Lie Algebras

Fulton B. Gonzalez

December 4, 2007

Contents

Introduction

Lie algebras are vector spaces endowed with a special non-associative multiplication called a *Lie bracket*. They arise naturally in the study of mathematical objects called *Lie groups*, which serve as groups of transformations on spaces with certain symmetries. An example of a Lie group is the group O(3) of rotations of the unit sphere $x^2 + y^2 + z^2 = 1$ in $\mathbb{R}^3$.

While the study of Lie algebras without Lie groups deprives the subject of much of its motivation, Lie algebra theory is nonetheless a rich and beautiful subject which will reward the physics and mathematics student wishing to study the structure of such objects, and who expects to pursue further studies in geometry, algebra, or analysis.

Lie algebras, and Lie groups, are named after Sophus Lie (pronounced "lee"), a Norwegian mathematician who lived in the latter half of the 19th century. He studied continuous symmetries (i.e., the Lie groups above) of geometric objects called manifolds, and their derivatives (i.e., the elements of their Lie algebras).

The study of the general structure theory of Lie algebras, and especially the important class of *simple* Lie algebras, was essentially completed by Élie Cartan and Wilhelm Killing in the early part of the 20th century. The concepts introduced by Cartan and Killing are extremely important and are still very much in use by mathematicians in their research today.

Chapter 1

Background Linear Algebra

This course requires some knowledge of linear algebra beyond what's normally taught in a beginning undergraduate course. In this section, we recall some general facts from basic linear algebra and introduce some additional facts needed in the course, including generalized eigenspaces and the Jordan canonical form of a linear map on a complex vector space. It is important that you familiarize yourself with these basic facts, which are also important in and of themselves.

1.1 Subspaces and Quotient Spaces

In what follows, all our vector spaces will be finite-dimensional over the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers. The notation $\mathbb{F}$ will denote either $\mathbb{R}$ or $\mathbb{C}$; when we talk about a vector space V over $\mathbb{F}$, we mean that V is a vector space either over $\mathbb{R}$ or $\mathbb{C}$.

A vector space can of course be defined over any algebraic field $\mathbb{F}$, and not just $\mathbb{R}$ or $\mathbb{C}$, but we will limit our vector spaces to these two fields in order to simplify the exposition. Many of our results carry over to vector spaces over arbitrary fields (the ones for $\mathbb{C}$ mostly carry over to algebraically closed fields), although the proofs may not necessarily be the same, especially for fields with prime characteristic.

Unless otherwise stated, all our vector spaces will be finite-dimensional.

Let V be a vector space over $\mathbb{F}$, and let A and B be any nonempty subsets of V and $\lambda \in \mathbb{F}$. We put

$$A + B = \{a + b \,|\, a \in A \text{ and } b \in B\} \tag{1.1}$$

and

$$\lambda A = \{\lambda a \,|\, a \in A\}. \tag{1.2}$$

For simplicity, if v is a vector in V, we put $v + B = \{v\} + B$. This is called the *translate* of B by the vector v. From (1.1) above, it's easy to see that

$$A + B = \bigcup_{a \in A} (a + B) = \bigcup_{b \in B} (b + A).$$

Note that a nonempty subset W of V is a vector subspace of V if and only if $W + W \subset W$ and $\lambda W \subset W$, for all $\lambda \in \mathbb{F}$.

Suppose now that W is a subspace of our vector space V. If v is any vector in V, the translate $v + W$ is called the *plane parallel to* W *through* v. For example, suppose that $V = \mathbb{R}^3$ (=3-dimensional Euclidean space) and W is the (x, y)-plane:

$$W = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \,|\, x, y \in \mathbb{R} \right\}$$

If

$$v = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$$

then $v + W$ is the plane $z = 3$.

As another example, if W is the one-dimensional subspace of $\mathbb{R}^3$ spanned by the vector

$$w = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}$$

(this is a straight line through the origin), and if

$$v = \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix},$$

then $v+W$ is the straight line in $\mathbb{R}^3$ through v and parallel to w; that is, it is the straight line specified by the parametric equations $x = 4 - t$, $y = 2t$, $z = 1 + 2t$.

Now let W be a subspace of a vector space V. Two translates $v_1 + W$ and $v_2 + W$ of W coincide if and only of $v_1 - v_2 \in W$. To see this, first assume that $v_1 + W = v_2 + W$. Then $v_1 = v_1 + 0 \in v_1 + W = v_2 + W$, so, $v_1 = v_2 + w$ for some $w \in W$, whence $v_1 - v_2 = w \in W$. Conversely, suppose that $v_1 - v_2 = w \in W$. Then $v_1 + W = v_2 + w + W = v_2 + W$.

The set of all translates of W by vectors in V is denoted by V/W, and is called the *quotient space* of V by W. (We pronounce V/W as "V *mod* W.") Thus

$V/W = \{v + W \mid v \in V\}$. V/W has a natural vector space structure whose vector addition is given by (1.1):

$$(v_1 + W) + (v_2 + W) = (v_1 + v_2) + (W + W) = v_1 + v_2 + W.$$

The scalar multiplication on V/W is given by

$$\lambda(v + W) = \lambda v + W.$$

Note that this definition of scalar multiplication is slightly different from the definition of scalar multiplication of sets in (1.2) above. (The reason being that $0 \cdot B = \{0\}$ for any nonempty subset B of V.) We will leave to the student the easy and routine verification that the operations above give rise to a vector space structure on V/W. Note that in V/W the zero vector is $0 + W = W$. (Later, when we study quotient spaces in greater detail, we'll just abuse notation and write the zero vector in V/W as 0.)

Proposition 1.1.1. *Let W be a subspace of V. Then* $\dim(V/W) = \dim V - \dim W$.

Proof. Let $B' = (w_1, \dots, w_m)$ be any basis of W, and extend this to a basis $B = (w_1, \dots, w_m, v_{m+1}, \dots, v_n)$ of V. We claim that $(v_{m+1} + W, \dots, v_n + W)$ is a basis of V/W. First we show that they span V/W. Let $v + W$ be an arbitrary element of V/W. Then $v = a_1 w_1 + \dots + a_m w_m + a_{m+1} v_{m+1} + \dots + a_n v_n$, for suitable scalars $a_1, \dots, a_n$. Then

$$\begin{aligned} v + W &= a_1 w_1 + \dots + a_m w_m + a_{m+1} v_{m+1} + \dots + a_n v_n + W \\ &= a_{m+1} v_{m+1} + \dots + a_n v_n + W \\ &= a_{m+1}(v_{m+1} + W) + \dots + a_n(v_n + W), \end{aligned}$$

so $v + W$ is a linear combination of $(v_{m+1} + W, \dots, v_n + W)$.

Next we show that $(v_{m+1} + W, \dots, v_n + W)$ is a linearly independent set in V/W. Suppose that $a_{m+1}(v_{m+1} + W) + \dots + a_n(v_n + W) = 0$. This is equivalent to $a_{m+1} v_{m+1} + \dots + a_n v_n + W = W$, so that $a_{m+1} v_{m+1} + \dots + a_n v_n \in W$. Since $w_1, \dots, w_m$ is a basis of W, we must have $a_{m+1} v_{m+1} + \dots + a_n v_n = b_1 w_1 + \dots + b_m w_m$, for suitable scalars $b_1, \dots, b_m$, and thus

$$-b_1 w_1 - \dots - b_m w_m + a_{m+1} v_{m+1} + \dots + a_n v_n = 0;$$

Since $(w_1, \dots, w_m, v_{m+1}, \dots, v_n)$ is a basis of V, we see that, in particular, $a_{m+1} = \dots = a_n = 0$. $\square$

It is easy to check that the sum of subspaces of V is also a subspace of V. Explicitly, if $W_1, \dots, W_k$ are subspaces of V, then $W_1 + \dots + W_k$ is also a subspace of V. This sum is called a *direct sum* if, for any vectors $w_1 \in W_1, \dots, w_k \in W_k$, the condition

$$w_1 + \dots + w_k = 0$$

implies that $w_1 = 0, \ldots, w_k = 0$. In this case, we will use the notation $W_1 \oplus \cdots \oplus W_k$ to denote the direct sum. Note that $(w_1, \ldots, w_m)$ is a linearly independent set if and only if the subspace sum $\mathbb{F}w_1 + \cdots + \mathbb{F}w_m$ is direct.

Exercise 1.1.2. Prove that if U and W are subspaces of V, then the sum $U + W$ is direct if and only if $U \cap W = \{0\}$.

Example 1.1.3. Let $\langle\ ,\ \rangle$ be an inner product on a real vector space V. If W is a subspace of V, put $W^\perp = \{v \in V \mid \langle v, w\rangle = 0 \text{ for all } w \in W\}$. The subspace $W^\perp$ is called the *orthogonal complement* of W in V. We have $V = W \oplus W^\perp$. (See [1], Theorem 6.29.)

Exercise 1.1.4. Let U and W be subspaces of V. Show that $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$. From this, show that $\dim(U \oplus W) = \dim U + \dim W$.

Given any subspace W of a vector space V, we can always find a subspace U of V such that $V = W \oplus U$. (U is called a *complementary subspace* to W.) The cases $W = \{0\}$ and $W = V$ being trivial, we can assume that $\{0\} \neq W \subsetneq V$. Take any basis $(w_1, \ldots, w_m)$ of W, extend this to a basis $(w_1, \ldots, w_m, v_{m+1}, \ldots, v_n)$ of V, and put $U = \mathbb{F}v_{m+1} + \cdots + \mathbb{F}v_n$. Then it is clear that $V = U \oplus W$. Since there are infinitely many ways to complete a basis of W to a basis of V, it is also clear that, unless $W = \{0\}$ or $W = V$, the choice of a complementary subspace to W is not unique.

1.2 Linear Maps

Let V and W be vector spaces over $\mathbb{F}$. The set of all linear maps from V to W will be denoted by $\mathcal{L}(V, W)$. $\mathcal{L}(V, W)$ has a natural vector space structure given by addition and scalar multiplication of linear maps: if S and T are in $\mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$, then the linear maps $S + T$ and λS in $\mathcal{L}(V, W)$ are given by

$$\begin{aligned}(S+T)(v) &= S(v) + T(v) \\ (\lambda S)(v) &= \lambda\, S(v) \qquad\qquad \text{for all } v \in V\end{aligned}$$

It is not hard to prove that with these operations, $\mathcal{L}(V, W)$ is a vector space over $\mathbb{F}$.

Fix a basis $(v_1, \ldots, v_n)$ of V. Then any $T \in \mathcal{L}(V, W)$ is completely determined by its effect on the basis vectors v_j. For if $v \in V$, then we can write $v = c_1v_1 + \cdots + c_nv_n$ for scalars $c_1, \ldots, c_n$, whence

$$T(v) = c_1T(v_1) + \cdots + c_nT(v_n). \tag{1.3}$$

Conversely, given any vectors $w_1, \ldots, w_n$ in W, there is a unique linear map $T \in \mathcal{L}(V, W)$ such that $T(v_1) = w_1, \ldots, T(v_n) = w_n$. This is because any vector

$v \in V$ can be written uniquely as $v = c_1 v_1 + \cdots c_n v_n$; if we define the map $T : V \to W$ by (1.3), then it is easy to see that $T \in \mathcal{L}(V, W)$.

Abstract linear algebra is inextricably bound to matrix theory since any linear map may be represented by an appropriate matrix, and since the algebra of linear maps corresponds to the algebra of matrices.

More precisely, let us fix bases $B = (v_1, \ldots, v_n)$ and $B' = (w_1, \ldots, w_m)$ of vector spaces V and W, respectively. Recall that any $T \in \mathcal{L}(V, W)$ is uniquely determined by the basis images $T(v_1), \ldots, T(v_n)$ in W. Each of these vectors $T(v_j)$ is a unique linear combination of $w_1, \ldots, w_m$:

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \qquad (j = 1, \ldots, n) \tag{1.4}$$

We define the *matrix* $M_{B,B'}(T)$ *of* T *with respect to these bases* to be the $m \times n$ matrix whose (i, j)-entry is a_{ij};

$$M_{B',B}(T) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \tag{1.5}$$

We will frequently denote this matrix by $M(T)$ if the bases B and B' are clear from the context of the discussion.

Let $T \in \mathcal{L}(V, W)$. The *kernel*, or *nullspace*, of T is the subspace of V given by

$$\ker T = \{v \in V \mid T(v) = 0\}$$

From linear algebra, we know that the linear map T is *injective*, or one-to-one, if and only if $\ker T = \{0\}$; in this case we say that T is a *linear isomorphism of* V *into* W.

The *range* of T is the subspace of W given by

$$T(V) = \{T(v) \mid v \in V\}.$$

We recall the definition that T is *surjective*, or *onto*, if $T(V) = W$.

The following is an easy to prove, yet important fact in linear algebra:

Theorem 1.2.1. *For any* $T \in \mathcal{L}(V, W)$, *we have*

$$\dim T(V) = \dim V - \dim(\ker T)$$

(See [1], Theorem 3.4, or better yet, prove it yourself!) In particular, if T is a linear isomorphism from V onto W, then $\dim V = \dim W$.

Making an abrupt and unforgivable change of notation for the moment, suppose that W is a subspace of a vector space V. The *quotient map* π from V onto the quotient space V/W is given by $\pi(v) = v + W$. It is obvious that π is linear and surjective, with kernel W. Using Theorem 1.2.1, this provides a completely trivial proof of Proposition 1.1.1.

1.3 The Matrix of a Linear Map

Again let us fix bases $B = \{v_1, \dots, v_n\}$ and $B' = \{w_1, \dots, w_m\}$ of V and W, respectively. From (1.4) and (1.5) we see that each $T \in \mathcal{L}(V,W)$ corresponds to a unique $m \times n$ matrix $M_{B',B}(T)$. The map $T \mapsto M_{B',B}(T)$ is from $\mathcal{L}(V,W)$ to the vector space $M_{m,n}(\mathbb{F})$ of $m \times n$ matrices with entries in $\mathbb{F}$ is easily checked to be linear (and onto), and hence is a linear isomorphism. Since $\dim M_{m,n} = mn$, we see that $\dim \mathcal{L}(V,W) = mn = nm = \dim V \cdot \dim W$.

Another useful property of the map $T \mapsto M_{B',B}(T)$ is that it is multiplicative. More precisely, suppose that V, W, and U are vector spaces with fixed bases B, B', and B'', respectively, and suppose that $T \in \mathcal{L}(V,W)$ and $S \in \mathcal{L}(W,U)$. Then the composite map $ST := S \circ T$ belongs to $\mathcal{L}(V,U)$, and we have

$$M_{B'',B}(ST) = M_{B'',B'}(S)\, M_{B',B}(T), \tag{1.6}$$

where the right hand side is a matrix product.

Exercise 1.3.1. Prove equation (1.6).

For simplicity, we'll denote the space of linear maps $\mathcal{L}(V,V)$ simply by $\mathcal{L}(V)$. An element of $\mathcal{L}(V)$ is called a *linear operator* on V.

Theorem 1.3.2. *Fix a basis $B = (v_1, \dots, v_n)$ of V. Suppose that $T \in \mathcal{L}(V)$. Then the following are equivalent:*

1. *T is one-to-one*
2. *T is onto*
3. *The matrix $M(T) := M_{B,B}(T)$ with respect to the basis B is nonsingular*

Proof. (1) $\iff$ (2): $\dim T(V) = \dim V - \dim(\ker T)$ so $\dim T(V) = \dim V$ if and only if $\dim(\ker T) = 0$.
(1) and (2) $\iff$ (3): If T is one-to-one and onto, it is invertible; that is, there is a unique linear map $S \in \mathcal{L}(V)$ such that $ST = TS = 1_V$, the identity map on V. If I_n is the identity $n \times n$ matrix, we see from (1.6) that

$$\begin{aligned} I_n &= M(ST) = M(S)M(T) \\ &= M(TS) = M(T)M(S), \end{aligned}$$

which shows that $M(T)$ is invertible. Conversely, assume $M(T)$ is an invertible $n \times n$ matrix; let S be the linear operator on V whose matrix is $M(T)^{-1}$. Then, again by (1.6),

$$\begin{aligned} I_n &= M(T)^{-1}M(T) = M(S)M(T) = M(ST) \\ &= M(T)M(T)^{-1} = M(T)M(S) = M(TS), \end{aligned}$$

which shows that $ST = TS = 1_V$. $\square$

If P and Q are $n \times n$ matrices with P nonsingular, the *conjugate* of Q by P is the $n \times n$ matrix PQP^{-1}. Suppose now that $B = (v_1, \dots, v_n)$ and $B' = (v_1', \dots, v_n')$ are two bases of the same vector space V. The *change of basis matrix* from B to B' is $M_{B',B}(1_V)$: it gives us the coefficients in the linear combination expressing each v_j as a linear combination of the v_i''s. From (1.6), this change of basis matrix is nonsingular, with inverse $M_{B,B'}(1_V)$. For simplicity, let us denote this change of basis matrix by S. Let $T \in \mathcal{L}(V)$. If $M(T) := M_{B,B}(T)$ is the matrix of T with respect to the basis B, then its matrix with respect to B' is given by conjugating $M(T)$ by A. Explicitly, by (1.6)

$$\begin{aligned} M_{B',B'}(T) &= M_{B',B}(1_V) M_{B,B}(T) M_{B,B'}(1_V) \\ &= S M(T) S^{-1}. \end{aligned} \tag{1.7}$$

Example 1.3.3. Suppose that T is the linear operator on $\mathbb{R}^3$ whose matrix with respect to the standard basis $B_0 = (e_1, e_2, e_3)$ of $\mathbb{R}^3$ is

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 3 & 1 \\ 0 & 2 & 0 \end{pmatrix}$$

Consider the vectors

$$v_1 = \begin{pmatrix} 4 \\ 8 \\ 7 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} -2 \\ -3 \\ -2 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 3 \\ 5 \\ 4 \end{pmatrix}$$

The 3×3 matrix S whose columns are v_1, v_2, v_3 is invertible; its inverse can be calculated using the Gauss-Jordan method and is found to be

$$\begin{pmatrix} 4 & -2 & 3 \\ 8 & -3 & 5 \\ 7 & -2 & 4 \end{pmatrix}^{-1} = \begin{pmatrix} -2 & 2 & -1 \\ 3 & -5 & 4 \\ 5 & -6 & 4 \end{pmatrix}$$

We therefore see that $B = (v_1, v_2, v_3)$ is a linearly independent set which thus forms a basis of $\mathbb{R}^3$, and the change of basis matrix from the standard basis B_0 to B is given by S. Hence the matrix of T with respect to the basis B is

$$\begin{aligned} SAS^{-1} &= \begin{pmatrix} 4 & -2 & 3 \\ 8 & -3 & 5 \\ 7 & -2 & 4 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 3 & 1 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} -2 & 2 & -1 \\ 3 & -5 & 4 \\ 5 & -6 & 4 \end{pmatrix} \\ &= \begin{pmatrix} -22 & 28 & -18 \\ -74 & 91 & -59 \\ -57 & 69 & -44 \end{pmatrix} \end{aligned}$$

The *transpose* of an $m \times n$ matrix $A = a_{ij}$ is the $n \times m$ matrix ${}^tA$ whose (i, j) entry is a_{ji}. Thus rows of A transform into the columns of ${}^tA$, and the columns of A transform into the rows of ${}^tA$. It's not hard to show that if A and B are

$m \times n$ matrices and α is a scalar, then ${}^t(A+B) = {}^tA + {}^tB$, ${}^t(\alpha A) = \alpha\, {}^tA$. A somewhat longer but completely straightforward calculation shows that if A is an $m \times n$ matrix and B is an $n \times k$ matrix, then ${}^t(AB) = {}^tB\, {}^tA$.

The *dual space* of a vector space V is $\mathcal{L}(V, \mathbb{F})$ (where $\mathbb{F}$ is viewed as a one-dimensional vector space), and is denoted V^*. Its elements are called *linear functionals* on V. Any basis $(v_1, \ldots, v_n)$ of V gives rise to a *dual basis* $(f_1, \ldots, f_n)$ of V^* where each f_i is given by

$$f_i(v_j) = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Let V and W be vector spaces over $\mathbb{F}$, and let $: V \to W$ be a linear map. The *transpose* of T is the map ${}^tT : W^* \to V^*$ given by ${}^tT(\lambda) = \lambda \circ T$, for all $\lambda \in W^*$. It is not hard to show that ${}^tT$ is a linear map from W^* to V^*. Suppose that $B = (v_1, \ldots, v_n)$ and $B' = (w_1, \ldots, w_m)$ are bases of V and W, respectively. Let $B^* = (f_1, \ldots, f_n)$ and $(B')^* = (h_1, \ldots, h_n)$ be the corresponding dual bases of V^* and W^*, respectively. We have the easily verified relation

$$M_{(B')^*, B^*}({}^tT) = {}^t(M_{B,B'}(T))$$

where the right hand side denotes the transpose of the matrix $M_{B.B'}(T)$.

Exercise 1.3.4. 1. Prove that T is injective iff ${}^tT$ is surjective.

2. Using Part (a), prove that the ranges of T and ${}^tT$ have the same dimension.

3. Use Part (b) to prove that the row space and the column space of an $m \times n$ matrix A over $\mathbb{F}$ have the same dimension. (This dimension is called the *rank* of A. The dimension of the range of a linear mapping T is called the *rank* of T.)

1.4 Determinant and Trace

The *determinant* of an $n \times n$ matrix A is defined in various ways. (Most definitions of the determinant in standard linear algebra texts are non-intuitive. Axler's book [1] develops all of linear algebra without resorting to the determinant until the very end, where it "comes naturally." For our purposes, since we're after different game, it's sufficient to provide two of the equivalent expressions of the determinant and state its most salient features.)

If $A = (a_{ij})$, let us recall that its determinant $\det A$ is the homogeneous degree n polynomial in the entries of A given by

$$\det A = \sum_{\sigma} \epsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

where the sum runs through all the permutations σ of $\{1, \ldots, n\}$, and $\epsilon(\sigma)$ denotes the sign of the permutation σ.

Let's also recall that the determinant $\det A$ can also be expanded using minors along a given row or column, as follows. For each pair of indices i, j in $\{1, \ldots, n\}$, let A_{ij} denote the $(n-1) \times (n-1)$ submatrix obtained from A by deleting the ith row and jth column. Then the *minor expansion* of $\det A$ along the ith row is

$$\det A = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \, \det A_{ik}$$

and that along the jth column is

$$\det A = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \, \det A_{kj}.$$

If A has real entries, $\det A$ has a geometrical significance. Let $v_1, \ldots, v_n$ denote the columns of A. These are vectors in $\mathbb{R}^n$, and $|\det A|$ turns out to be the n-dimensional volume of the parallelepiped whose sides are $v_1, \ldots, v_n$. This can be proved by induction on n, and can be seen at least for 3×3 matrices A, since $\det A$ is the triple scalar product $(v_1 \times v_2) \cdot v_3$. (The proof for 2×2 determinants is even easier and just uses cross products.)

Exercise 1.4.1 (**Graduate Exercise).** Prove this geometrical fact about the n-dimensional determinant.

The determinant is multiplicative in that, if A and B are square matrices of the same size, then $\det AB = \det A \cdot \det B$. We also recall that a square matrix A is nonsingular if and only if $\det A \neq 0$. It then follows by multiplicativity that $\det A^{-1} = (\det A)^{-1}$.

Let V be a vector space over $\mathbb{F}$ and suppose that $T \in \mathcal{L}(V)$. We define the *determinant of* T to be $\det M(T)$, where $M(T)$ is the matrix of T with respect to any basis B of V. The value of the determinant $\det T$ is independent of the choice of basis: if B' is any other basis of V and S is the change of basis matrix from B to B', then by (1.7), the matrix of T with respect to B' is $S\,M(T)\,S^{-1}$, and hence

$$\begin{aligned} \det(S\,M(T)\,S^{-1}) &= \det S \, \det M(T) \, \det S^{-1} \\ &= \det S \, \det M(T) \, (\det S)^{-1} \\ &= \det M(T) \end{aligned}$$

Theorem 1.4.2. *Let $T \in \mathcal{L}(V)$. Then T is invertible if and only if $\det T \neq 0$.*

Proof. By Theorem 1.3.2, T is invertible $\iff$ $M(T)$ is nonsingular $\iff$ $\det M(T) \neq 0 \iff \det T \neq 0$. $\square$

Another useful quantity associated to a linear operator $T \in \mathcal{L}(V)$ is its *trace*. If $A = (a_{ij})$ is an square $n \times n$ matrix, the *trace* of A is defined to be the sum of its diagonal entries: $\operatorname{tr} A = \sum_{i=1}^n a_{ii}$. The trace satisfies the following easily verified property:

Proposition 1.4.3. *Let A and B be $n \times n$ matrices. Then $\operatorname{tr} AB = \operatorname{tr} BA$.*

Proof. Let $A = (a_{ij}$ and $B = (b_{ij})$. Then from the definition of matrix product, $AB = (c_{ij})$, where

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}.$$

Likewise, $BA = (d_{ij})$, with

$$d_{ij} = \sum_{k=1}^n b_{ik} a_{kj}.$$

Hence

$$\operatorname{tr} AB = \sum_{i=1}^n c_{ii} = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki} \tag{1.8}$$

whereas

$$\operatorname{tr}(BA) = \sum_{i=1}^n d_{ii} = \sum_{i=1}^n \sum_{k=1}^n b_{ik} a_{ki}$$

If we interchange the indices i and k in the above sum, we see that it equals the sum in (1.8). $\square$

The *trace* of a linear operator $T \in \mathcal{L}(V)$ is, by definition, the trace of the matrix $M(T)$, where $M(T)$ is the matrix of T with respect to any basis of V. Now the matrix of T with respect to any other basis of V is given by $S\,M(T)\,S^{-1}$ for some matrix S, and it follows from Proposition 1.4.3 above that $\operatorname{tr}(S\,M(T)\,S^{-1}) = \operatorname{tr}(M(T)\,S^{-1}\,S) = \operatorname{tr} M(T)$. Thus the trace of T is a well-defined scalar, depending only on T and not on the choice of basis of V.

1.5 Eigenvalues and Invariant Subspaces

A scalar $\lambda \in \mathbb{F}$ is called an *eigenvalue* of a linear operator $T \in \mathcal{L}(V)$ if there is a *nonzero* vector $v \in V$ such that $T(v) = \lambda v$. The vector v is called an *eigenvector* of T corresponding to λ. If $\lambda \in \mathbb{C}$, the subspace $\ker(T - \lambda I_V) = \{v \in V \,|\, T(v) = \lambda v\}$ is called the *eigenspace of T corresponding to λ*.

Proposition 1.5.1. *Suppose that $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$. Then the following are equivalent:*

1. *λ is an eigenvalue of T*

2. $\ker(T - \lambda I_V) \neq \{0\}$

3. $\det(T - \lambda I_V) = 0$.

Proof. An easy exercise. $\square$

The polynomial $\det(\lambda I_V - T)$ in the indeterminate λ, with coefficients in $\mathbb{F}$, is called the *characteristic polynomial* of T; its roots in $\mathbb{F}$ are precisely the eigenvalues of T.

Linear operators on real vector spaces do not necessarily have real eigenvalues. For instance, consider the operator $T \in \mathcal{L}(\mathbb{R}^2)$ given by $Tx = Ax$ ($x \in \mathbb{R}^2$), where

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

Then the characteristic polynomial of T is $\det(\lambda I_2 - T) = \lambda^2 + 1$, which has no real roots. Thus T has no real eigenvalues.

On the other hand, a linear operator on a complex vector space has at least one eigenvalue.

Theorem 1.5.2. *Let V be a nonzero vector space over $\mathbb{C}$ and let $T \in \mathcal{L}(V)$. Then T has at least one eigenvalue.*

Proof. The characteristic polynomial $\det(\lambda I_V - T)$ is a polynomial in λ of degree $\dim V > 0$, and so by Gauss's Fundamental Theorem of Algebra, has at least one complex root, which is, by Proposition 1.5.1, an eigenvalue of T. $\square$

An easy consequence of the Fundamental Theorem of algebra is that any polynomial $p(z)$ of degree n has n complex roots, counting multiplicities, and has a unique linear factorization $p(z) = c \cdot \prod_{j=1}^{k} (z - \lambda_j)^{m_j}$, where $\lambda_1, \ldots, \lambda_k$ are the distinct roots of $p(z)$ and $m_1, \ldots, m_k$ are their respective multiplicities, with $m_1 + \cdots + m_k = n$. Applying this to the characteristic polynomial of T, we get $\det(\lambda I_V - T) = \prod_{j=1}^{k} (\lambda - \lambda_j)^{m_j}$. Here $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T and $m_1, \ldots, m_k$ are called their respective *multiplicities*.

It is often useful to study an operator $T \in \mathcal{L}(V)$ by examining its invariant subspaces. A subspace W is said to be *invariant under* T, or *T-invariant* if $T(w) \in W$ for all $w \in W$. Thus the restriction $T|_W$ of the map T to W belongs to $\mathcal{L}(W)$. If v is an eigenvector of T, then the one-dimensional subspace $\mathbb{F}v$ is obviously a T-invariant subspace.

Exercise 1.5.3. If two operators S and T in $\mathcal{L}(V)$ commute, then show that both the kernel $\ker S$ and the range $S(V)$ are T-invariant.

Since T commutes with $T - \lambda I_V$ for all $\lambda \in \mathbb{F}$, Exercise 1.5.3 implies, in particular, that each eigenspace $\ker(T - \lambda I_V)$ is T-invariant.

Suppose that W is a T-invariant subspace of V. Then T induces a well-defined map T' on the quotient space V/W given by $T'(v + W) = T(v) + W$. It is easy to check that $T' \in \mathcal{L}(V/W)$. We have the commutative diagram

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\pi"'] & V \arrow[d, "\pi"] \\ V/W \arrow[r, "T'"'] & V/W \end{tikzcd} \tag{1.9}$$

which says that $T'\pi = \pi T$ (as maps from V to V/W).

1.6 Upper Triangular Matrices

We would like to find a basis of V with respect to which the matrix of a given operator T is "nice," in some sense. Ideally, we want the matrix of T to be diagonal, if this is possible, or at least to have as many 0's as possible, arranged in an orderly fashion.

A square matrix is called *upper triangular* if all the entries below the main diagonal are 0. Such a matrix can then be represented as follows:

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Note that the determinant of any upper triangular matrix is the product of the diagonal entries. Thus, if A is the matrix above, then $\det A = \lambda_1 \cdots \lambda_n$.

Proposition 1.6.1. *Suppose that $T \in \mathcal{L}(V)$. Then the following are equivalent:*

1. *There is a basis of V for which the matrix of T is upper triangular*
2. *There is a nested sequence of subspaces $0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$ each of which is invariant under T.*

The proof is obvious. Note that the condition (2) implies that the subspace V_j must have dimension j.

It is a very useful fact that if V is a complex vector space, any linear map $T \in \mathcal{L}(V)$ has an upper triangular representation:

Theorem 1.6.2. *Let V be a vector space over $\mathbb{C}$ and let $T \in \mathcal{L}(V)$. Then there is a basis of V for which the matrix of T is upper triangular.*

Proof. By induction on $\dim V$. If $\dim V = 1$, there is nothing to prove. So let's assume that $\dim V = n > 1$. Then by Theorem 1.5.2, T has an eigenvalue λ. Let v_1 be an eigenvector corresponding to λ, and let W be the one-dimensional subspace $\mathbb{C}v_1$. Since W is T-invariant, we can consider the induced linear map T' on the complex vector space V/W given by $T'(v + W) = T(v) + W$. Now $\dim(V/W) = n - 1$, so by the induction hypothesis, there is a basis $(v_2 + W, \ldots, v_n + W)$ of V/W for which the matrix of T' is upper triangular. Note that for each j, $j = 2, \ldots, n$, $T'(v_j + W)$ is a linear combination of $v_2 + W, \ldots, v_j + W$; hence $T(v_j)$ is a linear combination of $v_1, v_2, \ldots, v_j$.

It remains to prove that $(v_1, v_2, \ldots, v_n)$ is a linearly independent set (and so is a basis of V). Once we prove this, it is clear that the matrix of T with respect to this basis is upper triangular.

Suppose that $\sum_{i=1}^{n} c_j v_j = 0$. Then $(\sum_{i=1}^{n} c_i v_i) + W = W \implies \sum_{i=2}^{n} (c_i v_i + W) = W \implies c_2 = \cdots = c_n = 0$, by the linear independence of $(v_2 + W, \ldots, v_n + W)$, so we end up with $c_1 v_1 = 0$, which clearly implies that $c_1 = 0$. $\square$

Note: Suppose that $n \geq 2$. Since the choice of the v_j's in the proof is not necessarily unique, the upper triangular matrix in the theorem above is not necessarily unique. What is unique, from the characteristic polynomial of T, are the diagonal entries and their multiplicities.

Let $p(z) = a_m z^m + a_{m-1} z^{m-1} + \cdots + a_1 z + a_0$ be any polynomial in the variable z. If $T \in \mathcal{L}(V)$, we put $p(T) = a_m T^m + a_{m-1} T^{m-1} + \cdots + a_1 T + a_0 I_V$. Then $p(T) \in \mathcal{L}(V)$, and if M is the matrix of T with respect to some basis of V, then $p(M) = a_m M^m + a_{m-1} M^{m-1} + \cdots + a_1 M + a_0 I_n$ is the matrix of $p(T)$ with respect to this basis.

Theorem 1.6.3. *(The Cayley-Hamilton Theorem) Suppose that V is a vector space over $\mathbb{C}$ and that $T \in \mathcal{L}(V)$. Let $p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1 \lambda + a_0$ be the characteristic polynomial of T. Then the linear map $p(T)$ is identically zero on V.*

Proof. By induction on $\dim V$. If $\dim V = 1$, then the conclusion is obvious: $T = \lambda_1 I_V$ for some $\lambda_1 \in \mathbb{C}$, the characteristic polynomial of T is $p(\lambda) = \lambda - \lambda_1$, and $p(T) = T - \lambda_1 I_V \equiv 0$.

So assume that $n > 1$ and that the theorem holds for all linear maps on all vector spaces of dimension $< n$. Suppose that $\dim V = n$ and that $T \in \mathcal{L}(V)$.

Choose a basis $(v_1, \ldots, v_n)$ of V for which T has upper triangular matrix

$$\begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \tag{1.10}$$

Then the characteristic polynomial $\det(\lambda I_V - T) = \prod_{j=1}^n (\lambda - \lambda_j)$, and the diagonal entries $\lambda_1, \ldots, \lambda_n$ above are the eigenvalues of T. (Of course, these λ_j are not necessarily distinct.) Let $V_1 = \mathbb{C}v_1$. Now V_1 is a T-invariant subspace, the quotient space $V' = V/V_1$ is easily seen to have basis $v_2 + V_1, \ldots, v_n + V_1$, and the matrix of the induced map $T' : V/V_1 \to V/V_1,\ v + V_1 \mapsto T(v) + V_1$ with respect to this basis is

$$\begin{pmatrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \tag{1.11}$$

The characteristic polynomial of T' is thus $(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$. Since V' has dimension $n - 1$, the induction hypothesis implies that $(T' - \lambda_2 I_{V'}) \cdots (T' - \lambda_n I_{V'}) \equiv 0$ on V'.

Thus for any $v \in V$, we have

$$\begin{aligned} (T' - \lambda_2 I_{V'}) \cdots (T' - \lambda_n I_{V'})(v + V_1) &= (T - \lambda_2 I_V) \cdots (T - \lambda_n I_V)(v) + V_1 \\ &= V_1 \end{aligned}$$

and so

$$(T - \lambda_2 I_V) \cdots (T - \lambda_n I_V)(v) \in V_1$$

Therefore $(T - \lambda_2 I_V) \cdots (T - \lambda_n I_V)(v) = cv_1$ for some $c \in \mathbb{C}$. Hence

$$\begin{aligned} (T - \lambda_1 I_V)(T - \lambda_2 I_V) \cdots (T - \lambda_n I_V)(v) &= (T - \lambda_1 I_V)(cv_1) \\ &= 0. \end{aligned}$$

Since $v \in V$ is arbitrary, we have shown that $(T - \lambda_1 I_V) \cdots (T - \lambda_n I_V) \equiv 0$ on V, proving the conclusion for V and completing the induction step. □

Note: The Cayley-Hamilton Theorem also holds for linear operators on real vector spaces. In order to see this, we note that it suffices to prove the Cayley-Hamilton Theorem for real square matrices, due to the correspondence between linear maps on V and square matrices of size $\dim V$. But then any real square matrix can be considered to be a complex matrix, which by the Cayley-Hamilton Theorem, satisfies its characteristic polynomial.

Exercise 1.6.4. Suppose that $T \in \mathcal{L}(V)$ is invertible. Show that there exists a polynomial $p(z)$ such that $T^{-1} = p(T)$.

Exercise 1.6.5. Suppose that V is a n-dimensional complex vector space, and that $T \in \mathcal{L}(V)$ has spectrum $\{4, 5\}$. Prove that

$$(T - 4I_V)^{n-1}(T - 5I_V)^{n-1} = 0.$$

1.7 Generalized Eigenspaces

The *spectrum* of a linear operator T on a vector space V over $\mathbb{F}$ is the collection of all eigenvalues of T in $\mathbb{F}$. We saw, from Theorem 1.5.2, that if V is complex, then any $T \in \mathcal{L}(V)$ has a nonempty spectrum.

In order to derive further nice properties about linear maps, it is useful to have at least one eigenvalue, so in this section, we'll assume that V is a vector space over $\mathbb{C}$ and that $T \in \mathcal{L}(V)$.

Choose a basis of V for which T has upper triangular matrix

$$\begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \tag{1.12}$$

As we had already observed, the characteristic polynomial $\det(\lambda I_V - T)$ equals $\prod_{j=1}^n (\lambda - \lambda_j)$, and the diagonal entries $\lambda_1, \ldots, \lambda_n$ above are the eigenvalues of T, which are not necessarily distinct. Note that the number of times each distinct eigenvalue λ_j appears in the diagonal of the matrix above equals its multiplicity m_j as a root of the characteristic polynomial of T.

As mentioned previously, the upper triangular representation (1.12) of T is not unique, except for the appearance of the eigenvalues (with the correct multiplicities) along the diagonal. Our goal is to obtain a particular upper triangular representation of T which is unique and useful in the sense that much of the behavior of T is apparent upon cursory examination of the matrix.

With this goal in mind, we define the generalized eigenspaces of T as follows. For any complex scalar λ, the *generalized eigenspace* of T corresponding to λ is the set

$$\{v \in V \mid (T - \lambda I_V)^k v = 0 \text{ for some } k \in \mathbb{Z}^+\} \tag{1.13}$$

Note that the eigenspace $\ker(T - \lambda I_V)$ is a subset of the generalized eigenspace (1.13) above. The following result shows that the generalized eigenspace is a subspace of V.

Theorem 1.7.1. *Fix $\lambda \in \mathbb{C}$. Then there is an integer $m \geq 0$ such that*

$$\begin{aligned}\{0\} \subsetneq \ker(T - \lambda I_V) \subsetneq \ker(T - \lambda I_V)^2 \subsetneq \cdots \subsetneq \ker(T - \lambda I_V)^m &= \ker(T - \lambda I_V)^{m+1} \\ &= \ker(T - \lambda I_V)^{m+2} = \cdots\end{aligned}$$

Proof. For simplicity, let $S = T - \lambda I_V$. If $v \in \ker S^k$, then $S^{k+1}(v) = S(S^k(v)) = 0$, so $v \in \ker S^{k+1}$, so it follows that $\ker S^k \subset \ker S^{k+1}$, and we get a nested chain of subspaces $\{0\} \subset \ker S \subset \ker S^2 \subset \cdots$

Since V is finite-dimensional, the chain stops increasing after some point. Let m be the smallest nonnegative integer such that $\ker S^m = \ker S^{m+1}$. Then $\ker S^{m+1} = \ker S^{m+2}$: if $v \in \ker S^{m+2}$, then $S(v) \in \ker S^{m+1} = \ker S^m$, so $S^m(Sv) = 0$, and so $v \in \ker S^{m+1}$.

Arguing in the same manner, we see that $\ker S^{m+2} = \ker S^{m+3}$, etc. $\square$

Note that the m in Theorem 1.7.1 must be $\leq \dim V$.

It follows that the generalized eigenspace of T corresponding to λ equals the kernel $\ker(T - \lambda I_V)^{\dim V}$, which is a subspace of V.

Corollary 1.7.2. *Let $S \in \mathcal{L}(V)$. Then there is an integer m such that*

$$V \supsetneq S(V) \supsetneq S^(V) \supsetneq \cdots \supsetneq S^m(V) = S^{m+1}(V) = S^{m+2}(V) = \cdots \tag{1.14}$$

Proof. This follows immediately from the proof of the preceding theorem, once we observe that $V \supset S(V) \supset S^2(V) \supset \cdots$, and that $\dim S^k(V) = \dim V - \dim\ker(S^k)$. $\square$

Exercise 1.7.3. Show that for any $T \in \mathcal{L}(V)$, we have

$$V = (\ker T^n) \oplus T^n(V),$$

where $n = \dim V$.

As an additional application of the upper triangular representation (1.12) of T, we can determine the dimension of each generalized eigenspace.

Proposition 1.7.4. *Suppose that $T \in \mathcal{L}(V)$ has characteristic polynomial given by $\det(\lambda I_V - T) = \prod_{j=1}^n (\lambda - \lambda_j)$ ($\lambda_1, \ldots, \lambda_n$ not necessarily distinct). Then the generalized eigenspace of T corresponding to λ_j has dimension equal to the multiplicity of λ_j.*

Proof. Note that m_j is the number of times λ_j appears in the upper triangular representation of T.

We prove this by induction on $\dim V$. If $\dim V = 1$, then the conclusion is trivial. Let $n \geq 2$ and assume that the conclusion holds for all linear maps on all complex vector spaces of dimension $< n$. Suppose that $\dim V = n$ and that $T \in \mathcal{L}(V)$. For each j, we let V_j denote the generalized eigenspace of T corresponding to λ_j.

Choose a basis $(v_1, \ldots, v_n)$ of V for which the matrix of T is upper triangular, of the form (1.10). Let $W = \mathbb{C}v_1$, and let T' be the induced linear map on $V' = V/W$: $T'(v + W) = T(v) + W$. Then, with respect to the basis $(v_2 + W, \ldots, v_n + W)$ of V/W, the matrix of T' is upper triangular and is given by (1.11).

Now the induction hypothesis says that the generalized eigenspace V'_j of T' corresponding to λ_j has dimension m_j if $\lambda_j \neq \lambda_1$ and has dimension $m_1 - 1$ if $\lambda_j = \lambda_1$.

We therefore consider the two cases $\lambda_j \neq \lambda_1$ and $\lambda_j = \lambda_1$ separately.

Let $\pi : V \to V/W$ be the quotient map. We first note that, for any j, $\pi(V_j) \subset V'_j$. Indeed, for each $v \in V_j$, we have $(T - \lambda_j I_V)^N v = 0$, for sufficiently large N, so $0 = \pi\left((T - \lambda_j I_V)^N v\right) = (T' - \lambda_j I_{V'})^N (\pi(v))$, whence $\pi(v) \in V'_j$. Thus π maps V_j into V'_j.

Now let us first consider the case $\lambda_j \neq \lambda_1$. We claim that π maps V_j isomorphically onto V'_j.

Suppose that $v \in V_j$ belongs to $\ker \pi$. Then $v = cv_1$ for some $c \in \mathbb{C}$. The condition $(T - \lambda_j I_V)^N v = 0$ (for some N) then implies that $c(\lambda_1 - \lambda_j)^N v_1 = 0$, so that $c = 0$, and so $v = 0$. Hence π maps V_j injectively into V'_j.

Next let us show that π maps V_j onto V'_j. Let $v' \in V'_j$, and write $v' = \pi(v)$ for some $v \in V$. By assumption $(T' - \lambda_j I_{V'})^N (\pi(v)) = 0$ for some N. This yields $\pi\left((T - \lambda_j I_V)^N v\right) = 0$, so $(T - \lambda_j I_V)^N v = cv_1$ for some $c \in \mathbb{C}$. Now let $u = v - c(\lambda_1 - \lambda_j)^{-N} v_1$. Then $(T - \lambda_j I_V)^N u = 0$, so $u \in V_j$, and clearly, $\pi(u) = \pi(v) = v'$. Thus $\pi(V_j) = V'_j$, and so we can conclude that $\dim V_j = m_j$.

Next we consider the case $\lambda_j = \lambda_1$. We claim that π maps V_1 onto V'_1 with kernel $\mathbb{C}v_1$. Since $\mathbb{C}v_1 = \ker \pi = V_1 \cap \ker \pi$, it suffices to prove that $\pi(V_1) = V'_1$. Let $v' \in V'_1$. We have $v' = \pi(v)$ for some $v \in V$. Now the condition $(T' - \lambda_1 I_{V'})^N v' = 0$ for some N implies that $(T - \lambda_1 I_V)^N v = av_1$ for some $a \in \mathbb{C}$. Hence $(T - \lambda_1 I_V)^{N+1} v = 0$, and thus $v \in V_1$. This shows that $v' \in \pi(V_1)$, and so $\pi(V_1) = V'_1$.

We conclude in this case that $\dim V_1 = \dim V'_1 + 1 = m_1$. This completes the proof of Proposition 1.7.4. $\square$

Exercise 1.7.5. Suppose that V is a complex vector space of dimension n and $T \in \mathcal{L}(V)$ such that

$$\ker T^{n-2} \subsetneq \ker T^{n-1}.$$

Prove that T has at most two distinct eigenvalues.

We now reorder the eigenvalues of T, if necessary, so that we now assume that the *distinct* eigenvalues of T are $\lambda_1, \dots, \lambda_k$. Let us again consider the characteristic polynomial

$$p(\lambda) = \det(\lambda I_V - T) \tag{1.15}$$

and its factorization

$$\prod_{j=1}^{k} (\lambda - \lambda_j)^{m_j} \tag{1.16}$$

Our objective now is to show that V is the direct sum of the generalized eigenspaces corresponding to each λ_j.

Lemma 1.7.6. *Let $p_1(z), \dots, p_k(z)$ be nonzero polynomials with coefficients in $\mathbb{C}$ sharing no common factor of degree ≥ 1. Then there are polynomials $q_1(z), \dots, q_k(z)$ such that $p_1(z)q_1(z) + \cdots + p_k(z)q_k(z) = 1$.*

Proof. (Optional) Let $\mathbb{C}[z]$ be the ring of polynomials in z with complex coefficients. Since $\mathbb{C}[z]$ is a Euclidean ring, it is a principal ideal domain, and so the ideal $\mathbb{C}[z]p_1(z) + \cdots + \mathbb{C}[z]p_k(z)$ is principal: $\mathbb{C}[z]p_1(z) + \cdots + \mathbb{C}[z]p_k(z) = \mathbb{C}[z]r(z)$, for some nonzero polynomial $r(z)$. Clearly, $r(z)$ divides all the $p_j(z)$, so $r(z)$ must be a degree 0 polynomial; i.e., a nonzero constant. Thus $\mathbb{C}[z]p_1(z) + \cdots + \mathbb{C}[z]p_k(z) = \mathbb{C}[z]$; in particular $1 \in \mathbb{C}[z]p_1(z) + \cdots + \mathbb{C}[z]p_k(z)$. $\square$

Theorem 1.7.7. *V is a direct sum of the generalized eigenspaces of the eigenvalues of T. More precisely,*

$$V = \bigoplus_{j=1}^{k} \ker(T - \lambda_j I_V)^{m_j}$$

Moreover, for each j, we have $\dim(\ker(T - \lambda_j I_V)^{m_j}) = m_j$. Thus the dimension of the generalized eigenspace corresponding to λ_j equals the number of times λ_j appears in any upper triangular matrix representation of T.

Proof. We let $p(\lambda)$ be the characteristic polynomial (1.15) of T, factored as in (1.16).

Suppose first that T has just one eigenvalue λ_1. Then $p(\lambda) = (\lambda - \lambda_1)^n$, and by the Cayley-Hamilton Theorem, $(T - \lambda_1 I_V)^n = 0$, so $V = \ker(T - \lambda_1 I_V)^n$, proving the theorem.

Thus we can assume that T has more than one eigenvalue.

For each j, let

$$p_j(\lambda) = \frac{p(\lambda)}{(\lambda - \lambda_j)^{m_j}} = \prod_{l \neq j} (\lambda - \lambda_l)^{m_l}.$$

Then by Lemma 1.7.6, there exist complex polynomials $q_1(\lambda), \dots, q_k(\lambda)$ such that $p_1(\lambda)q_1(\lambda) + \cdots + p_k(\lambda)q_k(\lambda) = 1$. Replacing λ by T, we have $p_1(T)q_1(T) + \cdots + p_k(T)q_k(T) = I_V$. (Strictly speaking, we're applying the well-defined algebra homomorphism $p(\lambda) \mapsto p(T)$ from $\mathbb{C}(\lambda)$ to $\mathcal{L}(V)$.)

For each j, let V_j be the image $V_j = p_j(T)q_j(T)(V)$. Then V_j is a subspace of V, and for each $v \in V$, we have

$$v = I_V v = p_1(T)q_1(T)(v) + \cdots + p_k(T)q_k(T)(v) \in V_1 + \cdots + V_k$$

Thus $V = V_1 + \cdots + V_k$.

(*Note:* The subspaces V_j here have not yet been proven to be generalized eigenspaces: that comes next!)

Now by the Cayley-Hamilton Theorem, $(T - \lambda_j I_V)^{m_j} V_j = (T - \lambda_j I_V)^{m_j} p_j(T) q_j(T)(V) = q_j(T) p(T)(V) = \{0\}$. This shows that $V_j \subset \ker(T - \lambda_j)^{m_j}$.

Note that each of the subspaces V_j is T-invariant, since $T(V_j) = T p_j(T) q_j(T)(V) = p_j(T) q_j(T) T(V) \subset p_j(T) q_j(T)(V) = V_j$. Moreover, for $i \neq j$, the restriction of $T - \lambda_i I_V$ to V_j is invertible. For, if $w \in V_j$ such that $(T - \lambda_i I_V) w = 0$, then $T(w) = \lambda_i w$, so $0 = (T - \lambda_j I_V)^{m_j}(w) = (\lambda_i - \lambda_j)^{m_j} w$, which implies that $w = 0$.

Next we prove that the sum $V = V_1 + \cdots V_k$ is direct. If this were not the case, there would exist vectors $v_1 \in V_1, \ldots, v_k \in V_k$, not all zero, such that $v_1 + \cdots + v_k = 0$. Assume that $v_j \neq 0$. Then since $T - \lambda_i I_V$ is invertible on V_j for $i \neq j$, we see that $p_j(T)$ is invertible on V_j and is identically zero on V_i for all other i. Thus

$$
\begin{aligned}
0 &= p_j(T)(v_1 + \cdots + v_k) \\
&= p_j(T)(v_1) + \cdots + p_j(T)(v_k) \\
&= p_j(T) v_j
\end{aligned}
$$

The last expression above is nonzero because $p_j(T)$ is invertible on V_j and $v_j \neq 0$. This contradiction shows that $V = V_1 \oplus \cdots \oplus V_k$.

Note that $\dim V_j \leq m_j$, for all j by Proposition 1.7.4, since V_j is contained in the generalized eigenspace of T corresponding to λ_j. Since we have a direct sum $V = V_1 \oplus \cdots \oplus V_k$ and $\dim V = m_1 + \cdots + m_n$, we must in fact have $\dim V_j = m_j$, for all j.

It remains to prove that $V_j = \ker(T - \lambda_j I_V)^{m_j}$ for all j. It will then follow from the above that $\dim(\ker(T - \lambda_j I_V)^{m_j}) = m_j$. Suppose $v \in \ker(T - \lambda_j I_V)^{m_j}$. Then write $v = v_1 + \cdots + v_k$, where each $v_j \in V_j$. Then

$$
\begin{aligned}
0 &= (T - \lambda_j I_V)^{m_j}(v) \\
&= (T - \lambda_j I_V)^{m_j} v_1 + \cdots + (T - \lambda_j I_V)^{m_j} v_{j-1} + (T - \lambda_j I_V)^{m_j} v_{j+1} + \cdots \\
&\qquad + \cdots + (T - \lambda_j I_V)^{m_j} v_n
\end{aligned}
$$

Now $(T - \lambda_j I_V)^{m_j} v_i \in V_i$ for each i; since the sum $V = \bigoplus_{i=1}^k V_i$ is direct, this forces $(T - \lambda_j I_V)^{m_j} v_i = 0$ for all $i \neq j$. Then, since $(T - \lambda_j I_V)^{m_j}$ is invertible on V_i, it follows that $v_i = 0$ whenever $i \neq j$. Thus $v = v_j$; this shows that $\ker(T - \lambda_j I_V)^{m_j} \subset V_j$, so in fact $\ker(T - \lambda_j I_V)^{m_j} = V_j$. $\square$

1.8 The Jordan Canonical Form

An operator $N \in \mathcal{L}(V)$ is said to be *nilpotent* if $N^m = 0$ for some positive integer m.

Exercise 1.8.1. Let V be a complex vector space. Prove that N is nilpotent if and only if the only eigenvalue of N is 0.

Suppose that V is a vector space over $\mathbb{F}$, and that $N \in \mathcal{L}(V)$ is nilpotent. It is not hard to show that there is a basis of V with respect to which the matrix of T has the form

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}; \tag{1.17}$$

here all the entries on or below the main diagonal are 0. (We call such a matrix *strictly upper triangular.*) If $\mathbb{F} = \mathbb{C}$, then this follows immediately from Theorem 1.6.2 and Exercise 1.8.1. For $\mathbb{F}$ aribitrary, we just consider a basis of $\ker N$, then extend this to a basis of $\ker N^2$, then extend that to a basis of $\ker N^3$, etc. If we continue this procedure, we end up with a basis of $V = \ker N^m$, for sufficiently large m. It is clear that the matrix of N with respect to this basis is strictly upper triangular.

Our goal in this section is to represent a linear operator $T \in \mathcal{L}(V)$ by a matrix which has as many 0's as possible. As a first step, we see that an immediate application of Theorem 1.7.7 is the following result.

Proposition 1.8.2. *Let T be a linear operator on a complex vector space V. Then there is a basis of V with respect to which T has the following matrix in block diagonal form*

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{pmatrix} \tag{1.18}$$

where each A_j is an upper triangular matrix of the form

$$\begin{pmatrix} \lambda_j & & * \\ & \ddots & \\ 0 & & \lambda_j \end{pmatrix} \tag{1.19}$$

Proof. By Theorem 1.7.7, we have $V = \bigoplus_{j=1}^k V_j$, where $V_j = \ker(T - \lambda_j I_V)^{m_j}$. Thus the restriction $(T - \lambda_j I_V)|_{V_j}$ is nilpotent and there is a basis of V_j for which the matrix of this restriction is strictly upper triangular, of the form (1.17). Hence the matrix of $T|V_j$ with respect to this basis is of the form (1.19). If we combine the bases of the V_j so obtained, we get a basis of V with respect to which the matrix of T has the form (1.18). □

We now try to modify the bases of the V_j so as to simplify the block diagonal matrix (1.18) further. As already noted each V_j is T-invariant, and in addition, the restriction $(T - \lambda_j I_V)|_{V_j}$ is nilpotent. Thus we need to find a suitable basis of V_j with respect to which the matrix of $(T - \lambda_j I_V)$ is suitably nice.

Proposition 1.8.3. *Let N be a nilpotent operator on a nonzero vector space V over $\mathbb{F}$. Then there are vectors $v_1, \ldots, v_k$ in V such that*

1. *the list $(v_1, Nv_1, \ldots, N^{r_1}v_1, v_2, Nv_2, \ldots, N^{r_2}v_2, \ldots, v_k, Nv_k, \ldots, N^{r_k}v_k)$ is a basis of V*
2. *$(N^{r_1}v_1, N^{r_2}v_2, \ldots, N^{r_k}v_k)$ is a basis of* $\ker N$.

Proof. The proof is by induction on $\dim V$. If $\dim V = 1$, then N must be the identically 0, so the proposition is trivially true.

Suppose that the proposition holds for all nilpotent linear operators on all vector spaces of dimension $< \dim V$, and let N be a nilpotent linear operator on V. Since the range $N(V)$ is an N-invariant subspace of dimension $< \dim V$, we apply the induction hypothesis to obtain vectors $w_1, \ldots, w_m$ such that

(a) $(w_1, Nw_1, \ldots, N^{s_1}w_1, \ldots, w_m, Nw_m, \ldots, N^{s_m}v_m)$ is a basis of $N(V)$; and

(b) $(N^{s_1}w_1, \ldots, N^{s_m}w_m)$ is a basis of $(\ker N) \cap N(V)$.

Pick vectors $v_1, \ldots, v_m$ such that $Nv_j = w_j$ for $1 \leq j \leq m$, and, if necessary, pick additional vectors $v_{m+1}, \ldots, v_k$ so as to complete the list in (b) above to a basis of $\ker N$.

Put $r_1 = s_1 + 1, \ldots, r_m = s_m + 1$ and let $r_{m+1} = \cdots = r_k = 1$. We first claim that the list

$$(v_1, Nv_1, \ldots, N^{r_1}v_1, v_2, Nv_2, \ldots, N^{r_2}v_2, \ldots, v_m, Nv_m, \ldots, N^{r_m}v_m, v_{m+1}, \ldots, v_k) \tag{1.20}$$

is linearly independent. Indeed, given the relation

$$\sum_{j=1}^{m}\sum_{l=0}^{r_j} a_{lj}N^l v_j \;+\; \sum_{t=m+1}^{k} b_t v_t \;=\; 0, \tag{1.21}$$

we apply the operator N to both sides to obtain

$$\begin{aligned} 0 &= \sum_{j=1}^{m}\sum_{l=0}^{r_j} a_{lj}N^{l+1}v_j \\ &= \sum_{j=1}^{m}\sum_{l=0}^{r_j} a_{lj}N^{l}w_j \\ &= \sum_{j=1}^{m}\sum_{l=0}^{s_j} a_{lj}N^{l}w_j \end{aligned}$$

It follows by condition (a) above that $a_{lj} = 0$ for all $0 \le l \le s_j$; $1 \le j \le m$; that is, all the coefficients a_{lj} in the last sum above vanish. From (1.21), this leaves us with the relation

$$\begin{aligned} 0 &= a_{r_1 1} N^{r_1} v_1 + \cdots + a_{r_m m} N^{r_m} v_m + b_{m+1} v_{m+1} + \cdots + b_k v_k \\ &= a_{r_1 1} N^{s_1} w_1 + \cdots + a_{r_m m} N^{s_m} w_m + b_{m+1} v_{m+1} + \cdots + b_k v_k. \end{aligned}$$

But by condition b and the choice of $v_{m+1}, \ldots, v + k$, we see that all the coefficients above also vanish. It follows that the list (1.20) - which coincides with the list in conclusion (1) of the proposition - is linearly independent.

The list (1.20) is also a basis, since by the induction hypothesis (a), $\dim N(V) = s_1 + \cdots + s_m + m$, and $\dim(\ker N) = k$, so $\dim V = (\sum_{i=1}^m s_i) + m + k = (\sum_{i=1}^m r_i) + k$, which equals the number of vectors in (1.20). Condition (2) in the statement of the proposition is satisfied by construction, since $\ker N$ has basis $(N^{s_1} w_1, \ldots, N^{s_m} w_m, v_{m+1}, \ldots, v_k) = (N^{r_1} v_1, \ldots, N^{r_m} v_m, v_{m+1}, \ldots, v_k)$. $\square$

Remark 1.8.4. The numbers k and $r_1, \ldots, r_k$ in Proposition 1.8.3 are unique in the following sense. Let us, without loss of generality, arrange the basis in Proposition 1.8.3 such that $r_1 \ge r_2 \ge \cdots \ge r_k$. Now suppose that

$$(u_1, \ldots, N^{l_1} u_1, u_2, \ldots, N^{l_2} u_2, \cdots, u_s, \cdots, N^{l_s} u_s)$$

is another basis of V satisfying the conclusions of the proposition, with $l_1 \ge \cdots \ge l_s$. Then $s = k$ and $l_1 = r_1, \ldots, l_k = r_k$. This can be proved by a simple induction argument, going from the range $N(V)$ to V. We leave the details to the student.

Suppose $T \in \mathcal{L}(V)$. We call a basis of V a *Jordan basis for* T if the matrix of T with respect to this basis has block diagonal form

$$\begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_m \end{pmatrix} \tag{1.22}$$

where each diagonal block A_j has the form

$$\begin{pmatrix} \lambda_j & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_j \end{pmatrix} \tag{1.23}$$

The matrix (1.22) is then called the *Jordan canonical matrix* of T. The blocks A_j are called *Jordan blocks*.

Theorem 1.8.5. *Let V be a nonzero complex vector space and let $T \in \mathcal{L}(V)$. Then V has a Jordan basis for T.*

Proof. Assume that T has characteristic polynomial $p(\lambda) = \prod_{j=1}^{l}(\lambda - \lambda_j)^{m_j}$, where $\lambda_1, \cdots, \lambda_l$ are the distinct eigenvalues of T. By Theorem 1.7.7, V is the direct sum

$$V = \bigoplus_{j=1}^{l} V_j,$$

where V_j is the generalized eigenspace $\ker(T - \lambda_j I_V)^{m_j}$. V_j is T-invariant and of course the restriction $N_j := (T - \lambda_j I_V)|_{V_j} = T|_{V_j} - \lambda_j I_{V_j}$ is nilpotent. But then we can apply Proposition 1.8.3 to N_j to obtain a basis of V_j for which the matrix of N has block form

$$\begin{pmatrix} R_1 & & & 0 \\ & R_2 & & \\ & & \ddots & \\ 0 & & & R_k \end{pmatrix} \tag{1.24}$$

where each diagonal block R_i is of the form

$$\begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \tag{1.25}$$

Each block R_i above corresponds to the list $(N_j^{r_i} v_i, \ldots, N_j v_i, v_i)$ corresponding to the vector v_i in the basis given in Part (1) of Proposition 1.8.3. The linear span of $(N_j^{r_i} v_i, \ldots, N_j v_i, v_i)$ is invariant under $T|_{V_j} = N_j + \lambda_j I_{V_j}$, and on this linear span, the matrix of $T|_{V_j}$ is of the form 1.23.

Putting these bases together, we obtain a Jordan basis for T. □

Corollary 1.8.6. *Let A be an $n \times n$ matrix with complex entries. Then there is an $n \times n$ matrix S such that SAS^{-1} is of the form* (1.22).

Remark 1.8.7. Since the generalized eigenspace corresponding to λ_j has dimension m_j, the multiplicity of λ_j, the size of the collection of blocks corresponding to λ_j in the Jordan canonical form (1.22) is unique. Then, by Remark 1.8.4, for each λ_j, the number of Jordan blocks and their respective sizes is also unique.

1.9 The Jordan-Chevalley Decomposition

Suppose that V is a vector space of dimension n over $\mathbb{F}$ and that $T \in \mathcal{L}(V)$. Since $\dim(\mathcal{L}(V)) = n^2$, the operators $I_V, T, T^2, \ldots T^{n^2}$ are linearly dependent in $\mathcal{L}(V)$. We thus have a relation

$$Q(T) := a_0 I_V + a_1 T + a_2 T^2 + \cdots + a_{n^2} T^{n^2} = 0 \tag{1.26}$$

such that not all coefficients a_i are 0.

A *monic* polynomial $p(z)$ is a polynomial in z whose highest degree coefficient is 1. Thus we may write $p(z) = z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0$.

Proposition 1.9.1. *Let $T \in \mathcal{L}(V)$, and let $p(z)$ be a monic polynomial of smallest positive degree such that $p(T) = 0$. If $s(z)$ is any polynomial such that $s(T) = 0$, then $p(z)$ divides $s(z)$.*

Proof. By the Euclidean algorithm (i.e., long division), we have

$$s(z) = q(z)p(z) + r(z),$$

where $q(z)$ and $r(z)$ are polynomials with $\deg r(z) < \deg p(z)$. Replacing z by T in the above we obtain

$$\begin{aligned} & s(T) = q(T)p(T) + r(T) \\ \Longrightarrow \qquad & \quad\; 0 = r(T), \end{aligned}$$

which by the minimality of p implies that $r(z) = 0$. □

It follows that there is only one such polynomial $p(z)$. We call this polynomial the *minimal polynomial* of $T \in \mathcal{L}(V)$, and denote it by $P_{\text{min}}(z)$. From (1.26), any minimal polynomial of T has degree $\leq n^2$, and better yet, by the Cayley-Hamilton Theorem, it must have degree $\leq n$.

The following is an immediate consequence of Proposition 1.9.1.

Corollary 1.9.2. *The minimal polynomial of T divides the characteristic polynomial of T.*

Proposition 1.9.3. *Let V be a complex vector space and $T \in \mathcal{L}(V)$. Then the roots of the minimal polynomial of T are precisely its eigenvalues.*

Proof. Let λ be an eigenvalue of T, and let v eigenvector corresponding to λ. Then since $v \neq 0$,

$$0 = P_{\text{min}}(T)(v) = P_{\text{min}}(\lambda)v \implies P_{\text{min}}(\lambda) = 0.$$

Conversely, suppose that λ is a root of P_{min}. Then by Proposition 1.9.2, λ is a root of the characteristic polynomial of T, whence λ is an eigenvalue of T. □

If the eigenvalues of T all have multiplicity 1; that is, if the characteristic polynomial of T is of the form $\chi(z) = \prod_{i=1}^n (z - \lambda_i)$, with the λ_i distinct, then, by Corollary 1.9.2 and Proposition 1.9.3, the characteristic polynomial coincides with the minimal polynomial. On the other hand, if T is scalar multiplication, $T = \lambda I_V$, then the minimal polynomial of T is $z - \lambda$, whereas its characteristic polynomial is $(z - \lambda)^n$.

