tr(r_1 r_2) = \boldsymbol{x}_1 \cdot \boldsymbol{x}_2$$

Group multiplication $M_3 = M_2 M_1$ has its counterpart in $u_3 = u_2 u_1$. These are homomorphic since if

$$\begin{aligned} M_1\boldsymbol{x} = \boldsymbol{x}' \quad & M_2\boldsymbol{x}' = \boldsymbol{x}'' \\ u_1 r u_1^{-1} = r' \quad & u_2 r' u_1^{-2} = r'' \\ (u_2u_1)r(u_2u_1)^{-1} = r'' & \end{aligned}$$

So the mapping $\phi(u_1) = M_1$ satisfies

$$\phi(u_2)\phi(u_1) = M_2 M_1 = \phi(u_2 u_1)$$

The kernel of this mapping consists of two elements. Let $u \in \mathcal{K}$, then $uru^{-1} = r$ for all r. By Schur's lemma u is a multiple of the unit matrix I_2. Since $\det(u) = 1$, only I_2 and $-I_2$ are allowed. According to Theorem 3.8 this implies that the mapping is two-to-one. It is interesting to see how this comes about explicitly. Define

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.3}$$

the Pauli spin matrices. Evidentally

$$\begin{aligned} r &= \sum_{i=1}^{3} x_i \sigma_i \\ r' = \sum_{i=1}^{3} x_i u \sigma_i u^{-1} &= \sum_{j=1}^{3} x_j' \sigma_j. \end{aligned}$$

The σ's are orthogonal in the sense that

$$\frac{1}{2} Tr(\sigma_i \sigma_j) = \delta_{ij}, \quad i, j = 1, 2, 3$$

Consequently

$$M(u)_{ij} = \frac{1}{2} Tr(\sigma_j u \sigma_k u^{-1}). \tag{3.4}$$

Obviously u and $-u$ yield the same M. The parameterization

$$u = \begin{bmatrix} \cos\frac{1}{2}\theta \exp\{\frac{1}{2}i(\psi+\phi)\} & \sin\frac{1}{2}\theta \exp\{\frac{1}{2}i(\psi-\phi)\} \\ -\sin\frac{1}{2}\theta \exp\{-\frac{1}{2}i(\psi-\phi)\} & \cos\frac{1}{2}\theta \exp\{-\frac{1}{2}i(\psi+\phi)\} \end{bmatrix}, \tag{3.5}$$

where

$$0 \le \theta \le \pi \quad 0 \le \psi \le 4\pi \quad 0 \le \phi \le 2\pi$$

substituted into (3.4) reproduces the rotation matrix given by (1.17). The factors of $1/2$ appearing in (3.5) are a symptom of the two-to-one mapping. For example, the replacement $\psi \to \psi + 2\pi$ changes $u \to -u$ but leaves M unchanged.

3.3 Invariant integration and compact groups

There are a few technical developments that require the notion of integration over a group. Integration over continuous groups is a generalization of summation over discrete groups, which is much easier to explain. For example, let $G(b_i)$ be a function that associates a number with each group element b_i. The sum over the group is simply

$$\sum_{\mathcal{G}} G(b_i). \tag{3.6}$$

This sum (if finite) defines a number with an interesting invariance property;

$$\sum_{\mathcal{G}} G(cb_i) = \sum_{\mathcal{G}} G(b_i c) = \sum_{\mathcal{G}} G(b_i), \tag{3.7}$$

for any element c. This follows from the rearrangement theorem, Theorem 3.4(1). Right- or left-multiplication by c simply rearranges the terms in the sum.

The generalization of (3.7) to integrals over continuous groups might look like this

$$\int_{b\in\mathcal{G}} G(cb)db = \int_{b\in\mathcal{G}} G(bc)db = \int_{b\in\mathcal{G}} G(b)db. \tag{3.8}$$

$G(b)$ is a continuous function of the element b. The question is, what does the symbol db mean? We might plausibly replace (3.6) with the ordinary Riemann integral

$$\int G(\boldsymbol{\beta})d\boldsymbol{\beta}, \tag{3.9}$$

where $\boldsymbol{\beta}$ is a parameter vector specifying the element b. Unfortunately, (3.9) as it stands lacks the required invariance property. With a slight modification, however, it can be made to satisfy (3.8), at least for a special class of groups. The following argument shows how this is done:

Let $\boldsymbol{\alpha}$ represent an element $a \in \mathcal{G}$ close to the identity, and let $\boldsymbol{\beta}_\circ$ represent another element $b_\circ$. Then

$$\boldsymbol{\beta} = \boldsymbol{f}(\boldsymbol{\beta}_\circ, \boldsymbol{\alpha}) \tag{3.10}$$

maps the neighborhood of the identity to some new neighborhood centered at $\boldsymbol{\beta}_\circ$. Such a mapping is called a *left translation by* $\boldsymbol{\beta}_\circ$. Conversely, the mapping

$$\boldsymbol{\alpha} = \boldsymbol{f}(\boldsymbol{\beta}_\circ^{-1}, \boldsymbol{\beta}) \tag{3.11}$$

maps a neighborhood of $\boldsymbol{\beta}_\circ$ back to a neighborhood of the identity. We obtain invariant integration by using (3.11) to transform every volume element $d\boldsymbol{\beta}$ in (3.9) back to the identity. This is done by replacing

$$d\boldsymbol{\beta} \to d\boldsymbol{\alpha} = \rho_L(\boldsymbol{\beta}) d\boldsymbol{\beta}, \tag{3.12}$$

where

$$\rho_L(\boldsymbol{\beta}_\circ) = \det \left| \frac{\partial \alpha^i}{\partial \beta^j} \right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_\circ} = \det \left| \frac{\partial f^i(\boldsymbol{\beta}_\circ^{-1}, \boldsymbol{\beta})}{\partial \beta^j} \right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_\circ} \tag{3.13}$$

(Note that in passing from (3.13) to (3.12) $\boldsymbol{\beta}_\circ$ has been replaced by $\boldsymbol{\beta}$, *ie.* the Jacobian has been evaluated *at* $\boldsymbol{\beta}$.) The invariance of the resulting integral is shown by the following theorem:

Theorem 3.9

$$\int_{\mathcal{G}} G(\boldsymbol{\beta}') \rho_L(\boldsymbol{\beta}) d\boldsymbol{\beta} = \int_{\mathcal{G}} G(\boldsymbol{\beta}) \rho_L(\boldsymbol{\beta}) d\boldsymbol{\beta} \tag{3.14}$$

where $\boldsymbol{\beta}' = \boldsymbol{f}(\boldsymbol{\gamma}, \boldsymbol{\beta})$, *and* $\boldsymbol{\gamma}$ *is any fixed element.*

The proof follows from the property of Jacobians,

$$\det \left| \frac{\partial \alpha^i}{\partial \beta'^j} \right| = \det \left| \frac{\partial \alpha^i}{\partial \beta^j} \right| \det \left| \frac{\partial \beta^k}{\partial \beta'^l} \right|$$

so

$$\rho_L(\boldsymbol{\beta}') d\boldsymbol{\beta}' = \rho_L(\boldsymbol{\beta}) \det \left| \frac{\partial \beta^i}{\partial \beta'^j} \right| d\boldsymbol{\beta}' = \rho_L(\boldsymbol{\beta}) d\boldsymbol{\beta}$$

and

$$\int_{\mathcal{G}} G(\boldsymbol{\beta}') \rho_L(\boldsymbol{\beta}) d\boldsymbol{\beta} = \int_{\mathcal{G}} G(\boldsymbol{\beta}') \rho_L(\boldsymbol{\beta}') d\boldsymbol{\beta}' = \int_{\mathcal{G}} G(\boldsymbol{\beta}) \rho_L(\boldsymbol{\beta}) d\boldsymbol{\beta}$$

This proof assumes implicitly that the group can be covered with a single parameterization. This in itself is not a serious limitation because, as we shall see, groups that cannot be so covered lead to divergent integrals. If the integrals exist, we can write (3.14) in more compact notation.

$$\int G(cb) d_L b = \int G(b) d_L b \tag{3.15}$$

There is an analogous set of formulas for right translation, *ie.* if

$$\rho_R(\boldsymbol{\beta}_\circ) = \det \left| \frac{\partial f^i(\boldsymbol{\beta}, \boldsymbol{\beta}_\circ^{-1})}{\partial \beta j} \right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_\circ} \tag{3.16}$$

Then

$$\int_{\mathcal{G}} G(\boldsymbol{\beta}')\rho_R(\boldsymbol{\beta})d\boldsymbol{\beta} = \int_{\mathcal{G}} G(\boldsymbol{\beta})\rho_R(\boldsymbol{\beta})d\boldsymbol{\beta}, \tag{3.17}$$

where now $\boldsymbol{\beta}' = \boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\gamma})$ for any fixed $\boldsymbol{\gamma}$, and

$$\int_{\mathcal{G}} G(bc)d_R b = \int_{\mathcal{G}} G(b)d_R b. \tag{3.18}$$

The combinations $d_L b \to \rho_L(\boldsymbol{\beta})d\boldsymbol{\beta}$ and $d_R b \to \rho_R(\boldsymbol{\beta})d\boldsymbol{\beta}$ are called the left- and right-invariant measures respectively. The sums over discrete groups (3.7) lead us to hope that they would be equal as in (3.8). This is not always true as the next example shows.

Example 3.7 *Invariant measures on a two parameter group*

Consider the group of 2×2 matrices

$$A(\boldsymbol{\alpha}) = \begin{vmatrix} e^{\alpha^1} & \alpha^2 \\ 0 & 1 \end{vmatrix}$$

In the notation of (3.10) the composition function for left translation is

$$\beta^1 = \beta^1_\circ + \alpha^1 \qquad \beta^2 = \alpha^2 e^{\beta^1_\circ} + \beta^2_\circ,$$

so that

$$\boldsymbol{\beta}_\circ^{-1} = \left(-\beta^1_\circ, -\beta^2_\circ e^{-\beta^1_\circ}\right).$$

Equation (3.11) becomes

$$\alpha^1 = -\beta^1_\circ + \beta^1 \qquad \alpha^2 = (\beta^2 - \beta^2_\circ)e^{-\beta^1_\circ},$$

so

$$\rho_L(\boldsymbol{\beta}_\circ) = \det \begin{vmatrix} 1 & 0 \\ 0 & e^{-\beta^1_\circ} \end{vmatrix} = e^{-\beta^1_\circ}.$$

Finally, $\rho_L(\boldsymbol{\beta}) = e^{-\beta^1}$.

Repeating the calculation for right translations gives

$$\beta^1 = \alpha^1 + \beta^1_\circ \qquad \beta^2 = \beta^2_\circ e^{\alpha^1} + \alpha^2$$

$$\alpha^1 = \beta^1 - \beta^1_\circ \qquad \alpha^2 = -\beta^2_\circ e^{-\beta^1_\circ + \beta^1} + \beta^2$$

$$\rho_R(\boldsymbol{\beta}_\circ) = \det \left| \begin{matrix} 1 & 0 \\ -\beta^2_\circ e^{-\beta^1_\circ + \beta^1} & 1 \end{matrix} \right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_\circ} = 1$$

3.3.1 Compact groups

Theorem (3.9) leaves two open questions: (1) When do the integrals converge? (2) When are the left- and right-invariant measures equal? Although the two questions are somewhat independent, it turns out that for *compact* groups, a category that includes most of the groups of interest in physics, the two measures are equal *and* the integrals converge.

In the context of real variables, a set is said to be compact if it is closed and bounded. The following statements are consequences:

1) If the set S is compact, every infinite countable sequence contains a subsequence that converges to an element in S.

2) If S is compact it can be covered by a finite number of open sets. (The Heine-Borel theorem.)

These concepts can be applied to groups so long as some norm or distance function is applicable. For example, if the group is described in terms of parameter vectors, then we can define

$$\|\boldsymbol{\alpha} - \boldsymbol{\beta}\| = d(\boldsymbol{\alpha}, \boldsymbol{\beta}), \tag{3.19}$$

where d is the distance function (3.2). A sequence of vectors $\{\boldsymbol{\alpha}_j\}$ converges if for any $\epsilon > 0$ there is an integer N such that $\|\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j\| < \epsilon$ for all $i, j > N$. The group is bounded if there is a constant M such that $\|\boldsymbol{\alpha}\| = d(\boldsymbol{\alpha}, \mathbf{0}) < M$ for all $\boldsymbol{\alpha} \in \mathcal{G}$. If a matrix $M(a)$ is used to represent $a \in \mathcal{G}$, then

$$\|M\| = \text{Max}|M_{ij}| \quad i, j \leq n \tag{3.20}$$

The sequence of matrices $\{M^{(k)}\}$ converges if each element $\{M^{(k)}_{ij}\}$ converges in the usual sense.

If the group $\mathcal{G}$ is compact, then integrals of the form (3.14) and (3.17) will converge, because the domain of integration is bounded. Under the same circumstances, the left- and right-invariant measures are equal as shown by the following theorem:

Theorem 3.10 *If $\mathcal{G}$ is compact then $\rho_L(\boldsymbol{\beta})$ defined by (3.13) and $\rho_R(\boldsymbol{\beta})$ defined by (3.16) are equal.*

Proof: Consider the transformation $\boldsymbol{\alpha} \to \boldsymbol{\alpha}'$ consisting of a left translation by $\boldsymbol{\beta}_\circ$ followed by a right translation by $\boldsymbol{\beta}_\circ^{-1}$.

$$\boldsymbol{\alpha}' = \boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\beta}_\circ^{-1}) = \boldsymbol{f}(\boldsymbol{f}(\boldsymbol{\beta}_\circ, \boldsymbol{\alpha}), \boldsymbol{\beta}_\circ^{-1})$$

$$d\boldsymbol{\alpha} = \det\left|\frac{\partial \alpha^i}{\partial \alpha'^j}\right| d\boldsymbol{\alpha}' = \det\left|\frac{\partial \alpha^i}{\partial \beta^j}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_\circ} \det\left|\frac{\partial \beta^k}{\partial \alpha'^l}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_\circ} d\boldsymbol{\alpha}'$$

$$\det\left|\frac{\partial\beta^k}{\partial\alpha'^l}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_\circ} = \det\left|\frac{\partial f^k(\boldsymbol{\beta},\boldsymbol{\beta}_\circ^{-1})}{\partial\beta^l}\right|^{-1}_{\boldsymbol{\beta}=\boldsymbol{\beta}_\circ} = \rho_R^{-1}(\boldsymbol{\beta}_\circ)$$

Let

$$\eta = \rho_L(\boldsymbol{\beta}_\circ)\rho_R^{-1}(\boldsymbol{\beta}_\circ),$$

then

$$d\boldsymbol{\alpha} = \eta d\boldsymbol{\alpha}'.$$

The *volume* of the group is

$$V(\mathcal{G}) = \int_{\mathcal{G}} d\boldsymbol{\alpha} = \int_{\mathcal{G}} \eta d\boldsymbol{\alpha}'. \tag{3.21}$$

Now perform this "round-trip" translation n times. The resulting volume integral has the same form as (3.21) but with η replaced by η^n. The equality must hold, in fact, even in the limit $n \to \infty$, because of property (1) of compact sets. If $\eta < 1$ or $\eta > 1$ we get the contradictory result that $V(\mathcal{G}) = 0$ or $V(\mathcal{G}) = \infty$. We are forced to conclude that $\eta = 1$, and thus

$$\rho_L(\boldsymbol{\beta}_\circ) = \rho_R(\boldsymbol{\beta}_\circ) \tag{3.22}$$

Example 3.1 fails this test because $\boldsymbol{\alpha}$ is unbounded, and the group is therefore noncompact. There are many noncompact groups that do have equal left and right measures, however. The additive group of real numbers falls into this category as in fact do all noncompact abelian groups.

Our formalism must be modified slightly in the case of compact groups that consist of several disconnected components. The group integral must be replaced by a sum of integrals over the various components.

$$\int_{\mathcal{G}} \to \sum_i \int_{\mathcal{G}_i}$$

With this modification the integrals are still invariant in the sense of Theorem 3.9.

It can be shown (Price, Dynkin and Oniscik) that if $\mathcal{G}$ is a compact linear Lie group, every element in a connected component $\mathcal{G}_i$ can be written in exponential form with a single parameterization. This provides an *a posteriori* justification for the assumption made in proving Theorem 3.9 that the group could be covered by a single parameterization.

In Section 2.5 we proved the simple result that unitary operators on finite-dimensional spaces have unitary matrix representations. The invariant integral provides a tool for proving a much stronger theorem for compact groups.

Theorem 3.11 *Let $\mathcal{O}$ be an operator representation of the compact Lie group $\mathcal{G}$ on the finite-dimensional vector space S. Then $\mathcal{O}$ is equivalent to a unitary representation on S.*

Proof: A unitary representation is one for which

$$< \mathcal{O}\Phi|\mathcal{O}\Psi > =< \Phi|\Psi >$$

for arbitrary $\Psi, \Phi \in S$. Even if this is not true for the natural scalar product $< | >$, we can use invariant integrtion to define a new scalar product such that

$$(\mathcal{O}\Phi, \mathcal{O}\Psi) = (\Phi, \Psi).$$

Let $\mathcal{O}(a)$ be an operator representation. Define

$$(\Phi, \Psi) = V^{-1}(\mathcal{G}) \int_{\mathcal{G}} < \mathcal{O}(a)\Phi|\mathcal{O}(a)\Psi > da, \tag{3.23}$$

where the group volume

$$V(\mathcal{G}) = \int_{\mathcal{G}} da = \int_{\mathcal{G}} d\boldsymbol{\alpha} = \int_{\mathcal{G}} \rho(\boldsymbol{\beta}) d\boldsymbol{\beta}.$$

The scalar product inside the integral is a continuous function of a and $\mathcal{G}$ is compact, so the integral is finite and the left- and right-invariant measures are equal. The new scalar product is positive definite, because the weight functions are positive definite. The normalization insures that $(,) = 1$ if $< | >= 1$. Unitarity is now a simple consequence of the invariance theorem.

$$(\mathcal{O}(b)\Phi, \mathcal{O}(b)\Psi) = V^{-1}(\mathcal{G}) \int_{\mathcal{G}} < \mathcal{O}(ba)\Phi|\mathcal{O}(ba)\Psi > da$$

$$= V^{-1}(\mathcal{G}) \int_{\mathcal{G}} < \mathcal{O}(a)\Phi|\mathcal{O}(a)\Psi > da = (\Phi, \Psi)$$

This shows that $\mathcal{O}$ is unitary with respect to the new scalar product. Now let $\{\omega_i\}$ be an orthonormal basis with respect to the new scalar product, and let $\{\chi_i\}$ be an orthonormal basis with respect to the old scalar product.

$$(\omega_i, \omega_j) = \delta_{ij} \qquad < \chi_i|\chi_j >= \delta_{ij}$$

$$\Phi = \sum \alpha_i \omega_i \qquad \Psi = \sum \beta_i \omega_i$$

Define the nonsingular operator $\mathcal{S}$ such that $\mathcal{S}\omega_i = \chi_i$.

$$< \mathcal{S}\Phi | \mathcal{S}\Psi >= \sum \alpha_i \beta_j^* < \mathcal{S}\omega_i | \mathcal{S}\omega_j >= \sum \alpha_i \beta_j^* < \chi_i | \chi_j >= \sum \alpha_i \beta_i^*$$

But

$$(\Phi, \Psi) = \sum \alpha_i \beta_j^* (\omega_i, \omega_j) = \sum \alpha_i \beta_i^*,$$

so $< \mathcal{S}\Phi | \mathcal{S}\Psi >= (\Phi, \Psi)$. Now

$$< \mathcal{S}\mathcal{O}\mathcal{S}^{-1}\Phi | \mathcal{S}\mathcal{O}\mathcal{S}^{-1}\Psi >= (\mathcal{O}\mathcal{S}^{-1}\Phi, \mathcal{O}\mathcal{S}^{-1}\Psi) = (\mathcal{S}^{-1}\Phi, \mathcal{S}^{-1}\Psi) =< \Phi | \Psi >$$

Thus the operator $\mathcal{O}' = \mathcal{S}\mathcal{O}\mathcal{S}^{-1}$ is unitary on S, and its matrix representations will also be unitary.

Unitary matrix representations have an important property that was alluded to in Section 2.5.

Theorem 3.12 *Reducible unitary matrix representations are completely reducible.*

Proof: Suppose that a similarity transformation brings a matrix representation into the form (2.53) for all $a \in \mathcal{G}$. The representation space S can be partitioned into matching subspaces S_1 and S_2 as follows:

$$M(a)S = \left| \begin{array}{cc} M_{11} & M_{12} \\ 0 & M_{22} \end{array} \right| \left| \begin{array}{c} S_1 \\ S_2 \end{array} \right| = \left| \begin{array}{c} M_{11}S_1 + M_{12}S_2 \\ M_{22}S_2 \end{array} \right|$$

Reducibility means that the subspace S_2 is invariant; M maps S_2 into S_2. If M is also completely reducible, then S_1 is also invariant. One way of saying this is that the following matrix product vanishes.

$$\left| \begin{array}{cc} S_1 & 0 \end{array} \right| \left| \begin{array}{cc} M_{11} & 0 \\ 0 & M_{22} \end{array} \right| \left| \begin{array}{c} 0 \\ S_2 \end{array} \right| = 0$$

Restated in scalar product notation this says that $< \Phi | \mathcal{O}(a)\Psi >= 0$ for all $a \in \mathcal{G}$, all $\Psi \in S_2$, and all $\Phi \in S_1$. If $\mathcal{O}$ is unitary, however,

$$0 =< \Phi | \mathcal{O}(a^{-1})\Psi >=< \mathcal{O}(a)\Phi | \Psi >$$

The first scalar product is zero because $\mathcal{O}(a^{-1})$, like $\mathcal{O}(a)$, maps S_2 into S_2. The fact that the second scalar product is also zero means that $\mathcal{O}(a)$ maps

S_1 into S_1, which is to say that $\mathcal{O}$ as well as its matrix representations are completely reducible.

It might happen that M_{11} and/or M_{22} can be reduced further, *ie.* the spaces S_1 and/or S_2 might have invariant subspaces. If this is the case we continue with further similarity transformations until M is in the form of (2.56).

3.4 The Lie Algebra

The previous section explored the connection between the local group and the structure of the group as a whole. This section describes the relationship between the local group and the Lie algebra. We will assume that the group elements have a well defined local parameterization. In order to conform with standard notation the components of the parameter vectors will be written with upper indices, $\boldsymbol{\alpha} = (\alpha^1, \ldots, \alpha^n)$. We will use the Einstein summation convention in which repeated upper and lower indices are summed over.

Let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ be two elements close to the identity. The composition function can be expanded in powers of the components, α^i and β^i.

$$f^i(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \alpha^i + \beta^i + c^i_{jk}\alpha^j\beta^k + \cdots \tag{3.24}$$

This expansion satisfies the group requirements,

$$\boldsymbol{f}(\boldsymbol{\alpha}, \mathbf{0}) = \boldsymbol{\alpha} \qquad \boldsymbol{f}(\boldsymbol{\beta}, \mathbf{0}) = \boldsymbol{\beta}$$

The constants c^i_{jk} uniquely specify the expansion through second order. The associativity condition requires that

$$\boldsymbol{f}(\boldsymbol{\gamma}, \boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\beta})) = \boldsymbol{f}(\boldsymbol{f}(\boldsymbol{\gamma}, \boldsymbol{\alpha}), \boldsymbol{\beta}). \tag{3.25}$$

(See Definition 3.1(2).) Differentiating with respect to γ gives

$$\frac{\partial f^i(\boldsymbol{\gamma}, \boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\beta}))}{\partial \gamma^k} = \frac{\partial f^i(\boldsymbol{f}(\boldsymbol{\gamma}, \boldsymbol{\alpha}), \boldsymbol{\beta}))}{\partial f^j(\boldsymbol{\gamma}, \boldsymbol{\alpha})} \frac{\partial f^j(\boldsymbol{\gamma}, \boldsymbol{\alpha})}{\partial \gamma^k}. \tag{3.26}$$

Set $\boldsymbol{\gamma} = 0$ and define

$$u^i_j(\boldsymbol{\alpha}) = \left.\frac{\partial f^i(\boldsymbol{\gamma}, \boldsymbol{\alpha})}{\partial \gamma^j}\right|_{\gamma=0} \qquad u^i_j \Psi^k_i = \delta^k_j \tag{3.27}$$

Then (3.26) becomes

$$\frac{\partial f^i(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha^j} = u^i{}_k(\boldsymbol{f}) \Psi^k{}_j(\boldsymbol{\alpha}) \tag{3.28}$$

Partial differential equations of this form have unique solutions if and only if all mixed derivatives are equal.

$$\frac{\partial^2 f^i}{\partial \alpha^l \partial \alpha^j} = \frac{\partial^2 f^i}{\partial \alpha^j \partial \alpha^l} \tag{3.29}$$

Applying this to (3.28) yields

$$\frac{\partial}{\partial\alpha^l}\left\{u^i_k(\boldsymbol{f})\Psi^k_j(\boldsymbol{\alpha})\right\} = \frac{\partial}{\partial\alpha^j}\left\{u^i_k(\boldsymbol{f})\Psi^k_l(\boldsymbol{\alpha})\right\}$$

We now carry out the differentiation and use (3.28) to eliminate the derivitives of $\boldsymbol{f}$. Some straightforward algebra leads to the following form of the uniqueness condition:

$$\left[\frac{\partial\Psi^c_j(\boldsymbol{\alpha})}{\partial\alpha^l} - \frac{\partial\Psi^c_l(\boldsymbol{\alpha})}{\partial\alpha^j}\right] u^l_a(\boldsymbol{\alpha})u^j_b(\boldsymbol{\alpha}) = \tag{3.30}$$

$$\left[\frac{\partial u^i_a(\boldsymbol{f})}{\partial f^m}u^m_b(\boldsymbol{f}) - \frac{\partial u^i_b(\boldsymbol{f})}{\partial f^m}u^m_a(\boldsymbol{f})\right]\Psi^c_i(\boldsymbol{f}) = f^c_{ab}$$

Since $\boldsymbol{\alpha}$ and $\boldsymbol{f}$ are independent variables, the right and left sides of (3.30) can be set equal to the constant f^c_{ab}. This constant can be calculated as follows: in the vicinity of the identity

$$u^i_j(\boldsymbol{\alpha}) \approx \delta^i_j + c^i_{jl}\alpha^l$$

$$\Psi^i_j(\boldsymbol{\alpha}) \approx \delta^i_j - c^i_{jl}\alpha^l,$$

where c^i_{jl} is the coefficient of the quadratic term in (3.24). Inserting Ψ into the left side of (3.30) and setting $\boldsymbol{\alpha} = 0$ gives

$$f^c_{ab} = c^c_{ab} - c^c_{ba} \tag{3.31}$$

We now define

$$X_a = -u^k_a(\boldsymbol{\beta})\frac{\partial}{\partial\beta^k}. \tag{3.32}$$

Setting $\boldsymbol{\beta} = 0$ in the right side of (3.30) and rearranging terms gives

$$[X_a, X_b] = f^c_{ab}X_c. \tag{3.33}$$

The I_a, called *infinitesimal group generators*, form a Lie algebra according to Definition ???

3.4.1 Local Transformation Groups

In Chapter 1 we showed that infinitesimal coordinate system rotations can be described in terms of three differential operators called the infinitesimal generators. These operators acting on functions of position coordinates induce the corresponding changes in the functions. The generators form a Lie

algebra whose commutation relations are determined by the structure of the group. In this section we show that this result is quite general and that it provides an alternate route to the Lie algebra.

Consider a general transformation of coordinate systems. Let x^i $i = 1, \ldots, N$ be the cordinates of a point expressed in terms of the coordinate system S, and let x'^i be the coordinates of the same point in the transformed coordinate system S'. The transformation from S to S' is a group operation parameterized by $\boldsymbol{\alpha} = (\alpha^1, \ldots, \alpha^n)$. The effect of this transformation on the components is given by the function,

$$x'^i = F^i(\boldsymbol{\alpha}, \boldsymbol{x}) \;\; i = 1, \ldots, N. \tag{3.34}$$

This is a generalization of the matrix transformation (1.5), but it allows for the possibility that F^i might be non-linear and/or inhomogeneous. The F^i's, called coordinate transformation functions, can form a local Lie group.

Definition 3.17 *Local Lie Transformation Group*

Let $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ be members of a local Lie group, and let $\boldsymbol{x}$ be a member of some real vector space V. The coordinate transformation functions, $x'^i = F^i(\boldsymbol{\alpha}, \boldsymbol{x})$, $i = 1, \ldots, N$, form a local Lie group if the following requirements are met:

(a) F^i is a real analytic function of its $n + N$ arguments.

(b) $F^i(\mathbf{0}, \boldsymbol{x}) = x^i$ for all $\boldsymbol{x} = (x^1, \ldots, x^N) \in V$.

(c) $F^i(\boldsymbol{\beta}, \boldsymbol{F}(\boldsymbol{\alpha}, \boldsymbol{x})) = F^i(\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha}), \boldsymbol{x})$ for all $\boldsymbol{x} \in V$.

We can now generalize the derivation of equations (1.14) and (1.33) so that they can be used for general coordinate transfromations. Let $\boldsymbol{x}$ refer to a point in space that has coordinates x^i in S and x'^i in S'. The function $\psi(\boldsymbol{x})$ must have the same numerical value in both coordinate systems, so

$$\psi'(x'^i) = \psi(x^i). \tag{3.35}$$

Since $\boldsymbol{\alpha}$ is a member of a local Lie group it must have an inverse,

$$x^i = F^i(\boldsymbol{\alpha}^{-1}, \boldsymbol{x}'),$$

and

$$\psi'(x'^i) = \psi(F^i(\boldsymbol{\alpha}^{-1}, \boldsymbol{x}')). \tag{3.36}$$

Close to the identity the group element $\delta\boldsymbol{\alpha}$ has an inverse $(\delta\boldsymbol{\alpha})^{-1} = -\delta\boldsymbol{\alpha}$. Expanding F^i to first order in this parameter gives

$$x^i \;\; = \;\; F^i(-\delta\boldsymbol{\alpha}, \boldsymbol{x}')$$

$$
\begin{aligned}
&= F^i(\mathbf{0}, \boldsymbol{x}') + \left.\frac{\partial F^i(\boldsymbol{\beta}, \boldsymbol{x}')}{\partial \beta^j}\right|_{\boldsymbol{\beta}=\mathbf{0}} (-\delta\alpha^j) + \cdots \\
&\approx x'^i - \delta\alpha^j \left.\frac{\partial F^i(\boldsymbol{\beta}, \boldsymbol{x}')}{\partial \beta^j}\right|_{\boldsymbol{\beta}=\mathbf{0}}.
\end{aligned}
$$

Substituting this into (3.36) and again expanding in $\delta\boldsymbol{\alpha}$ we find

$$
\begin{aligned}
\psi'(\boldsymbol{x}') &= \psi\left(\boldsymbol{x}' - \delta\alpha^j \left.\frac{\partial \boldsymbol{F}(\boldsymbol{\beta}, \boldsymbol{x}')}{\partial \beta^j}\right|_{\boldsymbol{\beta}=\mathbf{0}}\right) \\
&\approx \psi(\boldsymbol{x}') - \delta\alpha^j \left.\frac{\partial F^i(\boldsymbol{\beta}, \boldsymbol{x}')}{\partial \beta^j}\right|_{\boldsymbol{\beta}=\mathbf{0}} \frac{\partial}{\partial x'^i}\psi(\boldsymbol{x}')
\end{aligned}
$$

To lowest order, then, the change in ψ is

$$
\psi'(\boldsymbol{x}) - \psi(\boldsymbol{x}) = \delta\alpha^j X_j(\boldsymbol{x})\psi(\boldsymbol{x}), \tag{3.37}
$$

which uses the definitions,

$$
v_j^i(\boldsymbol{x}) = \left.\frac{\partial F^i(\boldsymbol{\beta}, \boldsymbol{x})}{\partial \beta^j}\right|_{\boldsymbol{\beta}=\mathbf{0}} \tag{3.38}
$$

and

$$
X_j(\boldsymbol{x}) = -v_j^i(\boldsymbol{x})\frac{\partial}{\partial x^i} \tag{3.39}
$$

The X_j are called the *generators of infinitesimal displacements*. They are the analogs for general coordinate transformations of the angular momentum operators in (1.33). These operators, together with the definition of their commutator

$$
[X_i, X_j] = X_i X_j - X_j X_i, \tag{3.40}
$$

form a Lie algebra according to Definition 1.5. This is guaranteed by the following two theorems usually called Lie's first and second theorems.

Theorem 3.13 *If F_i is a coordinate transformation function of a local Lie transformation group, then*

$$
\frac{\partial x'^i}{\partial \alpha^j} = \Psi_j^k(\boldsymbol{\alpha})v_k^i(\boldsymbol{x}'). \tag{3.41}
$$

Proof: The proof is based on the associativity condition (c) in Definition 3.17. Suppose $\boldsymbol{\beta}$ in (c) is replaced with an infinitesimal transformation $\delta\boldsymbol{\alpha}$, so that $\boldsymbol{\alpha}$ corresponds to the transformation from S to S', and $\delta\boldsymbol{\alpha}$ to the transformation from S' to S''. The difference

$$dx'^i = x''^i - x'^i$$

can be computed in two different ways. First,

$$x''^i = x'^i + dx'^i = F^i(\delta\boldsymbol{\alpha}, \boldsymbol{x}'),$$

$$dx'^i = \delta\alpha^k \left.\frac{\partial F^i(\boldsymbol{\beta}, \boldsymbol{x}')}{\partial \beta^k}\right|_{\boldsymbol{\beta}=\mathbf{0}} = \delta\alpha^k v_k^i(\boldsymbol{x}').$$

We could also transform from $S \to S''$ in a single step using the parameter vector

$$\boldsymbol{\alpha} + d\boldsymbol{\alpha} = \boldsymbol{f}(\delta\boldsymbol{\alpha}, \boldsymbol{\alpha}),$$

where

$$d\alpha^l = \delta\alpha^j \left.\frac{\partial f^l(\boldsymbol{\beta}, \boldsymbol{\alpha})}{\partial \beta^j}\right|_{\boldsymbol{\beta}=\mathbf{0}} = \delta\alpha^j u_l^j(\boldsymbol{\alpha}).$$

The inverse of the last equation is

$$\delta\alpha^k = d\alpha^j \Psi_j^k(\boldsymbol{\alpha}).$$

Finally

$$dx'^i = d\alpha^j \Psi_j^k(\boldsymbol{\alpha}) v_k^i(\boldsymbol{x}')$$

$$\frac{\partial x'^i}{\partial \alpha^j} = \Psi_j^k(\boldsymbol{\alpha}) v_k^i(\boldsymbol{x}'). \tag{3.42}$$

Theorem 3.14 *The $X_j(\boldsymbol{x})$ defined by (3.39) together with the commutator defined by (3.40) constitute a Lie algebra,* i.e.

$$[X_a, X_b] = f_{ab}^c X_c$$

with f_{ab}^c constant.

