Chapter 2
Structure of Classical Groups

Abstract In this chapter we study the structure of a classical group G and its Lie algebra. We choose a matrix realization of G such that the diagonal subgroup $H \subset G$ is a *maximal torus*; by elementary linear algebra every conjugacy class of semisimple elements intersects H. Using the unipotent elements in G, we show that the groups $\mathbf{GL}(n,\mathbb{C})$, $\mathbf{SL}(n,\mathbb{C})$, $\mathbf{SO}(n,\mathbb{C})$, and $\mathbf{Sp}(n,\mathbb{C})$ are connected (as Lie groups and as algebraic groups). We examine the group $\mathbf{SL}(2,\mathbb{C})$, find its irreducible representations, and show that every regular representation decomposes as the direct sum of irreducible representations. This group and its Lie algebra play a basic role in the structure of the other classical groups and Lie algebras. We decompose the Lie algebra of a classical group under the adjoint action of a maximal torus and find the invariant subspaces (called *root spaces*) and the corresponding characters (called *roots*). The commutation relations of the root spaces are encoded by the set of roots; we use this information to prove that the classical (trace-zero) Lie algebras are simple (or semisimple). In the final section of the chapter we develop some general Lie algebra methods (solvable Lie algebras, Killing form) and show that every semisimple Lie algebra has a root-space decomposition with the same properties as those of the classical Lie algebras.

2.1 Semisimple Elements

A semisimple matrix can be diagonalized, relative to a suitable basis. In this section we show that a maximal set of mutually commuting semisimple elements in a classical group can be simultaneously diagonalized by an element of the group.

R. Goodman, N.R. Wallach, *Symmetry, Representations, and Invariants*,
Graduate Texts in Mathematics 255, DOI 10.1007/978-0-387-79852-3_2,

2.1.1 Toral Groups

Recall that an (algebraic) torus is an algebraic group T isomorphic to $(\mathbb{C}^\times)^l$; the integer l is called the *rank* of T. The rank is uniquely determined by the algebraic group structure of T (this follows from Lemma 2.1.2 below or Exercises 1.4.5 #1).

Definition 2.1.1. A *rational character* of a linear algebraic group K is a regular homomorphism $\chi : K \longrightarrow \mathbb{C}^\times$.

The set $\mathcal{X}(K)$ of rational characters of K has the natural structure of an abelian group with $(\chi_1\chi_2)(k) = \chi_1(k)\chi_2(k)$ for $k \in K$. The identity element of $\mathcal{X}(K)$ is the *trivial character* $\chi_0(k) = 1$ for all $k \in K$.

Lemma 2.1.2. *Let T be an algebraic torus of rank l. The group $\mathcal{X}(T)$ is isomorphic to $\mathbb{Z}^l$. Furthermore, $\mathcal{X}(T)$ is a basis for $\mathcal{O}[T]$ as a vector space over $\mathbb{C}$.*

Proof. We may assume that $T = (\mathbb{C}^\times)^l$ with coordinate functions $x_1, \ldots, x_l$. Thus $\mathcal{O}[T] = \mathbb{C}[x_1, \ldots, x_l, x_l^{-1}, \ldots, x_l^{-1}]$. For $t = [x_1(t), \ldots, x_l(t)] \in T$ and $\lambda = [p_1, \ldots, p_l] \in \mathbb{Z}^l$ we set

$$t^\lambda = \prod_{k=1}^{l} x_k(t)^{p_k} \,. \tag{2.1}$$

Then $t \mapsto t^\lambda$ is a rational character of T, which we will denote by χ_λ. Since $t^{\lambda+\mu} = t^\lambda t^\mu$ for $\lambda, \mu \in \mathbb{Z}^l$, the map $\lambda \mapsto \chi_\lambda$ is an injective group homomorphism from $\mathbb{Z}^l$ to $\mathcal{X}(T)$. Clearly, the set of functions $\{\chi_\lambda : \lambda \in \mathbb{Z}^l\}$ is a basis for $\mathcal{O}[T]$ as a vector space over $\mathbb{C}$.

Conversely, let χ be a rational character of T. Then for $k = 1, \ldots, l$ the function

$$z \mapsto \varphi_k(z) = \chi(1, \ldots, z, \ldots, 1) \quad (z \text{ in } k\text{th coordinate})$$

is a one-dimensional regular representation of $\mathbb{C}^\times$. From Lemma 1.6.4 we have $\varphi_k(z) = z^{p_k}$ for some $p_k \in \mathbb{Z}$. Hence

$$\chi(x_1, \ldots, x_l) = \prod_{k=1}^{l} \varphi_k(x_k) = \chi_\lambda(x_1, \ldots, x_l) \,,$$

where $\lambda = [p_1, \ldots, p_l]$. Thus every rational character of T is of the form χ_λ for some $\lambda \in \mathbb{Z}^l$. $\square$

Proposition 2.1.3. *Let T be an algebraic torus. Suppose (ρ, V) is a regular representation of T. Then there exists a finite subset $\Psi \subset \mathcal{X}(T)$ such that*

$$V = \bigoplus_{\chi \in \Psi} V(\chi) \,, \tag{2.2}$$

where $V(\chi) = \{v \in V : \rho(t)v = \chi(t)v \quad \text{for all } t \in T\}$ is the χ weight space *of T on V. If $g \in \mathrm{End}(V)$ commutes with $\rho(t)$ for all $t \in T$, then $gV(\chi) \subset V(\chi)$.*

Proof. Since $(\mathbb{C}^\times)^l \cong \mathbb{C}^\times \times (\mathbb{C}^\times)^{l-1}$, the existence of the decomposition (2.2) follows from Lemma 1.6.4 by induction on l. The last statement is clear from the definition of $V(\chi)$. □

Lemma 2.1.4. *Let T be an algebraic torus. Then there exists an element $t \in T$ with the following property: If $f \in \mathcal{O}[T]$ and $f(t^n) = 0$ for all $n \in \mathbb{Z}$, then $f = 0$.*

Proof. We may assume $T = (\mathbb{C}^\times)^l$. Choose $t \in T$ such that its coordinates $t_i = x_i(t)$ satisfy

$$t_1^{p_1} \cdots t_l^{p_l} \neq 1 \quad \text{for all } (p_1, \ldots, p_l) \in \mathbb{Z}^l \setminus \{0\} . \tag{2.3}$$

This is always possible; for example we can take $t_1, \ldots, t_l$ to be algebraically independent over the rationals.

Let $f \in \mathcal{O}[T]$ satisfy $f(t^n) = 0$ for all $n \in \mathbb{Z}$. Replacing f by $(x_1 \cdots x_l)^r f$ for a suitably large r, we may assume that

$$f = \sum_{|K| \leq p} a_K x^K$$

for some positive integer p, where the exponents K are in $\mathbb{N}^l$. Since $f(t^n) = 0$ for all $n \in \mathbb{Z}$, the coefficients $\{a_K\}$ satisfy the equations

$$\sum_K a_K (t^K)^n = 0 \quad \text{for all } n \in \mathbb{Z} . \tag{2.4}$$

We claim that the numbers $\{t^K : K \in \mathbb{N}^l\}$ are all distinct. Indeed, if $t^K = t^L$ for some $K, L \in \mathbb{N}^l$ with $K \neq L$, then $t^P = 1$, where $P = K - L \neq 0$, which would violate (2.3). Enumerate the coefficients a_K of f as $b_1, \ldots, b_r$ and the corresponding character values t^K as $y_1, \ldots, y_r$. Then (2.4) implies that

$$\sum_{j=1}^{r} b_j (y_j)^n = 0 \qquad \text{for } n = 0, 1, \ldots, r-1 .$$

We view these equations as a homogeneous linear system for $b_1, \ldots, b_r$. The coefficient matrix is the $r \times r$ *Vandermonde matrix:*

$$V_r(y) = \begin{bmatrix} y_1^{r-1} & y_1^{r-2} & \cdots & y_1 & 1 \\ y_2^{r-1} & y_2^{r-2} & \cdots & y_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_r^{r-1} & y_r^{r-2} & \cdots & y_r & 1 \end{bmatrix} .$$

The determinant of this matrix is the *Vandermonde determinant* $\prod_{1 \leq i < j \leq r} (y_i - y_j)$ (see Exercises 2.1.3). Since $y_i \neq y_j$ for $i \neq j$, the determinant is nonzero, and hence $b_K = 0$ for all K. Thus $f = 0$. □

2.1.2 Maximal Torus in a Classical Group

If G is a linear algebraic group, then a torus $H \subset G$ is *maximal* if it is not contained in any larger torus in G. When G is one of the classical linear algebraic groups $\mathbf{GL}(n,\mathbb{C})$, $\mathbf{SL}(n,\mathbb{C})$, $\mathbf{Sp}(\mathbb{C}^n,\Omega)$, $\mathbf{SO}(\mathbb{C}^n,B)$ (where Ω is a nondegenerate skew-symmetric bilinear form and B is a nondegenerate symmetric bilinear form) we would like the subgroup H of diagonal matrices in G to be a maximal torus. For this purpose we make the following choices of B and Ω:

We denote by s_l the $l \times l$ matrix

$$s_l = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{2.5}$$

with 1 on the skew diagonal and 0 elsewhere. Let $n = 2l$ be even, set

$$J_+ = \begin{bmatrix} 0 & s_l \\ s_l & 0 \end{bmatrix}, \qquad J_- = \begin{bmatrix} 0 & s_l \\ -s_l & 0 \end{bmatrix},$$

and define the bilinear forms

$$B(x,y) = x^t J_+ y, \qquad \Omega(x,y) = x^t J_- y \quad \text{for } x,y \in \mathbb{C}^n . \tag{2.6}$$

The form B is nondegenerate and *symmetric*, and the form Ω is nondegenerate and *skew-symmetric*. From equation (1.8) we calculate that the Lie algebra $\mathfrak{so}(\mathbb{C}^{2l},B)$ of $\mathbf{SO}(\mathbb{C}^{2l},B)$ consists of all matrices

$$A = \begin{bmatrix} a & b \\ c & -s_l a^t s_l \end{bmatrix}, \qquad \begin{cases} a,\, b,\, c \in M_l(\mathbb{C}) , \\ b^t = -s_l b s_l , \quad c^t = -s_l c s_l \end{cases} \tag{2.7}$$

(thus b and c are *skew-symmetric* around the skew diagonal). Likewise, the Lie algebra $\mathfrak{sp}(\mathbb{C}^{2l},\Omega)$ of $\mathbf{Sp}(\mathbb{C}^{2l},\Omega)$ consists of all matrices

$$A = \begin{bmatrix} a & b \\ c & -s_l a^t s_l \end{bmatrix}, \qquad \begin{cases} a,\, b,\, c \in M_l(\mathbb{C}) , \\ b^t = s_l b s_l , \quad c^t = s_l c s_l \end{cases} \tag{2.8}$$

(b and c are *symmetric* around the skew diagonal).

Finally, we consider the orthogonal group on $\mathbb{C}^n$ when $n = 2l+1$ is odd. We take the symmetric bilinear form

$$B(x,y) = \sum_{i+j=n+1} x_i y_j \quad \text{for } x,y \in \mathbb{C}^n . \tag{2.9}$$

We can write this form as $B(x,y) = x^t S y$, where the $n \times n$ symmetric matrix $S = s_{2l+1}$ has block form

$$S = \begin{bmatrix} 0 & 0 & s_l \\ 0 & 1 & 0 \\ s_l & 0 & 0 \end{bmatrix} .$$

Writing the elements of $M_n(\mathbb{C})$ in the same block form and making a matrix calculation from equation (1.8), we find that the Lie algebra $\mathfrak{so}(\mathbb{C}^{2l+1}, B)$ of $\mathbf{SO}(\mathbb{C}^{2l+1}, B)$ consists of all matrices

$$A = \begin{bmatrix} a & w & b \\ u^t & 0 & -w^t s_l \\ c & -s_l u & -s_l a^t s_l \end{bmatrix} , \qquad \begin{cases} a, b, c \in M_l(\mathbb{C}) , \\ b^t = -s_l b s_l , \quad c^t = -s_l c s_l , \\ \text{and} \quad u, w \in \mathbb{C}^l . \end{cases} \tag{2.10}$$

Suppose now that G is $\mathbf{GL}(n,\mathbb{C})$, $\mathbf{SL}(n,\mathbb{C})$, $\mathbf{Sp}(\mathbb{C}^n,\Omega)$, or $\mathbf{SO}(\mathbb{C}^n,B)$ with Ω and B chosen as above. Let H be the subgroup of diagonal matrices in G; write $\mathfrak{g} = \mathrm{Lie}(G)$ and $\mathfrak{h} = \mathrm{Lie}(H)$. By Example 1 of Section 1.4.3 and (1.39) we know that $\mathfrak{h}$ consists of all diagonal matrices that are in $\mathfrak{g}$. We have the following case-by-case description of H and $\mathfrak{h}$:

1. When $G = \mathbf{SL}(l+1,\mathbb{C})$ (we say that G is of **type $\mathbf{A}_\ell$**), then

$$H = \{\mathrm{diag}[x_1, \ldots, x_l, (x_1 \cdots x_l)^{-1}] : x_i \in \mathbb{C}^\times\} ,$$
$$\mathrm{Lie}(H) = \{\mathrm{diag}[a_1, \ldots, a_{l+1}] : a_i \in \mathbb{C}, \quad \textstyle\sum_i a_i = 0\} .$$

2. When $G = \mathbf{Sp}(\mathbb{C}^{2l},\Omega)$ (we say that G is of **type $\mathbf{C}_\ell$**) or $G = \mathbf{SO}(\mathbb{C}^{2l},B)$ (we say that G is of **type $\mathbf{D}_\ell$**), then by (2.7) and (2.8),

$$H = \{\mathrm{diag}[x_1, \ldots, x_l, x_l^{-1}, \ldots, x_1^{-1}] : x_i \in \mathbb{C}^\times\} ,$$
$$\mathfrak{h} = \{\mathrm{diag}[a_1, \ldots, a_l, -a_l, \ldots, -a_1] : a_i \in \mathbb{C}\} .$$

3. When $G = \mathbf{SO}(\mathbb{C}^{2l+1},B)$ (we say that G is of **type $\mathbf{B}_\ell$**), then by (2.10),

$$H = \{\mathrm{diag}[x_1, \ldots, x_l, 1, x_l^{-1}, \ldots, x_1^{-1}] : x_i \in \mathbb{C}^\times\} ,$$
$$\mathfrak{h} = \{\mathrm{diag}[a_1, \ldots, a_l, 0, -a_l, \ldots, -a_1] : a_i \in \mathbb{C}\} .$$

In all cases H is isomorphic as an algebraic group to the product of l copies of $\mathbb{C}^\times$, so it is a torus of rank l. The Lie algebra $\mathfrak{h}$ is isomorphic to the vector space $\mathbb{C}^l$ with all Lie brackets zero. Define coordinate functions $x_1, \ldots, x_l$ on H as above. Then $\mathcal{O}[H] = \mathbb{C}[x_1, \ldots, x_l, x_1^{-1}, \ldots, x_l^{-1}]$.

Theorem 2.1.5. *Let G be $\mathbf{GL}(n,\mathbb{C})$, $\mathbf{SL}(n,\mathbb{C})$, $\mathbf{SO}(\mathbb{C}^n,B)$ or $\mathbf{Sp}(\mathbb{C}^{2l},\Omega)$ in the form given above, where H is the diagonal subgroup in G. Suppose $g \in G$ and $gh = hg$ for all $h \in H$. Then $g \in H$.*

Proof. We have $G \subset \mathbf{GL}(n,\mathbb{C})$. An element $h \in H$ acts on the standard basis $\{e_1, \ldots, e_n\}$ for $\mathbb{C}^n$ by $he_i = \theta_i(h)e_i$. Here the characters θ_i are given as follows in terms of the coordinate functions $x_1, \ldots, x_l$ on H:

1. $G = \mathbf{GL}(l,\mathbb{C})$: $\quad \theta_i = x_i$ for $i = 1, \ldots, l$.

2. $G = \mathbf{SL}(l+1,\mathbb{C})$: $\theta_i = x_i$ for $i = 1,\ldots,l$ and $\theta_{l+1} = (x_1 \cdots x_l)^{-1}$.
3. $G = \mathbf{SO}(\mathbb{C}^{2l},B)$ or $\mathbf{Sp}(\mathbb{C}^{2l},\Omega)$: $\theta_i = x_i$ and $\theta_{2l+1-i} = x_i^{-1}$ for $i = 1,\ldots,l$.
4. $G = \mathbf{SO}(\mathbb{C}^{2l+1},B)$: $\theta_i = x_i$, $\theta_{2l+2-i} = x_i^{-1}$ for $i = 1,\ldots,l$, and $\theta_{l+1} = 1$.

Since the characters $\theta_1,\ldots,\theta_n$ are all distinct, the weight space decomposition (2.2) of $\mathbb{C}^n$ under H is given by the one-dimensional subspaces $\mathbb{C}e_i$. If $gh = hg$ for all $h \in H$, then g preserves the weight spaces and hence is a diagonal matrix. □

Corollary 2.1.6. *Let G and H be as in Theorem* 2.1.5. *Suppose $T \subset G$ is an abelian subgroup (not assumed to be algebraic). If $H \subset T$ then $H = T$. In particular, H is a maximal torus in G.*

The choice of the maximal torus H depended on choosing a particular matrix form of G. We shall prove that if T is any maximal torus in G then there exists an element $\gamma \in G$ such that $T = \gamma H \gamma^{-1}$. We begin by conjugating individual semisimple elements into H.

Theorem 2.1.7. (Notation as in Theorem 2.1.5) *If $g \in G$ is semisimple then there exists $\gamma \in G$ such that $\gamma g \gamma^{-1} \in H$.*

Proof. When G is $\mathbf{GL}(n,\mathbb{C})$ or $\mathbf{SL}(n,\mathbb{C})$, let $\{v_1,\ldots,v_n\}$ be a basis of eigenvectors for g and define $\gamma v_i = e_i$, where $\{e_i\}$ is the standard basis for $\mathbb{C}^n$. Multiplying v_1 by a suitable constant, we can arrange that $\det\gamma = 1$. Then $\gamma \in G$ and $\gamma g \gamma^{-1} \in H$.

If $g \in \mathbf{SL}(n,\mathbb{C})$ is semisimple and preserves a nondegenerate bilinear form ω on C^n, then there is an eigenspace decomposition

$$\mathbb{C}^n = \bigoplus V_\lambda, \quad gv = \lambda v \quad \text{for } v \in V_\lambda . \tag{2.11}$$

Furthermore, $\omega(u,v) = \omega(gu,gv) = \lambda\mu\,\omega(u,v)$ for $u \in V_\lambda$ and $v \in V_\mu$. Hence

$$\omega(V_\lambda,V_\mu) = 0 \quad \text{if } \lambda\mu \neq 1 . \tag{2.12}$$

Since ω is nondegenerate, it follows from (2.11) and (2.12) that

$$\dim V_{1/\mu} = \dim V_\mu . \tag{2.13}$$

Let $\mu_1,\ldots,\mu_k$ be the (distinct) eigenvalues of g that are not ± 1. From (2.13) we see that $k = 2r$ is even and that we can take $\mu_i^{-1} = \mu_{r+i}$ for $i = 1,\ldots,r$.

Recall that a subspace $W \subset \mathbb{C}^n$ is ω *isotropic* if $\omega(u,v) = 0$ for all $u,v \in W$ (see Appendix B.2.1). By (2.12) the subspaces V_{μ_i} and V_{1/μ_i} are ω isotropic and the restriction of ω to $V_{\mu_i} \times V_{1/\mu_i}$ is nondegenerate. Let $W_i = V_{\mu_i} \oplus V_{1/\mu_i}$ for $i = 1,\ldots,r$. Then

(a) the subspaces V_1, V_{-1}, and W_i are mutually orthogonal relative to the form ω, and the restriction of ω to each of these subspaces is nondegenerate;
(b) $\mathbb{C}^n = V_1 \oplus V_{-1} \oplus W_1 \oplus \cdots \oplus W_r$;
(c) $\det g = (-1)^k$, where $k = \dim V_{-1}$.

Now suppose $\omega = \Omega$ is the skew-symmetric form (2.6) and $g \in \mathbf{Sp}(\mathbb{C}^{2l}, \Omega)$. From (a) we see that $\dim V_1$ and $\dim V_{-1}$ are even. By Lemma 1.1.5 we can find canonical symplectic bases in each of the subspaces in decomposition (b); in the case of W_i we may take a basis $v_1, \ldots, v_s$ for V_{μ_i} and an Ω-dual basis $v_{-1}, \ldots, v_{-s}$ for V_{1/μ_i}. Altogether, these bases give a canonical symplectic basis for $\mathbb{C}^{2l}$. We may enumerate it as $v_1, \ldots, v_l$, $v_{-1}, \ldots, v_{-l}$, so that

$$gv_i = \lambda_i v_i\ , \quad gv_{-i} = \lambda_i^{-1} v_{-i} \quad \text{for } i = 1, \ldots, l\ .$$

The linear transformation γ such that $\gamma v_i = e_i$ and $\gamma v_{-i} = e_{2l+1-i}$ for $i = 1, \ldots, l$ is in G due to the choice (2.6) of the matrix for Ω. Furthermore,

$$\gamma g \gamma^{-1} = \mathrm{diag}[\lambda_1, \ldots, \lambda_l, \lambda_l^{-1}, \ldots, \lambda_1^{-1}] \in H\ .$$

This proves the theorem in the symplectic case.

Now assume that G is the orthogonal group for the form B in (2.6) or (2.9). Since $\det g = 1$, we see from (c) that $\dim V_{-1} = 2q$ is even, and by (2.13) $\dim W_i$ is even. Hence n is odd if and only if $\dim V_1$ is odd. Just as in the symplectic case, we construct canonical B-isotropic bases in each of the subspaces in decomposition (b) (see Section B.2.1); the union of these bases gives an isotropic basis for $\mathbb{C}^n$. When $n = 2l$ and $\dim V_1 = 2r$ we can enumerate this basis so that

$$gv_i = \lambda_i v_i\ , \quad gv_{-i} = \lambda_i^{-1} v_{-i} \quad \text{for } i = 1, \ldots, l\ .$$

The linear transformation γ such that $\gamma v_i = e_i$ and $\gamma v_{-i} = e_{n+1-i}$ is in $\mathbf{O}(\mathbb{C}^n, B)$, and we can interchange v_l and v_{-l} if necessary to get $\det \gamma = 1$. Then

$$\gamma g \gamma^{-1} = \mathrm{diag}[\lambda_1, \ldots, \lambda_l, \lambda_l^{-1}, \ldots, \lambda_1^{-1}] \in H\ .$$

When $n = 2l + 1$ we know that $\lambda = 1$ occurs as an eigenvalue of g, so we can enumerate this basis so that

$$gv_0 = v_0\ , \quad gv_i = \lambda_i v_i\ , \quad gv_{-i} = \lambda_i^{-1} v_{-i} \quad \text{for } i = 1, \ldots, l\ .$$

The linear transformation γ such that $\gamma v_0 = e_{l+1}$, $\gamma v_i = e_i$, and $\gamma v_{-i} = e_{n+1-i}$ is in $\mathbf{O}(\mathbb{C}^n, B)$. Replacing γ by $-\gamma$ if necessary, we have $\gamma \in \mathbf{SO}(\mathbb{C}^n, B)$ and

$$\gamma g \gamma^{-1} = \mathrm{diag}[\lambda_1, \ldots, \lambda_l, 1, \lambda_l^{-1}, \ldots, \lambda_1^{-1}] \in H\ .$$

This completes the proof of the theorem. □

Corollary 2.1.8. *If T is any torus in G, then there exists $\gamma \in G$ such that $\gamma T \gamma^{-1} \subset H$. In particular, if T is a maximal torus in G, then $\gamma T \gamma^{-1} = H$.*

Proof. Choose $t \in T$ satisfying the condition of Lemma 2.1.4. By Theorem 2.1.7 there exists $\gamma \in G$ such that $\gamma t \gamma^{-1} \in H$. We want to show that $\gamma x \gamma^{-1} \in H$ for all $x \in T$. To prove this, take any function $\varphi \in \mathfrak{I}_H$ and define a regular function f on T by $f(x) = \varphi(\gamma x \gamma^{-1})$. Then $f(t^p) = 0$ for all $p \in \mathbb{Z}$, since $\gamma t^p \gamma^{-1} \in H$. Hence

Lemma 2.1.4 implies that $f(x) = 0$ for all $x \in T$. Since φ was any function in $\mathcal{I}_H$, we conclude that $\gamma x \gamma^{-1} \in H$. If T is a maximal torus then so is $\gamma T \gamma^{-1}$. Hence $\gamma T \gamma^{-1} = H$ in this case. □

From Corollary 2.1.8, we see that the integer $l = \dim H$ does not depend on the choice of a particular maximal torus in G. We call l the *rank* of G.