Exercise 1.9.4. Let T be the linear operator on a 6-dimensional complex vector space whose Jordan matrix is

$$\begin{pmatrix} \lambda_1 & 1 & & & & 0 \\ & \lambda_1 & 1 & & \ddots & \\ & & \lambda_1 & & & \\ & \ddots & & \lambda_2 & 1 & \\ & & & & \lambda_2 & \\ 0 & & & & & \lambda_2 \end{pmatrix} \tag{1.27}$$

Find the minimal polynomial of T. For any $T \in \mathcal{L}(V)$, formulate a theorem stating what the minimal polynomial is in terms of its Jordan matrix.

Exercise 1.9.5. Suppose that V is a vector space over $\mathbb{C}$, and that $T \in \mathcal{L}(V)$ has characteristic polynomial $\chi(z) = \prod_{i=1}^{k}(z-\lambda_i)^{m_i}$ and minimal polynomial $P_{\min}(z) = \prod_{i=1}^{k}(z-\lambda_i)^{r_i}$. Suppose that V_i is the generalized eigenspace corresponding to λ_i. Prove that

$$r_i = \min\{r \,|\, (T-\lambda_i I_V)^r|_{V_i} = 0\}.$$

Exercise 1.9.6. Suppose that $T \in \mathcal{L}(V)$ and $v \in V$. Prove that there is a unique monic polynomial $s(z)$ of lowest degree such that $s(T)\,v = 0$. Then prove that $s(z)$ divides the minimal polynomial of T.

Exercise 1.9.7. Give an example of an operator on $\mathbb{C}^4$ whose characteristic polynomial is $z(z-1)^2(z-2)$ and whose minimal polynomial is $z(z-1)(z-2)$.

Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. The operator T is said to be *semisimple* if the minimal polynomial of T is $(z-\lambda_1)\cdots(z-\lambda_k)$.

Proposition 1.9.8. *T is semisimple if and only if it is diagonalizable; that is, there is a basis of V such that the matrix of T with respect to this basis is diagonal.*

Proof. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T.

Suppose first that T is diagonalizable. Then there is a basis of V consisting of eigenvectors of T. Let $V_1 (= \ker(T-\lambda_1 I_V))$ be the eigenspace of T corresponding to λ_1, V_2 the eigenspace corresponding to λ_2, etc. Then we must have $V = V_1 \oplus \cdots \oplus V_k$. For each j, the restriction $(T-\lambda_j I_V)|_{V_j}$ is obviously identically 0. Hence $(T-\lambda_1 I_V)\cdots(T-\lambda_k I_V) \equiv 0$. so the minimal polynomial of T is $(z-\lambda_1)\cdots(z-\lambda_k)$, and T is semisimple.

Next, we assume that T is semisimple, and try to prove that T is diagonalizable. If V_j is the generalized eigenspace corresponding to λ_j, we have the direct decomposition

$$V = \bigoplus_{j=1}^{k} V_j$$

V_j is invariant under $T - \lambda_i I_V$, and if $i \neq j$, the restriction $(T - \lambda_i I_V)|V_j$ is invertible, for if $v \in V_j$ satisfies $(T - \lambda_i I_V)(v) = 0$, then $v \in V_j \cap V_i = \{0\}$. Thus the restriction of $\prod_{i \neq j}(T - \lambda_i I_V)$ to V_j is invertible. Since $\prod_{i=1}^{k}(T - \lambda_i I_V) = 0$, we see that $T - \lambda_j I_V \equiv 0$ on V_j, so T is just scalar multiplication by λ_j on V_j. □

Note: Most authors define a semisimple linear operator as one which is diagonalizable. But our definition allows us more flexibility, as the following proposition shows.

Proposition 1.9.9. *Suppose that $T \in \mathcal{L}(V)$ is semisimple, and that W is a subspace of V invariant under T. Then the restriction $T|_W$ is semisimple.*

Proof. Let $P_{\min}(z)$ denote the minimal polynomial of T. We have $P_{\min}(T|_W) = P_{\min}(T)|_W = 0$, so the minimal polynomial of $T|_W$ divides $P_{\min}(z)$. This minimal poynomial must then be of the form $\prod_{i \in J}(z - \lambda_i)$, where J is a subset of the set of eigenvalues of T. □

Consider now the Jordan matrix of the linear operator T on $\mathbb{C}^6$ in Exercise 1.9.4. Let S be the semisimple linear operator on C^6 whose matrix with respect to this Jordan basis of T is

$$\begin{pmatrix} \lambda_1 & & & & & 0 \\ & \lambda_1 & & & \ddots & \\ & & \lambda_1 & & & \\ & \ddots & & \lambda_2 & & \\ & & & & \lambda_2 & \\ 0 & & & & & \lambda_2 \end{pmatrix}$$

and let N be the nilpotent operator with matrix

$$\begin{pmatrix} 0 & 1 & & & & 0 \\ & 0 & 1 & & \ddots & \\ & & 0 & & & \\ & \ddots & & 0 & 1 & \\ & & & & 0 & \\ 0 & & & & & 0 \end{pmatrix}$$

Then $T = S + N$, and it is easy to check that S and N commute. Using its Jordan matrix, it is easy to see that, in fact, any linear operator T on V has a *Jordan-Chevalley decomposition* $T = S + N$, where S is semisimple and N is nilpotent, and S and N commute. S and N satisfy a few additional properties, given in Theorem 1.9.14 below.

Lemma 1.9.10. *(The Chinese Remainder Theorem) Suppose that $p_1(z), \ldots, p_m(z)$ are nonconstant polynomials which are pairwise relatively prime. If $r_1(z), \ldots, r_m(z)$ are any polynomials, then then is a polynomial $P(z)$ such that $P(z) \equiv r_j(z) \mod p_j(z)$, for all j.*

Proof. For each j, let $Q_j(z) = \prod_{i \neq j} p_i(z)$. Then there exist polynomials $A_j(z)$ and $B_j(z)$ such that $A_j(z)p_j(z) + B_j(z)Q_j(z) = 1$. Now put

$$P(z) = \sum_{i=1}^{m} r_i(z) B_i(z) Q_i(z).$$

For $i \neq j$, $p_j(z)$ divides $Q_i(z)$, so

$$\begin{aligned}
P(z) &\equiv \sum_{i=1}^{m} r_i(z) B_i(z) Q_i(z) && \mod p_j(z) \\
&\equiv r_j(z) B_j(z) Q_j(z) && \mod p_j(z) \\
&\equiv r_j(z) \left(1 - A_j(z) p_j(z)\right) && \mod p_j(z) \\
&\equiv r_j(z) && \mod p_j(z)
\end{aligned}$$

□

The Chinese Remainder Theorem, properly formulated, holds for all principal ideal domains, and in fact for all commutative rings with identity. The original form of the theorem, as it pertains to the integers, appeared in a third-century AD book by the mathematician Sun Tzu [11] (not the same guy who wrote *The Art of War*).

Proposition 1.9.11. *Suppose that S_1 and S_2 are two diagonalizable linear operators on V. Then $S_1S_2 = S_2S_1$ if and only if S_1 and S_2 are simultaneously diagonalizable; that is, if and only if there is a basis of V for which the matrices of S_1 and S_2 are both diagonal.*

Proof. Since diagonal matrices of the same size commute, it is clear that if S_1 and S_2 are simultaneously diagonalizable, then they commute.

Conversely, let us assume that S_1 and S_2 are diagonalizable linear operators such that $S_1S_2 = S_2S_1$. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of S_1, with respective eigenspaces $V_1, \ldots, V_k$. Then $V = \bigoplus_{i=1}^{k} V_i$. Since S_1 and S_2 commute, each eigenspace V_i is invariant under S_2. Then by Lemma 1.9.9, the restriction $S_2|_{V_i}$ is diagonalizable. Choose a basis of V_i for which the matrix of $S_2|_{V_i}$ is diagonal. Then, combining the bases of the V_i, we obtain a basis of V for which the matrices of S_1 and S_2 are both diagonal. □

In particular, this proposition says that $S_1 + S_2$ must be semisimple!

Exercise 1.9.12. Show that if $\{S_1, \ldots, s_m\}$ are pairwise commuting semisimple elements of $\mathcal{L}(V)$, then there exists a basis of V for which the matrices of $\{S_1, \ldots, s_m\}$ are all diagonal.

Lemma 1.9.13. *Let N_1 and N_2 be commuting nilpotent linear operators on a vector space V. Then $N_1 + N_2$ is nilpotent.*

Proof. Assume that $N_1^{m_1} = 0$ and $N_2^{m_2} = 0$. Since N_1 and N_2 commute, we can apply the binomial theorem to obtain, for any $m \in \mathbb{N}$,

$$(N_1 + N_2)^m = \sum_{k=0}^{m} \binom{m}{k} N_1^k N_2^{m-k}$$

It follows immediately that $(N_1 + N_2)^{m_1+m_2} = 0$. □

Theorem 1.9.14. *(The Jordan-Chevalley Decomposition) Let V be a complex vector space, and let $T \in \mathcal{L}(V)$. Then there exists a polynomial $p(z)$ such that if $q(z) = z - p(z)$, the following properties hold:*

1. *$S := p(T)$ is semisimple and $N := q(T)$ is nilpotent;*
2. *Any linear operator which commutes with T must commute with both S and N;*
3. *If S' and N' are commuting semisimple and nilpotent operators, respectively, such that $T = S' + N'$, the $S' = S$ and $N' = N$; and*
4. *If $A \subset B \subset V$ are subspaces, and if $T : B \to A$, then so do S and N.*

Proof. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T, and assume that the minimal polynomial of T is $P_{\min}(z) = \prod_{i=1}^{k}(z - \lambda_i)^{r_i}$. Now according to the Chinese Remainder Theorem, there exists a polynomial $p(z)$ such that, for each i,

$$p(z) \equiv \lambda_i \mod (z - \lambda_i)^{r_i} \tag{1.28}$$
$$p(z) \equiv 0 \mod z \tag{1.29}$$

(In case one of the λ_i's equals 0, then (1.28) implies (1.29), so the second condition above is superfluous. Condition (1.29) is really only needed to prove the technical conclusion (4) above.)

Let V_i be the generalized eigenspace of T corresponding to λ_i. Then we have the direct decomposition

$$V = \bigoplus_{i=1}^{k} V_i.$$

Each V_i is invariant under T, hence is invariant under $p(T)$. Now by Exercise 1.9.5, $(T - \lambda_i I_V)^{r_i}$ vanishes on V_i, so by the relation (1.28) above, we see that the operator $p(T) - \lambda_i I_V$ is identically 0 on V_i. Thus $p(T)v = \lambda_i v$ for all $v \in V_i$, and it follows that $p(T)$ is semisimple on V.

For any $v \in V_i$, we also have $q(T)v = (T - p(T))(v) = (T - \lambda_i I_V)v$, so (since $(T - \lambda_i I_V)^{r_i} = 0$ on V_i), we see that $q(T)$ is nilpotent on V_i, with $q(T)^{r_i} \equiv 0$ on V_i. Putting $R = \max_{1 \le i \le k} r_i$, we have $q(T)^R = 0$ on V, so $q(T)$ is a nilpotent operator.

Since $S = p(T)$ and $N = q(T)$ are polynomials in T, they commute, and in addition, any linear operator which commutes with T must commute with S and N. By 1.29, $p(z)$ and $q(z)$ have constant term equal to 0, so clearly S and N satisfy statement (4) above.

The only thing left to prove is the uniqueness statement (3). So let S' and N' be commuting semisimple and nilpotent linear operators, respectively, on V, such that $T = S' + N'$. Then S' and N' commute with T, and so must commute with both S and N. We then have

$$S - S' = N' - N$$

Since S and S' are commuting semisimple operators, the left hand side above is a semisimple operator by Lemma 1.9.11. On the other hand, since N and N' commute, the right hand side above is a nilpotent operator, by Lemma 1.9.13. The only eigenvalue of $S - S'$ is therefore 0, whence $S - S' = 0$. Therefore, $N' - N = 0$. $\square$

1.10 Symmetric Bilinear Forms

Let V be a vector space over $\mathbb{F}$. A *bilinear form* on V is a map

$$\begin{aligned} \langle\ ,\ \rangle : V \times V &\to \mathbb{F} \\ (v, w) &\mapsto \langle v, w \rangle \end{aligned} \tag{1.30}$$

which is linear in each of its two arguments:

$$\begin{aligned} \langle \alpha v + \beta v', w \rangle &= \alpha \langle v, w \rangle + \beta \langle v', w \rangle \\ \langle v, \alpha w + \beta w' \rangle &= \alpha \langle v, w \rangle + \beta \langle v, w' \rangle, \end{aligned}$$

for all $v, v', w, w' \in V$ and all $\alpha, \beta \in \mathbb{F}$.

Example 1.10.1. The dot product on $\mathbb{R}^n$ is a bilinear form. More generally, an inner product on a real vector space V is a bilinear form.

Example 1.10.2. Let A be any $n \times n$ matrix over $\mathbb{F}$. Using the matrix A, we can define a bilinear form on $\mathbb{F}^n$ by putting

$$\langle x, y \rangle = {}^t x A x \qquad \text{for all } x, y \in \mathbb{F}^n$$

As a special case, when $A = I_n$ and $\mathbb{F} = \mathbb{R}$, we obtain the dot product on $\mathbb{R}^n$.

Let $\langle\ ,\ \rangle$ be a bilinear form on V. Fix a basis $B = (v_1, \dots, v_n)$ of V. The *matrix of* $\langle\ ,\ \rangle$ *with respect to* B is the $n \times n$ matrix A whose (i, j) entry is $a_{ij} = \langle v_i, v_j \rangle$. This matrix A completely determines the bilinear form, since each vector in V is a unique linear combination of the basis vectors in B:

$$v = \sum_{i=1}^{n} x_i v_i \text{ and } w = \sum_{j=1}^{n} y_j v_j \implies \langle v, w \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{ij} y_j \tag{1.31}$$

Given the basis B of V, we have a *coordinate map* $[\,\cdot\,]_B$ *from* V *onto* $\mathbb{F}^n$ *with respect to* B: namely, if $v \in V$, then

$$[v]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \iff v = \sum_{i=1}^{n} x_i v_i.$$

The coordinate map $v \mapsto [v]_B$ is a linear isomorphism from V onto $\mathbb{F}^n$.

Exercise 1.10.3. If $T \in \mathcal{L}(V)$ with matrix $M_{B,B}(T)$ with respect to B, show that $[Tv]_B = M_{B,B}(T)[v]_B$, for all $v \in V$.

Now again let $\langle\ ,\ \rangle$ be a bilinear form on V, and let A be its matrix with respect to the basis B of V. Let $v = \sum_{i=1}^{n} x_i v_i$ and $w = \sum_{j=1}^{n} y_j v_j$ be vectors in V. Then by (1.30), we have

$$\begin{aligned} \langle v, w \rangle &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{ij} y_j \\ &= \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & & a_{1n} \\ & \cdots & \\ a_{n1} & & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= {}^t[v]_B\, A\, [w]_B \end{aligned} \tag{1.32}$$

Thus Example 1.10.2 essentially gives us *all* bilinear forms on V, once we've fixed a basis B of V.

A bilinear form $\langle\ ,\ \rangle$ is called *nondegenerate* if, whenever v is a nonzero vector in V, there is a $w \in V$ such that $\langle v, w \rangle \neq 0$. The choice of the vector w, which is necessarily nonzero, depends on the vector v.

Theorem 1.10.4. *Let $\langle\ ,\ \rangle$ be a bilinear form on V, and let A be its matrix with respect to a given basis B of V. Then $\langle\ ,\ \rangle$ is nondegenerate if and only if A is nonsingular.*

Proof. Suppose that A is nonsingular. Let v be a nonzero vector in V. Then $[v]_B$ is a nonzero vector in $\mathbb{F}^n$. Since A is nonsingular, its rows are linearly independent, so ${}^t[v]_B A$ is a nonzero row matrix. Hence there exists an element $y \in \mathbb{F}^n$ such that ${}^t[v]_B A\, y \neq 0$. If we let $w \in V$ be the vector such that $[w]_B = y$, then according to (1.32), we have $\langle v, w \rangle \neq 0$. Thus $\langle\ ,\ \rangle$ is nondegenerate.

Suppose next that A is singular. Then its rows are linearly dependent, so there is an $x \neq 0$ in $\mathbb{F}^n$ such that ${}^t x\, A = 0$. Let v be the vector in V such that $[v]_B = x$. Then according to (1.32), we get $\langle v, w \rangle = 0$ for all $w \in V$. This shows that $\langle\ ,\ \rangle$ is not nondegenerate; i.e., is *degenerate*. $\square$

Example 1.10.5. Let

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

Then the bilinear form on $\mathbb{R}^2$ given by $(x, y) = {}^t x\, A\, y$ is nondegenerate.

Example 1.10.6. Let J_n be the $2n \times 2n$ matrix which is given in block form as

$$J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}, \tag{1.33}$$

where the "0" in the matrix above refers to the zero $n \times n$ matrix. Note that J_n is nonsingular, with inverse $J_n^{-1} = -J_n$. J_n gives rise to a nondegenerate symmetric bilinear form on $\mathbb{F}^{2n}$ given by

$$\langle x, y \rangle = {}^t x\, J_n\, y \qquad \text{for all } x, y \in \mathbb{F}^{2n}$$

When $\mathbb{F} = \mathbb{R}$, we call this form the *standard symplectic form on* $\mathbb{R}^{2n}$.

Let A be a square matrix with entries in $\mathbb{F}$. A is said to be *symmetric* if ${}^t A = A$. If $A = (a_{ij})$, this is equivalent to the condition that $a_{ij} = a_{ji}$ for all i and j. A is called *skew-symmetric* if ${}^t A = -A$. This is equivalent to the condition $a_{ij} = -a_{ji}$ for all i, j. Note that the diagonal entries of a skew-symmetric matrix are all 0.

A *symmetric bilinear form* on a vector space V is a bilinear form $\langle\ ,\ \rangle$ on V such that $\langle v, w \rangle = \langle w, v \rangle$ for all v, $w \in V$. The dot product on $\mathbb{R}^n$, or more generally, any inner product on a real vector space, is an example of a (nondegenerate) symmetric bilinear form.

Again fix a basis $B = (v_1, \ldots, v_n)$ of V, and let A be the matrix of $\langle\ ,\ \rangle$ with respect to B. If $\langle\ ,\ \rangle$ is symmetric, then A is a symmetric $n \times n$ matrix: $a_{ij} = \langle v_i, v_j \rangle = \langle v_j, v_i \rangle = a_{ji}$.

Conversely, it is an easy calculation using (1.31) to show that if A is a symmetric matrix, then $\langle\ ,\ \rangle$ is a symmetric bilinear form on V.

Suppose that $\langle\ ,\ \rangle$ is a symmetric bilinear form on a vector space V over $\mathbb{F}$. For any subspace W of V, the *orthogonal complement of* W is the set

$$W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}.$$

It is easy to see that $W^{\perp}$ is a subspace of V. Note that $V^{\perp} = \{v \in V \mid \langle v, v' \rangle = 0 \text{ for all } v' \in V\}$, so $\langle\ ,\ \rangle$ is nondegenerate if and only if $V^{\perp} = \{0\}$.

For any $v \in V$, we let f_v be the linear functional on V given by $f_v(v') = \langle v, v' \rangle$, for all $v' \in V$.

Proposition 1.10.7. *Suppose that $\langle\ ,\ \rangle$ is nondegenerate. If W is a subspace of V, then the map $v \mapsto f_v|_W$ is a linear map of V onto the dual space W^*, with kernel $W^{\perp}$.*

Proof. The map $f : V \to V^*$ given by $v \mapsto f_v$ is easily seen to be linear. Its kernel is $V^{\perp} = \{0\}$, since $\langle\ ,\ \rangle$ is nondegenerate. Since $\dim V = \dim V^*$, f is onto, and we conclude that any element of V^* is of the form f_v, for a unique $v \in V$.

Next we prove that any linear functional on W can be extended to a linear functional on V. To be precise, suppose that $g \in W^*$. Choose any subspace U of V complementary to W, so that $V = W \oplus U$. Then define the function G on V by $G(w + u) = g(w)$ for all $w \in W$ and all $u \in U$. G is a well-defined linear functional on V such that $G|_W = g$.

The restriction map $f \mapsto f|W$ is a linear map from V^* to W^*, and the above shows that it is surjective. Since the map $v \mapsto f_v$ is a linear bijection from V onto V^*, we see that the composition $v \mapsto f_v|_W$ is a surjective linear map from V onto W^*. The kernel of this map is clearly $W^{\perp}$. $\square$

This proposition implies that

$$\dim W^{\perp} = \dim V - \dim W^* = \dim V - \dim W. \tag{1.34}$$

Later, we will make use of the following remarkable fact from linear algebra. (See [1], Theorem 7.9, where it is stated for more general normal matrices.)

Theorem 1.10.8. *Let A be a symmetric $n \times n$ matrix over $\mathbb{F}$. Then there exists an $n \times n$ matrix U such that UAU^{-1} is a diagonal matrix.*

Chapter 2

Lie Algebras: Definition and Basic Properties

2.1 Elementary Properties

Let $\mathcal{A}$ be a vector space over $\mathbb{F}$. $\mathcal{A}$ is said to be an *algebra* over $\mathbb{F}$ if there is a binary operation (a "product") on $\mathcal{A}$

$$\begin{aligned} \mathcal{A} \times \mathcal{A} &\to \mathcal{A} \\ (a, b) &\mapsto a.b \end{aligned}$$

such that for all $a, b, c \in \mathcal{A}$ and $\alpha, \beta \in \mathbb{F}$, we have

$$\begin{aligned} a.(b + c) &= a.b + a.c \\ (a + b).c &= a.c + b.c \\ \alpha(a.b) &= (\alpha a).b = a.(\alpha b). \end{aligned}$$

The algebra $\mathcal{A}$ is *commutative*, or *abelian*, if $a.b = b.a$ for all a, $b \in \mathcal{A}$. $\mathcal{A}$ is *associative* if $a.(b.c) = (a.b).c$ for all $a, b, c \in \mathcal{A}$.

There are numerous examples of algebras. Here are but a pitiful few:

1. $P[z_1, \dots, z_n]$, the algebra of polynomials in the variables $z_,, \dots, z_n$, wth coefficients in $\mathbb{F}$, is a commutative, associative algebra over $\mathbb{F}$.
2. The vector space $M_n(\mathbb{F})$ of $n \times n$ matrices with entries in $\mathbb{F}$ is an associative (but not commutative) algebra over $\mathbb{F}$, of dimension n^2.
3. The vector space $\mathcal{L}(V)$ is an associative algebra under composition of linear operators. If we fix a basis B of V and identify each $T \in \mathcal{L}(V)$

with its matrix $M_B(T)$, then by (1.6), we see that $\mathcal{L}(V)$ is the same as the algebra $M_{\dim V}(\mathbb{F})$.

4. The algebra $\mathbb{H}$ of quaternions is the vector space $\mathbb{R}^4$, in which each element is written in the form $\mathbf{a} = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ (with $a_0, \ldots, a_3 \in \mathbb{R}$), and in which the multiplication is defined by the following rule, extended distributively:

$$\begin{aligned} \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 &= -1 \\ \mathbf{ij} = -\mathbf{ji} &= \mathbf{k} \\ \mathbf{jk} = -\mathbf{kj} &= \mathbf{i} \\ \mathbf{ki} = -\mathbf{ik} &= \mathbf{j} \end{aligned} \tag{2.1}$$

   Then $\mathbb{H}$ is a 4-dimensional algebra over $\mathbb{R}$, or a 2-dimensional algebra over $\mathbb{C}$ (with basis $(1, \mathbf{j})$). It is associative but not commutative.

5. The exterior algebra ΛV of a vector space V over $\mathbb{F}$, with wedge multiplication, is an associative, noncommutative algebra over $\mathbb{F}$, of dimension $2^{\dim V}$.

Definition 2.1.1. Let $\mathfrak{g}$ be an algebra over $\mathbb{F}$, with product $[x, y]$. $\mathfrak{g}$ is called a *Lie algebra* if

$$[x, x] = 0 \qquad \text{for all } x \in \mathfrak{g} \tag{2.2}$$

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \qquad \text{for all } x, y, z \in \mathfrak{g} \tag{2.3}$$

The multiplication on the Lie algebra $\mathfrak{g}$ is called the *Lie bracket*. Any algebra product which satisfies (2.2) is said to be *anticommutative*. The identity (2.3) is called the *Jacobi identity*.

Note: We have adopted the common practice of Lie theorists (those who study Lie groups and Lie algebras) to use lowercase gothic letters to denote Lie algebras, and to use $\mathfrak{g}$ to denote a typical Lie algebra. I don't know exactly how this practice came about, but I suspect that it may have had something to do with Hermann Weyl, a German who was the leading practitioner of Lie theory (and many other types of mathematics) in the mid-20th century.

Proposition 2.1.2. *The condition* (2.2) *is equivalent to the condition that*

$$[x, y] = -[y, x]$$

for all $x, y \in \mathfrak{g}$.

Proof. Suppose that $[x, y] = -[y, x]$ for all $x, y \in \mathfrak{g}$. then $[x, x] = -[x, x] \implies [x, x] = 0$ for all $x \in \mathfrak{g}$.

Conversely, if $[x,x]=0$ for all $x \in \mathfrak{g}$, then for all $x,y \in \mathfrak{g}$,

$$\begin{aligned}0 &= [x+y, x+y] = [x,x]+[x,y]+[y,x]+[y,y] = [x,y]+[y,x] \\ &\implies [x,y] = -[y,x]\end{aligned}$$

□

Proposition 2.1.3. *The Jacobi identity is equivalent to its alter ego:*

$$[[x,y],z] + [[y,z],x] + [[z,x],y] = 0. \tag{2.4}$$

Proof. Easy, just multiply the Jacobi identity by -1 and use the preceding proposition. □

We now consider some examples of Lie algebras.

Example 2.1.4. An *abelian* Lie algebra is one in which the Lie bracket is commutative. Thus, in an abelian Lie algebra $\mathfrak{g}$, $[x,y]=[y,x]$, for all $x,y \in \mathfrak{g}$. But then, we also have $[x,y]=-[y,x]$, so $[x,y]=0$ for all $x,y \in \mathfrak{g}$. Thus the Lie bracket in an abelian Lie algebra is identically 0.

Conversely, if we define the bracket on a vector space V by $[x,y]=0$ for all $x,y \in V$, we see that the conditions (2.2) and (2.3) are immediately satisfied. Thus, any vector space may be endowed with the (obviously trivial) structure of an abelian Lie algebra.

Exercise 2.1.5. In $\mathbb{R}^3$, show that

$$\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}.$$

Then show that the cross product is a Lie bracket on the real vector space $\mathbb{R}^3$, so that $\mathbb{R}^3$ is a 3-dimensional Lie algebra over $\mathbb{R}$.

Example 2.1.6. Here's a great source of Lie algebras: take any associative algebra $\mathcal{A}$, and define a Lie bracket on $\mathcal{A}$ by putting

$$[x,y] := xy - yx$$

Of course we need to verify that $[x,y]$ is indeed a Lie bracket on $\mathcal{A}$. But the anticommutativity is obvious, and the verification of the Jacobi identity is a routine calculation:

$$\begin{aligned}&[x,[y,z]] + [y,[x,z]] + [z,[x,y]] \\ &= x(yz-zy) - (yz-zy)x + y(xz-zx) - (xz-zx)y + z(xy-yx) - (xy-yx)z \\ &= xyz - xzy - yzx + zyx + yxz - yzx - xzy + zxy + zxy - zyx - xyz + yxz \\ &= 0.\end{aligned}$$

The Lie bracket $[x, y]$ defined above is called the *commutator product* of x and y.

The associative algebra $M_n(\mathbb{F})$ of $n \times n$ matrices can therefore be given a Lie algebra structure. So equipped, we will refer to this Lie algebra as $\mathrm{gl}(n, \mathbb{F})$.

Example 2.1.7. Let V be an n-dimensional vector space over $\mathbb{F}$. As we have already seen, $\mathcal{L}(V)$ is an associative algebra of dimension n^2, and if we fix a basis B of V, $\mathcal{L}(V)$ is isomorphic (as an associative algebra) to $M_n(\mathbb{F})$ under the map $T \mapsto M_{B,B}(T)$. From the preceding example, $\mathcal{L}(V)$ has a Lie algebra structure, given by $[S, T] = ST - TS$. The notation $\mathrm{gl}(V)$ will denote the associative algebra $\mathcal{L}(V)$ when it is given this Lie algebra structure

Example 2.1.8. Here's another construction which produces lots of Lie algebras. Suppose that $\mathcal{A}$ is *any* algebra – not necessarily commutative or asociative – over $\mathbb{F}$. A *derivation* of $\mathcal{A}$ is a linear operator D on $\mathcal{A}$ which satisfies *Leibniz' rule* with respect to the product on $\mathcal{A}$:

$$D(a.b) = D(a).b + a.D(b) \qquad \text{for all } a,\ b \in \mathcal{A} \tag{2.5}$$

Let $\mathrm{Der}(\mathcal{A})$ be the set of all derivations of $\mathcal{A}$. We want to show that $\mathrm{Der}(\mathcal{A})$ has the structure of a Lie algebra, under an appropriate definition of the Lie bracket.

First, we routinely verify that $\mathrm{Der}(\mathcal{A})$ is a vector space, or more precisely, a subspace of $\mathcal{L}(\mathcal{A})$. To accomplish this, we just need to establish that:

1. If D_1 and D_2 belong to $\mathrm{Der}(\mathcal{A})$, then so does $D_1 + D_2$.
2. If $D \in \mathrm{Der}(\mathcal{A})$ and $\lambda \in \mathbb{F}$, then $\lambda D \in \mathrm{Der}(\mathcal{A})$.

Exercise 2.1.9. Verify (1) and (2) above.

Next, for any D_1 and D_2 in $\mathrm{Der}(\mathcal{A})$, we define $[D_1, D_2]$ to be the commutator product $D_1D_2 - D_2D_1$. Clearly $[D_1, D_2] \in \mathcal{L}(\mathcal{A})$. We now claim that $[D_1, D_2] \in \mathrm{Der}(\mathcal{A})$. This claim is verified by another routine calculation: for any $a, b \in \mathcal{A}$, we have

$$\begin{aligned}
[D_1, D_2](a.b) &= (D_1D_2 - D_2D_1)(a.b) \\
&= D_1D_2(a.b) - D_2D_1(a.b) \\
&= D_1(D_2(a).b + a.D_2(b)) - D_2(D_1(a).b + a.D_1(b)) \\
&= (D_1D_2(a)).b + D_2(a).D_1(b) + D_1(a).D_2(b) + a.(D_1D_2(b)) \\
&\qquad - (D_2D_1(a)).b - D_1(a).D_2(b) - D_2(a).D_1(b) - a.(D_2D_1(b)) \\
&= (D_1D_2 - D_2D_1)(a).b + a.(D_1D_2 - D_2D_1)(b) \\
&= [D_1, D_2](a).b + a.[D_1, D_2](b)
\end{aligned}$$

It is not difficult to verify that $\mathrm{Der}(\mathcal{A})$, equipped with the commutator product, is an algebra. In order to prove that it is a Lie algebra, we need to check that $[\ ,\ ]$ is anticommutative and satisfies the Jacobi identity. But clearly, $[D, D] = 0$ for all $D \in \mathrm{Der}(\mathcal{A})$, and the Jacobi identity is verified by exactly the same calculation as in Example 2.1.6.

Definition 2.1.10. Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$ and $\mathfrak{s}$ a vector subspace of $\mathfrak{g}$. We say that $\mathfrak{s}$ is a *Lie subalgebra of* $\mathfrak{g}$ if $\mathfrak{s}$ is closed under the Lie bracket in $\mathfrak{g}$. That is, $\mathfrak{s}$ is a Lie subalgebra of $\mathfrak{g}$ if $[x, y] \in \mathfrak{s}$ whenever x, $y \in \mathfrak{s}$.

Example 2.1.11. If $\mathcal{A}$ is an algebra over $\mathbb{F}$, then $\mathrm{Der}(\mathcal{A})$ is a Lie subalgebra of the Lie algebra $\mathrm{gl}(\mathcal{A})$.

Example 2.1.12. *Any* one-dimensional subspace $\mathbb{F}x$ of a Lie algebra $\mathfrak{g}$ is an abelian Lie subalgebra of $\mathfrak{g}$, for $[cx, dx] = cd[x, x] = 0$ for all c, $d \in \mathbb{F}$.

Example 2.1.13. Suppose that $\mathcal{A}$ is an associative algebra over $\mathbb{F}$, and that $\mathcal{B}$ is a *subalgebra* of $\mathcal{A}$. That is, $\mathcal{B}$ is a vector subspace of $\mathcal{A}$ such that $b_1 b_2 \in \mathcal{B}$ whenever b_1 and b_2 are in $\mathcal{B}$. Then by Example 2.1.6, $\mathcal{B}$ is a Lie algebra under the commutator product $[x, y] = xy - yx$, and is in fact a Lie subalgebra of $\mathcal{A}$.

For example, the vector space $T_n(\mathbb{F})$ of all upper triangular matrices

$$\begin{pmatrix} t_1 & & * \\ & \ddots & \\ 0 & & t_n \end{pmatrix}$$

is a subalgebra of the associative algebra $M_n(\mathbb{F})$. (For this, we just need to verify that the product of any two upper triangular matrices is upper triangular.) Hence, equipped with the commutator product, the Lie algebra $T_n(\mathbb{F})$ is a Lie subalgebra of the Lie algebra $\mathrm{gl}(n, \mathbb{F})$.

Definition 2.1.14. Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$ and let $\mathfrak{s}$ be a vector subspace of $\mathfrak{g}$. We say that $\mathfrak{s}$ is an *ideal* of $\mathfrak{g}$ if $[s, x] \in \mathfrak{s}$ whenever $s \in \mathfrak{s}$ and $x \in \mathfrak{g}$.

Thus, the ideal $\mathfrak{s}$ "absorbs" elements of $\mathfrak{g}$ under the Lie bracket. An ideal $\mathfrak{s}$ of a Lie algebra $\mathfrak{g}$ is obviously a Lie subalgebra of $\mathfrak{g}$.

Example 2.1.15. Let $\mathrm{sl}(n, \mathbb{F})$ denote the set of all $n \times n$ matrices X with entries in $\mathbb{F}$ such that $\mathrm{tr}(X) = 0$. Since the trace map

$$\begin{aligned} \mathrm{gl}(n, \mathbb{F}) &\to \mathbb{F} \\ X &\mapsto \mathrm{tr}(X) \end{aligned}$$

is a surjective linear functional on $\mathrm{gl}(n, \mathbb{F})$, we see that $\mathrm{sl}(n, \mathbb{F}) = \ker(\mathrm{tr})$ is a vector subspace of $\mathrm{gl}(n, \mathbb{F})$, of dimension $n^2 - 1$. We claim that $\mathrm{sl}(n, \mathbb{F})$ is an ideal of $\mathrm{gl}(n, \mathbb{F})$. For this, we just need to verify that $\mathrm{tr}[X, Y] = 0$ whenever $X \in \mathrm{sl}(n, \mathbb{F})$ and $Y \in \mathrm{gl}(n, \mathbb{F})$.

But in fact, it turns out that for *any* X and *any* Y in $\mathrm{gl}(n, \mathbb{F})$, we have $\mathrm{tr}[X, Y] = \mathrm{tr}(XY - YX) = \mathrm{tr}(XY) - \mathrm{tr}(YX) = 0$! Thus $\mathrm{sl}(n, \mathbb{F})$ is an ideal of $\mathrm{gl}(n, \mathbb{F})$.

Example 2.1.16. Let V be a nonzero vector space over $\mathbb{F}$, and let $\mathrm{sl}(V) = \{T \in \mathrm{gl}(V) \,|\, \mathrm{tr}(T) = 0\}$. Then just as in the preceding example, it is easy to prove that $\mathrm{sl}(V)$ is an ideal of $\mathrm{gl}(V)$.

The next proposition gives rise to a large class of the so-called classical simple Lie algebras.

Proposition 2.1.17. *Let S be a nonsingular $n \times n$ matrix over $\mathbb{F}$. Then let*

$$\mathfrak{g} := \{X \in gl(n, \mathbb{F}) \,|\, S\, {}^tX\, S^{-1} = -X\} \tag{2.6}$$

Then $\mathfrak{g}$ is a Lie subalgebra of $gl(n, \mathbb{F})$. Moreover, $\mathfrak{g} \subset sl\,(n, \mathbb{F})$.

Proof. It is straightforward to check that $\mathfrak{g}$ is a subspace of $\mathrm{gl}(n, \mathbb{F})$. What's important is to prove that $\mathfrak{g}$ is closed under the Lie bracket in $\mathrm{gl}(n, \mathbb{F})$. That is, we must prove that $[X, Y] \in \mathfrak{g}$ whenever X and Y are in $\mathfrak{g}$.

But then, for X, $Y \in \mathfrak{g}$,

$$\begin{aligned}
S({}^t[X,Y])S^{-1} &= S({}^t(XY - YX))\, S^{-1} \\
&= S({}^tY\, {}^tX - {}^tX\, {}^tY)\, S^{-1} \\
&= S({}^tY\, {}^tX)\, S^{-1} - S({}^tX\, {}^tY)\, S^{-1} \\
&= (S\, {}^tY\, S^{-1})(S\, {}^tX\, S^{-1}) - (S\, {}^tX\, S^{-1})\, (S\, {}^tY\, S^{-1}) \\
&= (-Y)(-X) - (-X)(-Y) \\
&= (YX - XY) \\
&= -[X, Y],
\end{aligned}$$

which shows that $[X, Y]$ indeed belongs to $\mathfrak{g}$.

For any $X \in \mathfrak{g}$, we have $\mathrm{tr}\,(S\, {}^tX\, S^{-1}) = \mathrm{tr}\,(-X)$, which gives $\mathrm{tr}\, X = -\mathrm{tr}\, X$, and so $\mathrm{tr}\, X = 0$. Thus $\mathfrak{g}$ is a subalgebra of $\mathrm{sl}\,(n, \mathbb{F})$. ☐

Example 2.1.18. When we let $S = I_n$ in Proposition 2.1.17, we obtain the Lie algebra $\mathrm{so}(n, \mathbb{F}) = \{X \in \mathrm{gl}(n, \mathbb{F}) \,|\, {}^tX = -X\}$. $\mathrm{so}(n, \mathbb{F})$ consists of all *skew-symmetric* matrices in $\mathbb{F}$.

By convention, the real Lie algebra $\mathrm{so}\,(n, \mathbb{R})$ is often simply written as $\mathrm{so}\,(n)$.

Example 2.1.19. Let J_n be the $2n \times 2n$ matrix given in block form by (1.33):

$$J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

If we apply Proposition 2.1.17 to $S = J_n$, then we obtain the *symplectic Lie algebra* $\mathrm{sp}(n, \mathbb{F})$, given by

$$\mathrm{sp}(n, \mathbb{F}) := \{X \in \mathrm{gl}(2n, \mathbb{F}) \,|\, J_n\, {}^tX\, J_n^{-1} = -X\}$$

Exercise 2.1.20. (Easy exercise.) Show that if X is a $2n \times 2n$ matrix, then $X \in \mathrm{sp}(n, \mathbb{F})$ if and only if X has block form

$$\begin{pmatrix} A & B \\ C & -{}^tA \end{pmatrix}, \tag{2.7}$$

where A is an arbitrary $n \times n$ matrix, and B and C are symmetric $n \times n$ matrices over $\mathbb{F}$.

Example 2.1.21. Let $n = p + q$, where $p,\, q \in \mathbb{Z}^+$. If in Proposition 2.1.17, we let S be the $(p+q) \times (p+q)$ matrix which in block form is given by

$$S = I_{p,q} := \begin{pmatrix} -I_p & 0_{p\times q} \\ 0_{q\times p} & I_q \end{pmatrix}, \tag{2.8}$$

then we obtain the Lie subalgebra

$$\mathrm{so}\,(p,q,\mathbb{F}) = \{X \in \mathrm{gl}\,(p+q,\mathbb{F}) \,|\, I_{p,q}\, {}^tX\, I_{p,q} = -X\} \tag{2.9}$$

of $\mathrm{sl}\,(p+q,\mathbb{F})$. When $\mathbb{F} = \mathbb{R}$, this Lie algebra is denoted simply by $\mathrm{so}\,(p,q)$. Note that $\mathrm{so}\,(p,0,\mathbb{F}) = \mathrm{so}\,(0,p,\mathbb{F}) = \mathrm{so}\,(p,\mathbb{F})$.

We recall that the *adjoint*, or *transposed conjugate*, of a complex matrix X is the matrix $X^* = {}^t\overline{X}$.

Proposition 2.1.22. *Let S be a nonsingular complex $n \times n$ matrix, and let*

$$\mathfrak{g} = \{X \in gl\,(n,\mathbb{C}) \,|\, S\, X^*\, S^{-1} = -X\}. \tag{2.10}$$

Then $\mathfrak{g}$ is a Lie subalgebra of $gl\,(n,\mathbb{C})$.

The easy proof, which is quite similar to that of Proposition 2.1.17, will be omitted.

Example 2.1.23. In Proposition 2.1.22, if we let $S = I_n$, we get the Lie algebra

$$\mathrm{u}\,(n) = \{X \in \mathrm{gl}\,(n,\mathbb{C}) \,|\, X^* = -X\} \tag{2.11}$$

of *skew-Hermitian matrices*. Intersecting this with $\mathrm{sl}\,(n,\mathbb{C})$, we get the Lie algebra

$$\mathrm{su}\,(n) = \mathrm{u}\,(n) \cap \mathrm{sl}\,(n,\mathbb{C}) \tag{2.12}$$

of skew-Hermitian matrices of trace 0. (Here we are using the easily checked fact that the intersection of two Lie subalgebras of a Lie algebra $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{g}$.)

Example 2.1.24. If $n = p + q$ and if, in Proposition 2.1.22, we let $S = I_{p,q}$, as in equation 2.8, then we obtain the Lie subalgebra

$$\mathrm{u}(p,q) = \{X \in \mathrm{gl}\,(p+q,\mathbb{C}) \mid I_{p,q}\, X^*\, I_{p,q} = -X\} \tag{2.13}$$

of $\mathrm{gl}\,(p+q,\mathbb{C})$. The intersection $\mathrm{u}(p,q) \cap \mathrm{sl}\,(p+q,\mathbb{C})$ is denoted by $\mathrm{su}(p,q)$.

Exercise 2.1.25. Show that $\mathrm{sl}(n,\mathbb{F})$, $\mathrm{so}(n,\mathbb{F})$, and $\mathrm{sp}(n,\mathbb{F})$ are all invariant under the transpose map $X \mapsto {}^tX$. Show that $\mathrm{u}(n)$, $\mathrm{su}(n)$, $\mathrm{u}(p,q)$, and $\mathrm{su}(p,q)$ are all invariant under the adjoint map $X \mapsto X^*$.

Example 2.1.26. Suppose that $\langle\ ,\ \rangle$ is a bilinear form on a vector space V over $\mathbb{F}$. Let $\mathfrak{g}$ denote the set of all $T \in \mathcal{L}(V)$ which satisfies Leibniz' rule with respect to $\langle\ ,\ \rangle$:

$$\langle T(v), w\rangle + \langle v, T(w)\rangle = 0 \qquad\qquad \text{for all } v,\ w \in V$$

It is easy to check that $\mathfrak{g}$ is a vector subspace of $\mathcal{L}(V)$. Let us show that $\mathfrak{g}$ is a Lie subalgebra of $\mathrm{gl}(V)$ under the commutator product. Suppose that S and T are in $\mathfrak{g}$. Then for any V, $w \in V$, we have

$$\begin{aligned}
\langle (ST - TS)(v), w\rangle &= \langle ST(v), w\rangle - \langle TS(v), w\rangle \\
&= -\langle T(v), S(w)\rangle + \langle S(v), T(w)\rangle \\
&= \langle v, TS(w)\rangle - \langle v, ST(w)\rangle \\
&= -\langle v, (ST - TS)(w)\rangle
\end{aligned}$$

This shows that $[S,T] = ST - TS \in \mathfrak{g}$. The Lie algebra $\mathfrak{g}$ is sometimes denoted $\mathrm{so}\,(V)$.

Chapter 3

Basic Algebraic Facts

In this section we explore the basic algebraic properties satisfied by all Lie algebras. Many of these properties, properly formulated, are shared by general algebras.

3.1 Structure Constants.

Suppose that $\mathfrak{g}$ is a Lie algebra over $\mathbb{F}$, and that $B = (x_1, \ldots, x_n)$ is a basis of $\mathfrak{g}$. Then there exist unique scalars c_{ij}^k $(1 \leq k \leq n)$ such that

$$[x_i, x_j] = \sum_{k=1}^{n} c_{ij}^k x_k \tag{3.1}$$

for all $1 \leq i, j \leq n$. The scalars c_{ij}^k are called the *structure constants of* $\mathfrak{g}$ *relative to the given basis* B. Since every element of $\mathfrak{g}$ is a unique linear combination of the basis vectors in B, we see that the structure constants completely determine the Lie bracket $[x, y]$, for any x, $y \in \mathfrak{g}$.

From the anticommutativity (2.2) and the Jacobi identity (2.3), we see that the structure constants c_{ij}^k satsify the relations

$$c_{ji}^k = -c_{ij}^k \tag{3.2}$$

$$\sum_{r=1}^{n} (c_{ir}^m c_{jk}^r + c_{jr}^m c_{ki}^r + c_{kr}^m c_{ij}^r) = 0, \tag{3.3}$$

for all i, j, k, m.

Conversely, suppose that there exist n^3 constants c_{ij}^k in $\mathbb{F}$ satisfying the relations (3.2) and (3.3). Then it can be shown, by a straightforward computation, that if $\mathfrak{g}$ is a vector space with basis $B = (x_1, \ldots, x_n)$ and we define a bilinear binary operation $[\ ,\ ]$ on $\mathfrak{g}$ via (3.1), then this binary operation is anticommutative and satisfies the Jacobi identity.

Thus a Lie algebra $\mathfrak{g}$ is completely determined by its structure constants and the relations (3.1) on a given basis $(x_1, \ldots, x_n)$.

In this course, we will not be making much use of structure constants.

3.2 Quotient Algebras, Homomorphisms, and Isomorphisms.

Let $\mathfrak{h}$ and $\mathfrak{u}$ be subalgebras of a Lie algebra $\mathfrak{g}$. Then the subspace $\mathfrak{h} \cap \mathfrak{u}$ is easily checked to be a Lie subalgebra of $\mathfrak{g}$. In addition, if one of them is an ideal of $\mathfrak{g}$, then $\mathfrak{h} + \mathfrak{u}$ is a subalgebra of $\mathfrak{g}$.

If U and W are nonempty subsets of $\mathfrak{g}$, we define $[U, W]$ to be the subspace spanned by all $[u, w]$, where $u \in U$, $w \in W$. Thus a subspace U of $\mathfrak{g}$ is a subalgebra if and only if $[U, U] \subset U$.

Let $\mathfrak{a}$ be an ideal of a Lie algebra $\mathfrak{g}$. The quotient space $\mathfrak{g}/\mathfrak{a}$ has a (natural) Lie algebra structure, in which the Lie bracket is defined by

$$[x + \mathfrak{a}, y + \mathfrak{a}] = [x, y] + \mathfrak{a} \tag{3.4}$$

The binary operation on $\mathfrak{g}/\mathfrak{a}$ given by 3.4 is well-defined: if $x + \mathfrak{a} = x_1 + \mathfrak{a}$ and $y + \mathfrak{a} = y_1 + \mathfrak{a}$, then $x_1 - x \in \mathfrak{a}$ and $y_1 - y \in \mathfrak{a}$, so

$$\begin{aligned}[x_1, y_1] &= [x + (x_1 - x), y + (y_1 - y)] \\ &= [x, y] + [x_1 - x, y] + x, y_1 - y] + [x_1 - x, y_1 - y]\end{aligned}$$

Since the last three terms above belong to $\mathfrak{a}$, we therefore have $[x_1, y_1] + \mathfrak{a} = [x, y] + \mathfrak{a}$. It is easy to verify that the binary operation (3.4) on $\mathfrak{g}/\mathfrak{a}$ is a Lie bracket on $\mathfrak{g}/\mathfrak{a}$. We call the quotient space $\mathfrak{g}/\mathfrak{a}$, equipped with this Lie bracket, the *quotient Lie algebra of* $\mathfrak{g}$ *by* $\mathfrak{a}$.

Example 3.2.1. We saw in Example 2.1.15 that $\mathrm{sl}(n, \mathbb{F})$ is an ideal of $\mathrm{gl}(n, \mathbb{F})$, of dimension $n^2 - 1$. The quotient Lie algebra $\mathfrak{k} = \mathrm{gl}(n, \mathbb{F})/\mathrm{sl}(n, \mathbb{F})$ is one-dimensional, and so must be abelian. The fact that $\mathfrak{k}$ is abelian is also easy to see because $[X, Y] \in \mathrm{sl}(n, \mathbb{F})$ for all X, $Y \in \mathrm{gl}(n, \mathbb{F})$.

Let $\mathfrak{g}$ and $\mathfrak{m}$ be Lie algebras. A *homomorphism* from $\mathfrak{g}$ to $\mathfrak{m}$ is a linear map $\varphi : \mathfrak{g} \to \mathfrak{m}$ such that

$$\varphi[x, y] = [\varphi(x), \varphi(y)],$$

for all x, $y \in g$. An *isomorphism* is a one-to-one, onto, homomorphism. We say that Lie algebras $\mathfrak{g}$ and $\mathfrak{m}$ are *isomorphic*, written $\mathfrak{g} \cong \mathfrak{m}$, if there exists an isomorphism $\varphi : \mathfrak{g} \to \mathfrak{m}$. An *automorphism* of $\mathfrak{g}$ is an isomorphism of $\mathfrak{g}$ onto $\mathfrak{g}$.

As an example, if $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then the natural projection

$$\begin{aligned} \pi : &\mathfrak{g} \to \mathfrak{g}/\mathfrak{a} \\ &x \mapsto x + \mathfrak{a} \end{aligned}$$

is a surjective Lie algebra homomorphism.

If $\varphi : \mathfrak{g} \to \mathfrak{m}$ is a homomorphism, then $\ker \varphi$ is an ideal of $\mathfrak{g}$. In fact, if $x \in \ker \varphi$ and $y \in \mathfrak{g}$, then

$$\varphi[x, y] = [\varphi(x), \varphi(y)] = [0, \varphi(y)] = 0,$$

so $[x, y] \in \ker \varphi$.

Exercise 3.2.2. Show that if $\mathfrak{g}$ is a two-dimensional Lie algebra over $\mathbb{F}$, then either $\mathfrak{g}$ is abelian, or $\mathfrak{g}$ has a basis (x, y) such that $[x, y] = y$. Thus, up to isomorphism, there are only two Lie algebras of dimension 2 over $\mathbb{F}$.

Exercise 3.2.3. Let so(3) denote the Lie algebra of all 3×3 skew-symmetric matrices over $\mathbb{R}$. (See Example 2.1.18.) Find an explicit isomorphism from so(3) to the Lie algebra $\mathbb{R}^3$, equipped with the cross product.

Theorem 3.2.4. *(The Homomorphism Theorem) Let $\mathfrak{a}$ be an ideal of $\mathfrak{g}$, and let $\pi : \mathfrak{g} \to \mathfrak{g}/\mathfrak{a}$ be the natural projection. Suppose that $\varphi : \mathfrak{g} \to \mathfrak{m}$ is a homomorphism such that $\mathfrak{a} \subset \ker \varphi$. Then there is a unique homomorphism $\widetilde{\varphi} : \mathfrak{g}/\mathfrak{a} \to \mathfrak{m}$ satisfying $\widetilde{\varphi} \circ \pi = \varphi$.*

Proof. Let $\widetilde{\varphi} : \mathfrak{g}/\mathfrak{a} \to \mathfrak{m}$ be given by $\widetilde{\varphi}(x + \mathfrak{a}) = \varphi(x)$, for all $x \in \mathfrak{a}$. Since $\varphi(\mathfrak{a}) = 0$, $\widetilde{\varphi}$ is well-defined. It is a homomorphism since $\widetilde{\varphi}[x + \mathfrak{a}, y + \mathfrak{a}] = \widetilde{\varphi}([x, y] + \mathfrak{a}) = \varphi[x, y] = [\varphi(x), \varphi(y)] = [\widetilde{\varphi}(x + \mathfrak{a}), \widetilde{\varphi}(y + \mathfrak{a})]$. And it is unique, since the condition $\widetilde{\varphi} \circ \pi = \varphi$ means that, for all $x \in \mathfrak{g}$, $\widetilde{\varphi}(x + \mathfrak{a}) = \widetilde{\varphi} \circ \pi(x) = \varphi(x)$. □

Corollary 3.2.5. *Let $\varphi : \mathfrak{g} \to \mathfrak{m}$ be a Lie algebra homomorphism. Then the image $\varphi(\mathfrak{g})$ is a Lie subalgebra of $\mathfrak{m}$, and the resulting map $\widetilde{\varphi} : \mathfrak{g}/\ker \varphi \to \varphi(\mathfrak{g})$ is a Lie algebra isomorphism. Thus, if φ is onto, then $\widetilde{\varphi}$ is an isomorphism of $\mathfrak{g}/\ker \varphi$ onto $\mathfrak{m}$.*

Proof. For any x, $y \in \mathfrak{g}$, we have $[\varphi(x), \varphi(y)] = \varphi[x, y] \in \varphi(\mathfrak{g})$, so $\varphi(\mathfrak{g})$ is a subalgbra of $\mathfrak{m}$. Put $\mathfrak{a} = \ker \varphi$ in Theorem 3.2.4. Then $\widetilde{\varphi}$ is injective since if $\widetilde{\varphi}(x + \mathfrak{a}) = 0$, then $\varphi(x) = 0$, so $x \in \mathfrak{a}$, and thus $x + \mathfrak{a} = \mathfrak{a}$. Thus the map $\widetilde{\varphi} : \mathfrak{g}/\ker \varphi \to \varphi(\mathfrak{g})$ is an isomorphism. □

If $\ker \varphi = \{0\}$, then $\mathfrak{g}$ is isomorphic to its image $\varphi(\mathfrak{g})$ in $\mathfrak{m}$. In this case, we say that $\mathfrak{g}$ is *embedded in* $\mathfrak{m}$.

Theorem 3.2.6. *(The Correspondence Theorem) Let $\varphi : \mathfrak{g} \to \mathfrak{m}$ be a surjective homomorphism.*

1. *If $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then $\varphi(\mathfrak{a})$ is an ideal of $\mathfrak{m}$.*
2. *If $\mathfrak{s}$ is an ideal of $\mathfrak{m}$, the $\varphi^{-1}(\mathfrak{s})$ is an ideal of $\mathfrak{g}$ which contains $\ker\varphi$.*
3. *The mappings $\mathfrak{a} \mapsto \varphi(\mathfrak{a})$ and $\mathfrak{s} \mapsto \varphi^{-1}(\mathfrak{s})$ are inverse mappings between the set of all ideals of $\mathfrak{g}$ which contain $\ker\varphi$ and the set of all ideals of $\mathfrak{m}$, so the two sets of ideals are in one-to-one correspondence.*
4. *$\mathfrak{g}/\mathfrak{a} \cong \mathfrak{m}/\varphi(\mathfrak{a})$ for all ideals $\mathfrak{a}$ of $\mathfrak{g}$ containing $\ker\varphi$.*
5. *The correspondence in (3) preserves inclusion:*
$$\ker\varphi \subset \mathfrak{a}_1 \subset \mathfrak{a}_2 \iff \varphi(\mathfrak{a}_1) \subset \varphi(\mathfrak{a}_2).$$

Proof.

1. For any $y \in \mathfrak{m}$ and $v \in \mathfrak{a}$, we have $y = \varphi(x)$ for some $x \in \mathfrak{g}$, so
$$\begin{aligned}[y, \varphi(v)] &= [\varphi(x), \varphi(v)] \\ &= \varphi[x, v] \in \varphi(\mathfrak{a}).\end{aligned}$$
Hence $[\mathfrak{m}, \varphi(\mathfrak{a})] \subset \varphi(\mathfrak{a})$, and $\varphi(\mathfrak{a})$ is an ideal of $\mathfrak{m}$.
2. Let $v \in \varphi^{-1}(\mathfrak{s})$. Then for any $x \in \mathfrak{g}$, we have
$$\varphi[x, v] = [\varphi(x), \varphi(v)] \in [\mathfrak{m}, \mathfrak{s}] \subset \mathfrak{s},$$
so $[x, v] \in \varphi^{-1}(\mathfrak{s})$.
3. We first claim that if $\mathfrak{a}$ is an ideal of $\mathfrak{g}$ containing $\ker\varphi$, then $\varphi^{-1}(\varphi(\mathfrak{a})) = \mathfrak{a}$. Since clearly $\mathfrak{a} \subset \varphi^{-1}(\varphi(\mathfrak{a}))$, it suffices to prove that $\varphi^{-1}(\varphi(\mathfrak{a})) \subset \mathfrak{a}$. But if $x \in \varphi^{-1}(\varphi(\mathfrak{a}))$, then $\varphi(x) = \varphi(v)$ for some $v \in \mathfrak{a}$, so $x - v \in \ker\varphi$, and hence $x \in v + \ker\varphi \subset \mathfrak{a} + \mathfrak{a} = \mathfrak{a}$.

   Next, it is clear from the surjectivity of φ that if A is any subset of $\mathfrak{m}$, then $\varphi(\varphi^{-1}(A)) = A$. Thus, in particular, if $\mathfrak{s}$ is an ideal of $\mathfrak{m}$, then $\varphi(\varphi^{-1}(\mathfrak{s})) = \mathfrak{s}$.

   From the above, we see that $\mathfrak{a} \mapsto \varphi(\mathfrak{a})$ is a bijection between the sets in question, with inverse $\mathfrak{s} \mapsto \varphi^{-1}(\mathfrak{s})$.
4. Consider the following diagram of homomorphisms
$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\varphi"] \arrow[d, "\pi"'] & \mathfrak{m} \arrow[d, "p"] \\ \mathfrak{g}/\mathfrak{a} \arrow[r, dotted, "\varphi_1"'] & \mathfrak{m}/\varphi(\mathfrak{a}) \end{tikzcd} \tag{3.5}$$

where π and p are projections. Now $p \circ \varphi$ is a homomorphism of $\mathfrak{g}$ onto $\mathfrak{m}/\varphi(\mathfrak{a})$ whose kernel is $\varphi^{-1}(\varphi(\mathfrak{a})) = \mathfrak{a}$. Hence by the Homomorphism Theorem (Theorem 3.2.4)and its corollary, there is a isomorphism φ_1 from $\mathfrak{g}/\mathfrak{a}$ onto $\mathfrak{m}/\varphi(\mathfrak{a})$ such that $\varphi_1 \circ \pi = p \circ \varphi$.

5. Obvious.

$\square$

Theorem 3.2.7. *(The Isomorphism Theorem) If $\mathfrak{g}$ is a Lie algebra, and $\mathfrak{s}$ a subalgebra and $\mathfrak{a}$ an ideal of $\mathfrak{g}$, then*

1. *$\mathfrak{s} \cap \mathfrak{a}$ is an ideal of $\mathfrak{s}$*
2. *$\mathfrak{s}/(\mathfrak{s} \cap \mathfrak{a}) \cong (\mathfrak{s} + \mathfrak{a})/\mathfrak{a}$*
3. *If $\mathfrak{b}$ is an ideal of $\mathfrak{g}$ such that $\mathfrak{b} \subset \mathfrak{a}$, then $(\mathfrak{g}/\mathfrak{b})/(\mathfrak{a}/\mathfrak{b}) \cong \mathfrak{g}/\mathfrak{a}$.*

Proof. 1. Easy.

2. We already know that $\mathfrak{s} + \mathfrak{a}$ is a subalgebra of $\mathfrak{g}$. Consider the diagram

$$\begin{tikzcd} \mathfrak{s} \arrow[r, "\pi"] \arrow[d, "i"'] & \mathfrak{s}/(\mathfrak{s} \cap \mathfrak{a}) \arrow[d, dotted, "i'"] \\ \mathfrak{s} + \mathfrak{a} \arrow[r, "p"] & (\mathfrak{s} + \mathfrak{a})/\mathfrak{a} \end{tikzcd}$$

where π and p are projections, and i is the inclusion map of $\mathfrak{s}$ into $\mathfrak{s} + \mathfrak{a}$. $p \circ i$ is obviously a homomorphism, and it is surjective, since any element of $(\mathfrak{s} + \mathfrak{a})/\mathfrak{a}$ is of the form $v + w + \mathfrak{a}$, where $v \in \mathfrak{s}$ and $w \in \mathfrak{a}$, and this element is of course the same as $v + \mathfrak{a} = p \circ i(v)$. The kernel of $p \circ i$ is $\{v \in \mathfrak{s} \,|\, v + \mathfrak{a} = \mathfrak{a}\} = \{v \in \mathfrak{s} \,|\, v \in \mathfrak{a}\} = \mathfrak{s} \cap \mathfrak{a}$. Thus by the Homomorphism Theorem, the resulting map $i' : \mathfrak{s}/(\mathfrak{s} \cap \mathfrak{a}) \to (\mathfrak{s} + \mathfrak{a})/\mathfrak{a}$ is an isomorphism.

3. Consider the map $h : \mathfrak{g}/\mathfrak{b} \to \mathfrak{g}/\mathfrak{a}$ given by $x + \mathfrak{b} \mapsto x + \mathfrak{a}$. h is well-defined, since $\mathfrak{b} \subset \mathfrak{a}$ and is easily checked to be a surjective Lie algebra homomorphism. Its kernel is the ideal $\{x + \mathfrak{b} \,|\, x + \mathfrak{a} = \mathfrak{a}\} = \{x + \mathfrak{b} \,|\, x \in \mathfrak{a}\} = \mathfrak{a}/\mathfrak{b}$ of $\mathfrak{g}/\mathfrak{b}$. Thus, by the Homomorphism Theorem, the algebras $(\mathfrak{g}/\mathfrak{b})/(\mathfrak{a}/\mathfrak{b})$ and $\mathfrak{g}/\mathfrak{a}$ are isomorphic.

$\square$

Exercise 3.2.8. (Graduate Exercise.) Suppose that φ is an *involution* of a Lie algebra $\mathfrak{g}$; i.e., an automorphism φ of $\mathfrak{g}$ such that $\varphi^2 = I_{\mathfrak{g}}$. Let $\mathfrak{h} = \{x \in \mathfrak{g} \,|\, \varphi(x) = x\}$ be the $+1$ eigenspace and $\mathfrak{q} = \{x \in \mathfrak{g} \,|\, \varphi(x) = -x\}$ the -1-eigenspace of φ, respectively. Prove that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{q}$, and that $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$ (so that $\mathfrak{h}$ is a subalgebra), $[\mathfrak{h}, \mathfrak{q}] \subset \mathfrak{q}$, $[\mathfrak{q}, \mathfrak{q}] \subset \mathfrak{h}$.

Definition 3.2.9. A *representation* of a Lie algebra $\mathfrak{g}$ on a vector space V is a Lie algebra homomorphism $\pi : \mathfrak{g} \to \mathrm{gl}(V)$. V is called the *representation space* of π. If there is a representation of $\mathfrak{g}$ on V, then we say that $\mathfrak{g}$ *acts on* V.

If π is a representation of $\mathfrak{g}$ on V, then of course $\pi(x)$ is a linear map on V for any $x \in \mathfrak{g}$.

Example 3.2.10. If V is any vector space, then $\mathrm{gl}(V)$ acts on V, via the identity map $\mathrm{id} : \mathrm{gl}(V) \to \mathrm{gl}(V)$. Any Lie subalgebra $\mathfrak{g}$ of $\mathrm{gl}(V)$ likewise acts on V, via the inclusion map $\iota : \mathfrak{g} \hookrightarrow \mathrm{gl}(V)$. This action is called the *standard representation* of $\mathfrak{g}$ on V.

Note that $\mathrm{sl}(n, \mathbb{F})$, $\mathrm{so}(n, \mathbb{F})$ and $\mathrm{sp}(n, \mathbb{F})$ ((Examples 2.1.15), 2.1.18, and 2.1.19, respectively) are Lie algebras of matrices acting on $\mathbb{F}^n$, $\mathbb{F}^n$, and $\mathbb{F}^{2n}$ under their respective standard representations.

Example 3.2.11. The *trivial representation* of $\mathfrak{g}$ on V is the map $\pi : \mathfrak{g} \to \mathrm{gl}(V)$ such that $\pi(x) = 0$, for all $x \in \mathfrak{g}$.

In *representation theory*, one studies representations of Lie algebras, and their associated Lie groups, on finite and infinite-dimensional vector spaces. Representation theory has intimate connections to number theory, physics, differential and symplectic geometry, and harmonic and geometric analysis, to mention but a few fields. It is an extremely active and vibrant field of mathematics.