Proof: Equation (3.42) is a set of coupled first-order partial differential equations, which if integrated with initial conditions $x^i = F^i(\mathbf{0}, \boldsymbol{x})$ would give the coordinate transformation functions (3.34). A necessary and sufficient condition that an equation of the form (3.42) has a unique solution is that all the mixed derivatives are equal:

$$\frac{\partial^2 x'^i}{\partial \alpha^j \partial \alpha^l} = \frac{\partial^2 x'^i}{\partial \alpha^l \partial \alpha^j}.$$

By steps that are analogous to the derivation of (3.30) we arrive at

$$v_k^i(\boldsymbol{x}')\left[\frac{\partial\Psi_j^k(\boldsymbol{\alpha})}{\partial\alpha^l}-\frac{\partial\Psi_l^k(\boldsymbol{x}')}{\partial\alpha^j}\right]u_a^l(\boldsymbol{\alpha})u_b^j(\boldsymbol{\alpha}) \tag{3.43}$$

$$=\left[\frac{\partial v_a^i(\boldsymbol{x}')}{\partial x'^m}v_b^m(\boldsymbol{x}')-\frac{\partial v_b^i(\boldsymbol{x}')}{\partial x'^m}v_a^m(\boldsymbol{x}')\right].$$

This equation is not exactly analogous to (3.30) because v_k^i is not a square matrix and hence not invertable. We can use the definition of f_{ab}^c from (3.30), however, to obtain (after setting $\boldsymbol{\alpha}$=$\mathbf{0}$)

$$v_k^i(\boldsymbol{x})f_{ab}^k=\left[\frac{\partial v_b^i(\boldsymbol{x})}{\partial x^m}v_a^m(\boldsymbol{x})-\frac{\partial v_a^i(\boldsymbol{x})}{\partial x^m}v_b^m(\boldsymbol{x})\right]. \tag{3.44}$$

Substituting (3.39) into (3.44) gives us another version of the Lie algebra

$$[X_a,X_b]=f_{ab}^cX_c. \tag{3.45}$$

3.4.2 Local Linear Lie Groups

The formulation of Lie groups and algebras in terms of parameter vectors has the advantage of generality; it provides a conceptual framework for studying groups without specifying the actual form of the group elements. It has the drawback, however, of being awkward computationally. For example, it is often impossible to find an explicit, closed-form expression for the composition functions. For this reason, most applications of Lie groups in physics make use of matrix groups or groups of transformation operators. Transformation groups lead to Lie algebras whose elements are partial derivatives as seen in the preceeding section. In this section we discuss local linear Lie groups, which are matrix groups defined in a neighborhood of the identity. There is a deep result known as Ado's theorem (Jacobson, 1962) that every finite-dimensional Lie algebra is isomorphic to a matrix Lie algebra. Thus every Lie group (with a finite number of parameters) is at least locally isomorphic to one of these matrix groups. Local linear groups therefore bridge the gap between matrix representations (Chapter 2) and the Lie algebra.

Definition 3.18 *Local Linear Lie Group*

Let $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ be members of a set of n-tuples with a composition function that satisfies the requirements of Definition 3.2 through 3.4 for a local Lie group in a neighborhood U of the identity. A local linear Lie

group is a set of $m \times m$ non-singular matrices $M(\boldsymbol{\alpha})$ satisfying the following requirements:

(1) $M(\boldsymbol{e}) = I_m$ (the unit matrix)

(2) The matrix elements of $M(\boldsymbol{\alpha})$ are analytic functions of the parameters $\alpha^1, \ldots, \alpha^n$. For all $\boldsymbol{\alpha} \in U$ the mapping from $\boldsymbol{\alpha}$ to $M(\boldsymbol{\alpha})$ is one to one.

(3) If $\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \boldsymbol{\gamma}$ and $\boldsymbol{f}$ satisfies the requirements of a local Lie group, then $M(\boldsymbol{\alpha})M(\boldsymbol{\beta}) = M(\boldsymbol{\gamma})$, where the operation on the left is matrix multiplication.

Local linear Lie groups have their own Lie algebra, which is obtained through the following definition:

Definition 3.19 *Infinitesimal matrix generators*

Let $M(\boldsymbol{\alpha})$ be a set of matrices that satisfy the postulates of a local linear Lie group. The infinitesimal matrix generators are given by

$$X_i = \left.\frac{\partial M(\boldsymbol{\alpha})}{\partial \alpha^i}\right|_{\boldsymbol{\alpha}=\mathbf{0}} \tag{3.46}$$

Theorem 3.15 *The X_i form a Lie algebra*

Proof: Associativity requires that

$$M(\boldsymbol{\gamma})M(\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\beta})) = M(\boldsymbol{f}(\boldsymbol{\gamma}, \boldsymbol{\alpha}))M(\boldsymbol{\beta})$$

Differentiate with respect to γ^i, and then set $\boldsymbol{\gamma} = \mathbf{0}$.

$$X_k M(\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\beta})) = \frac{\partial M(\boldsymbol{\alpha})}{\partial \alpha^j} u_k^j(\boldsymbol{\alpha}) M(\boldsymbol{\beta})$$

$$\frac{\partial M(\boldsymbol{\alpha})}{\partial \alpha^j} M(\boldsymbol{\beta}) = \Psi_j^k(\boldsymbol{\alpha}) X_k M(\boldsymbol{f}(\boldsymbol{\alpha}, \boldsymbol{\beta})) \tag{3.47}$$

As usual we require that mixed derivitaves be equal

$$\frac{\partial^2 M(\boldsymbol{\alpha})}{\partial \alpha^l \partial \alpha^j} M(\boldsymbol{\beta}) = \frac{\partial^2 M(\boldsymbol{\alpha})}{\partial \alpha^j \partial \alpha^l} M(\boldsymbol{\beta})$$

The now-familiar steps lead to an equation analogous to (3.30),

$$\left[\frac{\partial \Psi_j^k(\boldsymbol{\alpha})}{\partial \alpha^l} - \frac{\partial \Psi_l^k(\boldsymbol{\alpha})}{\partial \alpha^j}\right] u_b^j(\boldsymbol{\alpha}) u_a^l(\boldsymbol{\alpha}) X_k = \tag{3.48}$$

$$\left[X_a \frac{\partial M(\boldsymbol{f})}{\partial f^m} u_b^m(\boldsymbol{f}) - X_b \frac{\partial M(\boldsymbol{f})}{\partial f^m} u_a^m(\boldsymbol{f})\right] M^{-1}(\boldsymbol{f})$$

Substituting (3.30) and (3.39) and setting $\boldsymbol{f} = \mathbf{0}$ yields

$$[X_a, X_b] = f_{ab}^c X_c. \tag{3.49}$$

The constants come from (3.24) and (3.31).

One might object that we have used the symbol X_i in three different ways: infinitesimal group generators (3.32), generators of infinitesimal transformations (3.39), and finally infinitesimal matrix generators (3.46). In fact, these objects all form Lie algebras with the same structure constants, thus from the point of view of abstract algebras, Definition ???, they are identical. Many theorems apply equally to all three kinds of operators. The following theorem is specific to the matrix generators, however.

Theorem 3.16 *Let $M(\boldsymbol{\alpha})$ be denote the members of a local linear Lie group according to Definition 3.18. Then the matrices $X_i \;\; i = 1, \ldots, n$ defined by (3.46)*

$$(X_i)_{jk} = \left.\frac{\partial M_{jk}}{\partial \alpha^i}\right|_{\boldsymbol{\alpha}=\mathbf{0}}$$

form the basis of a n dimensional real *vector space.*

Proof: We assume that $M(\boldsymbol{\alpha})$ is a $m \times m$ complex matrix and decompose it into two real matrices, $M_{jk} = B_{jk} + \imath C_{jk}$. The B's and C's together form a set of $2m^2$ real functions of n real variables, $\alpha^1, \ldots, \alpha^n$. Since the mapping between $\boldsymbol{\alpha}$ and M is one to one it follows that we can pick out a set of n independent functions from this set of $2m^2$ functions and express the remaining $2m^2 - n$ B's and C's as functions of this select set. Call the n independent functions $D_1, \ldots, D_n$. Independence implies that the Jacobian

$$\det \left[\frac{\partial D_i}{\partial \alpha_j}\right]_{\boldsymbol{\alpha}=\mathbf{0}} \neq 0$$

This in turn implies that the system of n homogeneous linear equations

$$\lambda^j \left.\frac{\partial D_i}{\partial \alpha_j}\right|_{\boldsymbol{\alpha}=\mathbf{0}} = 0 \quad i = 1, \ldots, n$$

has only the trivial solution $\lambda^1 = \lambda^2 = \cdots = \lambda^n = 0$.

If the matrices $X_1, X_2, \ldots X_n$ were not all independent, then there would be a non-trivial solution to the system of $2m^2$ homogeneous linear equations

$$\lambda^p X_p = \lambda_p \left.\frac{\partial M_{ij}}{\partial \alpha^p}\right|_{\boldsymbol{\alpha}=\mathbf{0}} = 0 \quad i,j = 1,\ldots,m$$

However, n equations out of this set, those for which the real or imaginary part of M_{ij} is one of the D's, have already been shown to have no non-trivial solutions, Consequently there is no non-trivial solution to the complete set of equations for λ real, and the X_p's form an n-dimensional basis of a real vector space.

This argument does *not* prove that the X_i's themselves are real, only that they span a n dimensional vector space with real coefficients. Any matrix group with complex elements will have complex generators, and there are many cases in which the X_i's are independent over the real number field yet not independent over the complex numbers.

3.4.3 Parameter transformations

A group can be parameterized in many different ways. It is important to distinguish, therefore, between those aspects of the Lie algebra that are "inconsequential" in the sense that they depend on the particular parameterization and those that are "crucial" in that they are invariant under reparameterization of the group. Consider a general parameter transformation function

$$\boldsymbol{\alpha}'^i = \phi^i(\boldsymbol{\alpha}). \tag{3.50}$$

This can be thought of as an automorphism in which each abstract group element is mapped into itself, $a \to a$, $b \to b$, *etc.* but the corresponding parameters and composition functions are mapped isomorphically, $\boldsymbol{\alpha} \to \boldsymbol{\alpha}'$, $\boldsymbol{f}(\boldsymbol{\beta},\boldsymbol{\alpha}) \to \boldsymbol{f}'(\boldsymbol{\alpha}',\boldsymbol{\beta}')$. The composition function must retain the general form of (3.24), however, which restricts the possible transformations to two types:

(a) Linear transformations

$$\alpha'^i = M^i_j \alpha^j, \tag{3.51}$$

where M is a square, non-singular matrix.

(b) Non-linear transformations

$$\alpha'^i = \alpha^i + O(a^2) \tag{3.52}$$

If (3.51) is to preserve the multiplication table the transformed parameters must obey the rule

$$\boldsymbol{\gamma}' = \boldsymbol{f}'(\boldsymbol{\beta}', \boldsymbol{\alpha}') \tag{3.53}$$

or

$$M^i_j \gamma^j = M^i_j \beta^j + M^i_j \alpha^j + c'^i_{kl} M^k_m \beta^m M^l_n \alpha^n + \cdots \tag{3.54}$$

We have used (3.24) with $c \to c'$ to express the transformed composition function. Evidentally

$$c^d_{ab} = (M^{-1})^d_i \, c'^i_{kl} \, M^k_a M^l_b \tag{3.55}$$

ie. the c^i_{jl} as well as the structure constants and all the other coefficients in (3.24) transform like tensors. In terms of the Lie algebra this can be thought of as a change of basis; for if a new basis is defined by

$$X'_i = X_j (M^{-1})^j_i$$

the commutation relations become

$$[X'_i, X'_j] = f'^k_{ij} X'_k$$

where f' is obtained from (3.31) and (3.55). We have thus shown that all sets of structure constants related by the transformation (3.55) $f \to f'$, describe the same group (at least near the origin). We can use this freedom to choose a basis in which the commutation relations have the maximum possible simplicity and clarity. Chapter 4 is devoted to strategies for doing this.

Transformations of the form (3.52) leave the structure constants invariant because f is calculated at the origin. It will change the higher order coefficients in (3.24) and hence the form of the composition function. This freedom can be exploited to cast the composition function in a particularly convenient form as follows: let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ refer to specfic group elements that are close enough to the origin to have a valid local parameterization. Construct the exponential operators

$$A = \exp \alpha^a X_a \qquad B = \exp \beta^b X_b \tag{3.56}$$

where $\exp X$ is defined by the usual power series $\sum X^n/n!$. The product, $C = BA$, can also be written as an exponential

$$C = \exp \gamma^c X_c, \tag{3.57}$$

and $\boldsymbol{\gamma}$ can be calculated using the BCH formula (2.39)

$$\gamma^p = f^p(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \alpha^p + \beta^p + f^p_{ij} \beta^i \alpha^j + \frac{1}{6} f^p_{iq} f^q_{jk} (\beta^i \beta^j \alpha^k + \alpha^i \alpha^j \beta^k) + \cdots \tag{3.58}$$

The following comments are in order regarding (3.58):

(a) Equations (3.56) through (3.58) are valid for all three definitions of X_i, (3.32), (3.39), and (3.46), since they all have the same commutation relations and the BCH formula is valid for any associative operators.

(b) Equation (3.58) uniquely determines the composition function in terms of the structure constants, which in turn are determined by the structure of the group at the origin. Parameters determined in this way are said to be *normal* or sometimes *canonical*.

(c) We can define a family of group elements,

$$M(t) = \exp t\alpha^a X_a, \tag{3.59}$$

where t is a real variable. We can think of $M(t)$ as a trajectory passing through the identity at $t = 0$ and through the point $\exp \alpha^a X_a$ at $t = 1$. Clearly $[M(t), M(s)] = 0$, so (3.58) gives $M(t)M(s) = M(t + s)$. Such a family of elements is called a *one-parameter subgroup*, and all one-parameter subgroups are abelian.

Example 3.8 *The Euclidean group on a plane.*

The group of all translations and rotations in ordinary flat space is called the Euclidean group. In two dimensions it can be parameterized as follows:

$$\begin{vmatrix} x' \\ y' \end{vmatrix} = \begin{vmatrix} \cos\alpha^1 & \sin\alpha^1 \\ -\sin\alpha^1 & \cos\alpha^1 \end{vmatrix} \begin{vmatrix} x \\ y \end{vmatrix} + \begin{vmatrix} \alpha^2 \\ \alpha^3 \end{vmatrix}$$

Two successive transformations $\boldsymbol{\gamma} = \boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\alpha})$ are described by the composition function

$$\begin{aligned} f^1 &= \beta^1 + \alpha^1 \\ f^2 &= \alpha^2 \cos\beta^1 + \alpha^3 \sin\beta^1 + \beta^2 \\ f^3 &= -\alpha^2 \sin\beta^1 + \alpha^3 \cos\beta^1 + \beta^3 \end{aligned}$$

This can also be treated as a homogeneous matrix group by regarding x and y as elements of a three-component vector.

$$\begin{vmatrix} x' \\ y' \\ 1 \end{vmatrix} = \begin{vmatrix} \cos\alpha^1 & \sin\alpha^1 & \alpha^2 \\ -\sin\alpha^1 & \cos\alpha^1 & \alpha^3 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} x \\ y \\ 1 \end{vmatrix} \tag{3.60}$$

This parameterization is simple and straightforward, but not "normal." To find the normal parameterization, construct the Lie algebra of matrix

generators using (3.46),

$$X_1 = \begin{vmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \quad X_2 = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \quad X_3 = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix},$$

with commutation relations

$$[X_1, X_2] = -X_3 \quad [X_1, X_3] = X_2 \quad [X_2, X_3] = 0$$

The exponential operator (3.56) is

$$M(\boldsymbol{\alpha}) = \exp \alpha^a X_a = \begin{vmatrix} \cos\alpha^1 & \sin\alpha^1 & M_{13} \\ -\sin\alpha^1 & \cos\alpha^1 & M_{23} \\ 0 & 0 & 1 \end{vmatrix} \tag{3.61}$$

where

$$M_{13} = (\alpha^2 \sin\alpha^1 - \alpha^3 \cos\alpha^1 + \alpha^3)/\alpha^1$$
$$M_{23} = (\alpha^3 \sin\alpha^1 + \alpha^2 \cos\alpha^1 - \alpha^2)/\alpha^1.$$

This is the same as (3.60) except that $\alpha^2 \to M_{13}$ and $\alpha^3 \to M_{23}$, and the composition function is correspondingly more complicated. The two parameterizations are identical at the origin, however, because they are based on the same commutation relations.

An interesting question arises arises at this point: how general is this procedure? Can every group element be written as the exponential of a Lie algebra? This must be so for local linear Lie groups as can be seen by the following argument: if $M(\boldsymbol{\alpha})$ is a member of a local linear Lie group, then it must be possible to reach this element through a continuous variation of the parameters starting at the identity. One such path is the straight line $\boldsymbol{\alpha}t$. The corresponding matrix exponential, (3.59), defines the matrix $M(\boldsymbol{\alpha})$ when $t = 1$. On the other hand, a disconnected component of a group cannot, by definition, be reached through a continuous variation of parameters starting at the identity; and so they cannot be written in the form (3.59). But even for connected linear Lie groups, there can be elements that are not expressible as a single exponential as shown in the next example.

Example 3.9 *The group $SL(2, R)$.*

The group $SL(2, R)$ consists of all real 2×2 matrices with unit determinant. One such matrix is

$$M = \begin{vmatrix} r & 0 \\ 0 & r^{-1} \end{vmatrix}$$

The eigenvalues of M are r and r^{-1}. If there were a matrix A such that $e^A = M$, then according to Theorem 2.5 and its corollary, the eigenvalues of A, say λ_1 and λ_2, must satisfy $e^{\lambda_1} = r$ and $e^{\lambda_2} = r^{-1}$. Moreover, since $\det(M) = 1$, $\mathrm{tr}(A) = 0$ and $\lambda_1 = -\lambda_2$. The eigenvalues of a real, traceless matrix, however, must be pure real or pure imaginary. These conditions cannot all be satisfied simultaneously when $r < -1$, so that M cannot be obtained in this range by exponentiation.

What has gone wrong here? The matrix M can be connected to the identity by the following indirect path: The parameterization

$$N(t) = \begin{vmatrix} \cos \pi t & \sin \pi t \\ -\sin \pi t & \cos \pi t \end{vmatrix}$$

connects the identity $N(t = 0) = I_2$ with $N(t = 1) = -I_2$. Then

$$O(s) = -\begin{vmatrix} e^s & 0 \\ 0 & e^{-s} \end{vmatrix}$$

connects $O(s = 0) = -I_2$ with $O(s = \ln |r|) = M$. This suggests that M cannot be connected to the identity without reparameterizing, and for this reason it has no normal parameterization.

Incidentally, this example illustrates the fact that every element in the connected component of a matrix Lie group can be written as the product of a finite number of exponentials [Cornwell, 1984], since both $N(t)$ and $O(s)$ can be obtained by exponentiation. The group is not compact, however, so it cannot be covered with a single exponential.

Chapter 4

The catalog of algebras

We have seen how a Lie group with its non-countably infinite number of elements can be analyzed in terms of a finite structure called the Lie algebra. The Lie algebra in turn is characterized by a set of numbers, the structure constants. This is a step in the right direction, but the structure constants by themselves are difficult to interpret, partly because there are so many of them. If the algebra has dimension d there must be d^3 constants. Of these d^2 must be zero (since $C^k_{ii} = 0$), and half of the remaining $d^3 - d^2$ are redundant, (since $C^k_{ij} = -C^k_{ji}$); but this still leaves $\frac{1}{2}d^2(d-1)$ potentially non-trivial constants. The group $SU(3)$, for example, has an 8-dimensional algebra and 224 structure constants; SU(4) is 15-dimensional, and the number of structure constants is 1575. It is hard to get much insight from staring at a table of numbers of this size. Moreover, the structure constants are not of much use in practical calculations; usually one needs a matrix representation of the algebra. Both problems were studied by E. Cartan, H. Weyl, and others around the beginning of this century. They discovered that most[1] algebras can be characterized by a much smaller set of constants called the *simple roots*. The simple roots are l-dimensional vectors, where l is a small number, $l \leq d$, called the *rank* of the algebra; and each algebra has l simple roots. With these roots one can put the commutation relations into a particularly simple form and construct canonical matrix representations. The extent of the simplification can be seen by comparing the rank of $SU(3)$ and $SU(4)$, two and three respectively, with the number of structure constants given above. It is also possible to catalog all possible (semisimple) algebras. There are only four basic types plus a handfull of exceptional cases. The

[1] By "most" I mean the semisimple algebras, which include virtually all the algebras of interest in physics.

important properties of these algebras can be found tabulated in standard reference works.

Unfortunately, the argument leading up to the root vector decomposition is rather long and technical. We start with some new definitions and terminology.

Definition 4.1 *Subalgebra*
The subset $\mathcal{A}$ of elements of the Lie algebra $\mathcal{L}$ constitutes a ***subalgebra*** *if $[X, Y] \in \mathcal{A}$ for all $X, Y \in \mathcal{A}$. In more compact notation, $[\mathcal{A}, \mathcal{A}] \subseteq \mathcal{A}$.*

Every Lie algebra has two "trivial" subalgebras, the algebra itself and the zero element. A subalgebra that is not trivial in this sense is called a *proper* subalgebra.

Definition 4.2 *Invariant subalgebra*
A subalgebra $\mathcal{A}$ is ***invariant*** *if $[X, Y] \in \mathcal{A}$ for all $X \in \mathcal{A}$ and all $Y \in \mathcal{L}$,* i.e. *$[\mathcal{A}, \mathcal{L}] \subseteq \mathcal{A}$.*

An invariant subalgebra is sometimes called an *ideal*.

Lemma 4.1 *Let $\mathcal{A}$ and $\mathcal{B}$ be invariant subalgebras. Then $[\mathcal{A}, \mathcal{B}]$ is an invariant subalgebra contained in $\mathcal{A}$ and $\mathcal{B}$.*

Proof. Clearly $[\mathcal{A}, \mathcal{B}] \subseteq \mathcal{A} \cap \mathcal{B}$. Using the Jacobi identity we have $[[X, Y], Z] = -[[Y, Z], X] - [[Z, X], Y]$ for each $X \in \mathcal{A}$, $Y \in \mathcal{B}$, and $Z \in \mathcal{L}$. Thus $[[\mathcal{A}, \mathcal{B}], \mathcal{L}] \subseteq [\mathcal{A}, \mathcal{B}]$, proving that $[\mathcal{A}, \mathcal{B}]$ is an invariant subalgebra.

Definition 4.3 *Abelian algebras*
A Lie algebra (or subalgebra) $\mathcal{L}$ is said to be ***abelian*** *if $[\mathcal{L}, \mathcal{L}] = 0$. In any Lie algebra the set of all elements that commute with all the elements of $\mathcal{L}$ forms an abelian subalgebra called the* ***center*** *of the algebra.*

Definition 4.4 *Simple and semisimple algebras*
An algebra with no proper invariant subalgebras is ***simple****. An algebra with no proper invariant abelian subalgebras is* ***semisimple****.*

The presence of an invariant subalgebras can always be used to simplify the representation of the algebra. For this reason we are interested in extracting as many invariant subalgebras as possible. It is clear from the last lemma that

$$[\mathcal{L}, \mathcal{L}] \equiv \mathcal{L}^{(1)} \tag{4.1}$$

is an invariant subalgebra of $\mathcal{L}$. By iterating this procedure we can construct an entire sequence of invariant subalgebras.

$$\mathcal{L} \equiv \mathcal{L}^{(0)} \supseteq \mathcal{L}^{(1)} \supseteq \mathcal{L}^{(2)} \cdots \mathcal{L}^{(n-1)} \supseteq \mathcal{L}^{(n)} \cdots \tag{4.2}$$

$$[\mathcal{L}^{(i)}, \mathcal{L}^{(i)}] \equiv \mathcal{L}^{(i+1)} \tag{4.3}$$

Since the Lie algebra has a finite dimension, this sequence will either terminate on a $\mathcal{L}^{(n)}$ that consists of only the zero element, or it will arrive at a $\mathcal{L}^{(n)}$ that is identical with the preceding $\mathcal{L}^{(n-1)}$.

Definition 4.5 *The sequence (4.2) is called the* ***derived*** *series. If the derived series terminates with the element 0, the algebra is said to be* ***solvable****.*

It is easy to show that every subalgebra of a solvable algebra is solvable and that the direct sum (see the following definitions) of two invariant solvable subalgebras is solvable.

Another sequence of invariant subalgebras is constructed as follows:

$$[\mathcal{L}, \mathcal{L}] \equiv \mathcal{L}^2 \tag{4.4}$$

$$\mathcal{L} \equiv \mathcal{L}^1 \supseteq \mathcal{L}^2 \cdots \mathcal{L}^{n-1} \supseteq \mathcal{L}^n \cdots \tag{4.5}$$

$$[\mathcal{L}^i, \mathcal{L}] \equiv \mathcal{L}^{i+1} \tag{4.6}$$

Definition 4.6 *If $\mathcal{L}^n = 0$ for n sufficiently large, the algebra is said to be* ***nilpotent****. We occasionally use this term to refer to a single element: A is nilpotent if $A^n = 0$ for some integer n.*

Lemma 4.2 $[\mathcal{L}^i, \mathcal{L}^j] \subseteq \mathcal{L}^{i+j}$.

Proof. This statement follows from the definition above when $j = 1$. We assume that it is true for any i and some j and prove that it is also true for $j + 1$.

$$\begin{aligned}[\mathcal{L}^i, \mathcal{L}^{j+1}] \equiv [\mathcal{L}^i, [\mathcal{L}^j, \mathcal{L}]] &\equiv [[\mathcal{L}^i, \mathcal{L}], \mathcal{L}^j] + [[\mathcal{L}^i, \mathcal{L}^j], \mathcal{L}] \\ &\equiv [\mathcal{L}^{i+1}, \mathcal{L}^j] + [\mathcal{L}^{i+j}, \mathcal{L}] \subseteq \mathcal{L}^{i+j+1}\end{aligned}$$

Lemma 4.3 $\mathcal{L}^{(n)} \subseteq \mathcal{L}^n$

Proof. It follows from the definitions that $\mathcal{L}^{(0)} \equiv \mathcal{L}^1$ and $\mathcal{L}^{(1)} \equiv \mathcal{L}^2$. Now assume that $\mathcal{L}^{(n)} \subseteq \mathcal{L}^n$ and prove that this is also true for $n + 1$. If this were not true there would be an element $A \in \mathcal{L}^{(n)}$ such that $A \in \mathcal{L}^{(n+1)}$, but $A \notin \mathcal{L}^{n+1}$. This in turn means that there must be two more elements

$B, C \in \mathcal{L}^{(n)}$ such that $A = [B, C]$. From Lemma 1.1, however, $\mathcal{L}^{(n)}$ is contained in both $\mathcal{L}^n$ and $\mathcal{L}$, so A is also in $\mathcal{L}^{n+1}$, and thus we arrive at a contradiction.

One immediate consequence of this lemma is that all nilpotent algebras are solvable, although the converse is not always true. An algebra consisting of square matrices with zeros below the main diagonal is solvable, because the commutator of any two such matrices will have an additional diagonal line of zeros either on or above the main diagonal. It is also true (though far from obvious) that any solvable algebra can be transformed into this form. This is the content of Lie's theorem, which will be proved later on. A matrix with zeros on and below the main diagonal is itself nilpotent, *i.e.* $A^n = 0$. Such a matrix is called *nil-triangular*. An algebra consisting of nil-triangular matrices is nilpotent. The first derived algebra of a solvable algebra is nilpotent.

We will often write $\mathcal{L} = \mathcal{M} + \mathcal{K}$, meaning simply that every element of the Lie algebra $\mathcal{L}$ can be written as a linear combination of elements in $\mathcal{M}$ and $\mathcal{K}$. There are two special sums, however, that carry definite implications about the commutation relations of $\mathcal{M}$ and $\mathcal{K}$.

Definition 4.7 *Direct sum*

A Lie algebra $\mathcal{L} = \mathcal{M} \oplus \mathcal{K}$ is the ***direct sum*** *of two Lie algebras (or subalgebras) $\mathcal{M}$ and $\mathcal{K}$ if every element in $\mathcal{L}$ is a linear combination of elements in $\mathcal{M}$ and $\mathcal{K}$ and if all the elements of $\mathcal{M}$ commute with all elements of $\mathcal{K}$,* i.e. $[\mathcal{M}, \mathcal{K}] = 0$.[2]

Definition 4.8 *Semidirect sum*

A Lie algebra $\mathcal{L} = \mathcal{M} \wedge \mathcal{K}$ is the ***semidirect sum*** *of two subalgebras $\mathcal{M}$ and $\mathcal{K}$ if every element in $\mathcal{L}$ is a linear combination of elements in $\mathcal{M}$ and $\mathcal{K}$ and if one of the subalgebras $\mathcal{K}$ is an invariant subalgebra,* i.e. $[\mathcal{K}, \mathcal{K}] = \mathcal{K}$, $[\mathcal{M}, \mathcal{M}] = \mathcal{M}$, *and* $[\mathcal{M}, \mathcal{K}] = \mathcal{K}$.

Note that the complement of an invariant subalgebra like $\mathcal{K}$ is not necessarily a subalgebra. $\mathcal{L}$ is equal to the semidirect sum of $\mathcal{K}$ and its complement if and only if the complement forms a subalgebra. There is another way of decomposing $\mathcal{L}$ into $\mathcal{K}$ plus "something else," however, that guarantees that the "something else" is an algebra. Any finite dimensional Lie algebra can be written as a linear combination of basis elements. We choose bases σ_j with Roman indices to span $\mathcal{K}$, so that any element $X \in \mathcal{K}$ can be written

[2]This definition is not universally accepted. Sometimes the term "direct sum" implies that $\mathcal{M} \cap \mathcal{K} = 0$ without any assumptions about the commutation relations of $\mathcal{M}$ and $\mathcal{K}$.

$X = \sum x^j \sigma_j$. The remaining bases $\sigma_\mu \notin \mathcal{K}$ have Greek indices. (We will often use this convention.) Any element $Y \in \mathcal{L}$ is then

$$Y = \sum y^\mu \sigma_\mu + \sum y^j \sigma_j.$$

The vector space spanned by the σ_μ is called variously $\mathcal{L}$ mod $\mathcal{K}$, $\mathcal{L}$-$\mathcal{K}$, or $\mathcal{L}/\mathcal{K}$. Two elements are identical in $\mathcal{L}$ mod $\mathcal{K}$ if their y^μ's are identical (regardless of their y^j's); alternatively, two elements are identical in $\mathcal{L}$ mod $\mathcal{K}$ if their difference lies entirely in $\mathcal{K}$. We will call $\sum y^\mu \sigma_\mu$ the "projection" of Y onto $\mathcal{L}$ mod $\mathcal{K}$ and denote it with a bar.

$$\overline{Y} = \sum y^\mu \sigma_\mu$$

We define a new algebra on this space with the following rules:

$$\overline{aX} = a\overline{X}$$

$$\overline{(X+Y)} = \overline{X} + \overline{Y}$$

$$\overline{[\overline{X}, \overline{Y}]} = \overline{Z} \text{ where } [X, Y] = Z.$$

If the elements of $\mathcal{L}$ mod $\mathcal{K}$ constitute a subalgebra, then the algebra defined by these rules is just this algebra. If not, the barred commutator "discards" the component of $[\overline{X}, \overline{Y}]$ contained in $\mathcal{K}$. (Note that this definition requires that $\mathcal{K}$ be an ideal.)

Definition 4.9 *Factor algebra*

Let $\mathcal{K}$ be an invariant subalgebra of $\mathcal{L}$. The algebra of the barred elements given by the above rules is called the ***factor algebra*** *$\mathcal{L}/\mathcal{K}$. Any algebra isomorphic to this is also called $\mathcal{L}/\mathcal{K}$.*

The nomenclature surrounding factor algebras is potentially confusing. It is important to distinguish between the *vector space* $\mathcal{L}/\mathcal{K}$ and the *algebra* $\mathcal{L}/\mathcal{K}$, which may or may not be the same as the algebra of $\mathcal{L}$ restricted to the space $\mathcal{L}/\mathcal{K}$. We will use $\mathcal{L}$-$\mathcal{K}$ or $\mathcal{L}$ mod $\mathcal{K}$ as shorthand notation for the complement of $\mathcal{K}$ in $\mathcal{L}$ without implying that $\mathcal{L}$-$\mathcal{K}$ is a subalgebra. We reserve the term $\mathcal{L}/\mathcal{K}$ for the algebra defined on this space.

There is a close connection between factor algebras and factor groups, Definition (??). In fact, we could adopt a slightly different definition of the factor algebra by considering $\overline{X} + \mathcal{K}$ as a single element regardless of the contents of $\mathcal{K}$. In this way the definition of an element of $\mathcal{L}/\mathcal{K}$ is analagous to Definition (??) of a coset. In this notation the factor algebra is defined as follows:

$$a(\overline{X} + \mathcal{K}) = a\overline{X} + \mathcal{K}$$

$$(\overline{X} + \mathcal{K}) + (\overline{Y} + \mathcal{K}) = (\overline{X} + \overline{Y}) + \mathcal{K}$$

$$[(\overline{X} + \mathcal{K}), (\overline{Y} + \mathcal{K})] = \overline{[\overline{X}, \overline{Y}]} + \mathcal{K}$$

The algebra defined in this way is isomorphic with the algebra of Definition 4.9

Definition 4.10 *The largest possible invariant solvable subalgebra of $\mathcal{L}$ is called the* ***radical*** *of $\mathcal{L}$ or $\mathcal{G}$.*

It is necessary to show that this definition makes sense. First suppose there is *some* invariant solvable subalgebra $\mathcal{K}$. If we can identify an element $X \notin \mathcal{K}$ such that $[X, \mathcal{L}] \in \mathcal{K}$, we can enlarge $\mathcal{K}$ by including X. The new subalgebra $\mathcal{K}'$ will also be solvable and invariant. We proceed in this way until all the elements of $\mathcal{L}$ meeting these requirements are used up. If the remaining elements contain a separate invariant solvable subalgebra it can be combined with $\mathcal{K}$. It is immaterial which invariant solvable subalgebra we start with, so the procedure yields a unique result. The remaining elements, $\mathcal{L}$-$\mathcal{G}$, contain no invariant solvable subalgebras and *a fortiori*, no invariant abelian subalgebras. Is $\mathcal{L}$-$\mathcal{G}$ a subalgebra? It seems as if it has to be, yet the proof of this statement (Jacobson, Hausner and Schwartz) is quite difficult. The final result is called the "radical splitting theorem" or sometimes the "Levi decomposition theorem."

Theorem 4.4 *Let $\mathcal{G}$ be the radical of $\mathcal{L}$. Then $\mathcal{L}/\mathcal{G}$ is a semisimple subalgebra. Any algebra can be decomposed as a semidirect sum,*

$$\mathcal{L} = \mathcal{G} \wedge \mathcal{L}/\mathcal{G}. \tag{4.7}$$

This is the basic decomposition theorem. Any algebra can be divided up into a solvable and a semisimple part. These two subalgebras are then analyzed using rather different strategies.

4.1 Representations

In Section 2.3 we discussed the notion of matrix representations of groups. The same ideas can be applied to the Lie algebra. We start with a definition of matrix representations.

Let $X \in \mathcal{L}$ be an arbitrary element of an abstract Lie algebra, and suppose that $\Gamma(X)$ is its matrix representation.

Definition 4.11 *Lie algebra representations*

The matrices Γ *constitute a* ***representation*** *of the Lie algebra* $\mathcal{L}$ *if*

(1) There is a matrix $\Gamma(X)$ *for each* $X \in \mathcal{L}$.

(2) $[\Gamma(X), \Gamma(Y)] = \Gamma([X, Y])$ *for each* $X, Y \in \mathcal{L}$.

If there is a one-to-one correspondence between the elements of $\mathcal{L}$ *and their representation, the representation is said to be* ***faithful****.*

4.1.1 The Adjoint Representation

Our ultimate goal is to catalog all possible matrix representations of any Lie algebra. There is one special representation, however, that must be discussed at the outset. It is called the *adjoint representation* or *regular representation*, and it is an essential tool for many technical proofs.