2.1.3 Exercises

1. Verify that the Lie algebras of the orthogonal and symplectic groups are given in the matrix forms (2.7), (2.8), and (2.10).
2. Let $V_r(y)$ be the Vandermonde matrix, as in Section 2.1.2. Prove that

$$\det V_r(y) = \prod_{1 \le i < j \le r} (y_i - y_j) \ .$$

(HINT: Fix $y_2, \ldots, y_r$ and consider $\det V_r(y)$ as a polynomial in y_1. Show that it has degree $r-1$ with roots $y_2, \ldots, y_r$ and that the coefficient of y_1^{r-1} is the Vandermonde determinant for $y_2, \ldots, y_r$. Now use induction on r.)
3. Let H be a torus of rank n. Let $\mathcal{X}_*(H)$ be the set of all regular homomorphisms from $\mathbb{C}^\times$ into H. Define a group structure on $\mathcal{X}_*(H)$ by pointwise multiplication: $(\pi_1\pi_2)(z) = \pi_1(z)\pi_2(z)$ for $\pi_1, \pi_2 \in \mathcal{X}_*(H)$.
(a) Prove that $\mathcal{X}_*(H)$ is isomorphic to $\mathbb{Z}^n$ as an abstract group. (HINT: Use Lemma 1.6.4.)
(b) Prove that if $\pi \in \mathcal{X}_*(H)$ and $\chi \in \mathcal{X}(H)$ then there is an integer $\langle \pi, \chi \rangle \in \mathbb{Z}$ such that

$$\chi(\pi(z)) = z^{\langle \pi, \chi \rangle} \qquad \text{for all } z \in \mathbb{C}^\times \ .$$

(c) Show that the pairing $\pi, \chi \mapsto \langle \pi, \chi \rangle$ is additive in each variable (relative to the abelian group structures on $\mathcal{X}(H)$ and $\mathcal{X}_*(H)$) and is *nondegenerate* (this means that if $\langle \pi, \chi \rangle = 0$ for all χ then $\pi = 1$, and similarly for χ).
4. Let $G \subset \mathbf{GL}(n, \mathbb{C})$ be a classical group with Lie algebra $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{C})$ (for the orthogonal and symplectic groups use the bilinear forms (2.6) and (2.9)). Define $\theta(g) = (g^t)^{-1}$ for $g \in G$.
(a) Show that θ is a regular automorphism of G and that $d\theta(X) = -X^t$ for $X \in \mathfrak{g}$.
(b) Define $K = \{g \in G : \theta(g) = g\}$ and let $\mathfrak{k}$ be the Lie algebra of K. Show that $\mathfrak{k} = \{X \in \mathfrak{g} : d\theta(X) = X\}$.
(c) Define $\mathfrak{p} = \{X \in \mathfrak{g} : d\theta(X) = -X\}$. Show that $\mathrm{Ad}(K)\mathfrak{p} \subset \mathfrak{p}$, $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, $[\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}$, and $[\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$. (HINT: $d\theta$ is a derivation of $\mathfrak{g}$ and has eigenvalues ± 1.).
(d) Determine the explicit matrix form of $\mathfrak{k}$ and $\mathfrak{p}$ when $G = \mathbf{Sp}(\mathbb{C}^{2l}, \Omega)$, with Ω given by (2.6). Show that $\mathfrak{k}$ is isomorphic to $\mathfrak{gl}(l, \mathbb{C})$ in this case. (HINT: Write $X \in \mathfrak{g}$ in block form $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ and show that the map $X \mapsto A + iBs_l$ gives a Lie algebra isomorphism from $\mathfrak{k}$ to $\mathfrak{gl}(l, \mathbb{C})$.)

2.2 Unipotent Elements

Unipotent elements give an algebraic relation between a linear algebraic group and its Lie algebra, since they are exponentials of nilpotent elements and the exponential map is a polynomial function on nilpotent matrices. In this section we exploit this property to prove the connectedness of the classical groups.

2.2.1 Low-Rank Examples

We shall show that the classical groups $\mathbf{SL}(n,\mathbb{C})$, $\mathbf{SO}(n,\mathbb{C})$, and $\mathbf{Sp}(n,\mathbb{C})$ are generated by their unipotent elements. We begin with the basic case $G = \mathbf{SL}(2,\mathbb{C})$. Let $N^+ = \{u(z) : z \in \mathbb{C}\}$ and $N^- = \{v(z) : z \in \mathbb{C}\}$, where

$$u(z) = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad v(z) = \begin{bmatrix} 1 & 0 \\ z & 1 \end{bmatrix}.$$

The groups N^+ and N^- are isomorphic to the additive group of the field $\mathbb{C}$.

Lemma 2.2.1. *The group* $\mathbf{SL}(2,\mathbb{C})$ *is generated by* $N^+ \cup N^-$.

Proof. Let $g = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ with $ad - bc = 1$. If $a \neq 0$ we can use elementary row and column operations to factor

$$g = \begin{bmatrix} 1 & 0 \\ a^{-1}c & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix} \begin{bmatrix} 1 & a^{-1}b \\ 0 & 1 \end{bmatrix}.$$

If $a = 0$ then $c \neq 0$ and we can likewise factor

$$g = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} c & 0 \\ 0 & c^{-1} \end{bmatrix} \begin{bmatrix} 1 & c^{-1}d \\ 0 & 1 \end{bmatrix}.$$

Finally, we factor

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix},$$

$$\begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix} = \begin{bmatrix} 1 & -a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ (a^{-1}-1) & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ (a-1) & 1 \end{bmatrix},$$

to complete the proof. □

The orthogonal and symplectic groups of low rank are closely related to $\mathbf{GL}(1,\mathbb{C})$ and $\mathbf{SL}(2,\mathbb{C})$, as follows. Define a skew-symmetric bilinear form Ω on $\mathbb{C}^2$ by

$$\Omega(v,w) = \det[v, w],$$

where $[v,w] \in M_2(\mathbb{C})$ has columns v,w. We have $\det[e_1,e_1] = \det[e_2,e_2] = 0$ and $\det[e_1,e_2] = 1$, showing that the form Ω is nondegenerate. Since the determinant function is multiplicative, the form Ω satisfies

$$\Omega(gv,gw) = (\det g)\Omega(v,w) \quad \text{for } g \in \mathbf{GL}(2,\mathbb{C}) \;.$$

Hence g preserves Ω if and only if $\det g = 1$. This proves that $\mathbf{Sp}(\mathbb{C}^2,\Omega) = \mathbf{SL}(2,\mathbb{C})$.

Next, consider the group $\mathbf{SO}(\mathbb{C}^2,B)$ with B the bilinear form with matrix s_2 in (2.5). We calculate that

$$g^t s_2 g = \begin{bmatrix} 2ac & ad+bc \\ ad+bc & 2bd \end{bmatrix} \quad \text{for} \quad g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{SL}(2,\mathbb{C}) \;.$$

Since $ad - bc = 1$, it follows that $ad + bc = 2ad - 1$. Hence $g^t s_2 g = s_2$ if and only if $ad = 1$ and $b = c = 0$. Thus $\mathbf{SO}(\mathbb{C}^2,B)$ consists of the matrices

$$\begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix} \quad \text{for} \quad a \in \mathbb{C}^\times .$$

This furnishes an isomorphism $\mathbf{SO}(\mathbb{C}^2,B) \cong \mathbf{GL}(1,\mathbb{C})$.

Now consider the group $G = \mathbf{SO}(\mathbb{C}^3,B)$, where B is the bilinear form on $\mathbb{C}^3$ with matrix s_3 as in (2.5). From Section 2.1.2 we know that the subgroup

$$H = \{\mathrm{diag}[x,1,x^{-1}] \,:\, x \in \mathbb{C}^\times\}$$

of diagonal matrices in G is a maximal torus. Set $\widetilde{G} = \mathbf{SL}(2,\mathbb{C})$ and let

$$\widetilde{H} = \{\mathrm{diag}[x,x^{-1}] \,:\, x \in \mathbb{C}^\times\}$$

be the subgroup of diagonal matrices in $\widetilde{G}$.

We now define a homomorphism $\rho : \widetilde{G} \longrightarrow G$ that maps $\widetilde{H}$ onto H. Set

$$V = \{X \in M_2(\mathbb{C}) \,:\, \mathrm{tr}(X) = 0\}$$

and let $\widetilde{G}$ act on V by $\rho(g)X = gXg^{-1}$ (this is the adjoint representation of $\widetilde{G}$). The symmetric bilinear form

$$\omega(X,Y) = \frac{1}{2}\,\mathrm{tr}(XY)$$

is obviously invariant under $\rho(\widetilde{G})$, since $\mathrm{tr}(XY) = \mathrm{tr}(YX)$ for all $X,Y \in M_n(\mathbb{C})$. The basis

$$v_0 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad v_1 = \begin{bmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}, \quad v_{-1} = \begin{bmatrix} 0 & 0 \\ \sqrt{2} & 0 \end{bmatrix}$$

for V is ω isotropic. We identify V with $\mathbb{C}^3$ via the map $v_1 \mapsto e_1, v_0 \mapsto e_2$, and $v_{-1} \mapsto e_3$. Then ω becomes B. From Corollary 1.6.3 we know that any element of the subgroup N^+ or N^- in Lemma 2.2.1 is carried by the homomorphism ρ to a unipotent matrix. Hence by Lemma 2.2.1 we conclude that $\det(\rho(g)) = 1$ for all $g \in \widetilde{G}$. Hence $\rho(\widetilde{G}) \subset G$ by Lemma 2.2.1. If $h = \mathrm{diag}[x, x^{-1}] \in \widetilde{H}$, then $\rho(h)$ has the matrix $\mathrm{diag}[x^2, 1, x^{-2}]$, relative to the ordered basis $\{v_1, v_0, v_{-1}\}$ for V. Thus $\rho(\widetilde{H}) = H$.

Finally, we consider $G = \mathbf{SO}(\mathbb{C}^4, B)$, where B is the symmetric bilinear form on $\mathbb{C}^4$ with matrix s_4 as in (2.5). From Section 2.1.2 we know that the subgroup

$$H = \{\mathrm{diag}[x_1, x_2, x_2^{-1}, x_1^{-1}] : x_1, x_2 \in \mathbb{C}^\times\}$$

of diagonal matrices in G is a maximal torus. Set $\widetilde{G} = \mathbf{SL}(2,\mathbb{C}) \times \mathbf{SL}(2,\mathbb{C})$ and let $\widetilde{H}$ be the product of the diagonal subgroups of the factors of $\widetilde{G}$. We now define a homomorphism $\pi : \widetilde{G} \longrightarrow G$ that maps $\widetilde{H}$ onto H, as follows. Set $V = M_2(\mathbb{C})$ and let $\widetilde{G}$ act on V by $\pi(a,b)X = aXb^{-1}$. From the quadratic form $Q(X) = 2\det X$ on V we obtain the symmetric bilinear form $\beta(X,Y) = \det(X+Y) - \det X - \det Y$. Set

$$v_1 = e_{11}\,, \quad v_2 = e_{12}\,, \quad v_3 = -e_{21}\,, \text{ and } \quad v_4 = e_{22}\,.$$

Clearly $\beta(\pi(g)X, \pi(g)Y) = \beta(X,Y)$ for $g \in \widetilde{G}$. The vectors v_j are β-isotropic and $\beta(v_1, v_4) = \beta(v_2, v_3) = 1$. If we identify V with $\mathbb{C}^4$ via the basis $\{v_1, v_2, v_3, v_4\}$, then β becomes the form B.

Let $g \in \widetilde{G}$ be of the form (I, b) or (b, I), where b is either in the subgroup N^+ or in the subgroup N^- of Lemma 2.2.1. From Corollary 1.6.3 we know that $\pi(g)$ is a unipotent matrix, and so from Lemma 2.2.1 we conclude that $\det(\pi(g)) = 1$ for all $g \in \widetilde{G}$. Hence $\pi(\widetilde{G}) \subset \mathbf{SO}(\mathbb{C}^4, B)$. Given $h = (\mathrm{diag}[x_1, x_1^{-1}], \mathrm{diag}[x_2, x_2^{-1}]) \in \widetilde{H}$, we have

$$\pi(h) = \mathrm{diag}[x_1 x_2^{-1},\, x_1 x_2,\, x_1^{-1} x_2^{-1},\, x_1^{-1} x_2]\,.$$

Since the map $(x_1, x_2) \mapsto (x_1 x_2^{-1}, x_1 x_2)$ from $(\mathbb{C}^\times)^2$ to $(\mathbb{C}^\times)^2$ is surjective, we have shown that $\pi(\widetilde{H}) = H$.

2.2.2 Unipotent Generation of Classical Groups

The differential of a regular representation of an algebraic group G gives a representation of $\mathrm{Lie}(G)$. On the nilpotent elements in $\mathrm{Lie}(G)$ the exponential map is algebraic and maps them to unipotent elements in G. This gives an algebraic link from Lie algebra representations to group representations, provided the unipotent elements generate G. We now prove that this is the case for the following families of classical groups.

Theorem 2.2.2. *Suppse that G is $\mathbf{SL}(l+1,\mathbb{C})$, $\mathbf{SO}(2l+1,\mathbb{C})$, or $\mathbf{Sp}(l,\mathbb{C})$ with $l \geq 1$, or that G is $\mathbf{SO}(2l,\mathbb{C})$ with $l \geq 2$. Then G is generated by its unipotent elements.*

Proof. We have $G \subset \mathbf{GL}(n,\mathbb{C})$ (where $n = l+1, 2l$, or $2l+1$). Let G' be the subgroup generated by the unipotent elements of G. Since the conjugate of a unipotent element is unipotent, we see that G' is a normal subgroup of G. In the orthogonal or symplectic case we take the matrix form of G as in Theorem 2.1.5 so that the subgroup H of diagonal matrices is a maximal torus in G. To prove the theorem, it suffices by Theorems 1.6.5 and 2.1.7 to show that $H \subset G'$.

Type A: When $G = \mathbf{SL}(2,\mathbb{C})$, we have $G' = G$ by Lemma 2.2.1. For $G = \mathbf{SL}(n,\mathbb{C})$ with $n \geq 3$ and $h = \mathrm{diag}[x_1,\dots,x_n] \in H$ we factor $h = h'h''$, where

$$h' = \mathrm{diag}[x_1, x_1^{-1}, 1, \dots, 1]\,, \quad h'' = \mathrm{diag}[1, x_1x_2, x_3, \dots, x_n]\,.$$

Let $G_1 \cong \mathbf{SL}(2,\mathbb{C})$ be the subgroup of matrices in block form $\mathrm{diag}[a, I_{n-2}]$ with $a \in \mathbf{SL}(2,\mathbb{C})$, and let $G_2 \cong \mathbf{SL}(n-1,\mathbb{C})$ be the subgroup of matrices in block form $\mathrm{diag}[1, b]$ with $b \in \mathbf{SL}(n-1,\mathbb{C})$. Then $h' \in G_1$ and $h'' \in G_2$. By induction on n, we may assume that h' and h'' are products of unipotent elements. Hence h is also, so we conclude that $G = G'$.

Type C: Let Ω be the symplectic form (2.6). From Section 2.2.1 we know that $\mathbf{Sp}(\mathbb{C}^2,\Omega) = \mathbf{SL}(2,\mathbb{C})$. Hence from Lemma 2.2.1 we conclude that $\mathbf{Sp}(\mathbb{C}^2,\Omega)$ is generated by its unipotent elements. For $G = \mathbf{Sp}(\mathbb{C}^{2l},\Omega)$ with $l > 1$ and $h = \mathrm{diag}[x_1,\dots,x_l,x_l^{-1},\dots,x_1^{-1}] \in H$, we factor $h = h'h''$, where

$$h' = \mathrm{diag}[x_1,1,\dots,1,x_1^{-1}]\,, \qquad h'' = \mathrm{diag}[1,x_2,\dots,x_l,x_l^{-1},\dots,x_2^{-1},1]\,.$$

We split $\mathbb{C}^{2l} = V_1 \oplus V_2$, where $V_1 = \mathrm{Span}\{e_1, e_{2l}\}$ and $V_2 = \mathrm{Span}\{e_2,\dots,e_{2l-1}\}$. The restrictions of the symplectic form Ω to V_1 and to V_2 are nondegenerate. Define

$$\begin{aligned} G_1 &= \{g \in G : gV_1 = V_1 \ \text{ and } \ g = I \ \text{ on } \ V_2\}\,, \\ G_2 &= \{g \in G : g = I \ \text{ on } \ V_1 \ \text{ and } \ gV_2 = V_2\}\,. \end{aligned}$$

Then $G_1 \cong \mathbf{Sp}(1,\mathbb{C})$, while $G_2 \cong \mathbf{Sp}(l-1,\mathbb{C})$, and we have $h' \in G_1$ and $h'' \in G_2$. By induction on l, we may assume that h' and h'' are products of unipotent elements. Hence h is also, so we conclude that $G = G'$.

Types B and D: Let B be the symmetric form (2.9) on $\mathbb{C}^n$. Suppose first that $G = \mathbf{SO}(\mathbb{C}^3,B)$. Let $\widetilde{G} = \mathbf{SL}(2,\mathbb{C})$. In Section 2.2.1 we constructed a regular homomorphism $\rho : \widetilde{G} \longrightarrow \mathbf{SO}(\mathbb{C}^3,B)$ that maps the diagonal subgroup $\widetilde{H} \subset \widetilde{G}$ onto the diagonal subgroup $H \subset G$. Since every element of $\widetilde{H}$ is a product of unipotent elements, the same is true for H. Hence $G = \mathbf{SO}(3,\mathbb{C})$ is generated by its unipotent elements.

Now let $G = \mathbf{SO}(\mathbb{C}^4,B)$ and set $\widetilde{G} = \mathbf{SL}(2,\mathbb{C}) \times \mathbf{SL}(2,\mathbb{C})$. Let H be the diagonal subgroup of G and let $\widetilde{H}$ be the product of the diagonal subgroups of the factors of $\widetilde{G}$. In Section 2.2.1 we constructed a regular homomorphism $\pi : \widetilde{G} \longrightarrow \mathbf{SO}(\mathbb{C}^4,B)$ that maps $\widetilde{H}$ onto H. Hence the argument just given for $\mathbf{SO}(3,\mathbb{C})$ applies in this case, and we conclude that $\mathbf{SO}(4,\mathbb{C})$ is generated by its unipotent elements.

Finally, we consider the groups $G = \mathbf{SO}(\mathbb{C}^n, B)$ with $n \geq 5$. Embed $\mathbf{SO}(\mathbb{C}^{2l}, B)$ into $\mathbf{SO}(\mathbb{C}^{2l+1}, B)$ by the regular homomorphism

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a & 0 & b \\ 0 & 1 & 0 \\ c & 0 & d \end{bmatrix} . \tag{2.14}$$

The diagonal subgroup of $\mathbf{SO}(\mathbb{C}^{2l}, B)$ is isomorphic to the diagonal subgroup of $\mathbf{SO}(\mathbb{C}^{2l+1}, B)$ via this embedding, so it suffices to prove that every diagonal element in $\mathbf{SO}(\mathbb{C}^n, B)$ is a product of unipotent elements when n is even. We just proved this to be the case when $n = 4$, so we may assume $n = 2l \geq 6$. For

$$h = \mathrm{diag}[x_1, \ldots, x_l, x_l^{-1}, \ldots, x_1^{-1}] \in H$$

we factor $h = h'h''$, where

$$h' = \mathrm{diag}[x_1, x_2, 1, \ldots, 1, x_2^{-1}, x_1^{-1}] ,$$
$$h'' = \mathrm{diag}[1, 1, x_3, \ldots, x_l, x_l^{-1}, \ldots, x_3^{-1}, 1, 1] .$$

We split $\mathbb{C}^n = V_1 \oplus V_2$, where

$$V_1 = \mathrm{Span}\{e_1, e_2, e_{n-1}, e_n\} , \qquad V_2 = \mathrm{Span}\{e_3, \ldots, e_{n-2}\} .$$

The restriction of the symmetric form B to V_i is nondegenerate. If we set

$$G_1 = \{g \in G : gV_1 = V_1 \text{ and } g = I \text{ on } V_2\} ,$$

then $h \in G_1 \cong \mathbf{SO}(4, \mathbb{C})$. Let $W_1 = \mathrm{Span}\{e_1, e_n\}$ and $W_2 = \mathrm{Span}\{e_2, \ldots, e_{n-1}\}$. Set

$$G_2 = \{g \in G : g = I \text{ on } W_1 \text{ and } gW_2 = W_2\} .$$

We have $G_2 \cong \mathbf{SO}(2l-2, \mathbb{C})$ and $h'' \in G_2$. Since $2l - 2 \geq 4$, we may assume by induction that h' and h'' are products of unipotent elements. Hence h is also a product of unipotent elements, proving that $G = G'$. □

2.2.3 Connected Groups

Definition 2.2.3. A linear algebraic group G is *connected* (in the sense of algebraic groups) if the ring $\mathcal{O}[G]$ has no zero divisors.

Examples

1. The rings $\mathbb{C}[t]$ and $\mathbb{C}[t, t^{-1}]$ obviously have no zero divisors; hence the additive group $\mathbb{C}$ and the multiplicative group $\mathbb{C}^\times$ are connected. Likewise, the torus D_n of diagonal matrices and the group N_n^+ of upper-triangular unipotent matrices are connected (see Examples 1 and 2 of Section 1.4.2).

2. If G and H are connected linear algebraic groups, then the group $G \times H$ is connected, since $\mathcal{O}[G \times H] \cong \mathcal{O}[G] \otimes \mathcal{O}[H]$.

3. If G is a connected linear algebraic group and there is a surjective regular homomorphism $\rho : G \longrightarrow H$, then H is connected, since $\rho^* : \mathcal{O}[H] \longrightarrow \mathcal{O}[G]$ is injective.