3.3 Centers, Centralizers, Normalizers, and Simple Lie Algebras

The *center* of a Lie algebra $\mathfrak{g}$ is the set $\mathfrak{c} = \{c \in \mathfrak{g} \,|\, [c, v] = 0 \text{ for all } v \in \mathfrak{g}\}$. It is obvious that $\mathfrak{c}$ is an ideal of $\mathfrak{g}$. If A is a nonempty subset of $\mathfrak{g}$, then the *centralizer* of A is the set $\mathfrak{c}(A) = \{x \in \mathfrak{g} \,|\, [x, v] = 0 \text{ for all } v \in A\}$. It is easily checked that $\mathfrak{c}(A)$ is a subspace of $\mathfrak{g}$. Note that $\mathfrak{c} = \mathfrak{c}(\mathfrak{g})$.

Proposition 3.3.1. *Let A be a nonempty subset of $\mathfrak{g}$. Then its centralizer $\mathfrak{c}(A)$ is a subalgebra of $\mathfrak{g}$.*

Proof. This is an immediate consequence of the Jacobi identity. Let x, $y \in \mathfrak{c}(A)$ and let $a \in A$. Then

$$[[x, y], a] = -[[y, a], x] - [[a, x], y] = -[0, x] - [0, y] = 0.$$

$\square$

Proposition 3.3.2. *If $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then its centralizer $\mathfrak{c}(\mathfrak{a})$ is an ideal of $\mathfrak{g}$.*

Proof. Another immediate consequence of the Jacobi identity: let $c \in \mathfrak{c}(\mathfrak{a})$, $v \in \mathfrak{g}$, and $x \in \mathfrak{a}$. Then

$$[[c, v], x] = -[[v, x], c] - [[x, c], v]$$

But $c \in \mathfrak{c}$, so $[x, c] = 0$ and $[v, x] \in \mathfrak{a}$ so $[[v, x], c] = 0$. $\square$

Proposition 3.3.3. *Let φ be a surjective Lie algebra homomorphism of $\mathfrak{g}$ onto $\mathfrak{m}$. If $\mathfrak{c}$ denotes the center of $\mathfrak{g}$, then its image $\varphi(\mathfrak{c})$ lies in the center of $\mathfrak{m}$.*

Proof. $\mathfrak{m} = \varphi(\mathfrak{g})$, so $[\mathfrak{m}, \varphi(\mathfrak{c})] = [\varphi(\mathfrak{g}), \varphi(\mathfrak{c})] = \varphi([\mathfrak{g}, \mathfrak{c}]) = \varphi(\{0\}) = \{0\}$. $\square$

If $\mathfrak{s}$ is a subalgebra of $\mathfrak{g}$, its *normalizer* is the set $\mathfrak{n}(\mathfrak{s}) = \{x \in \mathfrak{g} \mid [x, v] \in \mathfrak{s} \text{ for all } v \in \mathfrak{s}\}$. $\mathfrak{n}(\mathfrak{s})$ is clearly a subspace of $\mathfrak{g}$ containing $\mathfrak{s}$, and the Jacobi identity shows that it is in fact a subalgebra of $\mathfrak{g}$:

Proposition 3.3.4. *$\mathfrak{n}(\mathfrak{s})$ is a subalgebra of $\mathfrak{g}$.*

Proof. Let x and y be in $\mathfrak{n}(\mathfrak{s})$, and let $s \in \mathfrak{s}$. Then

$$[[x, y], s] = -[[y, s], x] - [[s, x], y] \ \in \ -[\mathfrak{s}, x] - [\mathfrak{s}, y] \ \subset \mathfrak{s} + \mathfrak{s} = \mathfrak{s}.$$

$\square$

One more thing: it is easy to see that $\mathfrak{n}(\mathfrak{s})$ is the largest subalgebra of $\mathfrak{g}$ for which $\mathfrak{s}$ is an ideal.

Example 3.3.5. Let $\mathfrak{g} = \mathrm{sl}(2, \mathbb{C}) = \{X \in \mathrm{gl}(2, \mathbb{C}) \mid \mathrm{tr}(X) = 0\}$. Its *standard basis* is given by

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{3.6}$$

The commutation relations among these basis elements is given by

$$\begin{aligned} [h, e] &= he - eh = 2e \\ [h, f] &= hf - fh = -2f \\ [e, f] &= ef - fe = h \end{aligned} \tag{3.7}$$

From the above, we see that the one-dimensional algebra $\mathbb{C}h$ is its own normalizer. ($\mathbb{C}h$ is called a *Cartan subalgebra* of $\mathfrak{g}$.) We will see later on that the commutation relations (3.7) play a key role in the structure and representation theory of semisimple Lie algebras.

Definition 3.3.6. A Lie algebra $\mathfrak{g}$ is said to be *simple* if $\mathfrak{g}$ is non-abelian and $\mathfrak{g}$ has no ideals except $\{0\}$ and $\mathfrak{g}$.

Example 3.3.7. $\mathrm{sl}(2, \mathbb{C})$ is simple. Suppose that $\mathfrak{a} \neq \{0\}$ is an ideal of $\mathrm{sl}(2, \mathbb{C})$. Let $v \neq 0$ be an element of $\mathfrak{a}$, and write $v = \alpha e + \beta f + \gamma h$, where not all of α, β or γ are 0.

Assume that $\alpha \neq 0$. Then from the commutation relations (3.7), $[v, f] = \alpha h - 2\gamma f \in \mathfrak{a}$, and so $[[v, f], f] = -2\alpha f \in \mathfrak{a}$. Hence $f \in \mathfrak{a}$, and so $h = [e, f] \in \mathfrak{a}$ and also $e = 1/2[h, e] \in \mathfrak{a}$. Thus $\alpha \neq 0$ implies $\mathfrak{a} = \mathfrak{g}$. A similar argument shows that $\beta \neq 0$ implies $\mathfrak{a} = \mathfrak{g}$.

Finally, if $\gamma \neq 0$, then $[v, e] = -\beta h + 2\gamma e \in \mathfrak{a}$, so the argument in the preceding paragraph shows that $\mathfrak{a} = \mathfrak{g}$.

One of the goals of this course is to obtain Cartan's classification of all simple Lie algebras over $\mathbb{C}$. These consist of four (infinite) classes – the so-called classical simple Lie algebras – and five so-called exceptional Lie algebras.

3.4 The Adjoint Representation.

For each $x \in \mathfrak{g}$ define the linear operator $\mathrm{ad}\ x : \mathfrak{g} \to \mathfrak{g}$ by $\mathrm{ad}\ x\ (y) = [x, y]$, for all $y \in \mathfrak{g}$. The Jacobi identity shows that $\mathrm{ad}\ x$ is a derivation of $\mathfrak{g}$, since for all $u, v \in \mathfrak{g}$, we have

$$\begin{aligned}
\mathrm{ad}\ x\ [u, v] &= [x, [u, v]] \\
&= -[u, [v, x]] - [v, [x, u]] \\
&= [u, [x, v]] + [[x, u], v] \\
&= [u, \mathrm{ad}\ x\ (v)] + [\mathrm{ad}\ x\ (u), v].
\end{aligned}$$

Example 3.4.1. Let $\mathfrak{g} = \mathrm{sl}(2, \mathbb{C})$, with standard basis (e, f, h) given by (3.6). Using the commutation relations (3.7), we see that the matrices of $\mathrm{ad}\, e$, $\mathrm{ad}\, f$, and $\mathrm{ad}\, h$ with respect to the standard basis are:

$$\mathrm{ad}\, e = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \mathrm{ad}\, f = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 0 \end{pmatrix}, \qquad \mathrm{ad}\, h = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Proposition 3.4.2. *The map $x \mapsto$ $\mathrm{ad}\, x$ is a homomorphism of $\mathfrak{g}$ into the Lie algebra Der $\mathfrak{g}$.*

Proof. First, we show that $\mathrm{ad}\ (x + y) = \mathrm{ad}\ x + \mathrm{ad}\ y$. Now for all $z \in \mathfrak{g}$, we have

$$\begin{aligned}
\mathrm{ad}\ (x + y)\ (z) &= [x + y, z] = [x, z] + [y, z] \\
&= \mathrm{ad}\ x\ (z) + \mathrm{ad}\ y\ (z) \\
&= (\mathrm{ad}\ x + \mathrm{ad}\ y)(z).
\end{aligned}$$

Similarly, ad $(\alpha x) = \alpha(\text{ad } x)$, since for all $z \in \mathfrak{g}$, we have ad (αx) $(z) = [\alpha x, z] = \alpha[x, z] = \alpha$ ad x (z).

Finally, we prove that ad $[x, y] = [\text{ad } x, \text{ad } y]$ for all x and y in $\mathfrak{g}$. For any $z \in \mathfrak{g}$, we have

$$\begin{aligned}
\text{ad } [x, y] \ (z) &= [[x, y], z] \\
&= -[[y, z], x] - [[z, x], y] \\
&= [x, [y, z]] - [y, [x, z]] \\
&= \text{ad } x \ (\text{ad } y \ (z)) - \text{ad } y \ (\text{ad } x \ (z)) \\
&= (\text{ad } x \circ \text{ad } y - \text{ad } y \circ \text{ad } x)(z) \\
&= [\text{ad } x, \text{ad } y](z)
\end{aligned}$$

□

Since the map $\text{ad}\colon \mathfrak{g} \to \text{Der}\, \mathfrak{g} \subset \text{gl}(\mathfrak{g})$ is a homomorphism, we see that ad is a representation of $\mathfrak{g}$ on itself. For this reason, it is called the *adjoint representation* of $\mathfrak{g}$.

Let ad $\mathfrak{g}$ denote the subspace of Der $\mathfrak{g}$ consisting of all ad x, for all $x \in \mathfrak{g}$.

Proposition 3.4.3. *ad $\mathfrak{g}$ is an ideal of Der $\mathfrak{g}$.*

Proof. The proposition will follow once we prove that

$$[D, \text{ad}\, x] = \text{ad}\, (Dx) \tag{3.8}$$

for all $x \in \mathfrak{g}$ and $D \in \text{Der}\, \mathfrak{g}$. But for any $y \in \mathfrak{g}$, we have

$$\begin{aligned}
[D, \text{ad}\, x](y) &= (D \circ \text{ad}\, x \ - \ \text{ad}\, x \circ D)(y) \\
&= D[x, y] - [x, Dy] \\
&= [Dx, y] \\
&= \text{ad}\, (Dx) \ (y), \qquad (3.9)
\end{aligned}$$

since D is a derivation of $\mathfrak{g}$. □

We say that $\mathfrak{g}$ is *complete* if $\text{ad}\, \mathfrak{g} = \text{Der}\, \mathfrak{g}$.

The kernel of the adjoint representation $x \mapsto \text{ad}\, x$ of $\mathfrak{g}$ into $\text{Der}\, \mathfrak{g}$ consists of all $x \in \mathfrak{g}$ such that $\text{ad}\, x = 0$; i.e., all x such that $[x, y] = 0$ for all $y \in \mathfrak{g}$. This is precisely the center $\mathfrak{c}$ of $\mathfrak{g}$.

Corollary 3.4.4. *If the center $\mathfrak{c}$ of $\mathfrak{g}$ is $\{0\}$ (e.g., when $\mathfrak{g}$ is simple), then ad maps $\mathfrak{g}$ isomorphically onto the ideal $\text{ad}\, \mathfrak{g}$ of $\text{Der}\, \mathfrak{g}$, so $\mathfrak{g}$ is embedded in $\text{Der}\, \mathfrak{g}$.*

Proof. $\mathfrak{c}$ is the kernel of the adjoint representation, so by the Corollary 3.2.5, ad maps $\mathfrak{g}$ isomorphically onto $\operatorname{ad}\mathfrak{g}$. ☐

Proposition 3.4.5. *If $\mathfrak{c} = \{0\}$, then the centralizer $\mathfrak{c}(ad\ \mathfrak{g})$ in $Der\,\mathfrak{g}$ is $\{0\}$, hence $\mathfrak{c}(Der\,\mathfrak{g}) = \{0\}$, so $Der\,\mathfrak{g}$ is embedded in $Der\,(Der\,\mathfrak{g})$.*

Proof. Suppose that $D \in \mathfrak{c}(\operatorname{ad}\mathfrak{g})$. Then for all $x \in \mathfrak{g}$, we have by by (3.9), $0 = [D, \operatorname{ad} x] = \operatorname{ad}(Dx)$. Since $\mathfrak{c} = \{0\}$, we see that $Dx = 0$ for all $x \in \mathfrak{g}$, and hence $D = 0$. Now since the elements of the center $\mathfrak{c}(\operatorname{Der}\mathfrak{g})$ kill everything in $\operatorname{Der}\mathfrak{g}$, we see that $\mathfrak{c}(\operatorname{Der}\mathfrak{g}) \subset \mathfrak{c}(\operatorname{ad}\mathfrak{g})$, whence $\mathfrak{c}(\operatorname{Der}\mathfrak{g}) = \{0\}$. By the previous corollary, this implies that $\operatorname{Der}\mathfrak{g}$ is embedded in $\operatorname{Der}(\operatorname{Der}\mathfrak{g})$. ☐

So, amusingly, if the center $\mathfrak{c}$ of $\mathfrak{g}$ is $\{0\}$, we have an increasing chain of Lie algebras

$$\mathfrak{g} \subset \operatorname{Der}\mathfrak{g} \subset \operatorname{Der}(\operatorname{Der}\mathfrak{g}) \subset \operatorname{Der}(\operatorname{Der}(\operatorname{Der}\mathfrak{g})) \subset \cdots$$

According to a theorem by Schenkmann, this chain eventually stops growing. (See Schenkmann's paper [9] or Jacobson's book ([6] p.56, where it's an exercise) for details.)

Definition 3.4.6. Let $\mathfrak{a}$ and $\mathfrak{b}$ be ideals of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$. We say that $\mathfrak{g}$ is a *direct sum* of the ideals $\mathfrak{a}$ and $\mathfrak{b}$.

Exercise 3.4.7. If $\mathfrak{c} = \{0\}$, then show that $\mathfrak{g}$ is complete if and only if $\mathfrak{g}$ is a direct summand of any Lie algebra $\mathfrak{m}$ which contains $\mathfrak{g}$ as an ideal: $\mathfrak{m} = \mathfrak{g} \oplus \mathfrak{h}$, where $\mathfrak{h}$ is another ideal of $\mathfrak{m}$.

Exercise 3.4.8. If $\mathfrak{g}$ is simple, show that $\operatorname{Der}\mathfrak{g}$ is complete.

Chapter 4

Solvable Lie Algebras and Lie's Theorem

4.1 Solvable Lie Algebras

Definition 4.1.1. The *derived algebra* of a Lie algebra $\mathfrak{g}$ is $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$.

$\mathfrak{g}'$ is an ideal of $\mathfrak{g}$, since $\mathfrak{g}'$ is spanned by the products $[x, y]$, for all x, $y \in \mathfrak{g}$, and $[[x, y], z] \in [\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}'$, for all x, y, $z \in \mathfrak{g}$.

We could abbreviate the argument that $\mathfrak{g}'$ is an ideal by writing $[\mathfrak{g}', \mathfrak{g}] = [[\mathfrak{g}, \mathfrak{g}], \mathfrak{g}] \subset [\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}'$.

Theorem 4.1.2. *$\mathfrak{g}/\mathfrak{g}'$ is abelian, and for any ideal $\mathfrak{a}$, $\mathfrak{g}/\mathfrak{a}$ is abelian $\iff$ $\mathfrak{g}' \subset \mathfrak{a}$.*

Proof. For any x, $y \in \mathfrak{g}$, we have $[x + \mathfrak{g}', y + \mathfrak{g}'] = [x, y] + \mathfrak{g}' = \mathfrak{g}'$, so $\mathfrak{g}/\mathfrak{g}'$ is abelian.

Let us now prove the second assertion. Now

$$\begin{aligned} \mathfrak{g}/\mathfrak{a} \text{ is abelian} &\iff [x + \mathfrak{a}, y + \mathfrak{a}] = [x, y] + \mathfrak{a} = \mathfrak{a} \text{ for all } x, y \in \mathfrak{g} \\ &\iff [x, y] \in \mathfrak{a} \text{ for all } x, y \in \mathfrak{g} \\ &\iff \mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{a} \end{aligned}$$

□

Definition 4.1.3. A *characteristic ideal* of $\mathfrak{g}$ is an ideal $\mathfrak{a}$ such that $D(\mathfrak{a}) \subset \mathfrak{a}$ for every derivation D of $\mathfrak{g}$.

The derived ideal $\mathfrak{g}'$ is a characteristic ideal: for every derivation D, we have $D(\mathfrak{g}') = D[\mathfrak{g}, \mathfrak{g}] = [D(\mathfrak{g}), \mathfrak{g}] + [\mathfrak{g}, D(\mathfrak{g})] \subset [\mathfrak{g}, \mathfrak{g}] + [\mathfrak{g}, \mathfrak{g}] = [\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}'$.

Proposition 4.1.4. *Let $\mathfrak{h}$ be any vector subspace of the Lie algebra $\mathfrak{g}$ such that $\mathfrak{g}' \subset \mathfrak{h}$. Then $\mathfrak{h}$ is an ideal of $\mathfrak{g}$.*

Proof. We have $[\mathfrak{g}, \mathfrak{h}] \subset [\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}' \subset \mathfrak{h}$. □

Definition 4.1.5. The *derived series* of $\mathfrak{g}$ is

$$\mathfrak{g}^{(0)} = \mathfrak{g} \supset \mathfrak{g}^{(1)} = \mathfrak{g}' \supset \mathfrak{g}^{(2)} = \mathfrak{g}'' = [\mathfrak{g}', \mathfrak{g}'] \supset \cdots \supset \mathfrak{g}^{(i)} \supset \mathfrak{g}^{(i+1)} = [\mathfrak{g}^{(i)}, \mathfrak{g}^{(i)}] \supset \cdots \tag{4.1}$$

Proposition 4.1.6. *The derived series consists of a decreasing sequence of characteristic ideals of $\mathfrak{g}$.*

Proof. We need to prove that each $\mathfrak{g}^{(i)}$ is a characteristic ideal of $\mathfrak{g}$. This is done by induction on i, noting that there is nothing to prove for $i = 0$, and that we have already proved that $\mathfrak{g}^{(1)} = \mathfrak{g}'$ is a characteristic ideal in the remark before Definition 4.1.5. So assume that $\mathfrak{g}^{(i)}$ is a characteristic ideal of $\mathfrak{g}$. Then by the Jacobi identity and the induction hypothesis,

$$[\mathfrak{g}, \mathfrak{g}^{(i+1)}] = [\mathfrak{g}, [\mathfrak{g}^{(i)}, \mathfrak{g}^{(i)}]] \subset [\mathfrak{g}^{(i)}, [\mathfrak{g}, \mathfrak{g}^{(i)}]] + [\mathfrak{g}^{(i)}, [\mathfrak{g}^{(i)}, \mathfrak{g}]] \subset [\mathfrak{g}^{(i)}, \mathfrak{g}^{(i)}] = \mathfrak{g}^{(i+1)}.$$

It follows that $\mathfrak{g}^{(i+1)}$ is an ideal of $\mathfrak{g}$. Next, for any $D \in \operatorname{Der}\mathfrak{g}$, we have

$$\begin{aligned} D(\mathfrak{g}^{(i+1)}) &= D[\mathfrak{g}^{(i)}, \mathfrak{g}^{(i)}] \\ &= [D(\mathfrak{g}^{(i)}), \mathfrak{g}^{(i)}] + [\mathfrak{g}^{(i)}, D(\mathfrak{g}^{(i)})] \\ &\subset [\mathfrak{g}^{(i)}, \mathfrak{g}^{(i)}] + [\mathfrak{g}^{(i)}, \mathfrak{g}^{(i)}] \qquad \text{(by the induction hypothesis)} \\ &= \mathfrak{g}^{(i+1)}. \end{aligned}$$

□

Definition 4.1.7. $\mathfrak{g}$ is said to be *solvable* if $\mathfrak{g}^{(k)} = \{0\}$ for some $k \in \mathbb{Z}^+$.

Exercise 4.1.8. Show that $\mathfrak{g}$ is solvable if and only if there is a nested sequence of ideals $\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_m = \{0\}$ such that $\mathfrak{g}_{i+1}$ is an ideal of $\mathfrak{g}_i$ and $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ is abelian.

Note that no simple Lie algebra can be solvable. In fact, if $\mathfrak{g}$ is simple, then $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ is a nonzero ideal of $\mathfrak{g}$ (since $\mathfrak{g}$ is, by definition, non-abelian); hence $\mathfrak{g}' = \mathfrak{g}$. Thus $\mathfrak{g}'' = \mathfrak{g}' = \mathfrak{g}$, etc, and the derived series is constant. In particular, $\mathrm{sl}(2, \mathbb{C})$ is not solvable.

Example 4.1.9. Let $\mathfrak{g} = T_n(\mathbb{F})$ be the vector space of upper triangular $n \times n$ matrices over $\mathbb{F}$. If A and B are upper triangular matrices

$$A = \begin{pmatrix} s_1 & & & * \\ 0 & s_2 & & \\ & & \ddots & \\ 0 & & & s_n \end{pmatrix}, \qquad B = \begin{pmatrix} t_1 & & & * \\ 0 & t_2 & & \\ & & \ddots & \\ 0 & & & t_n \end{pmatrix}$$

then the product AB has the form

$$AB = \begin{pmatrix} s_1t_1 & & & * \\ 0 & s_2t_2 & & \\ & & \ddots & \\ 0 & & & s_nt_n \end{pmatrix}$$

and likewise, BA has the same form. Hence the commutator $AB - BA$ is strictly upper triangular

$$AB - BA = \begin{pmatrix} 0 & & & * \\ & 0 & & \\ & & \ddots & \\ 0 & & & 0 \end{pmatrix} \tag{4.2}$$

Thus the elements of $\mathfrak{g}'$ consist of strictly upper triangular matrices. With a bit of thought, one can see that the elements of $\mathfrak{g}''$ are matrices whose entries are 0's below the diagonal 2 steps above the main diagonal; that is, $\mathfrak{g}''$ consists of matrices (a_{ij}) such that $a_{ij} = 0$ whenever $i \geq j - 1$.

$$\begin{pmatrix} 0 & 0 & * & & * & * \\ & 0 & 0 & * & & * \\ & & 0 & 0 & & \\ & & & \ddots & \ddots & * \\ & & & & 0 & 0 \\ 0 & & & & & 0 \end{pmatrix}$$

The $\mathfrak{g}^{(3)}$ matrices have 0's below the diagonal 2^2 steps above the main diagonal. Generally, $\mathfrak{g}^{(i)}$ matrices have 0's below the diagonal 2^{i-1} steps above the main diagonal.

We can also use Exercise 4.1.8 to show that $T_n(\mathbb{F})$ is solvable. First, for any i, j, let E_{ij} be the $n \times n$ matrix whose (i, j) entry is a 1 and all of whose other entries are 0. Then $\{E_{ij}\}_{1 \leq i,j \leq n}$ is a basis of $\mathrm{gl}(n, \mathbb{F})$. The E_{ij} satisfy the multiplication rules

$$E_{ij}E_{kl} = \delta_{jk}E_{il}, \tag{4.3}$$

and so

$$[E_{ij}, E_{kl}] = E_{ij}E_{kl} - E_{kl}E_{ij} = \delta_{jk}E_{il} - \delta_{li}E_{kj} \tag{4.4}$$

Now

$$T_n(\mathbb{F}) = \bigoplus_{i \le j} (\mathbb{F} E_{ij}).$$

For each integer $r \ge 0$, let $\mathfrak{g}_r$ denote the subspace of $T_n(\mathbb{F})$ consisting of those matrices whose entries below the diagonal r steps above the main diagonal are 0. Then

$$\mathfrak{g}_r = \bigoplus_{k \le l-r} (\mathbb{F} E_{kl})$$

Note that $T_n(\mathbb{F}) = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_n \supset \mathfrak{g}_{n+1} = \{0\}$. We claim that $\mathfrak{g}_r$ is an ideal of $T_n(\mathbb{F})$ and that $\mathfrak{g}_r/\mathfrak{g}_{r+1}$ is abelian.

To prove that $\mathfrak{g}_r$ is an ideal of $T_n(\mathbb{F})$, we just need to prove that $[E_{ij}, E_{kl}] \in \mathfrak{g}_r$ whenever $E_{ij} \in T_n(\mathbb{F})$ and $E_{kl} \in \mathfrak{g}_r$. For this, we apply the commutation rule (4.4). The right hand side of (4.4) is nonzero only if $j = k$ or $l = i$. If $j = k$, then $i \le j = k \le l - r$, so $E_{il} \in \mathfrak{g}_r$. If $l = i$, then $k \le l - r = i - r \le j - r$, so $E_{kj} \in \mathfrak{g}_r$. Either way, we get $[E_{ij}, E_{kl}] \in \mathfrak{g}_r$.

The condition that $\mathfrak{g}_r/\mathfrak{g}_{r+1}$ is abelian is the same as the condition that $[\mathfrak{g}_r, \mathfrak{g}_r] \subset \mathfrak{g}_{r+1}$. For $r = 0$, the proof is the same as the argument leading up to equation (4.2).

For $r \ge 1$, we will show that if E_{ij} and E_{kl} are in $\mathfrak{g}_r$, then $[E_{ij}, E_{kl}] \in \mathfrak{g}_{r+1}$. For this, it suffices, in turn, to show that the matrix product $E_{ij}E_{kl}$ lies in $\mathfrak{g}_{r+1}$. (The argument that $E_{kl}E_{ij} \in \mathfrak{g}_{r+1}$ is the same.)

Now, by (4.3), $E_{ij}E_{kl}$ is nonzero if and only if $j = k$, in which case the product is E_{il}. But this means that $i \le j - r = k - r \le l - 2r \le l - (r+1)$, since $r \ge 1$. Thus $E_{il} \in \mathfrak{g}_{r+1}$.

We have thus shown that for all $r \ge 0$, $[\mathfrak{g}_r, \mathfrak{g}_r] \subset \mathfrak{g}_{r+1}$, and hence $\mathfrak{g} = T_n(\mathbb{F})$ is solvable. $\square$

We now make the the following observations about solvable Lie algebras. First, if $\mathfrak{g}$ is solvable, then so is any subalgebra $\mathfrak{h}$ of $\mathfrak{g}$. This is because if $\mathfrak{h}^{(i)}$ is the ith term in the derived series for $\mathfrak{h}$, then a simple induction argument shows that $\mathfrak{h}^{(i)} \subset \mathfrak{g}^{(i)}$ for all i. The second observation is that if $\mathfrak{g}$ is solvable, then so is any homomorphic image of $\mathfrak{g}$. In fact, suppose that $\varphi : \mathfrak{g} \to \mathfrak{m}$ is a Lie algebra homomorphism. Then the image $\mathfrak{q} = \varphi(\mathfrak{g})$ is a subalgebra of $\mathfrak{m}$, and it is easy to see, using another simple induction argument, that $\varphi(\mathfrak{g}^{(i)}) = \mathfrak{q}^{(i)}$ for all i.

Proposition 4.1.10. *If $\mathfrak{g}$ is a Lie algebra and $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then $\mathfrak{g}$ is solvable $\iff$ both $\mathfrak{a}$ and $\mathfrak{g}/\mathfrak{a}$ are solvable.*

Proof. If $\mathfrak{g}$ is solvable, then so is $\mathfrak{a}$, since the ideal $\mathfrak{a}$ is also a subalgebra of $\mathfrak{g}$. The quotient algebra $\mathfrak{g}/\mathfrak{a}$ is the homomorphic image of $\mathfrak{g}$ under the projection $\pi : \mathfrak{g} \to \mathfrak{g}/\mathfrak{a}$, so it must also be solvable.

Conversely, suppose that both $\mathfrak{a}$ and $\mathfrak{g}/\mathfrak{a}$ are solvable. Since $\mathfrak{g}/\mathfrak{a}$ is solvable, we must have $(\mathfrak{g}/\mathfrak{a})^{(k)} = \{0\}$, for some k. (The "0" here refers to the zero vector in $\mathfrak{g}/\mathfrak{a}$.) But $(\mathfrak{g}/\mathfrak{a})^{(k)} = \pi(\mathfrak{g})^{(k)} = \pi((\mathfrak{g})^{(k)})$. It follows that $\mathfrak{g}^{(k)} \subset \mathfrak{a}$, and from this, it follows that $\mathfrak{g}^{(k+r)} \subset \mathfrak{a}^{(r)}$, for all r. But then $\mathfrak{a}$ is solvable, so $\mathfrak{a}^{(m)} = \{0\}$ for some m, whence $\mathfrak{g}^{(k+m)} = \{0\}$. Therefore, $\mathfrak{g}$ is solvable. □

Corollary 4.1.11. *Suppose that* $\mathfrak{a}$ *and* $\mathfrak{b}$ *are solvable ideals of any Lie algebra* $\mathfrak{g}$. *Then* $\mathfrak{a}+\mathfrak{b}$ *is a solvable ideal of* $\mathfrak{g}$.

Proof. It is easy to see that $\mathfrak{a}+\mathfrak{b}$ is an ideal of $\mathfrak{g}$. Now $\mathfrak{b}$ is also an ideal of $\mathfrak{a}+\mathfrak{b}$, and by the Isomorphism Theorem (Theorem 3.2.7), $(\mathfrak{a}+\mathfrak{b})/\mathfrak{b} \cong \mathfrak{a}/(\mathfrak{a}\cap\mathfrak{b})$. But the quotient algebra $\mathfrak{a}/(\mathfrak{a}\cap\mathfrak{b})$ is solvable by the preceding proposition. Hence both $(\mathfrak{a}+\mathfrak{b})/\mathfrak{b}$ and $\mathfrak{b}$ are solvable, so again by the preceding proposition, we see that $\mathfrak{a}+\mathfrak{b}$ is solvable. □

Theorem 4.1.12. *Any (finite-dimensional) Lie algebra* $\mathfrak{g}$ *has a maximal solvable ideal* $\mathfrak{R}_s$, *which contains every solvable ideal of* $\mathfrak{g}$.

Proof. Since $\{0\}$ is a solvable ideal of $\mathfrak{g}$, the collection of all solvable ideals of $\mathfrak{g}$ is nonempty. In this collection, let $\mathfrak{R}_s$ be a solvable ideal of *maximal* dimension. If $\mathfrak{a}$ is any solvable ideal, then by Corollary 4.1.11, $\mathfrak{R}_s+\mathfrak{a}$ is a solvable ideal of $\mathfrak{g}$, whence by the maximality of $\mathfrak{R}_s$, we conclude that $\mathfrak{R}_s+\mathfrak{a}=\mathfrak{R}_s$, and so $\mathfrak{a}\subset\mathfrak{R}_s$. □

Definition 4.1.13. $\mathfrak{R}_s$ is called the *solvable radical* of $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ is said to be *semisimple* if $\mathfrak{R}_s = \{0\}$.

Corollary 4.1.14. *If* $\mathfrak{g}$ *is simple, then* $\mathfrak{g}$ *is semisimple.*

Proof. We had previously observed that since $\mathfrak{g}$ is simple, the derived series for $\mathfrak{g}$ is constant: $\mathfrak{g}^{(i)} = \mathfrak{g}$ for all i. Thus, $\mathfrak{g} \neq \mathfrak{R}_s$, but $\mathfrak{R}_s$ is an ideal of $\mathfrak{g}$, so this forces $\mathfrak{R}_s = \{0\}$. □

Are there semisimple Lie algebras which are not simple? Sure! For an example, we first introduce the notion of an external direct sum of Lie algebras.

Let V and W be vector spaces over $\mathbb{F}$. The Cartesian product $V\times W$ has the structure of a vector space, where addition and scalar multiplication are defined by

$$\begin{aligned}(v_1,w_1)+(v_2,w_2) &= (v_1+v_2, w_1+w_2)\\ \alpha\,(v,w) &= (\alpha v, \alpha w),\end{aligned}$$

for all $v, v_1, v_2 \in V$, all $w, w_1, w_2 \in W$, and all $\alpha \in \mathbb{F}$. Equipped with this vector space structure, we call $V_1 \times V_2$ the *external direct sum* of V_1 and V_2.

The external direct sum $V_1 \times V_2 \times \cdots \times V_k$ of k vector spaces is defined similarly.

Exercise 4.1.15. (Easy) Suppose that U' and U'' are subspaces of a vector space V such that $U' \cap U'' = \{0\}$. Show that the map $(u', u'') \mapsto u' + u''$ is a linear isomorphism from the external direct sum $U' \times U''$ onto the direct sum $U' \oplus U''$.

If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras, then their external direct sum is also a Lie algebra with Lie bracket defined by

$$[(x_1, y_1), (x_2, y_2)] = ([x_1, x_2], [y_1, y_2])$$

(In the above, for the sake of clarity, we're massively abusing notation, since $[x_1, x_2]$ refers to the bracket in $\mathfrak{g}$, $[y_1, y_2]$ refers to the bracket in $\mathfrak{h}$, and the bracket on the left hand side above is the one for $\mathfrak{g} \times \mathfrak{h}$. Strictly speaking, we should distinguish between the Lie brackets for $\mathfrak{g}$ and $\mathfrak{h}$ by denoting them by $[\ ,\ ]_{\mathfrak{g}}$ and $[\ ,\ ]_{\mathfrak{h}}$, respectively, but you can see how cumbersome this quickly gets.)

Exercise 4.1.16. (Straightforward) Show that the Lie bracket above makes $\mathfrak{g} \times \mathfrak{h}$ a Lie algebra.

Exercise 4.1.17. Let $\mathfrak{g}$ be the external direct sum $\mathrm{sl}(2, \mathbb{C}) \times \mathrm{sl}(2, \mathbb{C})$. Prove that the Lie algebra $\mathfrak{g}$ is semisimple but that $\mathfrak{g}$ is not simple. (*Hint:* The projections π_1 and π_2 of $\mathrm{sl}(2, \mathbb{C}) \times \mathrm{sl}(2, \mathbb{C})$ onto its first and second factors, respectively, are surjective homomorphisms.)

Corollary 4.1.18. *If $\mathfrak{g}$ is any Lie algebra and $\mathfrak{R}_s$ is its solvable radical, then the quotient algebra $\mathfrak{g}/\mathfrak{R}_s$ is semisimple.*

Proof. Let $\mathfrak{I}$ denote the solvable radical of $\mathfrak{g}/\mathfrak{R}_s$. Then, by the Correspondence Theorem (Theorem 3.2.6), we have $\mathfrak{I} = \overline{\mathfrak{R}}/\mathfrak{R}_s$, where $\overline{\mathfrak{R}}$ is an ideal of $\mathfrak{g}$ containing $\mathfrak{R}_s$. But since both $\overline{\mathfrak{R}}/\mathfrak{R}_s$ and $\mathfrak{R}_s$ are solvable, it follows from Corollary 4.1.11 that $\overline{\mathfrak{R}}$ is solvable. Since $\mathfrak{R}_s$ is maximal solvable, we conclude that $\overline{\mathfrak{R}} = \mathfrak{R}_s$, so $\mathfrak{I} = \{0\}$. This shows that $\mathfrak{g}/\mathfrak{R}_s$ is semisimple. $\square$

Exercise 4.1.19. Suppose that $\mathfrak{g}$ is solvable. Show that $\mathfrak{g}$ has no semisimple subalgebra $\neq \{0\}$.

4.2 Lie's Theorem

Let V be a nonzero vector space over $\mathbb{F}$. Let us recall that $\mathrm{gl}(V)$ is the Lie algebra of all linear operators on V (same as $\mathcal{L}(V)$), in which the Lie bracket is the commutator $[A,B]=AB-BA$. If we fix a basis B of V, then the map which takes any $T\in \mathrm{gl}(V)$ into its matrix $M(T)$ with respect to B is a Lie algebra isomorphism from $\mathrm{gl}(V)$ onto $\mathrm{gl}(n,\mathbb{F})$.

Our objective now is to prove Lie's Theorem, which says that, when V is a *complex* vector space, then any solvable subalgebra of $\mathrm{gl}(V)$ is essentially an algebra of upper triangular matrices; i.e., a subalgebra of $T_n(\mathbb{C})$ (wherein we identify an operator T with its matrix $M(T)$ under the isomorphism given above).

Let $\mathfrak{g}$ be a Lie subalgebra of $\mathrm{gl}(V)$, and suppose that f is a linear functional on $\mathfrak{g}$. The *joint eigenspace of* $\mathfrak{g}$ *corresponding to* f is the subset of V given by

$$V_f=\{v\in V\,|\,T(v)=f(T)\,v \text{ for all } T\in\mathfrak{g}\} \tag{4.5}$$

The joint eigenspace V_f is easily shown to be a subspace of V: supposing that $v_1,\,v_2\in V_f$ and $\alpha\in\mathbb{C}$, then $T(v_1+v_2)=T(v_1)+T(v_2)=f(T)\,v_1+f(T)\,v_2=f(T)\,(v_1+v_2)$; and similarly, $T(\alpha v_1)=\alpha T(v_1)=\alpha\,f(T)\,v_1=f(T)\,(\alpha v_1)$, for all $T\in\mathfrak{g}$.

Of course, for a given f, V_f could very well be the trivial subspace $\{0\}$ of V. Any nonzero element of a joint eigenspace of $\mathfrak{g}$ is called a *joint eigenvector* of $\mathfrak{g}$.

Any nonzero vector $v\in V$ which is an eigenvector of each $T\in\mathfrak{g}$ is necessarily a joint eigenvector of $\mathfrak{g}$. For this, we simply define the function $f:\mathfrak{g}\to\mathbb{F}$ by the requirement that

$$T\,v=f(T)\,v,$$

for all $T\in\mathfrak{g}$. It is easy to show that f is a linear functional on $\mathfrak{g}$, and that therefore v is a nonzero element of V_f.

The following important lemma is the key to Lie's Theorem.

Lemma 4.2.1. *(E.B. Dynkin) Let V be a nonzero vector space over $\mathbb{F}$, and let $\mathfrak{g}$ be a Lie subalgebra of gl(V). Suppose that $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, and that f is a linear functional on $\mathfrak{a}$. If V_f is the joint eigenspace of $\mathfrak{a}$ corresponding to f, then V_f is invariant under $\mathfrak{g}$. That is, $X(V_f)\subset V_f$ whenever $X\in\mathfrak{g}$.*

Proof. Let $X\in\mathfrak{g}$ and $v\in V_f$. We want to prove that $X(v)\in V_f$. That is, we want to prove that $T(X(v))=f(T)\,X(v)$ for any $T\in\mathfrak{a}$. For $v=0$, this result trivially holds, so we may assume that $v\neq 0$.

Note that for any $T \in \mathfrak{a}$,

$$\begin{aligned} T(X(v)) &= X(T(v)) + (TX - XT)(v) \\ &= X(T(v)) + [T, X](v) \\ &= X(f(T)\, v) + f([T, X])\, v \qquad\qquad \text{(since } [T, X] \in \mathfrak{a}) \\ &= f(T)\, X(v) + f([T, X])\, v \end{aligned} \tag{4.6}$$

The proof will be complete once we prove that $f([T, X]) = 0$.

Let $v_0 = v$, $v_1 = X(v)$, $v_2 = X^2(v), \ldots, v_j = X^j(v), \ldots$. Next, for each $j \geq 0$, let V_j be the subspace of V spanned by $(v_0, \ldots, v_j)$. Since V is finite-dimensional, there is an integer $k \geq 0$ such that $(v_0, \ldots, v_k)$ is linearly independent but $(v_0, \ldots, v_k, v_{k+1})$ is not. Let k be the smallest such integer.

We claim that for each j such that $0 \leq j \leq k$, the subspace V_j is invariant under *any* $T \in \mathfrak{a}$ and that the matrix of $T|_{V_j}$ with respect to the basis $(v_0, \ldots, v_j)$ of V_j is upper triangular of the form

$$\begin{pmatrix} f(T) & & * \\ & \ddots & \\ 0 & & f(T) \end{pmatrix} \tag{4.7}$$

If $k = 0$, then this is obvious, since $V_j = V_k = V_0 = \mathbb{F}v_0$, and $T(v_0) = f(T)\, v_0$, because $v_0 \in V_f$.

So assume that $k \geq 1$. Equation (4.6) says that for any $T \in \mathfrak{a}$,

$$T(v_1) = f(T)\, v_1 + f([T, X])\, v_0,$$

which shows that the subspace $V_1 = \mathbb{F}v_0 + \mathbb{F}v_1$ is invariant under T. Moreover, relative to the basis $(v_0,\, v_1)$ of V_1, the matrix of the restriction $T|_{V_1}$ is

$$\begin{pmatrix} f(T) & f([T, X]) \\ 0 & f(T) \end{pmatrix}$$

We will now use induction on j to prove the same thing for V_j, for any $j \leq k$. So assume that V_{j-1} is T-invariant, and that, for any $T \in \mathfrak{a}$, the matrix of the restriction $T|_{V_{j-1}}$ with respect to the basis $(v_0, \ldots, v_{j-1})$ of V_{j-1} is of the form 4.7. Now for any $T \in \mathfrak{a}$, we have,

$$\begin{aligned} T(v_j) &= T(X^j(v)) \\ &= T\, X\, (X^{j-1}(v)) \\ &= X\, T\, (X^{j-1}(v)) + [T, X]\, X^{j-1}(v) \\ &= X\, T(v_{j-1}) + [T, X](v_{j-1}) \\ &= X\big(f(T) v_{j-1} + \sum_{i<j-1} c_i v_i\big) \;+\; \big(f([T, X])\, v_{j-1} + \sum_{i<j-1} d_i v_i\big), \end{aligned}$$

by the induction hypothesis, where the c_i and the d_i are constants. The last expression above then equals

$$\begin{aligned} f(T)\,X(v_{j-1}) + \sum_{i<j-1} c_i X(v_i) + f([T,X])\,v_{j-1} + \sum_{i<j-1} d_i v_i \\ = f(T)\,v_j + \sum_{i<j-1} c_i v_{i+1} + f([T,X])\,v_{j-1} + \sum_{i<j-1} d_i v_i \\ = f(T)\,v_j + \big(\text{a linear combination of } (v_0,\dots,v_{j-1})\big) \end{aligned}$$

This proves our claim. In particular, V_k is invariant under any $T \in \mathfrak{a}$, and the matrix of $T|_{V_k}$ is of the form 4.7.

This means that for any $T \in \mathfrak{a}$, the trace of $T|_{V_k}$ is $(k+1)f(T)$. Hence, the trace of the restriction $[T,X]|_{V_k}$ is $(k+1)\,f([T,X])$. But then, this trace also equals

$$\operatorname{tr}(TX - XT)|_{V_k} = \operatorname{tr}(T|_{V_k}\,X|_{V_k}) - \operatorname{tr}(X|_{V_k}\,T|_{V_k}) = 0.$$

Thus $(k+1)\,f([T,X]) = 0$, whence $f([T,X]) = 0$, proving the lemma. $\square$

The following theorem can be construed as a generalization of Theorem 1.5.2, which states that any linear operator on a complex vector space has an eigenvector.

Theorem 4.2.2. *Let V be a nonzero vector space over $\mathbb{C}$, and let $\mathfrak{g}$ be a solvable Lie subalgebra of $gl(V)$. Then $\mathfrak{g}$ has a joint eigenvector.*

This theorem asserts that there exists a nonzero vector $v \in V$ and a linear functional f on $\mathfrak{g}$ such that $T(v) = f(T)\,v$, for all $T \in \mathfrak{g}$.

Proof. We prove the theorem by induction on $\dim \mathfrak{g}$. If $\dim \mathfrak{g} = 1$, then $\mathfrak{g} = \mathbb{C}T$, where T is a linear operator on V. By Theorem 1.5.2, T has an eigenvalue λ. Let v be an eigenvector corresponding to λ. For any $S \in \mathfrak{g}$, we have $S = cT$, so $S(v) = cT(v) = c\lambda\, v$, so we can put $f(cT) = c\lambda$. Clearly, $f \in \mathfrak{g}^*$.

Now assume that $\dim \mathfrak{g} = m$, and that any solvable Lie subalgebra of $\mathrm{gl}(V)$ of dimension $< m$ has a joint eigenvector. Consider the derived algebra $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$. Since $\mathfrak{g}$ is solvable, $\mathfrak{g}'$ is a proper ideal of $\mathfrak{g}$, so it is a subalgebra of $\mathrm{gl}(V)$ of dimension $< m$.

Next let $\mathfrak{h}$ be *any* vector subspace of $\mathfrak{g}$, of dimension $m-1$, such that $\mathfrak{g}' \subset \mathfrak{h}$. Such an $\mathfrak{h}$ of course exists. By Proposition 4.1.4, $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. Moreover, since $\mathfrak{g}$ is solvable, so is $\mathfrak{h}$. (See the observations made after Example 4.1.9.)

Thus, by the induction hypothesis, $\mathfrak{h}$ has a joint eigenvector. In other words, $\mathfrak{h}$ has a nonzero joint eigenspace V_μ, where μ is a linear functional on $\mathfrak{h}$.

Since $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, we conclude, using Lemma 4.2.1, that V_μ must be $\mathfrak{g}$-invariant. Let S be a nonzero element of $\mathfrak{g}$ not in $\mathfrak{h}$. Then, since $\dim \mathfrak{h} = m-1$, we have $\mathfrak{g} = \mathfrak{h} \oplus \mathbb{C}S$. The subspace V_μ is S-invariant, so the restriction $S|_{V_\mu}$ must have an eigenvalue λ. Let $v \in V_\mu$ be an eigenvector of $S|_{V_\mu}$ corresponding to λ.

For any $T \in \mathfrak{g}$, we have $T = cS + Y$, for unique $Y \in \mathfrak{h}$ and $c \in \mathbb{C}$. Define the map $f : \mathfrak{g} \to \mathbb{C}$ by $f(cS+Y) = c\lambda + \mu(Y)$. It is easy to prove that f is a linear functional on $\mathfrak{g}$. Moreover, if $T = cS + Y \in \mathfrak{g}$,

$$\begin{aligned} T(v) &= (cS+Y)(v) \\ &= c\,S(v) + Y(v) \\ &= c\lambda\, v + \mu(Y)\, v \\ &= (c\lambda + \mu(Y))\, v \\ &= f(T)\, v. \end{aligned}$$

This shows that v is a joint eigenvector of $\mathfrak{g}$, completing the induction step and proving the theorem. □

Theorem 4.2.3. *(Lie's Theorem) Let V be a nonzero complex vector space, and let $\mathfrak{g}$ be a solvable Lie subalgebra of $gl(V)$. Then V has a basis $(v_1, \dots, v_n)$ with respect to which every element of $\mathfrak{g}$ has an upper triangular matrix.*

Proof. The proof is by induction on $\dim V$. If $\dim V = 1$, there is nothing to prove. So assume that $\dim V = n > 1$, and that Lie's theorem holds for all complex vector spaces of dimension $< n$.

Now by Theorem 4.2.2, $\mathfrak{g}$ has a joint eigenvector v_1. Let $V_1 = \mathbb{C}\, v_1$. Then, for every $T \in \mathfrak{g}$, the subspace V_1 is T-invariant; let $\widetilde{T} : V/V_1 \to V/V_1$ be the induced linear map.

The map $T \mapsto \widetilde{T}$ is a Lie algebra homomorphism of $\mathfrak{g}$ into $\mathrm{gl}(V/V_1)$. It's clearly linear, and the relation $\widetilde{[S,T]} = [\widetilde{S}, \widetilde{T}]$ is easily verified by a simple computation. Since homomorphic images of solvable Lie algebras are solvable, the image $\widetilde{\mathfrak{g}}$ of this homomorphism is a solvable Lie subalgebra of $\mathrm{gl}(V/V_1)$.

Since $\dim(V/V_1) = n-1$, we can now apply the induction hypothesis to obtain a basis $(v_2 + V_1, \dots, v_n + V_1)$ of V/V_1 for which the elements of $\widetilde{\mathfrak{g}}$ are upper triangular.

The list $(v_1, v_2, \dots, v_n)$ is then a basis of V. For each $T \in \mathfrak{g}$, the matrix of $\widetilde{T} : V/V_1 \to V/V_1$ with respect to $(v_2+V_1, \dots, v_n+V_1)$ is upper triangular. Hence the matrix of T with respect to $(v_1, v_2, \dots, v_n)$ is upper triangular, proving the theorem. □

A *flag* in a vector space V is a sequence $(V_1, \ldots, V_k)$ of subspaces of V such that $V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_k$. We say that a linear operator $T \in \mathcal{L}(V)$ *stabilizes* the flag $(V_1, \ldots, V_k)$ if each V_i is T-invariant. Finally, a Lie subalgebra $\mathfrak{g}$ of $\mathrm{gl}(V)$ *stabilizes* the flag $(V_1, \ldots, V_k)$ if each $T \in \mathfrak{g}$ stabilizes the flag.

Corollary 4.2.4. *If $\mathfrak{g}$ is a solvable Lie subalgebra of $gl(V)$, then $\mathfrak{g}$ stabilizes some flag $(\{0\} = V_0, V_1, V_2, \ldots, V_n = V)$.*

Proof. Let $(v_1, \ldots, v_n)$ be a basis of V with respect to which the matrix of every element of $\mathfrak{g}$ is upper triangular. Then, for each i, let $V_i = \mathbb{C}v_1 + \cdots + \mathbb{C}v_i$. $\square$

Corollary 4.2.5. *(Lie's Abstract Theorem) Let $\mathfrak{g}$ be a solvable Lie algebra over $\mathbb{C}$, of dimension N. Then $\mathfrak{g}$ has a chain of ideals $\{0\} = \mathfrak{g}_0 \subsetneq \mathfrak{g}_1 \subsetneq \cdots \subsetneq \mathfrak{g}_N = \mathfrak{g}$.*

Proof. The adjoint representation $x \mapsto \operatorname{ad} x$ maps $\mathfrak{g}$ onto the solvable Lie subalgebra $\operatorname{ad}\mathfrak{g}$ of $\operatorname{Der}\mathfrak{g} \subset \mathrm{gl}(\mathfrak{g})$. Thus $\operatorname{ad}\mathfrak{g}$ stabilizes a flag $\{0\} = \mathfrak{g}_0 \subsetneq \mathfrak{g}_1 \subsetneq \cdots \subsetneq \mathfrak{g}_N = \mathfrak{g}$ in $\mathfrak{g}$. Each subspace $\mathfrak{g}_i$ therefore satisfies $\operatorname{ad} x\,(\mathfrak{g}_i) \subset \mathfrak{g}_i$, for all $x \in \mathfrak{g}$. This means that $\mathfrak{g}_i$ is an ideal of $\mathfrak{g}$. $\square$

In particular, Corollary 4.2.5 shows that if $\mathfrak{g}$ is a complex solvable Lie algebra and if $0 \le i \le \dim\mathfrak{g}$, then $\mathfrak{g}$ has an ideal of dimension i.

In Example 4.1.9, we saw that the Lie algebra $T_n(\mathbb{F})$ of all upper triangular $n \times n$ matrices over $\mathbb{F}$ is solvable. If a Lie algebra $\mathfrak{g}$ is solvable and complex, then the following shows that $\mathfrak{g}$ is in some sense just a subalgebra of $T_n(\mathbb{C})$. Thus $T_n(\mathbb{C})$ is *the* "prototypical" solvable complex Lie algebra. For this, we will need the following important theorem.

Theorem 4.2.6. *(Ado's Theorem) Let $\mathfrak{g}$ be any nonzero Lie algebra over $\mathbb{F}$. Then there exists a vector space V over $\mathbb{F}$ and an injective Lie algebra homomorhism φ of $\mathfrak{g}$ into $gl(V)$.*

We won't be needing Ado's Theorem in the sequel, so we omit its proof.

Now suppose that $\mathfrak{g}$ is a solvable complex Lie algebra. Using Ado's Theorem, we may therefore identify $\mathfrak{g}$ with a (solvable) Lie subalgebra of $\mathrm{gl}(V)$. Then, from Lie's theorem, there is a basis B of V with respect to which the matrix of every element of $\mathfrak{g}$ is upper triangular. Now, for every linear operator T on V, let $M(T)$ be its matrix with respect to B. Then the map $T \mapsto M(T)$ is a Lie algebra isomorphism of $\mathrm{gl}(V)$ onto $\mathrm{gl}(n, \mathbb{C})$. The image of $\mathfrak{g}$ under this isomorphism is a Lie subalgebra of $T_n(\mathbb{C})$. Thus $\mathfrak{g}$ may be identified with this Lie subalgebra of $T_n(\mathbb{C})$.

Chapter 5

Nilpotent Lie Algebras and Engel's Theorem

5.1 Nilpotent Lie Algebras

For any Lie algebra $\mathfrak{g}$ over $\mathbb{F}$, we define a sequence of subspaces of $\mathfrak{g}$ as follows. Let $\mathcal{C}^1(\mathfrak{g}) = \mathfrak{g}$, $\mathcal{C}^2(\mathfrak{g}) = \mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$, and inductively, let $\mathcal{C}^{i+1}(\mathfrak{g}) = [\mathcal{C}^i(\mathfrak{g}), \mathfrak{g}]$ for all i.

Proposition 5.1.1. *The subspaces $\mathcal{C}^i(\mathfrak{g})$ satisfy the following properties:*

1. *Each $\mathcal{C}^i(\mathfrak{g})$ is a characteristic ideal of $\mathfrak{g}$.*
2. $\mathcal{C}^1(\mathfrak{g}) \supset \mathcal{C}^2(\mathfrak{g}) \supset \cdots \supset \mathcal{C}^i(\mathfrak{g}) \supset \cdots$
3. $\mathcal{C}^i(\mathfrak{g})/\mathcal{C}^{i+1}(\mathfrak{g})$ *lies inside the center of* $\mathfrak{g}/\mathcal{C}^{i+1}(\mathfrak{g})$

Proof. We prove (1) by induction on i, the case $i = 1$ being trivial. Suppose that $\mathcal{C}^i(\mathfrak{g})$ is a characteristic ideal of $\mathfrak{g}$. Then $[\mathfrak{g}, \mathcal{C}^{i+1}(\mathfrak{g})] = [\mathfrak{g}, [\mathcal{C}^i(\mathfrak{g}), \mathfrak{g}]] \subset [\mathfrak{g}, \mathcal{C}^i(\mathfrak{g})] = \mathcal{C}^{i+1}(\mathfrak{g})$, proving that $\mathcal{C}^{i+1}(\mathfrak{g})$ is an ideal of $\mathfrak{g}$. Moreover, for any derivation D of $\mathfrak{g}$, we have

$$\begin{aligned} D(\mathcal{C}^{i+1}(\mathfrak{g})) &= D([\mathcal{C}^i(\mathfrak{g}), \mathfrak{g}]) \\ &\subset [D(\mathcal{C}^i(\mathfrak{g})), \mathfrak{g}] + [\mathcal{C}^i(\mathfrak{g}), D\mathfrak{g}] \\ &\subset [\mathcal{C}^i(\mathfrak{g}), \mathfrak{g}] + [C^i(\mathfrak{g}), \mathfrak{g}] \\ &= \mathcal{C}^{i+1}(\mathfrak{g}), \end{aligned}$$

proving that $\mathcal{C}^{i+1}(\mathfrak{g})$ is characteristic.

Note that since $\mathcal{C}^{i+1}(\mathfrak{g})$ is an ideal of $\mathfrak{g}$, it is also an ideal of $\mathcal{C}^i(\mathfrak{g})$.

Likewise, for (2), we prove the inclusion $\mathcal{C}^{i+1}(\mathfrak{g}) \subset C^i(\mathfrak{g})$ by induction on i, with the case $i = 1$ corresponding to the trivial inclusion $[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$. Assume, then, that $\mathcal{C}^{i+1}(\mathfrak{g}) \subset C^i(\mathfrak{g})$. Then $\mathcal{C}^{i+2}(\mathfrak{g}) = [\mathcal{C}^{i+1}(\mathfrak{g}), \mathfrak{g}] \subset [\mathcal{C}^i(\mathfrak{g}), \mathfrak{g}] = \mathcal{C}^{i+1}(\mathfrak{g})$.

Finally, for (3), let $x \in \mathcal{C}^i(\mathfrak{g})$. Then for any $y \in \mathfrak{g}$, we have $[x, y] \in \mathcal{C}^{i+1}(\mathfrak{g})$. Hence in the quotient algebra $\mathfrak{g}/\mathcal{C}^{i+1}(\mathfrak{g})$, we have

$$[x + \mathcal{C}^{i+1}(\mathfrak{g}), y + \mathcal{C}^{i+1}(\mathfrak{g})] = [x, y] + \mathcal{C}^{i+1}(\mathfrak{g}) = \mathcal{C}^{i+1}(\mathfrak{g})$$

It follows that every element of $\mathcal{C}^i(\mathfrak{g})/\mathcal{C}^{i+1}(\mathfrak{g})$ is an element of the center of $\mathfrak{g}/\mathcal{C}^{i+1}(\mathfrak{g})$. □

Definition 5.1.2. The *descending central series* for $\mathfrak{g}$ is the sequence of ideals $\mathfrak{g} = \mathcal{C}^1(\mathfrak{g}) \supset \mathcal{C}^2(\mathfrak{g}) \supset \cdots \supset \mathcal{C}^i(\mathfrak{g}) \supset \cdots$. (Since $\dim \mathfrak{g}$ is finite, it is clear that this series stabilizes after some point.) The Lie algebra $\mathfrak{g}$ is said to be *nilpotent* if $\mathcal{C}^k(\mathfrak{g}) = \{0\}$ for some k.

Note that the term "central" is appropriate since $\mathcal{C}^i(\mathfrak{g})/\mathcal{C}^{i+1}(\mathfrak{g}) \subset \mathfrak{c}(\mathfrak{g}/\mathcal{C}^{i+1}(\mathfrak{g}))$.

Definition 5.1.3. Let $\mathcal{C}_0 = \{0\}$, $\mathcal{C}_1 = \mathfrak{c}(\mathfrak{g})$ and, recursively, let $\mathcal{C}_i$ be the ideal in $\mathfrak{g}$ such that $\mathcal{C}_i/\mathcal{C}_{i-1} = \mathfrak{c}(\mathfrak{g}/\mathcal{C}_{i-1})$. (This ideal exists and is unique because of the Correspondence Theorem (Theorem 3.2.6).) The *ascending central series* is the sequence of ideals $\{0\} = \mathcal{C}_0 \subset \mathcal{C}_1 \subset \cdots \subset \mathcal{C}_i \subset \cdots$. (Since $\dim \mathfrak{g}$ is finite, this series stabilizes after some point.)

Proposition 5.1.4. *The Lie algebra $\mathfrak{g}$ is nilpotent if and only if $\mathcal{C}_s = \mathfrak{g}$ for some positive integer s.*

Proof. We may assume that $\mathfrak{g} \neq \{0\}$; otherwise, there is nothing to prove.

Suppose first that $\mathfrak{g}$ is nilpotent. Let k be the smallest (necessarily positive) integer such that $\mathcal{C}^k(\mathfrak{g}) = \{0\}$. For any integer i, with $0 \leq i \leq k$, we claim that $\mathcal{C}_i \supset C^{k-i}(\mathfrak{g})$. From this, it will follow that $\mathcal{C}_k \supset \mathcal{C}^0(\mathfrak{g}) = \mathfrak{g}$.

To prove the claim, we first note that $\mathcal{C}_0 = \{0\} = \mathcal{C}^k(\mathfrak{g})$, so the claim is certainly true for $k = 0$. Assume, inductively, that $\mathcal{C}_i \supset \mathcal{C}^{k-i}(\mathfrak{g})$. By statement (3) of the Isomorphism Theorem (Theorem 3.2.7) and its proof, the map $\varphi : x + \mathcal{C}^{k-i}(\mathfrak{g}) \mapsto x + \mathcal{C}_i$ is a surjective Lie algebra homomorphism of $\mathfrak{g}/\mathcal{C}^{k-i}(\mathfrak{g})$ onto $\mathfrak{g}/\mathcal{C}_i$, with kernel $C_i/\mathcal{C}^{k-i}(\mathfrak{g})$. By Proposition 3.3.3, φ maps the center of $\mathfrak{g}/\mathcal{C}^{k-i}(\mathfrak{g})$ into the center $\mathcal{C}_{i+1}/\mathcal{C}_i$ of $\mathfrak{g}/\mathcal{C}_i$. But by part (3) of Proposition 5.1.1, $\mathcal{C}^{k-i-1}(\mathfrak{g})/\mathcal{C}^{k-i}(\mathfrak{g})$ lies in the center of $\mathfrak{g}/\mathcal{C}^{k-i}(\mathfrak{g})$. Hence φ maps $\mathcal{C}^{k-i-1}(\mathfrak{g})/\mathcal{C}^{k-i}(\mathfrak{g})$ into $\mathcal{C}_{i+1}/\mathcal{C}_i$. Thus, if $x \in \mathcal{C}^{k-i-1}(\mathfrak{g})$, then $x + \mathcal{C}_i \in \mathcal{C}_{i+1}/\mathcal{C}_i$, and hence $x \in \mathcal{C}_{i+1}$. This shows that $\mathcal{C}^{k-i-1}(\mathfrak{g}) \subset \mathcal{C}_{i+1}$, completing the induction and proving the claim.

Next we assume that $\mathcal{C}_s = \mathfrak{g}$, for some s. Let k be the smallest integer such that $\mathcal{C}_k = \mathfrak{g}$. Since $\mathfrak{g} \neq \{0\}$, this k is necessarily positive. We now prove, by induction on i, that $\mathcal{C}^i(\mathfrak{g}) \subset \mathcal{C}_{k-i}$. When $i = 0$, this inclusion is just $\mathcal{C}^0(\mathfrak{g}) = \mathfrak{g} = \mathcal{C}_k$, which is already true. Now assume that for $i \geq 0$, $\mathcal{C}^i(\mathfrak{g}) \subset \mathcal{C}_{k-i}$. Then $\mathcal{C}^{i+1}(\mathfrak{g}) = [\mathfrak{g}, \mathcal{C}^i(\mathfrak{g})] \subset [\mathfrak{g}, \mathcal{C}_{k-i}] \subset \mathcal{C}_{k-i-1}$, the last inclusion arising from the condition that $\mathcal{C}_{k-i} \subset \mathfrak{c}(\mathfrak{g}/\mathcal{C}_{k-i-1})$. This completes the induction.

When $i = k$, we therefore obtain $\mathcal{C}^k(\mathfrak{g}) \subset \mathcal{C}_0 = \{0\}$. Hence $\mathfrak{g}$ is nilpotent. □

Exercise 5.1.5. Show that $\mathfrak{g}$ is nilpotent if and only if it has a central series which reaches $\{0\}$; that is, there is a descending sequence of ideals of $\mathfrak{g}$:

$$\mathfrak{g} \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_m = \{0\}$$

where $\mathfrak{g}_i/\mathfrak{g}_{i+1} \subset \mathfrak{c}(\mathfrak{g}/\mathfrak{g}_{i+1})$.

Exercise 5.1.6. Prove that if $\mathfrak{g}$ is nilpotent, then $\mathfrak{g}$ is solvable.

Example 5.1.7. Recall that $\mathfrak{g} = T_n(\mathbb{F})$ is the solvable Lie algebra of $n \times n$ upper triangular matrices over $\mathbb{F}$. Then $T_n(\mathbb{F}) = D_n(\mathbb{F}) \oplus U_n(\mathbb{F})$, where $D_n(\mathbb{F})$ is the vector space of diagonal $n \times n$ matrices and $U_n(\mathbb{F})$ is the vector space of strictly upper triangular $n \times n$ matrices. We saw in Example 4.1.9 that $\mathfrak{g}' \subset U_n(\mathbb{F})$.

In fact $\mathfrak{g}' = U_n(\mathbb{F})$, since $U_n(\mathbb{F})$ has basis $\{E_{ij}\}_{i<j}$, and

$$E_{ij} = [E_{ii}, E_{ij}].$$

The above equation also shows that $[D_n(\mathbb{F}), U_n(\mathbb{F})] = U_n(\mathbb{F})$. Hence $\mathcal{C}^2(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}'] \supset [D_n(\mathbb{F}), U_n(\mathbb{F})] = U_n(\mathbb{F}) = \mathfrak{g}'$, so $\mathcal{C}^2(\mathfrak{g}) = \mathfrak{g}'$, and it follows that $\mathcal{C}^i(\mathfrak{g}) = \mathfrak{g}'$ for all $i \geq 1$. Thus $T_n(\mathbb{F})$ is not nilpotent.

On the other hand, the Lie algebra $\mathfrak{h} = U_n(\mathbb{F})$ is nilpotent. Using the notation of Example 4.1.9, let $\mathfrak{g}_r$ denote the subspace of $T_n(\mathbb{F})$ consisting of those matrices with 0's below the diagonal r steps above the main diagonal. We claim, using induction on r, that $\mathcal{C}^r(\mathfrak{h}) = \mathfrak{g}_r$ for all $r \geq 1$. For $r = 1$, this is just the equality $\mathfrak{g}_1 = U_n(\mathbb{F}) = \mathfrak{h}$. Assuming that the claim is true for r, the corresponding equality for $r + 1$ will follow if we can show that $[\mathfrak{g}_1, \mathfrak{g}_r] = \mathfrak{g}_{r+1}$. Now $\mathfrak{g}_r$ is spanned by the elementary matrices E_{kl}, where $l \geq k + r$. Suppose that $E_{ij} \in \mathfrak{g}_1$ and $E_{kl} \in \mathfrak{g}_r$. Then, equation (4.4) says that

$$[E_{ij}, E_{kl}] = \delta_{jk} E_{il} - \delta_{li} E_{kj}.$$

If $j = k$, then $l \geq k + r = j + r \geq i + 1 + r$, so $E_{il} \in \mathfrak{g}_{r+1}$. If $l = i$, then $j \geq i + 1 = l + 1 \geq k + r + 1$, so $E_{kj} \in \mathfrak{g}_{r+1}$. Either way, the Lie bracket above belongs to $\mathfrak{g}_{r+1}$, and this shows that $[\mathfrak{g}_1, \mathfrak{g}_r] \subset \mathfrak{g}_{r+1}$.