A Lie algebra is a linear vector space (Definition 1.5) with dimension n equal to the number of parameters in the corresponding group. (*e.g.* Theorem 3.12) Therefore we are able to choose a set of n linearly independent elements $\sigma_a \in \mathcal{L},\ a = 1, \ldots, n$ to serve as a basis for $\mathcal{L}$. In the following definitions and proofs we will use the convention of summing over repeated upper and lower indices, so, for example, an arbitrary element of $\mathcal{L}$ is written

$$X = \sum_a x^a \sigma_a = x^a \sigma_a.$$

Definition 4.12 *The Adjoint Representation*

Let $\sigma_a,\ a = 1, \ldots, n$ *be a basis and* $X = x^a \sigma_a$ *be an element of* $\mathcal{L}$. *The* ***adjoint representation*** $\boldsymbol{R}(X)$ *is defined by*

$$[X, \sigma_a] = \sigma_b \boldsymbol{R}(X)^b{}_a \tag{4.8}$$

In terms of the structure constants,

$$[X, \sigma_b] = x^a [\sigma_a, \sigma_b] = x^a f^c_{ab} \sigma_c = \sigma_c \boldsymbol{R}(X)^c{}_b,$$

so

$$\boldsymbol{R}(X)^c{}_b = x^a f^c_{ab},$$

or

$$\boldsymbol{R}(\sigma_a)^c{}_b = f^c_{ab}. \tag{4.9}$$

Now starting with the basic commutation relation $[X, Y] = Y'$, we have

$$[X, Y] = [X, \sigma_b] y^b = \sigma_c \boldsymbol{R}(X)^c{}_b y^b = \sigma_c y'^c$$

or

$$\boldsymbol{R}(X)^c{}_b y^b = y'^c. \tag{4.10}$$

$\boldsymbol{R}(X)$ is thus a $n \times n$ matrix that acts on the space of the coefficients y^b.

Theorem 4.5 *The adjoint representation is in fact a representation,* ie $[\boldsymbol{R}(X), \boldsymbol{R}(Y)] = \boldsymbol{R}([X,Y])$ *for all* $X, Y \in \mathcal{L}$.

Proof:

$$[\boldsymbol{R}(X), \boldsymbol{R}(Y)]^a{}_g = x^b y^e (f^a_{bd} f^d_{eg} - f^a_{ed} f^d_{bg}) =$$
$$f^a_{gd} f^d_{eb} x^b y^e = -f^a_{gd} \boldsymbol{R}(X)^d{}_e y^e = y'^d f^a_{dg} = \boldsymbol{R}(Y')^a{}_g$$

We have made repeated use of (4.9) and (4.10). The second equality follows from the Jacobi identity.

The adjoint representation is not necessarily faithful. All elements in the center of the algebra, for example, are represented by the zero matrix. We will prove later that the adjoint representation of semisimple algebras is faithful.

We will also use the adjoint *operators*.

Definition 4.13 *The* ***adjoint operator*** $\hat{R}(X)$ *is defined by*

$$\hat{R}(X)Y = [X,Y]. \tag{4.11}$$

Obviously

$$\hat{R}(X)\sigma_a = [X, \sigma_a] = \sigma_b \boldsymbol{R}(X)^b{}_a. \tag{4.12}$$

This result and the previous theorem give

$$\hat{R}([X,Y]) = [\hat{R}(X), \hat{R}(Y)].$$

Clearly the space on which the operators $\hat{R}(X)$ act is the Lie algebra $\mathcal{L}$. The definition is identical, however, with the linear transformation $Ad(A)$ defined on *group* representations, Definition 2.3; so that Lemma 2.7 and Theorem 2.8 hold for algebras as well as groups.

We will often switch back and forth between adjoint operators and the matrices of the adjoint representation. For example, the expression $\hat{R}^p$ denotes p nested commutators

$$\hat{R}^p(X)Y = [X, [X, \cdots [X,Y] \cdots]] = Y',$$

whereas $\boldsymbol{R}^p$ simply denotes a matrix raised to the p-th power,

$$(\boldsymbol{R}^p(X))^b{}_a y^a = y'^b.$$

Nonetheless, the two statements are completely equivalent.

The operator analog of the eigenvalue equation,

$$(\boldsymbol{R}(X) - \alpha I)y = 0,$$

is

$$(\hat{R}(X) - \alpha\hat{I})Y = [X, Y] - \alpha Y = 0.$$

The identity operator $\hat{I}$ is just a notational gimmick. It does not correspond to an actual element in the algebra.

The adjoint representation is useful because it enables us to visualize the commutation relations in terms of the simple linear transformation (4.10). The subalgebra structure can then be made apparent by partitioning $\boldsymbol{R}$(X). Suppose, for example, the Lie algebra $\mathcal{L}$ has a subalgebra $\mathcal{A}$. The subalgebra is spanned by the basis σ_j labeled with Roman indices. The remaining basis elements σ_α are labeled with Greek indices. With these conventions (4.10) becomes

$$\boldsymbol{R}(\sigma_i)y \to \begin{bmatrix} f^k_{ij} & f^k_{i\beta} \\ 0 & f^\gamma_{i\beta} \end{bmatrix} \begin{bmatrix} y^j \\ y^\beta \end{bmatrix} = \begin{bmatrix} f^k_{ij}y^j + f^k_{i\beta}y^\beta \\ f^\gamma_{i\beta}y^\beta \end{bmatrix} \tag{4.13}$$

$$\boldsymbol{R}(\sigma_\alpha)y \to \begin{bmatrix} f^k_{\alpha j} & f^k_{\alpha\beta} \\ f^\gamma_{\alpha j} & f^\gamma_{\alpha\beta} \end{bmatrix} \begin{bmatrix} y^j \\ y^\beta \end{bmatrix} = \begin{bmatrix} f^k_{\alpha j}y^j + f^k_{\alpha\beta}y^\beta \\ f^\gamma_{\alpha j}y^j + f^\gamma_{\alpha\beta}y^\beta \end{bmatrix} \tag{4.14}$$

The zero in the lower left corner of $\boldsymbol{R}(\sigma_i)$ insures the the σ_i's form a subalgebra, *i.e.* that $[\mathcal{A}, \mathcal{A}] \subseteq \mathcal{A}$. We could also say that the y^j's form a subspace that is invariant under $\boldsymbol{R}(\sigma_i)$ or that $\boldsymbol{R}(\sigma_i)$ is reducible. If $\mathcal{A}$ is an invariant subalgebra, there are additional zeros: $f^\gamma_{\alpha j} = f^\gamma_{i\beta} = 0$ so that $[\mathcal{A}, \mathcal{L}] \subseteq \mathcal{A}$

$$\boldsymbol{R}(\sigma_i)y \to \begin{bmatrix} f^k_{ij} & f^k_{i\beta} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y^j \\ y^\beta \end{bmatrix} = \begin{bmatrix} f^k_{ij}y^j + f^k_{i\beta}y^\beta \\ 0 \end{bmatrix} \tag{4.15}$$

$$\boldsymbol{R}(\sigma_\alpha)y \to \begin{bmatrix} f^k_{\alpha j} & f^k_{\alpha\beta} \\ 0 & f^\gamma_{\alpha\beta} \end{bmatrix} \begin{bmatrix} y^j \\ y^\beta \end{bmatrix} = \begin{bmatrix} f^k_{\alpha j}y^j + f^k_{\alpha\beta}y^\beta \\ f^\gamma_{\alpha\beta}y^\beta \end{bmatrix} \tag{4.16}$$

In addition we could require that the complement of $\mathcal{A}$ form a subalgebra, so that $f^k_{\alpha\beta} = 0$.

$$\boldsymbol{R}(\sigma_\alpha)y \to \begin{bmatrix} f^k_{\alpha j} & 0 \\ 0 & f^\gamma_{\alpha\beta} \end{bmatrix} \begin{bmatrix} y^j \\ y^\beta \end{bmatrix} = \begin{bmatrix} f^k_{\alpha j}y^j \\ f^\gamma_{\alpha\beta}y^\beta \end{bmatrix} \tag{4.17}$$

Finally, to make both $\mathcal{A}$ and its complement into invariant subalgebras we need

$$\boldsymbol{R}(\sigma_i)y \to \begin{bmatrix} f_{ij}^k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y^j \\ y^\beta \end{bmatrix} = \begin{bmatrix} f_{ij}^k y^j \\ 0 \end{bmatrix} \tag{4.18}$$

$$\boldsymbol{R}(\sigma_\alpha)y \to \begin{bmatrix} 0 & 0 \\ 0 & f_{\alpha\beta}^\gamma \end{bmatrix} \begin{bmatrix} y^j \\ y^\beta \end{bmatrix} = \begin{bmatrix} 0 \\ f_{\alpha\beta}^\gamma y^\beta \end{bmatrix} \tag{4.19}$$

There is a final point to be made about (4.13). With a slight change of notation we can write

$$\boldsymbol{R}(\sigma_i) = \left[\begin{array}{c|c} \boldsymbol{R}_{11}(\sigma_i) & \boldsymbol{R}_{12}(\sigma_i) \\ \hline 0 & \boldsymbol{R}_{22}(\sigma_i) \end{array}\right].$$

Direct calculation yields

$$\boldsymbol{R}([\sigma_i,\sigma_j]) = \boldsymbol{R}(\sigma_i)\boldsymbol{R}(\sigma_j) - \boldsymbol{R}(\sigma_j)\boldsymbol{R}(\sigma_i) = \left[\begin{array}{c|c} \boldsymbol{R}_{11}([\sigma_i,\sigma_j]) & * \\ \hline 0 & \boldsymbol{R}_{22}([\sigma_i,\sigma_j]) \end{array}\right]$$

The * in the upper right corner indicates that the entry is not zero but also not relevant to the following observation. Both $\boldsymbol{R}_{11}$ and $\boldsymbol{R}_{22}$ are representations of the subalgebra $\mathcal{A}$. $\boldsymbol{R}_{11}$ maps the space of the y^j's into itself

$$\boldsymbol{R}_{11}([\sigma_i,\sigma_j])^k{}_j y^j = y'^k,$$

or in terms of the adjoint operators

$$\hat{R}_{11}([\sigma_i,\sigma_j])\mathcal{A} \subseteq \mathcal{A},$$

which is just what we expected. The surprising thing is that $\boldsymbol{R}_{22}$ is also a representation of $\mathcal{A}$, even though it acts on the complementary space, the space of the y^β's. In terms of the operators

$$\hat{R}_{22}([\sigma_i,\sigma_j])(\mathcal{L}-\mathcal{A}) \subseteq (\mathcal{L}-\mathcal{A})$$

Since several key theorems depend on this peculiar fact, we state it as a theorem.

Theorem 4.6 *If $\mathcal{A}$ is a subalgebra of $\mathcal{L}$ and $\boldsymbol{R}$ is the regular representation of $\mathcal{L}$, then $\boldsymbol{R}$ restricted to $\mathcal{A}$ forms a representation of $\mathcal{A}$ acting on the space $\mathcal{L}$-$\mathcal{A}$.*

4.1.2 The Jordan canonical form

The carrier space V for the representation $\Gamma(X)$ of the Lie algebra $\mathcal{L}$ is a linear vector space consisting of vectors v such that $\Gamma(X)v = v' \in V$ for all $v \in V$ and all $X \in \mathcal{L}$. For example, if the representation consists of $n \times n$ matrices, then the carrier space is composed of $n \times 1$ column vectors on which the matrices operate. The elements of the representation can be thought of as linear transformations on this space. We will say that V *carries* the representation.

The definitions of reducible, completely reducible, and irreducible group representations given in Section 2.5 apply equally to representations of Lie algebras. This is a good place to call attention to the fact that a reducible representation always leaves invariant a subspace of the carrier space. The statement that $\Gamma(X)$ is irreducible is equivalent to saying that there are no invariant subspaces in V.

The clearest way of displaying the subspace structure associated with a single matrix is to choose a carrier space so that the matrix assumes Jordan canonical form. This procedure is discussed in several standard texts. I have found the treatment in Friedman espcially helpful. I will state the general theorem of the Jordan canonical form and then sketch an outline of the proof. The reader is invited to fill in the missing details.

Theorem 4.7 *Any square matrix L can be brought into the following form with a unitary transformation:*

$$L = \begin{bmatrix} A_1 & 0 & 0 & 0 & \cdots \\ 0 & A_2 & 0 & 0 & \cdots \\ 0 & 0 & A_3 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 0 & 0 & A_k \end{bmatrix} \tag{4.20}$$

where

$$A_j = \begin{bmatrix} \alpha_j & 1 & 0 & 0 & \cdots & \cdots \\ 0 & \alpha_j & 1 & 0 & \cdots & \cdots \\ 0 & 0 & \alpha_j & 1 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 0 & 0 & \alpha_j & 1 \\ \cdots & \cdots & 0 & 0 & 0 & \alpha_j \end{bmatrix} \tag{4.21}$$

and α_j is an eigenvalue of L.

In this form the action of L on the carrier space is evident at a glance. The following features should be noted:

1) L has been completely reduced. The A_j's are irreducible submatrices with eigenvalues on the main diagonal.

2) The matrix $(A_j - \alpha_j I)$ is a "lowering operator," *ie.* if the basis element ξ_k consists of a column matrix with 1 in the k-th place and 0 elsewhere, then

$$(A_j - \alpha_j I)\xi_k = \xi_{k-1}.$$

It would be very convenient if we could analyze the Lie algebra in terms of this structure. Unfortunately this theorem only refers to a single matrix. There is no guarantee that an arbitrary set of matrices can be simultaneously transformed to Jordan canonical form. The following material explores the extent to which this is possible and should be regarded as an extension of the theory underlying the Jordan canonical form. We first review this theory as it pertains to a single matrix.

Let L be an arbitrary $n \times n$ matrix. The eigenvalue equation

$$(L - \alpha I)v = 0 \tag{4.22}$$

will have n (possibly degenerate) eigenvalue solutions. If the eigenvalues are all distinct, the corresponding eigenvectors will be linearly independent and span a n-dimension vector space V. These eigenvectors can be used to construct an unitary matrix U with which L can be put in diagonal form with the eigenvalues along the main diagonal. If the eigenvalues are not all distinct, however, the eigenvectors will in general not be linearly independent and will not span V. In this case we can enlarge the space by defining *generalized eigenvectors.*

Definition 4.14 *A vector v_k for which*

$$(L - \alpha I)^{k-1} v_k \neq 0$$

but for which

$$(L - \alpha I)^k v_k = 0$$

is called a ***generalized eigenvector*** *of rank k (k is an integer) corresponding to the eigenvalue α.*

It is easy to show that for any integer $j < k$,

$$(L - \alpha I)^j v_k = v_{k-j}, \tag{4.23}$$

and that eigenvectors of different rank are linearly independent. We group these eigenvectors into *null spaces.*

Definition 4.15 *The **null space** $\mathcal{N}_k$ is the space of all vectors v such that*

$$(L - \alpha I)^k v = 0.$$

Equation (4.23) can be rewritten in this notation as

$$(L - \alpha I)^j \mathcal{N}_k \subseteq \mathcal{N}_{k-j}. \tag{4.24}$$

Evidently

$$\mathcal{N}_1 \subseteq \mathcal{N}_2 \subseteq \mathcal{N}_3 \subseteq \cdots,$$

but in a finite dimensional space this sequence must eventually reach some $\mathcal{N}_k$ such that $\mathcal{N}_n \equiv \mathcal{N}_k$ for all $n \geq k$. Then k is called the *index* of the eigenvalue α. Consider a null space $\mathcal{N}_j$ where $j \leq k$. Then $\mathcal{N}_{j-1} \subseteq \mathcal{N}_j$, and we can decompose $\mathcal{N}_j$ into $\mathcal{N}_{j-1}$ and "everything left over,"

$$\mathcal{N}_j = \mathcal{N}_{j-1} + \mathcal{P}_{j-1}. \tag{4.25}$$

In the usual terminology $\mathcal{P}_{j-1}$ is called a "progenitor space."

Definition 4.16 *The **range** of $(L-\alpha I)^{j-1}$ is the set of all non-zero vectors of the form $(L - \alpha I)^{j-1} v_i$ where $v_i \in \mathcal{N}_j$. The **progenitorspace** $\mathcal{P}_{j-1}$ is spanned by the smallest possible set of v_i's such that the $(L-\alpha I)^{j-1} v_i$'s span the range.*

With this definition the decomposition (4.25) is unique.

Now suppose that the eigenvalue α has a multiplicity d and an index $k \leq d$. We decompose $\mathcal{N}_k$ as follows:

$$\begin{array}{lcl} \mathcal{N}_k & = & \mathcal{N}_{k-1} + \mathcal{P}_{k-1} \\ \mathcal{N}_{k-1} & = & \mathcal{N}_{k-2} + \mathcal{P}_{k-2} \\ \mathcal{N}_{k-2} & = & \mathcal{N}_{k-3} + \mathcal{P}_{k-3} \\ \cdots & \cdots & \cdots \\ \mathcal{N}_1 & = & \mathcal{N}_0 + \mathcal{P}_0 \end{array}$$

This procedure terminates at the space $\mathcal{N}_0$, which contains nothing. Then

$$\mathcal{N}_k = \mathcal{P}_{k-1} + \mathcal{P}_{k-2} + \cdots + \mathcal{P}_0. \tag{4.26}$$

If $p \in \mathcal{P}_i$ then $(L - \alpha I)p \in \mathcal{P}_{i-1}$; consequently, every vector $p_{k-1} \in \mathcal{P}_{k-1}$ gives rise to a chain of vectors:

$$\begin{array}{ll} (L - \alpha I)p_{k-1} & = p_{k-2} \in \mathcal{P}_{k-2} \\ (L - \alpha I)p_{k-2} & = p_{k-3} \in \mathcal{P}_{k-3} \\ \cdots & \cdots \\ (L - \alpha I)p_1 & = p_0 \in \mathcal{P}_0 \\ (L - \alpha I)p_0 & = 0 \end{array}$$

The space $\mathcal{P}_{k-1}$ cannot be empty (otherwise k would not be the index), so none of the other $\mathcal{P}_i$'s are empty. The carrier space for a matrix in Jordan canonical form is composed of these progenitor vectors. We start by finding a vector in $\mathcal{P}_{k-1}$ and then computing the other members of its chain. We then construct an unitary transformation U such that

$$Up_0 = \xi_1, \quad Up_1 = \xi_2, \quad Up_2 = \xi_3, \cdots \tag{4.27}$$

Where as usual, ξ_k is a column matrix with 1 in the k-th place and zero elsewhere. The order is important. In order to obtain the Jordan canonical form in the conventional format, it is necessary to start at the "bottom" of the chain, p_0, and work up to p_{k-1}. Equation (4.27) also gives a recipe for calculating U; the columns of U^{-1} are the progenitor vectors $p_0, p_1, \cdots$ arranged in the canonical order. In this new basis

$$U(L-\alpha I)U^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ & \cdots & \cdots & \cdots & \end{bmatrix}$$

$$U(L-\alpha I)U^{-1}\xi_j = \xi_{j-1}$$

$$U(L-\alpha I)U^{-1}\xi_1 = 0$$

The k transformed progenitor vectors, $\xi_1, \xi_2, \cdots, \xi_k$, constitute an invariant subspace. The operation of L on this subspace is represented by the matrix A_1 given in (4.21). The eigenvalue $\alpha_1 = \alpha$ is just the first eigenvalue that we happened to consider.

If there is another vector in $\mathcal{P}_{k-1}$, it will spawn another chain of vectors and a separate, linearly independent, invariant subspace. In the new basis this subspace is spanned by $\xi_{k+1}, \xi_{k+2}, \cdots, \xi_{2k}$. We proceed in this way until all the vectors in $\mathcal{P}_{k-1}$ have been used up. If there are additional vectors in $\mathcal{P}_{k-2}$, they will have their own chains and associated $k-1$ dimensional subspaces. By the time we have used up the contents of $\mathcal{P}_0$ we will have a set of d_α eigenvectors where d_α is the multiplicity of α.

If (4.22) has other eigenvalues we calculate their index and repeat the above procedure. The final result is a matrix in the form (4.20). Each A_i operates on an invariant subspace consisting of a single chain of progenitor vectors.

The following example illustrates these ideas.

Example 4.1 *Jordan Canonical Form*

Consider the following matrix:

$$L = \begin{bmatrix} 2 & 0 & 1 & 1 & 0 & -1 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

The secular equation

$$\det \parallel L - \alpha I \parallel = (2-\alpha)^6 = 0,$$

so the multiplicity of the root $\alpha = 2$ is six. The eigenvector equation

$$(L - 2I)v = 0$$

yields the constraints $v_3 = v_6$ and $v_2 = v_5 = 0$. (Note that the subscripts here indicate the *components* of the vector v, not the index.) Thus only three of the six eigenvectors are independent. Further calculation shows that

$$(L - 2I)^2 v = 0$$

gives one constraint, $v_3 = v_6$, and

$$(L - 2I)^3$$

is identically zero. Consequently the index of $\alpha = 2$ is three. There are three ordinary eigenvectors in $\mathcal{N}_1$, five generalized eigenvectors in $\mathcal{N}_2$, and six in $\mathcal{N}_3$.

The space $\mathcal{N}_3$ consists of all six-component vectors. The range $\mathcal{R}_2$ is found from

$$(L - 2I)^2 v = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix} = \begin{bmatrix} 0 \\ -v_3 + v_6 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The range is one-dimensional, and thus the progenitor space is spanned by

a single vector, which we will take to be

$$p_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

This vector stands at the head of a chain with two more vectors in it.

$$(L-2I)p_2 = p_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix} \qquad (L-2I)p_1 = p_0 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

We now compute the range and progenitor spaces $\mathcal{R}_1$ and $\mathcal{P}_1$ starting with the null space $\mathcal{N}_2$, *i.e.* those vectors with $v_3 = v_6$.

$$(L-2I)v = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_4 \\ v_5 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$\mathcal{R}_2$ is two-dimensional; the progenitor vectors are

$$p_1' = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad p_1'' = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Finally

$$(L-2I)p_1' = p_0' = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad (L-2I)p_1'' = p_0'' = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Since $p_0 = -p_0''$ and $p_1 = p_0' - p_1''$, only five of the seven progenitor vectors found so far are independent. The missing vector must be in $\mathcal{P}_0$. We argue that the range $\mathcal{R}_0$ must be identical with $\mathcal{N}_1$, since $(L-2I)^0 = I$. Thus we may use any of the ordinary eigenvectors to complete the set. We choose

$$p_0''' = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

The six generalized eigenvectors are arranged as follows: $v_{(1)} = p_0,\ v_{(2)} = p_1,\ v_{(3)} = p_2,\ v_{(4)} = p_0',\ v_{(5)} = p_1',\ v_{(6)} = p_0'''$. The subscript in parentheses anticipates notation we will use later to indicate the canonical ordering of the eigenvectors. The transformation matrices are

$$U^{-1} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad U = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Finally

$$ULU^{-1} = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

The matrix ULU^{-1} can obviously be partitioned into three submatrices. In the notation of (4.20) and (4.21)

$$A_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \qquad A_3 = [2]\,.$$

The A_1 submatrix operates on the space spanned by the three unprimed progenitors, $p_\circ, p_1$, and p_2. A_2 corresponds to $p_\circ'$ and p_1', and A_3 to $p_\circ'''$. All share the same eigenvalue $\alpha = 2$.

4.1.3 Simultaneous eigenvectors and Lie's theorem

This procedure cannot be generalized to a set of matrices such as a Lie algebra unless we can find vectors that are simultaneous eigenvectors of the entire set. The next few theorems explore situations in which this is possible. We start with the eigenvalue equation,

$$\Gamma(X)v = \alpha(X)v, \tag{4.28}$$

where $\Gamma(X)$ is a matrix representation of an arbitrary element X of a Lie algebra. Since α depends on X (unlike v) we often write it as a linear functional $\alpha(X)$.

Theorem 4.8 *If $\Gamma(X)$ is a representation of an abelian Lie algebra, then it has a simultaneous eigenvector v, that is, a solution to (4.28) $v \neq 0$ for all $X \in \mathcal{L}$. If $\Gamma(X)$ is irreducible then the carrier space is one-dimensional, and $\Gamma(X)$ is just multiplication by a scalar.*

Proof: Every square matrix has at least one eigenvector and eigenvalue. Let S be the set of all $v \in V$ such that

$$\Gamma(X)v = \alpha(X)v,$$

for at least one $X \in \mathcal{L}$. But then

$$\Gamma(X)\Gamma(Y)v = \Gamma(Y)\Gamma(X)v = \alpha(X)\Gamma(Y)v,$$

so that $\Gamma(Y)v \in S$. Thus S is invariant. If Γ is irreducible, then S must be the entire space V. We would come to the same conclusion no matter what element $X \in \mathcal{L}$ we started with. Consequently, V is spanned by the simultaneous eigenvectors.

It remains to prove that V is one-dimensional. If not, we could find two vectors v_1 and v_2 such that

$$\Gamma(X)v_1 = \alpha_1(X)v_1$$
$$\Gamma(X)v_2 = \alpha_2(X)v_2.$$

Repeating the above argument we conclude that the space S_1 spanned by v_1 and the space S_2 spanned by v_2 are both invariant. This is only consistent with irreducibility if S_1, S_2, and V are identical.

We can summarize by saying that the effect of $\Gamma(X)$ on any element of V is simply to multiply it by a scalar $\alpha(X)$. The carrier space is spanned by a single v, which is simultaneously an eigenvector for the entire representation.

Irreducibility plays a key role in this theorem, but the results are trivial and uninteresting: the only matrices satisfying the conditions of the theorem are in fact not matrices at all but just numbers. Let us see what can be proved without assuming irreducibility. We first choose a basis for the algebra, $\sigma_1, \sigma_2, \cdots, \sigma_n$, and solve the eigenvalue equation for the first basis element,

$$\Gamma(\sigma_1)v_1 = \alpha_1 v_1.$$

In general the multiplicity of the root α_1 may be larger than one, so that there may be more than one linearly independent eigenvector. The space spanned by these eigenvectors will be called V_1. Since

$$\Gamma(\sigma_1)\Gamma(\sigma_2)v_1 = \alpha_1\Gamma(\sigma_2)v_1,$$

V_1 is invariant under the action of $\Gamma(\sigma_2)$.

$$\Gamma(\sigma_2)V_1 \subseteq V_1$$

Consequently, the equation

$$\Gamma(\sigma_2)v_{12} = \alpha_2 v_{12},$$

has at least one solution, $v_{12} \in V_1$. Call the space of all such solutions V_{12} and repeat this procedure with the remaining σ_i's. Finally we are left with an invariant subspace $V_{12\ldots n}$ of vectors $v_{12\ldots n}$ having the property

$$\Gamma(\sigma_i)v_{12\ldots n} = \alpha_i v_{12\ldots n}.$$

By choosing an appropriate basis for the carrier space we can simultaneously partition all the $\Gamma(\sigma_i)$'s as follows:

$$\Gamma(\sigma_i)v_{12\ldots n} \rightarrow \left[\begin{array}{c|c} \alpha_i I & \Gamma_{12}(\sigma_i) \\ \hline 0 & \Gamma_{22}(\sigma_i) \end{array}\right] \left[\begin{array}{c} V_{12\ldots n} \\ \hline V - V_{12\ldots n} \end{array}\right]$$

Since

$$[\Gamma(\sigma_i), \Gamma(\sigma_j)] = \Gamma([\sigma_i, \sigma_j]) = 0,$$

it must be true that (Theorem 4.6)

$$[\Gamma_{22}(\sigma_i), \Gamma_{22}(\sigma_j)] = \Gamma_{22}([\sigma_i, \sigma_j]) = 0.$$

So the $\Gamma_{22}(\sigma_i)$'s also make up an abelian algebra, and the entire procedure can be repeated with them. Consequently, $\Gamma(X)$ can be transformed to upper triangular form with eigenvalues along the main diagonal.

These results are more or less obvious. Much less obvious is the fact that the same conclusions hold for any solvable Lie algebra. This subtle theorem was first proved by Lie; we prove it here using an argument that is modeled after the derivation of the Jordan canonical form.

Lemma 4.9 *Suppose that* $\Gamma(X)$ *is irreducible and that* $\mathcal{L}$ *contains an invariant subalgebra* $\mathcal{A}$*. If* $\mathcal{A}$ *possesses a simultaneous eigenvector, then all the* $\Gamma(A)$ *for* $A \in \mathcal{A}$ *consist of scalar multiplication.*

Proof: Define $\mathcal{N}_p$ as the set of all v that are solutions of

$$(\Gamma(A) - \alpha(A))^p v = 0 \tag{4.29}$$

for all $A \in \mathcal{A}$. Our assumptions guarantee that there is at least one vector in $\mathcal{N}_1$, so we can construct the nested sets, $\mathcal{N}_1 \subseteq \mathcal{N}_2 \subseteq \mathcal{N}_3 \subseteq \cdots$, as we did in the proof of Theorem (4.7). Now let X be an arbitrary element in $\mathcal{L}$. We would like to prove that

$$\Gamma(X)\mathcal{N}_p \subseteq \mathcal{N}_{p+1}. \tag{4.30}$$

The proof works by induction; we assume that

$$\Gamma(X)\mathcal{N}_{p-1} \subseteq \mathcal{N}_p$$

and show that (4.30) follows immediately. If $v \in \mathcal{N}_p$, $X \in \mathcal{L}$, and $A \in \mathcal{A}$, then

$$(\Gamma(A) - \alpha(A))\Gamma(X)v = \Gamma(X)(\Gamma(A) - \alpha(A))v + [\Gamma(A), \Gamma(X)]v. \tag{4.31}$$

In terms of the null spaces,

$$\begin{aligned}
&(\Gamma(A) - \alpha(A))\Gamma(X)\mathcal{N}_p \\
&\equiv \Gamma(X)(\Gamma(A) - \alpha(A))\mathcal{N}_p + [\Gamma(A), \Gamma(X)]\mathcal{N}_p && (4.32)\\
&\subseteq \Gamma(X)\mathcal{N}_{p-1} + [\Gamma(A), \Gamma(X)]\mathcal{N}_p && (4.33)\\
&\subseteq \mathcal{N}_p + [\Gamma(A), \Gamma(X)]\mathcal{N}_p && (4.34)\\
&\subseteq \mathcal{N}_p && (4.35)
\end{aligned}$$

Equation (4.24) was used to obtain (4.33), and the induction hypothesis was used in (4.34). Equation (4.35) makes use of the following argument: since $[A, X] \in \mathcal{A}$,

$$\{[\Gamma(A), \Gamma(X)] - \alpha([\Gamma(A), \Gamma(X)])\}\,\mathcal{N}_p \subseteq \mathcal{N}_{p-1}.$$

But $\alpha\mathcal{N}_p \equiv \mathcal{N}_p$, so $[\Gamma(A), \Gamma(X)]\mathcal{N}_p \subseteq \mathcal{N}_p$. The conclusion (4.30) follows from (4.35) and (4.24).

Now $\mathcal{N}_p$ cannot be larger than V, so there must be some index k such that

$$\Gamma(X)\mathcal{N}_p = \mathcal{N}_p.$$

for $p \geq k$. Moreover $\mathcal{N}_p$ cannot be a subset of V, because of the hypothesis of irreducibility. Therefore, $\mathcal{N}_p = V$ for all $p \geq k$. We will now show that in fact $k = 1$.

First note that all the eigenvalues of $\Gamma(A)$ are equal. This can be seen as follows: suppose there were an $\alpha'(A) \neq \alpha(A)$ and a corresponding eigenvector v'. Then

$$(\Gamma(A) - \alpha)^p v' = (\alpha' - \alpha)^p v' \neq 0.$$

This contradicts the conclusion of the previous paragraph that every $v \in V$ is a solution of (4.29) with $p \geq k$. Now consider the commutator $[\Gamma(A), \Gamma(X)]$. Its trace is zero, because it is a commutator. Therefore, the sum of its eigenvalues, all of which are equal, is zero.

$$[\Gamma(A), \Gamma(X)]v = \alpha([\Gamma(A), \Gamma(X)])v = 0,$$

even for $v \in \mathcal{N}_1$. Then (4.31) shows that if $v \in \mathcal{N}_1$, then $\Gamma(X)v \in \mathcal{N}_1$, so $\mathcal{N}_1$ is invariant, and $\mathcal{N}_1 = V$. To summarize,

$$\Gamma(A)v = \alpha(A)v \tag{4.36}$$

for all $v \in V$ and all $A \in \mathcal{A}$.

If Γ is not irreducible we are not allowed to claim that $\mathcal{N}_p$ is the entire space. It is an invariant subspace, however, and we can still argue that all the eigenvalues of $\Gamma(A)$ acting on $\mathcal{N}_p$ are equal. Consequently, (4.36) is true for all $v \in \mathcal{N}_p$ and all $A \in \mathcal{A}$.

In order to prove this lemma we had to assume that $\mathcal{A}$ had a simultaneous eigenvector. We will now show that this is always true if $\mathcal{L}$ is solvable.

Theorem 4.10 *Every solvable algebra has a simultaneous eigenvector.*

First notice that under the assumptions of Lemma (4.9) $\Gamma(X)\mathcal{N}_1 \subseteq \mathcal{N}_1$, so that $\Gamma(X)$ has an eigenvector in $\mathcal{N}_1$.

$$\Gamma(X)v = \alpha(X)v$$

Since $v \in \mathcal{N}_1$, it is also an simultaneous eigenvector of $\mathcal{A}$. Since $\mathcal{A}$ is a solvable ideal, the algebra we get by adding X to $\mathcal{A}$ is also solvable. Consequently, every solvable ideal of dimension l with a simultaneous eigenvector

can be enlarged to an algebra of dimension $l+1$ with a simultaneous eigenvector.

Let $\mathcal{L}$ be a solvable algebra whose first derived algebra is $\mathcal{L}^{(1)}$. We can decompose $\mathcal{L}$ by picking out a basis element $\sigma_1 \notin \mathcal{L}^{(1)}$ and writing

$$\mathcal{L} \equiv a_1\sigma_1 + \mathcal{M}.$$

$\mathcal{M}$ is an ideal because the commutator of *any* two elements from $\mathcal{L}$ is contained in $\mathcal{L}^{(1)} \subseteq \mathcal{M}$, and $\mathcal{M}$ is solvable because every subalgebra of a solvable algebra is solvable. $\mathcal{M}$ itself can be decomposed in the same way so long as there is a $\sigma_2 \notin \mathcal{L}^{(1)}$. This procedure is repeated until we use all the bases that are not in $\mathcal{L}^{(1)}$.

$$\mathcal{L} \equiv a_1\sigma_1 + a_2\sigma_2 + \cdots + \mathcal{L}^{(1)}.$$

$\mathcal{L}^{(1)}$ is decomposed by picking out, one by one, all those σi's such that $\sigma i \in \mathcal{L}^{(1)}$ but $\sigma i \notin \mathcal{L}^{(2)}$. We continue in this way until we arrive at the penultimate derived algebra $\mathcal{L}^{(p)}$, which is abelian.

$$\mathcal{L} \equiv a_1\sigma_1 + a_2\sigma_2 + \cdots + a_n\sigma_n + \mathcal{L}^{(p)}.$$

The proof now proceeds by induction. $\mathcal{L}^{(p)}$ is abelian, so by theorem (4.8) it has a simultaneous eigenvector. The algebra composed of

$$a_n\sigma_n + \mathcal{L}^{(p)}$$

has a simultaneous eigenvector and is also a solvable ideal. We now "reassemble" $\mathcal{L}$ by adding $\sigma_{n-1}, \sigma_{n-2}, \cdots$ one at a time. At each step we are adding an element to a solvable ideal with a simultaneous eigenvector, and thus the theorem is proved.

Several useful conclusions follow immediately from this result.

Lemma 4.11 *If $\Gamma(X)$ is irreducible and trace $\Gamma(X) = 0$ for all $X \in \mathcal{L}$, then $\mathcal{L}$ is semisimple.*

If $\mathcal{L}$ were not semisimple it would have an invariant abelian subalgebra, which, by Theorem 1.3, would have a simultaneous eigenvector. Then by Lemma 1.4 the matrices of this subalgebra would be scalar multipliers with zero trace.