Theorem 2.2.4. *Let G be a linear algebraic group that is generated by unipotent elements. Then G is connected as an algebraic group and as a Lie group.*

Proof. Suppose $f_1, f_2 \in \mathcal{O}[G]$, $f_1 \neq 0$, and $f_1 f_2 = 0$. We must show that $f_2 = 0$. Translating f_1 and f_2 by an element of G if necessary, we may assume that $f_1(I) \neq 0$. Let $g \in G$. Since g is a product of unipotent elements, Theorem 1.6.2 implies that there exist nilpotent elements $X_1, \ldots, X_r$ in $\mathfrak{g}$ such that $g = \exp(X_1) \cdots \exp(X_r)$. Define $\varphi(t) = \exp(tX_1) \cdots \exp(tX_r)$ for $t \in \mathbb{C}$. The entries in the matrix $\varphi(t)$ are polynomials in t, and $\varphi(1) = g$. Since X_j is nilpotent, we have $\det(\varphi(t)) = 1$ for all t. Hence the functions $p_1(t) = f_1(\varphi(t))$ and $p_2(t) = f_2(\varphi(t))$ are polynomials in t. Since $p_1(0) \neq 0$ while $p_1(t)p_2(t) = 0$ for all t, it follows that $p_2(t) = 0$ for all t. In particular, $f_2(g) = 0$. This holds for all $g \in G$, so $f_2 = 0$, proving that G is connected as a linear algebraic group. This argument also shows that G is arcwise connected, and hence connected, as a Lie group. □

Theorem 2.2.5. *The groups* $\mathbf{GL}(n,\mathbb{C})$, $\mathbf{SL}(n,\mathbb{C})$, $\mathbf{SO}(n,\mathbb{C})$, *and* $\mathbf{Sp}(n,\mathbb{C})$ *are connected (as linear algebraic groups and Lie groups) for all* $n \geq 1$.

Proof. The homomorphism $\lambda, g \mapsto \lambda g$ from $\mathbb{C}^\times \times \mathbf{SL}(n,\mathbb{C})$ to $\mathbf{GL}(n,\mathbb{C})$ is surjective. Hence the connectedness of $\mathbf{GL}(n,\mathbb{C})$ will follow from the connectedness of $\mathbb{C}^\times$ and $\mathbf{SL}(n,\mathbb{C})$, as in Examples 2 and 3 above. The groups $\mathbf{SL}(1,\mathbb{C})$ and $\mathbf{SO}(1,\mathbb{C})$ are trivial, and we showed in Section 2.2.1 that $\mathbf{SO}(2,\mathbb{C})$ is isomorphic to $\mathbb{C}^\times$, hence connected. For the remaining cases use Theorems 2.2.2 and 2.2.4. □

Remark 2.2.6. The regular homomorphisms $\rho : \mathbf{SL}(2,\mathbb{C}) \longrightarrow \mathbf{SO}(3,\mathbb{C})$ and $\pi : \mathbf{SL}(2,\mathbb{C}) \times \mathbf{SL}(2,\mathbb{C}) \longrightarrow \mathbf{SO}(4,\mathbb{C})$ constructed in Section 2.2.1 have kernels $\pm I$; hence $d\rho$ and $d\pi$ are bijective by dimensional considerations. Since $\mathbf{SO}(n,\mathbb{C})$ is connected, it follows that these homomorphisms are surjective. After we introduce the spin groups in Chapter 6, we will see that $\mathbf{SL}(2,\mathbb{C}) \cong \mathbf{Spin}(3,\mathbb{C})$ and $\mathbf{SL}(2,\mathbb{C}) \times \mathbf{SL}(2,\mathbb{C}) \cong \mathbf{Spin}(4,\mathbb{C})$.

We shall study regular representations of a linear algebraic group in terms of the associated representations of its Lie algebra. The following theorem will be a basic tool.

Theorem 2.2.7. *Suppose G is a linear algebraic group with Lie algebra* $\mathfrak{g}$. *Let* (π, V) *be a regular representation of G and* $W \subset V$ *a subspace.*

1. *If* $\pi(g)W \subset W$ *for all* $g \in G$ *then* $d\pi(A)W \subset W$ *for all* $A \in \mathfrak{g}$.
2. *Assume that G is generated by unipotent elements. If* $d\pi(X)W \subset W$ *for all* $X \in \mathfrak{g}$ *then* $\pi(g)W \subset W$ *for all* $g \in G$. *Hence V is irreducible under the action of G if and only if it is irreducible under the action of* $\mathfrak{g}$.

Proof. This follows by the same argument as in Proposition 1.7.7, using the exponentials of nilpotent elements of $\mathfrak{g}$ to generate G in part (2). □

Remark 2.2.8. In Chapter 11 we shall show that the algebraic notion of connectedness can be expressed in terms of the *Zariski topology*, and that a connected linear algebraic group is also connected relative to its topology as a Lie group (Theorem 11.2.9). Since a connected Lie group is generated by $\{\exp X : X \in \mathfrak{g}\}$, this will imply part (2) of Theorem 2.2.7 without assuming unipotent generation of G.

2.2.4 Exercises

1. (*Cayley Parameters*) Let G be $\mathbf{SO}(n,\mathbb{C})$ or $\mathbf{Sp}(n,\mathbb{C})$ and let $\mathfrak{g} = \mathrm{Lie}(G)$. Define $\mathcal{V}_G = \{g \in G : \det(I+g) \neq 0\}$ and $\mathcal{V}_\mathfrak{g} = \{X \in \mathfrak{g} : \det(I-X) \neq 0\}$. For $X \in \mathcal{V}_\mathfrak{g}$ define the Cayley transform $c(X) = (I+X)(I-X)^{-1}$. (Recall that $c(X) \in G$ by Exercises 1.4.5 #5.)
   (a) Show that c is a bijection from $\mathcal{V}_\mathfrak{g}$ onto $\mathcal{V}_G$.
   (b) Show that $\mathcal{V}_\mathfrak{g}$ is invariant under the adjoint action of G on $\mathfrak{g}$, and show that $gc(X)g^{-1} = c(gXg^{-1})$ for $g \in G$ and $X \in \mathcal{V}_\mathfrak{g}$.
   (c) Suppose that $f \in \mathcal{O}[G]$ and f vanishes on $\mathcal{V}_G$. Prove that $f = 0$.
   (HINT: Consider the function $g \mapsto f(g)\det(I+g)$ and use Theorem 2.2.5.)
2. Let $\rho : \mathbf{SL}(2,\mathbb{C}) \longrightarrow \mathbf{SO}(\mathbb{C}^3,B)$ as in Section 2.2.1. Let H (resp. $\widetilde{H}$) be the diagonal subgroup in $\mathbf{SO}(\mathbb{C}^3,B)$ (resp. $\mathbf{SL}(2,\mathbb{C})$). Let $\rho^* : \mathcal{X}(H) \longrightarrow \mathcal{X}(\widetilde{H})$ be the homomorphism of the character groups given by $\chi \mapsto \chi \circ \rho$. Determine the image of ρ^*. (HINT: $\mathcal{X}(H)$ and $\mathcal{X}(\widetilde{H})$ are isomorphic to the additive group $\mathbb{Z}$, and the image of ρ^* can be identified with a subgroup of $\mathbb{Z}$.)
3. Let $\pi : \mathbf{SL}(2,\mathbb{C}) \times \mathbf{SL}(2,\mathbb{C}) \longrightarrow \mathbf{SO}(\mathbb{C}^4,B)$ as in Section 2.2.1. Repeat the calculations of the previous exercise in this case. (HINT: Now $\mathcal{X}(H)$ and $\mathcal{X}(\widetilde{H})$ are isomorphic to the additive group $\mathbb{Z}^2$, and the image of π^* can be identified with a lattice in $\mathbb{Z}^2$.)

2.3 Regular Representations of $\mathbf{SL}(2,\mathbb{C})$

The group $G = \mathbf{SL}(2,\mathbb{C})$ and its Lie algebra $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$ play central roles in determining the structure of the classical groups and their representations. To find all the regular representations of G, we begin by finding all the irreducible finite-dimensional representations of $\mathfrak{g}$. Then we show that every such representation is the differential of an irreducible regular representation of G, thereby obtaining all irreducible regular representations of G. Next we show that an every finite-dimensional representation of $\mathfrak{g}$ decomposes as a direct sum of irreducible representations (the *complete reducibility* property), and conclude that every regular representation of G is completely reducible.

2.3.1 Irreducible Representations of $\mathfrak{sl}(2,\mathbb{C})$

Recall that a *representation* of a complex Lie algebra $\mathfrak{g}$ on a complex vector space V is a linear map $\pi : \mathfrak{g} \longrightarrow \mathrm{End}(V)$ such that

$$\pi([A,B]) = \pi(A)\pi(B) - \pi(B)\pi(A) \quad \text{for all } A,B \in \mathfrak{g} \,.$$

Here the Lie bracket $[A,B]$ on the left is calculated in $\mathfrak{g}$, whereas the product on the right is composition of linear transformations. We shall call V a $\mathfrak{g}$-*module* and write $\pi(A)v$ simply as Av when $v \in V$, provided that the representation π is understood from the context. Thus, even if $\mathfrak{g}$ is a Lie subalgebra of $M_n(\mathbb{C})$, an expression such as $A^k v$, for a nonnegative integer k, means $\pi(A)^k v$.

Let $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$. The matrices $x = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$, $y = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$, $h = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$ are a basis for $\mathfrak{g}$ and satisfy the commutation relations

$$[h,x] = 2x \,, \quad [h,y] = -2y \,, \quad [x,y] = h \,. \tag{2.15}$$

Any triple $\{x,y,h\}$ of nonzero elements in a Lie algebra satisfying (2.15) will be called a TDS (three-dimensional simple) triple.

Lemma 2.3.1. *Let V be a $\mathfrak{g}$-module (possibly infinite-dimensional) and let $v_0 \in V$ be such that $xv_0 = 0$ and $hv_0 = \lambda v_0$ for some $\lambda \in \mathbb{C}$. Set $v_j = y^j v_0$ for $j \in \mathbb{N}$ and $v_j = 0$ for $j < 0$. Then $yv_j = v_{j+1}$, $hv_j = (\lambda - 2j)v_j$, and*

$$xv_j = j(\lambda - j + 1)v_{j-1} \quad \textit{for } j \in \mathbb{N} \,. \tag{2.16}$$

Proof. The equation for yv_j follows by definition, and the equation for hv_j follows from the commutation relation (proved by induction on j)

$$hy^j v = y^j hv - 2jv \quad \text{for all } v \in V \text{ and } j \in \mathbb{N} \,. \tag{2.17}$$

From (2.17) and the relation $xyv = yxv + hv$ one proves by induction on j that

$$xy^j v = jy^{j-1}(h - j + 1)v + y^j xv \quad \text{for all } v \in V \text{ and } j \in \mathbb{N} \,. \tag{2.18}$$

Taking $v = v_0$ and using $xv_0 = 0$, we obtain equation (2.16). □

Let V be a finite-dimensional $\mathfrak{g}$-module. We decompose V into generalized eigenspaces for the action of h:

$$V = \bigoplus_{\lambda \in \mathbb{C}} V(\lambda), \quad \text{where } V(\lambda) = \textstyle\bigcup_{k \geq 1} \mathrm{Ker}(h - \lambda)^k \,.$$

If $v \in V(\lambda)$ then $(h-\lambda)^k v = 0$ for some $k \geq 1$. As linear transformations on V,

$$x(h-\lambda) = (h - \lambda - 2)x \quad \text{and} \quad y(h-\lambda) = (h - \lambda + 2)x \,.$$

Hence $(h-\lambda-2)^k xv = x(h-\lambda)^k v = 0$ and $(h-\lambda+2)^k yv = y(h-\lambda)^k v = 0$. Thus

$$xV(\lambda) \subset V(\lambda+2) \quad \text{and} \quad yV(\lambda) \subset V(\lambda-2) \quad \text{for all } \lambda \in \mathbb{C}\,. \tag{2.19}$$

If $V(\lambda) \neq 0$ then λ is called a *weight* of V with *weight space* $V(\lambda)$.

Lemma 2.3.2. *Suppose V is a finite-dimensional $\mathfrak{g}$-module and $0 \neq v_0 \in V$ satisfies $hv_0 = \lambda v_0$ and $xv_0 = 0$. Let k be the smallest nonnegative integer such that $y^k v_0 \neq 0$ and $y^{k+1}v_0 = 0$. Then $\lambda = k$ and the space $W = \mathrm{Span}\{v_0, yv_0, \ldots, y^k v_0\}$ is a $(k+1)$-dimensional $\mathfrak{g}$-module.*

Proof. Such an integer k exists by (2.19), since V is finite-dimensional and the weight spaces are linearly independent. Lemma 2.3.1 implies that W is invariant under x, y, and h. Furthermore, $v_0, yv_0, \ldots, y^k v_0$ are eigenvectors for h with respective eigenvalues $\lambda, \lambda-2, \ldots, \lambda-2k$. Hence these vectors are a basis for W. By (2.16),

$$0 = xy^{k+1}v_0 = (k+1)(\lambda-k)y^k v_0\,.$$

Since $y^k v_0 \neq 0$, it follows that $\lambda = k$. □

We can describe the action of $\mathfrak{g}$ on the subspace W in Lemma 2.3.2 in matrix form as follows: For $k \in \mathbb{N}$ define the $(k+1)\times(k+1)$ matrices

$$X_k = \begin{bmatrix} 0 & k & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2(k-1) & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3(k-2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & k \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad Y_k = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix},$$

and $H_k = \mathrm{diag}[k, k-2, \ldots, 2-k, -k]$. A direct check yields

$$[X_k, Y_k] = H_k\,, \quad [H_k, X_k] = 2X_k\,, \quad \text{and} \quad [H_k, Y_k] = -2Y_k\,.$$

With all of this in place we can classify the irreducible finite-dimensional modules for $\mathfrak{g}$.

Proposition 2.3.3. *Let $k \geq 0$ be an integer. The representation $(\rho_k, F^{(k)})$ of $\mathfrak{g}$ on $\mathbb{C}^{k+1}$ defined by*

$$\rho_k(x) = X_k\,, \quad \rho_k(h) = H_k\,, \quad \textit{and} \quad \rho_k(y) = Y_k$$

is irreducible. Furthermore, if (σ, W) is an irreducible representation of $\mathfrak{g}$ with $\dim W = k+1 > 0$, then (σ, W) is equivalent to $(\rho_k, F^{(k)})$. In particular, W is equivalent to W^* as a $\mathfrak{g}$-module.

Proof. Suppose that $W \subset F^{(k)}$ is a nonzero invariant subspace. Since $xW(\lambda) \subset W(\lambda+2)$, there must be λ with $W(\lambda) \neq 0$ and $xW(\lambda) = 0$. But from the echelon form of X_k we see that $\mathrm{Ker}(X_k) = \mathbb{C}e_1$. Hence $\lambda = k$ and $W(k) = \mathbb{C}e_1$. Since $Y_k e_j = e_{j+1}$ for $1 \leq j \leq k$, it follows that $W = F^{(k)}$.

Let (σ, W) be any irreducible representation of $\mathfrak{g}$ with $\dim W = k+1 > 0$. There exists an eigenvalue λ of h such that $xW(\lambda) = 0$ and $0 \neq w_0 \in W(\lambda)$ such that $hw_0 = \lambda w_0$. By Lemma 2.3.2 we know that λ is a nonnegative integer, and the space spanned by the set $\{w_0, yw_0, y^2 w_0, \ldots\}$ is invariant under $\mathfrak{g}$ and has dimension $\lambda+1$. But this space is all of W, since σ is irreducible. Hence $\lambda = k$, and by Lemma 2.3.1 the matrices of the actions of x, y, h with respect to the ordered basis $\{w_0, yw_0, \ldots, y^k w_0\}$ are X_k, Y_k, and H_k, respectively. Since W^* is an irreducible $\mathfrak{g}$-module of the same dimension as W, it must be equivalent to W. □

Corollary 2.3.4. *The weights of a finite-dimensional* $\mathfrak{g}$*-module* V *are integers.*

Proof. There are $\mathfrak{g}$-invariant subspaces $0 = V_0 \subset V_1 \subset \cdots \subset V_k = V$ such that the quotient modules $W_j = V_j / V_{j-1}$ are irreducible for $j = 1, \ldots, k-1$. The weights are the eigenvalues of h on V, and this set is the union of the sets of eigenvalues of h on the modules W_j. Hence all weights are integers by Proposition 2.3.3. □

2.3.2 *Irreducible Regular Representations of* $\mathbf{SL}(2, \mathbb{C})$

We now turn to the construction of irreducible regular representations of $\mathbf{SL}(2, \mathbb{C})$. Let the subgroups N^+ of upper-triangular unipotent matrices and N^- of lower-triangular unipotent matrices be as in Section 2.2.1. Set $d(a) = \mathrm{diag}[a, a^{-1}]$ for $a \in \mathbb{C}^\times$.

Proposition 2.3.5. *For every integer* $k \geq 0$ *there is a unique (up to equivalence) irreducible regular representation* (π, V) *of* $\mathbf{SL}(2, \mathbb{C})$ *of dimension* $k+1$ *whose differential is the representation* ρ_k *in Proposition* 2.3.3. *It has the following properties:*

1. *The semisimple operator* $\pi(d(a))$ *has eigenvalues* $a^k, a^{k-2}, \ldots, a^{-k+2}, a^{-k}$.
2. $\pi(d(a))$ *acts on by the scalar* a^k *on the one-dimensional space* V^{N^+} *of* N^+*-fixed vectors.*
3. $\pi(d(a))$ *acts on by the scalar* a^{-k} *on the one-dimensional space* V^{N^-} *of* N^-*-fixed vectors.*

Proof. Let $\mathcal{P}(\mathbb{C}^2)$ be the polynomial functions on $\mathbb{C}^2$ and let $V = \mathcal{P}^k(\mathbb{C}^2)$ be the space of polynomials that are homogeneous of degree k. Here it is convenient to identify elements of $\mathbb{C}^2$ with row vectors $x = [x_1, x_2]$ and have $G = \mathbf{SL}(2, \mathbb{C})$ act by multiplication on the right. We then can define a representation of G on V by $\pi(g)\varphi(x) = \varphi(xg)$ for $\varphi \in V$. Thus

$$\pi(g)\varphi(x_1, x_2) = \varphi(ax_1 + cx_2, bx_1 + dx_2) \quad \text{when } g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} .$$

In particular, the one-parameter subgroups $d(a)$, $u(z)$, and $v(z)$ act by

$$\begin{aligned}
\pi(d(a))\varphi(x_1,x_2) &= \varphi(ax_1, a^{-1}x_2)\,, \\
\pi(u(z))\varphi(x_1,x_2) &= \varphi(x_1, x_2+zx_1)\,, \\
\pi(v(z))\varphi(x_1,x_2) &= \varphi(x_1+zx_2, x_2)\,.
\end{aligned}$$

As a basis for V we take the monomials

$$\mathbf{v}_j(x_1,x_2) = \frac{k!}{(k-j)!}\,x_1^{k-j}x_2^j \quad \text{for } j = 0,1,\ldots,k\,.$$

From the formulas above for the action of $\pi(d(a))$ we see that these functions are eigenvectors for $\pi(d(a))$:

$$\pi(d(a))\mathbf{v}_j = a^{k-2j}\mathbf{v}_j\,.$$

Also, V^{N^+} is the space of polynomials depending only on x_1, so it consists of multiples of $\mathbf{v}_0$, whereas V^{N^-} is the space of polynomials depending only on x_2, so it consists of multiples of $\mathbf{v}_k$.

We now calculate the representation $\mathrm{d}\pi$ of $\mathfrak{g}$. Since $u(z) = \exp(zx)$ and $v(z) = \exp(zy)$, we have $\pi(u(z)) = \exp(z\mathrm{d}\pi(x))$ and $\pi(v(z)) = \exp(z\mathrm{d}\pi(y))$ by Theorem 1.6.2. Taking the z derivative, we obtain

$$\begin{aligned}
\mathrm{d}\pi(x)\varphi(x_1,x_2) &= \frac{\partial}{\partial z}\varphi(x_1,x_2+zx_1)\Big|_{z=0} = x_1\frac{\partial}{\partial x_2}\varphi(x_1,x_2)\,, \\
\mathrm{d}\pi(y)\varphi(x_1,x_2) &= \frac{\partial}{\partial z}\varphi(x_1+zx_2,x_2)\Big|_{z=0} = x_2\frac{\partial}{\partial x_1}\varphi(x_1,x_2)\,.
\end{aligned}$$

Since $\mathrm{d}\pi(h) = \mathrm{d}\pi(x)\mathrm{d}\pi(y) - \mathrm{d}\pi(y)\mathrm{d}\pi(x)$, we also have

$$\mathrm{d}\pi(h)\varphi(x_1,x_2) = \left(x_1\frac{\partial}{\partial x_1} - x_2\frac{\partial}{\partial x_2}\right)\varphi(x_1,x_2)\,.$$

On the basis vectors $\mathbf{v}_j$ we thus have

$$\begin{aligned}
\mathrm{d}\pi(h)\mathbf{v}_j &= \frac{k!}{(k-j)!}\left(x_1\frac{\partial}{\partial x_1} - x_2\frac{\partial}{\partial x_2}\right)(x_1^{k-j}x_2^j) = (k-2j)\mathbf{v}_j\,, \\
\mathrm{d}\pi(x)\mathbf{v}_j &= \frac{k!}{(k-j)!}\left(x_1\frac{\partial}{\partial x_2}\right)(x_1^{k-j}x_2^j) = j(k-j+1)\mathbf{v}_{j-1}\,, \\
\mathrm{d}\pi(x)\mathbf{v}_j &= \frac{k!}{(k-j)!}\left(x_2\frac{\partial}{\partial x_1}\right)(x_1^{k-j}x_2^j) = \mathbf{v}_{j+1}\,.
\end{aligned}$$

It follows from Proposition 2.3.3 that $\mathrm{d}\pi \cong \rho_k$ is an irreducible representation of $\mathfrak{g}$, and all irreducible representations of $\mathfrak{g}$ are obtained this way. Theorem 2.2.7 now implies that π is an irreducible representation of G. Furthermore, π is uniquely determined by $\mathrm{d}\pi$, since $\pi(u)$, for u unipotent, is uniquely determined by $\mathrm{d}\pi(u)$ (Theorem 1.6.2) and G is generated by unipotent elements (Lemma 2.2.1). □

2.3.3 Complete Reducibility of $\mathbf{SL}(2,\mathbb{C})$

Now that we have determined the irreducible regular representations of $\mathbf{SL}(2,\mathbb{C})$, we turn to the problem of finding all the regular representations. We first solve this problem for finite-dimensional representations of $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$.