On the other hand, if $l \geq k + r + 1$, then

$$E_{kl} = E_{k,k+1}\, E_{k+1,l} = [E_{k,k+1}, E_{k+1,l}] \in [\mathfrak{g}_1, \mathfrak{g}_r],$$

so $\mathfrak{g}_{r+1} \subset [\mathfrak{g}_1, \mathfrak{g}_r]$. We have thus proved that $[\mathfrak{g}_1, \mathfrak{g}_r] = \mathfrak{g}_{r+1}$.

Of course, when $r = n$, we get $\mathcal{C}^n(\mathfrak{h}) = \mathfrak{g}_{n+1} = \{0\}$. Thus $\mathfrak{h} = U_n(\mathbb{F})$ is nilpotent.

Exercise 5.1.8. Prove that the Lie algebras $T_n(\mathbb{F})$ and $U_n(\mathbb{F})$ both have centers of dimension 1.

Exercise 5.1.9. (From Wikipedia [12]) The *general Heisenberg Lie algebra* $\mathfrak{h}_n$ is defined as the vector space of $(n+2) \times (n+2)$ matrices over $\mathbb{F}$ which have block form

$$\begin{pmatrix} 0 & {}^t v & c \\ 0 & \mathbf{0}_n & w \\ 0 & 0 & 0 \end{pmatrix}, \tag{5.1}$$

where v, $w \in \mathbb{F}^n$, $c \in \mathbb{F}$, and $\mathbf{0}_n$ is the zero $n \times n$ matrix. Prove that $\mathfrak{h}_n$ is a nilpotent Lie algebra of dimension $2n+1$. What is the minimum k such that $\mathcal{C}^k(\mathfrak{g}) = \{0\}$? ($\mathfrak{h}_n$ is the Lie algebra of the *Heisenberg group*, which in the case $\mathbb{F} = \mathbb{R}$ is used in the description of n-dimension quantum mechanical systems.)

Exercise 5.1.10. Show that any non-abelian two-dimensional Lie algebra over $\mathbb{F}$ is solvable, but not nilpotent. (See Exercise 3.2.2.)

Let $\varphi : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homorphism. Then the image $\varphi(\mathfrak{g})$ is a Lie subalgebra of $\mathfrak{h}$, and it is clear that $\varphi(\mathcal{C}^i(\mathfrak{g})) = \mathcal{C}^i(\varphi(\mathfrak{g}))$, for all i. Thus, if $\mathfrak{g}$ is nilpotent, then so are all homomorphic images of $\mathfrak{g}$.

Exercise 5.1.11. Suppose that $\mathfrak{g}$ is nilpotent. Prove that:

1. All subalgebras of $\mathfrak{g}$ are nilpotent.
2. If $\mathfrak{g} \neq \{0\}$, then its center $\mathfrak{c} \neq \{0\}$.

Proposition 5.1.12. *Let $\mathfrak{g}$ be a Lie algebra, with center $\mathfrak{c}$. Then $\mathfrak{g}$ is nilpotent if and only if $\mathfrak{g}/\mathfrak{c}$ is nilpotent.*

Proof. Suppose that $\mathfrak{g}$ is nilpotent. Then $\mathfrak{g}/\mathfrak{c}$ is the homomorphic image of $\mathfrak{g}$ under the projection $\pi : \mathfrak{g} \to \mathfrak{g}/\mathfrak{c}$. Thus $\mathfrak{g}/\mathfrak{c}$ is nilpotent.

Conversely, suppose that $\mathfrak{g}/\mathfrak{c}$ is nilpotent. If $\pi : \mathfrak{g} \to \mathfrak{g}/\mathfrak{c}$ is the projector, then $\pi(\mathcal{C}^i(\mathfrak{g})) = \mathcal{C}^i(\mathfrak{g}/\mathfrak{c})$. By hypothesis, $\mathcal{C}^k(\mathfrak{g}/\mathfrak{c}) = \{\mathfrak{c}\}$ (the zero subspace in $\mathfrak{g}/\mathfrak{c}$) for some k. Thus, $\mathcal{C}^k(\mathfrak{g}) \subset \mathfrak{c}$, from which we conclude that $\mathcal{C}^{k+1}(\mathfrak{g}) = [\mathfrak{g}, \mathcal{C}^k(\mathfrak{g})] \subset [\mathfrak{g}, \mathfrak{c}] = \{0\}$. $\square$

Exercise 5.1.13. Prove or give a counterexample: suppose that $\mathfrak{a}$ is an ideal of $\mathfrak{g}$. If $\mathfrak{a}$ and $\mathfrak{g}/\mathfrak{a}$ are nilpotent, then $\mathfrak{g}$ is nilpotent. (See Proposition 4.1.10.)

5.2 Engel's Theorem

In this section, our objective is is to prove the following important result.

Theorem 5.2.1. *(Engel's Theorem) Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$. Then $\mathfrak{g}$ is nilpotent if and only if, for all $x \in \mathfrak{g}$, $\operatorname{ad} x$ is a nilpotent linear operator on $\mathfrak{g}$.*

It is easy to prove that if $\mathfrak{g}$ is a nilpotent Lie algebra, then $\operatorname{ad} x$ is a nilpotent linear transformation, for all $x \in \mathfrak{g}$. Indeed, for any $y \in \mathfrak{g}$ and $k \geq 1$, we have

$$(\operatorname{ad} x)^k(y) = \underbrace{[x, [x, [\cdots, [x}_{k \text{ times}}, y]]]] \in \mathcal{C}^k(\mathfrak{g}).$$

Thus, if $\mathcal{C}^k(\mathfrak{g}) = \{0\}$, then $(\operatorname{ad} x)^k = 0$ for all $x \in \mathfrak{g}$.

To prove the opposite implication, we make use of the following lemma, which like Dynkin's lemma (Lemma 4.2.1), asserts the existence of a common eigenvector.

Lemma 5.2.2. *(Engel) Let V be a vector space over $\mathbb{F}$. Suppose that $\mathfrak{g}$ is a subalgebra of $gl(V)$ consisting of nilpotent linear operators on V. Then there exists a vector $v \neq 0$ in V such that $X(v) = 0$ for all $X \in \mathfrak{g}$.*

Proof. We prove this lemma by induction on $\dim \mathfrak{g}$. If $\dim \mathfrak{g} = 0$, there is nothing to prove.

Therefore, we may assume that $n \geq 1$ and that the lemma holds for all Lie subalgebras, of dimension $< n$, of all $\mathrm{gl}(W)$, for all vector spaces W over $\mathbb{F}$. Then, let $\mathfrak{g}$ be an n-dimensional Lie subalgebra of $\mathrm{gl}(V)$, for some vector space V, such that the elements of $\mathfrak{g}$ are all nilpotent linear maps on V. The induction proceeds along several steps:

Step 1: $\mathfrak{g}$ acts on itself via the adjoint representation. We claim that $\operatorname{ad} X$ is a nilpotent linear map on $\mathfrak{g}$, for each $x \in \mathfrak{g}$.

Proof of Step 1: Since $\mathfrak{g} \subset \mathrm{gl}(V)$, the Lie bracket in $\mathfrak{g}$ is the commutator product: $[X, Y] = XY - YX$. For $X \in \mathcal{L}(V)$, let L_X denote left multiplication by X on $\mathcal{L}(V)$: $L_X(Y) := XY$. Likewise, let R_X denote right multiplication by X: $R_X(Y) := YX$. Thus $\operatorname{ad} X\,(Y) = (L_X - R_X)(Y)$. Now, as linear maps on $\mathcal{L}(V)$, L_X and R_X commute: $L_X \circ R_X(Y) = XYX = R_X \circ L_X(Y)$. Hence, by the binomial theorem,

$$\begin{aligned}
(\operatorname{ad} X)^m\,(Y) &= (L_X - R_X)^m(Y) \\
&= \sum_{j=0}^{m} L_X^j \circ R_X^{m-j}\,(Y) \\
&= \sum_{j=0}^{m} X^j\,Y\,X^{m-j},
\end{aligned}$$

for all X, $Y \in \mathfrak{g}$. Now each $X \in \mathfrak{g}$ is a nilpotent linear map on V, so $X^k = 0$ for some k. If we let $m = 2k$ in the above equation, we see that $(\operatorname{ad} X)^m(Y) = 0$ for all $Y \in \mathfrak{g}$. Hence $\operatorname{ad} X$ is nilpotent.

Step 2: Let $\mathfrak{m}$ be a maximal proper subalgebra of $\mathfrak{g}$. ($\mathfrak{m}$ could very well be $\{0\}$.) Then there exists an $X_0 \in \mathfrak{g} \setminus \mathfrak{m}$ such that $[X_0, \mathfrak{m}] \subset \mathfrak{m}$.

Proof of Step 2: $\mathfrak{m}$ acts on the *vector space* $\mathfrak{g}/\mathfrak{m}$ via ad: that is, for each $Z \in \mathfrak{m}$, define the map $\operatorname{ad}' Z$ on $\mathfrak{g}/\mathfrak{m}$ by

$$\operatorname{ad}' Z\,(Y + \mathfrak{m}) = \operatorname{ad} Z\,(Y) + \mathfrak{m}.$$

It is easy to check that $\operatorname{ad}' Z$ is a well-defined linear map on $\mathfrak{g}/\mathfrak{m}$. For each $Z \in \mathfrak{m}$, $\operatorname{ad}' Z$ is a nilpotent linear map, since

$$\left(\operatorname{ad}' Z\right)^m \,(Y + \mathfrak{m}) = (\operatorname{ad} Z)^m \,(Y) + \mathfrak{m} = \mathfrak{m},$$

for sufficiently large m.

Moreover, the map $Z \mapsto \operatorname{ad}' Z$ is easily seen to be a Lie algebra homomorphism from $\mathfrak{m}$ into $\operatorname{gl}(\mathfrak{g}/\mathfrak{m})$. (For this, one just needs to show that $\operatorname{ad}'[Z_1, Z_2] = [\operatorname{ad}' Z_1, \operatorname{ad}' Z_2]$ (for Z_1, $Z_2 \in \mathfrak{m}$), which follows immediately from the same relation for ad.) Thus $\operatorname{ad}' \mathfrak{m}$ is a Lie subalgebra of $\operatorname{gl}(\mathfrak{g}/\mathfrak{m})$ consisting of nilpotent linear maps on $\mathfrak{g}/\mathfrak{m}$.

Since $\dim(\operatorname{ad}' \mathfrak{m}) \le \dim \mathfrak{m} < \dim \mathfrak{g}$, the induction hypothesis says that there exists a nonzero element $X_0 + \mathfrak{m} \in \mathfrak{g}/\mathfrak{m}$ such that $\operatorname{ad}' Z(X_0 + \mathfrak{m}) = \mathfrak{m}$ for all $Z \in \mathfrak{m}$. This means that $X_0 \in \mathfrak{g}$, $X_0 \notin \mathfrak{m}$, and $\operatorname{ad} Z\,(X_0) \in \mathfrak{m}$ for all $Z \in \mathfrak{m}$.

Step 3: $\mathfrak{m} + \mathbb{F}X_0 = \mathfrak{g}$, and $\mathfrak{m}$ is an ideal of $\mathfrak{g}$.

Proof of Step 3: $\mathfrak{m} + \mathbb{F}X_0$ is a subalgebra of $\mathfrak{g}$, since

$$[\mathfrak{m} + \mathbb{F}X_0, \mathfrak{m} + \mathbb{F}X_0] \subset [\mathfrak{m}, \mathfrak{m}] + [\mathfrak{m}, \mathbb{F}X_0] + [\mathbb{F}X_0, \mathfrak{m}] \subset \mathfrak{m}. \tag{5.2}$$

Since $X_0 \notin \mathfrak{m}$, $\mathfrak{m} + \mathbb{F}X_0$ is a subalgebra of $\mathfrak{g}$ properly containing $\mathfrak{m}$. But $\mathfrak{m}$ is a maximal proper subalgebra of $\mathfrak{g}$; thus $\mathfrak{g} = \mathfrak{m} + \mathbb{F}X_0$. Equation (5.2) above shows that $\mathfrak{m}$ is an ideal of $\mathfrak{g}$.

Step 4: There exists a nonzero vector $v \in V$ such that $Xv = 0$ for all $X \in \mathfrak{g}$.

Proof of Step 4: Now $\mathfrak{m}$ is a subalgebra of $\operatorname{gl}(V)$ whose elements are all nilpotent. Since $\dim \mathfrak{m} < \dim \mathfrak{g}$, we can again apply the induction hypothesis to conclude that

$$W := \{v \in V \mid Zv = 0 \text{ for all } Z \in \mathfrak{m}\} \ne \{0\}.$$

W is a joint eigenspace of $\mathfrak{m}$, corresponding to the zero linear functional, so is a subspace of V. Moreover W is X_0-invariant: in fact, for any $w \in W$ and $Z \in \mathfrak{m}$, we have

$$\begin{aligned}
Z(X_0\,(w)) &= (ZX_0 - X_0Z)(w) + X_0Z\,(w) \\
&= [Z, X_0](w) + X_0(0) \\
&= 0,
\end{aligned}$$

since $[Z, X_0] \in \mathfrak{m}$. Thus $X_0(w) \in W$, and W is X_0-invariant. Now the restriction $X_0|_W$ is a nilpotent linear map on W, so X_0 annihilates a nonzero vector $w_0 \in W$: $X_0(w_0) = 0$. Since $\mathfrak{m}$ also annihilates w_0 and $\mathfrak{g} = \mathfrak{m} + \mathbb{F}X_0$, we see that $Y(w_0) = 0$ for all $Y \in \mathfrak{g}$. Putting $v = w_0$, our lemma is proved. $\square$

We are now ready to finish the proof of Engel's Theorem.

Proof of Engel's Theorem. It suffices to prove that if $\mathfrak{g}$ is a Lie algebra such that $\operatorname{ad} x$ is nilpotent for all $x \in \mathfrak{g}$, then $\mathfrak{g}$ is nilpotent. We will do this by induction on $\dim \mathfrak{g}$, the cases $\dim \mathfrak{g} = 0$ and $\dim \mathfrak{g} = 1$ being trivial.

So assume that the result given above holds for all Lie algebras of dimension $< n$, and that $\dim \mathfrak{g} = n$, with $\operatorname{ad} x$ nilpotent for all $x \in \mathfrak{g}$.

Then $\operatorname{ad} \mathfrak{g}$ is a Lie subalgebra of $\mathrm{gl}(\mathfrak{g})$ consisting of nilpotent linear maps, so by Lemma 5.2.2, there exists a nonzero element $z \in \mathfrak{g}$ such that $\operatorname{ad} x\,(z) = 0$ for all $x \in \mathfrak{g}$. This implies that z lies in the center $\mathfrak{c}$ of $\mathfrak{g}$, so $\mathfrak{c} \neq \{0\}$.

Consider the quotient algebra $\mathfrak{g}' = \mathfrak{g}/\mathfrak{c}$. Let $\operatorname{ad}_{\mathfrak{c}}$ denote its adjoint representation. Then

$$\operatorname{ad}_{\mathfrak{c}}(x + \mathfrak{c})\,(y + \mathfrak{c}) = [x + \mathfrak{c}, y + \mathfrak{c}] = [x, y] + \mathfrak{c}$$

For each $x \in \mathfrak{g}$, $\operatorname{ad}_{\mathfrak{c}}(x + \mathfrak{c})$ is a nilpotent linear map on $\mathfrak{g}/\mathfrak{c}$. Thus, by the induction hypothesis (since $\dim(\mathfrak{g}/\mathfrak{c}) < \dim \mathfrak{g}$), the Lie algebra $\mathfrak{g}/\mathfrak{c}$ is nilpotent. Then by Proposition 5.1.12, $\mathfrak{g}$ is nilpotent.

This completes the induction step and the proof of Engel's Theorem.

The following theorem is the nilpotent analogue of Lie's Theorem (Theorem 4.2.3). While Lie's Theorem only holds for complex vector spaces, the theorem below holds for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Theorem 5.2.3. *(Engel's Structure Theorem) Let $\mathfrak{g}$ be a Lie algebra consisting of nilpotent linear maps acting on a vector space V. Then there is a basis of V relative to which the matrix of every element of $\mathfrak{g}$ is strictly upper triangular.*

Proof. The proof is by induction on $\dim V$. If $\dim V = 0$ or $\dim V = 1$, this result is trivial, since any nilpotent linear map on V is just 0.

Assume that the result is true for dimension $n - 1$, and let V have dimension n. Now by Lemma 5.2.2, there exists a nonzero vector v_1 such that $X(v_1) = 0$ for all $X \in \mathfrak{g}$. Let $V_1 = \mathbb{F}v_1$. Since V_1 is $\mathfrak{g}$-invariant, $\mathfrak{g}$ acts on the quotient space V/V_1 via

$$X \cdot (v + V_1) = X(v) + V_1$$

for all $v \in V$ and $X \in \mathfrak{g}$. Each $X \in \mathfrak{g}$ is clearly a nilpotent linear map on V/V_1. Hence V/V_1 has a basis $(v_2 + V_1, \ldots, v_n + V_1)$ relative to which the matrix of each $X \in \mathfrak{g}$ is strictly upper triangular. Then $(v_1, v_2, \ldots, v_n)$ is a basis of V, and it is also clear that the matrix of each $X \in \mathfrak{g}$ with respect to this basis is strictly upper triangular. $\square$

Theorem 5.2.4. *Let $\mathfrak{g}$ be a solvable Lie algebra over $\mathbb{C}$. Then $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ is nilpotent.*

Proof. By Lie's Abstract Theorem (Theorem 4.2.5), there is a basis of $\mathfrak{g}$ with respect to which the matrix of $\operatorname{ad} x$ is upper triangular, for each $x \in \mathfrak{g}$. Thus, for any x, $y \in \mathfrak{g}$, the matrix of $\operatorname{ad}[x, y] = [\operatorname{ad} x, \operatorname{ad} y] = \operatorname{ad} x \circ \operatorname{ad} y - \operatorname{ad} y \circ \operatorname{ad} x$ with respect to this basis is strictly upper triangular. Since strictly upper triangular matrices correspond to nilpotent linear maps, it follows that $\operatorname{ad} w$ is nilpotent for each $w \in \mathfrak{g}'$. Therefore the restriction $\operatorname{ad} w|_{\mathfrak{g}'}$ is also nilpotent. By Engel's Theorem, this implies that $\mathfrak{g}'$ is nilpotent. $\square$

The converse holds for any field $\mathbb{F}$.

Theorem 5.2.5. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$ such that $\mathfrak{g}'$ is nilpotent. Then $\mathfrak{g}$ is solvable.*

Proof. Since $\mathfrak{g}'$ is nilpotent, by Exercise 5.1.6 it is solvable. Moreover, the quotient Lie algebra $\mathfrak{g}/\mathfrak{g}'$ is abelian, hence solvable. Thus, by Proposition 4.1.10, $\mathfrak{g}$ is solvable. $\square$

Chapter 6

Cartan's Criteria for Solvability and Semisimplicity

In this section we define an important symmetric bilinear form on a Lie algebra $\mathfrak{g}$ and derive conditions on this form which are necessary and sufficient for $\mathfrak{g}$ to be solvable, as well as conditions on the form which are necessary and sufficent for $\mathfrak{g}$ to be semisimple.

6.1 The Killing Form

For any elements x and y of $\mathfrak{g}$, the map $\operatorname{ad} x \circ \operatorname{ad} y$ is a linear operator on $\mathfrak{g}$, so we may consider its trace.

Definition 6.1.1. Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$. The *Killing form* on $\mathfrak{g}$ is the map

$$\begin{aligned} B\colon \mathfrak{g} \times \mathfrak{g} &\to \mathbb{F} \\ (x, y) &\mapsto \operatorname{tr}(\operatorname{ad} x \circ \operatorname{ad} y) \end{aligned} \tag{6.1}$$

Thus, $B(x, y) = \operatorname{tr}(\operatorname{ad} x \circ \operatorname{ad} y)$.

Note that $B(y, x) = \operatorname{tr}(\operatorname{ad} y \circ \operatorname{ad} x) = \operatorname{tr}(\operatorname{ad} x \circ \operatorname{ad} y) = B(x, y)$, so the Killing form B is symmetric.

Since the adjoint map ad and the trace are linear, it is also easy to see that

$B(x, y)$ is a bilinear form:

$$\begin{aligned} B(\alpha\, x_1 + \beta x_2, y) &= \operatorname{tr}(\operatorname{ad}(c\, x_1 + x_2) \circ \operatorname{ad} y) \\ &= \operatorname{tr}((\alpha \operatorname{ad} x_1 + \beta \operatorname{ad} x_2) \circ \operatorname{ad} y) \\ &= \alpha \operatorname{tr}(\operatorname{ad} x_1 \circ \operatorname{ad} y) + \beta \operatorname{tr}(\operatorname{ad} x_2 \circ \operatorname{ad} y) \\ &= \alpha\, B(x_1, y) + \beta B(x_2, y), \end{aligned}$$

for all $x_1,\ x_2,\ y \in \mathfrak{g}$, and all $\alpha,\ \beta \in \mathbb{F}$. (The linearity of B in the second argument follows from its the above and the fact that B is symmetric.)

Exercise 6.1.2. (Graduate Exercise.) Suppose that $\mathfrak{g}$ is the Lie algebra of a compact Lie group G. Prove that B is negative semidefinite; i.e., $B(x, x) \leq 0$ for all $x \in \mathfrak{g}$. If $\mathfrak{c} = \{0\}$, show that B is negative definite. (*Hint:* There exists an inner product Q on $\mathfrak{g}$ invariant under $\operatorname{ad} G$: $B(\operatorname{Ad} g\,(x), \operatorname{Ad} g\,(y)) = B(x, y)$ for all $x,\ y \in \mathfrak{g}$ and all $g \in G$.)

Our objective in this section is to prove the following theorems.

Theorem 6.1.3. *(Cartan's Criterion for Solvability) Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$. Then $\mathfrak{g}$ is solvable if and only if $B(x, y) = 0$ for all $x \in [\mathfrak{g}, \mathfrak{g}]$ and $y \in \mathfrak{g}$.*

Theorem 6.1.4. *(Cartan's Criterion for Semisimplicity) Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$. Then $\mathfrak{g}$ is semisimple if and only if the Killing form B is nondegenerate.*

We will first prove Theorem 6.1.3 for complex Lie subalgebras of $\mathrm{gl}\,(V)$, where V is a complex vector space. Then, in order to prove it for real Lie algebras, we will need to use the notion of complexification. Theorem 6.1.4 will then essentially be a corollary of Theorem 6.1.3.

Let us now develop the necessary machinery.

6.2 The Complexification of a Real Lie Algebra

A vector space V over $\mathbb{R}$ is said to have a *complex structure* if there is a $J \in \mathcal{L}(V)$ such that $J^2 = -I_V$. Note that, by definition, J is $\mathbb{R}$-linear, and that the condition $J^2 = -I_V$ means that it is invertible. Note also that J has no real eigenvalues, since its only possible eigenvalues are $\pm i$.

The operator J turns the vector space V into a complex vector space in which scalar multiplication by $z = \alpha + \beta i$ (with $\alpha,\ \beta \in \mathbb{R}$) is given by

$$(\alpha + \beta i)\, v = \alpha\, v + \beta\, Jv, \tag{6.2}$$

for all $v \in V$. The routine verification that V is indeed a complex vector space will be left to the reader.

Example 6.2.1. For any real vector space U, let V be the external direct sum $V = U \oplus U$. Then the linear operator J on V given by $J\,(u_1, u_2) = (-u_2, u_1)$ is a complex structure on V. Since $(u_1, u_2) = (u_1, 0) + J(u_2, 0) = (u_1, 0) + i(u_2, 0)$, it is often convenient to identify U with the subspace $\{(u, 0) \,|\, u \in U\}$, and thus write the element (u_1, u_2) as $u_1 + Ju_2$. Since V now has a complex vector space structure, we call V the *complexification of* U, and denote it by U^c.

Note: For the advanced student, complexification can be carried out using tensor products: $V = U \otimes_{\mathbb{R}} \mathbb{C}$, but we'll not go through this route.

Suppose that J is a complex structure on a real vector space V. Then V becomes a complex vector space, with scalar multiplcation given by (6.2) above. Since any $\mathbb{R}$-spanning set in V is also a $\mathbb{C}$-spanning set, it is obvious that V is finite-dimensional as a complex vector space. Now let $(u_1, \dots, u_n)$ be a $\mathbb{C}$-basis of V. Then $(u_1, \dots, u_n, Ju_1, \dots, Ju_n)$ is an $\mathbb{R}$-basis of V: in fact, any $v \in V$ can be written as a *unique* linear combination

$$v = \sum_{j=1}^{n} (\alpha_j + i\beta_j)\, u_j = \sum_{j=1}^{n} \alpha_j\, u_j \;+\; \sum_{j=1}^{n} \beta_j\, Ju_j \qquad\qquad (\alpha_j, \beta_j \in \mathbb{R})$$

Now if we let U be the real subspace $\mathbb{R}u_1 + \cdots + \mathbb{R}u_n$ of V, we see that $V = U \oplus JU$ (as a real vector space), and is thus easy to see that $V \cong U^c$. In particular, $\dim_{\mathbb{R}} V = 2n = 2 \dim_{\mathbb{C}} V$, so any real vector space with a complex structure is even-dimensional over $\mathbb{R}$. The subspace U is called a *real form* of V.

Of course, any complex vector space V is a real vector space equipped with a complex structure: $Jv = iv$, for all $v \in V$. In the future, we will nonetheless have occasion to complexify a *complex* vector space (considered as a real vector space) using the construction in Example 6.2.1.

So suppose that U is a complex vector space. Considering U as a real vector space, we can then complexify U in accordance with Example 6.2.1. Now the external direct sum $U^c = U \oplus U = U \times U$ is already a complex vector space, since each factor is a complex vector space. The complex structure J on $U \oplus U$ commutes with multiplication by i, since

$$J(i(u_1, u_2)) = J(iu_1, iu_2) = (-iu_2, iu_1) = i(-u_2, u_1) = iJ(u_1, u_2).$$

Thus J is a $\mathbb{C}$-linear map on $U \oplus U$. $U \oplus U$ decomposes into a direct sum of $\pm i$-eigenspaces of J:

$$(u_1, u_2) = \frac{1}{2}\,(u_1 + iu_2, u_2 - iu_1) + \frac{1}{2}\,(u_1 - iu_2, u_2 + iu_1),$$

so

$$U^c = \{(v, -iv) \,|\, v \in U\} \;\oplus\; \{(w, iw) \,|\, w \in U\}.$$

If U is a real form of complex vector space V, we define the *conjugation* τ_U *of* V *with respect to* U as follows: for any $v \in V$, we can write v uniquely as

$v = u_1 + iu_2$, where $u_1, u_2 \in U$; put $\tau_U(v) = u_1 - iu_2$. Then τ_U is an $\mathbb{R}$-linear map of V satisfying $\tau_U^2 = I_V$. It is easy to check that τ_U is conjugate-linear: $\tau_U(zv) = \overline{z}\,\tau_U(v)$, for all $v \in V$ and $z \in \mathbb{C}$.

Exercise 6.2.2. (i) Suppose that T is a $\mathbb{C}$-linear operator on a complex vector space V. Show that if T_R denotes T considered as an $\mathbb{R}$-linear operator on V, then $\mathrm{tr}\,(T_R) = 2\,\mathrm{Re}(\mathrm{tr}\,(T))$. (ii) Next suppose that T is an $\mathbb{R}$-linear operator on a real vector space U. Show that T has a unique natural extension T^c to a $\mathbb{C}$-linear map on U^c, and that $\mathrm{tr}\,(T_c) = \mathrm{tr}\,(T)$.

Now suppose that $\mathfrak{g}$ is a real Lie algebra equipped with a complex structure J. J is said to be *compatible* with the Lie bracket in $\mathfrak{g}$ if $[Jx, y] = J[x, y]$ for all $x, y \in \mathfrak{g}$. (Then, of course $[x, Jy] = J[x, y]$ for all x, y.) If $\mathfrak{g}$ is given the complex vector space from (6.2), then multiplication by complex scalars commutes with the Lie bracket, since

$$i\,[x, y] = J\,[x, y] = [Jx, y] = [ix, y] = [x, Jy] = [x, iy].$$

Hence, $\mathfrak{g}$ has the structure of a complex Lie algebra. Of course, the Lie bracket of any complex Lie algebra is compatible with its complex structure.

A *real form* of a complex Lie algebra $\mathfrak{g}$ is a real Lie subalgebra $\mathfrak{g}_0$ of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{g}_0 \oplus i\mathfrak{g}_0$. (A real form of $\mathfrak{g}$ (as vector space) is not necessarily a real Lie subalgebra of $\mathfrak{g}$. For example, $\mathbb{R}e + \mathbb{R}f + \mathbb{R}ih$ is a real form of the vector space $\mathrm{sl}\,(2, \mathbb{C})$ but is not a real Lie subalgebra.) It is easy to check that if τ denotes the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$, then $\tau[x, y] = [\tau x, \tau y]$, for all $x, y \in \mathfrak{g}$.

Exercise 6.2.3. Let $\mathrm{u}(n)$ denote the Lie algebra of *skew-Hermitian matrices*; i.e., $\mathrm{u}(n) = \{X \in \mathrm{gl}\,(n, \mathbb{C}) \mid {}^t\overline{X} = -X\}$. (See Example 2.1.23.) Prove that $\mathrm{u}(n)$ is a real form of $\mathrm{gl}\,(n, \mathbb{C})$. If τ denotes the conjugation of $\mathrm{gl}\,(n, \mathbb{C})$ with respect to $\mathrm{u}(n)$, show that $\tau(X) = -{}^t\overline{X}$ for all $X \in \mathrm{gl}\,(n, \mathbb{C})$.

Next, suppose that $\mathfrak{g}_0$ is a real Lie algebra The Lie bracket in $\mathfrak{g}_0$ can be extended to its vector space complexification $\mathfrak{g} = \mathfrak{g}_0^c = \mathfrak{g}_0 \oplus J\mathfrak{g}_0$ via

$$[x_1 + Jx_2, y_1 + Jy_2] = [x_1, y_1] - [x_2, y_2] + J([x_1, y_2] + [x_2, y_1]).$$

The operation above is $\mathbb{R}$-bilinear and can be routinely verified to be anticommutative and to satisfy the Jacobi identity. Moreover

$$\begin{aligned} J[x_1 + Jx_2, y_1 + Jy_2] &= -([x_1, y_2] + [x_2, y_1]) + J([x_1, y_1] - [x_2, y_2]) \\ &= [-x_2 + Jx_1, y_1 + Jy_2] \\ &= [J(x_1 + Jx_2), y_1 + Jy_2], \end{aligned}$$

and so it follows that this extension of the Lie bracket to $\mathfrak{g}$ is $\mathbb{C}$-bilinear. Thus the complexification $\mathfrak{g}$ has the structure of a complex Lie algebra, and of course, $\mathfrak{g}_0$ is a real form of $\mathfrak{g}$.

Example 6.2.4. It is obvious that $\mathrm{gl}\,(n,\mathbb{R})$ is a real form of $\mathrm{gl}\,(n,\mathbb{C})$. We can express this as $(\mathrm{gl}\,(\mathbb{R}^n))^c = \mathrm{gl}\,(\mathbb{C}^n)$. Now any real vector space V is (duh) a real form of its complexification V^c. If we fix a basis B of V, then the map $T \mapsto M_{B,B}(T)$ identifies $\mathrm{gl}\,(V)$ with $\mathrm{gl}\,(n,\mathbb{R})$. Complexifying this identification, we see that $(\mathrm{gl}\,(V))^c \cong (\mathrm{gl}\,(\mathbb{R}^n))^c \cong \mathrm{gl}\,(\mathbb{C}^n) \cong \mathrm{gl}\,(V^c)$. This identification of $(\mathrm{gl}\,(V))^c$ with $\mathrm{gl}\,(V^c)$ is concretely given by

$$(T_1 + iT_2)\,(v_1 + iv_2) = T_1v_1 - T_2v_2 + i(T_1v_2 + T_2v_1),$$

for all T_1, $T_2 \in \mathrm{gl}\,(V)$ and all v_1, $v_2 \in V$.

If $\mathfrak{g}$ is a Lie subalgebra of $\mathrm{gl}\,(V)$, then its complexification $\mathfrak{g}^c$ may be viewed as a complex Lie subalgebra of $\mathrm{gl}\,(V^c)$.

Exercise 6.2.5. Let $\mathfrak{g}$ be a complex Lie algebra, and let $\mathfrak{g}_R$ denote $\mathfrak{g}$, considered as a real Lie algebra. If B and B_R are the Killing forms on $\mathfrak{g}$ and $\mathfrak{g}_R$, respectively, show that $B_R(x,y) = 2\,\mathrm{Re}\,(B(x,y)$, for all $x, y \in \mathfrak{g}$. Then show that B is nondegenerate $\iff$ B_R is nondegenerate.

(*Note: If you don't know anything about Lie groups, you may safely skip this paragraph.*) A real Lie algebra $\mathfrak{u}$ is said to be *compact* if $\mathfrak{u}$ is the Lie algebra of a compact Lie goup U. Here are two interesting and useful facts about compact Lie algebras (cf. [4], Chapter 3):

1. If $\mathfrak{u}$ is compact, then $\mathfrak{u} = \mathfrak{c} \oplus [\mathfrak{u},\mathfrak{u}]$, where $\mathfrak{c}$ is the center of $\mathfrak{u}$ and $[\mathfrak{u},\mathfrak{u}]$ is compact and semisimple.
2. Any complex semisimple Lie algebra $\mathfrak{g}$ has a compact real form $\mathfrak{u}$. This remarkable fact is a cornerstone of representation theory. In a later section, we will consider how to obtain such a real form.

6.3 Cartan's Criterion for Solvability.

Lemma 6.3.1. *Let V be a vector space over $\mathbb{C}$, and let $X \in gl(V)$. If X is semisimple, then so is $ad\,X$. If X is nilpotent, then so is $ad\,X$.*

Proof. If X is nilpotent, then so is $\mathrm{ad}\,X$ by Step 1 in the proof of Lemma 5.2.2.

Suppose that X is semisimple. Let $B = (v_1, \ldots, v_n)$ be a basis of of V consisting of eigenvectors of X, corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively. Next, we abuse notation and let, for any i, j in $\{1, \ldots, n\}$, E_{ij} be the linear operator on V whose matrix with respect to B is the elementary matrix E_{ij}:

$$E_{ij}(v_k) := \delta_{jk}\, v_i$$

Then $(E_{ij})_{1\leq i,j\leq n}$ is a basis of $\mathrm{gl}\,(V)$. By matrix multiplication, we see that

$$[X, E_{ij}] = (\lambda_i - \lambda_j)\, E_{ij}. \tag{6.3}$$

Thus, each E_{ij} is an eigenvector of $\mathrm{ad}\, X$ with eigenvalue $\lambda_i - \lambda_j$, and so $\mathrm{gl}\,(V)$ has basis (E_{ij}) consisting of eigenvectors of $\mathrm{ad}\, X$. Thus $\mathrm{ad}\, X$ is semisimple. ☐

Lemma 6.3.2. *Suppose that $X \in gl(V)$ has Jordan-Chevalley decomposition $X = X_s + X_n$, with X_s semisimple and X_n nilpotent. Then the Jordan-Chevalley decomposition for $ad\, X$ is $ad\, X = ad\, X_s + ad\, X_n$.*

Proof. By the preceding lemma, $\mathrm{ad}\, X_s$ and $\mathrm{ad}\, X_n$ are semisimple and nilpotent linear operators on $\mathrm{gl}\,(V)$, respectively. Moreover, $\mathrm{ad}\, X_s$ and $\mathrm{ad}\, X_n$ commute:

$$\begin{aligned}[\mathrm{ad}\, X_s, \mathrm{ad}\, X_n] &= \mathrm{ad}\, [X_s, X_n] \\ &= \mathrm{ad}\, (X_s X_n - X_n X_s) \\ &= 0,\end{aligned}$$

since X_s and X_n commute. By the uniqueness of the Jordan-Chevalley decomposition, it follows that $\mathrm{ad}\, X = \mathrm{ad}\, X_s + \mathrm{ad}\, X_n$ is the Jordan-Chevalley decomposition of $\mathrm{ad}\, X$. ☐

Lemma 6.3.3. *Let $x_0, x_1, \ldots, x_n$ be distinct numbers in $\mathbb{F}$, and let $C_0, C_1, \ldots, C_n$ be any numbers in $\mathbb{F}$. Then there exists a polynomial $P(x)$ in the variable x, with coefficients in $\mathbb{F}$, of degree $\leq n$, such that $P(x_i) = C_i$, for all i.*

Proof. According to the *Lagrange Interpolation Formula*, this polynomial is given by

$$P(x) = \frac{\prod_{i\neq 0}(x - x_i)}{\prod_{i\neq 0}(x_0 - x_i)}\, C_0 \,+\, \frac{\prod_{i\neq 1}(x - x_i)}{\prod_{i\neq 1}(x_1 - x_i)}\, C_1 \,+\, \cdots +\, \frac{\prod_{i\neq n}(x - x_i)}{\prod_{i\neq n}(x_n - x_i)}\, C_n$$

It is easy to see that this $P(x)$ satisfies the properties asserted in the lemma. ☐

Exercise 6.3.4. (Graduate Exercise.) Show that any polynomial satisfying the conclusion of Lemma 6.3.3 is unique. (*Hint:* The formula above comes from a linear system whose coefficient matrix is Vandermonde.)

The following is a technical lemma whose proof features some “out of the box” thinking.

Lemma 6.3.5. *Let V be a vector space over $\mathbb{C}$, and let $A \subset B$ be subspaces of $gl(V)$. Let $\mathfrak{m} = \{X \in gl(V) \,|\, [X, B] \subset A\}$. Suppose that some $X \in \mathfrak{m}$ has the property that $tr\,(XY) = 0$ for all $Y \in \mathfrak{m}$. Then X is nilpotent.*

Proof. Let $S = (v_1, \dots, v_n)$ be a Jordan basis of V corresponding to X. If $X = X_s + X_n$ is the Jordan-Chevalley decomposition of X, then S consists of eigenvectors of X_s, and the matrix of X_s with respect to S is diagonal, of the form

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \tag{6.4}$$

The matrix of X_n with respect to S is strictly upper triangular, with some 1's right above the diagonal. We want to show that $\lambda_i = 0$ for all i. This will establish that $X_s = 0$, and so $X = X_n$.

Let $E \subset \mathbb{C}$ be the vector space over $\mathbb{Q}$ (= the rationals) spanned by $\lambda_1, \dots, \lambda_n$. We'll show that $E = \{0\}$. This will, of course, show that each $\lambda_i = 0$. If E^* denotes the dual space (over $\mathbb{Q}$) of E, standard linear algebra says that $\dim_{\mathbb{Q}} E^* = \dim_{\mathbb{Q}} E$. (See Subsection 1.3.) Thus it's sufficient to prove that $E^* = \{0\}$. That is, we will prove that any $\mathbb{Q}$-linear functional on E must vanish identically.

So let $f \in E^*$. Then let $Y \in \mathrm{gl}\,(V)$ be the linear map on V whose matrix with respect to the basis S above is the diagonal matrix

$$\begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix}$$

Now by equation (6.3), the semisimple operator X_s satisfies

$$\operatorname{ad} X_s\,(E_{ij}) = (\lambda_i - \lambda_j)\,E_{ij}. \tag{6.5}$$

For the same reason, the semisimple operator Y satisfies

$$\operatorname{ad} Y\,(E_{ij}) = (f(\lambda_i) - f(\lambda_j))\,E_{ij}. \tag{6.6}$$

According to Lemma 6.3.3, there exists a polynomial $G(x)$ in the variable x, with complex coefficients, such that

$$\begin{aligned} G(0) &= 0, & \\ G(\lambda_i) &= f(\lambda_i) & \text{for all } i = 1, \dots, n \\ G(\lambda_i - \lambda_j) &= f(\lambda_i) - f(\lambda_j) & \text{for all } i, j = 1, \dots, n \end{aligned}$$

This polynomial is well-defined, since if $\lambda_i - \lambda_j = \lambda_k - \lambda_l$, then $f(\lambda_i) - f(\lambda_j) = f(\lambda_i - \lambda_j) = f(\lambda_k - \lambda_l) = f(\lambda_k) - f(\lambda_l)$. There are at most $2\binom{n}{2} + n + 1 = n^2 + n + 1$ elements in the set $\{0\} \cup \{\lambda_i\}_{i=1}^n \cup \{\lambda_i - \lambda_j\}_{i,j=1}^n$, so $G(x)$ can be assumed to have degree $\leq n^2 + n$, but this does not matter.

Let us now compute the linear operator $G(\operatorname{ad} X_s)$ on $\mathrm{gl}(V)$. It suffices to do this on each elementary matrix E_{ij}. Now, by equation (6.5), $\operatorname{ad} X_s\,(E_{ij}) = (\lambda_i - \lambda_j)\,E_{ij}$, and so

$$\begin{aligned} G(\operatorname{ad} X_s)\,(E_{ij}) &= G(\lambda_i - \lambda_j)\,E_{ij} \\ &= (f(\lambda_i) - f(\lambda_j))\,E_{ij} \\ &= \operatorname{ad} Y\,(E_{ij}), \end{aligned}$$

the last equation coming from (6.6). It follows that $G(\operatorname{ad} X_s) = \operatorname{ad} Y$. (Note also that the condition $G(\lambda_i) = f(\lambda_i)$ for all i implies that $G(X_s) = Y$.)

Now by Lemma 6.3.2, the semisimple part of $\operatorname{ad} X$ is $\operatorname{ad} X_s$, which by Theorem 1.9.14 is a polynomial in $\operatorname{ad} X$ with zero constant term. Since the polynomial $G(x)$ also has zero constant term, we see that $\operatorname{ad} Y = G(\operatorname{ad} X_s)$ is a polynomial in $\operatorname{ad} X$ with zero constant term:

$$\operatorname{ad} Y = a_r\,(\operatorname{ad} X)^r + a_{r-1}\,(\operatorname{ad} X)^{r-1} + \cdots + a_1 \operatorname{ad} X.$$

Since, by hypothesis, $\operatorname{ad} X\,(B) \subset A$, it follows from the above (and the fact that $A \subset B$) that $\operatorname{ad} Y\,(B) \subset A$. Therefore, by the definition of $\mathfrak{m}$, we see that $Y \in \mathfrak{m}$.

Now by the hypothesis on X, we have $\operatorname{tr}(XY) = 0$. With respect to the basis S of V, the product XY has matrix

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix} = \begin{pmatrix} \lambda_1 f(\lambda_1) & & * \\ & \ddots & \\ 0 & & \lambda_n f(\lambda_n) \end{pmatrix}$$

and so $\operatorname{tr}(XY) = \sum_{i=1}^n \lambda_i f(\lambda_i)$.

Thus $0 = \sum_{i=1}^n \lambda_i f(\lambda_i)$. Applying the linear functional f to this equality, we get $0 = \sum_{i=1}^n f(\lambda_i)^2$. Since the $f(\lambda_i)$ are all in $\mathbb{Q}$, we conclude that $f(\lambda_i) = 0$ for all i. Thus $f = 0$, so $E^* = \{0\}$, so $E = \{0\}$, and so $\lambda_i = 0$ for all i.

We conclude that $X = X_n$, and the lemma is proved. □

Lemma 6.3.6. *Let V be a vector space over $\mathbb{F}$. If $X, Y, Z \in gl(V)$, then $tr([X,Y]Z) = tr(X[Y,Z])$.*

This follows from

$$\begin{aligned} \operatorname{tr}([X,Y]Z) &= \operatorname{tr}((XY - YX)Z) \\ &= \operatorname{tr}(XYZ - YXZ) \\ &= \operatorname{tr}(XYZ) - \operatorname{tr}(YXZ) \\ &= \operatorname{tr}(XYZ) - \operatorname{tr}(XZY) \\ &= \operatorname{tr}(X(YZ - ZY)) \\ &= \operatorname{tr}(X[Y,Z]). \end{aligned}$$

The following theorem gives the version of Cartan's solvability criterion (Theorem 6.1.3) for Lie subalgebras of $\mathrm{gl}(V)$.

Theorem 6.3.7. *(Cartan's Criterion for $gl(V)$, V complex.) Let V be a vector space over $\mathbb{C}$, and let $\mathfrak{g}$ be a Lie subalgebra of $gl(V)$. Then $\mathfrak{g}$ is solvable if and only if $tr(XY) = 0$ for all $X \in [\mathfrak{g}, \mathfrak{g}]$ and all $Y \in \mathfrak{g}$.*

Proof. Suppose that $\mathfrak{g}$ is solvable. Then by Lie's Theorem (Theorem 4.2.3), there is a basis S of V relative to which every element of $\mathfrak{g}$ has an upper triangular matrix. It follows that every element of $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ has a strictly upper triangular matrix relative to this basis. If $X \in \mathfrak{g}'$ and $Y \in \mathfrak{g}$, it is easy to see that the matrix of XY with respect to S is also strictly upper triangular. Thus $\mathrm{tr}(XY) = 0$.

Conversely, suppose that $\underline{\mathrm{tr}(XY) = 0 \text{ for all } X \in \mathfrak{g}' \text{ and } Y \in \mathfrak{g}}$. We want to prove that $\mathfrak{g}$ is solvable. By Theorem 5.2.5, it suffices to prove that $\mathfrak{g}'$ is nilpotent. For this, it suffices in turn to prove that every $X \in \mathfrak{g}'$ is a nilpotent linear operator on V. For then, by Step 1 in the proof of Engel's Lemma (Lemma 5.2.2), $\mathrm{ad}\, X$ is nilpotent, and so by Engel's Theorem (Theorem 5.2.1), $\mathfrak{g}'$ is nilpotent.

To this end, we will use Lemma 6.3.5 with $A = \mathfrak{g}'$ and $B = \mathfrak{g}$. The subspace $\mathfrak{m}$ will then be $\{Y \in \mathrm{gl}(V) \mid [Y, \mathfrak{g}] \subset \mathfrak{g}'\}$. Clearly, $\mathfrak{g}' \subset \mathfrak{m}$. To apply the lemma, we will need to prove that $\mathrm{tr}(XY) = 0$ for all $X \in \mathfrak{g}'$ and all $Y \in \mathfrak{m}$.

Now $\mathfrak{g}'$ is spanned by the brackets $[Z, W]$, for all Z, $W \in \mathfrak{g}$. Suppose that $Y \in \mathfrak{m}$. By Lemma 6.3.6, $\mathrm{tr}([Z, W]Y) = \mathrm{tr}(Z[W, Y]) = \mathrm{tr}([W, Y]Z)$. But $[W, Y] \in \mathfrak{g}'$ (by the definition of $\mathfrak{m}$). So, by our underlined hypothesis above, $\mathrm{tr}([W, Y]Z) = 0$.

This shows that $\mathrm{tr}(XY) = 0$ for all generators $X = [Z, W]$ of $\mathfrak{g}'$ and all $Y \in \mathfrak{m}$. Since the trace is linear, we conclude that $\mathrm{tr}(XY) = 0$ for all $X \in \mathfrak{g}'$ and $Y \in \mathfrak{m}$. Hence, by Lemma 6.3.5, each $X \in \mathfrak{g}'$ is nilpotent, and the theorem is proved. $\square$

Corollary 6.3.8. *(Cartan's Criterion for $gl(V)$, V real.) Let V be a vector space over $\mathbb{R}$, and let $\mathfrak{g}$ be a Lie subalgebra of $gl(V)$. Then $\mathfrak{g}$ is solvable if and only if $tr(XY) = 0$ for all $X \in \mathfrak{g}'$ and all $Y \in \mathfrak{g}$.*

Proof. Let $B = (v_1, \dots, v_n)$ be a fixed basis of V. Then B is also a complex basis of V^c. As remarked earlier, the map $T \mapsto M_{B,B}(T)$ is a Lie algebra isomorphism of $\mathrm{gl}(V)$ onto $\mathrm{gl}(n, \mathbb{R})$ and $\mathrm{gl}(V^c)$ onto $\mathrm{gl}(n, \mathbb{C})$. Thus it suffices to prove this corollary for Lie subalgebras $\mathfrak{g}$ of $\mathrm{gl}(n, \mathbb{R})$.

The derived algebra $(\mathfrak{g}^c)'$ is the linear span of elements of the form $[X_1 + iX_2, Y_1 + iY_2] = [X_1, X_2] - [Y_1, Y_2] + i([X_1, Y_2] + [X_2, Y_1])$, where the X_i, $Y_j \in \mathfrak{g}$, and from this it is not hard to see that $(\mathfrak{g}^c)' = (\mathfrak{g}')^c$. By induction, we conclude that $(\mathfrak{g}^c)^{(r)} = (\mathfrak{g}^{(r)})^c$. This in turn shows that $\mathfrak{g}$ is solvable $\iff$ $\mathfrak{g}^c$ is solvable.

Now if $\mathfrak{g}$ is solvable, then $\mathfrak{g}^c$ is a solvable Lie subalgebra of $\mathrm{gl}\,(n, \mathbb{C})$. Hence by Theorem 6.3.7, $\mathrm{tr}\,(XY) = 0$ for all $X \in (\mathfrak{g}^c)'$ and all $Y \in \mathfrak{g}^c$. In particular, $\mathrm{tr}\,(XY) = 0$ for all $X \in \mathfrak{g}'$ and $Y \in \mathfrak{g}$.

Conversely, suppose that $\mathrm{tr}\,(XY) = 0$ for all $X \in \mathfrak{g}'$ and $Y \in \mathfrak{g}$. We want to prove that $\mathrm{tr}\,(ZW) = 0$ for all $Z \in (\mathfrak{g}^c)'$ and all $W \in \mathfrak{g}^c$. But then we can resolve Z into its real and imaginary components: $Z = X_1 + iX_2$, where $X_1, X_2 \in \mathfrak{g}'$. Likewise, $W = Y_1 + iY_2$, with $Y_1, Y_2 \in \mathfrak{g}$. Thus $\mathrm{tr}\,(ZW) = \mathrm{tr}\,(X_1Y_1 - X_2Y_2) + i\mathrm{tr}\,(X_1Y_2 + X_2Y_1) = 0$. By Theorem 6.3.7, we conclude that $\mathfrak{g}^c$ is solvable, and hence $\mathfrak{g}$ also is. $\square$

We are now ready to prove Cartan's criterion for solvability.

Proof of Theorem 6.1.3: Suppose first that $\mathfrak{g}$ is solvable. Then $\mathrm{ad}\,\mathfrak{g}$ is solvable by Proposition 4.1.10, since $\mathrm{ad}\,\mathfrak{g} \cong \mathfrak{g}/\mathfrak{c}$, where $\mathfrak{c}$ is the center of $\mathfrak{g}$. Thus $\mathrm{ad}\,\mathfrak{g}$ is a solvable subalgebra of $\mathrm{gl}\,(\mathfrak{g})$, and so by the "only if" part of Theorem 6.3.7 (for $\mathbb{F} = \mathbb{C}$) or Corollary 6.3.8 (for $\mathbb{F} = \mathbb{R}$), we conclude that $B(x, y) = \mathrm{tr}\,(\mathrm{ad}\,x \circ \mathrm{ad}\,y) = 0$ for all $x \in [\mathfrak{g}, \mathfrak{g}]$ and all $y \in \mathfrak{g}$.

Conversely, suppose that $B(x, y) = 0$ for all $x \in \mathfrak{g}'$ and all $y \in \mathfrak{g}$. This translates to the condition that $\mathrm{tr}\,(\mathrm{ad}\,x \circ \mathrm{ad}\,y) = 0$ for all $\mathrm{ad}\,x \in \mathrm{ad}\,[\mathfrak{g}, \mathfrak{g}]$ and all $\mathrm{ad}\,y \in \mathrm{ad}\,\mathfrak{g}$. By the "if" part of Theorem 6.3.7 or Corollary 6.3.8, we conclude that $\mathrm{ad}\,\mathfrak{g}$ is a solvable subalgebra of $\mathrm{gl}\,(V)$. But $\mathrm{ad}\,\mathfrak{g} = \mathfrak{g}/\mathfrak{c}$, so, since $\mathfrak{c}$ is obviously solvable, Proposition 4.1.10 implies that $\mathfrak{g}$ is solvable. $\square$

The argument above can be summarized as follows:

$\mathfrak{g}$ is solvable $\iff$ $\mathrm{ad}\,\mathfrak{g} \cong \mathfrak{g}/\mathfrak{c}$ is a solvable subalgebra of $\mathrm{gl}\,(\mathfrak{g})$ $\iff$ $\mathrm{tr}\,(\mathrm{ad}\,x \circ \mathrm{ad}\,y) = 0$ for all $\mathrm{ad}\,x \in (\mathrm{ad}\,\mathfrak{g})'$ and all $\mathrm{ad}\,y \in \mathrm{ad}\,\mathfrak{g}$ $\iff$ $\mathrm{tr}\,(\mathrm{ad}\,x \circ \mathrm{ad}\,y) = 0$ for all $x \in \mathfrak{g}'$ and all $y \in \mathfrak{g}$ $\iff$ $B(x, y) = 0$ for all $x \in \mathfrak{g}'$ and all $y \in \mathfrak{g}$.

6.4 Cartan's Criterion for Semisimplicity.

We've already seen that the Killing form B on a Lie algebra $\mathfrak{g}$ over $\mathbb{F}$ is an $\mathbb{F}$-valued symmetric bilinear form. The following lemma gives an invariance property satisfied by B:

Lemma 6.4.1. *The Killing form B satisfies the property that*

$$B([x, y], z) = B(x, [y, z]) \tag{6.7}$$

for all x, y, $z \in \mathfrak{g}$.

Proof. By Lemma 6.3.6, we have

$$\operatorname{tr}([\operatorname{ad} x, \operatorname{ad} y] \circ \operatorname{ad} z) = \operatorname{tr}(\operatorname{ad} x \circ [\operatorname{ad} y, \operatorname{ad} z]),$$

and so

$$\operatorname{tr}(\operatorname{ad}[x, y] \circ \operatorname{ad} z) = \operatorname{tr}(\operatorname{ad} x \circ \operatorname{ad}[y, z]),$$

which implies the result. ☐

The *radical* of B is the subspace of $\mathfrak{g}$ given by $B^\perp := \{x \in \mathfrak{g} \mid B(x, y) = 0 \text{ for all } y \in \mathfrak{g}\}$. Note that by Theorem 1.10.4, B is nondegenerate if and only if $B^\perp = \{0\}$.

Corollary 6.4.2. *$B^\perp$ is an ideal of $\mathfrak{g}$.*

Proof. Let $x \in B^\perp$. Then for all $y \in \mathfrak{g}$, we claim that $[x, y] \in B^\perp$. But for any $z \in \mathfrak{g}$, we have $B([x, y], z) = B(x, [y, z]) = 0$, so we conclude that $[x, y] \in B^\perp$ ☐

Proposition 6.4.3. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$ and $\mathfrak{a}$ an ideal of $\mathfrak{g}$. If $B_\mathfrak{a}$ denotes the Killing form of the Lie algebra $\mathfrak{a}$, then $B_\mathfrak{a}(x, y) = B(x, y)$ for all x, $y \in \mathfrak{a}$. Thus, $B_\mathfrak{a}$ equals the restriction of B to $\mathfrak{a} \times \mathfrak{a}$.*

Proof. Let $\mathfrak{r}$ be any subspace of $\mathfrak{g}$ complementary to $\mathfrak{a}$, so that $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{r}$. Next let B' and B'' be bases of $\mathfrak{a}$ and $\mathfrak{r}$, respectively. If x and y belong to $\mathfrak{a}$, then $\operatorname{ad} x$ and $\operatorname{ad} y$ both map $\mathfrak{g}$ to $\mathfrak{a}$; thus, relative to the basis (B', B'') of $\mathfrak{g}$, the matrices of $\operatorname{ad} x$ and $\operatorname{ad} y$ have block form

$$\operatorname{ad} x = \begin{pmatrix} R_1 & S_1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \operatorname{ad} y = \begin{pmatrix} R_2 & S_2 \\ 0 & 0 \end{pmatrix},$$

respectively. In the above, R_1 is the matrix of the restriction $\operatorname{ad} x|_\mathfrak{a}$ with respect to the basis B' of $\mathfrak{a}$. Likewise, R_2 is the matrix of $\operatorname{ad} y|_\mathfrak{a}$ with respect to B'. Hence

$$\begin{aligned} B(x, y) &= \operatorname{tr} \operatorname{ad} x \circ \operatorname{ad} y \\ &= \operatorname{tr} \begin{pmatrix} R_1 R_2 & R_1 S_2 \\ 0 & 0 \end{pmatrix} \\ &= \operatorname{tr}(R_1 R_2) \\ &= \operatorname{tr} \operatorname{ad} x|_\mathfrak{a} \circ \operatorname{ad} y|_\mathfrak{a} \\ &= B_\mathfrak{a}(x, y). \end{aligned}$$

☐

Example 6.4.4. In Example 3.4.1, we saw that for the basis (e, f, h) of $\mathrm{sl}\,(2, \mathbb{C})$, we could represent $\mathrm{ad}\, e,\ \mathrm{ad}\, f$, and $\mathrm{ad}\, h$ by the following matrices relative to this basis:

$$\mathrm{ad}\, e = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \ \mathrm{ad}\, f = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 0 \end{pmatrix}, \ \mathrm{ad}\, h = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Thus,

$$B(e,e) = \mathrm{tr}\,(\mathrm{ad}\, e \circ \mathrm{ad}\, e) = \mathrm{tr} \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0$$

$$B(e,f) = \mathrm{tr}\,(\mathrm{ad}\, e \circ \mathrm{ad}\, f) = \mathrm{tr} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} = 4$$

$$B(e,h) = \mathrm{tr}\,(\mathrm{ad}\, e \circ \mathrm{ad}\, h) = \mathrm{tr} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -2 & 0 \end{pmatrix} = 0$$

$$B(f,f) = \mathrm{tr}\,(\mathrm{ad}\, f \circ \mathrm{ad}\, f) = \mathrm{tr} \begin{pmatrix} 0 & 0 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0$$

$$B(f,h) = \mathrm{tr}\,(\mathrm{ad}\, f \circ \mathrm{ad}\, f) = \mathrm{tr} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix} = 0$$

$$B(h,h) = \mathrm{tr}\,(\mathrm{ad}\, f \circ \mathrm{ad}\, f) = \mathrm{tr} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 8$$

Thus the matrix of the bilinear form B with respect to $(e,\ f,\ h)$ is

$$\begin{pmatrix} 0 & 4 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

The determinant of this matrix is $-128 \neq 0$. Thus, by Theorem 1.10.4, B is nondegenerate.

Lemma 6.4.5. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$, and let $\mathfrak{a}$ be an ideal of $\mathfrak{g}$. Then the derived series $\mathfrak{a} \supset \mathfrak{a}' \supset \mathfrak{a}^{(2)} \supset \cdots$ consists of ideals of $\mathfrak{g}$.*

Proof. This is an easy induction. Certainly, $\mathfrak{a} = \mathfrak{a}^{(0)}$ is an ideal of $\mathfrak{g}$ by hypothesis. Then, assuming that $\mathfrak{a}^{(r)}$ is an ideal of $\mathfrak{g}$, we have

$$\begin{aligned}
[\mathfrak{a}^{(r+1)}, \mathfrak{g}] &= [[\mathfrak{a}^{(r)}, \mathfrak{a}^{(r)}], \mathfrak{g}] \\
&= [[\mathfrak{a}^{(r)}, \mathfrak{g}], \mathfrak{a}^{(r)}] && \text{(by the Jacobi identity)} \\
&\subset [\mathfrak{a}^{(r)}, \mathfrak{a}^{(r)}] && \text{(by the induction hypothesis)} \\
&= \mathfrak{a}^{(r+1)},
\end{aligned}$$

so $\mathfrak{a}^{(r+1)}$ is an ideal of $\mathfrak{g}$. ☐

Lemma 6.4.6. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$. Then $\mathfrak{g}$ is semisimple if and only if $\mathfrak{g}$ has no abelian ideals $\mathfrak{a} \neq \{0\}$.*

Proof. Suppose that $\mathfrak{g}$ is semisimple. Any abelian ideal $\mathfrak{a}$ of $\mathfrak{g}$ is solvable, so $\mathfrak{a} \subset \mathcal{R}_s = \{0\}$, and thus $\mathfrak{a} = \{0\}$.

Conversely, suppose that $\mathfrak{g}$ is not semisimple. Then the solvable radical $\mathcal{R}_s \neq \{0\}$. Let $\mathcal{R}_s \supsetneq \mathcal{R}_s' \supsetneq \cdots \supsetneq \mathcal{R}_s^{(k)} \supsetneq \{0\}$ be the derived series for $\mathcal{R}_s$. By the preceding lemma, each of the $\mathcal{R}_s^{(i)}$ is an ideal of $\mathfrak{g}$. The last nonzero ideal $\mathcal{R}_s^{(k)}$ is thus a nonzero abelian ideal of $\mathfrak{g}$. Thus, $\mathfrak{g}$ has nonzero abelian ideals. ☐

We are now ready to prove Cartan's criterion for semisimplicity:

Proof of Theorem 6.1.4: To avoid the obvious triviality, we may assume that $\mathfrak{g} \neq \{0\}$. Suppose first that $\mathfrak{g}$ is a Lie algebra over $\mathbb{F}$ such that B is nondegenerate. To prove that $\mathfrak{g}$ is semisimple, it suffices, by Lemma 6.4.6, to prove that $\mathfrak{g}$ has no nonzero abelian ideals. Suppose that $\mathfrak{a}$ is an abelian ideal. then for $x \in \mathfrak{a}$ and y, $z \in \mathfrak{g}$, we have

$$(\operatorname{ad} x \circ \operatorname{ad} y)^2\,(z) = [x, [y, [x, [y, z]]]] \in [\mathfrak{a}, \mathfrak{a}] = \{0\},$$

so $(\operatorname{ad} x \circ \operatorname{ad} y)^2 = 0$. Thus, $\operatorname{ad} x \circ \operatorname{ad} y$ is nilpotent. This implies that $\operatorname{tr}(\operatorname{ad} x \circ \operatorname{ad} y) = 0$. (See equation 1.17.) Hence $B(x, y) = 0$ for all $x \in \mathfrak{a}$ and all $y \in \mathfrak{g}$. Therefore, $\mathfrak{a} \subset B^\perp = \{0\}$, and so $\mathfrak{a} = \{0\}$. Hence any abelian ideal of $\mathfrak{g}$ is $\{0\}$, and so $\mathfrak{g}$ is semisimple.

Conversely, suppose that $\mathfrak{g}$ is semisimple. We'll show in this case that $B^\perp = \{0\}$. Now by definition, $B(x, y) = 0$ for all $x \in B^\perp$ and $y \in \mathfrak{g}$. Hence $B(x, y) = 0$ for all $x \in B^\perp$ and $y \in [B^\perp, B^\perp]$. By Proposition 6.4.3, we see that $B_{B^\perp}(x, y) = 0$ for all $x \in B^\perp$ and $y \in [B^\perp, B^\perp]$. (Here the awkward notation $B_{B^\perp}$ denotes the Killing form on the ideal $B^\perp$ of $\mathfrak{g}$.) Then by Cartan's solvability criterion (Theorem 6.1.3), we see that $B^\perp$ is solvable. Thus $B^\perp \subset \mathcal{R}_s = \{0\}$, and so B is nondegenerate. ☐

Corollary 6.4.7. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{R}$. Then $\mathfrak{g}$ is semisimple $\iff$ $\mathfrak{g}^c$ is semisimple.*

Proof. Let B and B^c denote the Killing forms on $\mathfrak{g}$ and $\mathfrak{g}^c$, respectively. Then it suffices to prove that B is nondegenerate $\iff$ B^c is nondegenerate. Note that $B^c(X,Y) = B(X,Y)$ if X and Y are in $\mathfrak{g}$.

Using this last observation, it is not hard to see that $(B^\perp)^c = (B^c)^\perp$, so $B^\perp = \{0\} \iff (B^c)^\perp = \{0\}$. $\square$

Exercise 6.4.8. Suppose that $\mathfrak{g}$ is a Lie algebra over $\mathbb{C}$. Let $\mathfrak{g}_\mathbb{R}$ be the Lie algebra $\mathfrak{g}$ considered as a real Lie algebra. Prove that $\mathfrak{g}$ is semisimple $\iff$ $\mathfrak{g}_\mathbb{R}$ is semisimple.

We conclude this section by considering a slight variant of the Killing form, called the *trace form*, on Lie algebras of linear operators.

Let V be a vector space over $\mathbb{F}$, and let $\mathfrak{g}$ be a Lie subalgebra of $\mathrm{gl}\,(V)$. The *trace form* on $\mathfrak{g}$ is the symmetric bilinear form $(X,Y) \mapsto \mathrm{tr}\,(XY)$, for all X, $Y \in \mathfrak{g}$.