Now suppose $\mathcal{L}$ is solvable and $\Gamma(X)$ is irreducible. If $\Gamma(X)$ were traceless we would have a contradiction, because an algebra cannot be both solvable and semisimple. We can construct a traceless algebra $\mathcal{L}'$, however, by subtracting (diminsion of V)$^{-1}$ tr $\Gamma(X)$ from each $\Gamma(X)$. The first derived

algebras of $\mathcal{L}$ and $\mathcal{L}'$ are identical, so $\mathcal{L}'$ is solvable, and we are back to the same contradiction. There is only one way out: $\mathcal{L}'$ contains only the zero element, and $\Gamma(X)$ consists of multiples of the unit matrix. The unit matrix is reducible, however, unless it is one-dimensional. This brings us to Lie's theorem.

Theorem 4.12 *If $\mathcal{L}$ is solvable and $\Gamma(X)$ is irreducible, then the the carrier space is one-dimensional, and $\Gamma(X)$ consists of scalar multipliers for each $X \in \mathcal{L}$.*

Corollary 4.13 *Let Γ be a reducible representation of $\mathcal{L}$. Then $\mathcal{L}$ is solvable if and only if Γ can be put in upper triangular form.*

Proof: By definition a reducible representation can be similarity transformed so that it looks like (2.55) with irreducible submatrices along the diagonal. But by Lie's theorem the submatrices are all one-dimensional. This is called the *canonical representation* of a solvable algebra. To prove the converse, assume that $\Gamma(X)$ is upper triangular for all $X \in \mathcal{L}$. The commutator of any two upper triangular matrices must have zeros along the diagonal, so the representation of the first derived algebra will have this form. The representation of the second derived algebra will have an additional line of zeros just above the diagonal. Succeeding derived algebras are more and more "upper triangular." If the representation is $n \times n$ then the matrices of the $(n-1)$-th derived algebra can have at most a single element in the upper right corner. The n-th derived algebra must be zero.

There is not much more to say about solvable algebras in general. The reader is reminded of Theorem (4.4); any algebra can be decomposed into an invariant solvable subalgebra and a semisimple algebra. Semisimple algebras have a rich structure, a vast body of mathamatical lore pertaining to them, and many physical applications. Solvable algebras are mostly a curiosity, although we will use Lie's theorem in the next section to analyze the non-invariant solvable algebras that appear within semisimple algebras.

We conclude this section with a somewhat contrived example to illustrate Lie's theorem and the analysis of solvable algebras.

Example 4.2 *A 3×3 solvable algebra*

$$X_1 = \frac{1}{2}\begin{bmatrix} 3 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & -1 \end{bmatrix} \quad X_2 = \tfrac{1}{2}\begin{bmatrix} 1 & -1 & 2 \\ 3 & -3 & 2 \\ 0 & 0 & 2 \end{bmatrix} \quad X_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$X_4 = \frac{1}{2}\begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad X_5 = \tfrac{1}{2}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}$$

The commutation relations are as follows:

$$\begin{aligned} [X_1, X_2] &= 2X_3 - 2X_4 - 2X_5 & [X_1, X_3] &= 2X_3 - X_4 & [X_1, X_4] &= 0 \\ [X_1, X_5] &= -2X_5 & [X_2, X_3] &= -X_3 & [X_2, X_4] &= X_4 \\ [X_2, X_5] &= X_4 + 2X_5 & [X_3, X_4] &= 0 & [X_3, X_5] &= X_4 \\ [X_4, X_5] &= 0 \end{aligned}$$

This is solvable since $\mathcal{L}^{(1)} = X_3, X_4, X_5$, $\mathcal{L}^{(2)} = X_4$, and $\mathcal{L}^{(3)} = 0$. There is one simultaneous eigenvector, $v_1 = (1, 1, 0)$. Choose a matrix U so that

$$(UX_iU^{-1})(Uv_1) = \alpha_i(Uv_1) = X_i'\xi_1 = \alpha_i\xi_1.$$

A convenient choice is

$$U^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The transformed matrices are

$$X_1' = \frac{1}{2}\left[\begin{array}{c|cc} 2 & -1 & 0 \\ \hline 0 & 2 & 0 \\ 0 & -1 & -2 \end{array}\right] \quad X_2' = \tfrac{1}{2}\left[\begin{array}{c|cc} 0 & -1 & 2 \\ \hline 0 & -2 & 0 \\ 0 & 0 & 2 \end{array}\right] \quad X_3' = \left[\begin{array}{c|cc} 0 & 0 & 1 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

$$X_4' = \frac{1}{2}\left[\begin{array}{c|cc} 0 & -1 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \quad X_5' = \tfrac{1}{2}\left[\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ 0 & -1 & 0 \end{array}\right]$$

The 2×2 submatrices in the lower right are also solvable with a simultaneous eigenvector $(0, 1)$. Repeating the above procedure with

$$U'^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

yields

$$X_1'' = \frac{1}{2}\begin{bmatrix} 2 & 0 & -1 \\ 0 & -2 & -1 \\ 0 & 0 & 2 \end{bmatrix} \quad X_2'' = \tfrac{1}{2}\begin{bmatrix} 0 & 2 & -1 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \quad X_3'' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$X_4'' = \frac{1}{2}\begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad X_5'' = \tfrac{1}{2}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

4.2 The Secular Equation

The results of the previous section are based on the existence of simultaneous eigenvectors, *i.e.* eigenvectors of the entire algebra. These do not exist in general for non-solvable algebras, and we must introduce some new ideas.

Let $\mathcal{L}$ be a n-dimensional Lie algebra with basis $\sigma_i, i = 1 \cdots n$, so that an arbitrary element can be written $X = x^i \sigma_i$. The regular representation of X is a $n \times n$ matrix $\boldsymbol{R}(X)$. This matrix has n eigenvalues, which can be found by solving the secular equation

$$\det \|\boldsymbol{R}(X) - \alpha I\| = 0. \tag{4.37}$$

This is a polynomial in α of order n with real coefficients ϕ_j.

$$\det \|\boldsymbol{R}(X) - \alpha I\| = \sum_{j=0}^{n} (-\alpha)^{n-j} \phi_j = 0 \tag{4.38}$$

The ϕ_j's are function of the x_i and contain all the information about the structure of the algebra. Since the secular equation is invariant under a similarity transformation

$$\det \|U\boldsymbol{R}(X)U^{-1} - \alpha I\| = \det \|U\| \det \|\boldsymbol{R}(X) - \alpha I\| \det \|U^{-1}\|,$$

this information is encoded in the ϕ_j's in a way that is independent of the choice of basis.

Two of the ϕ_j's are trivial

$$\begin{aligned} \phi_0 &= 1 \\ \phi_n &= \det \|\boldsymbol{R}(X)\| = 0 \end{aligned} \tag{4.39}$$

The fact that $\boldsymbol{R}(X)$ is singular can be seen from (4.10). If we replace Y with X, then

$$\boldsymbol{R}(X)^c{}_b x^b = 0$$

and the determinant of $\boldsymbol{R}$ must vanish.

Example 4.3 *Nilpotent algebras*

An algebra in upper triangular form with zeros on and below the main diagonal is nilpotent. Then

$$\det \|\boldsymbol{R} - \alpha I\| = (-\alpha)^n = 0.$$

Example 4.4 *Invariant subalgebras*

We have seen (4.13) how the presence of an invariant subalgebra allows us to partition the regular representation in upper triangular form

$$\left[\begin{array}{c|c} \boldsymbol{R}_{11}(X) & \boldsymbol{R}_{12}(X) \\ \hline 0 & \boldsymbol{R}_{22}(X) \end{array}\right]$$

In this case the secular equation factors.

$$\det \|\boldsymbol{R}_{11}(X) - \alpha I\| = 0$$

$$\det \|\boldsymbol{R}_{22}(X) - \alpha I\| = 0,$$

and the two matrices $\boldsymbol{R}_{11}(X)$ and $\boldsymbol{R}_{22}(X)$ are analyzed separately.

At this point we encounter a famous technical problem. We started with a real Lie algebra, *i.e.* the x^i are real coefficients. The roots of the secular equation (4.38), however, are in general complex. Thus we are forced to extend our algebra over the complex number field. We do this by writing an arbitrary element $X = c^i\sigma_i$ with complex c^i and replacing $x^i \to c^i$ in all our formulas. This process, called "complexification," leads to a theory of complex Lie algebras. The difficulty is that the bases $\sigma_i, i = 1, \cdots n$ might be linearly independent over the field of real numbers and yet fail to be independent over the field of complex numbers. In this chapter we assume that the n bases are independent and in this way develop a theory of complex Lie algebras. The corresponding real forms are discussed in Chapter (??).

The secular equation can now be factored as follows:

$$\sum_{j=0}^{n}(-\alpha)^{n-j}\phi_j(c^i) = \alpha^{d_\circ}(\alpha-\alpha_1)^{d_1}(\alpha-\alpha_2)^{d_2}\cdots \tag{4.40}$$

The roots, $\alpha_1, \alpha_2, \cdots$ are complicated functions of the c^i. The integers $d_0, d_1, d_2, \cdots$ are the multiplicities of the roots. The multiplicity of the root 0 turns out to be particularly important in the theory that follows. Because $\phi_n = 0$, $\alpha = 0$ is always a root, consequently $d_\circ \geq 1$. In most cases, however, $d_\circ > 1$ for all $X \in \mathcal{L}$. Nilpotent algebras, Example 4.3, are an extreme case, since $d_\circ = n$. Now let p be the largest value of j in (4.38) such that $\phi_j \neq 0$. Then $d_\circ = n - p$. To some extent p will depend on the choice of X, however each algebra will have some maximum value of p.

Definition 4.17 *The* **rank** *of an algebra is $l = n - p$, where p is the largest integer such that ϕ_j in (4.38) is zero for all $j > p$ and all X.*

It can be shown that $\phi_p \neq 0$ for "almost all" of the elements in the algebra. (See Jacobson for a discussion of this point.) Occasionally $\phi_p = 0$ because of some specific set of c^i for which a fortuitous cancellation takes place. An element for which this does *not* happen is said to be *regular*. We will give a slightly different but equivalent definition presently.

We now choose a particular element $X \in \mathcal{L}$, solve the secular equation for $\boldsymbol{R}(X)$, and transform $\boldsymbol{R}(X)$ into Jordan canonical form:

$$U(c)\boldsymbol{R}(c^i\sigma_i)U^{-1}(c) = \begin{bmatrix} M_0 & 0 & 0 & 0 & \cdots \\ 0 & M_{\alpha_1} & 0 & 0 & \cdots \\ 0 & 0 & M_{\alpha_2} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 0 & 0 & M_{\alpha_k} \end{bmatrix} \tag{4.41}$$

where

$$M_{\alpha_j} = \begin{bmatrix} A^{(1)}_{\alpha_j} & 0 & 0 & 0 & \cdots \\ 0 & A^{(2)}_{\alpha_j} & 0 & 0 & \cdots \\ 0 & 0 & A^{(3)}_{\alpha_j} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 0 & 0 & A^{(k)}_{\alpha_j} \end{bmatrix} \tag{4.42}$$

and

$$A^{(i)}_{\alpha_j} = \begin{bmatrix} \alpha_j & 1 & 0 & \cdots & \cdots \\ 0 & \alpha_j & 1 & \cdots & \cdots \\ 0 & 0 & \alpha_j & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 0 & \alpha_j & 1 \\ \cdots & \cdots & 0 & 0 & \alpha_j \end{bmatrix}$$

Each $A^{(i)}_{\alpha_j}$ operates on an invariant subspace consisting of a single chain of progenitor vectors as in (4.21). Each M_{α_j} consists of all those A_{α_j} derived from the same eigenvalue α_j. The notation is cumbersome because there are usually several independent subspaces (indexed by i) corresponding to the same α_j.

The matrix U that transforms $\boldsymbol{R}(X)$ into Jordan canonical form also establishes a new set of basis elements. For example, suppose that v is a generalized eigenvector of $\boldsymbol{R}(X)$ corresponding to the eigenvalue α. As explained in Section 4.1 there are n vectors in all, and they are arranged in a definite order. Let us say that $v_{(k)}$ is the k-th generalized eigenvector. Then

$$U(\boldsymbol{R} - \alpha I)^{d_\alpha} v_{(k)} = U(\boldsymbol{R} - \alpha I)^{d_\alpha} U^{-1} U v_{(k)} = (U\boldsymbol{R}U^{-1} - \alpha I)^{d_\alpha} U v_{(k)} = 0$$

Clearly $Uv_{(k)}$ is a generalized eigenvector of the transformed $\boldsymbol{R}$. This means that $Uv_{(k)} = \xi_k$, where ξ_k is a column vector with one in the k-th place and zeros elsewhere. The columns of U^{-1} are the n generalized eigenvectors arranged in the appropriate order.

The transformed carrier space, the space of the ξ_k's, breaks up into a sum of invariant subspaces. There is one subspace with dimension d_{α_j} for each M_{α_j} in (4.41).

$$V = V_\circ \oplus V_{\alpha_1} \oplus V_{\alpha_2} \cdots \tag{4.43}$$

In the Lie algebra there are n linearly independent elements corresponding to the generalized eigenvectors.

$$v^i_{(k)}\sigma_i = \sigma'_k \tag{4.44}$$

The transformed algebra, the space of the σ'_k, also breaks up into subspaces.

$$\mathcal{L} = \mathcal{V}_\circ + \mathcal{V}_{\alpha_1} + \mathcal{V}_{\alpha_2} \cdots, \tag{4.45}$$

where $\mathcal{V}_{\alpha_i}$ consists of all elements of the form $c^k\sigma'_k$ where the implied sum ranges over all k such that ξ_k spans V_{α_i}. These are invariant subspaces in the sense that

$$\hat{R}(X)\mathcal{V}_{\alpha_i} = [X, \mathcal{V}_{\alpha_i}] \subseteq \mathcal{V}_{\alpha_i}.$$

The forgoing decomposition was based on a single, arbitrarily chosen element X. Of course we cannot simultaneously transform all elements to Jordan canonical form, but it will turn out that one is enough if it is wisely chosen. The point is that the new basis elements have some remarkable properties that will emerge from the following theorem. In order to make best use of this theorem it is important to choose X so that the number of distinct roots is as large as possible. This is equivalent to saying that $d_\circ$ is as small as possible or that X is regular.

Definition 4.18 *An element X is said to be **regular** if there is no element in $\mathcal{L}$ for which the number of distinct roots of the secular equation is larger.*

In the remainder of this development it will be assumed that some regular element X has been chosen, the roots and eigenvectors of $\boldsymbol{R}(X)$ have been calculated, and that σ_i refers to the new, transformed basis elements.

Theorem 4.14 *Let α and β be roots of the secular equation. If $\alpha + \beta$ is also a root, then*

$$[\mathcal{V}_\alpha, \mathcal{V}_\beta] \subseteq \mathcal{V}_{\alpha+\beta} \tag{4.46}$$

If $\alpha + \beta$ is not a root, then

$$[\mathcal{V}_\alpha, \mathcal{V}_\beta] = 0 \tag{4.47}$$

Proof: Clearly

$$(M_\alpha - I\alpha)^{d_\alpha} = 0.$$

Restating this in terms of the adjoint operators gives

$$(\hat{R}(X) - \hat{I}\alpha)^{d_\alpha}\mathcal{V}_\alpha = 0.$$

Consequently, if α and β are roots of $\boldsymbol{R}(X)$ and $\mathcal{V}_\alpha$ and $\mathcal{V}_\beta$ are the corresponding subspaces in the Lie algebra, there will always be integers i and j such that

$$(\hat{R}(X) - \hat{I}\alpha)^i\mathcal{V}_\alpha = (\hat{R}(X) - \hat{I}\beta)^j\mathcal{V}_\beta = 0$$

Now

$$(\hat{R}(X) - (\alpha+\beta)\hat{I})[\mathcal{V}_\alpha, \mathcal{V}_\beta] =$$

$$[(\hat{R}(X) - \alpha)\mathcal{V}_\alpha, \mathcal{V}_\beta] + [\mathcal{V}_\alpha, (\hat{R}(X) - \beta)\mathcal{V}_\beta]. \tag{4.48}$$

This equation makes use of (4.11) and the Jacobi identity. Now multiply repeatedly on the left by the factor $(\hat{R}(X) - (\alpha+\beta)\hat{I})$ and use (4.48) to simplify the right side of the equation. After repeating this process $k-1$ times we have

$$(\hat{R}(X) - (\alpha+\beta)\hat{I})^k[\mathcal{V}_\alpha, \mathcal{V}_\beta]$$

on the left, and on the right is a sum of terms of the form

$$[(\hat{R}(X) - \hat{I}\alpha)^p\mathcal{V}_\alpha, (\hat{R}(X) - \hat{I}\beta)^q\mathcal{V}_\beta]$$

where $p+q=k$. We can always find a k large enough so that either $p \geq i$ or $q \geq j$ for every term in the sum. Then

$$(\hat{R}(X) - (\alpha+\beta)\hat{I})^k[\mathcal{V}_\alpha, \mathcal{V}_\beta] = 0.$$

There are two ways this equation might be true: either $[\mathcal{V}_\alpha, \mathcal{V}_\beta] \subseteq \mathcal{V}_{\alpha+\beta}$ or $[\mathcal{V}_\alpha, \mathcal{V}_\beta] = 0$. The latter is guaranteed if $\alpha+\beta$ is not a root of $\boldsymbol{R}(X)$.

The subspace $\mathcal{V}_\circ$ is especially important.

Definition 4.19 *If X is regular then $\mathcal{V}_\circ$ is called a* ***Cartan subalgebra****.*

From now on we will use $H_i, \quad i = 1, \ldots, l$ for the basis elements that span $\mathcal{V}_\circ$, and (whenever possible) H to represent an arbitrary element in $\mathcal{V}_\circ$.

According to Theorem 4.14, $[\mathcal{V}_\circ, \mathcal{V}_\circ] \subseteq \mathcal{V}_\circ$, so the Cartan subalgebra is an algebra as advertised. Furthermore, $[\mathcal{V}_\circ, \mathcal{V}_\alpha] \subseteq \mathcal{V}_\alpha$, so each $\mathcal{V}_\alpha$ is invariant under $\mathcal{V}_\circ$. These two facts require that the regular representation

of all the elements of $\mathcal{V}_\circ$ have the same block diagonal structure as (4.41). Once the new basis has been constructed according to (4.44), $\boldsymbol{R}(H)$ has the following form for all $H \in \mathcal{V}_\circ$.

$$\boldsymbol{R}(H) = \begin{bmatrix} M_0(H) & 0 & 0 & 0 & \cdots \\ 0 & M_{\alpha_1}(H) & 0 & 0 & \cdots \\ 0 & 0 & M_{\alpha_2}(H) & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & 0 & 0 & M_{\alpha_k}(H) \end{bmatrix} \tag{4.49}$$

For the time being we will use α_i to label the root subspaces even though α_i is not in general a root of $\boldsymbol{R}(H)$. Nonetheless, $\boldsymbol{R}(H)$ acting on $V_\circ$ has only zero eigenvalues for all $H \in \mathcal{V}_\circ$. This is a consequence of the following theorem.

Theorem 4.15 *Let $H_\circ$ be a regular element of $\mathcal{L}$ and let $\mathcal{V}_\circ$ be the set of all $H \in \mathcal{L}$ such that*

$$\hat{R}^l(H_\circ)H = 0,$$

then

$$\hat{R}^l(H)\mathcal{V}_\circ = 0,$$

where H is any element in $\mathcal{V}_\circ$. $\mathcal{V}_\circ$ is a solvable algebra of rank l.

Proof: Consider the eigenvalues of the matrix $\boldsymbol{R}(H_\circ + zH)$, where z is a complex variable. The eigenvalues are clearly functions of z, and there is a body of mathematical lore about the analytic structure of these functions. The crucial theorem (Hausner and Schwartz) states that in the limit of $|z| \to 0$, each root of $\boldsymbol{R}(H_\circ + zH)$ is continuously transformed into one of the roots of $\boldsymbol{R}(H_\circ)$. Furthermore, $\boldsymbol{R}(H_\circ + zH)$ has no fewer distinct roots than $\boldsymbol{R}(H_\circ)$, so that by making $|z|$ small enough, we can bring at least one root of $\boldsymbol{R}(H_\circ + zH)$ arbitrarily close to each root of $\boldsymbol{R}(H_\circ)$.

This theorem is true for any square matrices, but since $\hat{R}(H_\circ + zH)(H_\circ + zH) = 0$, $\boldsymbol{R}(H_\circ + zH)$ has an eigenvalue that is zero for all z. Since $H_\circ$ is regular, $\boldsymbol{R}(H_\circ + zH)$ cannot have any more distinct roots than $\boldsymbol{R}(H_\circ)$. A corollary of the theorem quoted above states that when $\boldsymbol{R}(H_\circ)$ and $\boldsymbol{R}(H_\circ + zH)$ have the same number of distinct roots so that there is a one-to-one correspondence between the two sets of roots for small $|z|$, then the dimensions of the corresponding generalized eigenspaces are the same. Consequently

$$[\hat{R}(H_\circ + zH)]^l \mathcal{V}_\circ = 0$$

for all z. This can only be true if

$$\hat{R}^l(H)\mathcal{V}_\circ = 0.$$

Thus all the submatrices of $\boldsymbol{R}(H)$ acting on $V_\circ$ can be simultaneousy transformed to Jordan canonical form with zeros on the main diagonal. Such matrices are nilpotent, and the corresponding subalgebra is solvable. This completes the proof.

Since the Cartan subalgebra is solvable, all its representations will have simultaneous eigenvectors. Thus if α is one of the roots of (4.40), we can find at least one vector $c_{(\alpha)}$ such that

$$M_\alpha(H_i)c_{(\alpha)} = \alpha_i c_{(\alpha)}$$

for all $H_i \in \mathcal{V}_\circ$, or equivalently, we can define

$$E_\alpha = \sum_{\sigma_j \in \mathcal{V}_\alpha} c^j_{(\alpha)}\sigma_j$$

such that

$$[H_i, E_\alpha] = \alpha_i E_\alpha. \tag{4.50}$$

The notation here is confusing. In (4.40)-(4.42) α_i is the i-th eigenvalue of a specific element $\boldsymbol{R}(X)$. It then becomes the label for the subspace $\mathcal{V}_{\alpha_i}$ on which E_α is defined. Finally, in (4.50) α_i is the eigenvalue of $H_i \in \mathcal{V}_\circ$ acting on the space $\mathcal{V}_\alpha$.[3] We can partially avoid this confusion by constructing the vector

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_l), \tag{4.51}$$

and using it to label the eigenvectors ${E_\alpha}$[4]. So, for example, (4.50) becomes

$$[H_i, E_\alpha] = \alpha_i E_\alpha. \tag{4.52}$$

The root vectors $\boldsymbol{\alpha}$ are defined on a l-dimensional Euclidian space. Each eigenvector E_α can be thought of as a single point in this space. This geometrical depiction of the algebra is a great aid to the intuition, and we will use it extensively in analyzing the semisimple algebras. The notation V_α and $\mathcal{V}_\alpha$is also useful since theorems such as 4.14 hold for any choice of $X \in \mathcal{V}_\circ$. Thus $[\mathcal{V}_\alpha, \mathcal{V}_\beta] \subseteq \mathcal{V}_{\alpha+\beta}$ if $\boldsymbol{\alpha} + \boldsymbol{\beta}$ is also a root, and otherwise $[\mathcal{V}_\alpha, \mathcal{V}_\beta] = 0$. The following example illustrates these themes with a two-dimensional Cartan subalgebra.

[3] Don't blame me. This is standard notation.

[4] For reasons of typography we do not use boldface subscripts.

Example 4.5 $SU(3)$

The group $SU(3)$ consists of all 3×3 unitary, unimodular matrices. There are 8 independent generators, which are conventionally chosen to be $\sigma_i = \lambda_i/2$ where the λ_i are the traceless Hermitian matrices originally introduced by Gell-Mann.

$$\lambda_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \lambda_2 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \lambda_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\lambda_4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \lambda_5 = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix} \quad \lambda_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\lambda_7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

The commutation relations are

$$[\sigma_i, \sigma_j] = i f_{ijk} \sigma_k. \tag{4.53}$$

The f_{ijk} are completely antisymmetric under the exchange of indices. The non-vanishing elements are defined by

$$\begin{gathered} f_{123} = 1 \\ f_{147} = f_{246} = f_{257} = f_{345} = \tfrac{1}{2} \\ f_{156} = f_{367} = -\tfrac{1}{2} \\ f_{458} = f_{678} = \tfrac{\sqrt{3}}{2} \end{gathered} \tag{4.54}$$

This is a good point to review the various factors of i that enter into the calculation. The above definitions make the group generators Hermitian. Physicists always do this because the formalism of quantum mechanics requires Hermitian operators. In the case of $SU(3)$ (and other compact groups) this results in imaginary structure constants and real eigenvalues. Mathematicians, on the other hand, prefer to make the generators anti-Hermitian so that the structure constants are real. The eigenvalues then are pure imaginary. In either case the eigenvectors can be normalized so the the new Lie algebra has all real structure constants as in Table 4.4. Unfortunately, the new basis elements are linear combinations of Hermitian and anti-Hermitian operators. These problems are revisited in Chapter 5 where

we discuss the larger issue of abstracting real Lie algebras from complex ones.

The regular representation is a set of 8×8 matrices.

$$\boldsymbol{R}(\sigma_i)^k{}_j = if_{ijk}$$

The eigenvalues are listed in Table 4.1. All of the elements are regular with the exception of σ_8; however, it is customary to do the Jordan decomposition with σ_3. Thus σ_3 becomes the first element of the Cartan subalgebra, and we identify $H_1 = \sigma_3$. The eigenvectors of $\boldsymbol{R}(\sigma_3)$ are given in Table 4.2. Despite the degeneracy of the eigenvalues they are all independent, so $\boldsymbol{R}(\sigma_3)$ can be diagonalized as follows:

$$U\boldsymbol{R}(\sigma_3)U^{-1} = \begin{bmatrix} 0 & & & & & & & \\ & 0 & & & & & & \\ & & +\frac{1}{2} & & & & & \\ & & & -\frac{1}{2} & & & & \\ & & & & +\frac{1}{2} & & & \\ & & & & & -\frac{1}{2} & & \\ & & & & & & +1 & \\ & & & & & & & -1 \end{bmatrix}$$

The root space consists of eight one-dimensional subspaces, *i.e.* all the A_{α_j} in (4.42) are one-dimensional. We will show eventually that this is a general feature of semisimple algebras.

There is one more element of $\mathcal{V}_\circ$, $H_2 = \sigma_8$, which is also diagonal.

$$U\boldsymbol{R}(\sigma_8)U^{-1} = \begin{bmatrix} 0 & & & & & & & \\ & 0 & & & & & & \\ & & +\frac{\sqrt{3}}{2} & & & & & \\ & & & -\frac{\sqrt{3}}{2} & & & & \\ & & & & -\frac{\sqrt{3}}{2} & & & \\ & & & & & +\frac{\sqrt{3}}{2} & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{bmatrix}$$

The root vectors $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ are two-dimensional. They can be used to label the remaining six elements as listed in Table 4.3 and plotted in Figure 4.1.

We will eventually adopt a different notation for the E_α that looks less cluttered and shows the physical significance of each element. For the time

$\boldsymbol{R}(\sigma_i)$	Eigenvalues
$i = 1, \cdots, 7$	$0, 0, \pm 1/2, \pm 1/2, \pm 1$
$i = 8$	$0, 0, 0, 0, \pm\sqrt{3}/2, \pm\sqrt{3}/2$

Table 4.1: Eigenvalues of the basis elements in the regular representation.

Eigenvalues	Eigenvectors
0	$(0, 0, 1, 0, 0, 0, 0, 0)$
0	$(0, 0, 0, 0, 0, 0, 0, 1)$
$\frac{1}{2}$	$(0, 0, 0, 0, 0, 1, -i, 0)$
$\frac{1}{2}$	$(0, 0, 0, 1, i, 0, 0, 0)$
$-\frac{1}{2}$	$(0, 0, 0, 1, -i, 0, 0, 0)$
$-\frac{1}{2}$	$(0, 0, 0, 0, 0, 1, i, 0)$
1	$(1, i, 0, 0, 0, 0, 0, 0)$
-1	$(1, -i, 0, 0, 0, 0, 0, 0)$

Table 4.2: Eigenvalues and eigenvectors of $R(\sigma_3)$.

μ	α_1	α_2	New Elements
1	0	0	$H_1 = \sigma_3$
2	0	0	$H_2 = \sigma_8$
3	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$E_{(\frac{1}{2}, \frac{\sqrt{3}}{2})} = \sigma_4 + i\sigma_5$
4	$-\frac{1}{2}$	$\frac{-\sqrt{3}}{2}$	$E_{(-\frac{1}{2}, -\frac{\sqrt{3}}{2})} = \sigma_4 - i\sigma_5$
5	$\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	$E_{(\frac{1}{2}, -\frac{\sqrt{3}}{2})} = \sigma_6 - i\sigma_7$
6	$-\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$E_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})} = \sigma_6 + i\sigma_7$
7	1	0	$E_{(1,0)} = \sigma_1 + i\sigma_2$
8	-1	0	$E_{(-1,0)} = \sigma_1 - i\sigma_2$

Table 4.3: Eigenvalues of H_1 and H_2 together with the new basis elements.

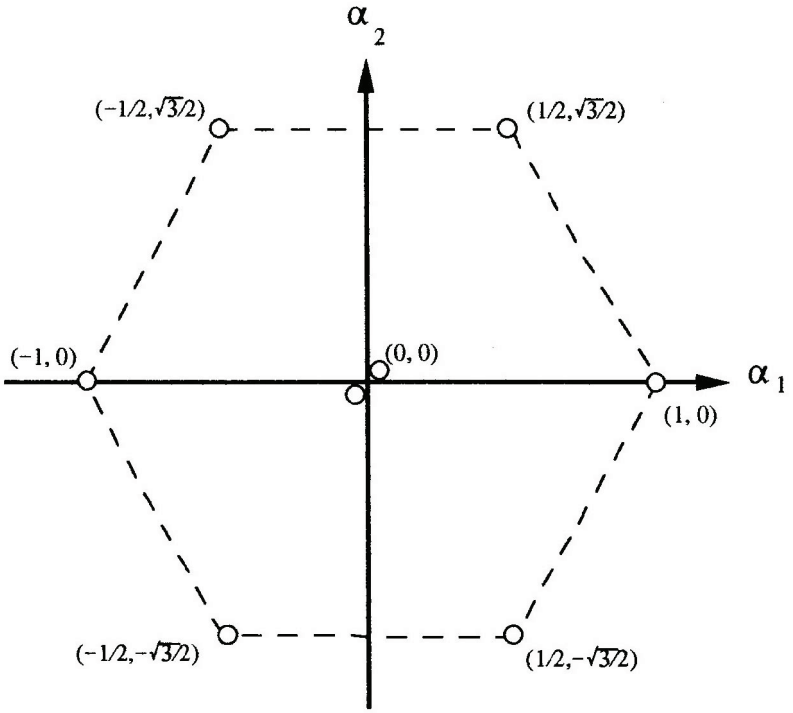


Figure 4.1: Rootspace diagram of the elements of $SU(3)$.

$$[H_1, E_{(\pm\frac{1}{2},\pm\frac{\sqrt{3}}{2})}] = \pm\frac{1}{2}E_{(\pm\frac{1}{2},\pm\frac{\sqrt{3}}{2})} \quad [H_2, E_{(\pm\frac{1}{2},\pm\frac{\sqrt{3}}{2})}] = \pm\frac{\sqrt{3}}{2}E_{(\pm\frac{1}{2},\pm\frac{\sqrt{3}}{2})}$$

$$[H_1, E_{(\pm\frac{1}{2},\mp\frac{\sqrt{3}}{2})}] = \pm\frac{1}{2}E_{(\pm\frac{1}{2},\mp\frac{\sqrt{3}}{2})} \quad [H_2, E_{(\pm\frac{1}{2},\mp\frac{\sqrt{3}}{2})}] = \mp\frac{\sqrt{3}}{2}E_{(\pm\frac{1}{2},\mp\frac{\sqrt{3}}{2})}$$

$$[H_1, E_{(\pm1,0)}] = \pm E_{(\pm1,0)} \qquad [H_2, E_{(\pm1,0)}] = 0$$

$$[H_1, H_2] = 0$$

$$[E_{(\frac{1}{2},\frac{\sqrt{3}}{2})}, E_{(-\frac{1}{2},-\frac{\sqrt{3}}{2})}] = H_1 + \sqrt{3}H_2$$

$$[E_{(\frac{1}{2},-\frac{\sqrt{3}}{2})}, E_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}] = H_1 - \sqrt{3}H_2$$

$$[E_{(1,0)}, E_{(-1,0)}] = 2H_1$$

$$[E_{(\frac{1}{2},\frac{\sqrt{3}}{2})}, E_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}] = 0 \qquad [E_{(\frac{1}{2},\frac{\sqrt{3}}{2})}, E_{(1,0)}] = 0$$

$$[E_{(-\frac{1}{2},-\frac{\sqrt{3}}{2})}, E_{(\frac{1}{2},-\frac{\sqrt{3}}{2})}] = 0 \qquad [E_{(-\frac{1}{2},-\frac{\sqrt{3}}{2})}, E_{(-1,0)}] = 0$$

$$[E_{(\frac{1}{2},-\frac{\sqrt{3}}{2})}, E_{(1,0)}] = 0 \qquad [E_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}, E_{(-1,0)}] = 0$$

$$[E_{(\frac{1}{2},\frac{\sqrt{3}}{2})}, E_{(\frac{1}{2},\frac{-\sqrt{3}}{2})}] = E_{(1,0)} \qquad [E_{(\frac{1}{2},\frac{\sqrt{3}}{2})}, E_{(-1,0)}] = -E_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}$$

$$[E_{(-\frac{1}{2},-\frac{\sqrt{3}}{2})}, E_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}] = -E_{(-1,0)} \quad [E_{(-\frac{1}{2},-\frac{\sqrt{3}}{2})}, E_{(1,0)}] = E_{(\frac{1}{2},-\frac{\sqrt{3}}{2})}$$

$$[E_{(\frac{1}{2},-\frac{\sqrt{3}}{2})}, E_{(-1,0)}] = E_{(-\frac{1}{2},-\frac{\sqrt{3}}{2})} \qquad [E_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}, E_{(1,0)}] = -E_{(\frac{1}{2},\frac{\sqrt{3}}{2})}$$

Table 4.4: Commutator relations for the new basis elements.

being we keep the root vectors in all their detail in order to illustrate Theorem 4.14. Note that the commutators (Table 4.4) fall into four general categories:
(1) $[\mathcal{V}_\circ, \mathcal{V}_\alpha] \subseteq \mathcal{V}_\alpha$ Commutators of this form are given by (4.52).
(2) $[\mathcal{V}_\circ, \mathcal{V}_\circ] = 0$ All members of the Cartan subalgebra commute.
(3) $[\mathcal{V}_\alpha, \mathcal{V}_{-\alpha}] \subseteq \mathcal{V}_\circ$ These define special elements in $\mathcal{V}_\circ$.
(4) $[\mathcal{V}_\alpha, \mathcal{V}_\beta] \subseteq \mathcal{V}_{\alpha+\beta}$
The last two categories will be reexamined in Section 4.4.