Theorem 2.3.6. *Let V be a finite-dimensional $\mathfrak{g}$-module with $\dim V > 0$. Then there exist integers $k_1,\ldots,k_r$ (not necessarily distinct) such that V is equivalent to $F^{(k_1)} \bigoplus F^{(k_2)} \bigoplus \cdots \bigoplus F^{(k_r)}$.*

The key step in the proof of Theorem 2.3.6 is the following result:

Lemma 2.3.7. *Suppose W is a $\mathfrak{g}$-module with a submodule Z such that Z is equivalent to $F^{(k)}$ and W/Z is equivalent to $F^{(l)}$. Then W is equivalent to $F^{(k)} \bigoplus F^{(l)}$.*

Proof. Suppose first that $k \neq l$. The lemma is true for W if and only if it is true for W^*. The modules $F^{(k)}$ are self-dual, and replacing W by W^* interchanges the submodule and quotient module. Hence we may assume that $k < l$. By putting h in upper-triangular matrix form, we see that the set of eigenvalues of h on W (ignoring multiplicities) is

$$\{k, k-2, \ldots, -k+2, -k\} \cup \{l, l-2, \ldots, -l+2, -l\} .$$

Thus there exists $0 \neq u_0 \in W$ such that $hu_0 = lu_0$ and $xu_0 = 0$. Since $k < l$, the vector u_0 is not in Z, so the vectors $u_j = y^j u_0$ are not in Z for $j = 0, 1, \ldots, l$ (since $xu_j = j(l-j+1)u_{j-1}$). By Proposition 2.3.3 these vectors span an irreducible $\mathfrak{g}$-module isomorphic to $F^{(l)}$ that has zero intersection with Z. Since $\dim W = k+l+2$, this module is a complement to Z in W.

Now assume that $k = l$. Then $\dim W(l) = 2$, while $\dim Z(l) = 1$. Thus there exist nonzero vectors $z_0 \in Z(l)$ and $w_0 \in W(l)$ with $w_0 \notin Z$ and

$$hw_0 = lw_0 + az_0 \quad \text{for some } a \in \mathbb{C} .$$

Set $z_j = y^j z_0$ and $w_j = y^j w_0$. Using (2.17) we calculate that

$$\begin{aligned} hw_j = hy^j w_0 &= -2jy^j w_0 + y^j hw_0 \\ &= -2jw_j + y^j(lw_0 + az_0) = (l-2j)w_j + az_j . \end{aligned}$$

Since $W(l+2) = 0$, we have $xz_0 = 0$ and $xw_0 = 0$. Thus equation (2.18) gives $xz_j = j(l-j+1)z_{j-1}$ and

$$\begin{aligned} xw_j &= jy^{j-1}(h-j+1)w_0 = j(l-j+1)y^{j-1}w_0 + ajy^{j-1}z_0 \\ &= j(l-j+1)w_{j-1} + ajz_{j-1} . \end{aligned}$$

It follows by induction on j that $\{z_j, w_j\}$ is linearly independent for $j = 0, 1, \ldots, l$. Since the weight spaces $W(l), \ldots, W(-l)$ are linearly independent, we conclude that

$$\{z_0, z_1, \ldots, z_l, w_0, w_1, \ldots, w_l\}$$

is a basis for W. Let X_l, Y_l, and H_l be the matrices in Section 2.3.1. Then relative to this basis the matrices for h, y, and x are

$$H = \begin{bmatrix} H_l & aI \\ 0 & H_l \end{bmatrix}, \quad Y = \begin{bmatrix} Y_l & 0 \\ 0 & Y_l \end{bmatrix}, \quad X = \begin{bmatrix} X_l & A \\ 0 & X_l \end{bmatrix},$$

respectively, where $A = \mathrm{diag}[0, a, 2a, \ldots, la]$. But

$$H = [X,Y] = \begin{bmatrix} H_l & [A,Y_l] \\ 0 & H_l \end{bmatrix}.$$

This implies that $[A,Y_l] = aI$. Hence $0 = \mathrm{tr}(aI) = (l+1)a$, so we have $a = 0$. The matrices H, Y, and X show that W is equivalent to the direct sum of two copies of $F^{(l)}$. □

Proof of Theorem 2.3.6. If $\dim V = 1$ the result is true with $r = 1$ and $k_1 = 0$. Assume that the theorem is true for all $\mathfrak{g}$-modules of dimension less than m, and let V be an m-dimensional $\mathfrak{g}$-module.

The eigenvalues of h on V are integers by Corollary 2.3.4. Let k_1 be the biggest eigenvalue. Then $k_1 \geq 0$ and $V(l) = 0$ for $l > k_1$, so we have an injective module homomorphism of $F^{(k_1)}$ into V by Lemma 2.3.1. Let Z be the image of $F^{(k_1)}$. If $Z = V$ we are done. Otherwise, since $\dim V/Z < \dim V$, we can apply the inductive hypothesis to conclude that V/Z is equivalent to $F^{(k_2)} \bigoplus \cdots \bigoplus F^{(k_r)}$. Let

$$T : V \longrightarrow F^{(k_2)} \bigoplus \cdots \bigoplus F^{(k_r)}$$

be a surjective intertwining operator with kernel Z. For each $i = 2,\ldots,r$ choose $v_i \in V(k_i)$ such that

$$\mathbb{C}Tv_i = 0 \bigoplus \cdots \bigoplus F^{(k_i)}(k_i) \bigoplus \cdots \bigoplus 0\,.$$

Let $W_i = Z + \mathrm{Span}\{v_i, yv_i, \ldots, y^{k_i}v_i\}$ and $T_i = T|_{W_i}$. Then W_i is invariant under $\mathfrak{g}$ and $T_i : W_i \longrightarrow F^{(k_i)}$ is a surjective intertwining operator with kernel Z. Lemma 2.3.7 implies that $W_i = Z \bigoplus U_i$ and T_i defines an equivalence between U_i and $F^{(k_i)}$. Now set $U = U_2 + \cdots + U_r$. Then

$$T(U) = T(U_2) + \cdots + T(U_r) = F^{(k_2)} \bigoplus \cdots \bigoplus F^{(k_r)}\,.$$

Thus $T|_U$ is surjective. Since $\dim U \leq \dim U_2 + \cdots + \dim U_r = \dim T(U)$, it follows that $T|_U$ is bijective. Hence $V = Z \bigoplus U$, completing the induction. □

Corollary 2.3.8. *Let (ρ,V) be a finite-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$. There exists a regular representation (π,W) of $\mathbf{SL}(2,\mathbb{C})$ such that $(\mathrm{d}\pi,W)$ is equivalent to (ρ,V). Furthermore, every regular representation of $\mathbf{SL}(2,\mathbb{C})$ is a direct sum of irreducible subrepresentations.*

Proof. By Theorem 2.3.6 we may assume that $V = F^{(k_1)} \oplus F^{(k_2)} \oplus \cdots \oplus F^{(k_r)}$. Each of the summands is the differential of a representation of $\mathbf{SL}(2,\mathbb{C})$ by Proposition 2.3.5. □

2.3.4 Exercises

1. Let $e_{ij} \in M_3(\mathbb{C})$ be the usual elementary matrices. Set $x = e_{13}$, $y = e_{31}$, and $h = e_{11} - e_{33}$.
   (a) Verify that $\{x,y,h\}$ is a TDS triple in $\mathfrak{sl}(3,\mathbb{C})$.
   (b) Let $\mathfrak{g} = \mathbb{C}x + \mathbb{C}y + \mathbb{C}h \cong \mathfrak{sl}(2,\mathbb{C})$ and let $U = M_3(\mathbb{C})$. Define a representation ρ of $\mathfrak{g}$ on U by $\rho(A)X = [A,X]$ for $A \in \mathfrak{g}$ and $X \in M_3(\mathbb{C})$. Show that $\rho(h)$ is diagonalizable, with eigenvalues ± 2 (multiplicity 1), ± 1 (multiplicity 2), and 0 (multiplicity 3). Find all $u \in U$ such that $\rho(h)u = \lambda u$ and $\rho(x)u = 0$, where $\lambda = 0,1,2$.
   (c) Let $F^{(k)}$ be the irreducible $(k+1)$-dimensional representation of $\mathfrak{g}$. Show that
   $$U \cong F^{(2)} \oplus F^{(1)} \oplus F^{(1)} \oplus F^{(0)} \oplus F^{(0)}$$
   as a $\mathfrak{g}$-module. (HINT: Use the results of (b) and Theorem 2.3.6.)
2. Let k be a nonnegative integer and let W_k be the polynomials in $\mathbb{C}[x]$ of degree at most k. If $f \in W_k$ set
   $$\sigma_k(g)f(x) = (cx+a)^k f\left(\frac{dx+b}{cx+a}\right) \quad \text{for } g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{SL}(2,\mathbb{C}) .$$
   Show that $\sigma_k(g)W_k = W_k$ and that (σ_k, W_k) defines a representation of $\mathbf{SL}(2,\mathbb{C})$ equivalent to the irreducible $(k+1)$-dimensional representation. (HINT: Find an intertwining operator between this representation and the representation used in the proof of Proposition 2.3.5.)
3. Find the irreducible regular representations of $\mathbf{SO}(3,\mathbb{C})$. (HINT: Use the homomorphism $\rho : \mathbf{SL}(2,\mathbb{C}) \longrightarrow \mathbf{SO}(3,\mathbb{C})$ from Section 2.2.1.)
4. Let $V = \mathbb{C}[x]$. Define operators E and F on V by
   $$E\varphi(x) = -\frac{1}{2}\frac{d^2\varphi(x)}{dx^2}, \qquad F\varphi(x) = \frac{1}{2}x^2\varphi(x) \quad \text{for } \varphi \in V .$$
   Set $H = [E,F]$.
   (a) Show that $H = -x(d/dx) - 1/2$ and that $\{E,F,H\}$ is a TDS triple.
   (b) Find the space $V^E = \{\varphi \in V : E\varphi = 0\}$.
   (c) Let $V_{\text{even}} \subset V$ be the space of *even* polynomials and $V_{\text{odd}} \subset V$ the space of *odd* polynomials. Let $\mathfrak{g} \subset \text{End}(V)$ be the Lie algebra spanned by E,F,H. Show that each of these spaces is invariant and irreducible under $\mathfrak{g}$. (HINT: Use (b) and Lemma 2.3.1.)

(d) Show that $V = V_{\text{even}} \oplus V_{\text{odd}}$ and that V_{even} is not equivalent to V_{odd} as a module for $\mathfrak{g}$. (HINT: Show that the operator H is diagonalizable on V_{even} and V_{odd} and find its eigenvalues.)

5. Let $X \in M_n(\mathbb{C})$ be a nilpotent and nonzero. By Exercise 1.6.4 #3 there exist $H, Y \in M_n(\mathbb{C})$ such that $\{X, Y, H\}$ is a TDS triple. Let $\mathfrak{g} = \text{Span}\{H, X, Y\} \cong \mathfrak{sl}(2, \mathbb{C})$ and consider $V = \mathbb{C}^n$ as a representation π of $\mathfrak{g}$ by left multiplication of matrices on column vectors.
(a) Show that π is irreducible if and only if the Jordan canonical form of X consists of a single block.
(b) In the decomposition of V into irreducible subspaces given by Theorem 2.3.6, let m_j be the number of times the representation $F^{(j)}$ occurs. Show that m_j is the number of Jordan blocks of size $j+1$ in the Jordan canonical form of X.
(c) Show that π is determined (up to isomorphism) by the eigenvalues (with multiplicities) of H on $\text{Ker}(X)$.
6. Let (ρ, W) be a finite-dimensional representation of $\mathfrak{sl}(2, \mathbb{C})$. For $k \in \mathbb{Z}$ set $f(k) = \dim\{w \in W : \rho(h)w = kw\}$.
(a) Show that $f(k) = f(-k)$.
(b) Let $g_{\text{even}}(k) = f(2k)$ and $g_{\text{odd}}(k) = f(2k+1)$. Show that g_{even} and g_{odd} are unimodal functions from $\mathbb{Z}$ to $\mathbb{N}$. Here a function ϕ is called *unimodal* if there exists k_0 such that $\phi(a) \le \phi(b)$ for all $a < b \le k_0$ and $\phi(a) \ge \phi(b)$ for all $k_0 \le a < b$. (HINT: Decompose W into a direct sum of irreducible subspaces and use Proposition 2.3.3.)

2.4 The Adjoint Representation

We now use the maximal torus in a classical group to decompose the Lie algebra of the group into eigenspaces, traditionally called *root spaces*, under the adjoint representation.

2.4.1 Roots with Respect to a Maximal Torus

Throughout this section G will denote a connected classical group of rank l. Thus G is $\mathbf{GL}(l, \mathbb{C})$, $\mathbf{SL}(l+1, \mathbb{C})$, $\mathbf{Sp}(\mathbb{C}^{2l}, \Omega)$, $\mathbf{SO}(\mathbb{C}^{2l}, B)$, or $\mathbf{SO}(\mathbb{C}^{2l+1}, B)$, where we take as Ω and B the bilinear forms (2.6) and (2.9). We set $\mathfrak{g} = \text{Lie}(G)$. The subgroup H of diagonal matrices in G is a maximal torus of rank l, and we denote its Lie algebra by $\mathfrak{h}$. In this section we will study the regular representation π of H on the vector space $\mathfrak{g}$ given by $\pi(h)X = hXh^{-1}$ for $h \in H$ and $X \in \mathfrak{g}$.

Let $x_1, \ldots, x_l$ be the coordinate functions on H used in the proof of Theorem 2.1.5. Using these coordinates we obtain an isomorphism between the group $\mathcal{X}(H)$ of rational characters of H and the additive group $\mathbb{Z}^l$ (see Lemma 2.1.2). Under this isomorphism, $\lambda = [\lambda_1, \ldots, \lambda_l] \in \mathbb{Z}^l$ corresponds to the character $h \mapsto h^{\lambda}$, where

$$h^{\lambda} = \prod_{k=1}^{l} x_k(h)^{\lambda_k}, \quad \text{for } h \in H \,. \tag{2.20}$$

For $\lambda, \mu \in \mathbb{Z}^l$ and $h \in H$ we have $h^{\lambda} h^{\mu} = h^{\lambda+\mu}$.

For making calculations it is convenient to fix the following bases for $\mathfrak{h}^*$:

(a) Let $G = \mathbf{GL}(l,\mathbb{C})$. Define $\langle \varepsilon_i, A\rangle = a_i$ for $A = \mathrm{diag}[a_1,\ldots,a_l] \in \mathfrak{h}$. Then $\{\varepsilon_1,\ldots,\varepsilon_l\}$ is a basis for $\mathfrak{h}^*$.

(b) Let $G = \mathbf{SL}(l+1,\mathbb{C})$. Then $\mathfrak{h}$ consists of all diagonal matrices of trace zero. With an abuse of notation we will continue to denote the restrictions to $\mathfrak{h}$ of the linear functionals in (a) by ε_i. The elements of $\mathfrak{h}^*$ can then be written uniquely as $\sum_{i=1}^{l+1} \lambda_i \varepsilon_i$ with $\lambda_i \in \mathbb{C}$ and $\sum_{i=1}^{l+1} \lambda_i = 0$. A basis for $\mathfrak{h}^*$ is furnished by the functionals

$$\varepsilon_i - \frac{1}{l+1}(\varepsilon_1 + \cdots + \varepsilon_{l+1}) \quad \text{for } i = 1,\ldots,l \,.$$

(c) Let G be $\mathbf{Sp}(\mathbb{C}^{2l},\Omega)$ or $\mathbf{SO}(\mathbb{C}^{2l},B)$. For $i = 1,\ldots,l$ define $\langle \varepsilon_i, A\rangle = a_i$, where $A = \mathrm{diag}[a_1,\ldots,a_l,-a_l,\ldots,-a_1] \in \mathfrak{h}$. Then $\{\varepsilon_1,\ldots,\varepsilon_l\}$ is a basis for $\mathfrak{h}^*$.

(d) Let $G = \mathbf{SO}(\mathbb{C}^{2l+1},B)$. For $A = \mathrm{diag}[a_1,\ldots,a_l,0,-a_l,\ldots,-a_1] \in \mathfrak{h}$ and $i = 1,\ldots,l$ define $\langle \varepsilon_i, A\rangle = a_i$. Then $\{\varepsilon_1,\ldots,\varepsilon_l\}$ is a basis for $\mathfrak{h}^*$.

We define $P(G) = \{\mathrm{d}\theta : \theta \in \mathfrak{X}(H)\} \subset \mathfrak{h}^*$. With the functionals ε_i defined as above, we have

$$P(G) = \bigoplus_{k=1}^{l} \mathbb{Z}\varepsilon_k \,. \tag{2.21}$$

Indeed, given $\lambda = \lambda_1\varepsilon_1 + \cdots + \lambda_l\varepsilon_l$ with $\lambda_i \in \mathbb{Z}$, let e^{λ} denote the rational character of H determined by $[\lambda_1,\ldots,\lambda_l] \in \mathbb{Z}^l$ as in (2.20). Every element of $\mathfrak{X}(H)$ is of this form, and we claim that $\mathrm{de}^{\lambda}(A) = \langle \lambda, A\rangle$ for $A \in \mathfrak{h}$. To prove this, recall from Section 1.4.3 that $A \in \mathfrak{h}$ acts by the vector field

$$X_A = \sum_{i=1}^{l} \langle \varepsilon_i, A\rangle\, x_i \frac{\partial}{\partial x_i}$$

on $\mathbb{C}[x_1, x_1^{-1},\ldots,x_l,x_l^{-1}]$. By definition of the differential of a representation we have

$$\mathrm{de}^{\lambda}(A) = X_A(x_1^{\lambda_1}\cdots x_l^{\lambda_l})(1) = \sum_{i=1}^{l} \lambda_i \langle \varepsilon_i, A\rangle = \langle \lambda, A\rangle$$

as claimed. This proves (2.21). The map $\lambda \mapsto \mathrm{e}^{\lambda}$ is thus an isomorphism between the additive group $P(G)$ and the character group $\mathfrak{X}(H)$, by Lemma 2.1.2. From (2.21) we see that $P(G)$ is a *lattice* (free abelian subgroup of rank l) in $\mathfrak{h}^*$, which is called the *weight lattice* of G (the notation $P(G)$ is justified, since all maximal tori are conjugate in G).

We now study the adjoint action of H and $\mathfrak{h}$ on $\mathfrak{g}$. For $\alpha \in P(G)$ let

$$\begin{aligned}\mathfrak{g}_\alpha &= \{X \in \mathfrak{g} : hXh^{-1} = h^\alpha X \text{ for all } h \in H\} \\ &= \{X \in \mathfrak{g} : [A,X] = \langle \alpha, A\rangle X \text{ for all } A \in \mathfrak{h}\} .\end{aligned}$$

(The equivalence of these two formulas for $\mathfrak{g}_\alpha$ is clear from the discussion above.) For $\alpha = 0$ we have $\mathfrak{g}_0 = \mathfrak{h}$, by the same argument as in the proof of Theorem 2.1.5. If $\alpha \neq 0$ and $\mathfrak{g}_\alpha \neq 0$ then α is called a *root* of H on $\mathfrak{g}$ and $\mathfrak{g}_\alpha$ is called a *root space*. If α is a root then a nonzero element of $\mathfrak{g}_\alpha$ is called a *root vector* for α. We call the set Φ of roots the *root system* of $\mathfrak{g}$. Its definition requires fixing a choice of maximal torus, so we write $\Phi = \Phi(\mathfrak{g}, \mathfrak{h})$ when we want to make this choice explicit. Applying Proposition 2.1.3, we have the *root space decomposition*

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha . \tag{2.22}$$

Theorem 2.4.1. *Let $G \subset \mathbf{GL}(n,\mathbb{C})$ be a connected classical group, and let $H \subset G$ be a maximal torus with Lie algebra $\mathfrak{h}$. Let $\Phi \subset \mathfrak{h}^*$ be the root system of $\mathfrak{g}$.*

1. $\dim \mathfrak{g}_\alpha = 1$ *for all* $\alpha \in \Phi$.
2. *If $\alpha \in \Phi$ and $c\alpha \in \Phi$ for some $c \in \mathbb{C}$ then $c = \pm 1$.*
3. *The symmetric bilinear form $(X,Y) = \mathrm{tr}_{\mathbb{C}^n}(XY)$ on $\mathfrak{g}$ is invariant:*

$$([X,Y],Z) = -(Y,[X,Z]) \quad \text{for } X,Y,Z \in \mathfrak{g} .$$

4. *Let $\alpha, \beta \in \Phi$ and $\alpha \neq -\beta$. Then $(\mathfrak{h}, \mathfrak{g}_\alpha) = 0$ and $(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$.*
5. *The form (X,Y) on $\mathfrak{g}$ is nondegenerate.*

Proof of (1): We shall calculate the roots and root vectors for each type of classical group. We take the Lie algebras in the matrix form of Section 2.1.2. In this realization the algebras are invariant under the transpose. For $A \in \mathfrak{h}$ and $X \in \mathfrak{g}$ we have $[A,X]^t = -[A,X^t]$. Hence if X is a root vector for the root α, then X^t is a root vector for $-\alpha$.

Type A: Let G be $\mathbf{GL}(n,\mathbb{C})$ or $\mathbf{SL}(n,\mathbb{C})$. For $A = \mathrm{diag}[a_1, \ldots, a_n] \in \mathfrak{h}$ we have

$$[A, e_{ij}] = (a_i - a_j)e_{ij} = \langle \varepsilon_i - \varepsilon_j, A\rangle e_{ij} .$$

Since the set $\{e_{ij} : 1 \le i,j \le n, i \neq j\}$ is a basis of $\mathfrak{g}$ modulo $\mathfrak{h}$, the roots are

$$\{\pm(\varepsilon_i - \varepsilon_j) : 1 \le i < j \le n\} ,$$

each with multiplicity 1. The root space $\mathfrak{g}_\lambda$ is $\mathbb{C}e_{ij}$ for $\lambda = \varepsilon_i - \varepsilon_j$.