Proposition 6.4.9. *Let $\mathfrak{g}$ be a Lie subalgebra of $gl(V)$. If $\mathfrak{g}$ is semisimple, then its trace form is nondegenerate.*

Proof. The proof is similar to that of Theorem 6.1.4. Let $\mathcal{I} = \{X \in \mathfrak{g} \mid \mathrm{tr}\,(XY) = 0 \text{ for all } Y \in \mathfrak{g}\}$. Then it follows easily from Lemma 6.3.6 that $\mathcal{I}$ is an ideal of $\mathfrak{g}$.

For any $X \in [\mathcal{I},\mathcal{I}]$ and $Y \in \mathcal{I}$, we have $\mathrm{tr}\,(XY) = 0$. Hence by Theorem 6.3.7 and Corollary 6.3.8, $\mathcal{I}$ is solvable, and so $\mathcal{I} = \{0\}$. Thus the trace form is nondegenerate. $\square$

Chapter 7

Semisimple Lie Algebras: Basic Structure and Representations

7.1 The Basic Structure of a Semisimple Lie Algebra

The rest of the text is essentially going to be devoted to the structure theory of semisimple and simple Lie algebras over $\mathbb{R}$ and $\mathbb{C}$. We start off with an important consequence of Cartan's criterion for semisimplicity.

We say that a Lie algebra $\mathfrak{g}$ is a *direct sum of ideals* if there exist ideals $\mathfrak{a}_1, \ldots, \mathfrak{a}_k$ of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_k$. Note that if $i \neq j$, then $[\mathfrak{a}_i, \mathfrak{a}_j] \subset \mathfrak{a}_i \cap \mathfrak{a}_j = \{0\}$.

Theorem 7.1.1. *Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{F}$. Then $\mathfrak{g}$ is a direct sum of simple ideals*

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k. \tag{7.1}$$

Any simple ideal of $\mathfrak{g}$ is one of the ideals $\mathfrak{g}_i$. Any ideal of $\mathfrak{g}$ is a direct sum of some of the $\mathfrak{g}_i$'s.

Proof. If $\mathfrak{g}$ is already simple, then we're done. So assume that $\mathfrak{g}$ is not simple. Then $\mathfrak{g}$ has ideals $\neq \{0\}$ and $\neq \mathfrak{g}$. Let $\mathfrak{g}_1$ be a nonzero ideal of $\mathfrak{g}$ of *minimal* dimension.

The subspace $\mathfrak{g}_1^\perp = \{x \in \mathfrak{g} \,|\, B(x, y) = 0 \text{ for all } y \in \mathfrak{g}_1\}$ is an ideal of $\mathfrak{g}$. in fact, if $x \in \mathfrak{g}_1^\perp$ and $y \in \mathfrak{g}$, then for any $z \in \mathfrak{g}_1$, we have by Lemma 6.4.1,

$$B([x,y],z) = B(x, \underbrace{[y,z]}_{\text{in } \mathfrak{g}_1}) = 0, \tag{7.2}$$

so $[x,y] \in \mathfrak{g}_1^\perp$.

Next we prove that $[\mathfrak{g}_1, \mathfrak{g}_1^\perp] = \{0\}$. For this, let $u \in \mathfrak{g}_1$ and $v \in \mathfrak{g}_1^\perp$. Then for any $w \in \mathfrak{g}$, we have

$$B([u,v],w) = B(u, \underbrace{[v,w]}_{\text{in } \mathfrak{g}_1^\perp}) = 0.$$

Since B is nondegenerate, we conclude that $[u,v] = 0$.

It follows that

$$[\mathfrak{g}_1 \cap \mathfrak{g}_1^\perp, \mathfrak{g}_1 \cap \mathfrak{g}_1^\perp] \subset [\mathfrak{g}_1, \mathfrak{g}_1^\perp] = \{0\},$$

and hence $\mathfrak{g}_1 \cap \mathfrak{g}_1^\perp$ is an abelian ideal of $\mathfrak{g}$. But since $\mathfrak{g}$ is semisimple, this means that $\mathfrak{g}_1 \cap \mathfrak{g}_1^\perp = \{0\}$.

In addition, since B is nondegenerate, equation (1.34) says that $\dim \mathfrak{g} = \dim \mathfrak{g}_1 + \dim \mathfrak{g}_1^\perp$. Together with our observation that $\mathfrak{g}_1 \cap \mathfrak{g}_1^\perp = \{0\}$, we see that

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_1^\perp. \tag{7.3}$$

Now, by Proposition 6.4.3, the Killing form on the ideal $\mathfrak{g}_1^\perp$ is the restriction of B to $\mathfrak{g}_1^\perp \times \mathfrak{g}_1^\perp$. But B is nondegenerate on $\mathfrak{g}_1^\perp$. In fact, if $x \in \mathfrak{g}_1^\perp$ satisfies $B(x, \mathfrak{g}_1^\perp) = \{0\}$, then we also have $B(x, \mathfrak{g}) = B(x, \mathfrak{g}_1 \oplus \mathfrak{g}_1^\perp) = \{0\}$, so $x = 0$. By Cartan's criterion for semisimplicity, we conclude that $\mathfrak{g}_1^\perp$ is semisimple.

Next we observe that $\mathfrak{g}_1$ is a simple ideal of $\mathfrak{g}$. In fact, by the decomposition (7.3), any ideal of $\mathfrak{g}_1$ is also an ideal of $\mathfrak{g}$. Then, by the minimality of $\dim \mathfrak{g}_1$, such an ideal is either $\{0\}$ or $\mathfrak{g}_1$.

We now apply the procedure above to the semisimple ideal $\mathfrak{g}_1^\perp$ in place of $\mathfrak{g}$ to produce ideals $\mathfrak{g}_2$ and $\mathfrak{g}''$ of $\mathfrak{g}_1^\perp$, with $\mathfrak{g}_2$ simple and $\mathfrak{g}''$ semisimple, such that

$$\mathfrak{g}_1^\perp = \mathfrak{g}_2 \oplus \mathfrak{g}''.$$

Then by (7.3),

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \mathfrak{g}''.$$

The decomposition above shows that both $\mathfrak{g}_2$ and $\mathfrak{g}''$ are ideals of $\mathfrak{g}$. We then apply the same procedure to $\mathfrak{g}''$, etc., to produce the direct sum (10.5) of simple ideals of $\mathfrak{g}$.

Now suppose that $\mathfrak{m}$ is an ideal of $\mathfrak{g}$. Let $I = \{i \mid \mathfrak{g}_i \subset \mathfrak{m}\}$ and let $J = \{1, \ldots, k\} \setminus I$. We claim that $\mathfrak{m} = \bigoplus_{i \in I} \mathfrak{g}_i$. Certainly, $\mathfrak{m} \supset \bigoplus_{i \in I} \mathfrak{g}_i$. Now suppose that $x \in \mathfrak{m} \setminus \bigoplus_{i \in I} \mathfrak{g}_i$. We decompose x according to the direct sum (10.5) to obtain

$$x = x' + x'',$$

where $x' \in \bigoplus_{i\in I} \mathfrak{g}_i$ and $x'' \in \bigoplus_{j\in J} \mathfrak{g}_j$. Since $x \notin \bigoplus_{i\in I} \mathfrak{g}_i$, we have $x'' \neq 0$. But $x' \in \bigoplus_{i\in I} \mathfrak{g}_i \subset \mathfrak{m}$, so we see that $x'' \in \mathfrak{m}$. If $[x'', \mathfrak{g}_j] = \{0\}$ for all $j \in J$, then $[x'', \mathfrak{g}] = \{0\}$, so x'' lies in the center $\mathfrak{c}$ of $\mathfrak{g}$, so $x'' = 0$, contrary to $x'' \neq 0$. Thus $[x'', \mathfrak{g}_j] \neq \{0\}$ for some $j \in J$. For this j, we conclude that $[\mathfrak{m}, \mathfrak{g}_j] \neq \{0\}$, and hence $\mathfrak{m} \cap \mathfrak{g}_j \neq \{0\}$. Since $\mathfrak{g}_j$ is simple, it follows that $\mathfrak{g}_j \subset \mathfrak{m}$, so $j \in I$, a contradiction. This shows that $\mathfrak{m} = \bigoplus_{i\in I} \mathfrak{g}_i$.

Finally, if $\mathfrak{m}$ is a simple ideal of $\mathfrak{g}$, then there is only one summand in $\mathfrak{m} = \bigoplus_{i\in I} \mathfrak{g}_i$, so $\mathfrak{m} = \mathfrak{g}_i$ for some i. □

Corollary 7.1.2. *Let* $\mathfrak{g}$ *be a semisimple Lie algebra over* $\mathbb{F}$. *Then* $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$.

Proof. By Theorem 7.1.1, $\mathfrak{g}$ is a direct sum of simple ideals: $\mathfrak{g} = \bigoplus_{i=1}^k \mathfrak{g}_i$. Since each $\mathfrak{g}_i$ is simple, we have $[\mathfrak{g}_i, \mathfrak{g}_i] = \mathfrak{g}_i$, and since the sum is direct, we have $[\mathfrak{g}_i, \mathfrak{g}_j] \subset \mathfrak{g}_i \cap \mathfrak{g}_j = \{0\}$ for $i \neq j$. Hence

$$\begin{aligned}
[\mathfrak{g}, \mathfrak{g}] &= \left[\bigoplus_{i=1}^k \mathfrak{g}_i, \bigoplus_{j=1}^k \mathfrak{g}_j\right] \\
&= \sum_{i,j} [\mathfrak{g}_i, \mathfrak{g}_j] \\
&= \sum_i \mathfrak{g}_i \\
&= \bigoplus_{i=1}^k \mathfrak{g}_i \\
&= \mathfrak{g}.
\end{aligned}$$

□

Exercise 7.1.3. Prove the converse to Theorem 7.1.1: If $\mathfrak{g}$ is a direct sum of simple ideals, then $\mathfrak{g}$ is semisimple.

Theorem 7.1.1 and Exercise 7.1.3 show that the study of semisimple Lie algebras over $\mathbb{F}$ essentially reduces to the study of *simple* Lie algebras over $\mathbb{F}$.

Corollary 7.1.4. *If* $\mathfrak{g}$ *is a semisimple Lie algebra over* $\mathbb{F}$, *then so are all ideals of* $\mathfrak{g}$ *and all homomorphic images of* $\mathfrak{g}$.

Proof. By Theorem 7.1.1, $\mathfrak{g}$ is a direct sum of simple ideals $\mathfrak{g} = \bigoplus_{i=1}^k \mathfrak{g}_i$, and any ideal of $\mathfrak{g}$ is a direct sum of some of the $\mathfrak{g}_i$. By Exercise 7.1.3, any such ideal must be semisimple.

If $\mathfrak{m}$ is a homomorphic image of $\mathfrak{g}$, then $\mathfrak{m} \cong \mathfrak{g}/\mathfrak{a}$, where $\mathfrak{a}$ is an ideal of $\mathfrak{g}$. Now $\mathfrak{a}$ is a direct sum $\mathfrak{a} = \bigoplus_{i \in I} \mathfrak{g}_i$, for some subset $I \subset \{1, \ldots, k\}$. Put $J = \{1, \ldots, n\} \setminus I$. Then $\mathfrak{m} \cong \bigoplus_{j \in J} \mathfrak{g}_j$, a semisimple Lie algebra by Exercise 7.1.3. $\square$

Theorem 7.1.5. *Let $\mathfrak{g}$ be a semsimple Lie algebra over $\mathbb{F}$. Then $\mathfrak{g}$ is complete; that is, $ad\, \mathfrak{g} = Der\, \mathfrak{g}$.*

Proof. We already know, by Proposition 3.4.3, that $\operatorname{ad} \mathfrak{g}$ is and ideal of $\operatorname{Der} \mathfrak{g}$. Let B_0 denote the Killng form on $\operatorname{Der} \mathfrak{g}$. The restriction of B_0 to $\operatorname{ad} \mathfrak{g} \times \operatorname{ad} \mathfrak{g}$ coincides with the Killing form B on $\operatorname{ad} \mathfrak{g} \cong \mathfrak{g}$, which is nondegenerate.

Now let
$$\mathfrak{m} = (\operatorname{ad} \mathfrak{g})^{\perp} = \{D \in \operatorname{Der} \mathfrak{g} \mid B_0(D, \operatorname{ad} \mathfrak{g}) = \{0\}\}.$$

$\mathfrak{m}$ is an ideal of $\operatorname{Der} \mathfrak{g}$, by exactly the same calculation as (7.2). Moreover, $\mathfrak{m} \cap \operatorname{ad} \mathfrak{g} = \{0\}$, since any $D \in \mathfrak{m} \cap \operatorname{ad} \mathfrak{g}$ must satisfy $B(D, \operatorname{ad} \mathfrak{g}) = \{0\}$, and so, since B is nondegenerate, $D = 0$.

For each $D \in \operatorname{Der} \mathfrak{g}$, let F_D denote the linear functional on $\operatorname{ad} \mathfrak{g}$ given by $F_D(\operatorname{ad} X) = B_0(D, \operatorname{ad} X)$ (for $X \in \mathfrak{g}$). The map $F : D \mapsto F_D$ is a surjective linear map from $\operatorname{Der} \mathfrak{g}$ onto the dual space $(\operatorname{ad} \mathfrak{g})^*$ – surjective, since the image of $\operatorname{ad} \mathfrak{g}$ under this map is $(\operatorname{ad} \mathfrak{g})^*$, by Proposition 1.10.7 and the nondegeneracy of B. The kernel of F is clearly the ideal $\mathfrak{m}$. Hence

$$\dim \mathfrak{m} = \dim \ker F = \dim \operatorname{Der} \mathfrak{g} - \dim (\operatorname{ad} \mathfrak{g})^* = \dim \operatorname{Der} \mathfrak{g} - \dim \operatorname{ad} \mathfrak{g}.$$

It follows that $\operatorname{Der} \mathfrak{g}$ is the direct sum of ideals

$$\operatorname{Der} \mathfrak{g} = \operatorname{ad} \mathfrak{g} \oplus \mathfrak{m}. \tag{7.4}$$

Now let $D \in \mathfrak{m}$. For any $X \in \mathfrak{g}$, equation 3.8 says that $[D, \operatorname{ad} X] = \operatorname{ad}(DX)$. But (7.4) shows that $[\mathfrak{m}, \operatorname{ad} \mathfrak{g}] = \{0\}$, so $\operatorname{ad}(DX) = 0$. Since $\mathfrak{c} = \{0\}$, ad is injective, so $DX = 0$ for all $X \in \mathfrak{g}$, whence $D = 0$. Thus $\mathfrak{m} = \{0\}$, and we finally conclude that $\operatorname{Der} \mathfrak{g} = \operatorname{ad} \mathfrak{g}$. $\square$

We conclude this section with an important theorem, whose proof we shall omit. (See [6], Chapter III, §9.)

Theorem 7.1.6. *(The Levi Decomposition) Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$, and let $\mathcal{R}_s$ be its solvable radical. Then $\mathfrak{g}$ is a direct sum of ideals $\mathfrak{g} = \mathcal{R}_s \oplus \mathcal{I}$, where the ideal $\mathcal{I}$ is semisimple.*

The semisimple ideal $\mathcal{I}$, which is not unique, is called a *Levi factor* of $\mathfrak{g}$. If $\mathcal{I}_1$ is another Levi factor of $\mathfrak{g}$, then there is an automorphism φ of $\mathfrak{g}$ such that $\varphi(\mathcal{I}) = \mathcal{I}_1$.

7.2 Simple Lie Algebras over $\mathbb{R}$

In this section we obtain a general characterization of simple Lie algebras over $\mathbb{R}$. It turns out that there are essentially two types, depending on their complexifications.

Theorem 7.2.1. *Let $\mathfrak{g}$ be a simple Lie algebra over $\mathbb{R}$. Then $\mathfrak{g}$ is exactly one of the following two types:*

1. *A real form of a simple Lie algebra over $\mathbb{C}$*
2. *A simple Lie algebra over $\mathbb{C}$, considered as a real Lie algebra.*

$\mathfrak{g}$ is of the second type if and only if its complexification $\mathfrak{g}^c$ is the direct sum of two simple ideals, both isomorphic (as real Lie algebras) to $\mathfrak{g}$.

Proof. We can assume that $\mathfrak{g} \neq \{0\}$. The Lie algebra $\mathfrak{g}$ is, of course, semisimple because of Exercise 7.1.3. Then by Lemma 6.4.7, the complexification $\mathfrak{g}^c$ of $\mathfrak{g}$ is semisimple. By Theorem 7.1.1, $\mathfrak{g}^c$ is the direct sum of simple ideals

$$\mathfrak{g}^c = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_m. \tag{7.5}$$

Now let σ denote the conjugation of $\mathfrak{g}^c$ with respect to its real form $\mathfrak{g}$. The image $\sigma(\mathfrak{g}_1)$ is closed with respect to multiplication by complex scalars, since if $z \in \mathbb{C}$ and $v \in \mathfrak{g}_1$, then $z\,\sigma(v) = \sigma(\overline{z}\,v) \in \sigma(\mathfrak{g}_1)$. Thus $\sigma(\mathfrak{g}_1)$ is a complex vector subspace of $\mathfrak{g}^c$. It is also an ideal of $\mathfrak{g}^c$ since

$$\begin{aligned}[]
[\sigma(\mathfrak{g}_1), \mathfrak{g}^c] &= [\sigma(\mathfrak{g}^c), \sigma(\mathfrak{g}_1)] \\
&= \sigma\left([\mathfrak{g}^c, \mathfrak{g}_1]\right) \\
&= \sigma(\mathfrak{g}_1).
\end{aligned}$$

Finally, $\sigma(\mathfrak{g}_1)$ is a *simple* ideal of $\mathfrak{g}^c$: if $\mathfrak{a}$ is any ideal of $\sigma(\mathfrak{g}_1)$, then $\sigma(\mathfrak{a})$ is an ideal of $\mathfrak{g}_1$, so $\sigma(\mathfrak{a}) = \mathfrak{g}_1$ or $\sigma(\mathfrak{a}) = \{0\}$. Since σ is bijective, this forces $\mathfrak{a} = \sigma(\mathfrak{g}_1)$ or $\mathfrak{a} = \{0\}$.

Thus $\sigma(\mathfrak{g}_1)$ must be one of the ideals $\mathfrak{g}_1, \ldots, \mathfrak{g}_m$. Suppose first that $\sigma(\mathfrak{g}_1) = \mathfrak{g}_1$. Then $\mathfrak{g}_1$ is σ-invariant. Let $\mathfrak{a} = \{v \in \mathfrak{g}_1 \,|\, \sigma(v) = v\}$. Clearly, $\mathfrak{a} = \mathfrak{g} \cap \mathfrak{g}_1$, so $\mathfrak{a}$ is an ideal of $\mathfrak{g}$. Each $x \in \mathfrak{g}_1$ can be written as

$$x = \frac{x + \sigma(x)}{2} + i\,\frac{i(\sigma(x) - x)}{2}$$

Both $(x + \sigma(x)/2$ and $i(\sigma(x) - x)/2$ belong to $\mathfrak{a}$, which shows that

$$\mathfrak{g}_1 = \mathfrak{a} \oplus i\mathfrak{a}$$

as real vector spaces. We conclude that $\mathfrak{a}$ is a nonzero ideal of $\mathfrak{g}$, whence $\mathfrak{a} = \mathfrak{g}$. Thus $\mathfrak{g}_1 = \mathfrak{g} \oplus i\mathfrak{g} = \mathfrak{g}^c$, and so $\mathfrak{g}$ is a real form of the complex simple Lie algebra $\mathfrak{g}_1$.

Suppose next that $\sigma(\mathfrak{g}_1) = \mathfrak{g}_j$ for some $j \geq 2$. Let $\mathfrak{h} = \mathfrak{g}_1 \oplus \sigma(\mathfrak{g}_1)$. $\mathfrak{h}$ is then a nonzero σ-invariant ideal of $\mathfrak{g}^c$. The same reasoning as that in the preceding paragraph then shows that $\mathfrak{h} = \mathfrak{g}^c$, and so

$$\mathfrak{g}^c = \mathfrak{g}_1 \oplus \sigma(\mathfrak{g}_1).$$

Thus $\mathfrak{g}^c$ is the direct sum of two simple (complex) ideals. The map

$$x \mapsto x + \sigma(x)$$

is then easily shown to be a real Lie algebra isomorphism from $\mathfrak{g}_1$ onto $\mathfrak{g}$. (See the exercise below.) Thus $\mathfrak{g}$ is isomorphic to a complex simple Lie algebra, considered as a real Lie algebra. $\square$

Exercise 7.2.2. In the last part of the proof of Theorem 7.2.1, show that $x \mapsto x + \sigma(x)$ is an real Lie algebra isomorphism of $\mathfrak{g}_1$ onto $\mathfrak{g}$.

The complete classification of complex simple Lie algebras was carried out by Cartan and Killing in the early part of the twentieth century. This also classifies the real simple Lie algebras of type (2) above. The classification of the real forms of complex simple Lie algebras is a much harder problem, and is related to the classification of symmetric spaces. This was also completed by Cartan in the 1930's.

7.3 Basic Representation Theory

In this section, we introduce some of the basic terminology and results of the representation theory of Lie algebras, such as the complete reducibility of $\mathfrak{g}$-modules when $\mathfrak{g}$ is semisimple, Schur's Lemma, and the representation theory of $\mathrm{sl}\,(2, \mathbb{C})$.

Definition 7.3.1. Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$. A vector space V over $\mathbb{F}$ is called a $\mathfrak{g}$-*module* if there is a representation π of $\mathfrak{g}$ on V.

Recall that we also say that $\mathfrak{g}$ *acts on* V.

Definition 7.3.2. Let $\pi_1 : \mathfrak{g} \to \mathrm{gl}\,(V)$ and $\pi_2 : \mathfrak{g} \to \mathrm{gl}\,(W)$ be representations of the Lie algebra $\mathfrak{g}$. A linear map $T : V \to W$ is said to *intertwine* π_1 and π_2 if $\pi_2(x) \circ T = T \circ \pi_1(x)$, for all $x \in \mathfrak{g}$. We also say that T is a $\mathfrak{g}$-*equivariant* linear map from the $\mathfrak{g}$-module V to the $\mathfrak{g}$-module W.

Thus T intertwines the representations π_1 and π_2 if, for all $x \in \mathfrak{g}$, the following diagram commutes:

$$\begin{tikzcd} V \arrow[r, "\pi_1(x)"] \arrow[d, "T"'] & V \arrow[d, "T"] \\ W \arrow[r, "\pi_2(x)"'] & W \end{tikzcd}$$

If T is a linear isomorphism, we call T a $\mathfrak{g}$-*module isomorphism*. In this case, it is clear that T^{-1} is also a $\mathfrak{g}$-module isomorphism from W onto V.

Definition 7.3.3. Let V be a $\mathfrak{g}$-module, via the representation π. A subspace U of V is called a $\mathfrak{g}$-*submodule* if W is invariant under all operators $\pi(x)$, for all $x \in \mathfrak{g}$. Thus the map $\pi_U : \mathfrak{g} \to \mathrm{gl}(U)$ given by $\pi_U(x) = \pi(x)|_U$ is a representation of $\mathfrak{g}$ on U.

If U is a $\mathfrak{g}$-submodule of V, we also say that U is a $\mathfrak{g}$-*invariant subspace* of V. Note that the sum and the intersection of $\mathfrak{g}$-invariant subspaces is a $\mathfrak{g}$-invariant subspace. In addition, if U is a $\mathfrak{g}$-invariant subspace, then the quotient space V/U is a $\mathfrak{g}$-module via the *quotient representation* π' given by

$$\pi'(x)\,(v+U) = \pi(x)\,(v) + U \tag{7.6}$$

for all $x \in \mathfrak{g}$ and all $v \in V$. (The relation $\pi'[x,y] = [\pi'(x), \pi'(y)]$ follows immediately from $\pi[x,y] = [\pi(x), \pi(y)]$.) We call V/U a *quotient module*.

Definition 7.3.4. A representation π of $\mathfrak{g}$ on a vector space V is said to be *irreducible* if V has no $\mathfrak{g}$-submodules other than $\{0\}$ and V. We also say that V is an *irreducible* $\mathfrak{g}$-module.

One more definition:

Definition 7.3.5. A representation π of $\mathfrak{g}$ on a vector space V is said to be *completely reducible* if, for any $\mathfrak{g}$-invariant subspace U of V, there exists a $\mathfrak{g}$-invariant subspace W of V such that $V = U \oplus W$.

Example 7.3.6. A Lie algebra $\mathfrak{g}$ acts on itself via the adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathrm{gl}(\mathfrak{g})$. A subspace $\mathfrak{a} \subset \mathfrak{g}$ is $\mathfrak{g}$-invariant if and only if $\mathfrak{a}$ is an ideal of $\mathfrak{g}$. The adjoint representation is completely reducible if and only if, for any ideal $\mathfrak{a}$ of $\mathfrak{g}$, there is another ideal $\mathfrak{b}$ of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$.

If $\mathfrak{g}$ is semisimple, then ad is completely reducible. In fact, by Theorem 7.1.1, $\mathfrak{g}$ is a direct sum of simple ideals $\mathfrak{g} = \bigoplus_{i=1}^{k} \mathfrak{g}_i$. If $\mathfrak{a}$ is any ideal of $\mathfrak{g}$, then $\mathfrak{a} = \oplus_{i \in I} \mathfrak{g}_i$, for some subset I of $\{1, \ldots, n\}$. Put $J = \{1, \ldots, n\} \setminus I$, and let $\mathfrak{b} = \oplus_{j \in J} \mathfrak{g}_j$. Then $\mathfrak{b}$ is an ideal of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$.

In the above example, there's nothing special about the representation ad. What's important is that $\mathfrak{g}$ is semisimple, as the following theorem shows:

Theorem 7.3.7. *(H. Weyl) Let $\mathfrak{g}$ be a semismple Lie algebra over $\mathbb{F}$. Then any representation π of $\mathfrak{g}$ is completely reducible.*

The proof, which we omit, can be found in several places, such as [5], §6.

Here is an equivalent characterization of complete reducibility.

Theorem 7.3.8. *Let V be a vector space over $\mathbb{F}$ and let π be a representation of a Lie algebra $\mathfrak{g}$ on V. Then π is completely reducible if and only if V is a direct sum of irreducible $\mathfrak{g}$-modules:*

$$V = V_1 \oplus \cdots \oplus V_m \tag{7.7}$$

Proof. Suppose that π is completely reducible. If V is already irreducible, then there is nothing to prove. Otherwise, choose a $\mathfrak{g}$-invariant subspace V_1 of V, of *minimum* positive dimension. Clearly, V_1 is an irreducible $\mathfrak{g}$-module. Since π is completely reducible, V_1 has a complementary $\mathfrak{g}$-invariant subspace W, so

$$V = V_1 \oplus W \tag{7.8}$$

If W is irreducible, then let $V_2 = W$, and we are done. If it isn't, there exists an $\mathfrak{g}$-invariant subspace V_2 of W, of minimum positive dimension. Then V_2 is irreducible. Moreover, V_2 has a complementary $\mathfrak{g}$-invariant subspace W' in V:

$$V = V_2 \oplus W' \tag{7.9}$$

We now claim that

$$W = V_2 \oplus (W' \cap W). \tag{7.10}$$

In fact, by equation (7.9) any $w \in W$ can be written as $w = v_2 + w'$, where $v_2 \in V_2$ and $w' \in W'$. Since $v_2 \in W$, it follows that $w' \in W$, so $w' \in W \cap W'$. Thus $W = V_2 + (W \cap W')$. The sum is direct, since $V_2 \cap (W \cap W') \subset V_2 \cap W' = \{0\}$. This proves (7.10), and so by (7.8),

$$V = V_1 \oplus V_2 \oplus (W \cap W') \tag{7.11}$$

The subspace $W'' = W \cap W'$ is an intersection of $\mathfrak{g}$-invariant subspaces, which is $\mathfrak{g}$-invariant. Thus V is the direct sum of $\mathfrak{g}$-invariant subspaces

$$V = V_1 \oplus V_2 \oplus W'' \tag{7.12}$$

If W'' is irreducible, put $V_3 = W''$ and we're done. If not, let V_3 be a $\mathfrak{g}$-invariant subspace of W'' of minimum positive dimension. Then V_3 is irreducible, and, just as we obtained the decomposition (7.12), we can write V as a direct sum of $\mathfrak{g}$-submodules

$$V = V_1 \oplus V_2 \oplus V_3 \oplus W^{(3)}. \tag{7.13}$$

If we continue this procedure, we will eventually reach the decomposition (7.7) above, since $\dim V$ is finite.

Conversely, suppose that π is a representation of $\mathfrak{g}$ on V, and that V is a direct sum (7.7) of irreducible $\mathfrak{g}$-modules. We want to prove that π is completely reducible. Let U be a $\mathfrak{g}$-invariant subspace of V, with $U \neq \{0\}$ and $U \neq V$.

Since $U \neq V$, there is a subspace V_{i_1} among the irreducible subspaces in (7.7) such that $V_{i_1} \nsubseteq U$. Thus $V_{i_1} \cap U$ is a proper $\mathfrak{g}$-invariant subspace of V_{i_1}; since V_{i_1} is irreducible, we conclude that $V_{i_1} \cap U = \{0\}$. Put $U_2 = U \oplus V_{i_1}$. If $U_2 = V$, then we can take V_{i_1} as our complementary $\mathfrak{g}$-invariant subspace. if $U_2 \neq V$, there is another irreducible subspace V_{i_2} in (7.7) such that $V_{i_2} \nsubseteq U_2$. Then $V_{i_2} \cap U_2 = \{0\}$, so we can let

$$U_3 = U_2 \oplus V_{i_2} = U \oplus V_{i_1} \oplus V_{i_2}.$$

If $U_3 = V$, then we can take our complementary $\mathfrak{g}$-invariant subspace to be $W = V_{i_1} \oplus V_{i_2}$. If $U_3 \neq V$, then there is a subspace V_{i_3} among the irreducible subspaces in (7.7) such that $V_{i_3} \nsubseteq U_3$, and so forth. Since V is finite-dimensional, this procedure ends after a finite number of steps, and we have

$$V = U \oplus V_{i_1} \oplus \cdots \oplus V_{i_r}.$$

The subspace $W = V_{i_1} \oplus \cdots \oplus V_{i_r}$ is then our $\mathfrak{g}$-invariant complementary subspace to U. $\square$

Some authors define complete reducibility by means of the decomposition (7.7). In general, neither this decomposition nor the complementary $\mathfrak{g}$-invariant subspace in the definition of complete reducibility is unique.

Example 7.3.9. Consider the representation π of $\mathrm{gl}\,(2,\mathbb{C})$ on itself via matrix multiplication:

$$\pi(X)\,(Y) = X\,Y$$

It is easy to see that π is indeed a representation, and that the representation space $\mathrm{gl}\,(2,\mathbb{C})$ decomposes into the direct sum of irreducible subspaces:

$$\mathrm{gl}\,(2,\mathbb{C}) = \left\{ \begin{pmatrix} z_1 & 0 \\ z_2 & 0 \end{pmatrix} \,\middle|\, z_1, z_2 \in \mathbb{C} \right\} \oplus \left\{ \begin{pmatrix} 0 & z_3 \\ 0 & z_4 \end{pmatrix} \,\middle|\, z_3, z_4 \in \mathbb{C} \right\}$$

$\mathrm{gl}\,(2,\mathbb{C})$ *also* decomposes into the following invariant irreducible subspaces

$$\mathrm{gl}\,(2,\mathbb{C}) = \left\{ \begin{pmatrix} z_1 & z_1 \\ z_2 & z_2 \end{pmatrix} \,\middle|\, z_1, z_2 \in \mathbb{C} \right\} \oplus \left\{ \begin{pmatrix} z_3 & -z_3 \\ z_4 & -z_4 \end{pmatrix} \,\middle|\, z_3, z_4 \in \mathbb{C} \right\}$$

Theorem 7.3.10. *(Schur's Lemma) Let V be a vector space over $\mathbb{F}$, and let π be an irreducible representation of $\mathfrak{g}$ on V. If $T \in \mathcal{L}(V)$ commutes with $\pi(x)$, for all $x \in \mathfrak{g}$, then either $T = 0$ or T is invertible. If $\mathbb{F} = \mathbb{C}$, then T is just scalar multiplication: $T = \lambda I_V$.*

Proof. First we observe that both the kernel and the range of T are $\mathfrak{g}$-invariant subspaces of V. In fact, if $v \in \ker T$, then $T(\pi(x)\,v) = \pi(x)\,(T(v)) = 0$, so $\pi(x)\,v \in \ker T$ for all $x \in V$. Likewise, $\pi(x)\,(T(V)) = T(\pi(x)(V)) \subset T(V)$.

Since $\ker T$ is $\mathfrak{g}$-invariant and π is irreducible, we have either $\ker T = \{0\}$ or $\ker T = V$. In the former case, T is invertible, and in the latter case, $T = 0$.

Suppose now that $\mathbb{F} = \mathbb{C}$. Then our linear operator T has an eigenvalue λ, so $\ker(T - \lambda I_V) \neq \{0\}$. But the operator $T - \lambda I_V$ commutes with $\pi(x)$, for all $x \in \mathfrak{g}$. Thus, by the preceding paragraph, $T - \lambda I_V = 0$, and so $T = \lambda I_V$. $\square$

Now we turn to a simple but important topic: the representation theory of the three-dimensional simple Lie algebra $\mathrm{sl}\,(2,\mathbb{C})$. As we will see later on, this is the "glue" by which the structure of any complex simple Lie algebra is built upon. Now by Weyl's Theorem (Theorem 7.3.7), any finite-dimensional representation of $\mathrm{sl}\,(2,\mathbb{C})$ is completely reducible, so to understand the (finite-dimensional) representation theory of this Lie algebra, it suffices, by Theorem 7.3.8, to study its irreducible representations.

Theorem 7.3.11. *(The Basic Representation Theorem for $sl\,(2,\mathbb{C})$.) Let $(e,\ f,\ h)$ be the standard basis of $\mathfrak{g} = sl\,(2,\mathbb{C})$, with*

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Let π be an irreducible representation of $\mathfrak{g}$ on a complex vector space V. Then there exists an eigenvector v_0 of $\pi(h)$, with eigenvalue λ, such that $\pi(e)\,v_0 = 0$. For each $j \in \mathbb{Z}^+$, let $v_j = (\pi(f))^j\,v_0$. Then the following properties hold:

1. *The eigenvalue λ is a nonnegative integer n*
2. $v_{n+1} = 0$
3. $(v_0, v_1, \ldots, v_n)$ *is a basis of V*
4. $\pi(f)\,v_j = v_{j+1}$
5. $\pi(h)\,v_j = (n - 2j)\,v_j$, *so each basis vector v_j is an eigenvector of $\pi(h)$*
6. $\pi(e)\,v_j = j(n - j + 1)v_{j-1}$.

Remark: This theorem implies, among other things, that $\pi(h)$ is a semisimple linear operator on V with integer eigenvalues n, $(n-2)$, $\ldots, -(n-2)$, $-n$. On the other hand, both $\pi(e)$ and $\pi(f)$ are nilpotent operators, since their matrices with respect to the basis $(v_0, v_1, \ldots, v_n)$ of V are strictly upper and lower triangular, respectively.

Proof. Since V is a complex vector space, the linear operator $\pi(h)$ has a complex eigenvalue μ. Let v be an eigenvector of $\pi(h)$ corresponding to μ. We claim that $\pi(e)\,v$ belongs to the eigenspace of $\pi(h)$ corresponding to the eigenvalue $\mu+2$. In fact,

$$\begin{aligned}\pi(h)\ (\pi(e)\,v) &= \pi(e)\ (\pi(h)\,v) + (\pi(h)\,\pi(e) - \pi(e)\,\pi(h))\ (v)\\ &= \mu\,\pi(e)\,v + [\pi(h),\pi(e)]\,v\\ &= \mu\,\pi(e)\,v + \pi[h,e]\,(v)\\ &= \mu\,\pi(e)\,v + 2\,\pi(e)\,v\\ &= (\mu+2)\ \pi(e)\,v.\end{aligned}$$

A similar argument then shows that $\pi(e)^2\,(v) = \pi(e)\ (\pi(e)\,v)$ belongs to the eigenspace of $\pi(h)$ corresponding to the eigenvalue $\mu+4$. In general, $\pi(e)^s\,v$ belongs to the eigenspace of $\pi(h)$ corresponding to the eigenvalue $\mu+2s$. Since $\pi(h)$ has only finitely many eigenvalues, we must have $\pi(e)^s\,v=0$ for some $s \in \mathbb{N}$. Let s be the *smallest* positive integer such that $\pi(e)^s\,v \neq 0$ but $\pi(e)^{s+1}\,v = 0$. Put $v_0 = \pi(e)^s\,v$. Then $\pi(e)\,v_0 = 0$, and v_0 is an eigenvector of $\pi(h)$. Let $\lambda\ (= \mu+2s)$ be its eigenvalue.

Now as prescribed in the statement of the theorem, for each $j \in \mathbb{N}$, we define $v_j = \pi(f)^j\,v_0$. This trivially gives conclusion (4). Let us now prove by induction that for each $j \in \mathbb{Z}^+$,

$$\pi(h)\,v_j = (\lambda - 2j)\,v_j. \tag{7.14}$$

If $j=0$, the equation above is just $\pi(h)\,v_0 = \lambda\,v_0$, which is true by the hypothesis on v_0. Assume, then, that equation 7.14 is true for v_j. Then

$$\begin{aligned}\pi(h)\,v_{j+1} &= \pi(h)\ (\pi(f)\,v_j)\\ &= \pi(f)\,(\pi(h)\,v_j) + [\pi(h),\pi(f)]\,(v_j)\\ &= (\lambda-2j)\,\pi(f)(v_j) + \pi[h,f]\,(v_j) \qquad \text{(by induction hypothesis)}\\ &= (\lambda-2j)\,v_{j+1} \,-\, 2\,\pi(f)\,(v_j)\\ &= (\lambda-2j)\,v_{j+1} \,-\, 2\,v_{j+1}\\ &= (\lambda-2(j+1))\,v_{j+1},\end{aligned}$$

which proves that equation (7.14) is true for v_{j+1}.

Thus, if $v_j \neq 0$, it must be an eigenvector of $\pi(h)$ corresponding to the eigenvalue $\lambda - 2j$. Since eigenvectors corresponding to distinct eigenvalues are linearly independent and V is finite-dimensional, we conclude that there must be a $j \in \mathbb{N}$ such that $v_j = 0$. Let n be the smallest nonnegative integer such that $v_n \neq 0$ but $v_{n+1} = 0$.

We will now prove by induction that for all $j \geq 1$,

$$\pi(e)\,v_j = j(\lambda - j + 1)\,v_{j-1}. \tag{7.15}$$

Let us first verify equation (7.15) for $j = 1$. We have

$$\begin{aligned}
\pi(e)\, v_1 &= \pi(e)\ (\pi(f)\, v_0) \\
&= \pi(f)\ (\pi(e)\, v_0)\ +\ \pi[e,f]\, v_0 \\
&= 0 + \pi(h)\, v_0 \\
&= \lambda\, v_0,
\end{aligned}$$

which is precisely equation (7.15) for $j = 1$.

Next assume that equation 7.15 holds for v_j (with $j \geq 1$). Then

$$\begin{aligned}
\pi(e)\, v_{j+1} &= \pi(e)\ (\pi(f)\, v_j) \\
&= \pi(f)\ (\pi(e)\, v_j)\ +\ \pi[e,f]\, v_j \\
&= j(\lambda - j + 1)\, \pi(f)\, v_{j-1}\ +\ \pi(h)\, v_j \qquad \text{(by induction hypothesis)} \\
&= j(\lambda - j + 1)\, v_j\ +\ (\lambda - 2j)\, v_j \\
&= (j+1)(\lambda - j)\, v_j \\
&= (j+1)(\lambda - (j+1) + 1)\, v_j,
\end{aligned}$$

proving (7.15) for v_{j+1}.

If we now apply (7.15) to the vector $v_{n+1} = 0$, we get

$$0 = \pi(e)\, v_{n+1} = (n+1)(\lambda - n)\, v_n.$$

Since $v_n \neq 0$, we conclude that $(n+1)(\lambda - n) = 0$, and so $\lambda = n$. This proves conclusion (1). Plugging in $\lambda = n$ to equation (7.14), we obtain conclusion (5); and plugging this into equation (7.15) gives us conclusion (6).

It remains to prove conclusion (3), that $(v_0, \dots, v_n)$ is a basis of V. These vectors are certainly linearly independent, since by (7.14), they are eigenvectors of $\pi(h)$ corresponding to distinct eigenvalues. From conclusions (4), (5), and (6), we also see that the $\mathbb{C}$-span of $(v_0, \dots, v_n)$ is invariant under $\pi(e)$, $\pi(f)$, and $\pi(h)$. Since (e, f, g) is a basis of $\mathfrak{g}$, we see that this linear span is a $\mathfrak{g}$-invariant subspace. Since V is an irreducible $\mathfrak{g}$-module, we conclude that this span is all of V. This proves conclusion (3) and finishes the proof of Theorem 7.3.11. $\square$

In Theorem 7.3.11, the nonnegative integer n is called the *highest weight* of the representation π. The vector v_0 is called a *highest weight vector* of π; the vectors $v_0, \dots, v_n$ are called *weight vectors*, and their eigenvalues n, $n-2$, $n-4$, $\dots$, $-(n-2)$, $-n$ are called the *weights* of π.

Theorem 7.3.11 says that, up to $\mathfrak{g}$-module isomorphism, any finite-dimensional representation of $\mathrm{sl}\,(2, \mathbb{C})$ is uniquely determined by its highest weight. It also says that the representation space V has a basis $(v_0, \dots, v_n)$ satisfying conditions (4)-(6) in the statement of the theorem.

Exercise 7.3.12. (*Converse of Theorem 7.3.11*) Fix a positive integer n, let V be a vector space over $\mathbb{C}$ with basis $(v_0, v_1, \ldots, v_n)$, and let π_n be the linear map from $\mathrm{sl}\,(2,\mathbb{C})$ to $\mathrm{gl}\,(V)$ defined on the basis $(e,\ f,\ h)$ of $\mathrm{sl}\,(2,\mathbb{C})$ by the relations (4)-(6) in Theorem 7.3.11. Prove that π_n is an irreducible representation of $\mathrm{sl}\,(2,\mathbb{C})$.

Actually, the representation π_n in Theorem 7.3.11 has an explicit realization. Namely, let V be the vector space of *homogeneous degree n polynomials* in two complex variables z and w, with complex coefficients. Thus the elements of V are polynomials of the form

$$p(z,w) = \alpha_n z^n + \alpha_{n-1} z^{n-1} w + \cdots + \alpha_1 z w^{n-1} + \alpha_0 w^n,$$

where $\alpha_n, \cdots, \alpha_0$ are complex numbers. The following $n+1$ degree n monomials

$$z^n,\ z^{n-1}w,\ \cdots, zw^{n-1},\ w^n$$

constitute a basis of V. For each matrix $X = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \in \mathrm{sl}\,(2,\mathbb{C})$, define the linear map $\pi(X)$ on V by

$$\left(\pi \begin{pmatrix} a & b \\ c & -a \end{pmatrix} p\right)(z,w) = (az+cw)\frac{\partial p}{\partial z} + (bz - aw)\frac{\partial p}{\partial w}$$

If $p(z,w)$ is homogeneous of degree n, it is clear that the right hand side above is also homogeneous of degree n. Since $\pi(X)$ is given by a linear differential operator, it is therefore clear that $\pi(X)$ is a linear operator on V.

Exercise 7.3.13. Prove that $X \mapsto \pi(X)$ is a Lie algebra homomorphism of $\mathrm{sl}\,(2,\mathbb{C})$ into $\mathrm{gl}\,(V)$. For this you need to show that $X \mapsto \pi(X)$ is linear (straightforward), and that $\pi[X,Y] = [\pi(X), \pi(Y)]$, for all $X, Y \in \mathrm{sl}\,(2,\mathbb{C})$. Show that this amounts to proving that for all polynomials p in two variables z and w,

$$\left[(a_1 z + c_1 w)\frac{\partial}{\partial z} + (b_1 z - a_1 w)\frac{\partial}{\partial w}, (a_2 z + c_2 w)\frac{\partial}{\partial z} + (b_2 z - a_2 w)\frac{\partial}{\partial w}\right](p)$$
$$= \big((b_1c_2 - c_1b_2)z + 2(c_1a_2 - a_1c_2)w\big)\frac{\partial p}{\partial z} + \big(2(a_1b_2 - b_1a_2)z - (b_1c_2 - c_1b_2)w\big)\frac{\partial p}{\partial w}$$

Exercise 7.3.14. Put $v_0 = z^n$ and for $j = 1, \ldots, n$, put $v_j = P(n,j) z^{n-j} w^j$, where $P(n,j) = n!/(n-j)!$. Then V has basis $(v_0, \ldots, v_n)$. Prove that these basis vectors satisfy the relations (4)-(6) in Theorem 7.3.11.

Chapter 8

Root Space Decompositions

Henceforth we will focus on complex simple and semisimple Lie algebras. These algebras have an incredibly rich structure, all brought to you by Theorem 7.3.11. As for *real* semisimple Lie algebras, it turns out that the theory of such algebras essentially amounts to finding the real forms of complex semisimple Lie algebras. Unfortunately (really!), the study of such algebras is beyond the scope of this course.

8.1 Cartan Subalgebras

We begin with a generalized binomial expansion for derivations.

Exercise 8.1.1. Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$ and let $D \in \operatorname{Der} \mathfrak{g}$. Prove that for any $x,\, y \in \mathfrak{g}$, $\lambda, \mu \in \mathbb{F}$, and $k \in \mathbb{N}$,

$$(D - (\lambda + \mu)\, I_{\mathfrak{g}})^k\, [x, y] = \sum_{r=0}^{k} \binom{k}{r} \left[(D - \lambda\, I_{\mathfrak{g}})^r\, x, (D - \mu\, I_{\mathfrak{g}})^{k-r}\, y\right]$$

Lemma 8.1.2. *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$. Then $\operatorname{Der} \mathfrak{g}$ contains the semisimple and nilpotent parts of all its elements.*

Proof. Let $D \in \operatorname{Der} \mathfrak{g}$, and let $D = S + N$ be its Jordan-Chevalley decomposition. Here S and N are commuting semisimple and nilpotent linear operators on $\mathfrak{g}$, respectively.

Since our field is $\mathbb{C}$, we can, by Theorem 1.7.7, decompose $\mathfrak{g}$ into a direct sum of generalized eigenspaces of D:

$$\mathfrak{g} = \bigoplus_{\lambda} \mathfrak{g}_{\lambda}, \tag{8.1}$$

where the sum is taken over all eigenvalues of D, and

$$\mathfrak{g}_\lambda = \{x \in \mathfrak{g} \,|\, (D - \lambda\, I_{\mathfrak{g}})^k\, x = 0 \text{ for some } k \in \mathbb{Z}^+\}$$

We now claim that $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subset \mathfrak{g}_{\lambda+\mu}$. Here $\mathfrak{g}_{\lambda+\mu}$ is understood to be $\{0\}$ if $\lambda + \mu$ is not an eigenvalue of D.

To prove the claim, we first note that by the remark after Theorem 1.7.1, $\mathfrak{g}_\lambda = \ker(D - \lambda\, I_{\mathfrak{g}})^n$, where $n = \dim \mathfrak{g}$. Suppose that $x \in \mathfrak{g}_\lambda$ and $y \in \mathfrak{g}_\mu$. Then by Exercise 8.1.1,

$$(D - (\lambda + \mu)\, I_{\mathfrak{g}})^{2n}\, [x, y] = \sum_{r=0}^{2n} \binom{2n}{r} [(D - \lambda\, I_{\mathfrak{g}})^r\, x, (D - \mu\, I_{\mathfrak{g}})^{2n-r}\, y].$$

At least one of the factors in each Lie bracket on the right hand side above vanishes. Hence the right hand side vanishes, and $[x, y]$ is annihilated by a power of $D - (\lambda+\mu)\, I_{\mathfrak{g}}$. Thus, $[x, y]$ belongs to the generalized eigenspace $\mathfrak{g}_{\lambda+\mu}$.

Assuming still that $x \in \mathfrak{g}_\lambda$ and $y \in \mathfrak{g}_\mu$, it is now easy to show that the semisimple operator S satisfies Leibniz' rule on $[x, y]$:

$$\begin{aligned} S[x, y] &= (\lambda + \mu)\, [x, y] \qquad (\text{since } [x, y] \in \mathfrak{g}_{\lambda+\mu}) \\ &= [\lambda x, y] + [x, \mu y] \\ &= [Sx, y] + [x, Sy]. \end{aligned}$$

If x and y are arbitrary elements of $\mathfrak{g}$, then we have $x = \sum_\lambda x_\lambda$ and $y = \sum_\lambda y_\lambda$, where $x_\lambda,\ y_\lambda \in \mathfrak{g}_\lambda$. Hence

$$\begin{aligned} S[x, y] &= S\left[\sum_\lambda x_\lambda\,,\ \sum_\mu y_\mu\right] \\ &= S\left(\sum_{\lambda,\mu} [x_\lambda, y_\mu]\right) \\ &= \sum_{\lambda,\mu} ([Sx_\lambda, y_\mu] + [x_\lambda, Sy_\mu]) \\ &= \left[S\left(\sum_\lambda x_\lambda\right), \sum_\mu y_\mu\right] + \left[\sum_\lambda x_\lambda, S\left(\sum_\mu y_\mu\right)\right] \\ &= [Sx, y] + [x, Sy] \end{aligned}$$

Hence Leibniz' rule holds for S on $\mathfrak{g}$, and so $S \in \operatorname{Der} \mathfrak{g}$. It follows that $N = D - S \in \operatorname{Der} \mathfrak{g}$. $\square$

Now if $\mathfrak{g}$ is a semisimple Lie algebra over $\mathbb{C}$, then we know by Theorem 7.1.5 that $\mathfrak{g}$ is complete; that is, $\operatorname{ad} \mathfrak{g} = \operatorname{Der} \mathfrak{g}$. Lemma 8.1.2 then implies that if

$x \in \mathfrak{g}$, the semisimple and nilpotent parts of $\operatorname{ad} x$ belong to $\operatorname{ad} \mathfrak{g}$. Since the map $\operatorname{ad} : \mathfrak{g} \to \operatorname{ad} \mathfrak{g}$ is injective, there must therefore exist *unique* elements x_s and x_n in $\mathfrak{g}$ such that

$$\operatorname{ad} x = \operatorname{ad} x_s + \operatorname{ad} x_n$$

is the Jordan-Chevalley decomposition of $\operatorname{ad} x$. Here $\operatorname{ad} x_s$ is semisimple and $\operatorname{ad} x_n$ is nilpotent. From the above, we obtain $\operatorname{ad} x = \operatorname{ad}(x_s + x_n)$, and so

$$x = x_s + x_n. \tag{8.2}$$

Definition 8.1.3. Equation 8.2 is called the *abstract Jordan-Chevalley decomposition* of $x \in \mathfrak{g}$. x_s is called the *semisimple part* of x, and x_n is called the *nilpotent part* of x.

Definition 8.1.3 begs the question: what if $\mathfrak{g} \subset \mathrm{gl}(V)$ for some complex vector space V? Does the abstract Jordan decomposition of any $X \in \mathfrak{g}$ coincide with its usual one? The answer is yes, as the following exercise shows.

First note that it is an immediate consequence of Lemma 6.3.1 and Lemma 1.9.9 (and its nilpotent counterpart) that if N is a nilpotent element of $\mathrm{gl}(V)$, then so is $\operatorname{ad} N$, and likewise, if S is a semisimple element of $\mathrm{gl}(V)$, then so is $\operatorname{ad} S$.

Exercise 8.1.4. Let V be a complex vector space, and let $\mathfrak{g}$ be a semisimple Lie subalgebra of $\mathrm{gl}(V)$. Let $X \in \mathfrak{g}$, and let $X = S + N$ be its Jordan-Chevalley decomposition, and $X = X_s + X_n$ its abstract Jordan-Chevalley decomposition. Using the following steps, prove that $S = X_s$ and $N = X_n$.

(a). We know that X_s and X_n belong to $\mathfrak{g}$. The problem is that S or N may not belong to $\mathfrak{g}$. Prove that S and N normalize $\mathfrak{g}$; that is, $[S, \mathfrak{g}] \subset \mathfrak{g}$ and $[N, \mathfrak{g}] \subset \mathfrak{g}$.

(b). Prove that $S - X_s$ and $N - X_n$ both belong to the centralizer of $\mathfrak{g}$ in $\mathrm{gl}(V)$.

(c). Prove that S commutes with X_s and N commutes with X_n.

(d). Since $\mathfrak{g}$ is semisimple, Weyl's theorem and Theorem 7.3.8 imply that V decomposes into a direct sum of $\mathfrak{g}$-invariant irreducible subspaces $V = V_1 \oplus \cdots \oplus V_m$. Prove that each subspace V_i is invariant under S and N.

(e). Since X_s, $X_n \in \mathfrak{g}$, each V_i is also X_s- and X_n-invariant. Show that $S - X_s$ and $N - X_n$ are scalar operators on each V_i. That is, prove that for each i, there is a $\lambda_i \in \mathbb{C}$ such that $(S - X_s)|_{V_i} = \lambda_i \, I_{V_i}$, and similarly for $N - X_n$.

(f). Prove that X_s is a semisimple linear operator on each V_i, and hence on V.

(g). Prove that for each $Y \in \mathfrak{g}$, $\operatorname{tr} Y|_{V_i} = 0$. Hence $\operatorname{tr} X|_{V_i} = \operatorname{tr} X_s|_{V_i} = \operatorname{tr} X_n|_{V_i} = 0$.

(h). Prove that $N = X_n$ and that $S = X_s$.

Example 8.1.5. In the case of $\mathfrak{g} = \mathrm{sl}\,(n, \mathbb{C})$, it is a lot easier to prove that the abstract and the regular Jordan-Chevalley decompositions coincide. Let $X \in \mathrm{sl}\,(n, \mathbb{C})$, and let $X = S + N$ be its regular Jordan-Chevalley decomposition. Since N is nilpotent, we have $\mathrm{tr}\, N = 0$. Hence $N \in \mathrm{sl}\,(n, \mathbb{C})$, and so $S = X - N \in \mathrm{sl}\,(n, \mathbb{C})$. Since $\mathrm{ad}\, S$ and $\mathrm{ad}\, N$ are semisimple and nilpotent, respectively, on $\mathrm{gl}\,(n, \mathbb{C})$, they are likewise semisimple and nilpotent, respectively, on $\mathrm{sl}\,(n, \mathbb{C})$. Hence $S = X_s$ and $N = X_n$.

Definition 8.1.6. Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$. A *toral subalgebra* of $\mathfrak{g}$ is a subalgebra consisting entirely of semisimple elements x. (This means $x = x_s$ in Definition 8.1.3.)

If $\mathfrak{g} \neq \{0\}$, does $\mathfrak{g}$ have any nonzero toral subalgebras? Yes! The reason is that there is at least one element x of $\mathfrak{g}$ such that $x_s \neq 0$. Otherwise, $x = x_n$ for all $x \in \mathfrak{g}$, so $\mathrm{ad}\, x$ is nilpotent for all $x \in \mathfrak{g}$, and hence by Engel's theorem (Theorem 5.2.1), $\mathfrak{g}$ is nilpotent. But no nilpotent Lie algebra is semisimple. (Why?) So choose $x \in \mathfrak{g}$ such that $x_s \neq 0$. Then $\mathfrak{a} = \mathbb{C}x_s$ is a one-dimensional toral subalgebra of $\mathfrak{g}$. Note that because $\mathfrak{a}$ is one-dimensional, it is obviously abelian.

Surprisingly, it turns out that all toral subalgebras are abelian.

Proposition 8.1.7. *Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$. Then any toral subalgebra $\mathcal{T}$ of $\mathfrak{g}$ is abelian.*

Proof. Let $x \in \mathcal{T}$. Since $\mathcal{T}$ is a subalgebra, $\mathcal{T}$ is $\mathrm{ad}\, x$-invariant. We want to show that $\mathrm{ad}\, x|_{\mathcal{T}} = 0$, since this will obviously imply that $\mathcal{T}$ is abelian. Now clearly, $\mathrm{ad}\, x = 0$ for $x = 0$, so let us assume that $x \neq 0$.

Now since $\mathrm{ad}\, x$ is semisimple, its restriction $\mathrm{ad}\, x|_{\mathcal{T}}$ is also semisimple, by Lemma 1.9.9. Thus it suffices to prove that any eigenvalue α of $\mathrm{ad}\, x|_{\mathcal{T}}$ is 0. So suppose that α is an eigenvalue of $\mathrm{ad}\, x|_{\mathcal{T}}$ and that $y \in \mathcal{T}$ is an eigenvector corresponding to α. Then $[x, y] = \mathrm{ad}\, x\,(y) = \alpha\, y$. Hence

$$\begin{aligned} \mathrm{ad}\, y\,[y, x] &= -\mathrm{ad}\, y\,[x, y] \\ &= -\mathrm{ad}\, y\,(\alpha\, y) \\ &= -\alpha\,[y, y] \\ &= 0. \end{aligned} \tag{8.3}$$

Now $y \in \mathcal{T}$, so $\mathrm{ad}\, y|_{\mathcal{T}}$ is semisimple. Thus, there is a basis $(e_1, \ldots, e_k)$ of $\mathcal{T}$ consisting of eigenvectors of $\mathrm{ad}\, y|_{\mathcal{T}}$, with respective eigenvalues $\alpha_1, \ldots, \alpha_k$.

We have, of course, $x = \sum_{i=1}^{k} \lambda_i\, e_i$, for some scalars λ_i, not all 0. Then

$\operatorname{ad} y\,(x) = [y, x] = \sum_{i=1}^{k} \lambda_i\,[y, e_i] = \sum_{i=1}^{k} \lambda_i \alpha_i\, e_i$, and so by equation (8.3),

$$\begin{aligned} 0 &= \operatorname{ad} y\,[y, x] \\ &= \operatorname{ad} y \left(\sum_{i=1}^{k} \lambda_i \alpha_i\, e_i \right) \\ &= \sum_{i=1}^{k} \lambda_i \alpha_i\,[y, e_i] \\ &= \sum_{i=1}^{k} \lambda_i \alpha_i^2\, e_i. \end{aligned}$$

It follows that $\lambda_i \alpha_i^2 = 0$ for all i, from which it follows that $\lambda_i \alpha_i = 0$ for all i. Hence $\alpha\, y = \operatorname{ad} x\,(y) = [x, y] = -[y, x] = -\sum_{i=1}^{k} \lambda_i \alpha_i\, e_i = 0$. Since $y \neq 0$, we conclude that $\alpha = 0$, proving the proposition. □

Definition 8.1.8. Let $\mathfrak{g}$ be a complex semisimple Lie algebra. A *Cartan subalgebra* of $\mathfrak{g}$ is a *maximal* toral subalgebra of $\mathfrak{g}$; i,e., a toral subalgebra which is not properly contained in any other toral subalgebra of $\mathfrak{g}$.

Any toral subalgebra of maximal dimension is obviously a Cartan subalgebra of $\mathfrak{g}$. It is an important result in Lie theory that any two Cartan subalgebras are conjugate under an automorphism of $\mathfrak{g}$.

Theorem 8.1.9. *Let $\mathfrak{h}_1$ and $\mathfrak{h}_2$ be Cartan subalgebras of a complex semisimple Lie algebra $\mathfrak{g}$. Then there exists an automorphism φ of $\mathfrak{g}$ such that $\varphi(\mathfrak{h}_1) = \mathfrak{h}_2$.*

In particular, any two Cartan subalgebras of $\mathfrak{g}$ have the same dimension, which is called the *rank* of $\mathfrak{g}$.

Theorem 8.1.9 says that the behavior of any two Cartan subalgebras of $\mathfrak{g}$ with respect to the adjoint representation is exactly the same. Since the proof of Theorem 8.1.9 requires some Lie group theory, we will omit it.

8.2 Root Space Decomposition

In this section, we examine the root space structure of a complex semisimple Lie algebra $\mathfrak{g}$. Knowledge of this structure is absolutely vital for any further study of semisimple Lie theory.

Lemma 8.2.1. *Let $\mathfrak{h}$ be a toral subalgebra of a complex semisimple Lie algebra $\mathfrak{g}$. Then there is a basis of $\mathfrak{g}$ relative to which the linear operators $\operatorname{ad} h$ all have diagonal matrices, for every $h \in \mathfrak{h}$.*

Proof. Let $h_1, \dots, h_l$ be a basis of $\mathfrak{h}$. then $\operatorname{ad} h_1, \dots, \operatorname{ad} h_k$ are semisimple linear operators on $\mathfrak{g}$. Moreover, these operators commute, since $[\operatorname{ad} h_i, \operatorname{ad} h_j] = \operatorname{ad}[h_i, h_j] = 0$. Hence by Exercise 1.9.12, there is a basis of $\mathfrak{g}$ relative to which each $\operatorname{ad} h_i$ has a diagonal matrix. If $h \in \mathfrak{h}$, then $\operatorname{ad} h$ is a linear combination of the $\operatorname{ad} h_i$, so the matrix of $\operatorname{ad} h$ relative to this basis is also diagonal. □

The basis obtained in Lemma 8.2.1 thus consists of joint eigenvectors of all the elements of $\operatorname{ad} \mathfrak{h}$.

Let us now assume that $\mathfrak{h}$ is a Cartan subalgebra of a complex semisimple Lie algebra $\mathfrak{g}$. Choose a basis of $\mathfrak{g}$ consisting of joint eigenvectors of $\operatorname{ad} h$. If v is an element of this basis, then for every $h \in \mathfrak{h}$, we have

$$\operatorname{ad} h\,(v) = \alpha(h)\, v, \tag{8.4}$$

where the complex coefficient $\alpha(h)$ clearly depends on $\mathfrak{h}$. The mapping $h \mapsto \alpha(h)$ is easily seen to be a linear functional on $\mathfrak{h}$, and it is also clear that v belongs to the *joint eigenspace* corresponding to $\alpha \in \mathfrak{h}^*$.

For each $\alpha \in \mathfrak{h}^*$, let $\mathfrak{g}_\alpha$ denote the joint eigenspace corresponding to α:

$$\mathfrak{g}_\alpha = \{v \in \mathfrak{g} \mid [h, v] = \alpha(h)\, v \text{ for all } h \in \mathfrak{h}\} \tag{8.5}$$

Of course, for a given α, $\mathfrak{g}_\alpha$ could very well be $\{0\}$. Nonetheless, each vector in our basis belongs to a unique joint eigenspace, and we obtain the direct sum decomposition

$$\mathfrak{g} = \bigoplus_\alpha \mathfrak{g}_\alpha \tag{8.6}$$

where the sum ranges over all $\alpha \in \mathfrak{h}^*$ such that $\mathfrak{g}_\alpha \neq \{0\}$.

The decomposition (8.6) is very important, and we will study its structure in some detail.

Observe that if 0 denotes the zero linear functional on $\mathfrak{h}$, then the joint eigenspace $\mathfrak{g}_0$ is the centralizer $\mathfrak{c}(\mathfrak{h})$ of $\mathfrak{h}$ in $\mathfrak{g}$. Since $\mathfrak{h}$ is abelian, we have $\mathfrak{h} \subset \mathfrak{g}_0$, so in particular $\mathfrak{g}_0 \neq \{0\}$.

Definition 8.2.2. Let $\mathfrak{h}$ be a Cartan subalgebra of a complex semisimple Lie algebra $\mathfrak{g}$. Any *nonzero* linear functional $\alpha \in \mathfrak{h}^*$ such that $\mathfrak{g}_\alpha \neq \{0\}$ is called a *root of* $\mathfrak{g}$ *relative to* $\mathfrak{h}$. If α is a root, the joint eigenspace $\mathfrak{g}_\alpha$ is called the root space corresponding to α. We denote the set of all roots by Δ.

Note that Δ is a *nonempty* finite set. It's nonempty since, otherwise $\mathfrak{g}_0 = \mathfrak{g}$, so $\{0\} \subsetneq \mathfrak{h} \subset \mathfrak{c}$, the center of $\mathfrak{g}$. But $\mathfrak{g}$ has center $\{0\}$.

Since $\mathfrak{g}$ is finite-dimensional and since each root space $\mathfrak{g}_\alpha$ has dimension ≥ 1, we see that the set Δ of roots is finite.

Definition 8.2.3. If $\mathfrak{h}$ is a Cartan subalgebra of a complex semisimple Lie algebra $\mathfrak{g}$, then we can rewrite the decomposition (8.6) as

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha. \tag{8.7}$$

We call (8.7) the *root space decomposition of* $\mathfrak{g}$ *relative to* $\mathfrak{h}$.

For the rest of this section, we assume that $\mathfrak{h}$ is a fixed Cartan subalgebra of a complex semisimple Lie algebra $\mathfrak{g}$.

Theorem 8.2.4. *The joint eigenspaces $\mathfrak{g}_\alpha$ satisfy the following peoperties:*

(a) For α, $\beta \in \mathfrak{h}^*$, $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$.

(b) If $\alpha \neq 0$ and $x \in \mathfrak{g}_\alpha$, then $\operatorname{ad} x$ is nilpotent.

(c) If α, $\beta \in \mathfrak{h}^*$ such that $\alpha + \beta \neq 0$, then $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = \{0\}$.

Proof. (a) Let $h \in \mathfrak{h}$, and let $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_\beta$. Then, since $\operatorname{ad} h$ is a derivation, we have

$$\begin{aligned} \operatorname{ad} h\,[x,y] &= [\operatorname{ad} h\,(x), y] + [x, \operatorname{ad} x\,(y)] \\ &= [\alpha(h)\,x, y] + [x, \beta(h)\,y] \\ &= (\alpha(h) + \beta(h))\,[x,y] \\ &= (\alpha+\beta)(h)\,[x,y]. \end{aligned}$$

This shows that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$.