4.3 The Cartan Metric and the Killing Form

It would be useful to have a scalar product defined for any two elements of a Lie algebra. This is not quite possible for reasons that will appear shortly; but we can define a symmetric bilinear form, which is the next best thing. The bilinear form, called the "Killing form," requires a metric tensor defined as follows:

Definition 4.20 *The Cartan Metric*

Let $\mathcal{L}$ be a Lie algebra with a basis σ_a, $a = 1, \ldots, n$. The Cartan metric g_{ab} is defined by the contraction

$$g_{ab} = f^d_{ac} f^c_{bd} \tag{4.55}$$

where the f's are the structure constants of the basis,

$$[\sigma_a, \sigma_b] = f^c_{ab} \sigma_c.$$

This metric can be used to lower indices; for example,

$$f_{abc} = g_{ad} f^d_{bc}$$

is useful because it is totally antisymmetric in all three indices as can easily be verified from the Jacobi relation. There is no guarantee that g_{ab} has an inverse, and in fact, g_{ab} is identically zero for an abelian algebra. However, when $\det g \neq 0$ one can define an inverse g^{ab} by

$$g^{ab} g_{bc} = \delta^a_c$$

(δ^a_c is the unit matrix as usual) and use it to raise indices.

Definition 4.21 *The Killing form*

The product of two basis elements is defined as

$$(\sigma_a, \sigma_b) = g_{ab}. \tag{4.56}$$

The product of two arbitrary elements $X = x^a \sigma_a$ and $Y = y^b \sigma_b$ is then

$$(X, Y) = (x^a \sigma_a, y^b \sigma_b) = g_{ab} x^a y^b \tag{4.57}$$

The product (X, Y) is called the *Killing form*[5]. It is not positive definite and hence not a scalar product in the usual sense, but it is a symmetric bilinear

[5]The notation $B(X, Y)$ is often used.

form in that $(X,Y) = (Y,X)$ and $(X, \alpha Y + \beta Z) = \alpha(X,Y) + \beta(X,Z)$. If in addition $\det g \neq 0$, it is said to be *non-degenerate*.

In common with ordinary scalar products, (X,Y) is invariant under a change of basis. To see this let $S_a{}^b \sigma_b = \sigma'_a$ be a new basis. Then $X = (xS^{-1})^a (S\sigma)_a = x'^a \sigma'_a$, and $g'_{ab} = (\sigma'_a, \sigma'_b)$. Finally

$$g'_{ab} x'^a y'^b = (xS^{-1})^a (yS^{-1})^b S_a{}^c S_b{}^d g_{cd} = x^a y^b g_{ab}$$

Another useful property is

$$(X,[Y,Z]) = (Y,[Z,X]) = (Z,[X,Y]) = f_{abc} x^a y^b z^c. \tag{4.58}$$

The Killing form can also be written in terms of the adjoint representation as follows:

$$(X,Y) = g_{ab} x^a y^b = f^d_{ac} f^c_{bd} x^a y^b = \boldsymbol{R}(X)^d_{\ c} \boldsymbol{R}(Y)^c_{\ d}.$$

Thus

$$(X,Y) = \mathrm{Trace}\{\boldsymbol{R}(X)\boldsymbol{R}(Y)\} \tag{4.59}$$

or

$$g_{ab} = \mathrm{Trace}[\boldsymbol{R}(\sigma_a)\boldsymbol{R}(\sigma_b)] \tag{4.60}$$

Suppose $\mathcal{A}$ is a subalgebra of $\mathcal{L}$ and X and Y are elements of $\mathcal{A}$. It is a peculiar property of the Killing form that (X,Y) has two different values depending on whether we think of $\mathcal{A}$ as a subalgebra of $\mathcal{L}$ or an algebra unto itself. This can easily be seen from (4.13). If σ_i and σ_j are bases in $\mathcal{A}$ then

$$(\sigma_i, \sigma_j) = \mathrm{trace} \left\{ \begin{bmatrix} f^k_{il} & f^k_{i\beta} \\ 0 & f^\gamma_{i\beta} \end{bmatrix} \begin{bmatrix} f^l_{jm} & f^l_{j\sigma} \\ 0 & f^\beta_{j\sigma} \end{bmatrix} \right\}$$

$$= \mathrm{trace} \begin{bmatrix} f^k_{il} f^l_{jm} & f^k_{il} f^l_{j\sigma} + f^k_{i\beta} f^\beta_{j\sigma} \\ 0 & f^\gamma_{i\beta} f^\beta_{j\sigma} \end{bmatrix}$$

$$= f^k_{il} f^l_{jk} + f^\gamma_{i\beta} f^\beta_{j\gamma} \tag{4.61}$$

The second term involves sums over Greek letters, which index the complement of $\mathcal{A}$. If $\mathcal{A}$ is an algebra (rather than a proper subalgebra) the

complement is empty, and this term is zero. The notation $(X,Y)_{\mathcal{A}}$ will refer to the first term in (4.61), *i.e.* the scalar product of X and Y restricted to $\mathcal{A}$. This refinement is not necessary if $\mathcal{A}$ is an invariant subalgebra, because the lower right partition of $\boldsymbol{R}(X)$ is zero, and $(X,Y) = (X,Y)_{\mathcal{A}}$.

Now let Y be an arbitrary element in $\mathcal{L}$, $X \in \mathcal{A}$, and $\mathcal{A}$ invariant. Then $Y = \sigma_i y^i + \sigma_\alpha y^\alpha$, and $X = \sigma_i x^i$. Using (4.15) and (4.16) we can calculate

$$(X,Y) = x^i f^k_{il}(y^j f^l_{jk} + y^\alpha f^l_{\alpha k}) \tag{4.62}$$

The implied sums involve only Roman indices, so $(X,Y) = (X,Y)_{\mathcal{A}}$.

The Killing form can be used to test for solvable subalgebras and "disect" them out. The following result is usually called "Cartan's criterion."

Theorem 4.16 *An algebra is solvable if and only if the Cartan metric is zero on the first derived algebra.*

Proof: If $\mathcal{L}$ is solvable it can be put in upper triangular form (Corollary 4.14). $\mathcal{L}^{(1)}$ has zeros along the diagonal, and the trace in (4.60) vanishes.

To prove the converse suppose that the Cartan metric is zero on $\mathcal{L}^{(1)}$ but that $\mathcal{L}$ is not solvable. Then the derived series terminates on some $\mathcal{L}^{(n)}$ so that

$$[\mathcal{L}^{(n)}, \mathcal{L}^{(n)}] \equiv \mathcal{L}^{(n+1)} \equiv \mathcal{L}^{(n)}.$$

Regard $\mathcal{L}^{(n)}$ as a Lie algebra in its own right. Let $\mathcal{V}^{(n)}_\circ$ be the Cartan subalgebra of $\mathcal{L}^{(n)}$ and $X_\circ$ an arbitrary element in $\mathcal{V}^{(n)}_\circ$. By assumption

$$0 = (X_\circ, X_\circ)_{\mathcal{L}^{(1)}} = (X_\circ, X_\circ)_{\mathcal{L}^{(n)}}$$

$$= \text{Trace}\{\boldsymbol{R}(X_\circ)\boldsymbol{R}(X_\circ)\}_{\mathcal{L}^{(n)}} = \sum_{\beta \in \mathcal{L}^{(n)}} d_\beta \beta^2.$$

The second equality follows because $\mathcal{L}^{(n)}$ is an invariant subalgebra of $\mathcal{L}^{(1)}$; β stands for the roots of $\boldsymbol{R}(X_\circ)$, and d_β is their multiplicity. Consequently all the roots of $\boldsymbol{R}(X_\circ)$ restricted to $\mathcal{L}^{(n)}$ are zero, and all the elements of $\mathcal{L}^{(n)}$ are in the Cartan subalgebra, which is solvable. This contradiction establishes the theorem.

We can prove a stronger result by noting that if $\det g = 0$, the equation $g_{ab}x^b = 0$ has solutions. This means that there are elements X such that $(X,Y) = 0$ for all $Y \in \mathcal{L}$. It is easy to prove that these X constitute an ideal $\mathcal{L}_\circ$: if $Z \in \mathcal{L}$ then (4.58) gives $([Z,X],Y) = ([Y,Z],X) = 0$, *i.e.* $[\mathcal{L}_\circ, \mathcal{L}] \subseteq \mathcal{L}_\circ$. If $\mathcal{L}_\circ$ is solvable, then $\mathcal{L}$ is not semisimple, because the

penultimate derived algebra of $\mathcal{L}^{(\circ)}$ would be abelian. If we assume $\mathcal{L}_\circ$ is not solvable, then we can repeat the proof of Theorem (4.16) with $\mathcal{L}_\circ$ rather than $\mathcal{L}$ and arrive at a contradiction. This proves that if $\mathcal{L}$ is semisimple, then g is nonsingular. The converse is also true. If $\mathcal{L}$ is not semisimple it has an invariant abelian subalgebra. Then the term f_{ij}^k in (4.15) is zero, and consequently the trace of both $\boldsymbol{R}(\sigma_i)\boldsymbol{R}(\sigma_j)$ and $\boldsymbol{R}(\sigma_i)\boldsymbol{R}(\sigma_\alpha)$ vanishes. Then g as given by (4.60) has some rows and columns of zeros, and so $\det g = 0$. We have just proved

Corollary 4.17 *$\mathcal{L}$ is semisimple if and only if the Cartan metric is nonsingular.*

If $\mathcal{L}$ is semisimple but not simple it can be decomposed into a direct sum of simple subalgebras. To see this let $\mathcal{A}$ be an invariant subalgebra of $\mathcal{L}$. As usual we use the basis σ_j to span $\mathcal{A}$ and σ_α to span the complement of $\mathcal{A}$. Since $\det g \neq 0$ it is possible to choose a new basis σ'_α for the complementary space so that $(\sigma_i, \sigma'_\alpha) = 0$.

$$\sigma'_\alpha = \sigma_\alpha - \sum_{i,j} \sigma_i g^{ij}(\sigma_j, \sigma_\alpha)$$

where $g_{ij} = (\sigma_i, \sigma_j)$, and g^{ij} is its inverse. As a result of this "Gram-Schmidt orthogonalization" the space $\mathcal{L}$-$\mathcal{A}$ is orthogonal to $\mathcal{A}$ in the sense that if $X \in \mathcal{A}$ and $Y \in \mathcal{L}$-$\mathcal{A}$, then $(X, Y) = 0$. Moreover, $\mathcal{L}$-$\mathcal{A}$ is an ideal, *i.e.* if Y' is an additional element in $\mathcal{L}$-$\mathcal{A}$, and $X' \in \mathcal{A}$, then

$$(X, [Y', Y]) = (Y', [Y, X]) = 0,$$

so $\mathcal{L}$-$\mathcal{A}$ is a subalgebra, and

$$(X', [Y, X]) = (Y, [X, X']) = 0,$$

so it is invariant. Evidentally the element $[X, Y]$ is orthogonal to everything in $\mathcal{A}$ and $\mathcal{L}$-$\mathcal{A}$. Since the scalar product is non-degenerate the commutator must be zero. If either $\mathcal{A}$ or $\mathcal{L}$-$\mathcal{A}$ contains an ideal this procedure can be repeated with it. Continue in this way until there are no proper ideals remaining. The subalgebras are then all simple, and their elements commute with those of all other subalgebras.

Corollary 4.18 *Any semisimple algebra can be decomposed into a direct sum of simple algebras.*

$$\mathcal{L} = \mathcal{S}_1 \oplus \mathcal{S}_2 \oplus \mathcal{S}_3 \cdots$$

The adjoint representation of these subalgebras will have the form of (4.18) and (4.19), so the adjoint representation of $\mathcal{L}$ is completely reducible, and the adjoint representation of each $\mathcal{S}_i$ is irreducible.

Theorem 4.19 *The regular representation of a semisimple algebra is faithful.*

Proof: Suppose $\boldsymbol{R}(X) = \boldsymbol{R}(Y)$. Then $\boldsymbol{R}(X) - \boldsymbol{R}(Y) = \boldsymbol{R}(X - Y) = 0$, and every element in the algebra is orthogonal to $(X - Y)$. Since the metric is non-singular, this is not possible unless $X = Y$.

4.4 Properties of the Roots

In the previous two sections we developed the root space decomposition and Killing form independently. The full power of these ideas appears when they are combined. In this section we will assume that the rootspace decomposition has been performed so that the elements can be labeled with the root vectors $\boldsymbol{\alpha}$. Furthermore, we will only deal with semisimple algebras so that the Killing form is non-degenerate. We first examine the structure of the Cartan metric on the root subspaces.

Theorem 4.20 *If $X_a \in \mathcal{V}_\alpha$ and $X_b \in \mathcal{V}_\beta$, $(X_a, X_b) = 0$ unless $\boldsymbol{\alpha} + \boldsymbol{\beta} = 0$.*

Proof: This is a simple consequence of Theorem (4.14). Suppose the index d in (4.55) refers to elements in $\mathcal{V}_\delta$ and c refers to elements in $\mathcal{V}_\gamma$. Then $f^d_{ac} = 0$ unless $\boldsymbol{\alpha}+\boldsymbol{\gamma} = \boldsymbol{\delta}$; $f^c_{bd} = 0$ unless $\boldsymbol{\beta}+\boldsymbol{\delta} = \boldsymbol{\gamma}$; and $g_{ab} = 0$ unless $\boldsymbol{\alpha} = -\boldsymbol{\beta}$.

Theorem 4.21 *The Cartan subalgebra is abelian.*

Proof: Let H and H' be elements in $\mathcal{V}_\circ$ and $X_\alpha \in \mathcal{V}_\alpha$, $\boldsymbol{\alpha} \neq 0$. Then

$$([H, H'], X_\alpha) = 0.$$

Now let H'' be any element in $\mathcal{V}_\circ$.

$$([H, H'], H'') = \text{trace}\{\boldsymbol{R}([H, H'])\boldsymbol{R}(H'')\} =$$

$$\text{trace}\{\boldsymbol{R}(H)\boldsymbol{R}(H')\boldsymbol{R}(H'') - \boldsymbol{R}(H')\boldsymbol{R}(H)\boldsymbol{R}(H'')\}$$

The Cartan subalgebra is solvable (Theorem 4.15), so by Lie's theorem it can be put in upper triangular form. The trace of a product of upper triangular matrices is independent of their order, so the difference vanishes. We have shown that $[H, H']$ is orthogonal to all elements and hence zero.

Theorem 4.22 *The space $\mathcal{V}_{-\alpha}$ is the adjoint (with respect to the Killing form) of $\mathcal{V}_\alpha$.*

Proof: We are assuming that the Killing form is non-degenerate. Thus for every element $X \in \mathcal{L}$ there is at least one other $X' \in \mathcal{L}$ for which $(X', X) \neq 0$. Theorem 4.20, however, requires that if $X \in \mathcal{V}_\circ$, then $X' \in \mathcal{V}_\circ$; and if $X \in \mathcal{V}_\alpha$, then $X' \in \mathcal{V}_{-\alpha}$. It follows that $\mathcal{V}_{-\alpha}$ is the adjoint of $\mathcal{V}_\alpha$ and that $\mathcal{V}_\circ$ is its own adjoint space. This argument also shows that the Kiilling

form is non-degenerate even when it is restricted to $\mathcal{V}_\circ$ or to any pair of subspaces $\mathcal{V}_\alpha$ and $\mathcal{V}_{-\alpha}$.

The argument leading up to (4.50) shows that for each root subspace $\mathcal{V}_\alpha$ there is "at least one" simultaneous eigenvector E_α such that

$$[H_i, E_\alpha] = \alpha_i E_\alpha$$

for all $H_i \in \mathcal{V}_\circ$. It will turn out that there is in fact only one. We are not entitled to assume this, however; so we will be coy and say that for each (putative) $E_\alpha \in \mathcal{V}_\alpha$, there is some corresponding $X_{-\alpha} \in \mathcal{V}_{-\alpha}$ such that $(E_\alpha, X_{-\alpha}) \neq 0$.

$$(E_\alpha, X_{-\alpha}) \equiv \lambda_\alpha$$

The value of λ_α must be decided by some convention. We will return to this point later.

What about the commutator of E_α and $X_{-\alpha}$? It must be an element of $\mathcal{V}_\circ$, so

$$[E_\alpha, X_{-\alpha}] \equiv \lambda_\alpha \alpha^i H_i \equiv \lambda_\alpha h_\alpha. \tag{4.63}$$

Here H_i denotes the usual set of basis elements for $\mathcal{V}_\circ$. Equation (4.63) defines a set of contravariant root components α^i and a specific $h_\alpha \in \mathcal{V}_\circ$. The following relationships follow from (4.63), (4.58), and (4.52).

$$(H_i, [E_\alpha, X_{-\alpha}]) = ([H_i, E_\alpha], X_{-\alpha}) =$$

$$\lambda_\alpha (H_i, \alpha^j H_j) = \lambda_\alpha \alpha^j (H_i, H_j) = (\alpha_i E_\alpha, X_{-\alpha}) = \lambda_\alpha \alpha_i$$

Since the "metric,"

$$h_{ij} = (H_i, H_j),$$

is non-degenerate, it has an inverse and can be used to raise and lower indices. For example, if $\boldsymbol{\alpha}, \boldsymbol{\beta} \neq 0$ are roots, and $[E_\alpha, X_{-\alpha}] = \lambda_\alpha h_\alpha$, $[E_\beta, X_{-\beta}] = \lambda_\beta h_\beta$, then

$$(h_\alpha, h_\beta) = \alpha^i \beta^j h_{ij} = \alpha^i \beta_i \equiv (\boldsymbol{\alpha}, \boldsymbol{\beta}). \tag{4.64}$$

We will use the notation $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ to describe the bilinear form in the l-dimensional space of the root components. In this notation,

$$[h_\beta, E_\alpha] = \alpha(h_\beta) E_\alpha = \beta^i [H_i, E_\alpha] = \beta^i \alpha_i E_\alpha = (\boldsymbol{\alpha}, \boldsymbol{\beta}) E_\alpha, \tag{4.65}$$

$$[h_\alpha, E_\beta] = \beta(h_\alpha) E_\beta = (\boldsymbol{\beta}, \boldsymbol{\alpha}) E_\beta,$$

and

$$\beta(h_\alpha) = \alpha(h_\beta) = (\boldsymbol{\alpha}, \boldsymbol{\beta}).$$

The regular representation of h_α and h_β will be in Jordan canonical form with the roots $\gamma(h_\alpha)$ and $\gamma(h_\beta)$ along the main diagonal. The product of h_α and h_β is then

$$(h_\alpha, h_\beta) = \sum_\gamma d_\gamma \gamma(h_\alpha)\gamma(h_\beta) = \sum_\gamma d_\gamma (\boldsymbol{\alpha}, \boldsymbol{\gamma})(\boldsymbol{\beta}, \boldsymbol{\gamma}) = (\boldsymbol{\alpha}, \boldsymbol{\beta}), \tag{4.66}$$

where we have used (4.64) and (4.65).

Theorem 4.23 *For all non-zero roots $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, the quantity $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is real and rational, and $(\boldsymbol{\alpha}, \boldsymbol{\alpha})$ is real, rational, and positive.*

Proof: Let $\mathcal{F}$ be the direct sum of all subspaces of the form $\mathcal{V}_{\beta+n\alpha}$ where n is an integer. Then

$$[\mathcal{V}_{\pm\alpha}, \mathcal{F}] \subseteq \mathcal{F},$$

and

$$[\mathcal{V}_\circ, \mathcal{F}] \subseteq \mathcal{F}.$$

Thus $\mathcal{F}$ is an invariant subalgebra. We can regard it as an algebra in its own right and calculate the adjoint representation of h_α. This will be a matrix in Jordan canonical form with roots $(\boldsymbol{\alpha}, \boldsymbol{\beta} + n\boldsymbol{\alpha})$ along the main diagonal. The trace of this matrix must be zero , however, because h_α can be written as a commutator. Thus

$$\text{trace}\{\boldsymbol{R}(h_\alpha)\} = \sum_n \{(\boldsymbol{\beta}, \boldsymbol{\alpha}) + n(\boldsymbol{\alpha}, \boldsymbol{\alpha})\} d_{\beta+n\alpha} = 0 \tag{4.67}$$

From (4.66)

$$(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = \sum_{\beta \subset \mathcal{F}} (\boldsymbol{\alpha}, \boldsymbol{\beta})^2. \tag{4.68}$$

Obviously $(\boldsymbol{\alpha}, \boldsymbol{\alpha})$ can't be negative. It can't be zero either for this would imply $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (h_\alpha, h_\beta) = 0$ for all $\boldsymbol{\beta}$. Combining (4.67) and (4.68) gives

$$(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = \sum_\beta \left[\frac{\sum n d_{\beta+n\alpha}}{\sum d_{\beta+n\alpha}} \right]^2 (\boldsymbol{\alpha}, \boldsymbol{\alpha})^2,$$

or

$$(\boldsymbol{\alpha}, \boldsymbol{\alpha})^{-1} = \sum_\beta \left[\frac{\sum n d_{\beta+n\alpha}}{\sum d_{\beta+n\alpha}} \right]^2.$$

Since $\sum n d_{\beta+n\alpha}$ and $\sum d_{\beta+n\alpha}$ are necessarily positive integers, $(\boldsymbol{\alpha}, \boldsymbol{\alpha})$ is real, rational, and positive, and $(\boldsymbol{\beta}, \boldsymbol{\alpha})$ is real and rational.

Incidentally, since

$$(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = \alpha^i \alpha^j h_{ij} > 0,$$

the metric h_{ij} is positive definite, and we can legitimately speak of a "scalar product" of two root vectors.

Theorem 4.24 *If $\boldsymbol{\alpha}$ is a non-zero root then the spaces $\mathcal{V}_\alpha$ and $\mathcal{V}_{-\alpha}$ are one-dimensional, and the spaces $\mathcal{V}_{n\alpha}$ are empty unless $n = \pm 1$ or 0.*

Proof: Let $\mathcal{F}$ be the space spanned by $X_{-\alpha}$, h_α, and all subspaces of the form $\mathcal{V}_{-n\alpha}$ where n is a positive integer. Then

$$[X_{-\alpha}, \mathcal{F}] \subseteq \mathcal{F},$$

$$[h_\alpha, \mathcal{F}] \subseteq \mathcal{F},$$

and

$$[E_\alpha, \mathcal{F}] \subseteq \mathcal{F}.$$

We repeat the argument from the previous theorem. The adjoint representation of h_α is a matrix in Jordan canonical form with eigenvalues along the main diagonal.

$$\text{trace}\{\boldsymbol{R}(h_\alpha)\} = -\alpha(h_\alpha) + 0 + \sum_n d_{-n\alpha} n\alpha(h_\alpha) = (\boldsymbol{\alpha}, \boldsymbol{\alpha})\{-1 + \sum_{n=1} n d_{-n\alpha}\} = 0$$

There is only one way to satisfy this equation, $d_{-\alpha} = 1$ and $d_{-n\alpha} = 0$ for $n \geq 2$.

We can repeat this argument with an $\mathcal{F}$ consisting of $X_{-\alpha}$, h_α, and all subspaces of the form $\mathcal{V}_{n\alpha}$. This leads to the conclusion that $d_\alpha = 1$ and $d_{n\alpha} = 0$ for $n \geq 2$. To put it bluntly, all the A_{α_j} in (4.42) are one-dimensional, and all the elements of $\mathcal{V}_\circ$ are diagonal in the adjoint representation.

It is possible to bypass the difficulties in this chapter with the following argument (Cornwell, Georgi): one *defines* the Cartan subalgebra as the largest set of commuting elements in the algebra. The claim is then made that since the matrices commute, they can be simultaneously diagonalized. This avoids all the tedious business about root space decomposition, but it leaves several important questions unanswered. For one thing, the Cartan subalgebra of non-semisimple algebras is not in general diagonal. The claim that commuting matrices can be simultaneously diagonalized is not true in general for non-Hermitian matrices, so Hermiticity has to be taken as an additional assumption. The root space decomposition also gives us a recipe

for finding the largest possible set of commuting matrices (How else would you know that you had found them all?), and a body of insight regarding why the semisimple algebras have such simple structures.

Since the $\mathcal{V}_\alpha$ subspaces are all one-dimensional, we can identify $X_{-\alpha}$ with $E_{-\alpha}$ and arrive at our canonical commutation relations.

$$[E_\alpha, E_{-\alpha}] = \lambda_\alpha h_\alpha \tag{4.69}$$
$$[h_\alpha, E_\beta] = (\boldsymbol{\alpha}, \boldsymbol{\beta}) E_\beta \tag{4.70}$$
$$[h_\alpha, h_\beta] = 0 \tag{4.71}$$

There is an extremely important simple subalgebra burried in these relations. Replacing $\beta \to \alpha$ gives

$$[E_\alpha, E_{-\alpha}] = \lambda_\alpha h_\alpha \tag{4.72}$$
$$[h_\alpha, E_\alpha] = (\boldsymbol{\alpha}, \boldsymbol{\alpha}) E_\alpha, \tag{4.73}$$

which is just the algebra of $O(3)$ or $SU(2)$, *i.e.* the algebra of angular momentum. Thus any semisimple algebra can be decomposed into a sum of angular momentum-like subalgebras. This is not the whole story, however, because they are not invariant subalgebras. We still need to study commutators of the form,

$$[E_\alpha, E_\beta] = N_{\alpha,\beta} E_{\alpha+\beta}, \tag{4.74}$$

for those cases where $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\alpha} + \boldsymbol{\beta}$ are all non-zero roots. The structure constants $N_{\alpha,\beta}$ like λ_α depend on the normalization of the E_α and h_α. We will assume the following:

(a) Since the metric $h_{ij} = (H_i, H_j)$ is positive definite, we are entitled to perform Grahm-Schmidt orthogonalization on the H_i so that $h_{ij} = \delta_{ij}$.

(b) $\lambda_\alpha = (E_\alpha, E_{-\alpha}) = 1$ for all pairs $\boldsymbol{\alpha}$ and $-\boldsymbol{\alpha}$.

(c) All the $N_{\alpha,\beta}$ are real, and $N_{-\alpha,-\beta} = -N_{\alpha,\beta}$.

We will assume these conventions in all subsequent calculations. Unfortunately, there are at least three other conventions in general use. (See Cornwell for a review of the different possibilities.)

The following example shows how the angular momentum commutation relations are restated with these conventions.

Example 4.6 *Angular Momentum*

The commutation relations,

$$[J_i, J_j] = \imath\hbar\epsilon_{ijk} J_k,$$

define the structure constants from which the adjoint representation is calculated.

$$\boldsymbol{R}(J_1) = \imath\hbar \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad \boldsymbol{R}(J_2) = \imath\hbar \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{R}(J_3) = \imath\hbar \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

It is customary to use the z component for the Cartan subalgebra.

$$H_1 = \frac{1}{\sqrt{2}\hbar}\boldsymbol{R}(J_3)$$

The eigenvalues of H_1 are $0, \pm\sqrt{1/2}$. The eigenvectors corresponding to $\pm\sqrt{1/2}$ are proportional to the familiar raising and lowering operators. We choose

$$E_\pm = \frac{1}{2\hbar}\boldsymbol{R}(J_1 \pm \imath J_2).$$

The following relations follow simply from the previous definitions.

$$(H_1, H_1) = 1 \qquad (E_+, E_-) = \lambda = 1$$

$$[E_+, E_-] = \alpha^1 H_1 = \frac{1}{\sqrt{2}} H_1 = h$$

Evidentally, $\alpha^1 = 1/\sqrt{2}$ and h (*i.e.* h_α)$= H_1/\sqrt{2}$.

$$[H_1, E_+] = \alpha_1 E_+ = \frac{1}{\sqrt{2}} E_+$$

$$[h, E_+] = (\boldsymbol{\alpha}, \boldsymbol{\alpha}) E_+ = \frac{1}{2} E_+$$

The scalar product $(\boldsymbol{\alpha}, \boldsymbol{\alpha})$ plays the role of a dimensionless structure constant. It is positive, real, and rational as promised; furthermore, $\hbar$ and $\imath$ have completely disappeared from the commutation relations.

We now proceed to derive the properties of the $N_{\alpha,\beta}$.

Theorem 4.25 *If $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\alpha}+\boldsymbol{\beta}$ are all non-zero roots, then $\boldsymbol{\gamma} = -\boldsymbol{\alpha}-\boldsymbol{\beta}$ is also a root, and*

$$N_{\alpha,\beta} = N_{\beta,\gamma} = N_{\gamma,\alpha}.$$

Proof:

$$[E_\alpha, [E_\beta, E_\gamma]] + [E_\beta, [E_\gamma, E_\alpha]] + [E_\gamma, [E_\alpha, E_\beta]] = 0$$

$$[E_\alpha, E_{-\alpha}]N_{\beta,\gamma} + [E_\beta, E_{-\beta}]N_{\gamma,\alpha} + [E_\gamma, E_{-\gamma}]N_{\alpha,\beta} = 0$$

$$H_\alpha N_{\beta,\gamma} + H_\beta N_{\gamma,\alpha} + H_\gamma N_{\alpha,\beta} = 0$$

However, $\boldsymbol{\alpha} + \boldsymbol{\beta} + \boldsymbol{\gamma} = 0$ implies that $H_\alpha + H_\beta + H_\gamma = 0$

$$(N_{\beta,\gamma} - N_{\alpha,\beta})H_\alpha + (N_{\gamma,\alpha} - N_{\alpha,\beta})H_\beta = 0$$

Finally, H_α and H_β are independent, so the theorem is proved.

It is possible to calculate the $N_{\alpha,\beta}$ structure constants using an extension of the proof of Theorem 4.23. As usual let $\mathcal{F}$ be the direct sum of all subspaces of the form $\mathcal{V}_{\beta+k\alpha}$. The integer k ranges from $-m$ to n where m is the largest positive integer such that $\boldsymbol{\beta} - m\boldsymbol{\alpha}$ is a root, and n is the largest positive integer such that $\boldsymbol{\beta} + n\boldsymbol{\alpha}$ is a root. Now apply the Jacobi identity to the elements E_α, $E_{\beta+k\alpha}$, and E_α.

$$[E_\alpha, [E_{\beta+k\alpha}, E_\alpha]] + [E_{\beta+k\alpha}, [E_\alpha, E_\alpha]] + [E_\alpha, [E_\alpha, E_{\beta+k\alpha}]] = 0$$

$$[E_\alpha, E_{\beta+(k-1)\alpha}]N_{\beta+k\alpha,-\alpha} - [E_{\beta+k\alpha}, H_\alpha] + [E_\alpha, E_{\beta+(k+1)\alpha}]N_{\alpha,\beta+k\alpha} = 0$$

$$N_{\alpha,\beta+(k-1)\alpha}N_{\beta+k\alpha,-\alpha} + N_{-\alpha,\beta+(k+1)\alpha}N_{\alpha,\beta+k\alpha} = -\alpha^i(\beta_i + k\alpha_i)$$

Using the results of Theorem 4.25 and the normalization convention $N_{-\alpha,-\beta} = -N_{\alpha,\beta}$, this becomes

$$N^2_{\alpha,\beta+(k-1)\alpha} = N^2_{\alpha,\beta+k\alpha} + \alpha^i(\beta_i + k\alpha_i)$$

We solve this recursion relation "from the top down." Our definition of n guarantees that

$$N_{\alpha,\beta+n\alpha} = 0,$$

so that

$$N^2_{\alpha,\beta+(n-1)\alpha} = \alpha^i(\beta_i + n\alpha_i).$$

A simple iterative calculation leads to

$$N^2_{\alpha,\beta+(k-1)\alpha} = (n-k+1)\left\{(\boldsymbol{\alpha},\boldsymbol{\beta}) + \frac{1}{2}(n+k)(\boldsymbol{\alpha},\boldsymbol{\alpha})\right\}. \tag{4.75}$$

This sequence must terminate at some negative value of k. First note that

$$[E_{-\alpha}, E_{\beta-m\alpha}] = N_{-\alpha,\beta-m\alpha}E_{\beta-(m+1)\alpha} = 0,$$

because $\boldsymbol{\beta} - (m+1)\boldsymbol{\alpha}$ is not a root. Consequently

$$N_{-\alpha,\beta-m\alpha} = -N_{\alpha,-\beta+m\alpha} = 0.$$

Use the results of Theorem 4.25 with $-\boldsymbol{\gamma} = \boldsymbol{\alpha}-\boldsymbol{\beta}+m\boldsymbol{\alpha}$ or $\boldsymbol{\gamma} = \boldsymbol{\beta}-(m+1)\boldsymbol{\alpha}$ yields

$$N_{\alpha,\beta-(m+1)\alpha} = 0.$$

Consequently the sequence terminates when $k = -m$. The recursion relation yields

$$(n+m+1)\left\{(\boldsymbol{\alpha},\boldsymbol{\beta}) + \frac{1}{2}(n-m)(\boldsymbol{\alpha},\boldsymbol{\alpha})\right\} = 0 \tag{4.76}$$

Since n and m are both positive integers,

$$n \geq -\frac{2(\boldsymbol{\alpha},\boldsymbol{\beta})}{(\boldsymbol{\alpha},\boldsymbol{\alpha})} = n - m \geq -m. \tag{4.77}$$

The ratio $-2(\boldsymbol{\alpha},\boldsymbol{\beta})/(\boldsymbol{\alpha},\boldsymbol{\alpha})$ is an integer that, like k, ranges from $-m$ to n. Consequently, if $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are both roots, then

$$\boldsymbol{\beta}' = \boldsymbol{\beta} - \frac{2(\boldsymbol{\alpha},\boldsymbol{\beta})}{(\boldsymbol{\alpha},\boldsymbol{\alpha})}\boldsymbol{\alpha} \tag{4.78}$$

is also a root. Substituting (4.78) back into (4.75) and setting $k = 1$ yields

$$N_{\alpha,\beta} = \sqrt{n(m+1)}\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})/2} \tag{4.79}$$

These theorems enable us to express the commutation rules for any semisimple algebra in terms of a discrete Euclidian space, the space of the root vectors. The dimension of this space is equal to the dimension of the Cartan subalgebra. The basis elements appear as points in a simple lattice structure controlled by the two integers, m and n. The lattice has some

simple symmetry properties: it is invariant under the exchange $\boldsymbol{\alpha} \to -\boldsymbol{\alpha}$, and the relative positions of the points on the lattice are given by (4.77) and (4.78). It is not suprising, in view of this, that it is possible to give a precise catalog of all semisimple algebras. This is the task to which we turn in the next section.

Example 4.7 *SU*(3) *Revisited*

The new basis elements in Table 4.3 need a new set of structure constants, which can be calculated from (4.54) by a simple linear mapping. A tedious but straightforward calculation using (4.55) then gives the Cartan metric tensor.

$$g_{\mu,\nu} = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 \end{bmatrix}$$

The elements $\mu, \nu = 1, 2$ refer to the Cartan subalgebra. Evidentally the metric

$$h_{i,j} = (H_i, H_j) = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

The remaining 6 elements are grouped in pairs of $\boldsymbol{\alpha}$ and $-\boldsymbol{\alpha}$. One sees the structure predicted by Theorem 4.20. The normalization is such that

$$(E_\alpha, E_{-\alpha}) = \lambda_\alpha = 6$$

The scalar products of the root vectors can be found from (4.64) or (4.65).

$$(h_3, h_3) = (h_5, h_5) = (h_7, h_7) = 1/3$$

$$(h_3, h_7) = (h_5, h_7) = -(h_3, h_5) = -1/6$$

Theorem 4.25 predicts that $-2(\boldsymbol{\alpha}, \boldsymbol{\beta})/(\boldsymbol{\alpha}, \boldsymbol{\alpha})$ is an integer, in this case ± 1.

In order to adhere to our normalization conventions and make $\lambda = 1$ we should redefine $E_\alpha \to E_\alpha/\sqrt{(}6)$, but this would not change the root vectors. Finally, we could redefine $H_i \to H_i/\sqrt{(}3)$. In this way all the elements of the Cartan tensor would be either 0 or 1. This last change would rescale the root vectors but leave the scalar products of root vectors unchanged.