Type C: Let $G = \mathbf{Sp}(\mathbb{C}^{2l}, \Omega)$. Label the basis for $\mathbb{C}^{2l}$ as $e_{\pm 1}, \ldots, e_{\pm l}$, where $e_{-i} = e_{2l+1-i}$. Let $e_{i,j}$ be the matrix that takes the basis vector e_j to e_i and annihilates e_k for $k \neq j$ (here i and j range over $\pm 1, \ldots, \pm l$). Set $X_{\varepsilon_i - \varepsilon_j} = e_{i,j} - e_{-j,-i}$ for $1 \le i,j \le l$, $i \neq j$. Then $X_{\varepsilon_i - \varepsilon_j} \in \mathfrak{g}$ and

$$[A, X_{\varepsilon_i - \varepsilon_j}] = \langle \varepsilon_i - \varepsilon_j, A\rangle X_{\varepsilon_i - \varepsilon_j} , \tag{2.23}$$

for $A \in \mathfrak{h}$. Hence $\varepsilon_i - \varepsilon_j$ is a root. These roots are associated with the embedding $\mathfrak{gl}(l,\mathbb{C}) \longrightarrow \mathfrak{g}$ given by $Y \mapsto \begin{bmatrix} Y & 0 \\ 0 & -s_l Y^t s_l \end{bmatrix}$ for $Y \in \mathfrak{gl}(l,\mathbb{C})$, where s_l is defined in (2.5). Set $X_{\varepsilon_i+\varepsilon_j} = e_{i,-j} + e_{j,-i}$, $X_{-\varepsilon_i-\varepsilon_j} = e_{-j,i} + e_{-i,j}$ for $1 \le i < j \le l$, and set $X_{2\varepsilon_i} = e_{i,-i}$ for $1 \le i \le l$. These matrices are in $\mathfrak{g}$, and

$$[A, X_{\pm(\varepsilon_i+\varepsilon_j)}] = \pm\langle \varepsilon_i + \varepsilon_j, A\rangle X_{\pm(\varepsilon_i+\varepsilon_j)}$$

for $A \in \mathfrak{h}$. Hence $\pm(\varepsilon_i + \varepsilon_j)$ are roots for $1 \le i \le j \le l$. From the block matrix form (2.8) of $\mathfrak{g}$ we see that

$$\{X_{\pm(\varepsilon_i-\varepsilon_j)}, X_{\pm(\varepsilon_i+\varepsilon_j)} : 1 \le i < j \le l\} \cup \{X_{\pm 2\varepsilon_i} : 1 \le i \le l\}$$

is a basis for $\mathfrak{g}$ modulo $\mathfrak{h}$. This shows that the roots have multiplicity one and are

$$\pm(\varepsilon_i - \varepsilon_j) \text{ and } \pm(\varepsilon_i + \varepsilon_j) \text{ for } 1 \le i < j \le l\,, \quad \pm 2\varepsilon_k \text{ for } 1 \le k \le l\,.$$

Type D: Let $G = \mathbf{SO}(\mathbb{C}^{2l}, B)$. Label the basis for $\mathbb{C}^{2l}$ and define $X_{\varepsilon_i-\varepsilon_j}$ as in the case of $\mathbf{Sp}(\mathbb{C}^{2l}, \Omega)$. Then $X_{\varepsilon_i-\varepsilon_j} \in \mathfrak{g}$ and (2.23) holds for $A \in \mathfrak{h}$, so $\varepsilon_i - \varepsilon_j$ is a root. These roots arise from the same embedding $\mathfrak{gl}(l,\mathbb{C}) \longrightarrow \mathfrak{g}$ as in the symplectic case. Set $X_{\varepsilon_i+\varepsilon_j} = e_{i,-j} - e_{j,-i}$ and $X_{-\varepsilon_i-\varepsilon_j} = e_{-j,i} - e_{-i,j}$ for $1 \le i < j \le l$. Then $X_{\pm(\varepsilon_i+\varepsilon_j)} \in \mathfrak{g}$ and

$$[A, X_{\pm(\varepsilon_i+\varepsilon_j)}] = \pm\langle \varepsilon_i + \varepsilon_j, A\rangle X_{\pm(\varepsilon_i+\varepsilon_j)}$$

for $A \in \mathfrak{h}$. Thus $\pm(\varepsilon_i + \varepsilon_j)$ is a root. From the block matrix form (2.7) for $\mathfrak{g}$ we see that

$$\{X_{\pm(\varepsilon_i-\varepsilon_j)} : 1 \le i < j \le l\} \cup \{X_{\pm(\varepsilon_i+\varepsilon_j)} : 1 \le i < j \le l\}$$

is a basis for $\mathfrak{g}$ modulo $\mathfrak{h}$. This shows that the roots have multiplicity one and are

$$\pm(\varepsilon_i - \varepsilon_j) \text{ and } \pm(\varepsilon_i + \varepsilon_j) \text{ for } 1 \le i < j \le l\,.$$

Type B: Let $G = \mathbf{SO}(\mathbb{C}^{2l+1}, B)$. We embed $\mathbf{SO}(\mathbb{C}^{2l}, B)$ into G by equation (2.14). Since $H \subset \mathbf{SO}(\mathbb{C}^{2l}, B) \subset G$ via this embedding, the roots $\pm\varepsilon_i \pm \varepsilon_j$ of $\mathrm{ad}(\mathfrak{h})$ on $\mathfrak{so}(\mathbb{C}^{2l}, B)$ also occur for the adjoint action of $\mathfrak{h}$ on $\mathfrak{g}$. We label the basis for $\mathbb{C}^{2l+1}$ as $\{e_{-l}, \ldots, e_{-1}, e_0, e_1, \ldots, e_l\}$, where $e_0 = e_{l+1}$ and $e_{-i} = e_{2l+2-i}$. Let $e_{i,j}$ be the matrix that takes the basis vector e_j to e_i and annihilates e_k for $k \ne j$ (here i and j range over $0, \pm 1, \ldots, \pm l$). Then the corresponding root vectors from type D are

$$X_{\varepsilon_i-\varepsilon_j} = e_{i,j} - e_{-j,-i}\,, \quad X_{\varepsilon_j-\varepsilon_i} = e_{j,i} - e_{-i,-j}\,,$$

$$X_{\varepsilon_i+\varepsilon_j} = e_{i,-j} - e_{j,-i}\,, \quad X_{-\varepsilon_i-\varepsilon_j} = e_{-j,i} - e_{-i,j}\,,$$

for $1 \le i < j \le l$. Define

$$X_{\varepsilon_i} = e_{i,0} - e_{0,-i}\,, \quad X_{-\varepsilon_i} = e_{0,i} - e_{-i,0}\,,$$

for $1 \le i \le l$. Then $X_{\pm\varepsilon_i} \in \mathfrak{g}$ and $[A, X_{\pm\varepsilon_i}] = \pm\langle\varepsilon_i, A\rangle X_{\varepsilon_i}$ for $A \in \mathfrak{h}$. From the block matrix form (2.10) for $\mathfrak{g}$ we see that $\{X_{\pm\varepsilon_i} : 1 \le i \le l\}$ is a basis for $\mathfrak{g}$ modulo $\mathfrak{so}(\mathbb{C}^{2l}, B)$. Hence the results above for $\mathfrak{so}(\mathbb{C}^{2l}, B)$ imply that the roots of $\mathfrak{so}(\mathbb{C}^{2l+1}, B)$ have multiplicity one and are

$$\pm(\varepsilon_i - \varepsilon_j) \text{ and } \pm(\varepsilon_i + \varepsilon_j) \text{ for } 1 \le i < j \le l\,, \quad \pm\varepsilon_k \text{ for } 1 \le k \le l\,.$$

Proof of (2): This is clear from the calculations above.

Proof of (3): Let $X, Y, Z \in \mathfrak{g}$. Since $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, we have

$$\begin{aligned}([X,Y], Z) &= \operatorname{tr}(XYZ - YXZ) = \operatorname{tr}(YZX - YXZ) \\ &= -\operatorname{tr}(Y[X,Z]) = -(Y, [X,Z])\,.\end{aligned}$$

Proof of (4): Let $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$, and $A \in \mathfrak{h}$. Then

$$0 = ([A,X], Y) + (X, [A,Y]) = \langle\alpha + \beta, A\rangle(X,Y)\,.$$

Since $\alpha + \beta \ne 0$ we can take A such that $\langle\alpha + \beta, A\rangle \ne 0$. Hence $(X,Y) = 0$ in this case. The same argument, but with $Y \in \mathfrak{h}$, shows that $(\mathfrak{h}, \mathfrak{g}_\alpha) = 0$.

Proof of (5): By (4), we only need to show that the restrictions of the trace form to $\mathfrak{h} \times \mathfrak{h}$ and to $\mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha}$ are nondegenerate for all $\alpha \in \Phi$. Suppose $X, Y \in \mathfrak{h}$. Then

$$\operatorname{tr}(XY) = \begin{cases} \sum_{i=1}^{n} \varepsilon_i(X)\varepsilon_i(Y) & \text{if } G = \mathbf{GL}(n,\mathbb{C}) \text{ or } G = \mathbf{SL}(n,\mathbb{C})\,, \\ 2\sum_{i=1}^{l} \varepsilon_i(X)\varepsilon_i(Y) & \text{otherwise.} \end{cases} \tag{2.24}$$

From this it is clear that the restriction of the trace form to $\mathfrak{h} \times \mathfrak{h}$ is nondegenerate.

For $\alpha \in \Phi$ we define $X_\alpha \in \mathfrak{g}_\alpha$ for types A, B, C, and D in terms of the elementary matrices $e_{i,j}$ as above. Then $X_\alpha X_{-\alpha}$ is given as follows (the case of $\mathbf{GL}(n,\mathbb{C})$ is the same as type A):

Type A: $X_{\varepsilon_i-\varepsilon_j}X_{\varepsilon_j-\varepsilon_i} = e_{i,i}$ for $1 \le i < j \le l+1$.

Type B: $X_{\varepsilon_i-\varepsilon_j}X_{\varepsilon_j-\varepsilon_i} = e_{i,i} + e_{-j,-j}$ and $X_{\varepsilon_i+\varepsilon_j}X_{-\varepsilon_j-\varepsilon_i} = e_{i,i} + e_{j,j}$ for $1 \le i < j \le l$. Also $X_{\varepsilon_i}X_{-\varepsilon_i} = e_{i,i} + e_{0,0}$ for $1 \le i \le l$.

Type C: $X_{\varepsilon_i-\varepsilon_j}X_{\varepsilon_j-\varepsilon_i} = e_{i,i} + e_{-j,-j}$ for $1 \le i < j \le l$ and $X_{\varepsilon_i+\varepsilon_j}X_{-\varepsilon_j-\varepsilon_i} = e_{i,i} + e_{j,j}$ for $1 \le i \le j \le l$.

Type D: $X_{\varepsilon_i-\varepsilon_j}X_{\varepsilon_j-\varepsilon_i} = e_{i,i} + e_{-j,-j}$ and $X_{\varepsilon_i+\varepsilon_j}X_{-\varepsilon_j-\varepsilon_i} = e_{i,i} + e_{j,j}$ for $1 \le i < j \le l$.

From these formulas it is evident that $\operatorname{tr}(X_\alpha X_{-\alpha}) \ne 0$ for all $\alpha \in \Phi$. □

2.4.2 Commutation Relations of Root Spaces

We continue the notation of the previous section ($G \subset \mathbf{GL}(n,\mathbb{C})$ a connected classical group). Now that we have decomposed the Lie algebra $\mathfrak{g}$ of G into root spaces under the action of a maximal torus, the next step is to find the commutation relations among the root spaces.

We first observe that

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta} \quad \text{for } \alpha, \beta \in \mathfrak{h}^* . \tag{2.25}$$

Indeed, let $A \in \mathfrak{h}$. Then

$$[A,[X,Y]] = [[A,X],Y] + [X,[A,Y]] = \langle \alpha+\beta, A\rangle [X,Y]$$

for $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_\beta$. Hence $[X,Y] \in \mathfrak{g}_{\alpha+\beta}$. In particular, if $\alpha+\beta$ is not a root, then $\mathfrak{g}_{\alpha+\beta} = 0$, so X and Y commute in this case. We also see from (2.25) that

$$[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{g}_0 = \mathfrak{h} .$$

When α, β, and $\alpha+\beta$ are all roots, then it turns out that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \neq 0$, and hence the inclusion in (2.25) is an equality (recall that $\dim \mathfrak{g}_\alpha = 1$ for all $\alpha \in \Phi$). One way to prove this is to calculate all possible commutators for each type of classical group. Instead of doing this, we shall follow a more conceptual approach using the representation theory of $\mathfrak{sl}(2,\mathbb{C})$ and the invariant bilinear form on $\mathfrak{g}$ from Theorem 2.4.1.

We begin by showing that for each root α, the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ is isomorphic to $\mathfrak{sl}(2,\mathbb{C})$.

Lemma 2.4.2. (Notation as in Theorem 2.4.1) *For each $\alpha \in \Phi$ there exist $e_\alpha \in \mathfrak{g}_\alpha$ and $f_\alpha \in \mathfrak{g}_{-\alpha}$ such that the element $h_\alpha = [e_\alpha, f_\alpha] \in \mathfrak{h}$ satisfies $\langle \alpha, h_\alpha \rangle = 2$. Hence*

$$[h_\alpha, e_\alpha] = 2e_\alpha , \quad [h_\alpha, f_\alpha] = -2f_\alpha ,$$

so that $\{e_\alpha, f_\alpha, h_\alpha\}$ is a TDS triple.

Proof. By Theorem 2.4.1 we can pick $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$ such that $(X,Y) \neq 0$. Set $A = [X,Y] \in \mathfrak{h}$. Then

$$[A,X] = \langle \alpha, A\rangle X , \quad [A,Y] = -\langle \alpha, A\rangle Y . \tag{2.26}$$

We claim that $A \neq 0$. To prove this take any $B \in \mathfrak{h}$ such that $\langle \alpha, B\rangle \neq 0$. Then

$$(A,B) = ([X,Y],B) = (Y,[B,X]) = \langle \alpha, B\rangle (Y,X) \neq 0 . \tag{2.27}$$

We now prove that $\langle \alpha, A\rangle \neq 0$. Since $A \in \mathfrak{h}$, it is a semisimple matrix. For $\lambda \in \mathbb{C}$ let

$$V_\lambda = \{v \in \mathbb{C}^n : Av = \lambda v\}$$

be the λ eigenspace of A. Assume for the sake of contradiction that $\langle \alpha, A\rangle = 0$. Then from (2.26) we see that X and Y would commute with A, and hence V_λ would be invariant under X and Y. But this would imply that

$$\lambda \dim V_\lambda = \operatorname{tr}_{V_\lambda}(A) = \operatorname{tr}_{V_\lambda}([X,Y]\big|_{V_\lambda}) = 0 .$$

Hence $V_\lambda = 0$ for all $\lambda \neq 0$, making $A = 0$, which is a contradiction.

Now that we know $\langle \alpha, A \rangle \neq 0$, we can rescale X, Y, and A, as follows: Set $e_\alpha = sX$, $f_\alpha = tY$, and $h_\alpha = stA$, where $s,t \in \mathbb{C}^\times$. Then

$$[h_\alpha, e_\alpha] = st\langle \alpha, A \rangle e_\alpha\ , \quad [h_\alpha, f_\alpha] = -st\langle \alpha, A \rangle f_\alpha\ ,$$
$$[e_\alpha, f_\alpha] = st[X,Y] = h_\alpha\ .$$

Thus any choice of s,t such that $st\langle \alpha, A \rangle = 2$ gives $\langle \alpha, h_\alpha \rangle = 2$ and the desired TDS triple. □

For future calculations it will be useful to have explicit choices of e_α and f_α for each pair of roots $\pm\alpha \in \Phi$. If $\{e_\alpha, f_\alpha, h_\alpha\}$ is a TDS triple that satisfies the conditions in Lemma 2.4.2 for a root α, then $\{f_\alpha, e_\alpha, -h_\alpha\}$ satisfies the conditions for $-\alpha$. So we may take $e_{-\alpha} = f_\alpha$ and $f_{-\alpha} = e_\alpha$ once we have chosen e_α and f_α. We shall follow the notation of Section 2.4.1.

Type A:
Let $\alpha = \varepsilon_i - \varepsilon_j$ with $1 \le i < j \le l+1$. Set $e_\alpha = e_{ij}$ and $f_\alpha = e_{ji}$. Then $h_\alpha = e_{ii} - e_{jj}$.

Type B:
(a) For $\alpha = \varepsilon_i - \varepsilon_j$ with $1 \le i < j \le l$ set $e_\alpha = e_{i,j} - e_{-j,-i}$ and $f_\alpha = e_{j,i} - e_{-i,-j}$. Then $h_\alpha = e_{i,i} - e_{j,j} + e_{-j,-j} - e_{-i,-i}$.
(b) For $\alpha = \varepsilon_i + \varepsilon_j$ with $1 \le i < j \le l$ set $e_\alpha = e_{i,-j} - e_{j,-i}$ and $f_\alpha = e_{-j,i} - e_{-i,j}$. Then $h_\alpha = e_{i,i} + e_{j,j} - e_{-j,-j} - e_{-i,-i}$.
(c) For $\alpha = \varepsilon_i$ with $1 \le i \le l$ set $e_\alpha = e_{i,0} - e_{0,-i}$ and $f_\alpha = 2e_{0,i} - 2e_{-i,0}$. Then $h_\alpha = 2e_{i,i} - 2e_{-i,-i}$.

Type C:
(a) For $\alpha = \varepsilon_i - \varepsilon_j$ with $1 \le i < j \le l$ set $e_\alpha = e_{i,j} - e_{-j,-i}$ and $f_\alpha = e_{j,i} - e_{-i,-j}$. Then $h_\alpha = e_{i,i} - e_{j,j} + e_{-j,-j} - e_{-i,-i}$.
(b) For $\alpha = \varepsilon_i + \varepsilon_j$ with $1 \le i < j \le l$ set $e_\alpha = e_{i,-j} + e_{j,-i}$ and $f_\alpha = e_{-j,i} - e_{-i,j}$. Then $h_\alpha = e_{i,i} + e_{j,j} - e_{-j,-j} - e_{-i,-i}$.
(c) For $\alpha = 2\varepsilon_i$ with $1 \le i \le l$ set $e_\alpha = e_{i,-i}$ and $f_\alpha = e_{-i,i}$. Then $h_\alpha = e_{i,i} - e_{-i,-i}$.

Type D:
(a) For $\alpha = \varepsilon_i - \varepsilon_j$ with $1 \le i < j \le l$ set $e_\alpha = e_{i,j} - e_{-j,-i}$ and $f_\alpha = e_{j,i} - e_{-i,-j}$. Then $h_\alpha = e_{i,i} - e_{j,j} + e_{-j,-j} - e_{-i,-i}$.
(b) For $\alpha = \varepsilon_i + \varepsilon_j$ with $1 \le i < j \le l$ set $e_\alpha = e_{i,-j} - e_{j,-i}$ and $f_\alpha = e_{-j,i} - e_{-i,j}$. Then $h_\alpha = e_{i,i} + e_{j,j} - e_{-j,-j} - e_{-i,-i}$.

In all cases it is evident that $\langle \alpha, h_\alpha \rangle = 2$, so e_α, f_α satisfy the conditions of Lemma 2.4.2.

We call h_α the *coroot* to α. Since the space $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ has dimension one, h_α is uniquely determined by the properties $h_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ and $\langle \alpha, h_\alpha \rangle = 2$. For $X,Y \in \mathfrak{g}$ let the bilinear form (X,Y) be defined as in Theorem 2.4.1. This form is nondegenerate on $\mathfrak{h} \times \mathfrak{h}$; hence we may use it to identify $\mathfrak{h}$ with $\mathfrak{h}^*$. Then (2.27) implies that

h_α is proportional to α. Furthermore, $(h_\alpha, h_\alpha) = \langle \alpha, h_\alpha \rangle (e_\alpha, f_\alpha) \neq 0$. Hence with $\mathfrak{h}$ identified with $\mathfrak{h}^*$ we have

$$\alpha = \frac{2}{(h_\alpha, h_\alpha)} h_\alpha \, . \tag{2.28}$$

We will also use the notation $\check{\alpha}$ for the coroot h_α.

For $\alpha \in \Phi$ we denote by $\mathfrak{s}(\alpha)$ the algebra spanned by $\{e_\alpha, f_\alpha, h_\alpha\}$. It is isomorphic to $\mathfrak{sl}(2,\mathbb{C})$ under the map $e \mapsto e_\alpha$, $f \mapsto f_\alpha$, $h \mapsto h_\alpha$. The algebra $\mathfrak{g}$ becomes a module for $\mathfrak{s}(\alpha)$ by restricting the adjoint representation of $\mathfrak{g}$ to $\mathfrak{s}(\alpha)$. We can thus apply the results on the representations of $\mathfrak{sl}(2,\mathbb{C})$ that we obtained in Section 2.3.3 to study commutation relations in $\mathfrak{g}$.

Let $\alpha, \beta \in \Phi$ with $\alpha \neq \pm\beta$. We observe that $\beta + k\alpha \neq 0$, by Theorem 2.4.1 (2). Hence for every $k \in \mathbb{Z}$,

$$\dim \mathfrak{g}_{\beta+k\alpha} = \begin{cases} 1 \text{ if } \beta + k\alpha \in \Phi \, , \\ 0 \text{ otherwise.} \end{cases}$$

Let

$$R(\alpha, \beta) = \{\beta + k\alpha \, : \, k \in \mathbb{Z}\} \cap \Phi \, ,$$

which we call the α *root string* through β. The number of elements of a root string is called the *length* of the string. Define

$$V_{\alpha,\beta} = \sum_{\gamma \in R(\alpha,\beta)} \mathfrak{g}_\gamma \, .$$

Lemma 2.4.3. *For every $\alpha, \beta \in \Phi$ with $\alpha \neq \pm\beta$, the space $V_{\alpha,\beta}$ is invariant and irreducible under* $\mathrm{ad}(\mathfrak{s}(\alpha))$.

Proof. From (2.25) we have $[\mathfrak{g}_\alpha, \mathfrak{g}_{\beta+k\alpha}] \subset \mathfrak{g}_{\beta+(k+1)\alpha}$ and $[\mathfrak{g}_{-\alpha}, \mathfrak{g}_{\beta+k\alpha}] \subset \mathfrak{g}_{\beta+(k-1)\alpha}$, so we see that $V_{\alpha,\beta}$ is invariant under $\mathrm{ad}(\mathfrak{s}(\alpha))$. Denote by π the representation of $\mathfrak{s}(\alpha)$ on $V_{\alpha,\beta}$.

If $\gamma = \beta + k\alpha \in \Phi$, then $\pi(h_\alpha)$ acts on the one-dimensional space $\mathfrak{g}_\gamma$ by the scalar

$$\langle \gamma, h_\alpha \rangle = \langle \beta, h_\alpha \rangle + k \langle \alpha, h_\alpha \rangle = \langle \beta, h_\alpha \rangle + 2k \, .$$

Thus by (2.29) we see that the eigenvalues of $\pi(h_\alpha)$ are integers and are either all even or all odd. Furthermore, each eigenvalue occurs with multiplicity one.

Suppose for the sake of contradiction that $V_{\alpha,\beta}$ is not irreducible under $\mathfrak{s}(\alpha)$. Then by Theorem 2.3.6, $V_{\alpha,\beta}$ contains nonzero irreducible invariant subspaces U and W with $W \cap U = \{0\}$. By Proposition 2.3.3 the eigenvalues of h_α on W are n, $n-2$, ..., $-n+2$, $-n$ and the eigenvalues of h_α on U are m, $m-2$, ..., $-m+2$, $-m$, where m and n are nonnegative integers. The eigenvalues of h_α on W and on U are subsets of the set of eigenvalues of $\pi(h_\alpha)$, so it follows that m and n are both even or both odd. But this implies that the eigenvalue $\min(m,n)$ of $\pi(h_\alpha)$ has multiplicity greater than one, which is a contradiction. □

Corollary 2.4.4. *If $\alpha, \beta \in \Phi$ and $\alpha + \beta \in \Phi$, then* $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$.

Proof. Since $\alpha+\beta \in \Phi$, we have $\alpha \neq \pm\beta$. Thus $V_{\alpha,\beta}$ is irreducible under $\mathfrak{s}_\alpha$ and contains $\mathfrak{g}_{\alpha+\beta}$. Hence by (2.16) the operator $E = \pi(e_\alpha)$ maps $\mathfrak{g}_\beta$ onto $\mathfrak{g}_{\alpha+\beta}$. □

Corollary 2.4.5. *Let $\alpha,\beta \in \Phi$ with $\beta \neq \pm\alpha$. Let p be the largest integer $j \geq 0$ such that $\beta + j\alpha \in \Phi$ and let q be the largest integer $k \geq 0$ such that $\beta - k\alpha \in \Phi$. Then*

$$\langle \beta, h_\alpha \rangle = q - p \in \mathbb{Z}\,,$$

and $\beta + r\alpha \in \Phi$ for all integers r with $-q \leq r \leq p$. In particular, $\beta - \langle \beta, h_\alpha\rangle\alpha \in \Phi$.