(b) We can assume that x is a nonzero element of $\mathfrak{g}_\alpha$. Let $\beta \in \mathfrak{h}^*$ such that $\mathfrak{g}_\beta \neq \{0\}$. Then by Part (a), $\operatorname{ad} x\,(\mathfrak{g}_\beta) \subset \mathfrak{g}_{\beta+\alpha}$. Applying Part (a) again, we have $(\operatorname{ad} x)^2\,(\mathfrak{g}_\beta) \subset \mathfrak{g}_{\beta+2\alpha}$. In general, for all $k \in \mathbb{Z}^+$,

$$(\operatorname{ad} x)^k\,(\mathfrak{g}_\beta) \subset \mathfrak{g}_{\beta+k\alpha}$$

Now the set of $\beta \in \mathfrak{h}^*$ such that $\mathfrak{g}_\beta \neq \{0\}$ is finite, so, since $\alpha \neq 0$, there is a nonnegative integer $k = k_\beta$ such that $\mathfrak{g}_{\beta+k_\beta\alpha} = \{0\}$. Then of course $\mathfrak{g}_{\beta+k_\beta\,\alpha} = \{0\}$, and so

$$(\operatorname{ad} x)^{k_\beta}\,(\mathfrak{g}_\beta) = \{0\}.$$

Let $N = \max\{k_\beta\}$, where the maximum is taken over all β such that $\mathfrak{g}_\beta \neq \{0\}$. Then we have $(\operatorname{ad} x)^N = 0$, and hence $\operatorname{ad} x$ is nilpotent.

(c) Let $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$. For any $h \in \mathfrak{h}$, we have by Lemma 6.4.1,

$$\begin{aligned} 0 &= B([h,x],y) + B(x,[h,y]) \\ &= B(\alpha(h)\,x,y) + B(x,\beta(h)\,y) \\ &= (\alpha(h)+\beta(h))\,B(x,y). \end{aligned} \tag{8.8}$$

Since $\alpha+\beta \neq 0$, we may choose an element $h_0 \in \mathfrak{h}$ such that $\alpha(h_0)+\beta(h_0) \neq 0$. If we plug in $h = h_0$ in (8.8), we obtain $B(x, y) = 0$. Since x and y are arbitrary in $\mathfrak{g}_\alpha$ and $\mathfrak{g}_\beta$, it follows that $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = \{0\}$.

□

Theorem 8.2.4 implies, in particular, that $[\mathfrak{g}_0, \mathfrak{g}_0] \subset \mathfrak{g}_0$, so that $\mathfrak{g}_0$ is a subalgebra of $\mathfrak{g}$.

Corollary 8.2.5. *The Killing form B is nondegenerate on $\mathfrak{g}_0$.*

Proof. The assertion is that $B|_{\mathfrak{g}_0 \times \mathfrak{g}_0}$ is nondegenerate. If $x \in \mathfrak{g}_0$, then by Theorem 8.2.4, we have $B(x, \mathfrak{g}_\alpha) = \{0\}$ for all $\alpha \neq 0$ in $\mathfrak{h}^*$. Suppose that $x \neq 0$. Then since B is nondegenerate, $B(x, \mathfrak{g}) \neq \{0\}$. Hence

$$\begin{aligned}\{0\} &\neq B(x, \mathfrak{g}) \\ &= B\big(x, \mathfrak{g}_0 \oplus \bigoplus_{\alpha\in\Delta} \mathfrak{g}_\alpha\big) \\ &= B(x, \mathfrak{g}_0) + \sum_{\alpha\in\Delta} B(x, \mathfrak{g}_\alpha) \\ &= B(x, \mathfrak{g}_0).\end{aligned}$$

This shows that B is nondegenerate on $\mathfrak{g}_0$. □

Lemma 8.2.6. *The Killing form B is nondegenerate on $\mathfrak{h}$.*

Proof. Suppose that $h \in \mathfrak{h}$ satisfies $B(h, \mathfrak{h}) = \{0\}$. We will prove that $B(h, \mathfrak{g}_0) = \{0\}$, so that by Lemma 8.2.5, it will follow that $h = 0$.

Let x be any element of $\mathfrak{g}_0 = \mathfrak{c}(\mathfrak{h})$, and let $x = x_s + x_n$ be its abstract Jordan-Chevalley decomposition. Since $\operatorname{ad} x_s$ and $\operatorname{ad} x_n$ are the semisimple and nilpotent parts of $\operatorname{ad} x$, they are polynomials in $\operatorname{ad} x$ with zero constant term. (See Theorem 1.9.14.) Since $\operatorname{ad} x(\mathfrak{h}) = \{0\}$, we conclude that $\operatorname{ad} x_s(\mathfrak{h}) = \{0\}$ and $\operatorname{ad} x_n(\mathfrak{h}) = \{0\}$ as well. Thus $x_s, x_n \in \mathfrak{c}(\mathfrak{h}) = \mathfrak{g}_0$. This means that $[x_s, \mathfrak{h}] = [x_n, \mathfrak{h}] = \{0\}$.

Next we claim that $x_s \in \mathfrak{h}$. Since x_s is a semisimple element of $\mathfrak{g}$ commuting with $\mathfrak{h}$, the subspace $\mathbb{C}x_s + \mathfrak{h}$ is a toral subalgebra of $\mathfrak{g}$ containg $\mathfrak{h}$. Since $\mathfrak{h}$ is *maximal* toral, this implies that $\mathbb{C}x_s + \mathfrak{h} = \mathfrak{h}$, and therefore $x_s \in \mathfrak{h}$. In particular, by the hypothesis on h, we have $B(x_s, h) = 0$.

Now x_n commutes with h, so $\operatorname{ad} x_n$ commutes with $\operatorname{ad} h$. Since $\operatorname{ad} x_n$ is nilpotent, this must also be true of $\operatorname{ad} x_n \circ \operatorname{ad} h$: $(\operatorname{ad} x_n)^N = 0 \implies (\operatorname{ad} x_n \circ \operatorname{ad} h)^N = (\operatorname{ad} x_n)^N \circ (\operatorname{ad} h)^N = 0$. Hence $B(x_n, h) = \operatorname{tr}\,(\operatorname{ad} x_n \circ \operatorname{ad} h) = 0$.

Together with $B(x_s, h) = 0$, we conclude that $B(x, h) = B(x_n, h) + B(x_s, h) = 0$. Since x is arbitrary in $\mathfrak{g}_0$, we get $B(\mathfrak{g}_0, h) = \{0\}$, from which we obtain $h = 0$. □

Lemma 8.2.7. *Let $\mathfrak{h}$ be a Cartan subalgebra of a complex semisimple Lie algebra $\mathfrak{g}$, and let $\mathfrak{g}_0 = \mathfrak{c}(\mathfrak{h})$. Then $\mathfrak{g}_0 = \mathfrak{h}$!*

Proof. The proof is carried out by proving the following successive assertions:

Step 1: $\mathfrak{g}_0$ is nilpotent. By Engel's theorem, it suffices to prove that for each $x \in \mathfrak{g}_0$, $(\operatorname{ad} x)|_{\mathfrak{g}_0}$ is nilpotent. If $x = x_s + x_n$ is the abstract Jordan decomposition of x, then the proof of Lemma 8.2.6 shows that $x_s \in \mathfrak{h}$ and $x_n \in \mathfrak{g}_0$. Hence $[x_s, \mathfrak{g}_0] = \{0\}$, and thus $(\operatorname{ad} x_s)|_{\mathfrak{g}_0} = 0$. On the other hand $\operatorname{ad} x_n$ is nilpotent on $\mathfrak{g}$, and so its restriction to $\mathfrak{g}_0$ is nilpotent. Therefore, $(\operatorname{ad} x)|_{\mathfrak{g}_0} = (\operatorname{ad} x_n)|_{\mathfrak{g}_0}$ is nilpotent.

Step 2: $\mathfrak{h} \cap [\mathfrak{g}_0, \mathfrak{g}_0] = \{0\}$. Note that $B(\mathfrak{h}, [\mathfrak{g}_0, \mathfrak{g}_0]) = B([\mathfrak{h}, \mathfrak{g}_0], \mathfrak{g}_0) = B(\{0\}, \mathfrak{g}_0) = \{0\}$. Thus if $h \in \mathfrak{h} \cap [\mathfrak{g}_0, \mathfrak{g}_0]$, it follows that $B(\mathfrak{h}, h) \subset B(\mathfrak{h}, [\mathfrak{g}_0, \mathfrak{g}_0]) = \{0\}$. Since B is nondegenerate on $\mathfrak{h}$, this forces $h = 0$.

Step 3: $\mathfrak{g}_0$ is abelian. Assume that $[\mathfrak{g}_0, \mathfrak{g}_0] \neq \{0\}$. From the descending central series for the nilpotent Lie algebra $\mathfrak{g}_0$, it is clear that if $\mathfrak{c}_0$ denotes the center of $\mathfrak{g}_0$, then $\mathfrak{c}_0 \cap [\mathfrak{g}_0, \mathfrak{g}_0] \neq \{0\}$. Let c be a nonzero element of $\mathfrak{c}_0 \cap [\mathfrak{g}_0, \mathfrak{g}_0]$, and let $c = c_s + c_n$ be its abstract Jordan-Chevalley decomposition. If $c_n = 0$, then from the proof of Lemma 8.2.6, we get $0 \neq c = c_s \in \mathfrak{h} \cap [\mathfrak{g}_0, \mathfrak{g}_0] = \{0\}$, a contradiction. Hence $c_n \neq 0$.

As with the proof of Lemma 8.2.6, we see that $c_n = c - c_s \in \mathfrak{g}_0$. Since $c \in \mathfrak{c}_0$ and $c_s \in \mathfrak{h} \subset \mathfrak{c}_0$, we see that $c_n \in \mathfrak{c}_0$. Thus, for all $x \in \mathfrak{g}_0$, $[c_n, x] = 0$ and therefore $[\operatorname{ad} c_n, \operatorname{ad} x] = 0$. Now $\operatorname{ad} c_n$ is a nilpotent linear operator, and we conclude that $\operatorname{ad} c_n \circ \operatorname{ad} x$ is also nilpotent for all $x \in \mathfrak{g}_0$. Hence, for all such x, $B(c_n, x) = \operatorname{tr}(\operatorname{ad} c_n \circ \operatorname{ad} x) = 0$. It follows that $B(c_n, \mathfrak{g}_0) = \{0\}$, contradicting the nondegeneracy of B on $\mathfrak{g}_0$.

This contradiction leads us to conclude that $[\mathfrak{g}_0, \mathfrak{g}_0] = \{0\}$.

Step 4: $\mathfrak{g}_0 = \mathfrak{h}$. Suppose that $x \in \mathfrak{g}_0 \setminus \mathfrak{h}$. If $x = x_s + x_n$ is its abstract Jordan-Chevalley decomposition, then we must have $x_n \neq 0$. Now $x_n \in \mathfrak{g}_0$, so, since $\mathfrak{g}_0$ is abelian, we conclude that $\operatorname{ad} x_n \circ \operatorname{ad} y$ is a nilpotent linear operator for all $y \in \mathfrak{g}_0$. Hence $B(x_n, y) = 0$ for all $y \in \mathfrak{g}_0$. Since B is nondegenerate on $\mathfrak{g}_0$, this forces $x_n = 0$, a contradiction. Thus $\mathfrak{g}_0 = \mathfrak{h}$, completing the proof of Lemma 8.2.7. □

Theorem 8.2.8. *(Root Space Decomposition) Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$ and let $\mathfrak{h}$ be a Cartan subalgebra. Then $\mathfrak{g}$ is a direct sum of $\mathfrak{h}$ and the root spaces of $\mathfrak{h}$:*

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha \tag{8.9}$$

Theorem 8.2.8 is an immediate consequence of equation (8.7) and Lemma 8.2.7.

Theorem 8.2.9. *Let $\mathfrak{h}$ be a Cartan subalgebra of a complex semisimple Lie algebra $\mathfrak{g}$. Then for each $\varphi \in \mathfrak{h}^*$, there is a unique element h_φ such that $\varphi(h) = B(h_\varphi, h)$ for all $h \in \mathfrak{h}$.*

Proof. This follows immediately from Proposition 1.10.7. ☐

Note that according to Proposition 1.10.7, we also have $h_{\alpha\varphi+\beta\psi} = \alpha h_\varphi + \beta h_\psi$, for all φ, $\psi \in \mathfrak{h}^*$ and all $\alpha, \beta \in \mathbb{C}$.

Definition 8.2.10. We transfer the Killing form B to the dual space $\mathfrak{h}^*$ by $B(\phi, \psi) = B(h_\phi, h_\psi)$ for all ϕ, $\psi \in \mathfrak{h}^*$.

Using the convention in Definition 8.2.10, we have $\varphi(h_\psi) = B(h_\varphi, h_\psi) = \psi(h_\varphi)$. By Lemma 8.2.6, we also see that B is nondegenerate on $\mathfrak{h}^*$.

Let us now investigate the root spaces $\mathfrak{g}_\alpha$.

Theorem 8.2.11. *(Theorem on Roots) Let $\mathfrak{h}$ be a Cartan subalgebra of a complex semisimple Lie algebra $\mathfrak{g}$, and let Δ be the set of roots of $\mathfrak{g}$ relative to $\mathfrak{h}$. Then:*

(a) There are $\dim \mathfrak{h}$ *linearly independent roots which thus form a basis of $\mathfrak{h}^*$.*

(b) If α is a root, then so is $-\alpha$.

(c) If α is a root and $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_{-\alpha}$, then $[x, y] = B(x, y)\, h_\alpha$.

(d) If α is a root, then $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathbb{C}\, h_\alpha$.

(e) For each $\alpha \in \Delta$, $\alpha(h_\alpha) = B(\alpha, \alpha) \neq 0$.

(f) Let α be a root, and let $h^*_\alpha = 2h_\alpha/\alpha(h_\alpha)$. If e_α is a nonzero element of $\mathfrak{g}_\alpha$, then there is an $f_\alpha \in \mathfrak{g}_{-\alpha}$ such that $(e_\alpha,\, f_\alpha,\, h^*_\alpha)$ is the basis of a three-dimensional simple Lie subalgebra of $\mathfrak{g}$ isomorphic to $sl(2, \mathbb{C})$.

Proof. (a) The claim here is that Δ spans the dual space $\mathfrak{h}^*$. Suppose, to the contrary, that Δ does *not* span $\mathfrak{h}^*$. Then there exists a nonzero vector $h \in \mathfrak{h}$ such that $\alpha(h) = 0$ for all $\alpha \in \Delta$. For each $x \in \mathfrak{g}_\alpha$, we therefore obtain $[h, x] = \alpha(h)\, x = 0$. Hence $[h, \mathfrak{g}_\alpha] = \{0\}$ for all $\alpha \in \Delta$. Since $\mathfrak{h}$ is abelian, we also have $[h, \mathfrak{h}] = \{0\}$. Hence by Theorem 8.2.8, $[h, \mathfrak{g}] = \{0\}$, and therefore h lies in the center $\mathfrak{c}$ of $\mathfrak{g}$. But since $\mathfrak{g}$ is seimisimple, $\mathfrak{c} = \{0\}$, so $h = 0$, a contradiction.

(b) Suppose that α is a root, but that $-\alpha$ isn't. Let x be a nonzero element of $\mathfrak{g}_\alpha$. Then by Theorem 8.2.4 (c), $B(x, \mathfrak{g}_\beta) = \{0\}$ for all $\beta \in \Delta$ (including $\beta = \alpha$!). For the same reason, we have $B(x, \mathfrak{g}_0) = B(x, \mathfrak{h}) = \{0\}$. Thus, by Theorem 8.2.8, we obtain $B(x, \mathfrak{g}) = \{0\}$. This contradicts the fact that B is nondegenerate on $\mathfrak{g}$.

(c) Suppose that α is a root and that $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_{-\alpha}$. Then according to Theorem 8.2.4 (a), $[x, y] \in \mathfrak{g}_0 = \mathfrak{h}$. We can determine the vector $[x, y]$ in $\mathfrak{h}$ as follows: For any $h \in \mathfrak{h}$, we have

$$\begin{aligned} B(h, [x, y]) &= B([h, x], y) \\ &= B(\alpha(h)\, x, y) \\ &= \alpha(h)\, B(x, y) \\ &= B(x, y)\, B(h, h_\alpha) \\ &= B(h, B(x, y)\, h_\alpha). \end{aligned}$$

Since B is nondegenerate on $\mathfrak{h}$ (Theorem 8.2.6), we see that $[x, y] = B(x, y)\, h_\alpha$.

(d) By part (c), it suffices to show that if $0 \neq x \in \mathfrak{g}_\alpha$, then there exists a $y \in \mathfrak{g}_{-\alpha}$ such that $B(x, y) \neq 0$. Suppose, to the contrary, that $B(x, \mathfrak{g}_{-\alpha}) = \{0\}$. Now by Theorem 8.2.4 (c), $B(x, \mathfrak{g}_\beta) = \{0\}$ for all $\beta \in \Delta$ such that $\beta \neq -\alpha$, and likewise $B(x, \mathfrak{h}) = \{0\}$. Thus $B(x, \mathfrak{g}) = \{0\}$, contradicting the nondegeneracy of B on $\mathfrak{g}$.

(e) Let $\alpha \in \Delta$, and suppose that $\alpha(h_\alpha) = 0$. By part (d), it is possible to choose vectors $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_{-\alpha}$ such that $B(x, y) \neq 0$. Let $\mathfrak{s} = \mathbb{C}\,x + \mathbb{C}\,y + \mathbb{C}\,h_\alpha$. Then, since $[x, y] = c\,h_\alpha$ (with $c = B(x, y) \neq 0$) and $[h_\alpha, x] = \alpha(h_\alpha)\,x = 0$, $[h_\alpha, y] = \alpha(h_\alpha)\,y = 0$, we see that $\mathfrak{s}$ is a solvable Lie subalgebra of $\mathfrak{g}$ with derived algebra $\mathfrak{s}' = \mathbb{C}h_\alpha$. The algebra $\operatorname{ad}\mathfrak{s}$ is a solvable Lie algebra of linear operators on $\mathfrak{g}$, so by Lie's Theorem (Theorem 4.2.3), $\mathfrak{g}$ has a basis relative to which every element of $\operatorname{ad}\mathfrak{s}$ has an upper triangular matrix. With respect to this basis, the elements of $\operatorname{ad}\mathfrak{s}'$ have strictly upper triangular matrices. In particular, this means that $\operatorname{ad} h_\alpha$ has a strictly upper triangular matrix, and so $\operatorname{ad} h_\alpha$ is nilpotent. But $\operatorname{ad} h_\alpha$ is semisimple, so $\operatorname{ad} h_\alpha = 0$, whence $h_\alpha = 0$, and therefore $\alpha = 0$, contradicting the fact that the elements of Δ are nonzero.

(f) Let $0 \neq e_\alpha \in \mathfrak{g}_\alpha$. Then by part (d) above, there exists an element $f'_\alpha \in \mathfrak{g}_{-\alpha}$ such that $B(e_\alpha, f'_\alpha) = c \neq 0$. Let $f_\alpha = 2f'_\alpha/(c\,\alpha(h_\alpha))$. Then $B(e_\alpha, f_\alpha) = 2/\alpha(h_\alpha)$. Hence, by part (c),

$$[e_\alpha, f_\alpha] = \frac{2\,h_\alpha}{\alpha(h_\alpha)} = h^*_\alpha. \tag{8.10}$$

Note that $\alpha(h^*_\alpha) = 2$. From equation (8.10), we obtain the commutation relations

$$[h^*_\alpha, e_\alpha] = \alpha(h^*_\alpha)\, e_\alpha = 2\, e_\alpha \tag{8.11}$$
$$[h^*_\alpha, f_\alpha] = -\alpha(h^*_\alpha)\, f_\alpha = -2\, f_\alpha, \tag{8.12}$$

which show that the span of $(e_\alpha, f_\alpha, h^*_\alpha)$ is a subalgebra of $\mathfrak{g}$ isomorphic to $\mathrm{sl}\,(2, \mathbb{C})$, the explicit isomorphism mapping e_α to e, f_α to f and h^*_α to h.

$\square$

An ordered triple (e', f', h') of elements of $\mathfrak{g}$ satisfying $[h', e'] = 2e'$, $[h', f'] = -2f'$, $[e', f'] = h'$ is called an *sl$_2$-triple*. Clearly, any sl$_2$-triple in $\mathfrak{g}$ is the basis of a Lie subalgebra of $\mathfrak{g}$ isomorphic to $\mathrm{sl}\,(2, \mathbb{C})$. Assertion (f) above states that $(e_\alpha,\ f_\alpha,\ h_\alpha^*)$ is an sl$_2$-triple.

Theorem 8.2.12. *Let Δ be the set of roots of a complex semisimple Lie algebra $\mathfrak{g}$ with respect to a Cartan subalgebra $\mathfrak{h}$. Suppose that $\alpha \in \Delta$. Then* $\dim \mathfrak{g}_\alpha = 1$. *The only multiples of α which are roots are $\pm\alpha$.*

Proof. We retain the notation of Theorem 8.2.11, Part (f). Thus $(e_\alpha, f_\alpha, h_\alpha^*)$ is the basis of a three-dimensional simple Lie subalgebra $\mathfrak{g}^{(\alpha)}$ of $\mathfrak{g}$ isomorphic to $\mathrm{sl}\,(2, \mathbb{C})$. Now $\mathfrak{g}^{(\alpha)}$ acts on $\mathfrak{g}$ via the adjoint representation, so by Weyl's Theorem (Theorem 7.3.7), $\mathfrak{g}$ is a direct sum of irreducible $\mathfrak{g}^{(\alpha)}$-invariant subspaces. By Theorem 7.3.11, each irreducible subspace has a basis consisting of eigenvectors of $\mathrm{ad}\, h_\alpha^*$, whose corresponding eigenvalues are all integers. Thus, $\mathfrak{g}$ has a basis consisting of eigenvectors of $\mathrm{ad}\, h_\alpha^*$ with integer eigenvalues. This implies in particular that *any* eigenvalue of $\mathrm{ad}\, h_\alpha^*$ is an integer.

Suppose that $\beta \in \Delta$. Then any nonzero vector $x_\beta \in \mathfrak{g}_\beta$ is an eigenvector of $\mathrm{ad}\, h_\alpha^*$, with eigenvalue

$$\beta(h_\alpha^*) = 2\,\frac{\beta(h_\alpha)}{\alpha(h_\alpha)} = 2\,\frac{B(\beta, \alpha)}{B(\alpha, \alpha)} \in \mathbb{Z}. \tag{8.13}$$

Suppose that $c\alpha$ is a root, for some $c \in \mathbb{C}$. Then by the integrality condition (8.13),

$$2\,\frac{B(c\alpha, \alpha)}{B(\alpha, \alpha)} = 2c \in \mathbb{Z},$$

and hence $c = m/2$ for some nonzero integer m. Thus any multiple of a root which is also a root must be a half integer multiple of that root. If $|c| > 1$, then, since $\alpha = (1/c)\beta$, and $|1/c| < 1$, this forces $1/c = \pm 1/2$; that is $c = \pm 2$. If $|c| < 1$, then we must have $c = \pm 1/2$.

We conclude, therefore, that the only multiples of α which can be roots are $\pm\alpha/2$ and $\pm 2\alpha$. It is clear that $\alpha/2$ and 2α *cannot both* be roots, since $\alpha/2 = (1/4)\,2\alpha$.

Let us now show that the root space $\mathfrak{g}_\alpha$ is one-dimensional and that 2α cannot be a root. For this, let $\mathfrak{s}$ be the subspace of $\mathfrak{g}$ given by

$$\mathfrak{s} = \mathbb{C} f_\alpha \oplus \mathbb{C} h_\alpha^* \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{2\alpha}.$$

It is straightforward to show, using Theorem 8.2.11, Part (d), and the fact that $k\alpha$ is not a root, for $k > 2$, that $\mathfrak{s}$ is a Lie subalgebra of $\mathfrak{g}$, and that $h_\alpha^* \in [\mathfrak{s}, \mathfrak{s}]$. Let $\mathrm{ad}_\mathfrak{s}$ denote the adjoint representation on $\mathfrak{s}$. Since $h_\alpha^* \in \mathfrak{s}'$, we see that $\mathrm{tr}\,(\mathrm{ad}_\mathfrak{s}\, h_\alpha^*) = 0$. But since $\alpha(h_\alpha^*) = 2$,

$$\begin{aligned}\mathrm{tr}\,(\mathrm{ad}_\mathfrak{s}\, h_\alpha^*) &= -\alpha(h_\alpha^*) + 0 + (\dim \mathfrak{g}_\alpha)\, \alpha(h_\alpha^*) + (\dim \mathfrak{g}_{2\alpha})\, 2\alpha(h_\alpha^*) \\ &= -2 + 0 + 2 \dim \mathfrak{g}_\alpha + 4 \dim \mathfrak{g}_{2\alpha}\end{aligned}$$

Since $\dim \mathfrak{g}_\alpha \geq 1$, we therefore conclude that $\dim \mathfrak{g}_\alpha = 1$ and $\dim \mathfrak{g}_{2\alpha} = 0$. In particular, $2\alpha \notin \Delta$.

From this, we also see that $\alpha/2$ cannot be a root. If it were, then by the argument above, $2(\alpha/2) = \alpha$ cannot be a root, a contradiction.

This completes the proof of Theorem 8.2.12. $\square$

The following theorem is a direct consequence of the representation theorem of $\mathrm{sl}\,(2, \mathbb{C})$ (Theorem 7.3.11) and is a basis of the theory of Weyl groups and general Coxeter groups.

Theorem 8.2.13. *Suppose that α and β are roots of a complex semsimple Lie algebra $\mathfrak{g}$ relative to a Cartan subalgebra $\mathfrak{h}$, with $\alpha \neq \beta$. Let q be the largest integer j such that $\beta + j\alpha$ is a root, and let p be the smallest integer k such that $\beta + k\alpha$ is a root. Then:*

(i) For every j between p and q, $\beta + j\alpha$ is a root.

(ii) $\beta(h^*_\alpha) = 2B(\beta, \alpha)/B(\alpha, \alpha) = -(p + q)$, so is an integer.

(iii) $\beta - \beta(h^*_\alpha)$ is a root.

(iv) $[\mathfrak{g}_\beta, \mathfrak{g}_\alpha] = \mathfrak{g}_{\beta+\alpha}$ if $\beta + \alpha$ is a root.

Remark: By Part (a), the set of roots $\{\beta + j\alpha \,|\, p \leq j \leq q\}$ forms a connected string, which we call the *α-string through β*.

Proof of Theorem 8.2.13: Note that $q \geq 0$ and $p \leq 0$.

As with the proof of the Theorem 8.2.12, we let $\mathfrak{g}^{(\alpha)}$ be the three-dimensional simple Lie subalgebra of $\mathfrak{g}$ with basis $(e_\alpha, f_\alpha, h^*_\alpha)$, which, by Theorem 8.2.11, Part (f), is isomorphic to $\mathrm{sl}\,(2, \mathbb{C})$, Now let V be the subspace of $\mathfrak{g}$ given by

$$V = \bigoplus_{j=p}^{q} \mathfrak{g}_{\beta+j\alpha}, \tag{8.14}$$

where if course $\mathfrak{g}_{\beta+j\alpha} = \{0\}$ if $\beta + j\alpha$ is not a root. It is clear that V is invariant under $\mathrm{ad}\, e_\alpha$, $\mathrm{ad}\, f_\alpha$, and $\mathrm{ad}\, h^*_\alpha$, and so V is $\mathrm{ad}\, \mathfrak{g}^{(\alpha)}$-invariant. By Weyl's theorem (Theorem 7.3.7), V therefore decomposes into a direct sum of irreducible $\mathrm{ad}\, \mathfrak{g}^{(\alpha)}$-invariant subspaces:

$$V = V_1 \oplus \cdots \oplus V_l \tag{8.15}$$

By Theorem 7.3.11, each invariant irreducible subspace V_i further decomposes into a direct sum of one-dimensional eigenspaces of $\mathrm{ad}\, h^*_\alpha$, with the eigenvalues of $\mathrm{ad}\, h^*_\alpha$ on each V_i being integers all having the same parity (i.e., all odd or all even), and corresponding to a symmetric string $\{k, k-2, \ldots, 2-k, -k\}$ about

0. Thus V has a basis consisting of eigenvectors of $\operatorname{ad} h_\alpha^*$, a fact already evident from (8.14).

From equation (8.14), the eigenvalues of $\operatorname{ad} h_\alpha^*$ on V are of the form $(\beta + j\alpha)(h_\alpha^*) = \beta(h_\alpha^*) + 2j$, so they all have the same parity; moreover, the corresponding eigenspaces $\mathfrak{g}_{\beta+j\alpha}$ are all at most one-dimensional. It follows that there is just one irreducible component in the sum (8.15); that is, V is already irreducible.

We now apply Theorem 7.3.11 to V. The largest eigenvalue of $\operatorname{ad} h_\alpha^*$ on V is $(\beta+q\alpha)(h_\alpha^*) = \beta(h_\alpha^*)+2q$, the smallest eigenvalue is $(\beta+p\alpha)(h_\alpha^*) = \beta(h_\alpha^*)+2p$, and they are negatives of each other: $\beta(h_\alpha^*) + 2p = -(\beta(h_\alpha^*) + 2q)$. This proves (ii). Moreover, the integers in the string $\beta(h_\alpha^*)+2j,\ p \leq j \leq q$ are all eigenvalues of $\operatorname{ad} h_\alpha^*$. Hence $\mathfrak{g}_{\beta+j\alpha} \neq \{0\}$ for all $p \leq j \leq q$. This proves (i).

Since $q \geq 0$ and $p \leq 0$, we have $p \leq p + q \leq q$, so by (ii), $\beta - \beta(h_\alpha^*)\alpha = \beta + (p + q)\alpha \in \Delta$. This proves (iii). Finally, if $\beta + \alpha \in \Delta$, then $q \geq 1$, so by Theorem 7.3.11, Part (6), $\operatorname{ad} e_\alpha (\mathfrak{g}_\beta) \neq \{0\}$. Hence $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \neq \{0\}$. This proves (iv).

Theorem 8.2.14. *Let α and β be roots, and let $\beta + k\alpha$ ($p \leq k \leq q$) be the α-string through β. Let x_α, $x_{-\alpha}$, and x_β be any vectors in $\mathfrak{g}_\alpha$, $\mathfrak{g}_{-\alpha}$, and $\mathfrak{g}_\beta$, respectively. Then*

$$[x_{-\alpha}, [x_\alpha, x_\beta]] = \frac{q(1-p)\,\alpha(h_\alpha)}{2}\, B(x_\alpha, x_{-\alpha})\, x_\beta. \tag{8.16}$$

Proof. Note that if any of x_α, $x_{-\alpha}$, or x_β is 0, then both sides in equation (8.16) equal 0. Hence we may assume that these vectors are all nonzero.

Let $(e_\alpha,\ f_\alpha,\ h_\alpha^*)$ be the sl_2-triple whose existence is guaranteed by Theorem 8.2.11, Part (f), and let $\mathfrak{s}_\alpha$ denote its linear span. Then $\mathfrak{s}_\alpha$ is a subalgebra of $\mathfrak{g}$ isomorphic to $\mathrm{sl}\,(2, \mathbb{C})$.

Since $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ are one-dimensional, there exist nonzero complex constants c_α and $c_{-\alpha}$ such that $x_\alpha = c_\alpha\, e_\alpha$ and $x_{-\alpha} = c_{-\alpha}\, f_\alpha$. By Theorem 8.2.11, Part (c), we can obtain a relation between c_α and $c_{-\alpha}$:

$$\begin{aligned} B(x_\alpha, x_{-\alpha}) h_\alpha &= [x_\alpha, x_{-\alpha}] \\ &= c_\alpha\, c_{-\alpha}\, [e_\alpha, f_\alpha] \\ &= c_\alpha\, c_{-\alpha}\, h_\alpha^* \\ &= 2\, \frac{c_\alpha\, c_{-\alpha}}{\alpha(h_\alpha)}\, h_\alpha. \end{aligned}$$

Hence

$$c_\alpha\, c_{-\alpha} = \frac{\alpha(h_\alpha)}{2}\, B(x_\alpha, x_{-\alpha}).$$

By the proof of Theorem 8.2.13, $\mathfrak{s}_\alpha$ acts irreducibly on the subspace $\oplus_{k=p}^{q}\mathfrak{g}_{\beta+j\alpha}$ of $\mathfrak{g}$ via the adjoint representation, with highest weight $n = (\beta + q\alpha)(h_\alpha^*) = -(p+q) + 2q = q - p$.

Let v_0 be any nonzero element of $\mathfrak{g}_{\beta+q\alpha}$. Then by the representation theorem for $\mathrm{sl}\,(2,\mathbb{C})$ (Theorem 7.3.11), we know that $(\mathrm{ad}\, f_\alpha)^q\,(v_0) = c\, x_\beta$, for some nonzero constant c. Now replace v_0 by $(1/c)\, v_0$, so that we may assume that $(\mathrm{ad}\, f_\alpha)^q\,(v_0) = x_\beta$.

According to the notation of Theorem 7.3.11, $(\mathrm{ad}\, f_\alpha)^q\,(v_0)$ is the vector v_q. From Part 6 of that theorem, we obtain the relation

$$\mathrm{ad}\, e_\alpha\,(v_q) = q(n - q + 1)v_{q-1}$$

and so

$$\mathrm{ad}\, f_\alpha \circ \mathrm{ad}\, e_\alpha\,(v_q) = q(n - q + 1)\, v_q.$$

Thus,

$$[f_\alpha, [e_\alpha, x_\beta]] = q(1-p)\, x_\beta.$$

Multiplying both sides above by $c_\alpha\, c_{-\alpha}$, we obtain

$$[x_{-\alpha}, [x_\alpha, x_\beta]] = q(1-p)\,\frac{\alpha(h_\alpha)}{2}\, B(x_\alpha, x_{-\alpha})\, x_\beta,$$

as desired. □

For the rest of this section, we assume that Δ is the set of roots of a complex semisimple Lie algebra $\mathfrak{g}$ relative to a Cartan subalgebra $\mathfrak{h}$. Our next objective is to show that the Killing form B is positive definite on the real linear span of the root vectors h_α, for all $\alpha \in \Delta$. This linear span, which we denote by $h_{\mathbb{R}}$, will thus be a real inner product space.

Theorem 8.2.15. *Let $B = (\alpha_1, \ldots, \alpha_l)$ be any basis of $\mathfrak{h}^*$ consisting of roots. (This is possible by Theorem 8.2.11, Part (a).) Then any root β is a linear combination of the α_j, with rational coefficients.*

Proof. Let $\beta \in \Delta$. Then we can certainly write $\beta = \sum_{j=1}^{l} c_j\, \alpha_j$, where $c_j \in \mathbb{C}$ for all j. Hence, for any i, we obtain

$$B(\beta, \alpha_i) = \sum_{j=1}^{l} c_j\, B(\alpha_j, \alpha_i).$$

By Theorem 8.2.11, Part (e), $B(\alpha_i, \alpha_i) \neq 0$. Hence

$$\frac{2\,B(\beta, \alpha_i)}{B(\alpha_i, \alpha_i)} = \sum_{j=1}^{n} c_j\, \frac{2\,B(\alpha_j, \alpha_i)}{B(\alpha_i, \alpha_i)} \tag{8.17}$$

For any i and j in $\{1, \ldots, n\}$, let $n_i = 2B(\beta, \alpha_i)/B(\alpha_i, \alpha_i)$ and let $A_{ji} = 2B(\alpha_j, \alpha_i)/B(\alpha_i, \alpha_i)$. Then by Theorem 8.2.13, Part (ii), all the n_i and A_{ji} are integers. Now the linear system (8.17) corresponds to the matrix equation

$$\begin{pmatrix} c_1 & \cdots & c_l \end{pmatrix} \begin{pmatrix} A_{11} & \cdots & A_{1l} \\ \vdots & \ddots & \vdots \\ A_{l1} & \cdots & A_{ll} \end{pmatrix} = \begin{pmatrix} n_1 & \cdots & n_l \end{pmatrix}. \tag{8.18}$$

The coefficient matrix (A_{ji}) of the above matrix equation is equals the product

$$\begin{pmatrix} B(\alpha_1, \alpha_1) & \cdots & B(\alpha_1, \alpha_l) \\ \vdots & \ddots & \vdots \\ B(\alpha_l, \alpha_1) & \cdots & B(\alpha_l, \alpha_l) \end{pmatrix} \begin{pmatrix} 2/B(\alpha_1, \alpha_1) & & 0 \\ & \ddots & \\ 0 & & 2/B(\alpha_l, \alpha_l) \end{pmatrix}$$

Since B is nondegenerate on $\mathfrak{h}^*$, the matrix on the left above is nonsingular by Theorem 1.10.4; the matrix on the right is clearly nonsingular since it's a diagonal matrix with nonzero entries on the diagonal. This shows that (A_{ji}) is a nonsingular matrix with integer entries. Its inverse $(A_{ji})^{-1}$ is therefore a matrix with *rational* entries. Now we can solve for the coefficients c_j in the system (8.18) by multiplying both sides on the right by$(A_{ji})^{-1}$:

$$\begin{pmatrix} c_1 & \cdots & c_l \end{pmatrix} = \begin{pmatrix} n_1 & \cdots & n_l \end{pmatrix} \begin{pmatrix} A_{11} & \cdots & A_{1l} \\ \vdots & \ddots & \vdots \\ A_{l1} & \cdots & A_{ll} \end{pmatrix}^{-1}.$$

This shows that $c_j \in \mathbb{Q}$, for all j. ☐

Lemma 8.2.16. *Let α and β be roots. Then $B(\alpha, \beta) \in \mathbb{Q}$.*

Proof. For any $h \in \mathfrak{h}$, the root space decomposition $\mathfrak{g} = \mathfrak{h} \oplus \sum_{\sigma \in \Delta} \mathfrak{g}_\sigma$ is a decomposition of $\mathfrak{g}$ into eigenspaces of $\operatorname{ad} h$ with eigenvalues 0, of multiplicity $\dim \mathfrak{h}$, and $\sigma(h)$ (for all $\sigma \in \Delta$), of multiplicity 1.

For α and β in Δ, let h_α and h_β be the corresponding vectors in $\mathfrak{h}$, in accordance with Definition 8.2.10. Then

$$\begin{aligned} B(\alpha, \beta) &= B(h_\alpha, h_\beta) \\ &= \operatorname{tr}\, (\operatorname{ad} h_\alpha \circ \operatorname{ad} h_\beta) \\ &= \sum_{\sigma \in \Delta} \sigma(h_\alpha) \sigma(h_\beta) \\ &= \sum_{\sigma \in \Delta} B(\sigma, \alpha)\, B(\sigma, \beta). \end{aligned} \tag{8.19}$$

By Theorem 8.2.11 Part (e), $B(\alpha, \alpha) \neq 0$ and $B(\beta, \beta) \neq 0$. Hence we can divide both sides of (8.19) by these to get

$$\frac{4\, B(\alpha, \beta)}{B(\alpha, \alpha)\, B(\beta, \beta)} = \sum_{\sigma \in \Delta} \frac{2\, B(\sigma, \alpha)}{B(\alpha, \alpha)} \cdot \frac{2\, B(\sigma, \beta)}{B(\beta, \beta)} \tag{8.20}$$

Now by Theorem 8.2.13, Part (ii), the terms in the sum in the right hand side of (8.20) are integers. Hence the left hand side of (8.20) is an integer.

We want to prove that $B(\alpha,\beta)$ is rational. If $B(\alpha,\beta)=0$, there is nothing to prove, so let us assume that $B(\alpha,\beta)\neq 0$. Then, again from Theorem 8.2.13, Part (ii), the fraction

$$\frac{4\,B(\alpha,\beta)^2}{B(\alpha,\alpha)\,B(\beta,\beta)} = \frac{2\,B(\beta,\alpha)}{B(\alpha,\alpha)}\cdot\frac{2\,B(\alpha,\beta)}{B(\beta,\beta)}$$

is an integer. Dividing this by the (nonzero) integer representing the left hand side of (8.20), we see that

$$B(\alpha,\beta) \;=\; \frac{\dfrac{4\,B(\alpha,\beta)^2}{B(\alpha,\alpha)\,B(\beta,\beta)}}{\dfrac{4\,B(\alpha,\beta)}{B(\alpha,\alpha)\,B(\beta,\beta)}}$$

is rational. □

Theorem 8.2.17. *Let $\mathfrak{h}_{\mathbb{R}}=\sum_{\alpha\in\Delta}\mathbb{R}\,h_\alpha$, the real vector space spanned by the vectors h_α ($\alpha\in\Delta$). Then the Killing form B is positive definite on $\mathfrak{h}_{\mathbb{R}}$.*

Proof. Certainly, Lemma 8.2.16 implies that $B(h_1,h_2)\in\mathbb{R}$ for all h_1, $h_2\in\mathfrak{h}_{\mathbb{R}}$. Thus B is a real-valued symmetric bilinear form on $\mathfrak{h}_{\mathbb{R}}$. We want to prove that $B(h,h)>0$ for all nonzero vectors h in $\mathfrak{h}_{\mathbb{R}}$.

For any $h\in\mathfrak{h}$, the root space decomposition (8.9) shows that the eigenvalues of the semisimple linear operator $\operatorname{ad} h$ on $\mathfrak{g}$ are 0 and $\alpha(h)$, for all $\alpha\in\Delta$. Thus, if $\alpha(h)=0$ for all roots α, it would follow that $\operatorname{ad} h=0$, and so $h=0$. Consequently, if $h\neq 0$ in $\mathfrak{h}$, then $\alpha(h)\neq 0$ for some root α.

Now suppose that $h\in\mathfrak{h}_{\mathbb{R}}$. Then $h=\sum_{\alpha\in\Delta}c_\alpha h_\alpha$ for some scalars $c_\alpha\in\mathbb{R}$. For any root σ, the scalar $\sigma(h)=\sum_{\alpha\in\Delta}c_\alpha\sigma(h_\alpha)=\sum_{\alpha\in\Delta}c_\alpha\,B(\sigma,\alpha)$ is real, by Lemma 8.2.16. Therefore, by the root space decomposition,

$$\begin{aligned} B(h,h) &= \operatorname{tr}\,(\operatorname{ad} h\circ\operatorname{ad} h)\\ &= \sum_{\sigma\in\Delta}\sigma(h)^2>0 \end{aligned}$$

whenever $h\neq 0$ in $\mathfrak{h}_{\mathbb{R}}$. □

We conclude from this theorem that $\mathfrak{h}_{\mathbb{R}}$ is a real inner product space, with inner product given by the Killing form B. The map $h\mapsto B(h,\cdot)$ (which takes h_α to α, for all $\alpha\in\Delta$) identifies $\mathfrak{h}_{\mathbb{R}}$ with its real dual $\mathfrak{h}_{\mathbb{R}}^*$.

Corollary 8.2.18. *Suppose α and β are roots such that $B(\beta,\alpha)<0$. Then $\beta+\alpha$ is a root. If $B(\beta,\alpha)>0$, then $\beta-\alpha$ is a root.*

Proof. Suppose that $B(\beta, \alpha) < 0$. Now by Theorem 8.2.13, Part (ii), $2\,B(\beta, \alpha)/B(\alpha, \alpha) = -(p+q)$. Since $B(\alpha, \alpha) = B(h_\alpha, h_\alpha) > 0$, this implies that $q > 0$. Hence by Part (i), $\beta + \alpha$ is a root. If $B(\beta, \alpha) > 0$, then $p < 0$, so $\beta - \alpha$ is a root. $\square$

8.3 Uniqueness of the Root Pattern

In this section our objective is to prove that if two complex semisimple Lie algebras $\mathfrak{g}$ and $\mathfrak{g}'$ have the same root pattern, then they are isomorphic. More precisely, we will prove the following theorem.

Theorem 8.3.1. *Suppose that $\mathfrak{g}$ and $\mathfrak{g}'$ are complex semisimple Lie algebras with Cartan subalgebras $\mathfrak{h}$ and $\mathfrak{h}'$, respectively. Let Δ be the set of roots of $\mathfrak{g}$ relative to $\mathfrak{h}$ and let Δ' be the set of roots of $\mathfrak{g}_1$ relative to $\mathfrak{h}'$. If $\varphi : \mathfrak{h} \to \mathfrak{h}'$ is a linear bijection such that ${}^t\varphi(\Delta') = \Delta$, then φ extends to an isomorphism of $\mathfrak{g}$ onto $\mathfrak{g}'$.*

(Note: Here $\mathfrak{g}'$ *does not* refer to the derived algebra of $\mathfrak{g}$. It's just some other complex semisimple Lie algebra.)

Recall that ${}^t\varphi$ is the linear map from the dual space $(\mathfrak{h}')^*$ into the dual space $\mathfrak{h}^*$ given by ${}^t\varphi(\lambda) = \lambda \circ \varphi$, for all $\lambda \in (\mathfrak{h}')^*$. Since $\Delta \subset \mathfrak{h}^*$ and $\Delta' \subset (\mathfrak{h}')^*$, the requirement ${}^t\varphi(\Delta') = \Delta$ in the theorem above makes sense.

We will follow the arguments in [4], Chapter 3, §5.

For now, we focus on a complex semisimple Lie algebra $\mathfrak{g}$, its Cartan subalgebra $\mathfrak{h}$, and the set of roots Δ of $\mathfrak{g}$ relative to $\mathfrak{h}$. For each $\alpha \in \Delta$, we choose fix a vector $e_\alpha \in \mathfrak{g}_\alpha$. We can choose the e_α to have the property that $B(e_\alpha, e_{-\alpha}) = 1$ for all $\alpha \in \Delta$. Then by Theorem 8.2.11 Part (c), $[e_\alpha, e_{-\alpha}] = h_\alpha$, for all $\alpha \in \Delta$.

Now let S be a subset of Δ. The *hull* $\overline{S}$ of S is the set of all roots of the form $\pm\alpha$, $\pm(\alpha+\beta)$, for all α, $\beta \in S$. Thus $\overline{S} \subset \Delta$.

Suppose that γ and δ are in the hull $\overline{S}$ with $\gamma + \delta \neq 0$, and suppose that either $\gamma + \delta \in \overline{S}$ or $\gamma + \delta \notin \overline{\Delta}$. We define the complex scalar $N_{\gamma,\delta}$ as follows. If $\gamma + \delta \in \Delta$, then $[e_\gamma, e_\delta] \in \mathfrak{g}_{\gamma+\delta}$, so we have $[e_\gamma, e_\delta] = N_{\gamma,\delta} e_{\gamma+\delta}$, where $N_{\gamma,\delta}$ is uniquely determined. If $\gamma + \delta \notin \Delta$, put $N_{\gamma,\delta} = 0$.

Thus $N_{\gamma\delta}$ is defined for $\gamma, \delta \in \overline{S}$ when and only when:

1. $\gamma + \delta \neq 0$, and
2. $\gamma + \delta \in \overline{S}$, or $\gamma + \delta \notin \Delta$.

Clearly, $N_{\delta,\gamma} = -N_{\gamma,\delta}$.

Proposition 8.3.2. *Suppose that* α, β, $\gamma \in \overline{S}$ *such that* $\alpha + \beta + \gamma = 0$. *Then* $N_{\alpha,\beta} = N_{\beta,\gamma} = N_{\gamma,\alpha}$.

Proof. Note that the constants $N_{\alpha,\beta}$, $N_{\beta,\gamma}$, and $N_{\gamma,\alpha}$ are all defined. By the Jacobi identity

$$\begin{aligned} [e_\alpha,[e_\beta,e_\gamma]] &= -[e_\beta,[e_\gamma,e_\alpha]] - [e_\gamma,[e_\alpha,e_\beta]] \\ \implies N_{\beta,\gamma}\, h_\alpha &= -N_{\gamma,\alpha}\, h_\beta \,-\, N_{\alpha,\beta}\, h_\gamma \end{aligned}$$

But $h_\alpha = -h_\beta - h_\gamma$, and so

$$-N_{\beta,\gamma}\, h_\beta \,-\, N_{\beta,\gamma}\, h_\gamma \;=\; -N_{\gamma,\alpha}\, h_\beta \,-\, N_{\alpha,\beta}\, h_\gamma$$

Now β and γ are linearly independent (otherwise $\beta = \pm\gamma$), and so h_β and h_γ are linearly independent. The last equation above thus establishes the proposition. □

Corollary 8.3.3. *Suppose that* α, $\beta \in \overline{S}$ *such that* $N_{\alpha\beta}$ *exists. Then*

$$N_{\alpha,\beta}N_{-\alpha,-\beta} = -\frac{q(1-p)}{2}\,\alpha(h_\alpha), \tag{8.21}$$

where, as usual, $\beta + j\alpha$ $(p \le j \le q)$ *is the* α*-string through* β.

Proof. By Theorem 8.2.14,

$$\begin{aligned} \frac{q(1-p)}{2}\,\alpha(h_\alpha)\, e_\beta &= [e_{-\alpha},[e_\alpha,e_\beta]] \\ &= N_{-\alpha,\alpha+\beta}\, N_{\alpha,\beta}\, e_\beta, \end{aligned}$$

and so $N_{-\alpha,\alpha+\beta}N_{\alpha,\beta} = q(1-p)\alpha(h_\alpha)/2$. Applying Proposition 8.3.2 to the triple $-\alpha$, $\alpha+\beta$, $-\beta$, we obtain $N_{-\alpha,\alpha+\beta} = N_{\alpha+\beta,-\beta} = N_{-\beta,-\alpha}$. Then using $N_{-\beta,-\alpha} = -N_{-\alpha,-\beta}$, we obtain the conclusion. □

Proposition 8.3.4. *Suppose that* α, β, γ, $\delta \in S$ *such that* $\alpha+\beta+\gamma+\delta = 0$. *Then* $N_{\alpha,\beta}N_{\gamma,\delta} + N_{\gamma,\alpha}N_{\beta,\delta} + N_{\alpha,\delta}N_{\beta,\gamma} = 0$.

Proof. Note that all the N's above are defined. Now since $\beta+\gamma+\delta = -\alpha$, we have

$$\begin{aligned} B(e_\alpha,[e_\beta,[e_\gamma,e_\delta]]) &= N_{\beta,\gamma+\delta}\, N_{\gamma,\delta}\, B(e_\alpha,e_{-\alpha}) \\ &= N_{\beta,\gamma+\delta}\, N_{\gamma,\delta}. \end{aligned}$$

Applying the Jacobi identity to the bracket on the left hand side, we obtain

$$N_{\beta,\gamma+\delta}N_{\gamma,\delta} + N_{\gamma,\delta+\beta}N_{\delta,\beta} + N_{\delta,\beta+\gamma}N_{\beta\,gamma} = 0.$$

But from Proposition 8.3.2, $N_{\beta,\gamma+\delta} = N_{\gamma+\delta,\alpha} = N_{\alpha,\beta}$, $N_{\gamma,\delta+\beta} = N_{\delta+\beta,\alpha} = N_{\alpha,\gamma}$, and $N_{\delta,\beta+\gamma} = N_{\beta+\gamma,\alpha} = N_{\alpha,\delta}$. This proves the proposition. □

For convenience, we will introduce what is called a lexicographic order to the real dual space $\mathfrak{h}_{\mathbb{R}}^*$, and thus to $\mathfrak{h}_{\mathbb{R}}$. An *ordered vector space* is a pair $(V, >)$ consisting of real vector space V and a total ordering $>$ on V which is preserved under vector addition and multiplication by positive scalars. Thus we require that for any u, v, and w in V and any positive scalar c,

$$u > v \implies u + w > v + w \text{ and } cu > cv.$$

Any vector $v > 0$ is called a *positive vector*; if $v < 0$, we call v a *negative vector*.

Let V be any real vector space. For a fixed basis $B = (v_1, \dots, v_n)$ of V, we can turn V into an ordered vector space by introducing the *lexicographic ordering relative to* B, defined as follows: for any nonzero vector $v \in V$, let us write $v = a_1 v_1 + \cdots + a_n v_n$. Let j be the smallest integer such that $a_j \neq 0$. By definition, $v > 0$ iff $a_j > 0$. Then, if v and w are any vectors in V, we define $v > w$ iff $v - w > 0$. It is straightforward to prove that $>$ is a total ordering on V which turns V into an ordered vector space.

Exercise 8.3.5. Suppose that $(V, >)$ is an ordered vector space. Prove that there is a basis B of V such that the total order $>$ is the lexicographic order relative to B.

We are now ready to prove Theorem 8.3.1.

Proof of Theorem 8.3.1: Let us first fix notation. For each $\alpha \in \Delta$, let α' be the unique element of Δ' such that ${}^t\varphi(\alpha') = \alpha$. Thus, in particular, $\alpha'(\varphi(h)) = \alpha(h)$ for all $h \in \mathfrak{h}$.

Let B and B' denote the Killing forms on $\mathfrak{g}$ and $\mathfrak{g}'$, respectively. The first thing we'll do is to prove that φ is an *isometry* with respect to B and B'. More precisely, we will prove that

$$B'(\varphi(h), \varphi(h_1)) = B(h, h_1) \tag{8.22}$$

for all h, $h_1 \in \mathfrak{h}$.

In fact, if α, $\beta \in \Delta$, then

$$\begin{aligned} B'(\varphi(h_\alpha), \varphi(h_\beta)) &= \operatorname{tr}(\operatorname{ad}\varphi(h_\alpha) \circ \operatorname{ad}\varphi(h_\beta)) \\ &= \sum_{\gamma' \in \Delta'} \gamma'(\varphi(h_\alpha))\, \gamma'(\varphi(h_\beta)) \\ &= \sum_{\gamma \in \Delta} \gamma(h_\alpha)\, \gamma(h_\beta) \\ &= B(h_\alpha, h_\beta) \end{aligned}$$

Since every vector in $\mathfrak{h}$ is a $\mathbb{C}$-linear combination of the h_α, equation (8.22) follows by linearity.

For all $\alpha \in \Delta$ and $h \in \mathfrak{h}$, we then see that $B'(\varphi(h_\alpha), \varphi(h)) = B(h_\alpha, h) = \alpha(h) = \alpha'(\varphi(h)) = B'(h_{\alpha'}, \varphi(h))$. Thus, since B' is nondegenerate on $\mathfrak{h}'$, we have $\varphi(h_\alpha) = h_{\alpha'}$. From this, we obtain $B(h_\alpha, h_\beta) = B'(h_{\alpha'}, h_{\beta'})$ for all α, $\beta \in \Delta$; moreover, $\varphi(\mathfrak{h}_\mathbb{R}) = \mathfrak{h}'_\mathbb{R}$.

The real dual space $\mathfrak{h}^*_\mathbb{R}$ is the $\mathbb{R}$-linear span $\sum_{\alpha\in\Delta} \mathbb{R}\alpha$, and likewise $(\mathfrak{h}'_\mathbb{R})^*$ is the $\mathbb{R}$-linear span of Δ'. Fix a basis E of $(\mathfrak{h}'_\mathbb{R})^*$, and let $>$ be the lexicographic order on $(\mathfrak{h}'_\mathbb{R})^*$ with respect to E. We also let $>$ be the lexicographic order on $\mathfrak{h}^*_\mathbb{R}$ with respect to its basis ${}^t\varphi(E)$. These orders are obviously compatible in the sense that $\lambda' > \mu'$ iff ${}^t\varphi(\lambda') > {}^t\varphi(\mu')$ for all λ', $\mu' \in (\mathfrak{h}'_\mathbb{R})^*$. In particular, $\alpha > \beta$ iff $\alpha' > \beta'$ for all α, $\beta \in \Delta$.

Let us apply the discussion prior to Proposition 8.3.2 to $S = \Delta$. Thus for all $\alpha \in \Delta$, we choose vectors $e_\alpha \in \mathfrak{g}_\alpha$ such that $B(e_\alpha, e_{-\alpha}) = 1$. Then, for each pair of roots α, $\beta \in \Delta$ such that $\alpha + \beta \neq 0$, the scalar $N_{\alpha,\beta} \in \mathbb{C}$ is defined by

$$[e_\alpha, e_\beta] = N_{\alpha,\beta}\, e_{\alpha+\beta}$$

if $\alpha + \beta \in \Delta$, and $N_{\alpha,\beta} = 0$ if $\alpha + \beta \notin \Delta$. The $N_{\alpha,\beta}$ satisfy the conclusions of Proposition 8.3.2, Corollary 8.3.3, and Proposition 8.3.4.

For each $\alpha' \in \Delta'$, we claim that there exists a vector $e_{\alpha'} \in \mathfrak{g}'_{\alpha'}$ such that

$$B'(e_{\alpha'}, e_{-\alpha'}) = 1 \tag{8.23}$$

$$[e_{\alpha'}, e_{\beta'}] = N_{\alpha,\beta}\, e_{\alpha'+\beta'} \qquad (\alpha' + \beta' \neq 0) \tag{8.24}$$

Assuming that the the vectors $e_{\alpha'}$ exist, we extend the linear map φ to all of $\mathfrak{g}$ by putting $\varphi(e_\alpha) = e_{\alpha'}$ for all α. Then, for all $h \in \mathfrak{h}$, we have

$$[\varphi(h), \varphi(e_\alpha)] = [\varphi(h), e_{\alpha'}] = \alpha'(\varphi(h))\, e_{\alpha'} = \alpha(h)\, e_{\alpha'} = \varphi[h, e_\alpha].$$

This relation and (8.24) then show that this extension of φ is our desired isomorphism of $\mathfrak{g}$ onto $\mathfrak{g}'$.

So everything hinges on proving the existence of the vectors $e_{\alpha'}$ satisfying equations (8.23) and (8.24).

We will do this by induction with respect to $>$. For each positive root $\rho \in \Delta$, let Δ_ρ denote the set of all $\alpha \in \Delta$ such that $-\rho < \alpha < \rho$. For $\rho' \in \Delta'$, define $\Delta'_{\rho'}$ similarly.

Order the positive roots in Δ by $\rho_1 < \rho_2 < \cdots < \rho_m$, and let ρ_{m+1} be any vector in $\mathfrak{h}^*_\mathbb{R}$ larger than ρ_m. By induction on j, we will prove the following claim:

Claim: *For each $\alpha \in \Delta_{\rho_j}$, a vector $e_{\alpha'} \in \mathfrak{g}'_{\alpha'}$ can be chosen such that equation* (8.23) *holds whenever $\alpha \in \Delta_{\rho_j}$ and equation* (8.24) *holds whenever α, β, $\alpha+\beta \in \Delta_{\rho_j}$.*

Since $\Delta_{\rho_{m+1}} = \Delta$, the proof will be complete by the $(m+1)$st induction step. Note that for all $j \le m$, $\Delta_{\rho_{j+1}} = \Delta_{\rho_j} \cup \{\rho_j, -\rho_j\}$.

For $j = 1$, we have $\Delta_{\rho_1} = \emptyset$, so the claim certainly holds and there is nothing to prove. So assume that the claim holds for Δ_{ρ_j}. We wish to prove it for $\Delta_{\rho_{j+1}}$. For this, we just need to define $e_{\rho'_j}$ and $e_{-\rho'_j}$ in an appropriate manner so that the claim still holds for $\Delta_{\rho_{j+1}}$.

If there are *no* α, $\beta \in \Delta_{\rho_j}$ such that $\rho_j = \alpha + \beta$, then we can choose $e_{\rho'_j}$ to be any nonzero vector in $\mathfrak{g}'_{\rho'_j}$, and then let $e_{-\rho'_j}$ be the vector in $\mathfrak{g}'_{-\rho'_j}$ satisfying $B'(e_{\rho'_j}, e_{-\rho'_j}) = 1$. In this case, the claim then holds for $\Delta_{\rho_{j+1}}$. In fact, if α, β, $\alpha + \beta \in \Delta_{\rho_{j+1}}$, then the only cases that are not necessarily covered by the induction hypothesis occur when α or β equals $\pm\rho_j$. If $\alpha = \rho_j$, then $\alpha + \beta$ cannot be $\pm\rho_j$, and so we would have $\rho_j = (\alpha + \beta) + (-\beta)$, where both $\alpha + \beta$ and $-\beta$ are in Δ_{ρ_j}. This is impossible, since ρ_j is not the sum of elements in Δ_{ρ_j}. The other possibilities $\alpha = -\rho_j$, $\beta = \pm\rho_j$ also cannot occur. Thus, the only cases that occur are already covered by the induction hypothesis, and so the claim holds for $\Delta_{\rho_{j+1}}$.

So suppose that there are roots α, $\beta \in \Delta_{\rho_j}$ such that $\alpha + \beta = \rho_j$. Note that any such pair of roots must be positive: if, for instance, $\beta < 0$, then $\alpha = \rho_j - \beta > \rho_j$, contradicting $\alpha \in \Delta_{\rho_j}$.

Among all such pairs of roots, let α, β be the pair such that α is as small as possible. Define $e_{\rho'_j} \in \mathfrak{g}'_{\rho'_j}$ by the condition

$$[e_{\alpha'}, e_{\beta'}] = N_{\alpha,\beta}\, e_{\rho'_j}. \tag{8.25}$$

Then let $e_{-\rho'_j}$ be the vector in $\mathfrak{g}'_{-\rho'_j}$ such that $B'(e_{\rho'_j}, e_{-\rho'_j}) = 1$.

To prove that the claim holds for $\Delta_{\rho_{j+1}}$, we just need to verify that

$$[e_{\gamma'}, e_{\delta'}] = N_{\gamma,\delta}\, e_{\gamma'+\delta'} \tag{8.26}$$

whenever γ, δ, $\gamma + \delta \in \Delta_{\rho_{j+1}}$. For this, we'll need to consider several cases:

1. γ, δ, *and* $\gamma+\delta$ *belong to* Δ_{ρ_j}. By the induction hypothesis, equation (8.26) holds and there is nothing to prove.

2. $\gamma + \delta = \rho_j$. In this case, we can assume that $\{\gamma, \delta\} \neq \{\alpha, \beta\}$. Note that γ and δ are positive. Now $\alpha + \beta - \gamma - \delta = 0$, and no two of the roots α, β, $-\gamma$, $-\delta$ have sum 0. Thus by Proposition 8.3.4 (for $S = \Delta$), we have

$$N_{\alpha,\beta}\, N_{-\gamma,-\delta} = -N_{-\gamma,\alpha}\, N_{\beta,-\delta} - N_{\alpha,-\delta}\, N_{\beta,-\gamma} \tag{8.27}$$

Moreover, by Corollary 8.3.3,

$$N_{\gamma,\delta}\, N_{-\gamma,-\delta} = -\frac{l(1-k)}{2}\, \gamma(h_\gamma), \tag{8.28}$$

where $\delta + s\gamma$, $k \leq s \leq l$ is the γ-string through δ.

Let $S = \{\alpha, \beta, -\gamma, -\delta\}$. As in the beginning of this section, for $\mu, \nu \in \overline{S}$ such that $\mu + \nu \neq 0$, we define the scalar $M_{\mu,\nu}$ under the condition that either $\mu + \nu \in \overline{S}$ or $\mu + \nu \notin \Delta$ by:

$$\begin{aligned} [e_{\mu'}, e_{\nu'}] &= M_{\mu,\nu}\, e_{\mu'+\nu'} && \text{if } \mu + \nu \in \overline{S} \\ M_{\mu,\nu} &= 0 && \text{if } \mu + \nu \notin \Delta \end{aligned} \tag{8.29}$$

By construction, we already have $M_{\alpha,\beta} = N_{\alpha,\beta}$. In addition, by the induction hypothesis, we also have $M_{\mu,\nu} = N_{\mu,\nu}$ whenever μ, ν, and $\mu+\nu$ are in Δ_{ρ_j}.

Now by Corollary 8.3.3, we have

$$M_{\alpha,\beta}\, M_{-\gamma,-\delta} = -M_{-\gamma,\alpha}\, M_{\beta,-\delta} - M_{\alpha,-\delta}\, M_{\beta,-\gamma}. \tag{8.30}$$

But then the all the terms on the right hand side above coincide with the corresponding terms on the right hand side of equation (8.27). Hence $M_{-\gamma,-\delta} = N_{-\gamma,-\delta}$. Also, by Corollary 8.3.3, we have

$$M_{\gamma,\delta}\, M_{-\gamma,-\delta} = -\frac{l'(1-k')}{2}\, \gamma'(h_{\gamma'}) \tag{8.31}$$

where $\delta' + s\gamma'$, $k' \leq s \leq l'$ is the γ'-string through δ'. But $\gamma'(h_{\gamma'}) = \gamma(h_\gamma)$, and by the hypothesis ${}^t\varphi(\Delta') = \Delta$, we have ${}^t\varphi(\delta' + s\gamma') = \delta + s\gamma$, for all s. Hence $k = k'$ and $l = l'$. It follows that the right hand side in (8.31) equals that in (8.28), whence $M_{\gamma,\delta}\, N_{-\gamma,-\delta} = N_{\gamma,\delta}\, N_{-\gamma,-\delta}$. It follows that $M_{\gamma,\delta} = N_{\gamma,\delta}$.