θ	$\frac{2(\boldsymbol{\alpha},\boldsymbol{\beta})}{(\boldsymbol{\alpha},\boldsymbol{\alpha})} = n$	$\frac{2(\boldsymbol{\alpha},\boldsymbol{\beta})}{(\boldsymbol{\beta},\boldsymbol{\beta})} = n'$	$\frac{(\boldsymbol{\alpha},\boldsymbol{\alpha})}{(\boldsymbol{\beta},\boldsymbol{\beta})}$
30°	3	1	1/3
	1	3	3
45°	2	1	1/2
	1	2	2
60°	1	1	1
90°	0	0	0/0

Table 4.5: The allowed angles between two roots α and β.

4.5 Classification of semisimple algebras

We have seen how semisimple algebras can be described in terms of a descrete, Euclidian root space endowed with a positive definite scalar product. The root vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ obey very restrictive conditions given by (4.77) and (4.78). To display the content of (4.77) we calculate the angle θ between any two root vectors.

$$0 \leq \cos^2\theta = \frac{(\boldsymbol{\alpha},\boldsymbol{\beta})(\boldsymbol{\alpha},\boldsymbol{\beta})}{(\boldsymbol{\alpha},\boldsymbol{\alpha})(\boldsymbol{\beta},\boldsymbol{\beta})} = \frac{n}{2}\cdot\frac{n'}{2} \leq 1 \tag{4.80}$$

The n and n' are integers, since $2(\boldsymbol{\alpha},\boldsymbol{\beta})/(\boldsymbol{\alpha},\boldsymbol{\alpha})$ must be an integer, but they are restricted to the range $0 \leq nn' \leq 4$. The only possibilities for n and n' are the so-called Cartan integers, $0, \pm1, \pm2, \pm3$, and ±4. The allowed angles are $0°, 30°, 45°, 60°, 90°, 120°, 135°, 150°$, and $180°$. The angles larger than $90°$ are redundant, however, because of the symmetry between positive and negative roots, and $0°$ is trivial. This leaves four interesting possibilities listed in Table 4.5.

The additional restriction imposed by (4.78) can be interpreted geometrically as follows: define

$$\boldsymbol{\beta} = \boldsymbol{\beta}_{\|} + \boldsymbol{\beta}_{\perp} = \left[\boldsymbol{\beta} - \frac{\boldsymbol{\alpha}(\boldsymbol{\alpha},\boldsymbol{\beta})}{(\boldsymbol{\alpha},\boldsymbol{\alpha})}\right] + \frac{(\boldsymbol{\alpha},\boldsymbol{\beta})\boldsymbol{\alpha}}{(\boldsymbol{\alpha},\boldsymbol{\alpha})}. \tag{4.81}$$

The $\|$ and $\perp$ signs refer to a plane passing through the origin and perpendicular to $\boldsymbol{\alpha}$ called the **$\boldsymbol{\alpha}$-hyperplane**. The vectors $\boldsymbol{\beta}_{\|}$ and $\boldsymbol{\beta}_{\perp}$ are the

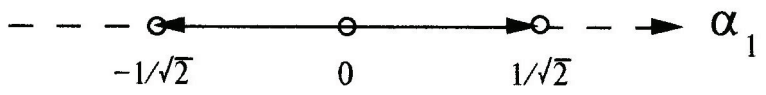


Figure 4.2: The root space diagram for all $l = 1$ algebras.

components of $\boldsymbol{\beta}$ parallel and perpendicular (respectively) to this plane. In this notation

$$\boldsymbol{\beta}' = \boldsymbol{\beta}_{\|} - \boldsymbol{\beta}_{\perp}, \tag{4.82}$$

i.e. $\boldsymbol{\beta}'$ is the image of $\boldsymbol{\beta}$ reflected in the $\boldsymbol{\alpha}$ hyperplane. If $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are roots then $\boldsymbol{\beta}'$ must also be a root. Thus every root defines a hyperplane, and the remaining roots must be symmetric with respect to a reflection in this plane. These are called Weyl reflections, and the set of all such reflections and their products makes up a discrete group called the **Weyl group**. We can restate (4.78) as follows: all root diagrams must be invariant under the Weyl group.

So how many ways are there to arrange a finite number of points in space so that these two conditions are satisfied? The answer to this question obviously depends on the dimensionality of the space. We can easily construct all possible root vector diagrams for $l = 1, 2$ using the rules presented above. This procedure was generalized to arbitrary dimension by Van der Waerden using a constructive argument that we will present eventually. Later Dynkin devised an elegant diagramatical proof that all spaces had been found. We will work through the Weyl-Cartan-Van der Waerden construction here. Dynkin diagrams will be discussed in the following section.

The case $l = 1$ case is particularly simple. All the roots must lie on a single line, so $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are redundant, and the only allowed Cartan angle is $\theta = 0°$. We know from Example 4.6 that the length of the root vectors $\sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})} = 1/\sqrt{2}$. The root space diagram is shown in Fig. 4.2. It is perhaps stretching the point to talk about a "hyperplane" in a one-dimensional space, but $\alpha = \pm 1/\sqrt{2}$ are mirror images of one another in a plane perpendicular to the axis and passing through the origin.

The first non-trivial root space diagrams appear for the case $l = 2$. We construct them by starting with the $l = 1$ diagram and adding one additional vector at one of the "official" Cartan angles. The remaining roots are found by taking all possible Weyl reflections. (One might expect this procedure to produce an infinite number of roots. The conditions in Table 4.5 guarantee

that this will not happen.) There are five distinct possibilities, which Cartan labeled G_2, B_2, C_2, D_2, and A_2.

$\mathbf{G_2}$. $\theta = 30°$

Suppose $\boldsymbol{\alpha}$ is a root vector with components $(1, 0)$. (We will return to the question of the normalization of the roots later.) There will be another root vector $\boldsymbol{\beta}$ at an angle of 30° to $\boldsymbol{\alpha}$. Choosing $n = 3$ and $n' = 1$ fixes the length of $\boldsymbol{\beta}$ at $\sqrt{3}$. Of course $-\boldsymbol{\alpha}$ and $-\boldsymbol{\beta}$ are also roots as are all possible Weyl reflections with respect to all other roots. This produces the star figure shown in Fig. 4.3 with the short and long vectors alternating at 30° intervals. The alternate choice, $n = 1$ and $n' = 3$, produces the same star rotated by 30°. A rotation of the root space diagram corresponds to a unitary transformation of the elements of the algebra and thus does not constitute a distinct algebra.

$\mathbf{B_2}$ and **$\mathbf{C_2}$.** $\theta = 45°$

The choice $n = 2$ and $n' = 1$ produces the square shown in Fig. 4.4. Cartan called this B_2. Alternatively $n = 1$, $n' = 2$ yields the rotated square C_2, Fig. 4.5. These two figures are also related by a similarity transformation, but this is not true in higher dimension, thus the separate names.

$\mathbf{A_2}$. $\theta = 60°$

The root space diagram is a regular hexagon. This is the algebra of the group $SU(3)$ (see Figures 4.1 and 4.6.) called A_2.

$\mathbf{D_2}$. $\theta = 90°$

When the second root vector is added at an angle of 90°, its length is indeterminate, and there are no commutation relations connecting the two vectors. Thus D_2 can be decomposed into a direct sum of two one-dimensional algebras. This only happens when $l = 2$, however.

Van der Waerden's procedure for generalizing these diagrams uses the following device: We introduce a set of unit vectors $\boldsymbol{e}_1 = (1, 0)$ and $\boldsymbol{e}_2 = (0, 1)$ with obvious generalization to higher dimension. Evidentally the roots of D_2, B_2, and C_2 can be written as linear combinations of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ with coefficients ± 1. These can be immediately generalized to higher dimension as follows:

$$\begin{array}{lll} D_l\colon & (\pm\boldsymbol{e}_i \pm \boldsymbol{e}_j); \pm 0\boldsymbol{e}_i & 1 \le i \ne j \le l \\ B_l\colon & (\pm\boldsymbol{e}_i \pm \boldsymbol{e}_j); \pm 1\boldsymbol{e}_i & \\ C_l\colon & (\pm\boldsymbol{e}_i \pm \boldsymbol{e}_j); \pm 2\boldsymbol{e}_i & \end{array} \tag{4.83}$$

These three root systems satisfy the two principal conditions and are invariant under the Weyl reflection group.

The remaining two algebras, A_2 and G_2, don't seem to fit this pattern, but they can be made to fit by embedding them in a three-dimensional space.

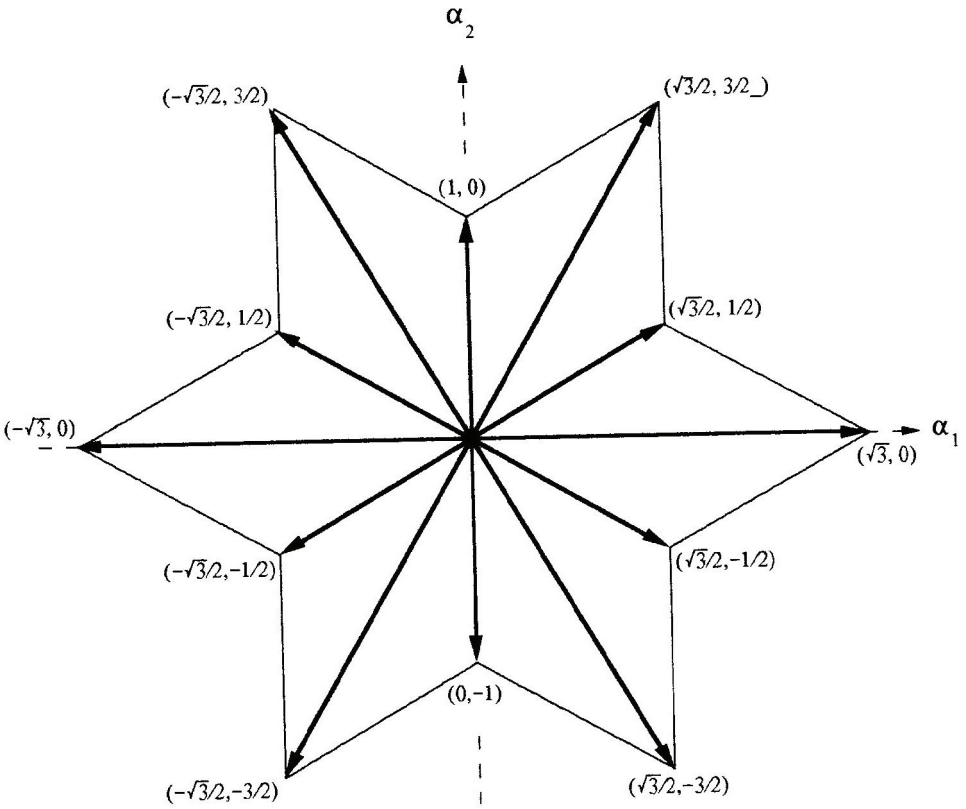


Figure 4.3: The root space diagram G_2.

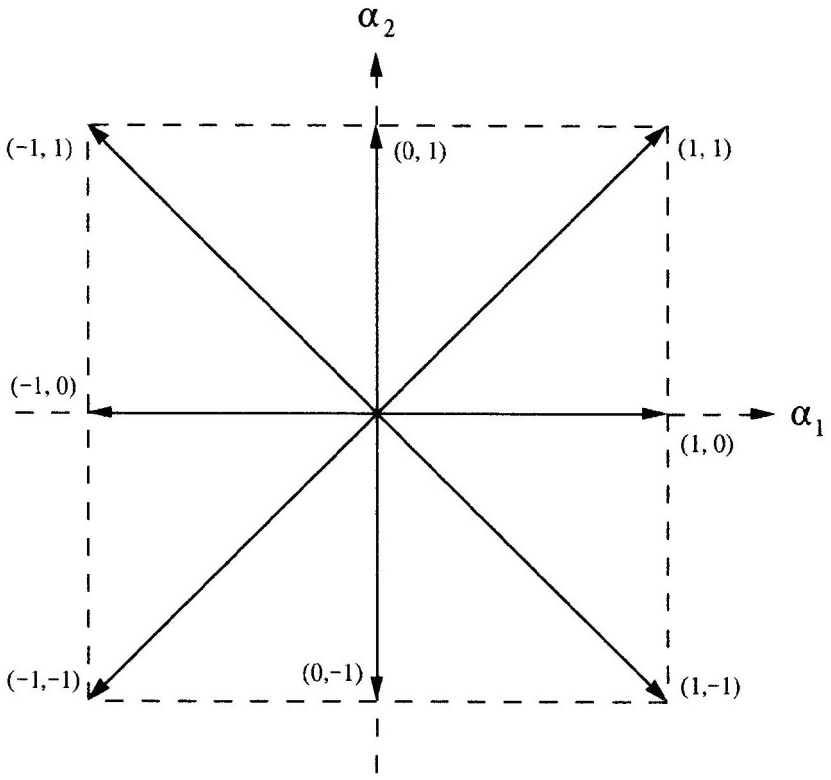


Figure 4.4: The root space diagram for B_2.

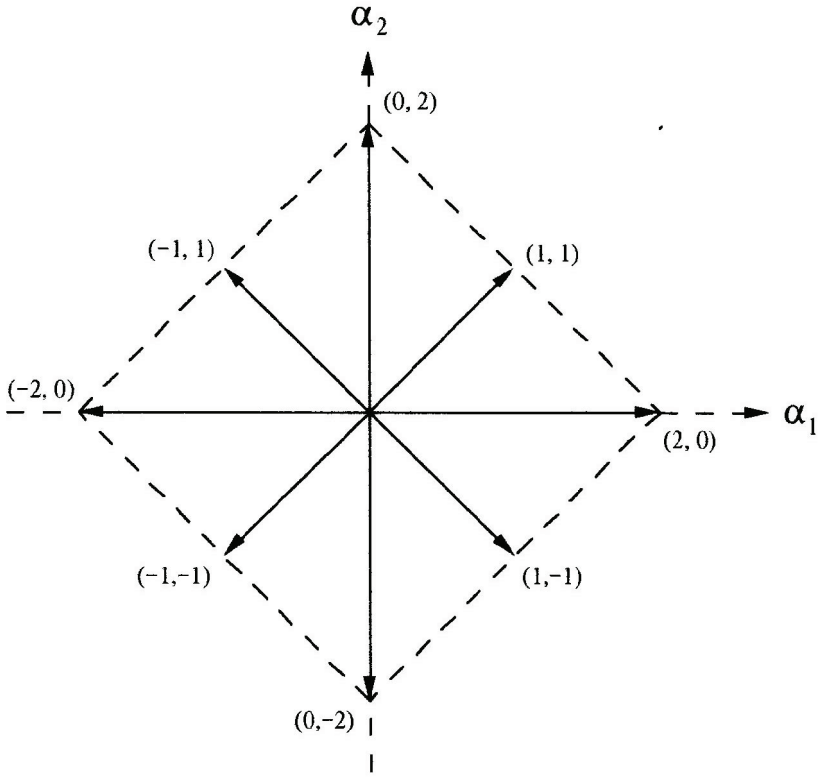


Figure 4.5: The root space diagram for C_2.

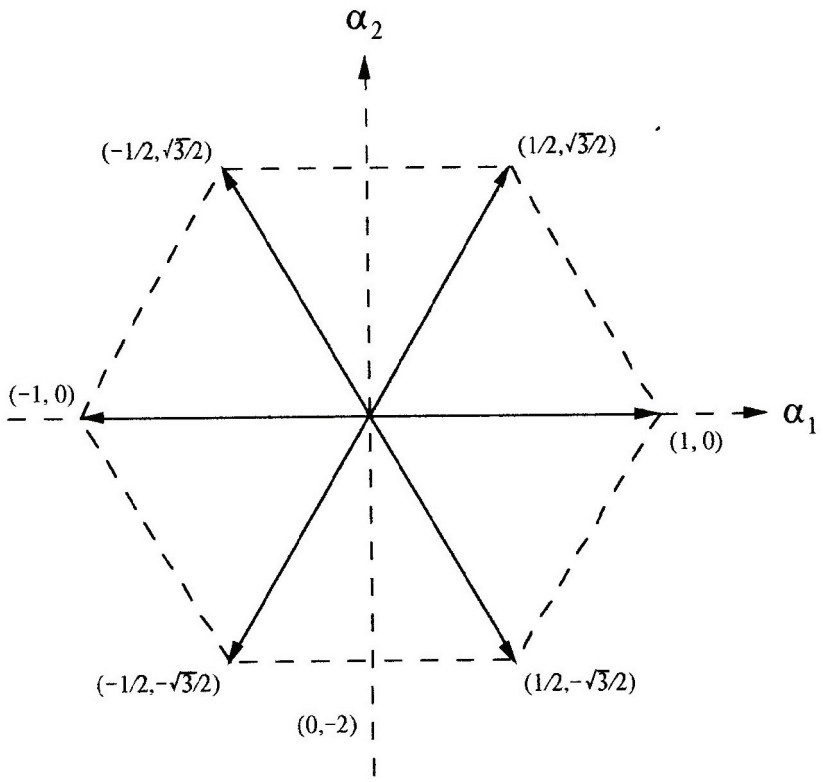


Figure 4.6: The root space diagram for A_2.

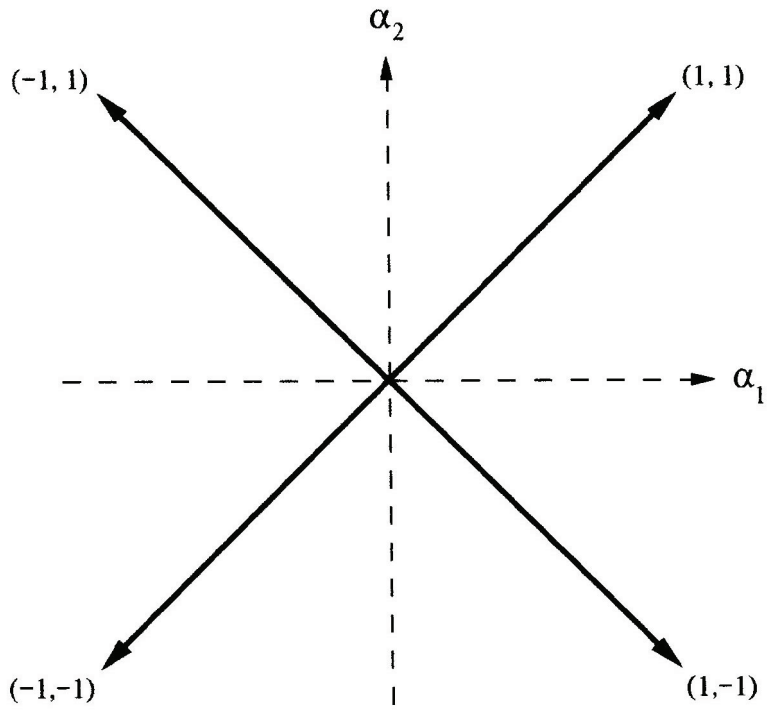


Figure 4.7: The root space diagram for D_2.

For example, the vectors of the form $(\boldsymbol{e}_i - \boldsymbol{e}_j); 1 \le i \ne j \le 3$ all lie in a plane perpendicular to the vector $\boldsymbol{R} = \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3$, and in this plane they point to the vertices of a regular hexagon. Thus we can construct A_l by taking all vectors of the form $(\boldsymbol{e}_i - \boldsymbol{e}_j); 1 \le i \ne j \le l+1$ and projecting them on the plane perpendicular to $\boldsymbol{R} = \sum \boldsymbol{e}_i$. The algebra G_2 is obtained by taking all vectors of the form $(\boldsymbol{e}_i - \boldsymbol{e}_j)$ and $\pm(\boldsymbol{e}_i + \boldsymbol{e}_j - 2\boldsymbol{e}_k)$ for $1 \le i \ne j \ne k \le 3$ and projecting them on the plane perpendicular to $\boldsymbol{R}$.

This procedure for constructing G_2 cannot be generalized to higher dimension. Thus G_2 is one of the "exceptional algebras" that cannot exist in all dimensions. There are four others called F_4, E_6, E_7, and E_8. They are constructed as follows:

$\boldsymbol{F_4}$. Starting with B_4 add the 16 root vectors

$$\frac{1}{2}(\pm\boldsymbol{e}_1 \pm \boldsymbol{e}_2 \pm \boldsymbol{e}_3 \pm \boldsymbol{e}_4)$$

$\boldsymbol{E_6}$. Starting with A_5 add the root vectors $\pm\sqrt{2}\boldsymbol{e}_7$ and some vectors of the form

$$\frac{1}{2}(\pm\boldsymbol{e}_1 \pm \boldsymbol{e}_2 \pm \boldsymbol{e}_3 \pm \boldsymbol{e}_4 \pm \boldsymbol{e}_5 \pm \boldsymbol{e}_6) \pm \boldsymbol{e}_7/\sqrt{2}.$$

In this case the $\pm$ signs are not all arbitrary. The first term must contain three positive and three negative signs.

$\boldsymbol{E_7}$. Starting with A_7 add roots of the form

$$\frac{1}{2}(\pm\boldsymbol{e}_1 \pm \boldsymbol{e}_2 \pm \boldsymbol{e}_3 \pm \boldsymbol{e}_4 \pm \boldsymbol{e}_5 \pm \boldsymbol{e}_6 \pm \boldsymbol{e}_7 \pm \boldsymbol{e}_8)$$

In this case four signs must be positive and four must be negative.

$\boldsymbol{E_8}$. Add to the root vectors of D_8 the vectors

$$\frac{1}{2}(\pm\boldsymbol{e}_1 \pm \boldsymbol{e}_2 \pm \boldsymbol{e}_3 \pm \boldsymbol{e}_4 \pm \boldsymbol{e}_5 \pm \boldsymbol{e}_6 \pm \boldsymbol{e}_7 \pm \boldsymbol{e}_8),$$

this time with an even number of positive signs.

In summary, there are four infinite chains of simple complex Lie algebras, A_l, B_l,C_l, and D_l, and five exceptional algebras G_2, F_4, E_6, E_7, and E_8. All semi-simple complex algebras can be decomposed into a direct sum of these simple algebras.

It is difficult to go much further with this constructive procedure. One can verify "by hand" that the various root systems described in this section satisfy (4.77) and (4.78) and are invariant under the Weyl group (see Gilman), but it is hard to prove convincingly that there are no other possibilities. It is also not clear how to construct matrix representations of these

algebras. We will deal with these issues in the next two sections. We first introduce the simple roots, which constitute the absolute minimum information necessary to completely specify a simple complex Lie algebra. The simple roots also yield a procedure for constructing a set of canonical matrix representations. We then present the elegant diagramatic proof of Dynkin that all possible simple complex algebras have been found. Dynkin's formalism also provides a procedure for decomposing an arbitrary semisimple algebra into a sum of simple algebras.

4.5.1 Simple roots

In the previous section we introduced a set of orthogonal unit vectors $\boldsymbol{e}_i$ in the l-dimensional root space. Every root $\boldsymbol{\alpha}$ can be expressed as a linear combination of these basis vectors.

$$\boldsymbol{\alpha} = \sum_{i=1}^{l} \beta^i \boldsymbol{e}_i$$

Definition 4.22 *Positive roots*
A non-zero root $\boldsymbol{\alpha}$ is said to be **positive** *if the first non-vanishing coefficient β^i is positive.*

The notation $\boldsymbol{\alpha} > 0$ will mean that $\boldsymbol{\alpha}$ is positive in the above sense.

Definition 4.23 *Lexicographic ordering of roots.*
If $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are non-zero roots and $(\boldsymbol{\alpha}-\boldsymbol{\beta}) > 0$, then we will write $\boldsymbol{\alpha} > \boldsymbol{\beta}$.

This is called "lexicographic" because it is similar to the way that words are arranged in a dictionary. A sequence of roots $\boldsymbol{\alpha}_1 < \boldsymbol{\alpha}_2 < \boldsymbol{\alpha}_3 < \cdots$ is arranged in order of increasing first component; if the first components are equal, then in order of increasing second component, *etc.* It should be clear that this ordering depends on the basis chosen. In the rest of this section we will assume that the basis $\boldsymbol{e}_i$ has been chosen and adhered to.

Definition 4.24 *Simple roots*
A non-zero root $\boldsymbol{\alpha}$ is said to be **simple** *if it is positive but cannot be expressed in the form $\boldsymbol{\alpha} = \boldsymbol{\beta} + \boldsymbol{\gamma}$ where $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are both positive roots.*

The simple roots can be used as a new set of basis vectors. This is a consequence of the following theorem.

Theorem 4.26 *A semisimple algebra of rank l has exactly l simple roots $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_l$. Moreover, if $\boldsymbol{\alpha}$ is any positive root*

$$\boldsymbol{\alpha} = \sum_{i=1}^{l} \kappa^i \boldsymbol{\alpha}_i \tag{4.84}$$

where $\kappa^1, \kappa^2, \ldots, \kappa^l$ are non-negative integers.

The proof of this theorem makes use of the following lemmas:

Lemma 4.27 *If $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are two simple roots, and $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$, then $\boldsymbol{\alpha} - \boldsymbol{\beta}$ is not a root, and $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leq 0$.*

Proof: If $\boldsymbol{\alpha} - \boldsymbol{\beta}$ were a positive root, we could write $\boldsymbol{\alpha} = (\boldsymbol{\alpha} - \boldsymbol{\beta}) + \boldsymbol{\beta}$, and $\boldsymbol{\alpha}$ would not be simple. On the other hand, if $\boldsymbol{\beta} - \boldsymbol{\alpha}$ were a positive root, we could write $\boldsymbol{\alpha} = (\boldsymbol{\beta} - \boldsymbol{\alpha}) + \boldsymbol{\alpha}$, and again $\boldsymbol{\alpha}$ would not be simple. Thus $\boldsymbol{\alpha} - \boldsymbol{\beta}$ cannot be a root.

The second assertion is a by-product of Theorem 4.25. In the proof of that theorem we constructed the chain $\mathcal{V}_{\beta+k\alpha}$. The integer k ranges from $-m$ to n where m is the largest integer such that $\boldsymbol{\beta} - m\boldsymbol{\alpha}$ is a root, and n is the largest integer such that $\boldsymbol{\beta} + n\boldsymbol{\alpha}$ is also a root. If $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are simple roots, however, $m = 0$, and (4.77) becomes $2(\boldsymbol{\alpha}, \boldsymbol{\beta}) = -n(\boldsymbol{\alpha}, \boldsymbol{\alpha}) \leq 0$.

Lemma 4.28 *The simple roots are linearly independent.*

Proof: Suppose there were q simple roots, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_q$, not all independent. Then we could find a set of q *positive* coefficients $c_1, c_2, \ldots, c_q$ such that

$$\boldsymbol{\gamma} = c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \cdots + c_p\boldsymbol{\alpha}_p = c_{p+1}\boldsymbol{\alpha}_{p+1} + c_{p+2}\boldsymbol{\alpha}_{p+2} + \cdots + c_q\boldsymbol{\alpha}_q$$

The scalar product $(\boldsymbol{\gamma}, \boldsymbol{\gamma})$ is positive definite; however, it can be written as a sum of terms of the form $c_i c_j(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j)$ where $i \leq p$ and $j > p$. All the scalar products $(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j)$ combine pairs of different simple roots and hence are negative by the previous lemma. This contradiction establishes the proof.

Equation (4.84) can now be proved by induction. Let $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots$ be the set of all positive roots ordered so that $\boldsymbol{\beta}_j < \boldsymbol{\beta}_{j+1}$. Then $\boldsymbol{\beta}_1$ is the smallest positive root, so it cannot possibly be written as a sum of positive roots. It is therefore simple, and (4.84) is trivially true. Now suppose the theorem has been proved for the first p positive roots. If $\boldsymbol{\beta}_{p+1}$ is simple, (4.84) is again true. If it is not simple, it can be written as a sum of two smaller

positive roots both of which satisfy (4.84). This proves that all positive roots can be written as a sum of simple roots with positive integer coefficients. Since the root space is l-dimensional and since the simple roots are linearly independent, there must be l simple roots to span the space.

The significance of this result depends on the commutation relation, (assuming $\boldsymbol{\alpha} + \boldsymbol{\beta}$ is a root)

$$[E_\alpha, E_\beta] = N_{\alpha,\beta} E_{\alpha+\beta}.$$

Once we know the simple roots and the corresponding elements in the algebra, all other positive roots are obtained by adding simple roots, and the corresponding elements are found by taking the commutators. The negative roots and their elements come from the replacements $\boldsymbol{\alpha} \to -\boldsymbol{\alpha}$ and $E_\alpha \to E_{-\boldsymbol{\alpha}} = E_\alpha^\dagger$. The previous theorem guarantees that all non-zero roots and their elements can be found in this way. Another way of looking at this is that the set of all elements with positive roots constitutes an algebra that is solvable and nilpotent. We call this algebra $\mathcal{P} = \mathcal{P}^1$, and construct the usual sequence:

$$[\mathcal{P}^1, \mathcal{P}^1] \equiv \mathcal{P}^2$$

$$\cdots$$

$$[\mathcal{P}^1, \mathcal{P}^j] \equiv \mathcal{P}^{j+1}$$

The elements that "disappear" in going from $\mathcal{P}^j \to \mathcal{P}^{j+1}$ are contained in the factor algebra $\mathcal{P}^j$ mod $\mathcal{P}^{j+1}$. The corresponding roots are said to lie in the j-th level. The first level contains the simple roots. The second level contains roots that can be written as a sum of two simple roots. The j-th level contains those roots for which the sum of the κ^i in (4.84) is j. The roots in the $(j+1)$-th level are larger (in the lexicographic sense) than the roots in the j-th level.

Example 4.8 *The algebra G_2*

The most complicated of the rank-2 algebras is G_2. The positive roots are listed in Table 4.6 in lexicographic order. There are two roots that cannot be expressed as a sum of two other roots. These are the simple roots, which we call $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$. The other roots are simple linear combinations of $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ as promised.

Example 4.9 *The algebra A_l*

$\sqrt{3}\boldsymbol{e}_1$	$2\boldsymbol{\alpha}_1 + 3\boldsymbol{\alpha}_2$
$\sqrt{3}/2\boldsymbol{e}_1 + \frac{3}{2}\boldsymbol{e}_2$	$\boldsymbol{\alpha}_1 + 3\boldsymbol{\alpha}_2$
$\sqrt{3}/2\boldsymbol{e}_1 + \frac{1}{2}\boldsymbol{e}_2$	$\boldsymbol{\alpha}_1 + 2\boldsymbol{\alpha}_2$
$\sqrt{3}/2\boldsymbol{e}_1 - \frac{1}{2}\boldsymbol{e}_2$	$\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$
$\sqrt{3}/2\boldsymbol{e}_1 - \frac{3}{2}\boldsymbol{e}_2$	$\boldsymbol{\alpha}_1$
$\boldsymbol{e}_2$	$\boldsymbol{\alpha}_2$

Table 4.6: The first column contains the positive roots of G_2 in lexicographic order. The second column expresses them in terms of the simple roots, α_1 and α_2.

The positive roots of A_l have the form $\boldsymbol{e}_i - \boldsymbol{e}_j$, $1 \leq i < j \leq l+1$. The only such roots that cannot be expressed as the sum of two positive roots are the $\boldsymbol{e}_i - \boldsymbol{e}_{i+1}$, $1 \leq i \leq l$. It is easy to construct an $l \times l$ representation with these roots.

$$[H_i]_{km} = \delta_{ik}\delta_{im}, \quad 1 \leq i \leq l \tag{4.85}$$
$$H_{l+1} = -(H_1 + H_2 + \cdots + H_l)$$

$$H_i = \begin{array}{c} \\ i \end{array} \overset{\begin{array}{c} \\ i \end{array}}{\begin{bmatrix} & \vdots & \\ \cdots & 1 & \cdots \\ & \vdots & \end{bmatrix}}$$

$$\left[E_{e_i - e_{i+1}}\right]_{km} = \delta_{ik}\delta_{i,m-1}, \quad 1 \leq i \leq l \tag{4.86}$$

$$E_{e_i - e_{i+1}} = \begin{array}{c} \\ i \end{array} \overset{\begin{array}{c} \\ i+1 \end{array}}{\begin{bmatrix} & \vdots & \\ \cdots & 1 & \cdots \\ & \vdots & \end{bmatrix}}$$

It is easy to verify that these matrices satisfy the commutation relations

$$[H_i, E_{e_j - e_{j+1}}] = \epsilon_i(j, j+1) E_{e_j - e_{j+1}}$$

where the root vector, $\boldsymbol{\epsilon}(j,k) = \boldsymbol{e}_j - \boldsymbol{e}_k$. One can construct the remaining elements by taking commutators; for example, the elements corresponding to second level roots are found as follows:

$$[E_{e_j - e_{j+1}}, E_{e_{j+1} - e_{j+2}}] = E_{e_j - e_{j+2}}$$

In general,

$$[H_i, E_{e_j - e_k}] = \epsilon_i(j,k) E_{e_j - e_k}.$$

This representation makes the concept of the level of the roots quite concrete. The matrices corresponding to roots of level p have a 1 in the p-th diagonal above the main diagonal and zeros elsewhere.

There are different choices of basis elements that are more useful for exploring the connection between the algebra and the groups that give rise to it. For example, we can choose a different basis for $\mathcal{V}_\circ$.

$$h_i = H_i - H_{i+1} \quad 1 \le i \le l$$

where the H_i are defined in (4.85). All elements are now real and traceless. Exponentiating a real traceless algebra yields a real, special (*i.e.* unimodular) group. Thus A_l is the algebra of $SL(l+1, r)$. A different group results from the choice

$$h_i = \imath(H_i - H_{i+1}) \quad 1 \le i \le l$$
$$e_\alpha = E_\alpha - E_\alpha^T$$
$$e'_\alpha = \imath(E_\alpha + E_\alpha^T)$$

In this definition $\boldsymbol{\alpha}$ stands for any positive root, and the superscript ^{T} is the usual matrix transpose. The new basis is not only traceless but also complex and anti-Hermitian, *i.e.* $E^\dagger = -E$. It exponentiates into $SU(l+1)$. Evidentally, different groups can share the same algebra ("same" in the sense of having the same root structure), depending on the choice of basis and the restrictions placed on the coefficient of the algebra. This theme will be explored more fully in the next chapter.