Proof. The largest eigenvalue of $\pi(h_\alpha)$ is the positive integer $n = \langle \beta, h_\alpha\rangle + 2p$. Since π is irreducible, Proposition 2.3.3 implies that the eigenspaces of $\pi(h_\alpha)$ are $\mathfrak{g}_{\beta+r\alpha}$ for $r = p, p-1, \ldots, -q+1, -q$. Hence the α-string through β is $\beta + r\alpha$ with $r = p, p-1, \ldots, -q+1, -q$. Furthermore, the smallest eigenvalue of $\pi(h)$ is $-n = \langle \beta, h_\alpha\rangle - 2q$. This gives the relation

$$-\langle \beta, h_\alpha\rangle - 2p = \langle \beta, h_\alpha\rangle - 2q\,.$$

Hence $\langle \beta, h_\alpha\rangle = q - p$. Since $p \geq 0$ and $q \geq 0$, we see that $-q \leq -\langle \beta, h_\alpha\rangle \leq p$. Thus $\beta - \langle \beta, h_\alpha\rangle\alpha \in \Phi$. □

Remark 2.4.6. From the case-by-case calculations for types **A–D** made above we see that

$$\langle \beta, h_\alpha\rangle \in \{0, \pm 1, \pm 2\} \quad \text{for all} \;\; \alpha, \beta \in \Phi\,. \tag{2.29}$$

2.4.3 Structure of Classical Root Systems

In the previous section we saw that the commutation relations in the Lie algebra of a classical group are controlled by the root system. We now study the root systems in more detail. Let Φ be the root system for a classical Lie algebra $\mathfrak{g}$ of type A_l, B_l, C_l, or D_l (with $l \geq 3$ for D_l). Then Φ spans $\mathfrak{h}^*$ (this is clear from the descriptions in Section 2.4.1). Thus we can choose (in many ways) a set of roots that is a basis for $\mathfrak{h}^*$. An optimal choice of basis is the following:

Definition 2.4.7. A subset $\Delta = \{\alpha_1, \ldots, \alpha_l\} \subset \Phi$ is a set of *simple roots* if every $\gamma \in \Phi$ can be written uniquely as

$$\gamma = n_1\alpha_1 + \cdots + n_l\alpha_l\,, \;\; \text{with} \;\; n_1, \ldots, n_l \;\; \text{integers all of the same sign.} \tag{2.30}$$

Notice that the requirement of uniqueness in expression (2.30), together with the fact that Φ spans $\mathfrak{h}^*$, implies that Δ is a basis for $\mathfrak{h}^*$. Furthermore, if Δ is a set of simple roots, then it partitions Φ into two disjoint subsets

$$\Phi = \Phi^+ \cup (-\Phi^+)\,,$$

where Φ^+ consists of all the roots for which the coefficients n_i in (2.30) are non-negative. We call $\gamma \in \Phi^+$ a *positive root*, relative to Δ.

We shall show, with a case-by-case analysis, that Φ has a set of simple roots. We first prove that if $\Delta = \{\alpha_1, \dots, \alpha_l\}$ is a set of simple roots and $i \neq j$, then

$$\langle \alpha_i, h_{\alpha_j} \rangle \in \{0, -1, -2\} .$$

Indeed, we have already observed that $\langle \alpha, h_\beta \rangle \in \{0, \pm 1, \pm 2\}$ for all roots α, β. Let $H_i = h_{\alpha_i}$ be the coroot to α_i and define

$$C_{ij} = \langle \alpha_j, H_i \rangle . \tag{2.31}$$

Set $\gamma = \alpha_j - C_{ij}\alpha_i$. By Corollary 2.4.5 we have $\gamma \in \Phi$. If $C_{ij} > 0$ this expansion of γ would contradict (2.30). Hence $C_{ij} \leq 0$ for all $i \neq j$.

Remark 2.4.8. The integers C_{ij} in (2.31) are called the *Cartan integers*, and the $l \times l$ matrix $C = [C_{ij}]$ is called the *Cartan matrix* for the set Δ. Note that the diagonal entries of C are $\langle \alpha_i, H_i \rangle = 2$.

If Δ is a set of simple roots and $\beta = n_1\alpha_1 + \cdots + n_l\alpha_l$ is a root, then we define the *height* of β (relative to Δ) as

$$\mathrm{ht}(\beta) = n_1 + \cdots + n_l .$$

The positive roots are then the roots β with $\mathrm{ht}(\beta) > 0$. A root β is called the *highest root* of Φ, relative to a set Δ of simple roots, if

$$\mathrm{ht}(\beta) > \mathrm{ht}(\gamma) \quad \text{for all roots } \ \gamma \neq \beta .$$

If such a root exists, it is clearly unique.

We now give a set of simple roots and the associated Cartan matrix and positive roots for each classical root system, and we show that there is a highest root, denoted by $\widetilde{\alpha}$ (in type D_l we assume $l \geq 3$). We write the coroots H_i in terms of the elementary diagonal matrices $E_i = e_{i,i}$, as in Section 2.4.1. The Cartan matrix is very sparse, and it can be efficiently encoded in terms of a *Dynkin diagram*. This is a graph with a node for each root $\alpha_i \in \Delta$. The nodes corresponding to α_i and α_j are joined by $C_{ij}C_{ji}$ lines for $i \neq j$. Furthermore, if the two roots are of different lengths (relative to the inner product for which $\{\varepsilon_i\}$ is an orthonormal basis), then an inequality sign is placed on the lines to indicate which root is longer. We give the Dynkin diagrams and indicate the root corresponding to each node in each case. Above the node for α_i we put the coefficient of α_i in the highest root.

Type A ($G = \mathbf{SL}(l+1, \mathbb{C})$): Let $\alpha_i = \varepsilon_i - \varepsilon_{i+1}$ and $\Delta = \{\alpha_1, \dots, \alpha_l\}$. Since

$$\varepsilon_i - \varepsilon_j = \alpha_i + \cdots + \alpha_{j-1} \quad \text{for } 1 \leq i < j \leq l+1 ,$$

we see that Δ is a set of simple roots. The associated set of positive roots is

$$\Phi^+ = \{\varepsilon_i - \varepsilon_j : 1 \leq i < j \leq l+1\} \tag{2.32}$$

and the highest root is $\widetilde{\alpha} = \varepsilon_1 - \varepsilon_{l+1} = \alpha_1 + \cdots + \alpha_l$ with $\mathrm{ht}(\widetilde{\alpha}) = l$. Here $H_i = E_i - E_{i+1}$. Thus the Cartan matrix has $C_{ij} = -1$ if $|i-j| = 1$ and $C_{ij} = 0$ if $|i-j| > 1$. The Dynkin diagram is shown in Figure 2.1.

Fig. 2.1 Dynkin diagram of type A_l.

Type B ($G = \mathbf{SO}(2l+1, \mathbb{C})$): Let $\alpha_i = \varepsilon_i - \varepsilon_{i+1}$ for $1 \le i \le l-1$ and $\alpha_l = \varepsilon_l$. Take $\Delta = \{\alpha_1, \ldots, \alpha_l\}$. For $1 \le i < j \le l$, we can write $\varepsilon_i - \varepsilon_j = \alpha_i + \cdots + \alpha_{j-1}$ as in type A, whereas

$$\begin{aligned}\varepsilon_i + \varepsilon_j &= (\varepsilon_i - \varepsilon_l) + (\varepsilon_j - \varepsilon_l) + 2\varepsilon_l \\ &= \alpha_i + \cdots + \alpha_{l-1} + \alpha_j + \cdots + \alpha_{l-1} + 2\alpha_l \\ &= \alpha_i + \cdots + \alpha_{j-1} + 2\alpha_j + \cdots + 2\alpha_l \,.\end{aligned}$$

For $1 \le i \le l$ we have $\varepsilon_i = (\varepsilon_i - \varepsilon_l) + \varepsilon_l = \alpha_i + \cdots + \alpha_l$. These formulas show that Δ is a set of simple roots. The associated set of positive roots is

$$\Phi^+ = \{\varepsilon_i - \varepsilon_j,\, \varepsilon_i + \varepsilon_j \,:\, 1 \le i < j \le l\} \cup \{\varepsilon_i \,:\, 1 \le i \le l\} \,. \tag{2.33}$$

The highest root is $\widetilde{\alpha} = \varepsilon_1 + \varepsilon_2 = \alpha_1 + 2\alpha_2 + \cdots + 2\alpha_l$ with $\mathrm{ht}(\widetilde{\alpha}) = 2l - 1$. The simple coroots are

$$H_i = E_i - E_{i+1} + E_{-i-1} - E_{-i} \quad \text{for } 1 \le i \le l-1 \,,$$

and $H_l = 2E_l - 2E_{-l}$, where we are using the same enumeration of the basis for $\mathbb{C}^{2l+1}$ as in Section 2.4.1. Thus the Cartan matrix has $C_{ij} = -1$ if $|i-j| = 1$ and $i, j \le l-1$, whereas $C_{l-1,l} = -2$ and $C_{l,l-1} = -1$. All other nondiagonal entries are zero. The Dynkin diagram is shown in Figure 2.2 for $l \ge 2$.

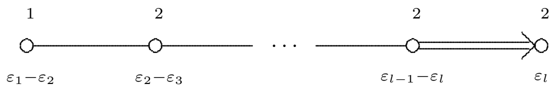

Fig. 2.2 Dynkin diagram of type B_l.

Type C ($G = \mathbf{Sp}(l, \mathbb{C})$): Let $\alpha_i = \varepsilon_i - \varepsilon_{i+1}$ for $1 \le i \le l-1$ and $\alpha_l = 2\varepsilon_l$. Take $\Delta = \{\alpha_1, \ldots, \alpha_l\}$. For $1 \le i < j \le l$ we can write $\varepsilon_i - \varepsilon_j = \alpha_i + \cdots + \alpha_{j-1}$ and $\varepsilon_i + \varepsilon_l = \alpha_i + \cdots + \alpha_l$, whereas for $1 \le i < j \le l-1$ we have

$$\begin{aligned}\varepsilon_i + \varepsilon_j &= (\varepsilon_i - \varepsilon_l) + (\varepsilon_j - \varepsilon_l) + 2\varepsilon_l \\ &= \alpha_i + \cdots + \alpha_{l-1} + \alpha_j + \cdots + \alpha_{l-1} + \alpha_l \\ &= \alpha_i + \cdots + \alpha_{j-1} + 2\alpha_j + \cdots + 2\alpha_{l-1} + \alpha_l \,.\end{aligned}$$

For $1 \le i < l$ we have $2\varepsilon_i = 2(\varepsilon_i - \varepsilon_l) + 2\varepsilon_l = 2\alpha_i + \cdots + 2\alpha_{l-1} + \alpha_l$. These formulas show that Δ is a set of simple roots. The associated set of positive roots is

$$\Phi^+ = \{\varepsilon_i - \varepsilon_j,\, \varepsilon_i + \varepsilon_j \,:\, 1 \le i < j \le l\} \cup \{2\varepsilon_i \,:\, 1 \le i \le l\}\,. \tag{2.34}$$

The highest root is $\widetilde{\alpha} = 2\varepsilon_1 = 2\alpha_1 + \cdots + 2\alpha_{l-1} + \alpha_l$ with $\mathrm{ht}(\widetilde{\alpha}) = 2l-1$. The simple coroots are

$$H_i = E_i - E_{i+1} + E_{-i-1} - E_{-i} \quad \text{for} \;\; 1 \le i \le l-1\,,$$

and $H_l = E_l - E_{-l}$, where we are using the same enumeration of the basis for $\mathbb{C}^{2l+1}$ as in Section 2.4.1. The Cartan matrix has $C_{ij} = -1$ if $|i-j| = 1$ and $i, j \le l-1$, whereas now $C_{l-1,l} = -1$ and $C_{l,l-1} = -2$. All other nondiagonal entries are zero. Notice that this is the transpose of the Cartan matrix of type B. If $l \ge 2$ the Dynkin diagram is shown in Figure 2.3. It can be obtained from the Dynkin diagram of type B_l by reversing the arrow on the double bond and reversing the coefficients of the highest root. In particular, the diagrams B_2 and C_2 are identical. (This low-rank coincidence was already noted in Exercises 1.1.5 #8; it is examined further in Exercises 2.4.5 #6.)

Fig. 2.3 Dynkin diagram of type C_l.

Type D ($G = \mathbf{SO}(2l,\mathbb{C})$ with $l \ge 3$): Let $\alpha_i = \varepsilon_i - \varepsilon_{i+1}$ for $1 \le i \le l-1$ and $\alpha_l = \varepsilon_{l-1} + \varepsilon_l$. For $1 \le i < j \le l$ we can write $\varepsilon_i - \varepsilon_j = \alpha_i + \cdots + \alpha_{j-1}$ as in type A, whereas for $1 \le i < l-1$ we have

$$\varepsilon_i + \varepsilon_{l-1} = \alpha_i + \cdots + \alpha_l\,, \qquad \varepsilon_i + \varepsilon_l = \alpha_i + \cdots + \alpha_{l-2} + \alpha_l\,.$$

For $1 \le i < j \le l-2$ we have

$$\begin{aligned} \varepsilon_i + \varepsilon_j &= (\varepsilon_i - \varepsilon_{l-1}) + (\varepsilon_j - \varepsilon_l) + (\varepsilon_{l-1} + \varepsilon_l) \\ &= \alpha_i + \cdots + \alpha_{l-2} + \alpha_j + \cdots + \alpha_{l-1} + \alpha_l \\ &= \alpha_i + \cdots + \alpha_{j-1} + 2\alpha_j + \cdots + 2\alpha_{l-2} + \alpha_{l-1} + \alpha_l\,. \end{aligned}$$

These formulas show that Δ is a set of simple roots. The associated set of positive roots is

$$\Phi^+ = \{\varepsilon_i - \varepsilon_j,\, \varepsilon_i + \varepsilon_j \,:\, 1 \le i < j \le l\}\,. \tag{2.35}$$

The highest root is $\widetilde{\alpha} = \varepsilon_1 + \varepsilon_2 = \alpha_1 + 2\alpha_2 + \cdots + 2\alpha_{l-2} + \alpha_{l-1} + \alpha_l$ with $\mathrm{ht}(\widetilde{\alpha}) = 2l-3$. The simple coroots are

$$H_i = E_i - E_{i+1} + E_{-i-1} - E_{-i} \quad \text{for} \;\; 1 \le i \le l-1,$$

and $H_l = E_{l-1} + E_l - E_{-l} - E_{-l+1}$, with the same enumeration of the basis for $\mathbb{C}^{2l}$ as in type C. Thus the Cartan matrix has $C_{ij} = -1$ if $|i-j| = 1$ and $i, j \le l-1$, whereas $C_{l-2,l} = C_{l,l-2} = -1$. All other nondiagonal entries are zero. The Dynkin diagram is shown in Figure 2.4. Notice that when $l = 2$ the diagram is not connected (it is the diagram for $\mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{sl}(2,\mathbb{C})$; see Remark 2.2.6). When $l = 3$ the diagram is the same as the diagram for type A_3. This low-rank coincidence was already noted in Exercises 1.1.5 #7; it is examined further in Exercises 2.4.5 #5.

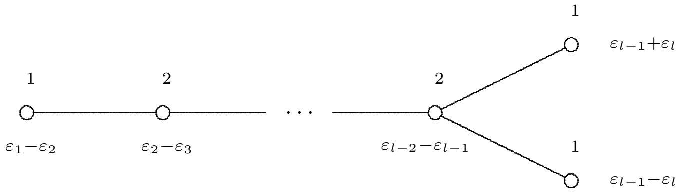

Fig. 2.4 Dynkin diagram of type D_l.

Remark 2.4.9. The Dynkin diagrams of the four types of classical groups are distinct except in the cases $A_1 = B_1 = C_1$, $B_2 = C_2$, and $A_3 = D_3$. In these cases there are corresponding Lie algebra isomorphisms; see Section 2.2.1 for the rank-one simple algebras and see Exercises 2.4.5 for the isomorphisms $\mathfrak{so}(\mathbb{C}^5) \cong \mathfrak{sp}(\mathbb{C}^4)$ and $\mathfrak{sl}(\mathbb{C}^4) \cong \mathfrak{so}(\mathbb{C}^6)$. We will show in Chapter 3 that all systems of simple roots are conjugate by the Weyl group; hence the Dynkin diagram is uniquely defined by the root system and does not depend on the choice of a simple set of roots. Thus the Dynkin diagram completely determines the Lie algebra up to isomorphism.

For a root system of types A or D, in which all the roots have squared length two (relative to the trace form inner product on $\mathfrak{h}$), the identification of $\mathfrak{h}$ with $\mathfrak{h}^*$ takes roots to coroots. For root systems of type B or C, in which the roots have two lengths, the roots of type B_l are identified with the coroots of type C_l and vice versa (e.g., ε_i is identified with the coroot to $2\varepsilon_i$ and vice versa). This allows us to transfer results known for roots to analogous results for coroots. For example, if $\alpha \in \Phi^+$ then

$$H_\alpha = m_1 H_1 + \cdots + m_l H_l, \tag{2.36}$$

where m_i is a nonnegative integer for $i = 1, \ldots, l$.

Lemma 2.4.10. *Let Φ be the root system for a classical Lie algebra $\mathfrak{g}$ of rank l and type A, B, C, or D (in the case of type D assume that $l \geq 3$). Let the system of simple roots $\Delta \subset \Phi$ be chosen as above. Let Φ^+ be the positive roots and let $\widetilde{\alpha}$ be the highest root relative to Δ. Then the following properties hold:*

1. *If $\alpha, \beta \in \Phi^+$ and $\alpha + \beta \in \Phi$, then $\alpha + \beta \in \Phi^+$.*
2. *If $\beta \in \Phi^+$ and β is not a simple root, then there exist $\gamma, \delta \in \Phi^+$ such that $\beta = \gamma + \delta$.*
3. *$\widetilde{\alpha} = n_1\alpha_1 + \cdots + n_l\alpha_l$ with $n_i \geq 1$ for $i = 1, \ldots, l$.*
4. *For any $\beta \in \Phi^+$ with $\beta \neq \widetilde{\alpha}$ there exists $\alpha \in \Phi^+$ such that $\alpha + \beta \in \Phi^+$.*
5. *If $\alpha \in \Phi^+$ and $\alpha \neq \widetilde{\alpha}$, then there exist $1 \leq i_1, i_2, \ldots, i_r \leq l$ such that $\alpha = \widetilde{\alpha} - \alpha_{i_1} - \cdots - \alpha_{i_r}$ and $\widetilde{\alpha} - \alpha_{i_1} - \cdots - \alpha_{i_j} \in \Phi$ for all $1 \leq j \leq r$.*

Proof. Property (1) is clear from the definition of a system of simple roots. Properties (2)–(5) follow on a case-by-case basis from the calculations made above. We leave the details as an exercise. □

We can now state the second structure theorem for $\mathfrak{g}$.

Theorem 2.4.11. *Let* $\mathfrak{g}$ *be the Lie algebra of* $\mathbf{SL}(l+1,\mathbb{C})$, $\mathbf{Sp}(\mathbb{C}^{2l},\Omega)$, *or* $\mathbf{SO}(\mathbb{C}^{2l+1},B)$ *with* $l \geq 1$, *or the Lie algebra of* $\mathbf{SO}(\mathbb{C}^{2l},B)$ *with* $l \geq 3$. *Take the set of simple roots* Δ *and the positive roots* Φ^+ *as in Lemma* 2.4.10. *The subspaces*

$$\mathfrak{n}^+ = \bigoplus_{\alpha\in\Phi^+} \mathfrak{g}_\alpha\,, \quad \mathfrak{n}^- = \bigoplus_{\alpha\in\Phi^+} \mathfrak{g}_{-\alpha}$$

are Lie subalgebras of $\mathfrak{g}$ *that are invariant under* $\mathrm{ad}(\mathfrak{h})$. *The subspace* $\mathfrak{n}^+ + \mathfrak{n}^-$ *generates* $\mathfrak{g}$ *as a Lie algebra. In particular,* $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$. *There is a vector space direct sum decomposition*

$$\mathfrak{g} = \mathfrak{n}^- + \mathfrak{h} + \mathfrak{n}^+\,. \tag{2.37}$$

Furthermore, the iterated Lie brackets of the root spaces $\mathfrak{g}_{\alpha_1},\ldots,\mathfrak{g}_{\alpha_l}$ *span* $\mathfrak{n}^+$, *and the iterated Lie brackets of the root spaces* $\mathfrak{g}_{-\alpha_1},\ldots,\mathfrak{g}_{-\alpha_l}$ *span* $\mathfrak{n}^-$.

Proof. The fact that $\mathfrak{n}^+$ and $\mathfrak{n}^-$ are subalgebras follows from property (1) in Lemma 2.4.10. Equation (2.37) is clear from Theorem 2.4.1 and the decomposition

$$\Phi = \Phi^+ \cup (-\Phi^+)\,.$$

For $\alpha \in \Phi$ let $h_\alpha \in \mathfrak{h}$ be the coroot. From the calculations above it is clear that $\mathfrak{h} = \mathrm{Span}\{h_\alpha : \alpha \in \Phi\}$. Since $h_\alpha \in [\mathfrak{g}_\alpha,\mathfrak{g}_{-\alpha}]$ by Lemma 2.4.2, we conclude from (2.37) that $\mathfrak{n}^+ + \mathfrak{n}^-$ generates $\mathfrak{g}$ as a Lie algebra.

To verify that $\mathfrak{n}^+$ is generated by the simple root spaces, we use induction on the height of $\beta \in \Phi^+$ (the simple roots being the roots of height 1). If β is not simple, then $\beta = \gamma + \delta$ for some $\gamma,\delta \in \Phi^+$ (Lemma 2.4.10 (2)). But we know that $[\mathfrak{g}_\gamma, \mathfrak{g}_\delta] = \mathfrak{g}_\beta$ from Corollary 2.4.4. Since the heights of γ and δ are less than the height of β, the induction continues. The same argument applies to $\mathfrak{n}^-$. ☐

Remark 2.4.12. When $\mathfrak{g}$ is taken in the matrix form of Section 2.4.1, then $\mathfrak{n}^+$ consists of all strictly upper-triangular matrices in $\mathfrak{g}$, and $\mathfrak{n}^-$ consists of all strictly lower-triangular matrices in $\mathfrak{g}$. Furthermore, $\mathfrak{g}$ is invariant under the map $\theta(X) = -X^t$ (negative transpose). This map is an automorphism of $\mathfrak{g}$ with $\theta^2 = \mathrm{Identity}$. Since $\theta(H) = -H$ for $H \in \mathfrak{h}$, it follows that $\theta(\mathfrak{g}_\alpha) = \mathfrak{g}_{-\alpha}$. Indeed, if $[H,X] = \alpha(H)X$ then

$$[H,\theta(X)] = \theta([-H,X]) = -\alpha(H)\theta(X)\,.$$

In particular, $\theta(\mathfrak{n}^+) = \mathfrak{n}^-$.