3. $\gamma + \delta = -\rho_j$. Then $-\gamma - \delta = \rho_j$. By Case 2, we have $[e_{-\gamma'}, e_{-\delta'}] = N_{-\gamma,-\delta}\, e_{\rho_j'}$. Let $S = \{\gamma, \delta, -\rho_j\}$. Then define the scalars $M_{\mu,\nu}$ for $\overline{S}$ in a manner analogous to (8.29). Then

   $$M_{\gamma,\delta}\, M_{-\gamma,-\delta} = -\frac{l(1-k)}{2}\, \gamma(h_\gamma) = N_{\gamma,\delta}\, N_{-\gamma,-\delta},$$

   where $\delta + s\gamma$, $k \leq s \leq l$ is the γ-string through δ. Since $M_{-\gamma,-\delta} = N_{-\gamma,-\delta}$, it follows from the above that $M_{\gamma,\delta} = N_{\gamma,\delta}$. Thus, $[e_{\gamma'}, e_{\delta'}] = N_{\gamma,\delta}\, e_{\gamma'+\delta'}$.

4. *One of γ or δ is $\pm\rho_j$.* Suppose, for instance, that $\gamma = -\rho_j$. Then by $-\rho_j \leq -\rho_j + \delta \leq \rho_j$, we obtain $\delta > 0$; since $\delta \leq \rho_j$, we then conclude that $-\rho_j + \delta \leq 0$; since $-\rho_j + \delta$ is a root, we must have $-\rho_j + \delta < 0$. Thus, in fact $\delta < \rho_j$.

   From this we obtain that $\rho_j = \delta + (\rho_j - \delta)$, where both $\rho_j - \delta$ and δ lie in Δ_{ρ_j}. Now let $S = \{\delta, \rho_j - \delta, -\rho_j\}$, and define scalars $M_{\mu,\nu}$ for $\mu, \nu \in \overline{S}$ as in (8.29). From Case 2, we have $M_{\delta,\rho_j-\delta} = N_{\delta,\rho_j-\delta}$. Then by Proposition 8.3.2, we have

   $$M_{\delta,\rho_j-\delta} = M_{\rho_j-\delta,-\rho_j} = M_{-\rho_j,\delta}.$$

Since we also have

$$N_{\delta,\rho_j-\delta} = N_{\rho_j-\delta,-\rho_j} = N_{-\rho_j,\delta},$$

we conclude that $M_{-\rho_j,\delta} = N_{-\rho_j,\delta}$. The cases $\gamma = \rho_j$, $\delta = \pm\rho_j$ are treated in a similar fashion.

This completes the induction step and concludes the proof of Theorem 8.3.1. $\square$

Chapter 9

The Classical Simple Complex Lie Algebras

In this chapter, we study the four classes of classical complex simple Lie algebras. These are all matrix algebras, and their root space decompositions can be expressed quite explicitly.

9.1 Type A_l

Let l be an integer ≥ 1, and let $\mathfrak{g} = \mathrm{sl}(l+1, \mathbb{C})$. (We use $\mathrm{sl}(l+1, \mathbb{C})$ instead of $\mathrm{sl}(l, \mathbb{C})$ because the rank of $\mathrm{sl}(l+1, \mathbb{C})$ turns out to be l.)

For any j, $k \in \{1, \ldots, l+1\}$, let E_{jk} be the elementary $(l+1) \times (l+1)$ matrix whose (j,k)-entry is 1, and all of whose other entries are 0. Then

$$E_{jk}\, E_{rs} = \delta_{kr}\, E_{js} \tag{9.1}$$

and

$$[E_{jk}, E_{rs}] = \delta_{kr}\, E_{js} - \delta_{sj}\, E_{rk}. \tag{9.2}$$

Let $\mathfrak{h}$ be the subspace consisting of all diagonal matrices in $\mathfrak{g}$. Then $\mathfrak{h}$ has dimension l, and has basis $(E_{11} - E_{22}, \ldots, E_{ll} - E_{l+1,l+1})$. The Lie algebra $\mathfrak{g}$ itself has basis $\{E_{jj} - E_{j+1,j+1}\} \cup \{E_{jk}\}_{1 \leq j \neq k \leq l+1}$, suitably ordered.

We can therefore write

$$\mathrm{sl}\,(l+1, \mathbb{C}) \,=\, \mathfrak{h} \oplus \bigoplus_{j \neq k} \mathbb{C} E_{jk} \tag{9.3}$$

Let $H \in \mathfrak{h}$, with

$$h = \begin{pmatrix} h_1 & & 0 \\ & \ddots & \\ 0 & & h_{l+1} \end{pmatrix}$$

For any j, $1 \le j \le l+1$, let α_j be the linear functional on $\mathfrak{h}$ given by $\alpha_j(H) = h_j$. Then if $j \neq k$, it is easy to see (directly or from (9.2)) that for all $h \in \mathfrak{h}$, we have

$$[h, E_{jk}] = (h_j - h_k)\, E_{jk} = (\alpha_j - \alpha_k)(h)\, E_{jk}. \tag{9.4}$$

Since $\mathfrak{h}$ is abelian, (9.3) and (9.4) show that $\operatorname{ad} h$ is a semisimple linear operator on $\mathfrak{g}$, for all $H \in \mathfrak{h}$. Equation (9.4) also shows that $\mathfrak{h}$ is maximal abelian in $\mathfrak{g}$.

Therefore, if we can show that $\mathfrak{g}$ is simple, it will follow that $\mathfrak{h}$ is a Cartan subalgebra and (9.3) is the root space decomposition of $\mathfrak{g}$ with respect to $\mathfrak{h}$. The roots are given by

$$\alpha_{jk}(H) = (\alpha_j - \alpha_k)(H) \qquad\qquad (H \in \mathfrak{h}) \tag{9.5}$$

for all $1 \le j \neq k \le l+1$.

Note first that $\mathfrak{g}$ is not abelian, since $[E_{12}, E_{21}] = E_{11} - E_{22} \neq 0$. Let $\mathfrak{I}$ be a nonzero ideal of $\mathfrak{g}$. We wish to show that $\mathfrak{I} = \mathfrak{g}$.

First we claim that $\mathfrak{I} \cap \mathfrak{h} \neq \{0\}$. Let X be a nonzero element of $\mathfrak{I}$. If $X \in \mathfrak{h}$, there is noting to prove, so we can assume that $X \notin \mathfrak{h}$. Write $X = \sum_{j,k} a_{jk} E_{jk}$, with $a_{qp} \neq 0$ for some indices $q \neq p$.

Now $\mathfrak{I}$ contains the element $W = [E_{pp} - E_{qq}, X]$. Using (9.2), we see that

$$W = \sum_k a_{pk}\, E_{pk} - \sum_j a_{jp}\, E_{jp} - \sum_k a_{qk}\, E_{qk} + \sum_j a_{jq}\, E_{jq}$$

$\mathfrak{I}$ thus also contains

$$\begin{aligned}
[E_{pp} - E_{qq}, W] = {} & \sum_k a_{pk}\, E_{pk} - a_{pp}\, E_{pp} + a_{pq}\, E_{pq} - a_{pp}\, E_{pp} + \sum_j a_{jp}\, E_{jp} \\
& + a_{qp}\, E_{qp} + a_{qp}\, E_{qp} + \sum_k a_{qk}\, E_{qk} - a_{qq}\, E_{qq} + a_{pq}\, E_{pq} \\
& - a_{qq}\, E_{qq} + \sum_j a_{jq}\, E_{jq} \\
= {} & \sum_k a_{pk}\, E_{pk} + \sum_j a_{jp}\, E_{jp} + \sum_k a_{qk}\, E_{qk} + \sum_j a_{jq}\, E_{jq} \\
& - 2a_{pp}\, E_{pp} - 2a_{qq}\, E_{qq} + 2a_{pq}\, E_{pq} + 2a_{qp}\, E_{qp}
\end{aligned}$$

Hence $\mathfrak{I}$ contains

$$\begin{aligned}
Z &= W - [E_{pp} - E_{qq}, W] \\
&= -2 \sum_j a_{jp}\, E_{jp} - 2 \sum_k a_{qk}\, E_{qk} + 2a_{pp}\, E_{pp} + 2a_{qq}\, E_{qq} \\
&\qquad - 2a_{pq}\, E_{pq} - 2a_{qp}\, E_{qp}
\end{aligned}$$

$\mathfrak{I}$ will therefore contain

$$\begin{aligned}
Z' &= [Z, E_{pq}] \\
&= -2 \sum_j a_{jp}\, E_{jq} + 2a_{qp}\, E_{pp} - 2a_{qp}\, E_{qq} + 2 \sum_k a_{qk}\, E_{pk} + 2a_{pp}\, E_{pq} - 2a_{qq}\, E_{pq} \\
&\qquad - 2a_{qp}\, E_{qq} + 2a_{qp}\, E_{pp} \\
&= 2 \sum_j a_{jp}\, E_{jq} + 2 \sum_k a_{qk}\, E_{pk} + 4a_{qp}\, (E_{pp} - E_{qq}) + 2(a_{pp} - a_{qq})\, E_{pq}
\end{aligned}$$

Finally, $\mathfrak{I}$ contains

$$\begin{aligned}
Z'' &= [Z', E_{pq}] \\
&= -2a_{qp}\, E_{pq} + 2a_{qp}\, E_{pq} + 4a_{qp}\, E_{pq} + 4a_{qp}\, E_{pq} \\
&= 8a_{qp}\, E_{pq}
\end{aligned}$$

Thus $E_{pq} \in \mathfrak{I}$. Hence $E_{pp} - E_{qq} = [E_{pq}, E_{qp}] \in \mathfrak{I}$. This shows that $\mathfrak{h} \cap \mathfrak{I} \neq \{0\}$.

So let H be a nonzero vector in $\mathfrak{I} \cap \mathfrak{h}$. Write $H = \sum_j a_{jj}\, E_{jj}$. Since $\operatorname{tr} H = \sum_j a_{jj} = 0$, there exist p and q such that $a_{pp} \neq a_{qq}$. Then $\mathfrak{I}$ contains $E_{pq} = (a_{pp} - a_{qq})^{-1}\, [H, E_{pq}]$, and also $E_{pp} - E_{qq} = [E_{pq}, E_{qp}]$, as well as $E_{qp} = [E_{qq} - E_{pp}, E_{qp}]$.

Next, for all $j \neq p, q$, $\mathfrak{I}$ contains $E_{pj} = [E_{pp} - E_{qq}, E_{pj}]$ as well as $E_{jq} = [E_{jq}, E_{qq} - E_{pp}]$. Thus $\mathfrak{I}$ contains $E_{pp} - E_{jj}$ and $E_{jj} - E_{qq}$. Finally, if $k \neq j, q$, we see that $\mathfrak{I}$ contains $E_{jk} = [E_{jj} - E_{qq}, E_{jk}]$. Since $\mathfrak{I}$ also contains E_{jq}, we see that $\mathfrak{I}$ contains all vectors E_{jk}, for all $j \neq k$. It also thus contains the vectors $E_{jj} - E_{kk}$. Since all these vectors span $\mathfrak{g}$, we finally conclude that $\mathfrak{I} = \mathfrak{g}$. It follows that $\mathfrak{g}$ is simple. Since $\dim \mathfrak{h} = l$, we see that $\mathfrak{g}$ has rank l.

Let us now calculate the Killing form on $\mathfrak{g}$. For simplicity, let $n = l + 1$ and let $\mathrm{su}\,(n) = \mathrm{u}\,(n) \cap \mathrm{sl}\,(n, \mathbb{C}) = \{Y \in \mathrm{sl}\,(n, \mathbb{C}) \mid Y^* = -Y\}$. Any $X \in \mathrm{sl}\,(n, \mathbb{C})$ can be written

$$X = \frac{X - X^*}{2} + i\frac{X + X^*}{2i} \in \mathrm{su}\,(n) + i\mathrm{su}\,(n)$$

Moreover, since $i\mathrm{su}\,(n)$ consists of all Hermitian matrices of trace 0, we see that $\mathrm{su}\,(n) \cap i\mathrm{su}\,(n) = \{0\}$. Thus $\mathrm{sl}\,(n, \mathbb{C})$ equals the complexification $\mathrm{su}\,(n)^c$. (See Exercise 6.2.3.)

Suppose that $H \in \mathfrak{h} \cap i\mathrm{su}\,(n)$. Thus H is a diagonal matrix with real entries, and with trace 0. Then by (9.5),

$$\begin{aligned}
B(H,H) &= \sum_{j \neq k} (\alpha_j(H) - \alpha_k(H))^2 \\
&= \sum_{j,k} (\alpha_j(H) - \alpha_k(H))^2 \\
&= \sum_{j,k} \alpha_j(H)^2 + \sum_{j,k} \alpha_k(H)^2 - 2 \sum_{j,k} \alpha_j(H)\alpha_k(H) \\
&= n \sum_{j} \alpha_j(H)^2 + n \sum_{k} \alpha_k(H)^2 \\
&= 2n\,\mathrm{tr}\,(H^2) \\
&= 2(l+1)\,\mathrm{tr}\,(H^2)
\end{aligned}$$

Now suppose that $X \in i\,\mathrm{su}\,(n)$. Then X is a Hermitian matrix, and so there exists a (unitary) $n \times n$ matrix u such uXu^{-1} is a real diagonal matrix. Now the map $\varphi : Y \mapsto uYu^{-1}$ is an automorphism of $\mathrm{sl}\,(n, \mathbb{C})$. This automorphism preserves the Killing form (that is $B(\varphi(Y), \varphi(Y')) = B(Y, Y')$ for all Y, $Y' \in \mathrm{sl}\,(n, \mathbb{C})$). Hence

$$\begin{aligned}
B(X,X) &= B(uXu^{-1}, uXu^{-1}) \\
&= 2n\,\mathrm{tr}\,(uXu^{-1})^2 \\
&= 2n\,\mathrm{tr}\,(uX^2u^{-1}) \\
&= 2n\,\mathrm{tr}\,(X^2).
\end{aligned}$$

We can now polarize the above equation expression for B. Suppose that X and Y are matrices in $i\,\mathrm{su}\,(n)$. Then

$$\begin{aligned}
4\,B(X,Y) &= B(X+Y, X+Y) - B(X-Y, X-Y) \\
&= 2n\,\mathrm{tr}\,(X+Y)^2 - 2n\,\mathrm{tr}\,(X-Y)^2 \\
&= 4\,(2n)\,\mathrm{tr}\,(XY)
\end{aligned}$$

Hence $B(X,Y) = 2n\,\mathrm{tr}\,(XY)$. Now $\mathrm{sl}\,(n, \mathbb{C})$ is the complexification of the real vector space $i\,\mathrm{su}\,(n)$, and both $B(X,Y)$ and $\mathrm{tr}\,(XY)$ are $\mathbb{C}$-bilinear in X and Y. It follows that

$$B(X,Y) = 2n\,\mathrm{tr}\,(XY) \tag{9.6}$$

for all X, $Y \in \mathrm{sl}\,(n, \mathbb{C})$.

Chapter 10

Root Systems

Suppose that $\mathfrak{h}$ is a Cartan subalgebra of a complex semisimple Lie algebra $\mathfrak{g}$, and that Δ is the set of roots of $\mathfrak{g}$ relative to $\mathfrak{h}$. In Theorem 8.2.17, we saw the the Killing form B is an inner product on the real vector space $\mathfrak{h}_{\mathbb{R}} = \sum_{\alpha\in\Delta} \mathbb{R}\, h_\alpha$. The extension of B to the real dual space $\mathfrak{h}_{\mathbb{R}}^* = \sum_{\alpha\in\Delta} \mathbb{R}\,\alpha$, given in Definition 8.2.10 by $B(\alpha,\beta) = B(h_\alpha, h_\beta)$, is thus an inner product on $\mathfrak{h}_{\mathbb{R}}^*$.

In this chapter, we will study abstract root systems in inner product spaces. These are finite sets of vectors essentially satisfying conditions (ii) and (iii) in Theorem 8.2.13. Our objective is to classify such root systems. This will allow us to classify all finite-dimensional complex simple Lie algebras, up to isomorphism.

10.1 Abstract Root Systems

Throughout this chapter, we will assume that E is an inner product space, with inner product $(\ ,\ )$ and norm $\|\ \ \|$.

Definition 10.1.1. An isometry of E is a linear operator T on E which preserves the inner product: $(Tv, Tw) = (v, w)$ for all v, $w \in E$. We denote the group of all isometries on E by $\mathrm{O}(E)$. A *reflection* in E is an isometry of E whose fixed-point subspace has codimension one.

Remark 10.1.2. If $E = \mathbb{R}^n$ with the standard inner product, then any isometry on E corresponds to multiplication by a given $n\times n$ orthogonal matrix u: $x \mapsto ux$ for all $x \in \mathbb{R}^n$. The group of all isometries on $\mathbb{R}^n$ may be identified with the orthogonal group $\mathrm{O}(n)$ of all $n \times n$ orthogonal matrices.

Let R be a reflection on E, and let W be its fixed-point subspace. We choose a nonzero vector u in its orthogonal complement $W^\perp$. Since $\dim W^\perp = 1$, u is unique up to a scalar factor. Since R is an isometry, $Ru \in W^\perp$; since Ru is not fixed by R and has the same length as u, we have $Ru = -u$.

Any $v \in E$ may be written as $v = w + tu$, for unique $w \in W$ and $t \in \mathbb{R}$. Then

$$\begin{aligned} Rv &= w - tu \\ &= (w + tu) - 2t\,u \\ &= v - 2\,\frac{(v,u)}{(u,u)}\,u. \end{aligned} \tag{10.1}$$

Thus any reflection R may be expressed in the form (10.1), for some nonzero vector $u \in E$.

Conversely, given any nonzero vector $u \in E$, the map $R : v \mapsto v - (2(v,u)/(u,u))\,u$ clearly fixes the subspace $u^\perp = \{w \in E \,|\, (u,w) = 0\}$ and maps u to $-u$, hence is a reflection. (The fact that R preserves inner products is an immediate consequence of the Pythagorean identity $\|w + tu\|^2 = \|w\|^2 + \|tu\|^2$, for all $t \in \mathbb{R}$ and $w \in u^\perp$.)

It is obvious that any reflection R is its own inverse; that is $R^2 = I_E$. Thus, R is an *involution.* (There are, of course, involutions which are not reflections.)

Definition 10.1.3. A subset Δ of E is called a *root system* if it satisfies the following conditions:

(R1) Δ is finite, $0 \notin \Delta$, and Δ spans E.

(R2) If $\alpha \in \Delta$, then $-\alpha \in \Delta$.

(R3) For any α, $\beta \in \Delta$, the scalar $c(\beta, \alpha) = 2(\beta,\alpha)/(\alpha,\alpha)$ is an integer.

(R4) If $\alpha \in \Delta$, the reflection

$$r_\alpha \colon \beta \mapsto \beta - c(\beta,\alpha)\,\alpha \tag{10.2}$$

maps Δ into Δ.

The elements of Δ are called *roots.*

Suppose that $\alpha \in \Delta$. Then the integrality condition (R3) implies that, besides $\pm\alpha$, the only multiples of α which could also be roots are $\pm 2\,\alpha$ and $\pm\,\alpha/2$. In fact, if $c\alpha \in \Delta$, then $2c = (c\alpha, \alpha)/(\alpha,\alpha) \in \mathbb{Z}$ and $2/c = 2(\alpha, c\alpha)/(c\alpha, c\alpha) \in \mathbb{Z}$. Then $c = m/2 = 2/n$, for some integers m and n. If $|c| < 1$, then $n = \pm 4$, and if $|c| > 1$, then $n = \pm 1$. Of course, we could also have $m = \pm 2$, $n = \pm 2$. Hence $c = \pm 1/2$, ± 1, ± 2. Finally, $\alpha/2$ and 2α cannot *both* be roots; otherwise we would have $\alpha/2 = (1/4)\,(2\alpha)$, contrary to the fact that the only multiples of any root are half-integer multiples of that root.

Definition 10.1.4. A root system Δ is said to be *reduced* if it satisfies the condition

(R5) If $\alpha \in \Delta$, the only multiples of α which are roots are $\pm\alpha$.

Given a root system Δ, we put

$$\Delta_i = \{\alpha \in \Delta \,|\, \alpha/2 \notin \Delta\} \tag{10.3}$$
$$\Delta_u = \{\alpha \in \Delta \,|\, 2\alpha \notin \Delta\} \tag{10.4}$$

It is straightforward to check that Δ_i and Δ_u are root systems in their own right, and they are obviously both reduced. Δ_i is called the set of *indivisible roots* in Δ, and Δ_u is called the set of *unmultipliable roots* in Δ.

Example 10.1.5. Let $\mathfrak{g}$ be a complex semisimple Lie algebra, let $\mathfrak{h}$ be a Cartan subalgebra, and let Δ be the set of roots of $\mathfrak{g}$ relative to $\mathfrak{h}$. Then the Killing form B is an inner product on the real vector space $E = \mathfrak{h}^*_{\mathbb{R}}$, and Δ is a root system in E: Condition (R1) follows from the definition of a root in $\mathfrak{g}$ and Theorem 8.2.11(a). Condition (R2) follows from Theorem 8.2.11(b). The integrality condition (R3) follows from Theorem 8.2.13(ii), and condition (R4) is from statement (iii) of the same theorem. Finally, Theorem 8.2.12 shows that Δ is reduced.

Definition 10.1.6. For any $\alpha \in \Delta$, the orthogonal complement $\pi_\alpha := \alpha^\perp$ is called a *root hyperplane*. The reflection r_α given by (10.2) is called the *Weyl reflection in the hyperplane π_α*. The *Weyl group* W is the subgroup of $\mathrm{O}(E)$ generated by all r_α, $\alpha \in \Delta$.

Since the reflection r_α is its own inverse, r_α actually maps Δ *onto* Δ. Thus r_α permutes Δ. It follows that every element $\sigma \in W$ permutes Δ. Now since Δ spans E, every $\sigma \in W$ is completely detemined (as a linear operator on E) by its restriction to Δ. The restriction map $\sigma \mapsto \sigma|_\Delta$ is thus an injective group homomorphism from W into the group $\mathfrak{S}(\Delta)$ of all permutations of the finite set Δ. We conclude from this that W is a *finite* subgroup of $\mathrm{O}(E)$.

Theorem 10.1.7. *Suppose that α and β are linearly independent elements of Δ. Then*

1. $c(\beta,\alpha)\, c(\alpha,\beta) = 0,\ 1,\ 2,$ *or* 3.
2. *If* $(\alpha,\beta) = 0$*, then* $c(\beta,\alpha) = c(\alpha,\beta) = 0$.
3. *If* $(\alpha,\beta) > 0$ *and* $\|\alpha\| \le \|\beta\|$*, then* $c(\alpha,\beta) = 1$ *and* $c(\beta,\alpha) = (\beta,\beta)/(\alpha,\alpha) = 1,\ 2,$ *or* 3.
4. *If* $(\alpha,\beta) < 0$ *and* $\|\alpha\| \le \|\beta\|$*, then* $c(\alpha,\beta) = -1$ *and* $c(\beta,\alpha) = -(\beta,\beta)/(\alpha,\alpha) = -1,\ -2,$ *or* -3.

5. *If $(\alpha, \beta) > 0$ and $\|\alpha\| \leq \|\beta\|$, the $-2(\alpha, \beta) + (\beta, \beta) = 0$. If $(\alpha, \beta) < 0$ and $\|\alpha\| \leq \|\beta\|$, then $2(\alpha, \beta) + (\beta, \beta) = 0$.*

Proof. 1. Since α and β are linearly independent, the Cauchy-Schwartz Inequality implies that

$$(\alpha, \beta)^2 < (\alpha, \alpha)\,(\beta, \beta).$$

Hence

$$c(\alpha, \beta)\, c(\beta, \alpha) = 4\, \frac{(\alpha, \beta)^2}{(\alpha, \alpha)\,(\beta, \beta)} \;<\; 4$$

Since $c(\alpha, \beta)$ and $c(\beta, \alpha)$ are integers, we must have $c(\alpha, \beta)\, c(\beta, \alpha) = 0,\ 1,\ 2,\ \text{or}\ 3$.

2. Obvious.

3. If $(\alpha, \beta) > 0$ and $\|\alpha\| \leq \|\beta\|$, then

$$\frac{2\,(\alpha, \beta)}{(\beta, \beta)} \;<\; \frac{2\,(\beta, \alpha)}{(\alpha, \alpha)}$$

So $c(\alpha, \beta) < c(\beta, \alpha)$. By Part (1), $c(\alpha, \beta)\, c(\beta, \alpha) = 1,\ 2,\ \text{or}\ 3$. Since $c(\alpha, \beta)$ is the smaller integer factor, we must have $c(\alpha, \beta) = 1$, whence $2(\alpha, \beta) = (\beta, \beta)$. Thus $c(\beta, \alpha) = 2(\beta, \alpha)/(\alpha, \alpha) = (\beta, \beta)/(\alpha, \alpha)$.

4. Replace α by $-\alpha$ and apply Part (3) above.

5. Suppose that $(\alpha, \beta) > 0$ and $\|\alpha\| \leq \|\beta\|$. Then by Part (3), $c(\alpha, \beta) = 1$, so $-2\,(\alpha, \beta) + (\beta, \beta) = 0$. On the other hand, if $(\alpha, \beta) < 0$ and $\|\alpha\| \leq \|\beta\|$, then by Part (4), $c(\alpha, \beta) = -1$, whence $2(\alpha, \beta) + (\beta, \beta) = 0$.

□

Theorem 10.1.8. *Suppose that α and β are linearly independent roots and $\|\alpha\| \leq \|\beta\|$. Let θ be the angle between α and β. Then we have the following table:*

$c(\alpha, \beta)\, c(\beta, \alpha)$	$c(\alpha, \beta)$	$c(\beta, \alpha)$	$\|\beta\|^2/\|\alpha\|^2$	$\cos\theta$	θ
0	0	0	?	0	$\pi/2$
1	1	1	1	$1/2$	$\pi/3$
1	-1	-1	1	$-1/2$	$2\pi/3$
2	1	2	2	$\sqrt{2}/2$	$\pi/4$
2	-1	-2	2	$-\sqrt{2}/2$	$3\pi/4$
3	1	3	3	$\sqrt{3}/2$	$\pi/6$
3	-1	-3	3	$-\sqrt{3}/2$	$5\pi/6$

Proof. Here $\cos\theta = (\alpha,\beta)/(\|\alpha\|\,\|\beta\|)$, so $|\cos\theta| = \sqrt{c(\alpha,\beta)\,c(\beta,\alpha)}/2$. Now by Theorem 10.1.7. $c(\alpha,\beta)\,c(\beta,\alpha) = 0$, 1, 2, or 3 and $\|\beta\|^2/\|\alpha\|^2 = |c(\beta,\alpha)|$. Moreover, $\cos\theta$ has the same sign as $c(\alpha,\beta)$. This gives us the table above. □

Lemma 10.1.9. *Suppose that α and β are linearly independent roots. If $(\alpha,\beta) < 0$, then $\alpha+\beta$ is a root. If $(\alpha,\beta) > 0$, then $\alpha-\beta$ is a root.*

Proof. We can assume that $\|\alpha\| \leq \|\beta\|$. If $(\alpha,\beta) < 0$, it follows from Theorem 10.1.7 that $c(\alpha,\beta) = -1$. Hence, by Condition (R4) in Definition 10.1.3, $\alpha - c(\alpha,\beta)\,\beta = \alpha+\beta$ is a root.

If $(\alpha,\beta) > 0$, then we can replace β by $-\beta$ in the argument above to conclude that $\alpha+\beta$ is a root. □

Theorem 10.1.10. *Suppose that α and β are linearly independent elements of Δ. Let q be the largest integer j such that $\beta+j\alpha \in \Delta$, and let p be the smallest integer j such that $\beta+j\alpha\in\Delta$. Then for any integer j such that $p \leq j \leq q$, we have $\beta+j\alpha\in\Delta$. Moreover $c(\beta,\alpha) = -(p+q)$.*

Remark: While the conclusion of the theorem above is similar to that of Theorem 8.2.13, it requires a different proof because, in the present abstract setting, Δ is not necessarily the root system of a complex semisimple Lie algebra $\mathfrak{g}$ relative to a Cartan subalgebra $\mathfrak{h}$. In particular, we cannot use the representation theory of $\mathrm{sl}\,(2,\mathbb{C})$ as we had done to prove Theorem 8.2.13.

As in Chapter 8, we call the set $\{\beta+j\alpha \,|\, p\leq j\leq q\}$ the *α-string through β.*

Proof. Suppose that there is an integer j_0, $p < j_0 < q$, such that $\beta+j_0\alpha\notin\Delta$. Let $j_1 = \max\{j < j_0 \,|\, \beta+j\alpha\in\Delta\}$ and $j_2 = \min\{j > j_0\,|\,\beta+j\alpha\in\Delta\}$. Clearly, $j_1 \geq p$ and $j_2 \leq q$. Liekwise, it is clear that $\beta+j_1\alpha\in\Delta$ but $\beta+(j_1+1)\alpha\notin\Delta$. By Lemma 10.1.9, $(\beta+j_1\alpha,\alpha)\geq 0$. Likewise, $\beta+j_2\alpha\in\Delta$ but $\beta+(j_2-1)\alpha\notin\Delta$. Hence, again by Lemma 10.1.9, $(\beta+j_2\alpha,\alpha)\leq 0$. We conclude that

$$(\beta,\alpha)+j_2\,(\alpha,\alpha) \leq (\beta,\alpha)+j_1\,(\alpha,\alpha),$$

which contradicts the fact that $j_1 < j_2$. This shows that $\{\beta+j\alpha\,|\,p\leq j\leq q\}\subset\Delta$.

For each integer j such that $p\leq j\leq q$, let $a_j = \beta+j\alpha$. Then $r_\alpha(a_j) = a_{-j-c(\beta,\alpha)}$. Thus r_α reverses the sequence $a_p, a_{p+1},\dots,a_{q-1},a_q$. In particular $r_\alpha(a_p) = a_q$, and hence

$$\begin{aligned} r_\alpha(\beta+p\alpha) = \beta+q\,\alpha &\implies (\beta - c(\beta,\alpha)\,\alpha) - p\,\alpha = \beta+q\,\alpha \\ &\implies c(\beta,\alpha) = -(p+q) \end{aligned}$$

□

For further insights into the nature of the root system Δ and its Weyl group W, we will need to introduce an ordering on the inner product space E. So fix a basis B of E, and let $>$ be the lexicographic ordering on E relative to B. Let Δ^+ be the set of positive roots with respect to this ordering. Then $\Delta = \Delta^+ \cup (-\Delta^+)$.

Lemma 10.1.11. *Let $\{v_1, \dots, v_m\}$ be a set of positive vectors in an ordered real inner product space E such that $(v_i, v_j) \leq 0$ for all $i \neq j$. Then the set $\{v_1, \dots, v_m\}$ is linearly independent.*

Proof. Suppose instead that the vectors are linearly dependent. Let $c_1 v_1 + \cdots + c_m v_m = 0$ be a linear dependence relation *with the fewest possible nonzero coefficients*. By reordering the vectors if necessary, we can assume that the coefficients $c_1, \dots, c_k$ are nonzero and the rest zero. Note that $2 \leq k \leq m$. Note also that $c_1, \dots, c_k$ cannot all be of the same sign, since the sum of positive vectors is positive, and the product of a positive scalar with a positive vector is a positive vector. By reordering $v_1, \dots, v_k$ again, we can assume that $c_1, \dots, c_r$ are positive (where $1 \leq r < k$), and $c_{r+1}, \dots, c_k$ are negative. This results in

$$\sum_{i=1}^{r} c_i v_i = \sum_{j=r+1}^{k} (-c_j) v_j$$

Let w be the vector represented by both sides of the above equation. Then $w \neq 0$, for in fact, w is a positive vector. However, from the above, we obtain

$$(w, w) = -\sum_{1 \leq i \leq r} \sum_{r < j \leq k} c_i c_j (v_i, v_j) \leq 0,$$

But $(w, w) > 0$, which gives us a contradiction. ∎

Definition 10.1.12. A root $\alpha \in \Delta$ is said to be *simple* if α is positive and α is not the sum of two positive roots. The collection Φ of all simple roots is called a *simple system of roots*.

The simple system Φ depends, of course, on the ordering chosen on E. Note that any simple root must be indivisible. Thus Φ is a subset of the system Δ_i of indivisible roots.

Theorem 10.1.13. *Let Φ be a simple system of roots, and put $\Phi = \{\alpha_1, \dots, \alpha_l\}$. Then*

1. *$\alpha_i - \alpha_j \notin \Phi$.*
2. *If $i \neq j$, then $(\alpha_i, \alpha_j) \leq 0$.*

3. *Every positive root $\gamma \in \Delta^+$ is a linear combination of simple roots*

$$\gamma = \sum_{i=1}^{l} n_i \, \alpha_i \tag{10.5}$$

where each $n_i \in \mathbb{Z}^+$. Every negative root γ is given by a linear combination of the form (10.5)*, where each $n_i \in -\mathbb{Z}^+$.*

4. *$(\alpha_1, \ldots, \alpha_l)$ is a basis of E.*

5. *If γ is a positive root, and $\gamma \notin \Phi$, then there is a simple root α_i such that $\gamma - \alpha_i$ is a positive root.*

Proof. 1. We can assume that $i \neq j$. Suppose that $\beta = \alpha_i - \alpha_j \in \Delta$. If $\beta \in \Delta^+$, then $\alpha_i = \alpha_j + \beta$, contradicting the fact that α_i is simple. If $\beta \in -\Delta^+$, then $\alpha_j = \alpha_i + (-\beta)$, contradicting the fact that α_j is simple.

2. If $(\alpha_i, \alpha_j) > 0$, then by Lemma 10.1.9, $\alpha_i - \alpha_j$ would be a root, contradicting Part 1 above.

3. Assume that $\gamma \in \Delta^+$. If γ is already simple, there is nothing to prove. So assume that γ is not simple. Then $\gamma = \rho + \mu$, where ρ and μ are positive roots. If ρ and μ are both simple, then we're done. If either ρ or μ is not simple, we can break it down further into the sum of positive roots. If either of those summands is not simple, decompose it further, and so forth.

   This procedure can be viewed as a "branching process" where each positive root β which is not simple has two branches below it, corresponding to two positive roots whose sum is β. This process produces a tree with γ as the topmost vertex, and where each downward path from γ must end since there are only a finite number of positive roots, and a positive root cannot recur along that path. The lowest vertices correspond to roots which can no longer be decomposed; i.e., simple roots. Clearly, γ is the sum of all these simple roots, and so

$$\gamma = \sum_{i=1}^{l} n_i \, \alpha_i$$

   where $n_i \in \mathbb{Z}^+$ for all i.

   If γ is negative, then $-\gamma$ is a linear combination of the simple roots α_i with coefficients which are all nonnegative integers.

4. Since Δ spans E, Part 3 above shows that Φ spans E. Then Part 1 and Lemma 10.1.11 show that the simple roots are linearly independent.

5. Suppose that $\gamma \in \Delta^+ \setminus \Phi$. Then we have $\gamma = \sum_i n_i\,\alpha_i$, where $n_i \in \mathbb{Z}^+$ for all i and at least one n_i is positive. Hence

$$\begin{aligned} 0 &< (\gamma, \gamma) \\ &= (\gamma, \sum_i n_i \alpha_i) \\ &= \sum_i n_i\,(\gamma, \alpha_i) \end{aligned}$$

Thus $n_i > 0$ and $(\gamma, \alpha_i) > 0$ for some i, and hence by Lemma 10.1.9, $\gamma - \alpha_i$ must be a root. This root must be positive, since it is a linear combination of simple roots with nonnegative integer coefficients.

□

Here's converse to Theorem 10.1.13, Part 3.

Proposition 10.1.14. *If Φ is a set of roots having $\dim E$ elements for which Theorem 10.1.13, Part 3 holds, then Φ is a simple system of roots relative to an appropriate lexicographic order on E.*

Proof. Let $\Phi = \{\alpha_1, \ldots, \alpha_l\}$ (where $l = \dim E$), and let $>$ be the lexicographic order relative to the basis $(\alpha_1, \ldots, \alpha_l)$ of E. Then a root γ is positive if and only if it is a linear combination of the α_i with nonnegative integer coefficients. It is also clear from that equation 10.5 that each α_i is a simple root. Now, according to Theorem 10.1.13, Part 4, the set of simple roots corresponding to any lexicographic order on E has l elements. Thus Φ is a simple system of roots. □

Lemma 10.1.15. *Let $\Phi = \{\alpha_1, \ldots, \alpha_l\}$ be a simple system of roots. Then every positive root γ can be written as $\gamma = \alpha_{i_1} + \alpha_{i_2} + \cdots + \alpha_{i_k}$, where each initial partial sum $\alpha_{i_1} + \cdots + \alpha_{i_j}$ $(1 \le j \le k)$ is a root.*

Proof. For every positive root $\gamma = \sum_{i=1}^{l} n_i\,\alpha_i$, the *height* of γ is defined to be the positive number $\operatorname{ht}\gamma = \sum_{i=1}^{l} n_i$. We prove this lemma by induction on $\operatorname{ht}\gamma$. If $\operatorname{ht}\gamma = 1$, then γ is simple and there is nothing to prove. So assume that $m > 1$ and that the lemma's conclusion holds for all positive roots of height $< m$. Now suppose that γ is a positive root of height m. According to Theorem 10.1.13, Part 5, $\gamma - \alpha_i$ is a positive root for some i. Now apply the induction hypothesis to the root $\gamma - \alpha_i$, which has height $m - 1$. Then $\gamma - \alpha_i = \alpha_{i_1} + \cdots + \alpha_{i_{m-1}}$, where each initial partial sum is a root. Then $\gamma = \alpha_{i_1} + \cdots + \alpha_{i_{m-1}} + \alpha_i$. Thus γ satisfies the conclusion of the lemma, completing the induction step as well as the proof. □

Definition 10.1.16. The root system Δ is *decomposable* if Δ is a union $\Delta = \Delta_1 \cup \Delta_2$, with $\Delta_1 \neq \emptyset$, $\Delta_2 \neq \emptyset$ and $\Delta_1 \perp \Delta_2$.

Note that Δ_1 and Δ_2 must be disjoint, since $\Delta_1 \perp \Delta_2$.

Definition 10.1.17. If Φ is a simple system of roots in Δ, we say that Φ is *decomposable* if Φ is a union $\Phi = \Phi_1 \cup \Phi_2$, with $\Phi_1 \neq \emptyset$, $\Phi_2 \neq \emptyset$ and $\Phi_1 \perp \Phi_2$.

Again, we see that Φ_1 and Φ_2 must be disjoint. The two notions of decomposability given above are compatible, as the following proposition shows. (This is not a trivial fact!)

Proposition 10.1.18. *Let Φ be a simple system of roots in Δ. Then Δ is decomposable if and only if Φ is decomposable.*

Proof. Suppose that Δ is decomposable, with $\Delta = \Delta_1 \cup \Delta_2$. For $i = 1, 2$, let $\Phi_i = \Delta_i \cap \Phi$. Then neither Φ_1 nor Φ_2 can be empty. For if, say $\Phi_1 = \emptyset$, then $\Phi_2 = \Phi$, which implies that $\Delta_1 \perp \Phi$. Since Φ is a basis of E, we conclude that $\Delta_1 \perp \Delta$, and so $\Delta_1 = \emptyset$, contradiction.

Conversely, suppose that Φ is decomposable, with $\Phi = \Phi_1 \cup \Phi_2$. We arrange the elements of Φ so that $\Phi_1 = \{\alpha_1, \ldots, \alpha_r\}$ and $\Phi_2 = \{\alpha_{r+1}, \ldots, \alpha_l\}$. Now let $\gamma \in \Delta$. We claim that γ is a linear combination of elements of Φ_1 or γ is a linear combination of elements of Φ_2. To prove this claim, we may assume that γ is positive. Now suppose, to the contrary, that γ is a linear combination

$$\gamma = \sum_{i=1}^{r} n_i\,\alpha_i \; + \sum_{i=r+1}^{l} n_i\,\alpha_i \qquad\qquad (n_i \geq 0)$$

where both sums on the right are nonzero. According to Lemma 10.1.15, we can rewrite the sum above as $\gamma = \alpha_{i_1} + \cdots + \alpha_{i_k}$, where each initial partial sum is a root. Without loss of generality, we can assume that $\alpha_{i_1} \in \Phi_1$. Let s be the smallest integer such that $\alpha_{i_s} \in \Phi_2$. Then $\beta = \alpha_{i_1} + \cdots + \alpha_{i_{s-1}} + \alpha_{i_s}$ is a root. Now consider the root $r_{\alpha_{i_s}}\beta$. This root equals

$$r_{\alpha_{i_s}}(\alpha_{i_1} + \cdots + \alpha_{i_{s-1}} + \alpha_{i_s}) = \alpha_{i_1} + \cdots + \alpha_{i_{s-1}} - \alpha_{i_s},$$

which is not a linear combination of simple roots with nonnegative integer coefficients, a contradiction. This proves the claim.

Using the claim, we now let Δ_1 be the set of roots which are linear combinations of elements of Φ_1, and let Δ_2 be the set of roots which are linear combinations of elements of Φ_2. Then $\Delta_1 \neq \emptyset$, $\Delta_2 \neq \emptyset$, $\Delta_1 \perp \Delta_2$, and $\Delta = \Delta_1 \cup \Delta_2$. Thus Δ is decomposable. $\square$

The following theorem shows that *indecomposable* root systems correspond to complex simple Lie algebras, and vice versa.

Theorem 10.1.19. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra, let $\mathfrak{h}$ be a Cartan subalgebra, and let Δ be the set of roots of $\mathfrak{g}$ relative to $\mathfrak{h}$. Then Δ is indecomposable if and only if $\mathfrak{g}$ is simple.*

Proof. Suppose first that Δ is decomposable. We will prove that $\mathfrak{g}$ is not simple. Indeed, we have $\Delta = \Delta_1 \cup \Delta_2$, with $\Delta_1 \neq \emptyset$, $\Delta_2 \neq \emptyset$ and $\Delta_1 \perp \Delta_2$. For $i = 1, 2$, let

$$\mathfrak{g}_i = \sum_{\alpha \in \Delta_i} \mathbb{C}\, h_\alpha + \sum_{\alpha \in \Delta_i} \mathfrak{g}_\alpha$$

Let Φ be any simple system of roots in Δ. Then Φ is decomposable, and for $i = 1, 2$, let $\Phi_i = \Phi \cap \Delta_i$. For $i = 1, 2$, let $\mathfrak{h}_i = \sum_{\alpha \in \Delta_i} \mathbb{C}\, h_\alpha$. Then clearly $\mathfrak{h}_i = \sum_{\alpha \in \Phi_i} \mathbb{C}\, h_\alpha$. Since the vectors h_α, $\alpha \in \Phi$ form a basis of $\mathfrak{h}$, we conclude that $\mathfrak{h} = \mathfrak{h}_1 \oplus \mathfrak{h}_2$.

Next observe that $\alpha \pm \beta$ is never a root whenever $\alpha \in \Delta_1$ and $\beta \in \Delta_2$. This implies that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \{0\}$ for any $\alpha \in \Delta_1$ and $\beta \in \Delta_2$. Moreover, $\beta(\mathfrak{h}_1) = \{0\}$ for all $\beta \in \Delta_2$ and $\alpha(\mathfrak{h}_2) = \{0\}$ for all $\alpha \in \Delta_1$. From all this, we conclude that $[\mathfrak{g}_1, \mathfrak{g}_2] = \{0\}$.

The root space decomposition (8.9) of $\mathfrak{g}$ shows that

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2.$$

But both $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are also clearly subalgebras of $\mathfrak{g}$. Hence $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are nonzero ideals of $\mathfrak{g}$, and so $\mathfrak{g}$ is not simple.

Conversely, suppose that $\mathfrak{g}$ is not simple. Then by Theorem 7.1.1, $\mathfrak{g}$ is the direct sum of two nonzero semisimple ideals $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$. Let $\mathfrak{h}_1'$ and $\mathfrak{h}_2'$ be Cartan subalgebras of $\mathfrak{g}_1$ and $\mathfrak{g}_2$, respectively, and let Δ_1' and Δ_2' be the corresponding respective root systems. From the root space decompositions of $\mathfrak{g}_1$ and $\mathfrak{g}_2$, it is clear that $\mathfrak{h}' = \mathfrak{h}_1' \oplus \mathfrak{h}_2'$ is a Cartan subalgebra of $\mathfrak{g}$, and that the system of roots of $\mathfrak{g}$ relative to $\mathfrak{h}'$ is $\Delta' = \Delta_1' \cup \Delta_2'$.

Now suppose that $\beta \in \Delta_2'$ and $h \in \mathfrak{h}_1'$. Then if $0 \neq x_\beta \in (\mathfrak{g}_1)_\beta$, we have $0 = [h, x_\beta] = \beta(h)\, x_\beta$. Thus $\beta(h) = 0$. This implies that $B(\beta, \alpha) = B(h_\beta, h_\alpha) = \beta(h_\alpha) = 0$ for all $\alpha \in \Delta_1'$. Hence $\Delta_1' \perp \Delta_2'$.

Now by Theorem 8.1.9, there exists an automorphism φ of $\mathfrak{g}$ which maps $\mathfrak{h}$ onto $\mathfrak{h}'$. By hypothesis, Δ is the system of roots of $\mathfrak{g}$ relative to $\mathfrak{h}$. Suppose that $\alpha \in \Delta'$. Then for any $0 \neq x \in \mathfrak{g}_\alpha$ and $h \in \mathfrak{h}$, we have

$$\begin{aligned}[h, \varphi^{-1}(x)] &= \varphi^{-1}([\varphi(h), x] \\ &= \alpha(\varphi(h))\, \varphi^{-1}(x).\end{aligned}$$

This shows that $\alpha \circ \varphi = {}^t\varphi(\alpha) \in \Delta$. Hence ${}^t\varphi(\Delta') \subset \Delta$. Replacing φ by φ^{-1}, we also obtain ${}^t(\varphi^{-1})(\Delta) = ({}^t\varphi)^{-1}(\Delta) \subset \Delta'$. Hence ${}^t\varphi(\Delta') = \Delta$.

Finally, suppose that $\alpha \in \Delta'$. Let h'_α be the vector in $\mathfrak{h}'$ such that $B(h'_\alpha, h') = \alpha(h')$ for all $h' \in \mathfrak{h}'$. Then, since B is preserved under the automorphism φ, we have

$$\begin{aligned} B(\varphi^{-1}(h'_\alpha), h) &= B(h'_\alpha, \varphi(h)) \\ &= \alpha(\varphi(h)) \\ &= {}^t\varphi(\alpha)(h) \\ &= B(h_{{}^t\varphi(\alpha)}, h) \end{aligned}$$

for all $h \in \mathfrak{h}$. Thus $\varphi^{-1}(h'_\alpha) = h_{{}^t\varphi(\alpha)}$. Therefore, for all α, $\beta \in \Delta'$, we have

$$\begin{aligned} B(\alpha, \beta) &= B(h'_\alpha, h'_\beta) \\ &= B(\varphi^{-1}(h'_\alpha), \varphi^{-1}(h'_\beta)) \\ &= B(h_{{}^t\varphi(\alpha)}, h_{{}^t\varphi(\beta)}) \\ &= B({}^t\varphi(\alpha), {}^t\varphi(\beta)) \end{aligned}$$

Hence the map ${}^t\varphi : \Delta' \to \Delta$ preserves B. Put $\Delta_1 = {}^t\varphi(\Delta'_1)$ and $\Delta_2 = {}^t\varphi(\Delta'_2)$. We obtain $\Delta = \Delta_1 \cup \Delta_2$, $\Delta_1 \neq \emptyset$, $\Delta_2 \neq \emptyset$, and $\Delta_1 \perp \Delta_2$. This shows that Δ is decomposable. $\square$

Example 10.1.20. Let E be a two-dimensional inner product space. We will show that, up to isometry, there are only three possible indecomposable simple systems of roots Φ on E. Suppose that $\Phi = \{\alpha_1, \alpha_2\}$. Then $(\alpha_1, \alpha_2) \neq 0$, since Φ is indecomposable. By Theorem 10.1.13 Part 2, we have $(\alpha_1, \alpha_2) < 0$. We may assume that $\|\alpha_1\| \leq \|\alpha_2\|$. Then by Theorem 10.1.7 Part 4, $c(\alpha_1, \alpha_2) = -1$ and $c(\alpha_2, \alpha_1) = -1$, -2, or -3.

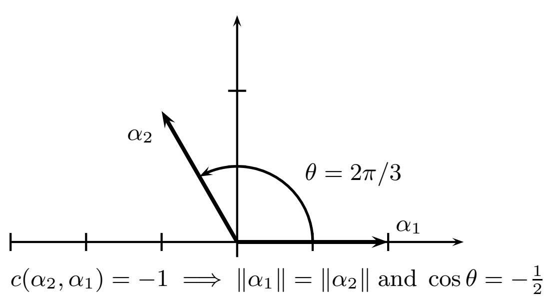
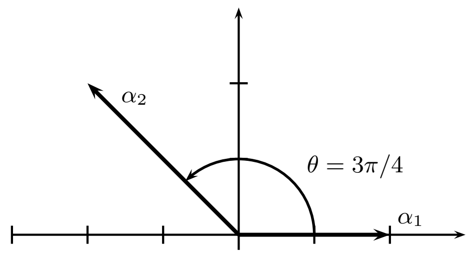
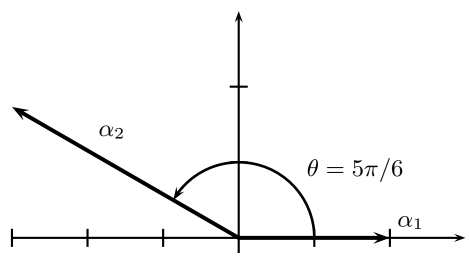


Figure 1



$c(\alpha_2, \alpha_1) = -2 \implies \|\alpha_2\| = \sqrt{2}\,\|\alpha_1\|$ and $\cos\theta = -\frac{\sqrt{2}}{2}$

Figure 2



$c(\alpha_2, \alpha_1) = -3 \implies \|\alpha_2\| = \sqrt{3}\,\|\alpha_1\|$ and $\cos\theta = -\frac{\sqrt{3}}{2}$

Figure 3

Let $\Phi = \{\alpha_1, \ldots, \alpha_l\}$ be a simple system of roots in Δ. We introduce a partial ordering $\prec$ on Δ as follows: if α, $\beta \in \Delta$, then $\alpha \prec \beta$ if and only if $\beta - \alpha = \sum_i n_i\alpha_i$ where each $n_i \in \mathbb{Z}^+$ and at least one n_i is positive. It is clear that $\prec$ is indeed a partial order on Δ. Of course, $\prec$ depends on the choice of Φ.

Recall that the simple system Φ was obtained via a lexicographic order $<$ on E. Since each simple root α_i is a positive root under $<$, it is clear that if α and β are roots such that $\alpha \prec \beta$, then $\alpha < \beta$. The converse is not true, as there are vectors in Δ which are not comparable under $\prec$.

Theorem 10.1.21. *(The Highest Root) Suppose that Δ is indecomposable. Then there is a unique element $\delta \in \Delta$ which is maximal under $\prec$. This root δ satisfies the following properties:*

1. *$\alpha \prec \delta$ for all $\alpha \in \Delta$, $\alpha \neq \delta$.*
2. *$\delta = \sum_{i=1}^{l} n_i\alpha_i$ where each n_i is a positive integer.*
3. *$(\delta, \alpha_i) \geq 0$ for all i.*

Proof. It is obvious that for each $\alpha \in \Delta$, there is an element γ, maximal with respect to $\prec$, such that $\alpha \prec \gamma$. Any such γ is clearly a positive root. Now let δ be a maximal element with respect to $\prec$. For each i, $\delta + \alpha_i$ is not a root, so by Lemma 10.1.9, $(\delta, \alpha_i) \geq 0$. This proves assertion 3. Since δ is positive, there is at least one i such that $(\delta, \alpha_i) > 0$.

By Theorem 10.1.13, we have $\delta = \sum_{i=1}^{l} n_i\, \alpha_i$, with $n_i \in \mathbb{Z}^+$. We claim that all n_i are positive integers. If not all n_i are positive, then the set $S = \{\alpha_i \,|\, n_i = 0\}$ is a nonempty, proper subset of Φ. Let $S' = \Phi \setminus S$. Then $\delta = \sum_{\alpha_j \in S'} n_j \alpha_j$. By Theorem 10.1.13 Part 2, we see that for any $\alpha_i \in S$,

$$(\delta, \alpha_i) = \sum_{\alpha_j \in S'} n_j\, (\alpha_j, \alpha_i) \leq 0.$$

Since we also have $(\delta, \alpha_i) \geq 0$, it follows that $(\delta, \alpha_i) = 0$ and that $(\alpha_j, \alpha_i) = 0$ for all $\alpha_j \in S'$ and all $\alpha_i \in S$. This shows that $S \perp S'$ and so Φ is decomposable, with $\Phi = S \cup S'$, a contradiction. This contradiction proves assertion 2.

Suppose that δ' is maximal element $\neq \delta$. Then by the above $\delta' = \sum_i m_i \alpha_i$, where each m_i is a positive integer. Since $(\delta, \alpha_i) \geq 0$ for all i and > 0 for at least one i,

$$(\delta, \delta') = \sum_{i=1}^{l} m_i (\delta, \alpha_i) > 0.$$

It follows that $\delta - \delta' \in \Delta$. If $\delta - \delta'$ is a positive root, then $\delta' \prec \delta$, contradicting the fact that δ' is maximal. If $\delta - \delta'$ is a negative root, then $\delta \prec \delta'$, contradicting the fact that δ is maximal. This shows that there can only be one maximal element.

Finally, since each $\alpha \in \Delta$ is $\prec$ some maximal element, and there is only one maximal element, we must have $\alpha \prec \delta$, proving assertion 1. $\square$

The theorem above is a special case (just as Theorem 7.3.11) of the *theorem of the highest weight*, a theorem about finite-dimensional irreducible representations of complex semsimple Lie algebras.

10.2 Cartan Matrices and Dynkin Diagrams

Let Δ be a root system on an inner product space E, let $>$ be a lexicographic order on E with respect to some basis of E, and let Φ be a simple system of roots corresponding to $>$. Assume that Φ is indecomposable. We list the elements of Φ as $\alpha_1, \ldots, \alpha_l$. The *Cartan matrix* corresponding to this listing of Φ is the $l \times l$ matrix $(c(\alpha_i, \alpha_j))_{1 \leq i,j \leq l}$. Of course, any other listing of the elements of Φ results in a Cartan matrix obtained from this one by applying a permutation

σ to its rows and columns. If we fix a listing $(\alpha_1, \ldots, \alpha_l)$ of Φ, we will abuse terminology and call the matrix $(c(\alpha_i, \alpha_j))$ the *Cartan matrix of* Φ.

The Cartan matrix $(c(\alpha_i, \alpha_j))$ satisfies the following properties:

1. It is nonsingular.
2. The diagonal entries are all equal to 2.
3. The off-diagonal entries have values $0, -1, -2$, or -3. If $c(\alpha_i, \alpha_j) = -2$ or -3, then $c(\alpha_j, \alpha_i) = -1$.

For example, suppose that $\dim E = 2$ and $\Phi = \{\alpha_1, \alpha_2\}$ with $\|\alpha_1\| \leq \|\alpha_2\|$. Then according to Example 10.1.20, the only possible Cartan matrices $(c(\alpha_i, \alpha_j))$ are:

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$

which corresponds to Figure 1;

$$\begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix},$$

which corresponds to Figure 2; and

$$\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix},$$

which corresponds to Figure 3.

If Δ is reduced, the Cartan matrix of Φ completely determines the root system Δ. The procedure for recovering the set of positive roots Δ^+ from Φ is carried out inductively on the height of each root, as follows. First, the roots of height one are precisely the elements of Φ. For any α, $\beta \in \Phi$, Theorem 10.1.10 implies that $c(\beta, \alpha) = -q$, where $\beta + j\alpha$ $(0 \leq j \leq q)$ is the α-string through β. In particular, $\beta + \alpha$ is a root if $c(\beta, \alpha) \neq 0$. This gives us all the roots of height 2. Since $c(\cdot, \cdot)$ is linear in the first argument, we also obtain the numbers $c(\beta, \alpha)$ for all roots β of height 2 and all roots α of height 1. Inductively, suppose that all roots β of height $\leq k$ have been determined, as well as the numbers $c(\beta, \alpha)$ for all roots β of height $\leq k$ and for all roots α of height one. Now $c(\beta, \alpha) = -(p + q)$, where $\beta + j\alpha$ $(p \leq j \leq q)$ is the α-string through β. The integer p is already known, since we know all the roots of height $\leq k$. Hence we also know the integer q, and in particular whether $\beta + \alpha$ is a root. By Lemma 10.1.15, this determines all the roots of height $k + 1$. This completes the induction step. Now $\Delta = \Delta^+ \cup -\Delta^+$, and we conclude from the above that Δ is determined by the simple system Φ and its Cartan matrix.

Theorem 10.2.1. *Let Δ and Δ' be indecomposable reduced root systems on inner product spaces E and E', respectively. Suppose that $\Phi \subset \Delta$ and $\Phi' \subset \Delta'$ are simple systems of roots whose Cartan matrices are equal. Then there exists a positive number d and an isometry T from E onto E' such that $\Delta' = d\, S(\Delta)$.*

Proof. By abuse of notation, we use $(\ ,\ )$ for the inner product on both E and E'. If $\Phi = \{\alpha_1, \ldots, \alpha_l\}$ and $\Phi' = \{\alpha'_1, \ldots, \alpha'_l\}$, the hypothesis is that $c(\alpha_i, \alpha_j) = c(\alpha'_i, \alpha'_j)$ for all i and j. Let $k = (\alpha'_1, \alpha'_1)/(\alpha_1, \alpha_1)$. Since Φ is indecomposable, there is a $j_2 \neq 1$ such that $(\alpha_1, \alpha_{j_2}) \neq 0$. Since $c(\alpha_{j_2}, \alpha_1) = c(\alpha'_{j_2}, \alpha'_1)$, it follows that $(\alpha'_{j_2}, \alpha'_1) = k\,(\alpha_{j_2}, \alpha_1)$. Then since $c(\alpha_1, \alpha_{j_2}) = c(\alpha'_1, \alpha'_{j_2})$, we conclude that $(\alpha'_{j_2}, \alpha'_{j_2}) = k\,(\alpha_{j_2}, \alpha_{j_2})$. Inductively, assume that $m < l$ and that $j_1 = 1, j_2, \ldots, j_m$ are integers in $\{1, \ldots, l\}$ such that $(\alpha'_{j_r}, \alpha'_{j_s}) = k\,(\alpha_{j_r}, \alpha_{j_s})$ for all $1 \le r,\ s \le m$. Since Φ is indecomposable, there is a root $\alpha_{j_{m+1}}$, with $j_{m+1} \notin \{j_1, \ldots, j_m\}$ such that $(\alpha_{j_{m+1}}, \alpha_{j_s}) \neq 0$ for some j_s. Then repeatedly applying the same argument as that for for j_2 above, we conclude that $(\alpha'_{j_{m+1}}, \alpha'_{j_r}) = k\,(\alpha_{j_{m+1}}, \alpha_{j_r})$ for all $j_r,\ 1 \le r \le m+1$.

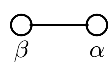
Let T be the linear map from E onto E' given on the basis Φ of E by $T(\alpha_i) = \alpha'_i$, for all i. Then by linearity, we conclude from the above that $(Tv, Tw) = k(v, w)$ for all $v,\ w \in E$. If we put $d = k^{-1/2}$ and $S = (1/d)\,T$, we see that S is an isometry of E onto E', and of course $T = d\,S$.

Since $c(\alpha_i, \alpha_j) = c(\alpha'_i, \alpha'_j)$ for all i, j, the construction of Δ from Φ (and of Δ' from Φ') shows us that $T(\Delta) = \Delta'$. Hence $d\,S(\Delta) = \Delta'$. $\square$

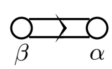
Let S be a set of positive roots such that $(\alpha, \beta) \le 0$ for all $\alpha,\ \beta \in S$. Then by Lemma 10.1.11, S is a linearly independent set. (S could, for example, be our simple system Φ.) The *Coxeter-Dynkin diagram* (or more simply, the *Dynkin diagram*) of S is a directed graph whose vertices are the elements of S, and whose directed edges are defined as follows:

1. *No edge:* if $\alpha, \beta \in S$ and $(\alpha, \beta) = 0$, then α and β are not connected by an edge.

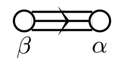
2. *Simple edge:* If $\alpha \neq \beta$ in S and $c(\alpha, \beta)c(\beta, \alpha) = 1$ then α and β are connected by a single undirected bond. Note here that $\|\alpha\| = \|\beta\|$ and that $c(\alpha, \beta) = c(\beta, \alpha) = -1$.



3. *Double edge:* If $\alpha \neq \beta$ in S, $\|\alpha\| \le \|\beta\|$, and $c(\alpha, \beta)\,c(\beta, \alpha) = 2$, then α and β are connected by a double bond from β to α. Note that $\|\beta\|^2/\|\alpha\|^2 = 2$, $c(\alpha, \beta) = -1$, and $c(\beta, \alpha) = -2$.



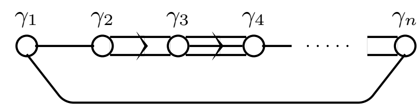
4. *Triple edge:* If $\alpha \neq \beta$ in S, $\|\alpha\| \le \|\beta\|$ and $c(\alpha, \beta)c(\beta, \alpha) = 3$, then α and β are connected by a triple bond from β to α. (Here $\|\beta\|^2/\|\alpha\|^2 = 3$.)



Suppose that $S = \Phi$, the simple system of roots. Then the Cartan matrix of Φ clearly determines its Dynkin diagram. Conversely, the Dynkin diagram of Φ determines the Cartan matrix of Φ. In fact, for any roots α and β in Φ, the integers $c(\alpha, \beta)$ and $c(\beta, \alpha)$ are uniquely determined by the type of edge (or lack of one) in the Dynkin diagram between α and β.

Proposition 10.2.2. *A Coxeter-Dynkin graph is a tree; i.e., it has no circuits.*

Proof. Suppose, to the contrary, that there are circuits. Let $\gamma_1, \dots, \gamma_n$ be the vertices of a *minimal* circuit.



Since the circuit is minimal, no root γ_i is connected to a root γ_j in the circuit unless $j \equiv (i+1) \mod n$ or $j \equiv (i-1) \mod n$.

Suppose now that γ_i and γ_j are consecutive roots in the circuit. We claim that

$$\frac{1}{2}(\gamma_i, \gamma_i) + 2(\gamma_i, \gamma_j) + \frac{1}{2}(\gamma_j, \gamma_j) \leq 0 \tag{10.6}$$

To show this, we may assume that $\|\gamma_i\| \leq \|\gamma_j\|$. Then obviously,

$$\frac{1}{2}(\gamma_i, \gamma_i) - \frac{1}{2}(\gamma_j, \gamma_j) \leq 0.$$

But by Theorem 10.1.7, Part 5, we have

$$2(\gamma_i, \gamma_j) + (\gamma_j, \gamma_j) = 0.$$

Adding the left hand sides of the last two relations above, we obtain inequality (10.6). Thus, in particular, $(1/2)(\gamma_i, \gamma_i) + 2(\gamma_i, \gamma_{i+1}) + (1/2)(\gamma_{i+1}, \gamma_{i+1}) \leq 0$ for all $i = 1, \dots, n$, where the index $i+1$ is counted modulo n. Adding these inequalities, we obtain

$$\begin{aligned} 0 &\geq \sum_{i\equiv 1}^{n} \left(\frac{1}{2}(\gamma_i, \gamma_i) + 2(\gamma_i, \gamma_{i+1}) + \frac{1}{2}(\gamma_{i+1}, \gamma_{i+1}) \right) \\ &= \sum_{i=1}^{n} (\gamma_i, \gamma_i) + 2 \sum_{i\equiv 1}^{n} (\gamma_i, \gamma_{i+1}) \end{aligned} \tag{10.7}$$

On the other hand,

$$
\begin{aligned}
0 &\leq \left(\sum_{i=1}^{n} \gamma_i, \sum_{i=1}^{n} \gamma_i \right) \\
&= \sum_{i=1}^{n} (\gamma_i, \gamma_i) + \sum_{i \neq j} (\gamma_i, \gamma_j) \\
&= \sum_{i=1}^{n} (\gamma_i, \gamma_i) + 2 \sum_{i \equiv 1}^{n} (\gamma_i, \gamma_{i+1}),
\end{aligned}
\tag{10.8}
$$

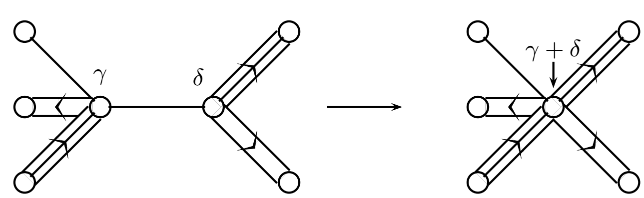
by our remark at the beginning of the proof about adjacent vertices. Inequalities (10.7) and (10.8) imply that $\sum_{i=1}^{n} \gamma_i = 0$. But this is a contradiction since the γ_i are linearly independent. $\square$

A *tree* is a graph without any circuits. Proposition 10.2.2 thus asserts that any Dynkin diagram is a tree.

A *subdiagram* of a Dynkin diagram consists of a subset of the vertex set of the Dynkin diagram and all the existing edges between them. It is clear that any subdiagram is also a Dynkin diagram.