Example 4.10 *The algebra D_l*

The positive roots of D_l have the form $\boldsymbol{e}_i \pm \boldsymbol{e}_j$, $1 \le i < j \le l$. There are $l-1$ simple roots $\boldsymbol{e}_i - \boldsymbol{e}_{i+1}$, $1 \le i \le l-1$, and one simple root, $\boldsymbol{e}_{l-1} + \boldsymbol{e}_l$. There is a matrix representations of this algebra consisting of $2l \times 2l$ matrices with a peculiar symmetry: define $\tilde{A}$ as the transpose of A with respect to the minor (*i.e.* lower left to upper right) diagonal. The required symmetry

is then $\tilde{A} = -A$. The Cartan subalgebra can be constructed using the H_i (with $1 \leq i \leq l$) from (4.85).

$$\boldsymbol{H}_i = \left[\begin{array}{c|c} H_i & 0 \\ \hline 0 & -\tilde{H}_i \end{array}\right] \tag{4.87}$$

The matrices corresponding to the positive roots are partitioned as follows:

$$\boldsymbol{E}_i = \left[\begin{array}{c|c} A & B \\ \hline 0 & -\tilde{A} \end{array}\right] \tag{4.88}$$

The first $l-1$ simple roots, $\boldsymbol{e}_i - \boldsymbol{e}_{i+1}$, correspond to matrices with $B = 0$ and $A = E_{e_i - e_{i+1}}$ from (4.86). The remaining simple root, $\boldsymbol{e}_{l-1} + \boldsymbol{e}_l$, requires $A = 0$ and

$$B = \left[\begin{array}{c|c} 0 & 0 \\ \hline C & 0 \end{array}\right], \tag{4.89}$$

where C is the 2×2 matrix

$$C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{4.90}$$

The $2l \times 2l$ dimensionality is required to accomodate the "extra" root, $\boldsymbol{e}_{l-1} + \boldsymbol{e}_l$, and the symmetry prevents the appearance of other unwanted roots. None of the classical groups discussed in Section 2.1 possess this symmetry in any obvious way; however, the metric tensor for the group $SO(l, l)$,

$$g = \begin{bmatrix} I_l & 0 \\ 0 & -I_l \end{bmatrix}, \tag{4.91}$$

does satisfy $\tilde{g} = -g$ suggesting that D_l might be "burried" in $SO(l, l)$. To prove that this is so, choose the transformation matrix S in (2.18) so that the metric becomes

$$g_{ij} = \delta_{l+1-i,j} = \delta_{i,l+1-j}, \tag{4.92}$$

i.e. g has $+1$'s on the minor diagonal and zeros elsewhere. Let M in (2.14) be an infinitesimal transformation of the form $M = I + X$. Then (2.14) becomes

$$(I + X)g(I + X)^T = g.$$

Keeping only terms of first order in X gives

$$Xg + gX^T = 0$$

$$X_i{}^k g_{kj} + g_{ik}(X^T)^k{}_j = 0 \tag{4.93}$$

$$X_{i,l+1-j} = -X_{j,l+1-i}, \tag{4.94}$$

which is the required symmetry. This symmetry is not preserved by similarity transformations in general; however, the group multiplication table is, so in this sense D_l is the algebra of $SO(l, l)$ even though (4.92) is not the usual metric for this group.

Example 4.11 *The algebra B_l*

The positive roots of B_l have the form $\boldsymbol{e}_i \pm \boldsymbol{e}_j$ and $\boldsymbol{e}_i$ where $1 \le i < j \le l$. The simple roots are identical with those of D_l exceps that the l-th simple roots is $\boldsymbol{e}_l$ rather than $\boldsymbol{e}_{l-1} + \boldsymbol{e}_l$. The corresponding matrices have dimension $(2l + 1)$.

$$\boldsymbol{H}_i = \left[\begin{array}{c|c} H_i & 0 \\ \hline 0 & -\tilde{H}_i \end{array}\right] \tag{4.95}$$

$$\boldsymbol{E}_{e_i - e_{i+1}} = \left[\begin{array}{c|c} E_{e_i - e_{i+1}} & 0 \\ \hline 0 & -\tilde{E}_{e_i - e_{i+1}} \end{array}\right] \tag{4.96}$$

$$\boldsymbol{E}_l = \left[\begin{array}{c|c|c} 0 & 1 & 0 \\ \hline & & -1 \\ \hline 0 & & 0 \end{array}\right] \tag{4.97}$$

This is the algebra of $SO(l+1,l)$.

Example 4.12 *The algebra C_l*

The positive roots of C_l have the form $\boldsymbol{e}_i \pm \boldsymbol{e}_j$ and $2\boldsymbol{e}_i$ where $1 \leq i < j \leq l$. The simple roots are the same as B_l except $2\boldsymbol{e}_l$ replaces $\boldsymbol{e}_l$. The corresponding matrices are $2l \times 2l$. The Cartan subalgebra is given by (4.87), and the $l-1$ simple roots by (4.88) with $B=0$. The remaining simple root $2\boldsymbol{e}_l$ requires $A=0$ in (4.88) and $B=0$ except for a 1 in the lower left corner. The other positive roots preserve the form of (4.88) but with the symmetry $\tilde{B} = +B$. This is the symmetry of the symplectic group $Sp(2l)$ as can be seen as follows: it is possible to choose an S in (2.18) that transforms the symplectic metric (???) into the form

$$g = \begin{bmatrix} 0 & \tilde{I}_l \\ -\tilde{I}_l & 0 \end{bmatrix}, \tag{4.98}$$

where $\tilde{I}_l$ is a $l \times l$ matrix with $+1$'s on the minor diagonal and zeros elsewhere. Substituting this metric into (4.93) yields (4.93) if i and j are both in the first or third "quadrants," *i.e.* if i and j are both $\leq l$ or $\geq l+1$. If $i \leq l$ and $j \geq l+1$ or vice versa then (4.93) becomes

$$X_{i,l+1-j} = X_{j,l+1-i},$$

which is the required symmetry.

4.5.2 Dynkin diagrams

Let $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_l$ be the simple roots of a simple or semisimple algebra. It will be convenient to write them as unit vectors.

$$\boldsymbol{u}_i = \frac{\boldsymbol{\alpha}_i}{\sqrt{(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_i)}} \tag{4.99}$$

The cosine of the angle between any two simple roots is given by (4.80) and Lemma 4.27.

$$(\boldsymbol{u}_i, \boldsymbol{u}_j) = -\sqrt{n/4} \qquad n = 0, 1, 2, 3 \tag{4.100}$$

To each root we assign a *weight* w_i as follows: the shortest root in the space is assigned a weight $w = 1$. If $\boldsymbol{\alpha}$ is the shortest root, then the weight of any other root $\boldsymbol{\alpha}_i$ is given by

$$w_i = \frac{(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_i)}{(\boldsymbol{\alpha}, \boldsymbol{\alpha})}. \tag{4.101}$$

Normally the weights will be 1, 2, or 3 as shown in Table 4.5. (The case $w = 0/0$ will not be a problem as we shall see presently.)

The Dynkin diagrams are drawn as follows:

1. Draw a dot for each simple root. Label the dot with the name of the root ($\boldsymbol{u}_1, \boldsymbol{u}_2$, *etc.*) and the coresponding weight.

2. Connect each pair of dots with n lines where n is given by (4.100).

Buried in the proof of Theorem **??** are certain restrictions on the topologies of these diagrams. We display these restrictions in the form of a few simple theorems that will allow us to enumerate all possible "legal" diagrams. (The treatment here closely follows that of Gilmore.)

Example 4.13 *The $l = 2$ algebras.*

The Dynkin diagrams for all the $l = 2$ algebras discussed in the previous section are given in Fig. 4.8.

The algebras B_2 and C_2 differ only by the exchange $w_1 \leftrightarrow w_2$, which as we saw, corresponds to a 45° rotation of the root diagram. The diagram for D_2 consists of two disconnected dots reflecting the fact that D_2 is a direct sum of two $l = 1$ algebras. This is an example of the following general theorem:

Theorem 4.29 *If a diagram consists of several disconnected components, then each component describes a simple algebra, and the entire algebra is a direct sum of the component algebras.*

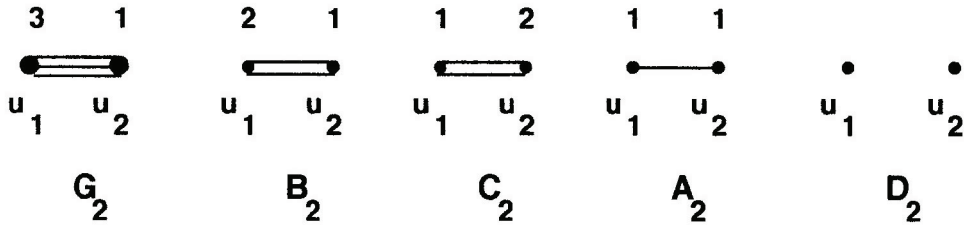


Figure 4.8: Dynkin diagrams for the $l = 2$ algebras.

Proof: Suppose the Dynkin diagram consists of two disconnected components, $\mathcal{A}$ and $\mathcal{B}$. Let $\boldsymbol{\alpha}$ be any simple root in $\mathcal{A}$ and $\boldsymbol{\beta}$ be any simple root in $\mathcal{B}$. Then $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0$. It is easy to show that $\boldsymbol{\alpha}+\boldsymbol{\beta}$ cannot be a root. After all, $\boldsymbol{\alpha} + \boldsymbol{\beta}$ is a member of the $\boldsymbol{\beta}$ chain containing $\boldsymbol{\alpha}$, $\boldsymbol{\beta} + k\boldsymbol{\alpha}$, with $k = 1$. According to the proof of Theorem **??**, $-m \leq k \leq n$, but equation (4.77) reduces to $n = m$. Lemma 4.27 assures us that $\boldsymbol{\alpha} - \boldsymbol{\beta}$ cannot be a root, consequently $n = m = 0$. As a consequence, $[E_\alpha, E_\beta] = 0$. Now suppose that $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$ where $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \mathcal{A}$ are simple roots. Then

$$[E_\alpha, E_\beta] \propto [[E_{\alpha_1}, E_{\alpha_2}], E_\beta] \propto [[E_\beta, E_{\alpha_1}], E_{\alpha_2}] + [[E_{\alpha_2}, E_\beta], E_{\alpha_1}] = 0.$$

Proceeding in this way we can show that $[E_\alpha, E_\beta] = 0$ for any $\boldsymbol{\alpha}$ (not necessarily simple) in $\mathcal{A}$ and any $\boldsymbol{\beta}$ in $\mathcal{B}$. Since all elements of $\mathcal{A}$ commute with all elements of $\mathcal{B}$, we are entitled to combine them as a direct sum according to Definition **??**.

Incidentally, when diagrams are disconnected the weights must be assigned relative to the shortest root in each diagram. The ratio of the lengths of two roots in disconnected diagrams is always 0/0.

Theorem 4.30 *There are no loops;* i.e. *if the line(s) connecting any pair of dots is(are) severed, the diagram falls into two pieces.*

Proof: Since the scalar product is positive definite,

$$\left(\sum_{i=1}^{n} \boldsymbol{u}_i, \sum_{j=1}^{n} \boldsymbol{u}_j \right) = n + 2 \sum_{i>j} (\boldsymbol{u}_i, \boldsymbol{u}_j) > 0. \tag{4.102}$$

If $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ are connected, then $2(\boldsymbol{u}_i, \boldsymbol{u}_j) \leq -1$, so the number of connected pairs must be less than n. This rules out all loops. Equation (4.102) also limits a single diagram to no more than two double lines or one triple line.

Theorem 4.31 *The total number of lines connected to any point is at most three.*

Proof: Let the roots $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots$ be connected to the root $\boldsymbol{u}$. The $\boldsymbol{v}_i$ cannot be connected to one another (that would make loops), so they make up a set of orthonormal unit vectors. Consequently

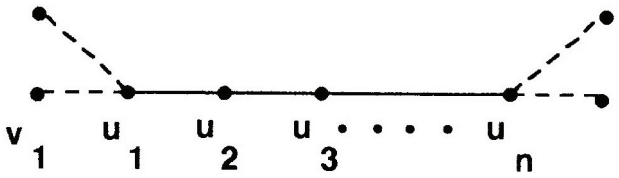
$$\sum_{i=1}(\boldsymbol{u}, \boldsymbol{v}_i)^2 = \sum_{i=1} \frac{n_i}{4} \leq 1 \tag{4.103}$$

The equality can be ruled out, however; because if $\sum_{i=1}(\boldsymbol{u}, \boldsymbol{v}_i)^2 = 1$, then the simple roots $\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots$ would not be linearly independent, and this contradicts Lemma 4.28.

Evidentally, there can only be on connected Dynkin diagram with a triple line: the exceptional algebra G_2, (see Fig. 4.8).

Theorem 4.32 *Let $\boldsymbol{u}_i$ be a set of dots connected by single lines. The diagram obtained by shrinking the entire chain to a single dot $\boldsymbol{u}$ is a valid diagram if and only if the original diagram was.*

Proof: Consider the following diagram.



The dashed lines extending from $\boldsymbol{u}_1$ and $\boldsymbol{u}_n$ represent arbitrary connections with 0, 1, or 2 lines. Let $\boldsymbol{u} = \sum_{i=1}^{n} \boldsymbol{u}_i$. Now:

$$2(\boldsymbol{u}_i, \boldsymbol{u}_j) = 0 \quad j > i+1 \text{ since there are no loops}$$

$$2(\boldsymbol{u}_i, \boldsymbol{u}_{i+1}) = -1 \text{ from (4.100)}$$

Consequently (4.102) yields $(\boldsymbol{u}, \boldsymbol{u}) = 1$, so $\boldsymbol{u}$ is a unit vector like the $\boldsymbol{u}_i$. Finally

$$(\boldsymbol{v}_1, \boldsymbol{u}) = (\boldsymbol{v}_1, \boldsymbol{u}_1),$$

so roots like $\boldsymbol{v}_1$ connect to $\boldsymbol{u}$ in exactly the same way they would connect to the end of the chain $\boldsymbol{u}_1$.

We can now enumerate the possible diagrams.

1. It is possible to have one pair of roots connected with a triple line. This is the special algebra G_2 shown in Fig. 4.8. We cannot connect it to

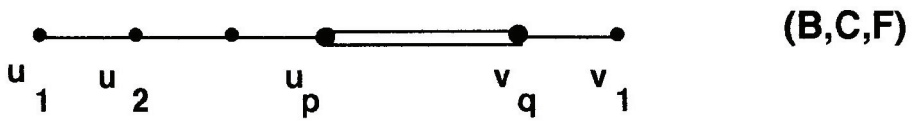


Figure 4.9: The most general diagram with one double line.

any other roots, however, for this would violate the "no more than three lines" rule.

2. It is possible to have one pair of roots connected with a double line with additional connections at one or both ends. The general diagram is shown in Fig. 4.9. There must be no other double lines or "forks," however, because the resulting diagram could be shrunk to one in which four lines met at one point.

The integers p and q are not completely arbitrary as can be seen from the following argument: Define

$$\boldsymbol{u} = \sum_{i=1}^{p} i\boldsymbol{u}_i \qquad \boldsymbol{v} = \sum_{j=1}^{q} j\boldsymbol{v}_j.$$

Then

$$(\boldsymbol{u}, \boldsymbol{u}) = \sum_{i=1}^{p} i^2 - \sum_{i=1}^{p-1} i(i+1) = p(p+1)/2,$$

and likewise,

$$(\boldsymbol{v}, \boldsymbol{v}) = q(q+1)/2.$$

Finally,

$$(\boldsymbol{u}, \boldsymbol{v}) = pq(\boldsymbol{u}_p, \boldsymbol{v}_q) = -pq/\sqrt{2}.$$

The Schwartz inequality now gives us a condition on p and q.

$$(\boldsymbol{u}, \boldsymbol{v})^2 < (\boldsymbol{u}, \boldsymbol{u})(\boldsymbol{v}, \boldsymbol{v})$$

$$2 < \left(1 + \frac{1}{p}\right)\left(1 + \frac{1}{q}\right) \tag{4.104}$$

We can satisfy this equation by setting $q = 1$ and $p \geq 1$. This gives the two root spaces B_n and C_n. (Fig. 4.10)

The solution $p = q = 2$ gives two diagrams for the same algebra F_4. (Fig. 4.11)

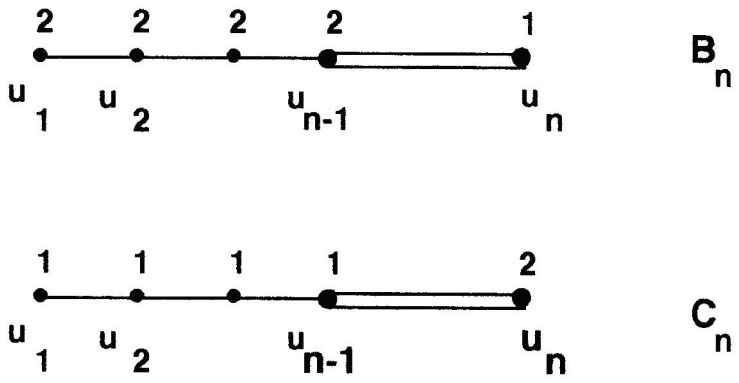


Figure 4.10: The root spaces B_n and C_n.

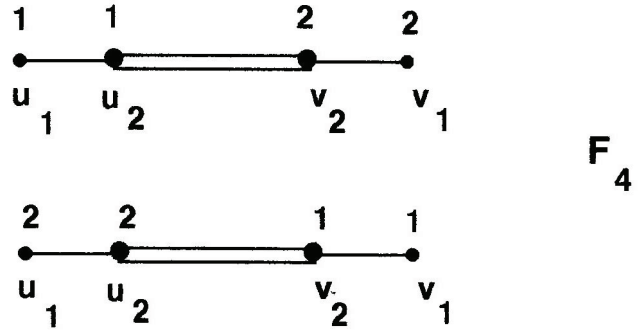


Figure 4.11: The root space diagram for F_4.

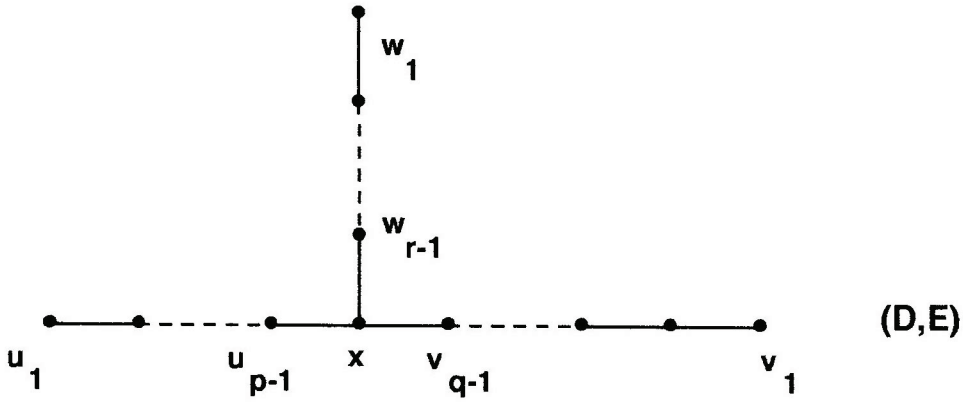


Figure 4.12: The most general diagram with one fork.

3. If there are no double or triple lines in a diagram we are entitled to have one fork. The most general diagram of this sort is shown in Fig. 4.12.

Again the integers p, q, and r are not completely arbitrary. The argument is similar to the previous case. Let

$$\boldsymbol{u} = \sum_{i=1}^{p-1} i\boldsymbol{u}_i \qquad \boldsymbol{v} = \sum_{j=1}^{q-1} j\boldsymbol{v}_j \qquad \boldsymbol{w} = \sum_{k=1}^{r-1} k\boldsymbol{w}_k$$

$$(\boldsymbol{u},\boldsymbol{u}) = (p-1)p/2 \qquad (\boldsymbol{v},\boldsymbol{v}) = (q-1)q/2 \qquad (\boldsymbol{w},\boldsymbol{w}) = (r-1)r/2$$

$$(\boldsymbol{u},\boldsymbol{v}) = (\boldsymbol{u},\boldsymbol{w}) = (\boldsymbol{v},\boldsymbol{w}) = 0$$

We assume $(\boldsymbol{x},\boldsymbol{x}) = 1$. The sum of the squares of the direction cosines of $\boldsymbol{x}$ along the three orthogonal axes, $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$, must be less than 1.

$$\frac{(\boldsymbol{x},\boldsymbol{u})^2}{(\boldsymbol{u},\boldsymbol{u})} = \frac{(p-1)^2/4}{p(p-1)/2} = \frac{1}{2}\left(1-\frac{1}{p}\right)$$

$$\frac{1}{2}\left(1-\frac{1}{p}+1-\frac{1}{q}+1-\frac{1}{r}\right) < 1$$

$$\frac{1}{p}+\frac{1}{q}+\frac{1}{r} > 1 \tag{4.105}$$

If we asume that $p \le q \le r$ (to avoid double counting) there are four solutions to (4.105) given by Table 4.7

The resulting diagrams are shown in Fig. 4.13.

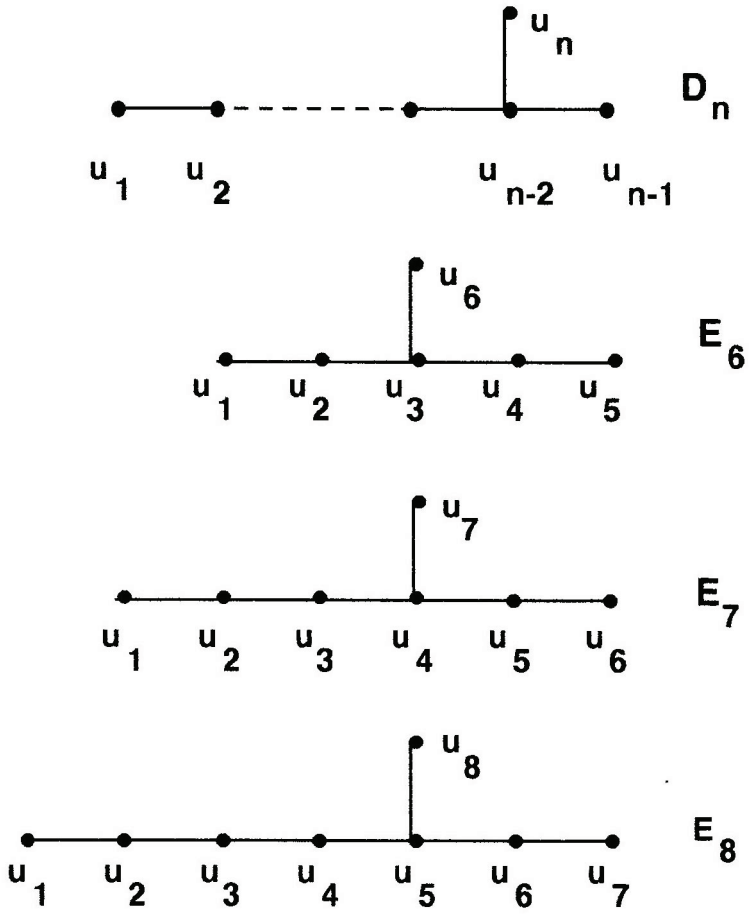


Figure 4.13: Dynkin diagrams for D_n and the exceptional algebras E_6, E_7, and E_8.

Figure 4.14: The Dynkin diagram for all algebras of the class A_n.

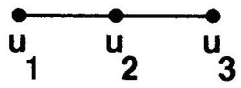
p	q	r	Algebra
$p \geq 1$	2	2	D_n
3	2	2	E_6
4	3	2	E_7
5	3	2	E_8

Table 4.7: The integer solutions to (4.105) and their corresponding algebras.

4. Finally it is possible to have diagrams without any double or triple lines and without any forks. This corresponds to the algebra A_n. The general diagram is shown in Fig. 4.14.

In summary, we have shown that the four series of Lie algebras A_n, B_n, C_n, and D_n, together with the exceptional algebras G_2, F_4, E_6, E_7, and E_8, account for all possible simple complex Lie algebras. Futhermore, a semisimple algebra can be decomposed into simple algebras by drawing its Dynkin diagram.

This completes our search for the "essence" of a Lie algebra. At least for semisimple algebras it consists of a pattern of simple root vectors. To return to the example with which we began this (exceptionally long) chapter, the properties of the algebra $SU(4)$ with its 1575 structure constants are completely specified by the diagram for A_3!



Chapter 5

Real Algebras and Their Groups

Groups are usually parameterized with real parameters. This leads to real Lie algebras, *i.e.* algebras defined on the field of real numbers with real structure constants. In order to cary out the analysis in Chapter 4, however, it was necessary to extend the algebra to the complex number field. The reverse procedure, reconstructing the real algebra from a complex one, turns out to be ambiguous. The central problem is this: matrix representations of Lie algebras, even real algebras, usually entail complex matrices. For this reason, a complex group element like e^{zX} could be made real by setting the imaginary part of z to zero, setting the real part to zero and absorbing the i into X, or by some combination of these two procedures. The Weyl canonical form, which was the final result of the analysis of Chapter 4, has real structure constants. One way to pose the question then is, how many other real algebras can we get from the Weyl form by suitably restricting the number field and what are the properties of the corresponding groups? From a practical point of view, it turns out to be more convenient to start from another real algebra called the "compact real form" and construct the other real algebras from it. The Weyl form and the compact real form stand at opposite ends of the spectrum of compactness. We will eventually show how to construct all the intermediate algebras and develop a quantity called the "character" to measure where an algebra stands on this scale.

5.1 Compact groups and compact algebras

Compact matrix groups are bounded in two related ways: (1) The individual matrix elements cannot exceed some maximum value. (2) The volume of parameter space is finite. This invariably comes about because of some periodic or recursive relationship between the two as shown in the next example.

Example 5.1 *The unitary matrix groups*

The unitary groups $U(n)$ and $O(n)$ preserve the metric $g_{ij} = \delta_{ij}$. In this case (2.14) becomes

$$\delta_{ij} = \sum_l M_i^{(*)l} M_j{}^l,$$

so

$$|M_i{}^l|^2 \leq 1.$$

The matrix elements are all bounded so the groups are compact. Now let $M_i{}^l = \delta_i^l + tX_i{}^l$ where t is infinitesimal. To first order in t:

$$\delta_{ij} = (\delta_i^k + tX_i^{(*)\,k})\delta_{kl}(\delta_j{}^l + tX_j{}^l)$$

$$0 = X_i^{(*)\;j} + X_j{}^i$$

As a consequence, the Lie algebras $u(n)$ and $o(n)$ consist of anti-Hermitian matrices. Conversely, group elements obtained by exponentiating an anti-Hermitian algebra are unitary, since if

$$U(X) = e^X$$

with X anti-Hermitian, then

$$U^\dagger(X) = e^{X^\dagger} = e^{-X} = U^{-1}(X).$$

We can construct a convenient basis for these groups in the defining representation. Let S, A, and D be $n \times n$ matrices.

$$[S(k,l)]_{ij} = \frac{i}{\sqrt{2}}(\delta_{ki}\delta_{lj} + \delta_{kj}\delta_{li}) \quad 1 \leq k < l \leq n \tag{5.1}$$

$$[A(k,l)]_{ij} = \frac{1}{\sqrt{2}}(\delta_{ki}\delta_{lj} - \delta_{kj}\delta_{li}) \quad 1 \leq k < l \leq n \tag{5.2}$$

$$[D(m)]_{ij} = i\delta_{mi}\delta_{mj} \quad 1 \leq m \leq n \tag{5.3}$$

These are simple anti-Hermitian matrices. They obviously form a complete basis for $u(n)$ with $(n^2-n)/2$ A's, $(n^2-n)/2$ S's, and n D's. (The algebra $o(n)$ is spanned by the real antisymmetric matrices A alone.) The structure constants are all real, so they constitute a real Lie algebra. We will call these n^2 elements (or $(n^2-n)/2$ elements in the case of $o(n)$) collectively X_i. They have the important property that

$$-\text{Trace}\,(X_iX_j) = \delta_{ij} \tag{5.4}$$

It is instructive to construct one-parameter subgroups with these basis elements. First note that

$$A^2(k,l) = S^2(k,l) = -\frac{1}{2}I(k,l) \qquad D^2(m) = iD(m), \tag{5.5}$$

where $I(k,l)$ is a diagonal matrix with 1's at the k'th and l'th places on the diagonal and zeros elsewhere. Consequently

$$e^{tS(k,l)} = I(k,l)\cos t/\sqrt{2} + \sqrt{2}S(k,l)\sin t/\sqrt{2} - I(k,l) + I$$
$$e^{tA(k,l)} = I(k,l)\cos t/\sqrt{2} + \sqrt{2}A(k,l)\sin t/\sqrt{2} - I(k,l) + I$$
$$e^{tD(m)} = iD(m)(1-e^{it}) + I$$

These functions are all bounded and periodic, thus compact. Note that it is exactly the anti-Hermitian character of S, A, and D that provides the necessary sign reversals so that the series converge to sin's and cos's rather than hyperbolic functions.

The unimodular group $SU(n)$ has the additional constraint that $\det|e^X| = 1$, so by Theorem 2.5, $\text{Trace}(X) = 0$. The elements A and S are already traceless, but the diagonal elements are harder to construct. The set

$$[T(m)]_{ij} = \frac{i}{\sqrt{2}}(\delta_{mi}\delta_{mj} - \delta_{m+1,i}\delta_{m+1,j}) \quad 1 \le m \le n-1 \tag{5.6}$$

is complete and traceless, but it does not satisfy the orthogonality condition (5.4). We can, however, form a positive definite scalar product,

$$< X_i, X_j >= -\text{Trace}(X_iX_j), \tag{5.7}$$

(where the X_i are any of the S, T, or A defined above) and use it to construct an orthonormal set using the Gram-Schmidt procedure. The resulting $n-1$ elements can be combined with the A's and S's to make a complete set of n^2-1 orthonormal basis elements for $su(n)$.

$$< X_i, X_j >= \delta_{ij} \tag{5.8}$$

This condition in turn leads to an important condition on the structure constants. Let

$$[X_{_i}, X_j] = \sum_k f_{ij}^k X_k.$$

Then

$$\text{Trace}\,(X_l[X_i, X_j]) = -f_{ij}^l$$

The cyclic property of the trace requires that

$$f_{ij}^l = -f_{il}^j = -f_{lj}^i = f_{jl}^i = f_{li}^j = -f_{ji}^l \tag{5.9}$$

i.e. the structure constants are antisymmetric with respect to the exchange of any pair of indices. The Cartan-Killing form is negative definite, *i.e.* $(X_i, X_i) < 0$, for such an algebra because

$$(X_i, X_i) = \sum_{j,k} f_{ij}^k f_{ik}^j = -\sum_{j,k} (f_{ij}^k)^2,$$

cf. (4.55) and (4.56).

As an exercise one can prove that if the elements X_i of the algebra are $n \times n$ matrices, then

$$(X_i, X_j) = 2(n-1)\text{Trace}(X_i X_j)$$

so that

$$(X_i, X_j) = -2(n-1)\delta_{ij}$$

This simple result can be generalized with the help of the following definition:

Definition 5.1 *A* **compact algebra** *is a real Lie algebra with a negative definite Cartan-Killing form.*

If the Killing form is negative definite we can always construct a positive definite scalar product as in (5.7) and use the Gram-Schmidt procedure to create an orthonormal set of basis elements. If the algebra is real to start with, the new basis will still have real structure constants. Then $(f_{ij}^k)^2$ can't be negative. If the algebra has no invariant abelian subalgebras then

$$\sum_{j,k} (f_{ij}^k)^2 > 0$$

for all i. We have proved

Theorem 5.1 *If $\mathcal{L}$ is a real, compact, semi-simple Lie algebra then its basis can be made orthonormal, and in this basis the structure constants are antisymmetric under the exchange of any pair of indices.*

We can now prove the so-called Peter-Weyl theorem.

Theorem 5.2 *A real, connected, semi-simple Lie group $\mathcal{G}$ is compact if and only if its corresponding real Lie algebra is compact.*

Proof: First supppose that the real Lie algebra is compact. As shown above, the structure constants are completely antisymmetric. The regular representation therefore consists of real antisymmetric matrices, which exponentiate to a unitary group. The group is thus a subgroup, at least, of $O(n)$. Subgroups of compact groups are compact, however, so $\mathcal{G}$ is compact.

Now assume that $\mathcal{G}$ is compact. According to Theorem 4.??, $\mathcal{G}$ is equivalent to a unitary group, so its Lie algebra consists of anti-Hermitian matrices. We have just learned how to choose a basis for such an algebra so that the structure constants are antisymmetric and the Cartan-Killing form is negative definite. The C-K form is invariant under a change of basis, however, so the theorem is proved.

There is a potential source of confusion that arises here, because physicists like to make compact groups by exponentiating Hermitian matrices, *i.e.* e^{iX}, whereas mathematicians prefer to exponentiate anti-Hermitian matrices, e^X. As we saw in connection with angular momentum, Example 4.6, the e^{iX} prescription leads to a complex Lie algebra (in fact the structure constants are pure imaginary) and a positive definite metric. The e^X prescription produces a real Lie algebra and a negative definite metric. All the results in this chapter so far **assume the e^X prescription.**

In a sense the choice of prescription is trivial (for example, in dealing with angular momentum one can always think of iJ_i as a set of anti-Hermitian matrices and use all the previous results verbatim). This is partly because all the elements are multiplied by the same factor of i. The serious problem arises when an algebra, even a real algebra, is defined on the field of *complex* numbers. Then we have the option of absorbing some of the i's into *some* elements of the algebra changing their character from anti-Hermitian to Hermitian. A good example of this is the Weyl canonical form from Chapter 4.

5.2 The Weyl canonical form and the compact real form

The Weyl form summarized in Table ??? always describes a real algebra. In the regular representation it consists of real, symmetric matrices as the next theorem shows.

Theorem 5.3 *If H_i, E_α, and $E_{-\alpha}$ are eigenelements in the Weyl canonical basis, and if the normalization is chosen so that $(H_i, H_j) = \delta_{ij}$ and $(E_\alpha, E_{-\alpha}) = 1$, then in the regular representation $E_\alpha^T = E_{-\alpha}$ and $H_i^T = H_i$.*

Proof: First note that (H_i, H_i) and $(E_\alpha, E_{-\alpha})$ (for each pair of roots, $\boldsymbol{\alpha}$ and $-\boldsymbol{\alpha}$) are the only non-zero Killing forms. This follows from Theorem 4.20. If this normalization is adopted then the Cartan metric has the form,

$$g = \left[\begin{array}{ccc|cc|cc} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ \hline & & & & 1 & & \\ & & & 1 & & & \\ \hline & & & & & 1 & \\ & & & & & 1 & \\ \hline & & & & & & \ddots \end{array}\right] \tag{5.10}$$

Now let σ_a, σ_b, and σ_c be any three basis elements in the Weyl canonical form. Then from (4.58)

$$(\sigma_a, [\sigma_b, \sigma_c]) = (\sigma_b, [\sigma_c, \sigma_a]) = (\sigma_c, [\sigma_a, \sigma_b])$$

It is clear from (5.10) that for each σ_a there is one and only one element, say σ_d for which (σ_a, σ_d) is non zero. Then

$$f_{bc}^d(\sigma_a, \sigma_d) = f_{ca}^e(\sigma_b, \sigma_e) = f_{ab}^g(\sigma_c, \sigma_g).$$

Here σ_d is the unique element that makes (σ_a, σ_d) non-zero, etc. With the assumed normalization then

$$f_{bc}^d = f_{ca}^e = f_{ab}^g.$$

The statement that $E_\alpha^T = E_{-\alpha}$ is equivalent to the statement that $f^i_{cj} = f^j_{bi}$, where b refers to E_α, c refers to $E_{-\alpha}$, and i and j refer to all those elements for which f is non-zero (see equation 4.9). This follows trivially from the above equality. Furthermore if i refers to H_i then $f^a_{ib} = 0$ unless $a = b$, so H_i is diagonal.

Example 5.2 *Angular Momentum*

The basis elements H_1, E_+, and E_- defined in Example 4.6 have the following non-zero commutation relations:

$$[H_1, E_\pm] = \pm\frac{1}{\sqrt{2}}E_\pm$$

$$[E_+, E_-] = \frac{1}{\sqrt{2}}H_1$$

In terms of this new basis the regular representation is

$$\boldsymbol{R}(H_1) = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \boldsymbol{R}(E_+) = \tfrac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{5.11}$$

$$\boldsymbol{R}(E_-) = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Clearly the Weyl canonical form is always non-compact; in fact, maximally non-compact as we will show later. It can be made real and compact, however, by absorbing a few i's in the right places. Consider the basis given by

$$iH_i, \qquad \frac{i(E_\alpha + E_{-\alpha})}{\sqrt{2}}, \qquad \frac{(E_\alpha - E_{-\alpha})}{\sqrt{2}}. \tag{5.12}$$

In the regular representation H_i and $E_{\pm\alpha}$ are real and meet the conditions of Theorem 5.3. Therefore the elements (5.12) are anti-Hermitian. The metric tensor is diagonal, and its elements are normalized so that $g_{\mu\nu} = -\delta_{\mu\nu}$. It is easy to show that the structure constants are all real. Since the regular representation is faithful (at least for semi-simple algebras) (5.12) defines a real, compact group. This leads to the following definition:

Definition 5.2 *A real algebra constructed from the Weyl canonical form with the basis (5.12) using real coefficients is called the* **compact real form**, $\mathcal{L}_c$. *Its metric is diagonal, and in fact,* $g_{\mu\nu} = -\delta_{\mu\nu}$.