2.4.4 Irreducibility of the Adjoint Representation

Now that we have the root space decompositions of the Lie algebras of the classical groups, we can prove the following fundamental result:

Theorem 2.4.13. *Let G be one of the groups* $\mathbf{SL}(\mathbb{C}^{l+1})$, $\mathbf{Sp}(\mathbb{C}^{2l})$, $\mathbf{SO}(\mathbb{C}^{2l+1})$ *with* $l \geq 1$, *or* $\mathbf{SO}(\mathbb{C}^{2l})$ *with* $l \geq 3$. *Then the adjoint representation of G is irreducible.*

Proof. By Theorems 2.2.2 and 2.2.7 it will suffice to show that $\mathrm{ad}(\mathfrak{g})$ acts irreducibly on $\mathfrak{g} = \mathrm{Lie}(G)$. Let Φ, Φ^+, Δ, and $\widetilde{\alpha}$ be as in Lemma 2.4.10.

Suppose U is a nonzero $\mathrm{ad}(\mathfrak{g})$-invariant subspace of $\mathfrak{g}$. We shall prove that $U = \mathfrak{g}$. Since $[\mathfrak{h}, U] \subset U$ and each root space $\mathfrak{g}_\alpha$ has dimension one, we have a decomposition

$$U = \left(U \cap \mathfrak{h}\right) \oplus \left(\bigoplus_{\alpha \in S} \mathfrak{g}_\alpha\right) ,$$

where $S = \{\alpha \in \Phi : \mathfrak{g}_\alpha \subset U\}$. We claim that

(1) S is nonempty.

Indeed, if $U \subset \mathfrak{h}$, then we would have $[U, \mathfrak{g}_\alpha] \subset U \cap \mathfrak{g}_\alpha = 0$ for all $\alpha \in \Phi$. Hence $\alpha(U) = 0$ for all roots α, which would imply $U = 0$, since the roots span $\mathfrak{h}^*$, a contradiction. This proves (1). Next we prove

(2) $U \cap \mathfrak{h} \neq 0$.

To see this, take $\alpha \in S$. Then by Lemma 2.4.2 we have $h_\alpha = -[f_\alpha, e_\alpha] \in U \cap \mathfrak{h}$. Now let $\alpha \in \Phi$. Then we have the following:

(3) If $\alpha(U \cap \mathfrak{h}) \neq 0$ then $\mathfrak{g}_\alpha \subset U$.

Indeed, $[U \cap \mathfrak{h}, \mathfrak{g}_\alpha] = \mathfrak{g}_\alpha$ in this case.

From (3) we see that if $\alpha \in S$ then $-\alpha \in S$. Set $S^+ = S \cap \Phi^+$. If $\alpha \in S^+$ and $\alpha \neq \tilde{\alpha}$, then by Lemma 2.4.10 (3) there exists $\gamma \in \Phi^+$ such that $\alpha + \gamma \in \Phi$. Since $[\mathfrak{g}_\alpha, \mathfrak{g}_\gamma] = \mathfrak{g}_{\alpha+\gamma}$ by Corollary 2.4.4, we see that $\mathfrak{g}_{\alpha+\gamma} \subset U$. Hence $\alpha + \gamma \in S^+$ and has a height greater than that of α. Thus if $\beta \in S^+$ has maximum height among the elements of S^+, then $\beta = \tilde{\alpha}$. This proves that $\widetilde{\alpha} \in S^+$. We can now prove

(4) $S = \Phi$.

By (3) it suffices to show that $S^+ = \Phi^+$. Given $\alpha \in \Phi^+$ choose $i_1, \ldots, i_r$ as in Lemma 2.4.10 (5) and set

$$\beta_j = \widetilde{\alpha} - \alpha_{i_1} - \cdots - \alpha_{i_j} \quad \text{for } \ j = 1, \ldots, r \ .$$

Write $F_i = f_{\alpha_i}$ for the element in Lemma 2.4.2. Then by induction on j and Corollary 2.4.4 we have

$$\mathfrak{g}_{\beta_j} = \mathrm{ad}(F_{i_j}) \cdots \mathrm{ad}(F_{i_1}) \mathfrak{g}_{\widetilde{\alpha}} \subset U \quad \text{for } \ j = 1, \ldots, r \ .$$

Taking $j = r$, we conclude that $\mathfrak{g}_\alpha \subset U$, which proves (4). Hence $U \cap \mathfrak{h} = \mathfrak{h}$, since $\mathfrak{h} \subset [\mathfrak{n}^+, \mathfrak{n}^-]$. This shows that $U = \mathfrak{g}$. ☐

Remark 2.4.14. For any Lie algebra $\mathfrak{g}$, the subspaces of $\mathfrak{g}$ that are invariant under $\mathrm{ad}(\mathfrak{g})$ are the *ideals* of $\mathfrak{g}$. A Lie algebra is called *simple* if it is not abelian and it has no proper ideals. (By this definition the one-dimensional Lie algebra is not simple, even though it has no proper ideals.) The classical Lie algebras occurring in Theorem 2.4.13 are thus simple. Note that their Dynkin diagrams are connected graphs.

Remark 2.4.15. A Lie algebra is called *semisimple* if it is a direct sum of simple Lie algebras. The low-dimensional orthogonal Lie algebras excluded from Theorem 2.4.11 and Theorem 2.4.13 are $\mathfrak{so}(4,\mathbb{C}) \cong \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{sl}(2,\mathbb{C})$, which is semisimple (with a disconnected Dynkin diagram), and $\mathfrak{so}(2,\mathbb{C}) \cong \mathfrak{gl}(1,\mathbb{C})$, which is abelian (and has no roots).

2.4.5 Exercises

1. For each type of classical group write out the coroots in terms of the ε_i (after the identification of $\mathfrak{h}$ with $\mathfrak{h}^*$ as in Section 2.4.1). Show that for types A and D the roots and coroots are the same. Show that for type B the coroots are the same as the roots for C and vice versa.
2. Let G be a classical group. Let Φ be the root system for G, $\alpha_1,\dots,\alpha_l$ the simple roots, and Φ^+ the positive roots as in Lemma 2.4.10. Verify that the calculations in Section 2.4.3 can be expressed as follows:
(a) For G of type A_l, $\Phi^+ \setminus \Delta$ consists of the roots
$$\alpha_i+\cdots+\alpha_j \quad \text{for } 1 \le i < j \le l\,.$$
(b) For G of type B_l with $l \ge 2$, $\Phi^+ \setminus \Delta$ consists of the roots
$$\begin{aligned} \alpha_i+\cdots+\alpha_j &\qquad \text{for } 1 \le i < j \le l\,,\\ \alpha_i+\cdots+\alpha_{j-1}+2\alpha_j+\cdots+2\alpha_l &\qquad \text{for } 1 \le i < j \le l\,. \end{aligned}$$
(c) For G of type C_l with $l \ge 2$, $\Phi^+ \setminus \Delta$ consists of the roots
$$\begin{aligned} \alpha_i+\cdots+\alpha_j &\qquad \text{for } 1 \le i < j \le l\,,\\ \alpha_i+\cdots+\alpha_{j-1}+2\alpha_j+\cdots+2\alpha_{l-1}+\alpha_l &\qquad \text{for } 1 \le i < j < l\,,\\ 2\alpha_i+\cdots+2\alpha_{l-1}+\alpha_l &\qquad \text{for } 1 \le i < l\,. \end{aligned}$$
(d) For G of type D_l with $l \ge 3$, $\Phi^+ \setminus \Delta$ consists of the roots
$$\begin{aligned} \alpha_i+\cdots+\alpha_j &\quad \text{for } 1 \le i < j < l\,,\\ \alpha_i+\cdots+\alpha_l &\quad \text{for } 1 \le i < l-1\,,\\ \alpha_i+\cdots+\alpha_{l-2}+\alpha_l &\quad \text{for } 1 \le i < l-1\,,\\ \alpha_i+\cdots+\alpha_{j-1}+2\alpha_j+\cdots+2\alpha_{l-2}+\alpha_{l-1}+\alpha_l &\quad \text{for } 1 \le i < j < l-1\,. \end{aligned}$$
Now use (a)–(d) to prove assertions (2)–(5) in Lemma 2.4.10.

3. (Assumptions and notation as in Lemma 2.4.10.) Let $S \subset \Delta$ be any subset that corresponds to a *connected* subgraph of the Dynkin diagram of Δ. Use the previous exercise to verify that $\sum_{\alpha \in S} \alpha$ is a root.
4. (Assumptions and notation as in Lemma 2.4.2 and Lemma 2.4.10.) Let $1 \leq i, j \leq l$ with $i \neq j$ and let C_{ij} be the Cartan integers.
(a) Show that the α_j root string through α_i is $\alpha_i, \ldots, \alpha_i - C_{ji}\alpha_j$. (HINT: Use the fact that $\alpha_i - \alpha_j$ is not a root and the proof of Corollary 2.4.5.)
(b) Show that $[e_{\alpha_j}, e_{-\alpha_i}] = 0$ and
$$\begin{aligned} \mathrm{ad}(e_{\alpha_j})^k(e_{\alpha_i}) &\neq 0 \quad \text{for } k = 0, \ldots, -C_{ji} , \\ \mathrm{ad}(e_{\alpha_j})^k(e_{\alpha_i}) &= 0 \quad \text{for } k = -C_{ji} + 1 . \end{aligned}$$
(HINT: Use (a) and Corollary 2.4.4.)
5. Consider the representation ρ of $\mathbf{SL}(4,\mathbb{C})$ on $\bigwedge^2 \mathbb{C}^4$, where $\rho(g)(v_1 \wedge v_2) = gv_1 \wedge gv_2$ for $g \in \mathbf{SL}(4,\mathbb{C})$ and $v_1, v_2 \in \mathbb{C}^4$. Let $\Omega = e_1 \wedge e_2 \wedge e_3 \wedge e_4$ and let B be the nondegenerate symmetric bilinear form such that $a \wedge b = B(a,b)\Omega$ for $a, b \in \bigwedge^2 \mathbb{C}^4$, as in Exercises 1.1.5 #6 and #7.
(a) Let $g \in \mathbf{SL}(4,\mathbb{C})$, $X \in \mathfrak{sl}(4,\mathbb{C})$, and $a, b \in \bigwedge^2 \mathbb{C}^4$. Show that
$$B(\rho(g)a, \rho(g)b) = B(a,b) \quad \text{and} \quad B(\mathrm{d}\rho(X)a, b) + B(a, \mathrm{d}\rho(X)b) = 0 .$$
(b) Use $\mathrm{d}\rho$ to obtain a Lie algebra isomorphism $\mathfrak{sl}(4,\mathbb{C}) \cong \mathfrak{so}(\bigwedge^2 \mathbb{C}^4, B)$. (HINT: $\mathfrak{sl}(4,\mathbb{C})$ is a simple Lie algebra.)
(c) Show that $\rho : \mathbf{SL}(4,\mathbb{C}) \longrightarrow \mathbf{SO}(\bigwedge^2 \mathbb{C}^4, B)$ is surjective, and $\mathrm{Ker}(\rho) = \{\pm I\}$. (HINT: For the surjectivity, use (b) and Theorem 2.2.2. To determine $\mathrm{Ker}(\rho)$, use (b) to show that $\mathrm{Ad}(g) = I$ for all $g \in \mathrm{Ker}(\rho)$, and then use Theorem 2.1.5.)
6. Let B be the symmetric bilinear form on $\bigwedge^2 \mathbb{C}^4$ and ρ the representation of $\mathbf{SL}(4,\mathbb{C})$ on $\bigwedge^2 \mathbb{C}$ as in the previous exercise. Let $\omega = e_1 \wedge e_4 + e_2 \wedge e_3$. Identify $\mathbb{C}^4$ with $(\mathbb{C}^4)^*$ by the inner product $(x,y) = x^t y$, so that ω can also be viewed as a skew-symmetric bilinear form on $\mathbb{C}^4$. Define
$$\mathcal{L} = \{a \in \textstyle\bigwedge^2 \mathbb{C}^4 \, : \, B(a, \omega) = 0\} .$$
Then $\rho(g)\mathcal{L} \subset \mathcal{L}$ for all $g \in \mathbf{Sp}(\mathbb{C}^4, \omega)$ and $\bigwedge^2 \mathbb{C}^4 = \mathbb{C}\omega \oplus \mathcal{L}$. Furthermore, if β is the restriction of the bilinear form B to $\mathcal{L} \times \mathcal{L}$, then β is nondegenerate (see Exercises 1.1.5 #8).
(a) Let $\varphi(g)$ be the restriction of $\rho(g)$ to the subspace $\mathcal{L}$, for $g \in \mathbf{Sp}(\mathbb{C}^4, \omega)$. Use $\mathrm{d}\varphi$ to obtain a Lie algebra isomorphism $\mathfrak{sp}(\mathbb{C}^4, \omega) \cong \mathfrak{so}(\mathbb{C}^5, \beta)$. (HINT: $\mathfrak{sp}(\mathbb{C}^4, \omega)$ is a simple Lie algebra.)
(b) Show that $\varphi : \mathbf{Sp}(\mathbb{C}^4, \omega) \longrightarrow \mathbf{SO}(\mathcal{L}, \beta)$ is surjective and $\mathrm{Ker}(\varphi) = \{\pm I\}$. (HINT: For the surjectivity, use Theorem 2.2.2. To determine $\mathrm{Ker}(\varphi)$, use (a) to show that $\mathrm{Ad}(g) = I$ for all $g \in \mathrm{Ker}(\varphi)$, and then use Theorem 2.1.5.)

2.5 Semisimple Lie Algebras

We will show that the structural features of the Lie algebras of the classical groups studied in Section 2.4 carry over to the class of *semisimple* Lie algebras. This requires some preliminary general results on Lie algebras. These results will be used again in Chapters 11 and 12, but the remainder of the current chapter may be omitted by the reader interested only in the classical groups (in fact, it turns out that there are only five *exceptional* simple Lie algebras, traditionally labeled E_6, E_7, E_8, F_4, and G_2, that are not Lie algebras of classical groups).

2.5.1 Solvable Lie Algebras

We begin with a Lie-algebraic condition for nilpotence of a linear transformation.

Lemma 2.5.1. *Let V be a finite-dimensional complex vector space and let $A \in \mathrm{End}(V)$. Suppose there exist $X_i, Y_i \in \mathrm{End}(V)$ such that $A = \sum_{i=1}^{k} [X_i, Y_i]$ and $[A, X_i] = 0$ for all i. Then A is nilpotent.*

Proof. Let Σ be the spectrum of A, and let $\{P_\lambda\}_{\lambda\in\Sigma}$ be the resolution of the identity for A (see Lemma B.1.1). Then $P_\lambda X_i = X_i P_\lambda = P_\lambda X_i P_\lambda$ for all i, so

$$P_\lambda [X_i, Y_i] P_\lambda = P_\lambda X_i P_\lambda Y_i P_\lambda - P_\lambda Y_i P_\lambda X_i P_\lambda = [P_\lambda X_i P_\lambda, P_\lambda Y_i P_\lambda] .$$

Hence $\mathrm{tr}(P_\lambda [X_i, Y_i] P_\lambda) = 0$ for all i, so we obtain $\mathrm{tr}(P_\lambda A) = 0$ for all $\lambda \in \Sigma$. However, $\mathrm{tr}(P_\lambda A) = \lambda \dim V_\lambda$, where

$$V_\lambda = \{v \in V : (A - \lambda)^k v = 0 \quad \text{for some } k\} .$$

It follows that $V_\lambda = 0$ for all $\lambda \neq 0$, so that A is nilpotent. □

Definition 2.5.2. A finite-dimensional representation (π, V) of a Lie algebra $\mathfrak{g}$ is *completely reducible* if every $\mathfrak{g}$-invariant subspace $W \subset V$ has a $\mathfrak{g}$-invariant complementary subspace U. Thus $W \cap U = \{0\}$ and $V = W \oplus U$.

Theorem 2.5.3. *Let V be a finite-dimensional complex vector space. Suppose $\mathfrak{g}$ is a Lie subalgebra of $\mathrm{End}(V)$ such that V is completely reducible as a representation of $\mathfrak{g}$. Let $\mathfrak{z} = \{X \in \mathfrak{g} : [X, Y] = 0$ for all $Y \in \mathfrak{g}\}$ be the center of $\mathfrak{g}$. Then*

1. *every $A \in \mathfrak{z}$ is a semisimple linear transformation;*
2. $[\mathfrak{g}, \mathfrak{g}] \cap \mathfrak{z} = 0$ *;*
3. $\mathfrak{g}/\mathfrak{z}$ *has no nonzero abelian ideal.*

Proof. Complete reducibility implies that $V = \bigoplus_i V_i$, where each V_i is invariant and irreducible under the action of $\mathfrak{g}$. If $Z \in \mathfrak{z}$ then the restriction of Z to V_i commutes with the action of $\mathfrak{g}$, hence is a scalar by Schur's lemma (Lemma 4.1.4). This proves (1). Then (2) follows from (1) and Lemma 2.5.1.

To prove (3), let $\mathfrak{a} \subset \mathfrak{g}/\mathfrak{z}$ be an abelian ideal. Then $\mathfrak{a} = \mathfrak{h}/\mathfrak{z}$, where $\mathfrak{h}$ is an ideal in $\mathfrak{g}$ such that $[\mathfrak{h},\mathfrak{h}] \subset \mathfrak{z}$. But by (2) this implies that $[\mathfrak{h},\mathfrak{h}] = 0$, so $\mathfrak{h}$ is an abelian ideal in $\mathfrak{g}$. Let $\mathcal{B}$ be the associative subalgebra of $\mathrm{End}(V)$ generated by $[\mathfrak{h},\mathfrak{g}]$. By Lemma 2.5.1 we know that the elements of $[\mathfrak{h},\mathfrak{g}]$ are nilpotent endomorphisms of V. Since $[\mathfrak{h},\mathfrak{g}] \subset \mathfrak{h}$ is abelian, it follows that the elements of $\mathcal{B}$ are nilpotent endomorphisms. If we can prove that $\mathcal{B} = 0$, then $\mathfrak{h} \subset \mathfrak{z}$ and hence $\mathfrak{a} = 0$, establishing (3).

We now turn to the proof that $\mathcal{B} = 0$. Let $\mathcal{A}$ be the associative subalgebra of $\mathrm{End}(V)$ generated by $\mathfrak{g}$. We claim that

$$\mathcal{A}\mathcal{B} \subset \mathcal{B}\mathcal{A} + \mathcal{B}\,. \tag{2.38}$$

Indeed, for $X,Y \in \mathfrak{g}$ and $Z \in \mathfrak{h}$ we have $[X,[Y,Z]] \in [\mathfrak{g},\mathfrak{h}]$ by the Jacobi identity, since $\mathfrak{h}$ is an ideal. Hence

$$X[Y,Z] = [Y,Z]X + [X,[Y,Z]] \in \mathcal{B}\mathcal{A} + \mathcal{B}\,. \tag{2.39}$$

Let $b \in \mathcal{B}$ and suppose that $Xb \in \mathcal{B}\mathcal{A} + \mathcal{B}$. Then by (2.39) we have

$$X[Y,Z]b = [Y,Z]Xb + [X,[Y,Z]]b \in [Y,Z]\mathcal{B}\mathcal{A} + \mathcal{B} \subset \mathcal{B}\mathcal{A} + \mathcal{B}\,.$$

Now (2.38) follows from this last relation by induction on the degree (in terms of the generators from $\mathfrak{g}$ and $[\mathfrak{h},\mathfrak{g}]$) of the elements in $\mathcal{A}$ and $\mathcal{B}$.

We next show that

$$(\mathcal{A}\mathcal{B})^k \subset \mathcal{B}^k\mathcal{A} + \mathcal{B}^k \tag{2.40}$$

for every positive integer k. This is true for $k = 1$ by (2.38). Assuming that it holds for k, we use (2.38) to get the inclusions

$$\begin{aligned}(\mathcal{A}\mathcal{B})^{k+1} = (\mathcal{A}\mathcal{B})^k(\mathcal{A}\mathcal{B}) \subset (\mathcal{B}^k\mathcal{A} + \mathcal{B}^k)(\mathcal{A}\mathcal{B}) &\subset \mathcal{B}^k\mathcal{A}\mathcal{B} \\ &\subset \mathcal{B}^k(\mathcal{B}\mathcal{A} + \mathcal{B}) \subset \mathcal{B}^{k+1}\mathcal{A} + \mathcal{B}^{k+1}\,.\end{aligned}$$

Hence (2.40) holds for all k.

We now complete the proof as follows. Since $\mathcal{B}^k = 0$ for k sufficiently large, the same is true for $(\mathcal{A}\mathcal{B})^k$ by (2.40). Suppose $(\mathcal{A}\mathcal{B})^{k+1} = 0$ for some $k \geq 1$. Set $\mathcal{C} = (\mathcal{A}\mathcal{B})^k$. Then $\mathcal{C}^2 = 0$. Set $W = \mathcal{C}V$. Since $\mathcal{A}\mathcal{C} \subset \mathcal{C}$, the subspace W is $\mathcal{A}$-invariant. Hence by complete reducibility of V relative to the action of $\mathfrak{g}$, there is an $\mathcal{A}$-invariant complementary subspace U such that $V = W \oplus U$. Now $\mathcal{C}W = \mathcal{C}^2V = 0$ and $\mathcal{C}U \subset \mathcal{C}V = W$. But $\mathcal{C}U \subset U$ also, so $\mathcal{C}U \subset U \cap W = \{0\}$. Hence $\mathcal{C}V = 0$. Thus $\mathcal{C} = 0$. It follows (by downward induction on k) that $\mathcal{A}\mathcal{B} = 0$. Since $I \in \mathcal{A}$, we conclude that $\mathcal{B} = 0$. □

For a Lie algebra $\mathfrak{g}$ we define the *derived algebra* $\mathcal{D}(\mathfrak{g}) = [\mathfrak{g},\mathfrak{g}]$ and we set $\mathcal{D}^{k+1}(\mathfrak{g}) = \mathcal{D}(\mathcal{D}^k(\mathfrak{g}))$ for $k = 1,2,\ldots$. One shows by induction on k that $\mathcal{D}^k(\mathfrak{g})$ is invariant under all derivations of $\mathfrak{g}$. In particular, $\mathcal{D}^k(\mathfrak{g})$ is an ideal in $\mathfrak{g}$ for each k, and $\mathcal{D}^k(\mathfrak{g})/cD^{k+1}(\mathfrak{g})$ is abelian.

Definition 2.5.4. $\mathfrak{g}$ is *solvable* if there exists an integer $k \geq 1$ such that $\mathcal{D}^k\mathfrak{g} = 0$.

It is clear from the definition that a Lie subalgebra of a solvable Lie algebra is also solvable. Also, if $\pi : \mathfrak{g} \longrightarrow \mathfrak{h}$ is a surjective Lie algebra homomorphism, then

$$\pi(\mathcal{D}^k(\mathfrak{g})) = \mathcal{D}^k(\mathfrak{h}) .$$

Hence the solvability of $\mathfrak{g}$ implies the solvability of $\mathfrak{h}$. Furthermore, if $\mathfrak{g}$ is a nonzero solvable Lie algebra and we choose k such that $\mathcal{D}^k(\mathfrak{g}) \neq 0$ and $\mathcal{D}^{k+1}(\mathfrak{g}) = 0$, then $\mathcal{D}^k(\mathfrak{g})$ is an abelian ideal in $\mathfrak{g}$ that is invariant under all derivations of $\mathfrak{g}$.