Lemma 10.2.3. *In a Dynkin diagram, suppose that roots γ and δ are joined by a simple edge. Then the configuration resulting from the deletion of γ and δ and replacement by the single root $\gamma + \delta$, and then joining $\gamma + \delta$ to all roots connected to γ or δ by the same types of edges as γ or δ is also a Dynkin diagram.*

Example:



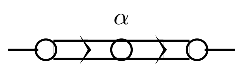
Proof. Note first that since γ and δ are connected, we have $(\gamma, \delta) \leq 0$ and thus $\gamma + \delta$ is a root. Moreover, since $c(\gamma, \delta) = c(\delta, \gamma) = -1$, we have $(\gamma, \gamma) = (\delta, \delta)$ and $2(\gamma, \delta) + (\gamma, \gamma) = 1$. Hence $(\gamma + \delta, \gamma + \delta) = (\gamma, \gamma)$.

Let S be the collection of roots β in the Dynkin diagram such that $\beta \neq \gamma$, $\beta \neq \delta$, and β is connected to γ or δ. Then by Proposition 10.2.2, each $\beta \in S$ is connected to just one of γ or δ, and not both.

So let $\beta \in S$. Without loss of generality, we can assume that β is connected to γ. Then $(\delta, \beta) = 0$, and so $c(\gamma + \delta, \beta) = c(\gamma, \beta)$. Moreover, $c(\beta, \gamma + \delta) = 2(\beta, \gamma+\delta)/(\gamma+\delta, \gamma+\delta) = 2(\beta, \gamma)/(\gamma, \gamma) = c(\beta, \gamma)$. Hence $c(\gamma+\delta, \beta)\, c(\beta, \gamma+\delta) = c(\gamma, \beta)\, c(\beta, \gamma)$. This shows that the number of bonds in the edge joining β and $\gamma + \delta$ is the same as the number of bonds in the edge joining β and γ. Finally, since $\|\gamma + \delta\| = \|\gamma\|$, the direction of the edge joining β and $\gamma + \delta$ is the same as the direction of the edge joining β and γ. $\square$

Lemma 10.2.4. *At each vertex of a Dynkin diagram, at most three bonds come together.*

Example: A vertex α of the type below cannot occur:



Proof. Let α be a vertex in the Dynkin diagram, and suppose that $\alpha_1, \dots, \alpha_k$ are the vertices in the diagram connected to α. Since the diagram has no circuits, we must have $(\alpha_i, \alpha_j) = 0$ for all $i \neq j$ in $\{1, \dots, k\}$. Let $\overline{\alpha} = \alpha/\|\alpha\|$ and $\overline{\alpha}_i = \alpha_i/\|\alpha_i\|$ for $i = 1, \dots, n$. The number of bonds between α and α_i is $c(\alpha, \alpha_i)\, c(\alpha_i, \alpha) = 4(\overline{\alpha}, \overline{\alpha}_i)^2$. Hence the number of bonds through α is $4\sum_{i=1}^k (\overline{\alpha}, \overline{\alpha}_i)^2$.

Since $\overline{\alpha}, \overline{\alpha}_1, \dots, \overline{\alpha}_k$ are linearly independent, there is a linear combination $\overline{\alpha}_0$ of $\overline{\alpha}, \overline{\alpha}_1, \dots, \overline{\alpha}_k$ such that $\overline{\alpha}_0 \perp \overline{\alpha}_i$ for all $i = 1, \dots, n$ and $\overline{\alpha}_0$ is a unit vector. Clearly, $(\overline{\alpha}, \overline{\alpha}_0) \neq 0$. Thus $\overline{\alpha}_0, \overline{\alpha}_1, \dots, \overline{\alpha}_k$ are orthonormal and

$$\overline{\alpha} = \sum_{i=0}^{k} (\overline{\alpha}, \overline{\alpha}_i)\overline{\alpha}_i$$

$$1 = (\overline{\alpha}, \overline{\alpha})^2 = \sum_{i=0}^{k} (\overline{\alpha}, \overline{\alpha}_i)^2$$

Since $(\overline{\alpha}, \overline{\alpha}_0) \neq 0$, we conclude that

$$\sum_{i=1}^{k} (\overline{\alpha}, \overline{\alpha}_i)^2 \; < \; 1$$

and thus $4\sum_{i=1}^k (\overline{\alpha}, \overline{\alpha}_i)^2 \; < \; 4$. Thus the number of bonds through α is at most 3. $\square$

A Dynkin diagram with exactly two vertices and a triple edge between them is called G_2. It corresponds to one of the five exceptional complex simple Lie algebras.

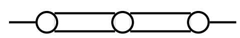
Theorem 10.2.5. *In a connected Dynkin diagram:*

(i) If a triple edge is present, then the diagram must be G_2.

(ii) If a double edge occurs, then all other edges are simple.

Proof. Assertion (i) is immediate from Lemma 10.2.4. For assertion (ii), suppose first that two double bonds occur. Then part of the diagram would look like:

where the arrows on the double edges could go either way. Using Lemma 10.2.3, this reduces to



which is impossible by Lemma 10.2.4. If the diagram has a double and a triple edge, a similar argument results in a contradiction. □

A *branch point* in a Dynkin diagram is a vertex which is connected by an edge to more than two vertices.

Theorem 10.2.6. *If α is a branch point in a connected Dynkin diagram, then it is the only branch point, and it must be of the following form:*

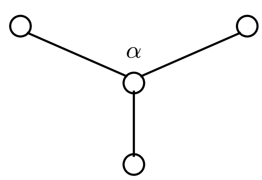
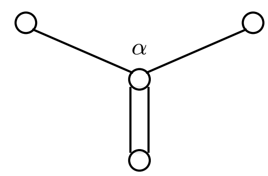


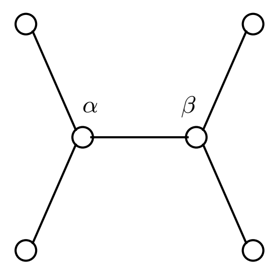
Figure 4

Proof. It is clear from Lemma 10.2.4 that the branch point α is adjacent to exactly three vertices, and that α is linked to these three vertices by simple edges. If a double edge is present somewhere else in the diagram, then by Theorem 10.2.5, all the other edges in the diagram are simple; applying Lemma 10.2.3, we would obtain

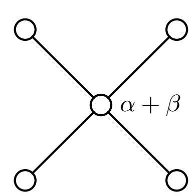


Thus, no double edges are present.

Next, suppose that β is another branch point in the diagram. Again applying Lemma 10.2.3, repeatedly if necessary, we would obtain

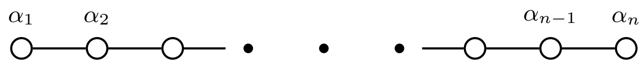


which, upon applying Lemma 10.2.3 one more time, results in



contradicting Lemma 10.2.4. □

Lemma 10.2.7. *Let*



be a simple chain. Then

$$\left\| \sum_{i=1}^{n} i\alpha_i \right\|^2 = \frac{n(n+1)}{2}\,(\alpha_1, \alpha_1).$$

Proof. This is an easy induction on n. If $n = 1$, then the above equation reduces to the trivial identity $\|\alpha_1\|^2 = (1 \cdot 2/2)\,(\alpha_1, \alpha_1)$. So assume that the identity holds for a simple chain of n vertices. We prove it for $n+1$ vertices. Now

$$\begin{aligned}
\left\| \sum_{i=1}^{n+1} i\alpha_i \right\|^2 &= \left\| \sum_{i=1}^{n} i\alpha_i + (n+1)\alpha_{n+1} \right\|^2 \\
&= \left\| \sum_{i=1}^{n} i\alpha_i \right\|^2 + 2\sum_{i=1}^{n} i(n+1)(\alpha_i, \alpha_{n+1}) + (n+1)^2\,(\alpha_{n+1}, \alpha_{n+1}) \\
&= \frac{n(n+1)}{2}\|\alpha_1\|^2 + (n+1)\left[2\sum_{i=1}^{n} i(\alpha_i, \alpha_{n+1}) + (n+1)\,(\alpha_{n+1}, \alpha_{n+1})\right]
\end{aligned}$$

(by the induction hypothesis)

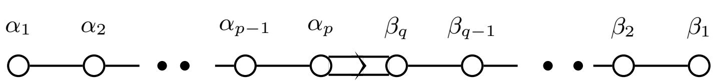
$$= \frac{n(n+1)}{2}\,(\alpha_1, \alpha_1) + (n+1)\,[2n\,(\alpha_n, \alpha_{n+1}) + (n+1)(\alpha_{n+1}, \alpha_{n+1})]$$

But $2(\alpha_n, \alpha_{n+1}) + (\alpha_{n+1}, \alpha_{n+1}) = 0$. Hence the last expression above equals

$$\frac{n(n+1)}{2}\,(\alpha_1, \alpha_1) + (n+1)\,(\alpha_{n+1}, \alpha_{n+1}) = \frac{(n+1)(n+2)}{2}\,(\alpha_1, \alpha_1),$$

the last equality arising from the fact that $(\alpha_1, \alpha_1) = (\alpha_i, \alpha_i)$ for all i, since all edges are simple. This proves the induction step and the lemma. ☐

Theorem 10.2.8. *If a connected Dynkin diagram has a double edge, it must be a chain of the form*



For this chain, there are only three possibilities: (i) $p = 1$, or (ii) $q = 1$, or (iii) $p = q = 2$.

Proof. If a double edge is present, then by Theorem 10.2.5, all other edges are simple, and by by Lemma 10.2.6, there are no branch points. Thus the diagram must be a chain of the form above. Now let

$$\sigma = \frac{1}{p+1} \sum_{i=1}^{p} i\,\alpha_i \;+\; \frac{1}{q} \sum_{j=1}^{q} j\,\beta_j.$$

Then $\sigma \neq 0$ and so

$$\begin{aligned}
0 < \|\sigma\|^2 &= \frac{1}{(p+1)^2} \left\| \sum_{i=1}^{p} i\,\alpha_i \right\|^2 \;+\; \frac{1}{q^2} \left\| \sum_{j=1}^{q} j\,\beta_j \right\|^2 \;+\; 2\,\frac{p}{p+1}\,\frac{q}{q}\,(\alpha_p, \beta_q) \\
&= \frac{1}{(p+1)^2}\,\frac{p(p+1)}{2}\,(\alpha_p, \alpha_p) \;+\; \frac{1}{q^2}\,\frac{q(q+1)}{2}\,(\beta_q, \beta_q) + \frac{2p}{p+1}\,(\alpha_p, \beta_q)
\end{aligned}$$

(by Lemma 10.2.7)

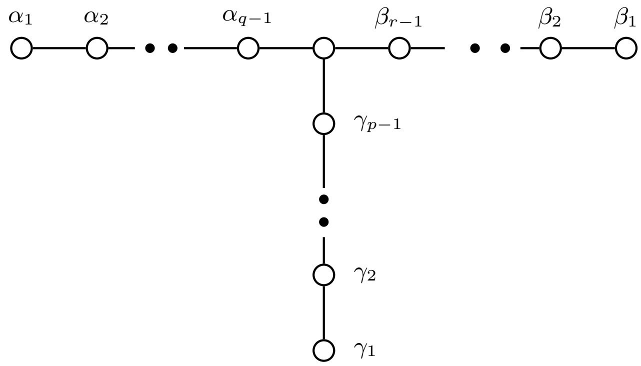
$$= \frac{p}{p+1}\,(\beta_q, \beta_q) \;+\; \frac{q+1}{2q}\,(\beta_q, \beta_q) \;-\; \frac{2p}{p+1}\,(\beta_q, \beta_q)$$

(since $(\alpha_p, \alpha_p) = 2(\beta_q, \beta_q)$ and $(\alpha_p, \beta_q) = -(\beta_q, \beta_q)$)

$$\begin{aligned}
&= \left(\frac{q+1}{2q} - \frac{p}{p+1} \right) (\beta_q, \beta_q) \\
&= \frac{p+q+1-pq}{2(p+1)q}\,(\beta_q, \beta_q)
\end{aligned}$$

It follows that $0 < p + q + 1 - pq = 2 - (p-1)(q-1)$. Thus $(p-1)(q-1) < 2$. Since $p \geq 1$ and $q \geq 1$, the only possibilities are (i), (ii), and (ii) above. $\square$

Theorem 10.2.9. *If a connected Dynkin diagram has a branch point, then the diagram must be of the form below:*



Assume that $2 \leq p \leq q \leq r$. Then the only possibilities for p, q, and r are: (i) $p = q = 2$, or (ii) $p = 2$, $q = 3$, and $r = 3, 4, 5$.

Proof. Assuming the existence of a branch point, Lemma 10.2.6 says that there is only one branch point and all edges are simple. Thus the Dynkin diagram nust be of the type given in the figure above.

In the figure above, we have put $\alpha_q = \beta_r = \gamma_p$. Note also that all the roots above have the same length. Let

$$\alpha = \frac{1}{q} \sum_{i=1}^{q} i\alpha_i, \qquad \beta = \frac{1}{r} \sum_{i=1}^{r} i\beta_i, \qquad \gamma = \frac{1}{p} \sum_{i=1}^{p} i\gamma_i$$

and put

$$\omega = \alpha + \beta + \gamma - 2\gamma_p$$

Then by the hypothesis on p, q, and r, and by the linear independence of the roots in the diagram, we have $\omega \neq 0$. Hence

$$\begin{aligned} 0 < \|\omega\|^2 = \|\gamma\|^2 + \|\alpha\|^2 + \|\beta\|^2 + 4\|\gamma_p\|^2 + 2(\alpha, \beta) + 2(\alpha, \gamma) + 2(\beta, \gamma) \\ - 4(\gamma, \gamma_p) - 4(\alpha, \gamma_p) - 4(\beta, \gamma_p) \end{aligned} \tag{10.9}$$

We will calculate the right hand side as a multiple of (γ_p, γ_p). The first three terms can be evaluated using Lemma 10.2.7. The cross term (α, β) is calculated as follows:

$$\begin{aligned}
2(\alpha,\beta) &= \frac{2}{qr}\left((q-1)\alpha_{q-1}+q\alpha_q,(r-1)\beta_{r-1}+r\beta_r\right)\\
&= \frac{2}{qr}\left[(q-1)r(\alpha_{q-1},\beta_r)+q(r-1)(\alpha_q,\beta_{r-1})+qr(\alpha_q,\beta_r)\right]\\
&= \frac{2}{qr}\left[(q-1)r\left(-\frac{1}{2}(\beta_r,\beta_r)\right)+q(r-1)\left(-\frac{1}{2}(\beta_{r-1},\beta_{r-1})\right)+qr(\beta_r,\beta_r)\right]\\
&= \frac{2}{qr}\left[-\frac{(q-1)r}{2}-\frac{q(r-1)}{2}+qr\right](\gamma_p,\gamma_p)\\
&= \left[\frac{1}{q}+\frac{1}{r}\right](\gamma_p,\gamma_p) \qquad (10.10)
\end{aligned}$$

Similarly,

$$\begin{aligned}
2(\alpha,\gamma) &= \left[\frac{1}{q}+\frac{1}{p}\right](\gamma_p,\gamma_p)\\
2(\beta,\gamma) &= \left[\frac{1}{r}+\frac{1}{p}\right](\gamma_p,\gamma_p) \qquad (10.11)
\end{aligned}$$

Let us now calculate the last three terms on the right hand side of (10.9):

$$\begin{aligned}
(\gamma,\gamma_p) &= \frac{1}{p}\left((p-1)\gamma_{p-1}+p\gamma_p,\gamma_p\right)\\
&= \frac{p-1}{p}(\gamma_{p-1},\gamma_p)+(\gamma_p,\gamma_p)\\
&= \frac{(p-1)}{p}\left(-\frac{1}{2}(\gamma_p,\gamma_p)\right)+(\gamma_p,\gamma_p)\\
&= \left[\frac{1}{2}+\frac{1}{2p}\right](\gamma_p,\gamma_p) \qquad (10.12)
\end{aligned}$$

Likewise,

$$\begin{aligned}
(\alpha,\gamma_p) &= \left[\frac{1}{2}+\frac{1}{2q}\right](\gamma_p,\gamma_p)\\
(\beta,\gamma_p) &= \left[\frac{1}{2}+\frac{1}{2r}\right](\gamma_p,\gamma_p) \qquad (10.13)
\end{aligned}$$

We can now calculate the right hand side or (10.9). Applying Lemma 10.2.7 to evaluate $\|\alpha\|^2$, $\|\beta\|^2$, and $\|\gamma\|^2$, and using (10.10), (10.11), (10.12), (10.13), we obtain

$$
\begin{aligned}
(\omega, \omega) = {} & \frac{1}{q^2}\,\frac{q(q+1)}{2}\,(\gamma_p, \gamma_p) + \frac{1}{r^2}\,\frac{r(r+1)}{2}\,(\gamma_p, \gamma_p) + \frac{1}{p^2}\,\frac{p(p+1)}{2}\,(\gamma_p, \gamma_p) \\
& + 4(\gamma_p, \gamma_p) + \left[\frac{1}{q} + \frac{1}{r}\right](\gamma_p, \gamma_p) + \left[\frac{1}{q} + \frac{1}{p}\right](\gamma_p, \gamma_p) \\
& + \left[\frac{1}{r} + \frac{1}{p}\right](\gamma_p, \gamma_p) - 4\left[\frac{1}{2} + \frac{1}{2p}\right](\gamma_p, \gamma_p) - 4\left[\frac{1}{2} + \frac{1}{2q}\right](\gamma_p, \gamma_p) \\
& - 4\left[\frac{1}{2} + \frac{1}{2r}\right](\gamma_p, \gamma_p) \\
= {} & (\gamma_p, \gamma_p)\left[\frac{1}{2} + \frac{1}{2q} + \frac{1}{2} + \frac{1}{2r} + \frac{1}{2} + \frac{1}{2p} + 4 + \frac{1}{q} + \frac{1}{r} + \frac{1}{q} + \frac{1}{p} + \frac{1}{r} + \frac{1}{p}\right. \\
& \left. - 2 - \frac{2}{p} - 2 - \frac{2}{q} - 2 - \frac{2}{r}\right] \\
= {} & \frac{1}{2}(\gamma_p, \gamma_p)\left[\frac{1}{p} + \frac{1}{q} + \frac{1}{r} - 1\right]
\end{aligned}
$$

It follows that

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} - 1 > 0. \tag{10.14}$$

But $2 \leq p \leq q \leq r$, and so $1/p \geq 1/q \geq 1/r$. Thus $3/p - 1 > 0$, and therefore $p < 3$. We conclude that $p = 2$. Thus we obtain

$$
\begin{aligned}
& \frac{1}{2} + \frac{1}{q} + \frac{1}{r} - 1 > 0 \\
& \quad \implies \frac{1}{q} + \frac{1}{r} > \frac{1}{2}
\end{aligned}
$$

But since $1/q \geq 1/r$, we obtain

$$\frac{2}{q} - \frac{1}{2} > 0$$

and so $q < 4$. But $2 = p \leq q$. There are thus two possibilities for q:

Case 1: $q = 2$. Thus $p = q = 2$. In this case, r can be any integer ≥ 2, since the inequality (10.14) would always be satisfied.

Case 2: $q = 3$. Inequality (10.14) then becomes

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{r} - 1 > 0$$

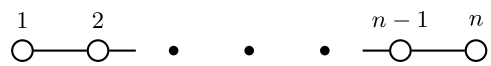
which implies that

$$\frac{1}{r} > \frac{1}{6},$$

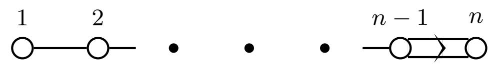
whence $r < 6$, and so r must equal 3, 4, or 5. This finishes the proof of Theorem 10.2.9. $\square$

Theorem 10.2.10. *Let Δ be an indecomposable root system in an inner product space E, and let Φ be a simple system of roots in Δ. Then the only possible Dynkin diagrams for Φ are:*

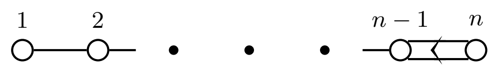
A_n



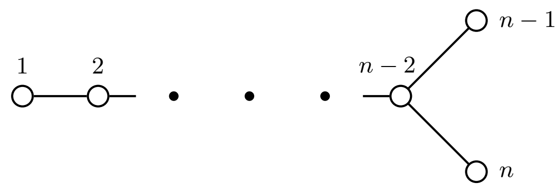
B_n



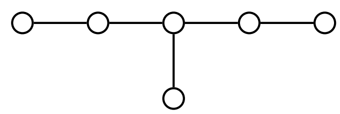
C_n



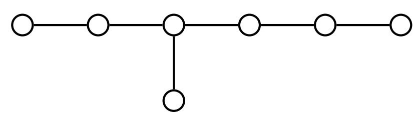
D_n



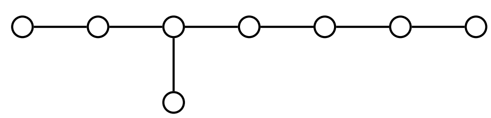
E_6



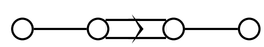
E_7



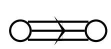
E_8



F_4



G_2



Proof. Since Φ is indecomposable, its Dynkin diagram must be connected. If the Dynkin diagram is a simple chain, then it must be of type A_n. Otherwise, the diagram will contain either a multiple edge or a branch point, but not both.

If it contains a triple edge, then, by Theorem 10.2.5, it must be G_2. If it contains a double edge, then all other edges are simple (again by Theorem 10.2.5). Thus, according to Theorem 10.2.8, there are only three possibilities:

$$\begin{aligned} p = 1 &\implies C_n \\ q = 1 &\implies B_n \\ p = q = 2 &\implies F_4 \end{aligned}$$

Finally, if the Dynkin diagram contains a branch point, then by Theorem 10.2.9, the only possibilities are:

$$\begin{aligned} p = q = 2 &\implies D_n \\ p = 2,\ q = 3,\ r = 3 &\implies E_6 \\ p = 2,\ q = 3,\ r = 4 &\implies E_7 \\ p = 2,\ q = 3,\ r = 5 &\implies E_8 \end{aligned}$$

$\square$

Chapter 11

Uniqueness Questions and the Weyl Group

In this chapter, we further explore basic questions about abstract root systems. One such question, which was left unanswered in Chapter 10, is the following: suppose that Δ is a root system in a real inner product space E, and that Φ and Φ' are two simple systems in Δ. Are the Dynkin diagrams of Φ and Φ' the same. Of course, one expects the answer to be "yes," and it is partly the objective of this chapter to provide a proof. For this, we will need a keener exploration of the Weyl group of Δ. At the end of this chapter, we will also classify nonreduced root systems.

In this chapter, we fix a real inner product space E and and a root system Δ in E.

11.1 Properties of the Weyl Group

In this section, we assume that the root system Δ is *reduced*. A *positive system of roots in* Δ is the set of positive roots in Δ relative to a lexicographic order $>$ on E. Such a positive system (once $>$ is given) will be denoted by Δ^+. Generally, there is more than one way to order E to produce a given positive system Δ^+.

Given a positive system Δ^+, we defined a simple root to be a positive root β which is not the sum of two positive roots, and the simple system Φ to be the set of all simple roots. We also saw that Φ is a basis of E and that any $\gamma \in \Delta^+$

can be written as a linear combination

$$\gamma = \sum_{\alpha \in \Phi} n_\alpha \alpha \tag{11.1}$$

where each $n_\alpha \in \mathbb{Z}^+$.

The following housecleaning lemma shows that Φ is the only subset of Δ^+ which satisfies the property (11.1).

Lemma 11.1.1. *Suppose that $\Phi' \subset \Delta^+$ is a basis of E and satisfies* (11.1). *Then $\Phi' = \Phi$.*

Proof. Put $\Phi = \{\alpha_1, \ldots, \alpha_l\}$ and $\Phi' = \{\beta_1, \ldots, \beta_l\}$. Since Φ and Φ' satisfy (11.1), there exist nonnegative integers c_{ij} and d_{ij} (for $1 \leq i, j \leq l$) such that

$$\beta_j = \sum_{i=1}^{l} c_{ij} \alpha_i, \qquad \alpha_j = \sum_{i=1}^{l} d_{ij} \beta_i$$

Since (c_{ij}) and (d_{ij}) are the change of basis matrices from Φ to Φ' and vice versa, we have $(d_{ij}) = (c_{ij})^{-1}$. Thus in particular,

$$\sum_{j=1}^{l} a_{1j} b_{j1} = 1$$

Since the a_{1j} and the b_{j1} are in $\mathbb{Z}^+$, we conclude that there exists a j_1 such that $a_{ij_1} = b_{j_1 1} = 1$. We claim that $a_{1j} = 0$ for all $j \neq j_1$. Suppose, to the contrary, that $a_{1j_2} > 0$ for some $j_2 \neq j_1$. Since (b_{ij}) is nonsingular, there is an i, with $i \neq 1$, such that $b_{j_2,i} > 0$. This would then imply that $\sum_{j=1}^{l} a_{1j} b_{ji} \geq a_{1j_2} b_{j_2 i} > 0$, contradicting the fact that $\sum_{j=1}^{l} a_{1j} b_{ji} = 0$.

Thus, the first row of (a_{ij}) contains only one nonzero entry, and that entry is a 1. A similar argument shows that every row of (a_{ij}) contains exactly one 1 and $l - 1$ zeros. Since each column of (a_{ij}) has at least one nonzero entry, this shows that (a_{ij}) is a permutation matrix. Thus (b_{ij}) is the inverse permutation matrix, and so $\beta_1, \ldots, \beta_l$ is a permutation of $\alpha_1, \ldots, \alpha_l$. This implies that $\{\beta_1, \ldots, \beta_l\} = \{\alpha_1, \ldots, \alpha_l\}$; i.e., that $\Phi = \Phi'$. ☐

Let us recall that the Weyl group W is the subgroup of $\mathrm{O}(E)$ generated by the reflections r_α, $\alpha \in \Delta$. We saw that W permutes the elements of Δ.

Lemma 11.1.2. *Suppose that Φ is a simple system of roots in Δ and $\alpha \in \Phi$. Then $r_\alpha(\alpha) = -\alpha$. If $\beta \in \Delta^+$ is not proportional to α, then $r_\alpha(\beta) \in \Delta^+$.*

Proof. $r_\alpha(\alpha) = -\alpha$ follows from the definition of r_α. Suppose $\beta \in \Delta^+$ is not proportional to α. Then

$$\beta = \sum_{\gamma \in \Phi} n_\gamma \gamma$$

where each $n_\gamma \in \mathbb{Z}^+$ and $n_\gamma > 0$ for at least one $\gamma \neq \alpha$. We obtain

$$\begin{aligned} r_\alpha(\beta) &= \sum_{\gamma\in\Phi} n_\gamma r_\alpha(\gamma) \\ &= \sum_{\gamma\in\Phi} n_\gamma \gamma - \left(\sum_{\gamma\in\Phi} n_\gamma c(\gamma,\alpha)\right)\alpha \end{aligned}$$

The coefficient of γ on the right hand side above is still n_γ, which is a positive integer, so the right hand side must still be a positive root. □

For the remainder of this section, we fix a simple system of roots $\Phi = \{\alpha_1, \dots, \alpha_l\}$ in Δ. For each i, the reflection r_{α_i} is called a *simple reflection*. Let W_0 be the subgroup of W generated by the simple reflections.

Lemma 11.1.3. *Each $\beta \in \Delta$ is the image of a simple root α_i under an element of W_0.*

Proof. Since $r_{\alpha_i}(\alpha_i) = -\alpha_i$ for all i, it suffices to prove the lemma for any positive root β. Thus the claim is that for any $\beta \in \Delta^+$, there is an element $w \in W_0$ and an $\alpha_i \in \Phi$ such that $\beta = w(\alpha_i)$. We prove the claim by induction on the height $\mathrm{ht}(\beta)$. If $\mathrm{ht}(\beta) = 1$, then β is a simple root, and we have $\beta = \alpha_i$ for some i, so we can put $w = e := I_E$, the identity element of W.

Assume that $k > 1$ and that the lemma holds for all $\beta \in \Delta^+$ of height $< k$. Let β be a positive root of height k. Then β is not a simple root. Write $\beta = \sum_{i=1}^l n_i \alpha_i$. Now $0 < (\beta,\beta) = \sum_{i=1}^l n_i(\beta,\alpha_i)$, and therefore $(\beta,\alpha_i) > 0$ for some i. For this i, the root

$$r_{\alpha_i}(\beta) = \beta - \frac{2(\beta,\alpha_i)}{(\alpha_i,\alpha_i)}\,\alpha_i$$

is positive, by Lemma 11.1.2. Moreover, the above shows that $\mathrm{ht}(r_{\alpha_i}(\beta)) = \mathrm{ht}(\beta) - c(\beta,\alpha_i) < \mathrm{ht}(\beta)$. By the induction hypothesis, there exists a $w_1 \in W_0$ and an $\alpha_j \in \Phi$ such that $r_{\alpha_i}(\beta) = w_1(\alpha_j)$. Since $r_{\alpha_i}^2 = e$, we have $\beta = r_{\alpha_i} w_1(\alpha_j)$, with $r_{\alpha_i} w_1 \in W_0$. □

Theorem 11.1.4. $W_0 = W$.

Thus W is generated by the simple relections α_i.

Proof. W is of course generated by the reflections r_β, for all $\beta \in \Delta^+$. We'll show that $r_\beta \in W_0$. By Lemma 11.1.3, there exists a $w \in W_0$ and an $\alpha_i \in \Phi$ such that $\beta = w(\alpha_i)$. We claim that $r_\beta = w r_{\alpha_i} w^{-1}$. This will of course prove that $r_\beta \in W_0$.

Now for any $v \in E$, we have

$$
\begin{aligned}
w\, r_{\alpha_i} w^{-1}(v) &= w\big(w^{-1}v - \frac{2(w^{-1}v, \alpha_i)}{(\alpha_i, \alpha_i)}\, \alpha_i\big) \\
&= w\big(w^{-1}v - \frac{2(v, w(\alpha_i))}{(w(\alpha_i), w(\alpha_i))}\, \alpha_i\big)
\end{aligned}
$$

(since w is an isometry)

$$
\begin{aligned}
&= v - \frac{2(v, w(\alpha_i))}{(w(\alpha_i), w(\alpha_i))}\, w(\alpha_i) \\
&= v - \frac{2(v, \beta)}{(\beta, \beta)}\, \beta \\
&= r_\beta(v).
\end{aligned}
$$

This shows that $r_\beta = w r_{\alpha_i} w^{-1}$ and finishes the proof. $\square$

The following theorem answers the question posed at the beginning of this chapter.

Theorem 11.1.5. *(a) For any $w \in W$, the image $w(\Phi)$ is a simple system of roots in Δ.*

(b) Given any simple system of roots Φ_1 in Δ, there is a $w \in W$ such that $w(\Phi) = \Phi_1$.

Proof. (a) Let $\beta \in \Delta$. Put $\gamma = w^{-1}(\beta)$. Then $\gamma \in \Delta$, and so we can write $\gamma = \sum_{i=1}^{l} n_i \alpha_i$, where either $n_i \in \mathbb{Z}^+$ for all i or $n_i \in -\mathbb{Z}^+$ for all i. It follows that

$$
\beta = w(\gamma) = \sum_{i=1}^{l} n_i\, w(\alpha_i)
$$

where either $n_i \in \mathbb{Z}^+$ for all i or $n_i \in -\mathbb{Z}^+$ for all i. Thus $w(\Phi) = \{w(\alpha_1), \dots, w(\alpha_l)\}$ satisfies the hypotheses of Proposition 10.1.14, and hence is a simple system of roots.

(b) By Proposition 10.1.14 and Lemma 11.1.1, there is a one-to-one correspondence between simple systems of roots and positive systems of roots in Δ, relative to some lexicographic order on E. Therefore, to prove part (b), it suffices to prove that if Δ_1 is any positive system of roots in E, then there is a $w \in W$ such that $w(\Delta^+) = \Delta_1$. We will do this by downward induction on the number of elements of $\Delta^+ \cap \Delta_1$.

For any subset S of Δ, let $n(S)$ denote the number of elements of S. Since $n(\Delta)$ is even, we can put $n(\Delta) = 2N$. Then $n(\Delta^+) = N$.

If $n(\Delta^+ \cap \Delta_1) = N$, then $\Delta^+ = \Delta_1$, so we can put $w = e$.

Next, we prove the induction step. So let $M < N$, assume that assertion (b) holds for all positive systems of roots Δ' in Δ such that $n(\Delta^+ \cap \Delta') > M$. Suppose that Δ_1 is a positive system of roots in Δ such that $n(\Delta^+ \cap \Delta_1) = M$.

Now since $n(\Delta^+ \cap \Delta_1) < N$, we see that Δ^+ is not a subset of Δ_1. It follows that Φ is not a subset of Δ_1: otherwise, all nonnegative integer linear combinations of Φ which are roots would also belong to Δ_1, so $\Delta^+ \subset \Delta_1$. Hence there exists $\alpha_i \in \Phi$ such that $\alpha_i \notin \Delta_1$, whence $-\alpha_i \in \Delta_1$.

Now from Lemma 11.1.2, we have

$$r_{\alpha_i}(\Delta^+) = \big(\Delta^+ \setminus \{\alpha_i\}\big) \cup \{-\alpha_i\}.$$

Hence

$$(r_{\alpha_i}\Delta^+) \cap \Delta_1 = \big(\Delta^+ \cap \Delta_1\big) \cup \{-\alpha_i\}$$

so that $n((r_{\alpha_i}(\Delta^+)) \cap \Delta_1) = n(\Delta^+ \cap \Delta_1) + 1 = M + 1$. This implies that $n(\Delta^+ \cap (r_{\alpha_i}(\Delta_1))) = M + 1$. Now $r_{\alpha_i}(\Delta_1)$ is also a positive system of roots, by Part (a) above and the one-to-one correspondence between positive systems and simple systems of roots. Hence, by the induction hypothesis, there is a $w_1 \in W$ such that

$$r_{\alpha_i}(\Delta_1) = w_1(\Delta^+).$$

This implies that

$$\Delta_1 = r_{\alpha_i} w_1(\Delta^+).$$

This completes the proof of Part (b) of the Theorem.

$\square$

Corollary 11.1.6. *There is an element $w^* \in W$ such that $w^*(\Delta^+) = -\Delta^+$.*

Proof. $-\Phi$ is a simple system of roots according to Proposition 10.1.14. Its corresponding "positive" system of roots is of course $-\Delta^+$. Thus, according to the proof of Part (b) of Theorem 11.1.5, there is a $w^* \in W$ such that $w^*(\Delta^+) = -\Delta^+$ (and $w^*(\Phi) = -\Phi$). $\square$

Remarks on the uniqueness of Dynkin dagrams for complex simple Lie algebras. Let $\mathfrak{g}$ be a complex simple Lie algebra, let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$, and let Δ be the set of roots of $\mathfrak{g}$ relative to $\mathfrak{h}$. Suppose that Φ and Φ_1 are simple systems of roots in Δ. According to Theorem 11.1.5, there is a $w \in W$ such that $w(\Phi) = \Phi_1$. Since w is an isometry of $E = \mathfrak{h}_{\mathbb{R}}^*$, we see that the Cartan matrices, and hence the Dynkin diagrams, of Φ and Φ_1 are the same.

Next let $\mathfrak{h}'$ be another Cartan subalgebra of $\mathfrak{g}$ and let Δ' be the set of roots of $\mathfrak{g}$ relative to $\mathfrak{h}'$. According to Theorem 8.1.9, there exists an automorphism φ of $\mathfrak{g}$ such that $\varphi(\mathfrak{h}) = \mathfrak{h}'$. An easy argument similar to that in the proof of Theorem

10.1.19 shows that ${}^t\varphi(\Delta') = \Delta$. Suppose that Φ' is any simple system of roots in Δ'. Then ${}^t\varphi(\Phi')$ is a simple system of roots in Δ: in fact, any element of Δ is an integer linear combination of the elements of ${}^t\varphi(\Phi')$, with all coefficients in $\mathbb{Z}^+$ or in $-\mathbb{Z}^+$. The proof of Theorem 10.1.19 also shows that ${}^t\varphi$ is an isometry of $(\mathfrak{h}')^*_{\mathbb{R}}$ onto $\mathfrak{h}^*_{\mathbb{R}}$. Thus the Cartan matrices, and the Dynkin diagrams of Φ' and ${}^t\varphi(\Phi')$ coincide.

The above shows that for any choice of Cartan subalgebra of $\mathfrak{g}$ and any choice of a simple system of roots, the Dynkin diagram which results is the same. Thus, any simple Lie algebra over $\mathbb{C}$ corresponds to exactly one of the nine Dynkin diagrams in Theorem 10.2.10.

Conversely, suppose that $\mathfrak{g}$ and $\mathfrak{g}'$ are complex simple Lie algebras with the same Dynkin diagram. Fix Cartan subalgebras $\mathfrak{h}$ and $\mathfrak{h}'$ of $\mathfrak{g}$ and $\mathfrak{g}'$, respectively, let Δ and Δ' be the respective root systems, and let Φ and Φ' be simple systems of roots in Δ and Δ', respectively. Put $\Phi = \{\alpha_1, \dots, \alpha_l\}$, $\Phi' = \{\alpha'_1, \dots, \alpha'_l\}$. By hypothesis, the Cartan matrices of Φ and Φ' are the same. We can therefore order Φ and Φ' so that $c(\alpha_i, \alpha_j) = c(\alpha'_i, \alpha'_j)$ for all i and j. Now according to Section 10.2, the Cartan matrix of a simple system of roots completely determines the root system. The map $\psi : \alpha'_i \mapsto \alpha_i$ $(1 \le i \le l)$ therefore extends to an $\mathbb{R}$-linear bijection of $(\mathfrak{h}')^*_{\mathbb{R}}$ onto $\mathfrak{h}^*_{\mathbb{R}}$ such that $\psi(\Delta') = \Delta$. Now put $\varphi = {}^t\psi$. Then φ is an $\mathbb{R}$-linear bijection from $\mathfrak{h}_{\mathbb{R}}$ onto $\mathfrak{h}'_{\mathbb{R}}$ with ${}^t\varphi = \psi$. Since $\mathfrak{h}_{\mathbb{R}}$ and $\mathfrak{h}'_{\mathbb{R}}$ are real forms of $\mathfrak{h}$ and $\mathfrak{h}'$, respectively, φ extends naturally to a $\mathbb{C}$-linear bijection of $\mathfrak{h}$ onto $\mathfrak{h}'$. Since ${}^t\varphi(\Delta') = \Delta$, we therefore conclude from Theorem 8.3.1 that φ extends to a Lie algebra isomorphism of $\mathfrak{g}$ onto $\mathfrak{g}'$.

This shows that if a Dynkin diagram in Theorem 10.2.10 corresponds to a simple Lie algebra $\mathfrak{g}$, then that Lie algebra must be unique up to isomorphism.

In the next chapter, we will prove, using a constructive argument, that any connected Dynkin diagram is the Dynkin diagram of a complex simple Lie algebra $\mathfrak{g}$. By the above, $\mathfrak{g}$ is unique up to isomorphism.

Definition 11.1.7. Let $w \in W$. The *length* $l(w)$ of w is the minimal number of simple reflections r_{α_i} such that $w = r_{\alpha_{i_1}} \cdots r_{\alpha_{i_k}}$.

For instance, $l(r_{\alpha_i}) = 1$ for any simple reflection r_{α_i}. By definition, $l(e) = 0$.

Now, for any $w \in W$, let $n(w) = n(\Delta^+ \cap w^{-1}(-\Delta^+))$. Thus $n(w)$ is the number of positive roots α such that $w(\alpha)$ is a negative root. For example, $n(e) = 0$. Moreover, according to Lemma 11.1.2, $n(r_{\alpha_i}) = 1$ for any simple reflection r_{α_i}.

The following remarkable theorem states that $n(w)$ and $l(w)$ are the same:

Theorem 11.1.8. *For any $w \in W$, we have $n(w) = l(w)$.*

Appendix A

The Proof of Weyl's Theorem

In this appendix, we will prove Weyl's Theorem (Theorem 7.3.7) on the complete reducibility of representations of semisimple Lie algebras over $\mathbb{F}$. Weyl's original proof employed a simple argument for compact Lie groups, and was extended to arbitrary Lie algebras using a complexification argument, the so-called "Weyl unitary trick." The argument presented here, due to Serre, uses no Lie group theory. (See [10] and [5], §6.)

A.1 The Contragredient Representation

Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$, V a vector space over $\mathbb{F}$, and let π be a representation of $\mathfrak{g}$ on V. We define a representation π^* of $\mathfrak{g}$ on the dual space V^*, called the *contragredient representation* as follows. If $f \in V^*$ and $x \in \mathfrak{g}$, we put $\pi^*(x)\, f = -\,{}^t\pi(x)(f)$.

Proposition A.1.1. *π^* is a representation of $\mathfrak{g}$ on V^*.*

Proof. It is clear that $\pi^*(x)$ is a linear operator on V^*, so that $\pi^*(x) \in \mathrm{gl}\,(V^*)$. Moreover, if $x, \in \mathfrak{g}$, $\alpha, \beta \in \mathbb{F}$, $f \in V^*$, and $v \in V$, then

$$\begin{aligned}
(\pi^*(\alpha x + \beta y)\, f)\,(v) &= -\,(\,{}^t\pi(\alpha x + \beta y)\, f)\,(v) \\
&= -f(\pi(\alpha x + \beta y)(v)) \\
&= -\alpha\, f(\pi(x)\, v) - \beta\, f(\pi(y)\, v) \\
&= -\alpha\,(\,{}^t\pi(x)\, f)(v) - \beta\,(\,{}^t\pi(y)\, f)(v) \\
&= (\alpha\, \pi^*(x) + \beta\, \pi^*(y))\, f\,(v)
\end{aligned}$$

This shows that $\pi^*(\alpha x + \beta y) = \alpha\pi^*(x) + \beta\pi^*(y)$, so π^* is a linear map from $\mathfrak{g}$ into $\mathrm{gl}\,(V^*)$. Finally, we prove that $\pi^*[x, y] = [\pi^*(x), \pi^*(y)]$ for all x, $y \in \mathfrak{g}$. If $f \in V^*$ and $v \in V$, then

$$\begin{aligned}
(\pi^*[x,y]\,f)(v) &= -f(\pi[x,y]\,v)\\
&= -f([\pi(x), \pi(y)](v))\\
&= f(\pi(y) \circ \pi(x)(v)) - f(\pi(x) \circ \pi(y)(v))\\
&= ({}^t\pi(y) f)(\pi(x)(v)) - ({}^t\pi(y) f)(\pi(x)(v))\\
&= ({}^t\pi(x) \circ^t \pi(y) f)(v) - ({}^t\pi(y) \circ^t \pi(x) f)(v)\\
&= [{}^t\pi(x),\, {}^t\pi(y)](f)(v)\\
&= [\pi^*(x), \pi^*(y)](f)(v).
\end{aligned}$$

□

Proposition A.1.2. *π is irreducible if and only if π^* is irreducible.*

Proof. Let us first assume that π is irreducible. We'll prove that π^* is irreducible. Suppose that Λ is a subspace of V^* invariant under $\pi^*(x)$, for all $x \in \mathfrak{g}$. Let $W = \{v \in V \,|\, \lambda(v) = 0 \text{ for all } \lambda \in \Lambda\}$. It is easy to show, using appropriate dual bases of V and V^*, that $\dim W = \dim V - \dim \Lambda$.

It is straightforward to show that W is a $\mathfrak{g}$-invariant subspace of V. In fact, for any $\lambda \in \Lambda$ and $x \in \mathfrak{g}$, we have $\lambda(\pi(x)\,W) = -(\pi^*(x)(\lambda))(W) = \{0\}$, since $\pi^*(x)\lambda \in \Lambda$. This implies that $\pi(x)\,W \subset W$. Since π is irreducible, either $W = \{0\}$ or $W = V$. If $W = \{0\}$, then $\Lambda = V^*$. If $W = V$, then $\Lambda = \{0\}$; if $W = \{0\}$, then $\Lambda = V^*$. This shows that π^* is irreducible.

To prove the converse, we just need to note that V can be identified with the second dual space $(V^*)^*$ via the map $v \mapsto \epsilon_v$, where $\epsilon_v(f) = f(v)$ for all $f \in V^*$. Under this identification, $(\pi^*)^*(x) = \pi(x)$, for all $x \in \mathfrak{g}$. Thus, by the first part of the proof, π^* irreducible $\implies$ $(\pi^*)^* = \pi$ is irreducible. □

Now suppose that π_1 and π_2 are representations of $\mathfrak{g}$ on vector spaces V and W, respectively. Let us recall from §1.2 that $\mathcal{L}(V, W)$ denotes the vector space of all linear maps from V to W.

Proposition A.1.3. *$\mathfrak{g}$ acts on $\mathcal{L}(V, W)$ via the operators $\pi(x)$, where*

$$\pi(x)\,T = \pi_2(x) \circ T - T \circ \pi_1(x), \tag{A.1}$$

for all $T \in \mathcal{L}(V, W)$ and all $x \in \mathfrak{g}$.

Proof. It is straightforward to show that each $\pi(x)$ is a linear operator on $\mathcal{L}(V, W)$ and that the map $x \mapsto \pi(x)$ is a linear map from $\mathfrak{g}$ into $\mathrm{gl}(\mathcal{L}(V, W))$.

Thus we just need to prove that $\pi[x,y] = [\pi(x), \pi(y)]$, for all $x, y \in \mathfrak{g}$. Note that if $T \in \mathcal{L}(V_1, V_2)$, then

$$\pi(x)\,T = (L_{\pi_2(x)} - R_{\pi_1(x)})(T), \tag{A.2}$$

where $L_{\pi_2(x)}$ represents left multiplication and $R_{\pi_1(x)}$ right multiplication. Then

$$\begin{aligned}\pi[x,y] &= L_{\pi_2[x,y]} - R_{\pi_1[x,y]}\\ &= L_{[\pi_2(x),\pi_2(y)]} - R_{[\pi_1(x),\pi_1(y)]}\\ &= L_{\pi_2(x)} \circ L_{\pi_2(y)} - L_{\pi_2(y)} \circ L_{\pi_2(x)}\\ &\qquad\qquad - R_{\pi_1(y)} \circ R_{\pi_1(x)} + R_{\pi_1(x)} \circ R_{\pi_1(y)}\end{aligned} \tag{A.3}$$

Note that the last two terms above come about because, for all $T \in \mathcal{L}(V_1, V_2)$, $R_{\pi_1(x)\pi_1(y)}(T) = T\pi_1(x)\pi_1(y) = R_{\pi_1(y)} \circ R_{\pi_1(x)}(T)$, and likewise $R_{\pi_1(x)\pi_1(y)}(T) = R_{\pi_1(y)} \circ R_{\pi_1(x)}(T)$.

On the other hand,

$$\begin{aligned}[\pi(x), \pi(y)] &= \pi(x) \circ \pi(y) - \pi(y) \circ \pi(x)\\ &= (L_{\pi_2(x)} - R_{\pi_1(x)}) \circ (L_{\pi_2(y)} - R_{\pi_1(y)})\\ &\qquad\qquad - (L_{\pi_2(y)} - R_{\pi_1(y)}) \circ (L_{\pi_2(x)} - R_{\pi_1(x)})\\ &= L_{\pi_2(x)} \circ L_{\pi_2(y)} + R_{\pi_1(x)} \circ R_{\pi_1(y)}\\ &\qquad\qquad - L_{\pi_2(y)} \circ L_{\pi_2(x)} - R_{\pi_1(y)} \circ R_{\pi_1(x)},\end{aligned}$$

since $L_{\pi_2(x)}$ commutes with $R_{\pi_1(y)}$ and $R_{\pi_1(x)}$ commutes with $L_{\pi_2(y)}$. $\square$

We also observe that an element $T \in \mathcal{L}(V_1, V_2)$ intertwines π_1 and π_2 if and only if $\pi_2(x)T = T\pi_1(x)$ for all $x \in \mathfrak{g}$; that is, if and only if $\pi(x)\,T = 0$. Thus the set of all intertwining maps is the subspace

$$\bigcap_{x \in \mathfrak{g}} \ker \pi(x),$$

of $\mathcal{L}(V, W)$. As we've already observed, $\mathfrak{g}$ acts trivially on this subspace.

A.2 Casimir Operators

Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{F}$, and let V be a vector space over $\mathbb{F}$. If π is a representation of $\mathfrak{g}$ on V, then $\ker \pi$ is an ideal of $\mathfrak{g}$. π is said to be *faithful* if $\ker \pi = \{0\}$. In this case, $\mathfrak{g} \cong \pi(\mathfrak{g})$, and so it is often convenient to identify $\mathfrak{g}$ with its homomorphic image $\pi(\mathfrak{g}) \subset \mathrm{gl}\,(V)$.

Example A.2.1. If $\mathfrak{g}$ is semisimple, then its center $\mathfrak{c}$ is $\{0\}$, so the adjoint representation ad (whose kernel is $\mathfrak{c}$) is faithful.

Suppose that $\mathfrak{g}$ is semisimple and π is a faithful representation of $\mathfrak{g}$ on V. Then according to Proposition 6.4.9, the *trace form*

$$B_\pi(x, y) = \mathrm{tr}\,(\pi(x) \circ \pi(y)) \tag{A.4}$$

is a nondegenerate symmetric bilinear form on $\mathfrak{g}$.

Fix any basis $(x_1, \ldots, x_n)$ of $\mathfrak{g}$. For each $i, j = 1, \ldots, n$, put $b_{ij} = B_\pi(x_i, x_j)$. Then, by Theorem 1.10.4, the $n \times n$ symmetric matrix $B = (b_{ij})$ is nonsingular. Let $B^{-1} = (b^{ij})$ be its inverse. Thus,

$$\sum_{j=1}^{n} b_{ij}\, b^{jk} = \delta_{ik}$$

for all $i, k \in \{1, \ldots, n\}$.

Definition A.2.2. The *Casimir operator of* π is the element $\Omega_\pi \in \mathrm{gl}\,(V)$ given by

$$\Omega_\pi = \sum_{i,j} b^{ij}\, \pi(x_i) \circ \pi(x_j). \tag{A.5}$$

Although it might seem that Ω_π depends on the choice of basis of $\mathfrak{g}$, it can be shown by a straightforward calculation that this is not really the case. As a matter of fact, let $(y_1, \ldots, y_n)$ be another basis of $\mathfrak{g}$, and let $S = (s_{ij})$ be the change of basis matrix from $(x_1, \ldots, x_n)$ to $(y_1, \ldots, y_n)$. If $A = (a_{ij})$ is the matrix of B_π with respect to $(y_1, \ldots, y_n)$ (so that $a_{ij} = B_\pi(y_i, y_j)$) then for all i, j,

$$\begin{aligned} a_{ij} &= B_\pi(y_i, y_j) \\ &= B_\pi\left(\sum_{k=1}^{n} s_{ki} x_k, \sum_{l=1}^{n} s_{lj} x_l\right) \\ &= \sum_{k=1}^{n}\sum_{l=1}^{n} s_{ki}\, b_{kl}\, s_{lj}. \end{aligned}$$

This shows that $A = {}^tS\,B\,S$. Hence $A^{-1} = S^{-1}\,B^{-1}\,({}^tS)^{-1}$, and so $B^{-1} = S\,A^{-1}\,{}^tS$. But S^{-1} is the change of basis matrix from $(y_1, \ldots, y_n)$ to $(x_1, \ldots, x_n)$.

Thus, if $A^{-1} = (a^{ij})$, we obtain

$$\begin{aligned}
\Omega_\pi &= \sum_{i,j} b^{ij}\, \pi(x_i) \circ \pi(x_j) \\
&= \sum_{i,j} \left(\sum_{k,l} s_{ik}\, a^{kl}\, s_{jl} \right) \pi(x_i) \circ \pi(x_j) \\
&= \sum_{k,l} a^{kl}\, \pi \left(\sum_i s_{ik} x_i \right) \circ \pi \left(\sum_l s_{jl} x_j \right) \\
&= \sum_{k,l} a^{kl}\, \pi(y_k) \circ \pi(y_l)
\end{aligned}$$

This shows that Ω_π is independent of the choice of basis of $\mathfrak{g}$ used to define it.

For each $x \in \mathfrak{g}$, let f_x be the linear functional on $\mathfrak{g}$ given by $f_x(y) = B_\pi(x, y)$ $(y \in \mathfrak{g})$. Since B_π is nondegenerate, the proof of Proposition 1.10.7 shows that the map $v \mapsto f_v$ is a linear bijection of V onto V^*.

Let $(x_1, \ldots x_n)$ be a basis of $\mathfrak{g}$. Then let $\lambda_1, \ldots, \lambda_n$ be a dual basis of $\mathfrak{g}^*$. There exist vectors $x^1, \ldots, x^n$ in $\mathfrak{g}$ such that $\lambda_j = f_{x^j}$ for all j. Explicitly, we have $\delta_{ij} = \lambda_i(x_j) = f_{x^i}(x_j) = B_\pi(x^i, x_j)$, for all i, j. We call $(x^1, \ldots, x^n)$ a *dual basis to* $(x_1, \ldots, x_n)$ *with respect to the nondegenerate form* B_π.

Note that for each i,

$$x^i = \sum_{j=1}^n b^{ij} x_j. \tag{A.6}$$

In fact, for each k,

$$B_\pi \left(\sum_{j=1}^n b^{ij} x_j, x_k \right) = \sum_{j=1}^n b^{ij}\, b_{jk} = \delta_{ik} = B(x^i, x_k).$$

Since B_π is nondegenerate, this establishes (A.6). Thus the Casimir operator (A.5) can be written as

$$\sum_{i=1}^n \pi(x_i) \circ \pi(x^i) \tag{A.7}$$

The next lemma says that the Casimir operator is a $\mathfrak{g}$-equivariant linear operator on V.

Lemma A.2.3. *The Casimir operator commutes with all operators $\pi(x)$; that is,*

$$\pi(x) \circ \Omega_\pi = \Omega_\pi \circ \pi(x) \tag{A.8}$$

for all $x \in \mathfrak{g}$.

Proof. Fix dual bases $(x_1, \ldots, x_n)$ and $(x^1, \ldots, x^n)$ of $\mathfrak{g}$ with respect to B_π. For each $x \in \mathfrak{g}$, we have $[x, x_j] = \sum_{i=1}^n c_{ij}(x)\, x_i$ and $[x.x^j] = \sum_{i=1}^n c^{ij}(x)\, x^i$, for some constants $c_{ij}(x)$ and $c^{ij}(x)$. We can relate these constants using the invariance of B_π (see Lemma 6.3.6):

$$B_\pi([x, y], z) = -B_\pi(y, [x, z])$$

for all x, y, $z \in \mathfrak{g}$. Thus

$$\begin{aligned} c_{ij}(x) &= B([x, x_j], x^i) \\ &= -B(x_j, [x, x^i]) \\ &= -c^{ji}(x). \end{aligned}$$

Hence, using Leibniz' rule for the commutator bracket, we obtain

$$\begin{aligned} [\pi(x), \Omega_\pi] &= \left[x\ ,\ \sum_{j=1}^n \pi(x_j) \circ \pi(x^j)\right] \\ &= \sum_{j=1}^n \big([\pi(x), \pi(x_j)] \circ \pi(x^j) + \pi(x_j) \circ [\pi(x), \pi(x^j)]\big) \\ &= \sum_{j=1}^n \big(\pi[x, x_j] \circ \pi(x^j) + \pi(x_j) \circ \pi[x, x^j]\big) \\ &= \sum_{j=1}^n \left(\sum_{i=1}^n c_{ij}(x)\, \pi(x_i) \circ \pi(x^j) + \sum_{i=1}^n c^{ij}(x)\, \pi(x_j) \circ \pi(x^i)\right) \\ &= \sum_{i,j} c_{ij}(x)\, \pi(x_i) \circ \pi(x^j) + \sum_{i,j} c^{ij}(x)\, \pi(x_j) \circ \pi(x^i) \\ &= \sum_{i,j} c_{ij}(x)\, \pi(x_i) \circ \pi(x^j) + \sum_{i,j} c^{ji}(x)\, \pi(x_i) \circ \pi(x^j) \\ &= \sum_{i,j} (c_{ij}(x) + c^{ji}(x))\, \pi(x_i) \circ \pi(x^j) \\ &= 0. \end{aligned}$$

This shows that $\pi(x) \circ \Omega_\pi = \Omega_\pi \circ \pi(x)$. ∎

Corollary A.2.4. *Let π be an irreducible representation of a complex Lie algebra $\mathfrak{g}$ on a complex vector space V. Then $\Omega_\pi = (\dim \mathfrak{g} / \dim V)\, I_V$.*

Proof. By Schur's Lemma (Theorem 7.3.10) and Lemma A.2.3, Ω_π is a scalar operator: $\Omega_\pi = \lambda\, I_V$, for some $\lambda \in \mathbb{C}$. Let $m = \dim V$. Then, using the notation of Lemma A.2.3,

$$\begin{aligned}
m\lambda &= \operatorname{tr}(\Omega_\pi) \\
&= \operatorname{tr}\left(\sum_{j=1}^{n} \pi(x_j)\circ \pi(x^j)\right) \\
&= \sum_{j=1}^{n} B_\pi(x_j, x^j) \\
&= n.
\end{aligned}$$

Hence $\lambda = n/m$. □

Lemma A.2.5. *Suppose that π is a representation of a semisimple Lie algebra $\mathfrak{g}$ over $\mathbb{F}$ on a vector space over $\mathbb{F}$. Then $tr(\pi(x)) = 0$ for all $x \in \mathfrak{g}$. In particular, if V is one-dimensional, then π is the trivial representation.*

Proof. The lemma follows immediately from the fact that $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ and the fact that $\operatorname{tr}([\pi(x), \pi(y)]) = 0$ for all x, $y \in \mathfrak{g}$. □

A.3 The Proof of Weyl's Theorem

We are now ready to prove Weyl's Theorem on complete reducibility.

Proof of Theorem 7.3.7: We are assuming that $\mathfrak{g}$ is a nonzero semisimple Lie algebra over $\mathbb{F}$, V is a vector space over $\mathbb{F}$, and that π is a representation of $\mathfrak{g}$ on V. Suppose that W is a $\mathfrak{g}$-invariant subspace of V. We want to produce a *complementary* $\mathfrak{g}$-invariant subspace; i.e., a $\mathfrak{g}$-invariant subspace U such that $V = W \oplus U$.

Of course, if $W = \{0\}$ or $W = V$, there is nothing to prove, so let us assume that $\{0\} \subsetneq W \subsetneq V$.

We will tackle the simplest case first. Suppose that W has codimension one (i.e., $\dim W = \dim V - 1$) and is irreducible. Let Ω_π be the Casimir operator corresponding to π. Since W is invariant under $\pi(x)$ for all $x \in \mathfrak{g}$, we see that W is Ω_π-invariant. Then, since W is irreducible and Ω_π commutes with all $\pi(x)$, it follows that $\Omega_\pi|_W = \lambda\, I_W$, where $\lambda = \dim\mathfrak{g}/\dim W$ is a nonzero constant.

By Lemma A.2.5, the quotient representation π' of $\mathfrak{g}$ on the one-dimensional vector space V/W is trivial. Hence $\pi'(x)\,(v + W) = W$ for all $x \in \mathfrak{g}$ and all $v \in W$. This implies that $\pi(x)\,(V) \subset W$ for all $x \in \mathfrak{g}$, and therefore $\Omega_\pi(V) \subset W$. Since the restriction $\Omega_\pi|W$ is given by multiplication by a nonzero scalar, we have in fact $\Omega_\pi\,(V) = W$.

It follows that $\ker \Omega_\pi$ is a one-dimensional subspace of V, and that $V = W \oplus \ker \Omega_\pi$. Since Ω_π commutes with all $\pi(x)$, we see that $\ker \Omega_\pi$ is a $\mathfrak{g}$-invariant subspace; it is thus the complementary $\mathfrak{g}$-invariant subspace to W.

We now consider the next-simplest case. Suppose that W is a $\mathfrak{g}$-invariant codimension one subspace of V. We no longer assume that W is irreducible. In this case, we will use induction on $\dim W$ to obtain a $\mathfrak{g}$-invariant complementary subspace to W.

If $\dim W = 1$ (so that $\dim V = 2$), then certainly W is irreducible, so the previous case ensures the existence of a complementary $\mathfrak{g}$-invariant subspace. So assume that $\dim W = n > 1$, and that Weyl's Theorem holds for all invariant codimension one subspaces of dimension $< n$.

Of course, if W is irreducible, we're back to the previous case, so assume that W is not irreducible. Then W has an invariant subspace U such that $\{0\} \subsetneq U \subsetneq W$. We can choose U to have minimal positive dimension, so that U is irreducible.

The quotient module W/U is a codimension one submodule of the quotient module V/U. Since $\dim W/U < n$, the induction hypothesis applies, and there thus exists a complementary $\mathfrak{g}$-invariant subspace $\mathbb{F}(v + U)$ to W/U in V/U. Since $\mathbb{F}(v+U)$ is one-dimensional, $\mathfrak{g}$ acts trivially on it. Hence $\pi(x)\, v \in U$ for all $x \in \mathfrak{g}$. The subspace $U_1 = U \oplus \mathbb{F}v$ is therefore $\mathfrak{g}$-invariant, and U is a codimension one invariant irreducible subspace of U_1. We can therefore apply the previous case to conclude that there exists a nonzero vector $u_1 \in U_1$ such that $\mathbb{F}u_1$ is $\mathfrak{g}$-invariant, and $U_1 = U \oplus \mathbb{F}u_1$. By Lemma A.2.5, $\pi(x)\, u_1 = 0$ for all $x \in \mathfrak{g}$. Note that u_1 cannot be in W: if $u_1 \in W$, then $U_1 \subset W$, so $v \in W$, and hence $\mathbb{F}(v + U)$ cannot be a complementary subspace to W/U. It follows that $V = W \oplus \mathbb{F}u_1$, and the complementary $\mathfrak{g}$-invariant subspace to W is $\mathbb{F}u_1$.

We can finally tackle the general case. Suppose that π is a representation of $\mathfrak{g}$ on V and that W is a $\mathfrak{g}$-invariant subspace of V. We can assume that $\{0\} \subsetneq W \subsetneq V$. Let π_W denote the sub-representation $x \mapsto \pi(x)|_W$ of $\mathfrak{g}$ on W. Then let σ be the representation of $\mathfrak{g}$ on $\mathcal{L}(V, W)$ given by Proposition A.1.3:

$$\sigma(x)\,(T) = \pi_W(x)T - T\pi(x) \qquad\qquad (T \in \mathcal{L}(V, W),\ x \in \mathfrak{g}) \tag{A.9}$$

Now let $\mathcal{S}$ denote the subspace of $\mathcal{L}(V, W)$ consisting of all $T \in \mathcal{L}(V, W)$ such that $T|_W = \lambda\, I_W$, for some $\lambda \in \mathbb{F}$. Then let $\mathcal{T}$ be the subspace of $\mathcal{S}$ consisting of those $T \in \mathcal{L}(V, W)$ such that $T|_W = 0$.

Since W is a $\mathfrak{g}$-invariant subspace, (A.9) shows that $\sigma(x)\, \mathcal{S} \subset \mathcal{T}$, and hence $\mathcal{S}$ is invariant under $\sigma(x)$, for all $x \in \mathfrak{g}$. Let $S_0 \in \mathcal{L}(V, W)$ be any linear map such that $S_0|_W = I_W$. (Such exist.) Then $\mathcal{S} = \mathcal{T} \oplus \mathbb{F}\, S_0$, and so $\mathcal{T}$ is a codimension one invariant subspace of $\mathcal{S}$. Thus, by the codimension one case, $\mathcal{S}$ has a complementary $\mathfrak{g}$-invariant subspace to $\mathcal{T}$:

$$\mathcal{S} = \mathcal{T} \oplus \mathbb{F}T_0,$$

for some $T_0 \in \mathcal{S}$. By scaling T_0, we may as well assume that $T_0|_W = I_W$. By Lemma A.2.5, $\sigma(x)(T_0) = 0$ for all $x \in \mathfrak{g}$. This imples that T_0 is a linear mapping from V to W which intertwines π and π_W. Let $U = \ker T_0$. Then U is a $\mathfrak{g}$-invariant subspace of V such that $U \cap W = \{0\}$; since T_0 is the identity map on W, the range $T_0(V)$ equals W, and thus $\dim U = \dim V - \dim W$. Hence $V = W \oplus U$, and U is our complementary $\mathfrak{g}$-invariant subspace. $\square$

Bibliography

[1] S. Axler *Linear Algebra Done Right*, 2nd ed., Springer, New York, 1997.

[2] K. Erdman and M. Wildon, *Introduction to Lie Algebras*, 2nd ed., Springer, New York, 2007.

[3] B. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, 2nd. ed, Springer, New York 2004.

[4] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, American Math. Society, Providence, 2004.

[5] J. Humphreys, *Introduction to Lie Algebras and Representation Theory*, 6th ed., Springer-Verlag, Berlin and Heidelberg, 1997.

[6] N. Jacobson, *Lie Algebras*, Dover Publications, New York, 1979.

[7] J.P. Serre, *Complex Semisimple Lie Algebras*, Springer, New York, 2001.

[8] R.D. Schafer, *An Introduction to Non-associative algebras*, Academic Press, New York, 1966.

[9] E.B. Schenkmann, *A theory of subinvariant Lie algebras*, American Journal of Mathematics 73 (1951), pp. 453-474.

[10] J.P. Serre, *Lie Algebras and Lie Groups*, 2nd. ed., Springer-Verlag, 1992.

[11] Wikipedia, *The Chinese Remainder Theorem.*

[12] Wikipedia, *The Heisenberg Group.*