In the Weyl canonical form, the metric tensor is non-negative and symmetric. Equation (5.12) can be thought of as a complex linear transformation that diagonalizes it. Here is another choice:

$$H_i, \qquad \frac{(E_\alpha + E_{-\alpha})}{\sqrt{2}}, \qquad \frac{(E_\alpha - E_{-\alpha})}{\sqrt{2}} \tag{5.13}$$

This is a *real* linear transformation that also produces a real Lie algebra. In this basis

$$g = \left[\begin{array}{ccc|c|c|c} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ \hline & & & 1 & & \\ & & & & -1 & \\ \hline & & & & 1 & \\ & & & & & -1 \\ \hline & & & & & \ddots \end{array}\right] \tag{5.14}$$

In general, a real matrix g can be diagonalized by choosing another basis, $\sigma'_a = S_a{}^b \sigma_b$, where S is a non-singular real matrix. (See Definition 4.21). The new metric is

$$g' = SgS^T \tag{5.15}$$

It is well known (Korn and Korn) that a transformation of this form can diagonalize g and bring it into a standard form in which the diagonal elements are +1,-1, or 0. Furthermore, it is not possible to change the number of +1's, -1's, or 0's. The metric for semisimple algebras must be non-singular, so there can be no 0's in the diagonal form. We can thus define an invariant quantity called the *character* χ of a metric by transforming the metric to standard form and counting the number of +1's and −1's:

$$\chi = (\text{number of } +1\text{'s}) - (\text{number of -1's}) \tag{5.16}$$

The character of the compact real form $\mathcal{L}_c$, is

$$\chi(\mathcal{L}_c) = -(\text{dimension of the algebra}). \tag{5.17}$$

The character of the Weyl canonical form can be inferred from (5.14).

$$\chi(\text{Weyl canonical form}) = +(\text{rank}). \tag{5.18}$$

The real algebras defined by the $\mathcal{L}_c$ and the Weyl canonical form represent opposite ends of the campactness spectrum. The Weyl canonical form is minimally compact in the sense that it contains no compact subalgebras except for the angular momentum-type subalgebras consisting of $E_{\pm\alpha}$ and h_α. Other real forms will have values of χ that are intermediate between (5.17) and (5.18) and correspondingly larger compact subgroups.

Apart from some special cases, the character function is non-degenerate, *i.e.* algebras with the same dimension and character are equivalent. Among the complex simple Lie algebras there are only fourteen instances in which the character function does not provide a unique classification of the real forms. These occur only in the systems A_{2n-1} and D_n and mostly in algebras of high dimensionality. Gilmore provides a table of these algebras and an algorithm for deciding when this degeneracy occurs.

It is possible to transform from the diagonal Weyl canonical form (5.14) to $\mathcal{L}_c$ by a complex linear transformation called the *Weyl unitary trick*, a pretentious name for multiplying by i!

$$\begin{aligned} H_i &\to iH_i \\ \frac{(E_\alpha + E_{-\alpha})}{\sqrt{2}} &\to \frac{i(E_\alpha + E_{-\alpha})}{\sqrt{2}} \\ \frac{(E_\alpha - E_{-\alpha})}{\sqrt{2}} &\to \frac{(E_\alpha - E_{-\alpha})}{\sqrt{2}} \end{aligned} \tag{5.19}$$

Clearly it can change the character of an algebra and change a compact algebra into a non-compact one and vice versa. If we were to multiply indiscriminantly by i's, however, we would generate complex structure constants. The problem of finding all possible real algebras concealed in a given complex algebra can be restated as follows: how many ways are there to play the Weyl unitary trick on the compact real form without getting complex structure constants?

Some insight into this question can be gained by observing that if the basis elements (5.13) are grouped into two sets, $\mathcal{K}$ and $\mathcal{P}$, with

$$\begin{aligned} H_i, \quad (E_\alpha + E_{-\alpha})/\sqrt{2} \quad &\subseteq \mathcal{P} \\ (E_\alpha - E_{-\alpha})/\sqrt{2} \quad &\subseteq \mathcal{K}, \end{aligned}$$

then

$$\begin{aligned} [\mathcal{K}, \mathcal{K}] &\subseteq \mathcal{K} \\ [\mathcal{K}, \mathcal{P}] &\subseteq \mathcal{P} \end{aligned}$$

$$[\mathcal{P}, \mathcal{P}] \subseteq \mathcal{K}. \tag{5.20}$$

This maneuver is called the *Cartan decomposition.* Now we can do the Weyl unitary trick on $\mathcal{P}$, $\mathcal{P} \to i\mathcal{P}$. The i's cancel in (5.20), and the structure constants remain real. Thus we can transform the Weyl canonical form into $\mathcal{L}_c$ and vice versa. This is the basic paradigm for finding real forms. To find all the real forms we must first find all possible Cartan decompositions. (The correspondance is not one-to-one. Different decompositions can produce equivalent real forms, but every real form is associated with a Cartan decomposition.)

Example 5.3 *Real forms of Angular Momentum*

The basis for $\mathcal{L}_c$ is given by

$$\sigma_1 = iH_1, \quad \sigma_2 = i\frac{E_+ + E_-}{\sqrt{2}} \quad \sigma_3 = \frac{E_+ - E_-}{\sqrt{2}}$$

then

$$[\sigma_i, \sigma_j] = -\frac{1}{2}\epsilon_{ijk}\sigma_k. \tag{5.21}$$

In terms of this new basis the regular representation is

$$\boldsymbol{R}(\sigma_1) = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \boldsymbol{R}(\sigma_2) = \tfrac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \tag{5.22}$$

$$\boldsymbol{R}(\sigma_3) = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The matrices are real and anti-Hermitian. If we now choose $\sigma_1 \subseteq \mathcal{K}$ and $\sigma_2, \sigma_3 \subseteq \mathcal{P}$ and perform the Weyl unitary trick, $\sigma_1 \to \sigma_1$, $\sigma_2 \to i\sigma_2$, and $\sigma_3 \to i\sigma_3$, the regular representation becomes

$$\boldsymbol{R}(\sigma_1) = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \boldsymbol{R}(\sigma_2) = \tfrac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \tag{5.23}$$

$$\boldsymbol{R}(E_-) = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

The matrices are still real. (At first sight it is surprising that we can multiply a real matrix by i and get another real matrix, but the regular representation depends on the basis chosen. The representations (5.22) and (5.23) are based on different bases.) The first element σ_1 is unchanged and thus anti-Hermitian. It represents a trivial compact subgroup. The remaining elements are now real and symmetric. It is evident from the commutation relations (5.21) that there is nothing special about σ_1. We could have chosen any one of the three σ's to represent $\mathcal{K}$ and the remaining two to populate $\mathcal{P}$. There are thus three different Cartan decompositions producing three isomorphic, non-compact real forms. We will have more to say about isomorphism in the next section.

There is a trick for finding all possible decompositions that was probably inspired by the following observation: Consider a transformation ψ that simply takes the complex conjugate of an element of the compact real form, $\mathcal{L}_c$. Then if $X \in \mathcal{K}$, $\psi(X) = X$, and if $Y \in \mathcal{P}$, then $\psi(Y) = -Y$. This transformation has three important properties: (1) It maps the Lie algebra onto itself; $\psi(\mathcal{L}) = \mathcal{L}$. (2) $\psi(\psi(X)) = X$ for all $X \in \mathcal{L}$. (3) All the basis elements in the Cartan decomposition are eigenelements of ψ with eigenvalues ± 1. Mappings of this sort are called *involutive automorphisms.* The task of finding all possible Cartan decompositions is thus equivalent to finding all the involutive automorphisms.

5.3 Automorphism

Definition 5.3 *Lie group automorphism*

Let ϕ be a mapping of the Lie group $\mathcal{G}$ onto itself, $\phi(\mathcal{G}) = \mathcal{G}$, such that

$$\phi(a)\phi(b) = \phi(ab)$$

for all $a, b \in \mathcal{G}$. Then ϕ is called a **Lie group automorphism**. *If we think of this mapping in terms of conjugations*

$$\phi_d(a) = dad^{-1}$$

then the set of all elements like d that map $\mathcal{G} \to \mathcal{G}$ constitute a group called Aut($\mathcal{G}$), the group of Lie group automorphisms.

Definition 5.4 *Lie algebra automorphism.*

Let ψ be a mapping of the Lie algebra $\mathcal{L}$ onto itself, $\psi(\mathcal{L}) = \mathcal{L}$, such that

$$\psi(\alpha X + \beta Y) = \alpha\psi(X) + \beta\psi(Y)$$

$$\psi([X,Y]) = [\psi(X), \psi(Y)]$$

for all $X, Y \in \mathcal{L}$. *(The coefficients* α *and* β *might be real or complex.) Then* ψ *is called a* **Lie algebra automorphism.** *The product of two automorphisms* ψ *and* ϕ *is defined by* $(\psi\phi)(X) = \psi(\phi(X))$. *The group of Lie algebra automorphisms can be defined in the same way as the Lie group automorphisms,* i.e. *the set of all elements* d *such that*

$$\psi(X) = dXd^{-1}$$

maps $\mathcal{L} \to \mathcal{L}$. *This group is called Aut(*$\mathcal{L}$*).*

Definition 5.5 *Inner and outer automorphism.*

If the elements represented by d *in the two previous definitions are contained in the group* $\mathcal{G}$*, the automorphisms are said to be "inner." We will use the notation Int(*$\mathcal{G}$*) and Int(*$\mathcal{L}$*) for group and algebra automorphisms respectively. Those automorphisms that are not inner are said to be "outer."*

The relationship between $\mathrm{Aut}(\mathcal{G})$ and $\mathrm{Int}(\mathcal{G})$ turns out to be pivotal in cataloging the possible real forms of Lie algebras. The following statements are more or less obvious: (1) $\mathrm{Int}(\mathcal{G})$ is identical with the group $\mathcal{G}$. (2) $\mathrm{Int}(\mathcal{G})$ is an invariant subgoup of $\mathrm{Aut}(\mathcal{G})$. (3) Both groups contain the identity element. We can invoke the formalism of factor groups from Section 3.2 particularly theorems **??** and **??**. The factor group $\mathrm{Aut}(\mathcal{G})/\mathrm{Int}(\mathcal{G})$ consists of the distinct cosets of $\mathrm{Int}(\mathcal{G})$. The particular coset containing the identity, call it $\mathrm{Aut}_0(\mathcal{G})$, is identical to $\mathrm{Int}(\mathcal{G})$=$\mathcal{G}$. In summary

$$\mathrm{Aut}_0(\mathcal{G}) = \mathrm{Int}(\mathcal{G}) = \mathcal{G}$$

Similar arguments can be made for the Lie algebra automorphism groups, $\mathrm{Aut}(\mathcal{L})$ and $\mathrm{Int}(\mathcal{L})$. For example, if $a = e^{\epsilon X}$ is an element of $\mathrm{Int}(\mathcal{L})$, and X and Y are elements of $\mathcal{L}$, then the automorphism $\psi_a(Y) = aYa^{-1}$ is connected to the identity in the limit $\epsilon \to 0$. Consequently, $\mathrm{Aut}_0(\mathcal{L})$=$\mathrm{Int}(\mathcal{L})$. Furthermore, the factor group $\mathrm{Aut}(\mathcal{L})/\mathrm{Int}(\mathcal{L})$ has the same coset structure as $\mathrm{Aut}(\mathcal{G})/\mathrm{Int}(\mathcal{G})$.

Now let Y_i be a set of basis elements for $\mathcal{L}$. Then

$$\psi_a(Y_i) = aY_ia^{-1} \approx Y_i + \epsilon[X, Y_i] = Y_i + \epsilon \sum_j \gamma_i{}^j(X)Y_j.$$

At least in the case of semi-simple algebras, this represents a mapping of $\mathcal{L}$ onto itself. We are not entitled to claim that $\mathrm{Int}(\mathcal{L})$=$\mathcal{L}$, since $\mathrm{Int}(\mathcal{L})$ is a

group and $\mathcal{L}$ is an algebra, but the automorphism group $\mathrm{Int}(\mathcal{L})$ induces an *iso*morphism, $\mathcal{L} \to \mathcal{L}$.

We are now in a position to prove an important result that holds for all $\mathrm{Aut}(\mathcal{L})$ and not just the connected component. It turns out that the Killing form is invariant under algebra automorphisms.

Theorem 5.4 *If ψ is any Lie algebra automorphism, then*

$$(\psi(X), \psi(Y)) = (X, Y)$$

for all $X, Y \in \mathcal{L}$. Here (X, Y) is the Killing form definition (???).

Proof: Let $\sigma_1, \sigma_2, \ldots, \sigma_n$ form a basis for $\mathcal{L}$. Then $\psi(\sigma_1), \psi(\sigma_2), \ldots, \psi(\sigma_n)$ will also be a basis. Thus there exists a non-singular $n \times n$ matrix S such that $\psi(\sigma_i) = \sum_j S_{ij}\sigma_j$. Since ψ is an automorphism,

$$\psi([X, \sigma_j]) = [\psi(X), \psi(\sigma_j)] = \psi(\sum_k \sigma_k R(X)^k{}_j) = \sum_k \psi(\sigma_k) R(X)^k{}_j,$$

so that

$$\sum_l S_{jl}[\psi(X), \sigma_l] = \sum_{k,p} S_{kp}\sigma_p R(X)^k{}_j.$$

But

$$[\psi(X), \sigma_l] = \sum_q \sigma_q R(\psi(X))^q{}_l.$$

Combining the last two equations gives

$$\sum_{l,q} S_{jl}\sigma_p R(\psi(X))^p{}_l = \sum_{k,p} R(X)_{jk} S_{kp}\sigma_p.$$

It follows that

$$R(\psi(X)) = S^{-1} R(X) S. \tag{5.24}$$

Thus in the regular representation an automorphism can always be represented by a similarity transformation. The rest is simple:

$$(\psi(X), \psi(Y)) = \mathrm{trace}\{R(\psi(X)), R(\psi(Y))\} = \mathrm{trace}\{S^{-1}R(X)SS^{-1}R(Y)S\}$$

$$= \mathrm{trace}\{S^{-1}R(X)R(Y)S\} = (X, Y)$$

The last equality follows from the invariance of the trace under a similarity transformation.

The importance of this result lies in the fact that all the internal structure of the Dykin diagrams that were developed in the previous chapter was

computed using the Killing form. Consequently, Lie algebra automorphisms *do not change the internal structure of the Dynkin diagrams.* The only possible automorphisms are those that at most re-label some of the lines and/or vertices. There are only a few ways of doing this, hence the factor groups $\mathrm{Aut}(\mathcal{L})/\mathrm{Int}(\mathcal{L})$ and $\mathrm{Aut}(\mathcal{G})/\mathrm{Int}(\mathcal{G})$ have a simple coset structure corresponding to the few allowable changes to the Dynkin diagrams.

Definition 5.6 *An automorphism ψ is said to be* **involutive** *if $\psi^2 = I$,* i.e. *if $\psi(\psi(X)) = X$ for evey element X in the group or algebra.*

Theorem 5.5 *(Cartan's theorem) Let $\mathcal{L}$ be a semisimple complex Lie algebra and ψ be an involutive automorphism. Then $\mathcal{L}$ can be decomposed into two subspaces, $\mathcal{L} = \mathcal{K} + \mathcal{P}$ such that $\psi(\mathcal{K}) = +\mathcal{K}$ and $\psi(\mathcal{P}) = -\mathcal{P}$. $\mathcal{K}$ is a subalgebra, and $\mathcal{P}$ is an orthogonal complementary subspace.*

Proof: Let $\sigma_1, \sigma_2, \ldots, \sigma_n$ be a basis of $\mathcal{L}$. The action of ψ can be represented by a matrix T.

$$\psi(\sigma_i) = \sum_j T_{ij}\sigma_j$$

Since ψ is involutive, $T^2 = I$ or $(T - I)(T + I) = 0$. It is not hard to show that a matrix that can be factored like this can be diagonalized, and its eigenvalues are all ± 1. We can thus choose a new basis of $\mathcal{L}$ so that it is composed of eigenelements of T. Then let $\mathcal{K}$ be the space spanned by the basis elements with eigenvalue +1, and let $\mathcal{P}$ be the space spanned by those with eigenvalue -1. The two subspaces are orthogonal, since

$$(\mathcal{K}, \mathcal{P}) = (T\mathcal{K}, T\mathcal{P}) = -(\mathcal{K}, \mathcal{P}) = 0.$$

The subspace $\mathcal{K}$ is closed under commutation, since if $K, K' \in \mathcal{K}$, then

$$([K, K'], \mathcal{P}) = (T[K, K'], T\mathcal{P}) = ([TK, TK'], T\mathcal{P}) = -([K, K'], \mathcal{P}) = 0$$

so

$$[\mathcal{K}, \mathcal{K}] \subseteq \mathcal{K}$$

In the same way we can show that

$$[\mathcal{K}, \mathcal{P}] \subseteq \mathcal{P}$$

$$[\mathcal{P}, \mathcal{P}] \subseteq \mathcal{K}$$

This is just the *Cartan decomposition*, (**??**). Thus we have proved the following corollary:

Corollary 5.6 *For every involutive automorphism there is a real Lie algebra obtained by performing the corresponding Cartan decomposition on $\mathcal{L}_c$ and replacing $\mathcal{P} \to i\mathcal{P}$.*

Since $\mathcal{L}_c$ was compact to start with, $\mathcal{K}$ is still compact, but the entire algebra $\mathcal{K} + i\mathcal{P}$ is non-compact. $\mathcal{K}$ is thus the maximal compact subalgebra. Furthermore, $\mathcal{L}_c$ is a real algebra, and the i's all cancel in (**??**), so the new algebra is also real.

The Cartan decomposition is especially important because in fact *all* the real forms of a complex semisimple algebra can be found in this way. We refer the reader to Cornwell for a list of references and history of the proof of this remarkable theorem.

5.4 The Catalog of Automorphisms

It turns out to be useful to think of the automorphisms of $\mathcal{L}_c$ in the following form:

$$\psi = \phi^{-1}\theta\phi \tag{5.25}$$

Here it is assumed that ϕ is an element of the group and hence an inner automorphism. Its purpose is to rearrange the basis elements into some convenient form so that θ, the "chief" automorphism can be represented as simply as possible. Different ψ's corresponding to the same θ and different ϕ's are isomorphic, and have the same character, since the trace of a matrix is invariant under a similarity transformation. We are really interested in cataloging the various chief automorphisms. Even so, many θ's are equivalent to one another. The $\theta's$ in turn can be "inner" or "outer" as discussed previously. We begin with the inner automorphisms.

5.4.1 Inner automorphisms

Assume that $\mathcal{L}_c$ is described with the following basis elements: ih_{α_i} $(i = 1, 2, \ldots, l)$ followed by $i(E_\alpha + E_{-\alpha})/\sqrt{2}$ and $(E_\alpha - E_{-\alpha})/\sqrt{2}$ for each $\boldsymbol{\alpha}$ contained in the set of positive roots. (l is the dimension of the Cartan subalgebra.) All chief inner involutive automorphisms $\boldsymbol{\theta}$ can then be represented with the following diagonal matrix.

$$\boldsymbol{\theta} = \left[\begin{array}{c|c|c|c|c} I_l & & & & \\ \hline & \pm I_{\alpha_1} & & & \\ \hline & & \pm I_{\alpha_2} & & \\ \hline & & & \pm I_{\alpha_3} & \\ \hline & & & & \ddots \end{array}\right] \equiv I_{k,p} \tag{5.26}$$

The submatrix I_l is a $l \times l$ unit matrix that maps $ih_{\alpha_i} \rightarrow ih_{\alpha_i}$. The remaining 2×2 unit matrices $\pm I_{\alpha_i}$ transform $i(E_{\alpha_i}+E_{-\alpha_i})/\sqrt{2} \rightarrow \pm i(E_{\alpha_i}+E_{-\alpha_i})/\sqrt{2}$ and $(E_{\alpha_i}-E_{-\alpha_i})/\sqrt{2} \rightarrow \pm(E_{\alpha_i}-E_{-\alpha_i})/\sqrt{2}$. The choice of + or - is arbitrary for each matrix I_{α_j}, where $\boldsymbol{\alpha_j}$ is a *simple* positive root. If $\boldsymbol{\alpha_j}$ is a sum of simple roots, *e.g.* if $\boldsymbol{\alpha}_j = \boldsymbol{\alpha}_m + \boldsymbol{\alpha}_n$, then the sign of I_{α_j} must equal the product of the sign of I_{α_m} and the sign of I_{α_n}. There are thus 2^l possible choices, all corresponding to different but possibly isomorphic algebras. When performing the Cartan decomposition, the k elements with + signs are grouped into $\mathcal{K}$ and the remaining $\surd$ elements with - signs make up $\mathcal{P}$. We will refer to this as the "$I_{k,p}$ algorithm."

It is easy to show that (5.26) defines an inner involutive automorphism. The fact that *all* such automorphisms can be obtained from $\boldsymbol{\theta}$ by a similarity transformation was first proved by Gantmacher (1939).

Example 5.4 *The real form of* $su(2)$

The basis defined in Example 5.3 is consistent with (5.26). Except for the identity there is only one automorphism, $\sigma_1 \rightarrow \sigma_1$, $\sigma_2 \rightarrow -\sigma_2$, and $\sigma_3 \rightarrow -\sigma_3$. The Weyl unitary trick then leads to one of the three equivalent non-compact real forms mentioned in Example 5.2.

Example 5.5 *The real form of* $su(3)$

There are three positive roots, $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, and $\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$. The choice of signs for $\pm I_{\alpha_1}$ and $\pm I_{\alpha_2}$ is arbitrary. Thus there are four automorphisms, one of which is the identity. The remaining three are $+--$, $-+-$, and $--+$. They all have a character $\chi = -\mathrm{Tr}\,\boldsymbol{\theta} = 0$. In this case all three forms yield the algebra of $su(2,1)$.

5.4.2 Dynkin diagrams and outer automorphisms

It was claimed (so far without proof) that there was an automorphism associated with interchanging the simple positive roots in a Dynkin diagram.

In fact, any interchange between pairs of roots that keeps the "scalar product," $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (h_\alpha, h_\beta)$ between each pair of roots $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ unchanged constitutes an involutive automorphism so long as certain sign conventions are observed. It turns out that all the permutations associated with Weyl reflections, Section 4.5, generate inner automorphisms (Cornwell, Gantmacher), which are already contained in the previous algorithm (Section 5.4.1). The only new automorphisms are associated with permutations of simple roots in the Dynkin diagrams. The details are given in the following theorem.

Theorem 5.7 *Consider a linear transformation τ that maps the roots of a simple or semi-simple complex Lie algebra onto one another in such a way that it satisfies the following three conditions:*

1. *If $\boldsymbol{\alpha}$ is a root, then so is $\tau(\boldsymbol{\alpha})$.*
2. *If $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$, then $\tau(\boldsymbol{\alpha}) \neq \tau(\boldsymbol{\beta})$,* i.e. *no roots are lost and no new roots are created in the mapping.*
3. *The mapping is linear, i.e.*

$$\tau(x\boldsymbol{\alpha} + y\boldsymbol{\beta}) = x\tau(\boldsymbol{\alpha}) + y\tau(\boldsymbol{\beta}) \tag{5.27}$$

for any complex numbers x and y.

Further, assume the following sign conventions:

$$\psi_\tau(h_\alpha) = h_{\tau(\alpha)} \tag{5.28}$$

and

$$\psi_\tau(E_\alpha) = \chi_\alpha E_{\tau(\alpha)} \tag{5.29}$$

Here $\chi_\alpha = +1$ if $\boldsymbol{\alpha}$ is a simple positive root, $\chi_{-\alpha} = \chi_\alpha$, and finally

$$\chi_{\alpha+\beta} = \chi_\alpha \chi_\beta N_{\tau(\alpha),\tau(\beta)}/N_{\alpha,\beta} \tag{5.30}$$

Then τ generates an involutive automorphism, which we will call ψ_τ.

Proof: Lie algebra automorphisms must preserve commutation relations (see Definition 5.4). So, for example, (**??**) with $\lambda_\alpha = 1$ gives

$$\psi_\tau(h_\alpha) = \psi_\tau([E_\alpha, E_{-\alpha}]) = h_{\tau(\alpha)} \equiv [E_{\tau(\alpha)}, E_{\tau(-\alpha)}] =$$

$$\chi_\alpha \chi_{-\alpha} [\psi_\tau(E_\alpha), \psi_\tau(E_{-\alpha})] = [\psi_\tau(E_\alpha), \psi_\tau(E_{-\alpha})]$$

Thus if we make the natural definition,

$$h_{\tau(\alpha)} \equiv [E_{\tau(\alpha)}, E_{\tau(-\alpha)}], \tag{5.31}$$

then

$$\psi_\tau([E_\alpha, E_{-\alpha}]) = [\psi_\tau(E_\alpha), \psi_\tau(E_{-\alpha})], \tag{5.32}$$

as required by the definition of automorphism.

The same argument works for (**??**).

$$\psi_\tau(E_{\alpha+\beta}) = \chi_{\alpha+\beta} E_{\tau(\alpha+\beta)} = \chi_{\alpha+\beta} E_{\tau(\alpha)+\tau(\beta)} =$$

$$\chi_{\alpha+\beta}[E_{\tau(\alpha)}, E_{\tau(\beta)}]/N_{\tau(\alpha),\tau(\beta)} = \chi_\alpha \chi_\beta [E_{\tau(\alpha)}, E_{\tau(\beta)}]/N_{\alpha,\beta} =$$

$$[\psi_\tau(E_\alpha), \psi_\tau(E_\beta)]/N_{\alpha,\beta} = \psi_\tau([E_\alpha, E_\beta])/N_{\alpha,\beta}$$

We can conclude that

$$\psi_\tau([E_\alpha, E_\beta]) = [\psi_\tau(E_\alpha), \psi_\tau(E_\beta)] \tag{5.33}$$

Finally (**??**) is transformed as follows:

$$\psi_\tau([h_\alpha, E_\beta]) = \psi_\tau([[E_\alpha, E_{-\alpha}], E_\beta]) =$$

$$\psi_\tau([[E_\beta, E_\alpha], E_{-\alpha}]) = N_{\beta,\alpha} \psi_\tau([E_{\beta+\alpha}, E_{-\alpha}]) =$$

$$N_{\beta,\alpha}[\psi_\tau(E_{\beta+\alpha}), \psi_\tau(E_{-\alpha})] = N_{\beta,\alpha}\chi_{\beta+\alpha}\chi_{-\alpha}[E_{\tau(\beta+\alpha)}, E_{\tau(-\alpha)}] =$$

$$\frac{N_{\beta,\alpha}\chi_{\beta+\alpha}\chi_{-\alpha}}{N_{\tau(\beta),\tau(\alpha)}}[[E_{\tau(\beta)}, E_{\tau(\alpha)}], E_{\tau(-\alpha)}] =$$

$$\frac{N_{\beta,\alpha}\chi_{\beta+\alpha}\chi_{-\alpha}}{N_{\tau(\beta),\tau(\alpha)}}[h_{\tau(\alpha)}, E_{\tau(\beta)}] = [\psi(h_\alpha), \psi(E_\beta)]$$

In the last step we used the fact that

$$\frac{N_{\beta,\alpha}\chi_{\beta+\alpha}\chi_{-\alpha}\chi_\beta}{N_{\tau(\beta),\tau(\alpha)}} = 1$$

Again we conclude that

$$\psi_\tau([h_\alpha, E_\beta]) = [\psi(h_\alpha), \psi(E_\beta)] \tag{5.34}$$

Since ψ_τ is an automorphism, we can use Theorem 5.4 to derive the following result:

Theorem 5.8

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\tau(\boldsymbol{\alpha}), \tau(\boldsymbol{\beta})) \tag{5.35}$$

Proof: From (**??**)

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (h_\alpha, h_\beta) = (\psi(h_\alpha), \psi(h_\beta)) =$$

$$(h_{\tau(\alpha)}, h_{\tau(\beta)}) = (\tau(\boldsymbol{\alpha}), \tau(\boldsymbol{\beta}))$$

The possible automorphisms are easy to enumerate with reference to the diagrams in Figure 4.9-4.14.

$\boldsymbol{A_n}$ Figure 4.14

All the roots have the same weights, but if we were to exchange one pair of roots, *e.g.* $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$, that would bring $\boldsymbol{u_1}$ adjacent to $\boldsymbol{u_3}$. This does not preserve the scalar product since $(\boldsymbol{u_2}, \boldsymbol{u_3})$ =?? whereas $(\boldsymbol{u_1}, \boldsymbol{u_3})$ =??. It is possible to exchange all the roots simultaneously $\boldsymbol{u_i} \leftrightarrow \boldsymbol{u_{n+1-i}}$. Thus there is only one chief outer involutary isomorphism. To put it another way, the factor group, $\mathrm{Aut}(\mathcal{L})/\mathrm{Int}(\mathcal{L})$ has only two elements, the identity and this one transformation.

$\boldsymbol{B_n}$, $\boldsymbol{C_n}$ Figure 4.10

The reflection that worked for A_n fails here because $\boldsymbol{u_1}$ and $\boldsymbol{u_n}$ have different weights. There are no automorphisms for these algebras.

$\boldsymbol{D_n}$, $\boldsymbol{E_6}$, $\boldsymbol{E_7}$, and $\boldsymbol{E_8}$ Figure 4.13

In general, there is one symmetry corresponding to the interchange $\boldsymbol{u_{n-1}} \leftrightarrow \boldsymbol{u_n}$. Because of the exceptional symmerty in the case $n = 4$, however, there are three possible exchanges, $\boldsymbol{u_1} \leftrightarrow \boldsymbol{u_3}$, $\boldsymbol{u_1} \leftrightarrow \boldsymbol{u_4}$, and $\boldsymbol{u_3} \leftrightarrow \boldsymbol{u_4}$. There is a single reflection allowed for E_6, i.e. $\boldsymbol{u_1} \leftrightarrow \boldsymbol{u_5}$ together with $\boldsymbol{u_2} \leftrightarrow \boldsymbol{u_4}$. There is no symmetry associated with the exceptional algebras, E_7, and E_8, and consequently, no outer automorphisms.

$\boldsymbol{G_2}$ Figure 4.8 and $\boldsymbol{F_4}$ Figure 4.11

No relflections are allowed becuase of the asymmetry of the weights.

Example 5.6 *Outer automorphism of su(3)*

The algebra $su(3)$ is A_2. The exchange $\alpha_1 \leftrightarrow \alpha_2$ produces the following automorphism:

$$\psi_\tau(ih_{\alpha_1}) = ih_{\alpha_2}, \quad \psi_\tau(ih_{\alpha_2}) = ih_{\alpha_1}$$

$$\psi_\tau(i(E_{\alpha_1} + E_{-\alpha_1})) = i(E_{\alpha_2} + E_{-\alpha_2})$$

$$\psi_\tau((E_{\alpha_1} - E_{-\alpha_1})) = (E_{\alpha_2} - E_{-\alpha_2})$$

$$\psi_\tau(i(E_{\alpha_2} + E_{-\alpha_2})) = i(E_{\alpha_1} + E_{-\alpha_1})$$

$$\psi_\tau((E_{\alpha_2} - E_{-\alpha_2})) = (E_{\alpha_1} - E_{-\alpha_1})$$

$$\psi_\tau(i(E_{\alpha_1+\alpha_2} + E_{-(\alpha_1+\alpha_2)}) = i(E_{(\alpha_1+\alpha_2)} + E_{-(\alpha_1+\alpha_2)})$$

$$\psi_\tau((E_{\alpha_1+\alpha_2} - E_{-(\alpha_1+\alpha_2)})) = (E_{(\alpha_1+\alpha_2)} - E_{-(\alpha_1+\alpha_2)})$$

The last two lines made use of the fact that $N_{\tau(\alpha_1),\tau(\alpha_2)} = N_{\alpha_2,\alpha_1} = -N_{\alpha_1,\alpha_2}$. The matrix $\boldsymbol{\theta}$ is obviously non-diagonal and its character $\chi = -\mathrm{Tr}\,\boldsymbol{\theta} = +2$. This turns out to be the algebra $sl(3,R)$.

5.4.3 Summary of algebras

We have discussed three involutive automorphisms: (1) complex conjugation, (2) inner automorphisms given by the $I_{k,p}$ algorithm (5.26), and (3) outer automorphisms produced by permutations of Dynkin diagrams. It can be shown that (2) and (3) together account for all possible involutive automorphisms, so (1) is in a sense redundant. It is simple to use, however, and it always produces one real algebra, the maximally non-compact Weyl canonical form. The other two algorithms are only relevant to the extent that there are additional real algebras. We will now give a summary of how they work in cases more complicated than the simple examples given so far. In the following summary it is assumed that the Cartan decomposition has already been performed and that the basis elements have been rearranged so that the k elements of the subalgebra $\mathcal{K}$ come first followed by the p elements of $\mathcal{P}$. From now on we work entirely in the regular representation, which can be partitioned as follows:

$$\mathcal{K} = \left[\begin{array}{c|c} A & 0 \\ \hline 0 & C \end{array}\right] \qquad \mathcal{P} = \left[\begin{array}{c|c} 0 & B \\ \hline -B^\dagger & 0 \end{array}\right] \tag{5.36}$$

Since the compact real form is anti-Hermitian, A is a $k \times k$ anti-Hermitian submatrix, and B is $p \times p$ and anti-Hermitian.

The inner automorphism (5.26) is a diagonal transformation, *i.e.* it does not rearrange the basis elements. After applying the Weyl unitary trick, $\mathcal{P} \to i\mathcal{P}$, and redefining the regular representation (*cf.* Example 5.3), $\mathcal{K}$ is unchanged and $\mathcal{P}$ becomes

$$i\mathcal{P} = \left[\begin{array}{c|c} 0 & B \\ \hline +B^\dagger & 0 \end{array}\right] \tag{5.37}$$

There are as many different ways of performing this algorithm as there are ways to choose the signs in equation (5.26). With a few exceptions, all choices with the same value of k yield isomorphic algebras, however. (Of course, $k + p = n$, the number of independent generators of the algebra.)

Outer automorphisms require that a new basis be chosen to diagonalize the transformation τ of Theorem 5.7. In terms of the new basis the regular

Root space	$\mathcal{L}_c$	Real Algebra	χ
A_{n-1}	$su(n)$	$sl(n,r)$	$n-1$
	$su(2n)$	$su^*(2n)$	$-1n-1$
	$su(k,p)$	$su(k,p)$	$-(k-p)^2+1$
B_n	$so(n)$	$so(k,p)$	$[(k+p)-(k-p)^2]/2$
C_n	$usp(2n)$	$sp(2n,r)$	n
	$usp(2n)$	$usp(2k,2p)$	$[-2(k+p)-4(k-p)^2]/2$
D_n	$so(n)$	$so(k,p)$	$[(k+p)-(k-p)^2]/2$
	$so(2n)$	$so^*(2n)$	$-n$

Table 5.1: The complex simple Lie algebras and their real forms.

representation looks like (5.36) and the transition to the real form is given by (5.37).

The resulting real forms obtained with these techniques are summarized in Table 5.1 for the classical complex simple Lie algebras. The exceptional algebras are summarized in several standard reference works.