Remark 2.5.5. The archetypical example of a solvable Lie algebra is the $n \times n$ upper-triangular matrices $\mathfrak{b}_n$. Indeed, we have $\mathcal{D}(\mathfrak{b}_n) = \mathfrak{n}_n^+$, the Lie algebra of $n \times n$ upper-triangular matrices with zeros on the main diagonal. If $\mathfrak{n}_{n,r}^+$ is the Lie subalgebra of $\mathfrak{n}_n^+$ consisting of matrices $X = [x_{ij}]$ such that $x_{ij} = 0$ for $j - i \leq r - 1$, then $\mathfrak{n}_n^+ = \mathfrak{n}_{n,1}^+$ and $[\mathfrak{n}_n^+, \mathfrak{n}_{n,r}^+] \subset \mathfrak{n}_{n,r+1}^+$ for $r = 1, 2, \ldots$. Hence $\mathcal{D}^k(\mathfrak{b}_n) \subset \mathfrak{n}_{n,k}^+$, and so $\mathcal{D}^k(\mathfrak{b}_n) = 0$ for $k > n$.

Corollary 2.5.6. *Suppose $\mathfrak{g} \subset \mathrm{End}(V)$ is a solvable Lie algebra and that V is completely reducible as a $\mathfrak{g}$-module. Then $\mathfrak{g}$ is abelian. In particular, if V is an irreducible $\mathfrak{g}$-module, then* $\dim V = 1$.

Proof. Let $\mathfrak{z}$ be the center of $\mathfrak{g}$. If $\mathfrak{z} \neq \mathfrak{g}$, then $\mathfrak{g}/\mathfrak{z}$ would be a nonzero solvable Lie algebra and hence would contain a nonzero abelian ideal. But this would contradict part (3) of Theorem 2.5.3, so we must have $\mathfrak{z} = \mathfrak{g}$. Given that $\mathfrak{g}$ is abelian and V is completely reducible, we can find a basis for V consisting of simultaneous eigenvectors for all the transformations $X \in \mathfrak{g}$; thus V is the direct sum of invariant one-dimensional subspaces. This implies the last statement of the corollary. □

We can now obtain Cartan's trace-form criterion for solvability of a Lie algebra.

Theorem 2.5.7. *Let V be a finite-dimensional complex vector space. Let $\mathfrak{g} \subset \mathrm{End}(V)$ be a Lie subalgebra such that* $\mathrm{tr}(XY) = 0$ *for all $X, Y \in \mathfrak{g}$. Then $\mathfrak{g}$ is solvable.*

Proof. We use induction on $\dim \mathfrak{g}$. A one-dimensional Lie algebra is solvable. Also, if $[\mathfrak{g}, \mathfrak{g}]$ is solvable, then so is $\mathfrak{g}$, since $\mathcal{D}^{k+1}(\mathfrak{g}) = \mathcal{D}^k([\mathfrak{g}, \mathfrak{g}])$. Thus by induction we need to consider only the case $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$.

Take any maximal proper Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Then $\mathfrak{h}$ is solvable, by induction. Hence the natural representation of $\mathfrak{h}$ on $\mathfrak{g}/\mathfrak{h}$ has a one-dimensional invariant subspace, by Corollary 2.5.6. This means that there exist $0 \neq Y \in \mathfrak{g}$ and $\mu \in \mathfrak{h}^*$ such that

$$[X, Y] \equiv \mu(X) Y \qquad (\mathrm{mod}\ \mathfrak{h})$$

for all $X \in \mathfrak{h}$. But this commutation relation implies that $\mathbb{C}Y + \mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$. Since $\mathfrak{h}$ was chosen as a maximal subalgebra, we must have $\mathbb{C}Y + \mathfrak{h} = \mathfrak{g}$. Furthermore, $\mu \neq 0$ because we are assuming $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$.

Given the structure of $\mathfrak{g}$ as above, we next determine the structure of an arbitrary irreducible $\mathfrak{g}$-module (π, W). By Corollary 2.5.6 again, there exist $w_0 \in W$ and $\sigma \in \mathfrak{h}^*$ such that

$$\pi(X) w_0 = \sigma(X) w_0 \quad \text{for all } X \in \mathfrak{h} .$$

Set $w_k = \pi(Y)^k w_0$ and $W_k = \mathbb{C}w_k + \cdots + \mathbb{C}w_0$. We claim that for $X \in \mathfrak{h}$,

$$\pi(X)w_k \equiv (\sigma(X) + k\mu(X))w_k \qquad (\mathrm{mod}\ W_{k-1}) \tag{2.41}$$

(where $W_{-1} = \{0\}$). Indeed, this is true for $k = 0$ by definition. If it holds for k then $\pi(\mathfrak{h})W_k \subset W_k$ and

$$\begin{aligned}\pi(X)w_{k+1} &= \pi(X)\pi(Y)w_k = \pi(Y)\pi(X)w_k + \pi([X,Y])w_k \\ &\equiv (\sigma(X) + (k+1)\mu(X))w_{k+1} \qquad (\mathrm{mod}\ W_k)\ .\end{aligned}$$

Thus (2.41) holds for all k. Let m be the smallest integer such that $W_m = W_{m+1}$. Then W_m is invariant under $\mathfrak{g}$, and hence $W_m = W$ by irreducibility. Thus $\dim W = m+1$ and

$$\mathrm{tr}(\pi(X)) = \sum_{k=0}^{m} \sigma(X) + k\mu(X) = (m+1)\Big(\sigma(X) + \frac{m}{2}\mu(X)\Big)$$

for all $X \in \mathfrak{h}$. However, $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$, so $\mathrm{tr}(\pi(X)) = 0$. Thus

$$\sigma(X) = -\frac{m}{2}\mu(X) \quad \text{for all } X \in \mathfrak{h}\ .$$

From (2.41) again we get

$$\mathrm{tr}(\pi(X)^2) = \sum_{k=0}^{m}\Big(k - \frac{m}{2}\Big)^2 \mu(X)^2 \quad \text{for all } X \in \mathfrak{h}\ . \tag{2.42}$$

We finally apply these results to the given representation of $\mathfrak{g}$ on V. Take a composition series $\{0\} = V_0 \subset V_1 \subset \cdots \subset V_r = V$, where each subspace V_j is invariant under $\mathfrak{g}$ and $W_i = V_i/V_{i-1}$ is an irreducible $\mathfrak{g}$-module. Write $\dim W_i = m_i + 1$. Then (2.42) implies that

$$\mathrm{tr}_V(X^2) = \mu(X)^2 \sum_{i=1}^{r}\sum_{k=0}^{m_i}\Big(k - \frac{1}{2}m_i\Big)^2$$

for all $X \in \mathfrak{h}$. But by assumption, $\mathrm{tr}_V(X^2) = 0$ and there exists $X \in \mathfrak{h}$ with $\mu(X) \neq 0$. This forces $m_i = 0$ for $i = 1, \ldots, r$. Hence $\dim W_i = 1$ for each i. Since $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$, this implies that $\mathfrak{g}V_i \subset V_{i-1}$. If we take a basis for V consisting of a nonzero vector from each W_i, then the matrices for $\mathfrak{g}$ relative to this basis are strictly upper triangular. Hence $\mathfrak{g}$ is solvable, by Remark 2.5.5. □

Recall that a finite-dimensional Lie algebra is *simple* if it is not abelian and has no proper ideals.

Corollary 2.5.8. *Let $\mathfrak{g}$ be a Lie subalgebra of* $\mathrm{End}(V)$ *that has no nonzero abelian ideals. Then the bilinear form* $\mathrm{tr}(XY)$ *on $\mathfrak{g}$ is nondegenerate, and $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$ (Lie algebra direct sum), where each $\mathfrak{g}_i$ is a simple Lie algebra.*

Proof. Let $\mathfrak{r} = \{X \in \mathfrak{g} : \mathrm{tr}(XY) = 0 \quad \text{for all } Y \in \mathfrak{g}\}$ be the radical of the trace form. Then $\mathfrak{r}$ is an ideal in $\mathfrak{g}$, and by Cartan's criterion $\mathfrak{r}$ is a solvable Lie algebra. Suppose

$\mathfrak{r} \neq 0$. Then $\mathfrak{r}$ contains a nonzero abelian ideal $\mathfrak{a}$ that is invariant under all derivations of $\mathfrak{r}$. Hence $\mathfrak{a}$ is an abelian ideal in $\mathfrak{g}$, which is a contradiction. Thus the trace form is nondegenerate.

To prove the second assertion, let $\mathfrak{g}_1 \subset \mathfrak{g}$ be an irreducible subspace for the adjoint representation of $\mathfrak{g}$ and define

$$\mathfrak{g}_1^{\perp} = \{X \in \mathfrak{g} : \operatorname{tr}(XY) = 0 \quad \text{for all } Y \in \mathfrak{g}_1\} .$$

Then $\mathfrak{g}_1^{\perp}$ is an ideal in $\mathfrak{g}$, and $\mathfrak{g}_1 \cap \mathfrak{g}_1^{\perp}$ is solvable by Cartan's criterion. Hence $\mathfrak{g}_1 \cap \mathfrak{g}_1^{\perp} = 0$ by the same argument as before. Thus $[\mathfrak{g}_1, \mathfrak{g}_1^{\perp}] = 0$, so we have the decomposition

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_1^{\perp} \quad \text{(direct sum of Lie algebras)} .$$

In particular, $\mathfrak{g}_1$ is irreducible as an $\operatorname{ad}\mathfrak{g}_1$-module. It cannot be abelian, so it is a simple Lie algebra. Now use induction on $\dim \mathfrak{g}$. □

Corollary 2.5.9. *Let V be a finite-dimensional complex vector space. Suppose $\mathfrak{g}$ is a Lie subalgebra of* $\operatorname{End}(V)$ *such that V is completely reducible as a representation of $\mathfrak{g}$. Let $\mathfrak{z} = \{X \in \mathfrak{g} : [X,Y] = 0$ for all $Y \in \mathfrak{g}\}$ be the center of $\mathfrak{g}$. Then the derived Lie algebra $[\mathfrak{g},\mathfrak{g}]$ is semisimple, and $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}] \oplus \mathfrak{z}$.*

Proof. Theorem 2.5.3 implies that $\mathfrak{g}/\mathfrak{z}$ has no nonzero abelian ideals; hence $\mathfrak{g}/\mathfrak{z}$ is semisimple (Corollary 2.5.8). Since $\mathfrak{g}/\mathfrak{z}$ is a direct sum of simple algebras, it satisfies $[\mathfrak{g}/\mathfrak{z}, \mathfrak{g}/\mathfrak{z}] = \mathfrak{g}/\mathfrak{z}$. Let $p : \mathfrak{g} \longrightarrow \mathfrak{g}/\mathfrak{z}$ be the natural surjection. If $u, v \in \mathfrak{g}$ then $p([u,v]) = [p(u), p(v)]$. Since p is surjective, it follows that $\mathfrak{g}/\mathfrak{z}$ is spanned by the elements $p([u,v])$ for $u, v \in \mathfrak{g}$. Thus $p([\mathfrak{g},\mathfrak{g}]) = \mathfrak{g}/\mathfrak{z}$. Now Theorem 2.5.3 (2) implies that the restriction of p to $[\mathfrak{g},\mathfrak{g}]$ gives a Lie algebra isomorphism with $\mathfrak{g}/\mathfrak{z}$ and that $\dim([\mathfrak{g},\mathfrak{g}]) + \dim \mathfrak{z} = \dim \mathfrak{g}$. Hence $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}] \oplus \mathfrak{z}$. □

Let $\mathfrak{g}$ be a finite-dimensional complex Lie algebra.

Definition 2.5.10. The *Killing form* of $\mathfrak{g}$ is the bilinear form $B(X,Y) = \operatorname{tr}(\operatorname{ad}X \operatorname{ad}Y)$ for $X, Y \in \mathfrak{g}$.

Recall that $\mathfrak{g}$ is *semisimple* if it is the direct sum of simple Lie algebras. We now obtain *Cartan's criterion* for semisimplicity.

Theorem 2.5.11. *The Lie algebra $\mathfrak{g}$ is semisimple if and only if its Killing form is nondegenerate.*

Proof. Assume that $\mathfrak{g}$ is semisimple. Since the adjoint representation of a simple Lie algebra is faithful, the same is true for a semisimple Lie algebra. Hence a semisimple Lie algebra $\mathfrak{g}$ is isomorphic to a Lie subalgebra of $\operatorname{End}(\mathfrak{g})$. Let

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$$

(Lie algebra direct sum), where each $\mathfrak{g}_i$ is a simple Lie algebra. If $\mathfrak{m}$ is an abelian ideal in $\mathfrak{g}$, then $\mathfrak{m} \cap \mathfrak{g}_i$ is an abelian ideal in $\mathfrak{g}_i$, for each i, and hence is zero. Thus $\mathfrak{m} = 0$. Hence B is nondegenerate by Corollary 2.5.8.

Conversely, suppose the Killing form is nondegenerate. Then the adjoint representation is faithful. To show that $\mathfrak{g}$ is semisimple, it suffices by Corollary 2.5.8 to show that $\mathfrak{g}$ has no nonzero abelian ideals.

Suppose $\mathfrak{a}$ is an ideal in $\mathfrak{g}$, $X \in \mathfrak{a}$, and $Y \in \mathfrak{g}$. Then $\operatorname{ad}X \operatorname{ad}Y$ maps $\mathfrak{g}$ into $\mathfrak{a}$ and leaves $\mathfrak{a}$ invariant. Hence

$$B(X,Y) = \operatorname{tr}(\operatorname{ad}X|_{\mathfrak{a}} \operatorname{ad}Y|_{\mathfrak{a}}) . \tag{2.43}$$

If $\mathfrak{a}$ is an abelian ideal, then $\operatorname{ad}X|_{\mathfrak{a}} = 0$. Since B is nondegenerate, (2.43) implies that $X = 0$. Thus $\mathfrak{a} = 0$. $\square$

Corollary 2.5.12. *Suppose* $\mathfrak{g}$ *is a semisimple Lie algebra and* $D \in \operatorname{Der}(\mathfrak{g})$. *Then there exists* $X \in \mathfrak{g}$ *such that* $D = \operatorname{ad}X$.

Proof. The derivation property $D([Y,Z]) = [D(Y),Z] + [Y,D(Z)]$ can be expressed as the commutation relation

$$[D, \operatorname{ad}Y] = \operatorname{ad}D(Y) \quad \text{for all } Y \in \mathfrak{g} . \tag{2.44}$$

Consider the linear functional $Y \mapsto \operatorname{tr}(D \operatorname{ad}Y)$ on $\mathfrak{g}$. Since the Killing form is nondegenerate, there exists $X \in \mathfrak{g}$ such that $\operatorname{tr}(D\operatorname{ad}Y) = B(X,Y)$ for all $Y \in \mathfrak{g}$. Take $Y, Z \in \mathfrak{g}$ and use the invariance of B to obtain

$$\begin{aligned} B(\operatorname{ad}X(Y), Z) &= B(X,[Y,Z]) = \operatorname{tr}(D\operatorname{ad}[Y,Z]) = \operatorname{tr}(D[\operatorname{ad}Y, \operatorname{ad}Z]) \\ &= \operatorname{tr}(D\operatorname{ad}Y\operatorname{ad}Z) - \operatorname{tr}(D\operatorname{ad}Z\operatorname{ad}Y) = \operatorname{tr}([D,\operatorname{ad}Y]\operatorname{ad}Z) . \end{aligned}$$

Hence (2.44) and the nondegeneracy of B give $\operatorname{ad}X = D$. $\square$

For the next result we need the following formula, valid for any elements Y, Z in a Lie algebra $\mathfrak{g}$, any $D \in \operatorname{Der}(\mathfrak{g})$, and any scalars λ, μ:

$$\big(D - (\lambda+\mu)\big)^k [Y,Z] = \sum_r \binom{k}{r} [(D-\lambda)^r Y,\, (D-\mu)^{k-r} Z] . \tag{2.45}$$

(The proof is by induction on k using the derivation property and the inclusion–exclusion identity for binomial coefficients.)

Corollary 2.5.13. *Let* $\mathfrak{g}$ *be a semisimple Lie algebra. If* $X \in \mathfrak{g}$ *and* $\operatorname{ad}X = S + N$ *is the additive Jordan decomposition in* $\operatorname{End}(\mathfrak{g})$ *(with S semisimple, N nilpotent, and* $[S,N] = 0$*), then there exist* $X_s, X_n \in \mathfrak{g}$ *such that* $\operatorname{ad}X_s = S$ *and* $\operatorname{ad}X_n = N$.

Proof. Let $\lambda \in \mathbb{C}$ and set

$$\mathfrak{g}_\lambda(X) = \bigcup_{k \ge 1} \operatorname{Ker}(\operatorname{ad}X - \lambda)^k$$

(the generalized λ eigenspace of $\operatorname{ad}X$). The Jordan decomposition of $\operatorname{ad}X$ then gives a direct-sum decomposition

$$\mathfrak{g} = \bigoplus_{\lambda} \mathfrak{g}_\lambda(X) ,$$

and S acts by λ on $\mathfrak{g}_\lambda(X)$. Taking $D = \operatorname{ad} X$, $Y \in \mathfrak{g}_\lambda(X)$, $Z \in \mathfrak{g}_\mu(X)$, and k sufficiently large in (2.45), we see that

$$[\mathfrak{g}_\lambda(X), \mathfrak{g}_\mu(X)] \subset \mathfrak{g}_{\lambda+\mu}(X) . \tag{2.46}$$

Hence S is a derivation of $\mathfrak{g}$. By Corollary 2.5.12 there exists $X_s \in \mathfrak{g}$ such that $\operatorname{ad} X_s = S$. Set $X_n = X - X_s$. □

2.5.2 Root Space Decomposition

In this section we shall show that every semisimple Lie algebra has a *root space decomposition* with the properties that we established in Section 2.4 for the Lie algebras of the classical groups. We begin with the following Lie algebra generalization of a familiar property of nilpotent linear transformations:

Theorem 2.5.14 (Engel). *Let V be a nonzero finite-dimensional vector space and let $\mathfrak{g} \subset \operatorname{End}(V)$ be a Lie algebra. Assume that every $X \in \mathfrak{g}$ is a nilpotent linear transformation. Then there exists a nonzero vector $v_0 \in V$ such that $Xv_0 = 0$ for all $X \in \mathfrak{g}$.*

Proof. For $X \in \operatorname{End}(V)$ write L_X and R_X for the linear transformations of $\operatorname{End}(V)$ given by left and right multiplication by X, respectively. Then $\operatorname{ad} X = L_X - R_X$ and L_X commutes with R_X. Hence

$$(\operatorname{ad} X)^k = \sum_j \binom{k}{j} (-1)^{k-j} \big(L_X\big)^j \big(R_X\big)^{k-j}$$

by the binomial expansion. If X is nilpotent on V then $X^n = 0$, where $n = \dim V$. Thus $\big(L_X\big)^j \big(R_X\big)^{2n-j} = 0$ if $0 \le j \le 2n$. Hence $(\operatorname{ad} X)^{2n} = 0$, so $\operatorname{ad} X$ is nilpotent on $\operatorname{End}(V)$.

We prove the theorem by induction on $\dim \mathfrak{g}$ (when $\dim \mathfrak{g} = 1$ the theorem is clearly true). Take a proper subalgebra $\mathfrak{h} \subset \mathfrak{g}$ of maximal dimension. Then $\mathfrak{h}$ acts on $\mathfrak{g}/\mathfrak{h}$ by the adjoint representation. This action is by nilpotent linear transformations, so the induction hypothesis implies that there exists $Y \notin \mathfrak{h}$ such that

$$[X, Y] \equiv 0 \bmod \mathfrak{h} \quad \text{for all } X \in \mathfrak{h} .$$

Thus $\mathbb{C}Y + \mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$, since $[Y, \mathfrak{h}] \subset \mathfrak{h}$. But $\mathfrak{h}$ was chosen maximal, so we must have $\mathfrak{g} = \mathbb{C}Y + \mathfrak{h}$. Set

$$W = \{ v \in V : Xv = 0 \ \text{ for all } X \in \mathfrak{h} \} .$$

By the induction hypothesis we know that $W \ne 0$. If $v \in W$ then

$$XYv = YXv + [X,Y]v = 0$$

for all $X \in \mathfrak{h}$, since $[X,Y] \in \mathfrak{h}$. Thus W is invariant under Y, so there exists a nonzero vector $v_0 \in W$ such that $Yv_0 = 0$. It follows that $\mathfrak{g}v_0 = 0$. □

Corollary 2.5.15. *There exists a basis for V in which the elements of $\mathfrak{g}$ are represented by strictly upper-triangular matrices.*

Proof. This follows by repeated application of Theorem 2.5.14, replacing V by $V/\mathbb{C}v_0$ at each step. □

Corollary 2.5.16. *Suppose $\mathfrak{g}$ is a semisimple Lie algebra. Then there exists a nonzero element $X \in \mathfrak{g}$ such that* $\operatorname{ad}X$ *is semisimple.*

Proof. We argue by contradiction. If $\mathfrak{g}$ contained no nonzero elements X with $\operatorname{ad}X$ semisimple, then Corollary 2.5.13 would imply that $\operatorname{ad}X$ is nilpotent for all $X \in \mathfrak{g}$. Hence Corollary 2.5.15 would furnish a basis for $\mathfrak{g}$ such that $\operatorname{ad}X$ is strictly upper triangular. But then $\operatorname{ad}X \operatorname{ad}Y$ would also be strictly upper triangular for all $X,Y \in \mathfrak{g}$, and hence the Killing form would be zero, contradicting Theorem 2.5.11. □

For the rest of this section we let $\mathfrak{g}$ be a semisimple Lie algebra. We call a subalgebra $\mathfrak{h} \subset \mathfrak{g}$ a *toral subalgebra* if $\operatorname{ad}X$ is semisimple for all $X \in \mathfrak{h}$. Corollary 2.5.16 implies the existence of nonzero toral subalgebras.

Lemma 2.5.17. *Let $\mathfrak{h}$ be a toral subalgebra. Then* $[\mathfrak{h},\mathfrak{h}] = 0$.

Proof. Let $X \in \mathfrak{h}$. Then $\mathfrak{h}$ is an invariant subspace for the semisimple transformation $\operatorname{ad}X$. If $[X,\mathfrak{h}] \neq 0$ then there would exist an eigenvalue $\lambda \neq 0$ and an eigenvector $Y \in \mathfrak{h}$ such that $[X,Y] = \lambda Y$. But then

$$(\operatorname{ad}Y)(X) = -\lambda Y \neq 0\,, \quad (\operatorname{ad}Y)^2(X) = 0\,,$$

which would imply that $\operatorname{ad}Y$ is not a semisimple transformation. Hence we must have $[X,\mathfrak{h}] = 0$ for all $X \in \mathfrak{h}$. □

We shall call a toral subalgebra $\mathfrak{h} \subset \mathfrak{g}$ a *Cartan subalgebra* if it has maximal dimension among all toral subalgebras of $\mathfrak{g}$. From Corollary 2.5.16 and Lemma 2.5.17 we see that $\mathfrak{g}$ contains nonzero Cartan subalgebras and that Cartan subalgebras are abelian. We fix a choice of a Cartan subalgebra $\mathfrak{h}$. For $\lambda \in \mathfrak{h}^*$ let

$$\mathfrak{g}_\lambda = \{Y \in \mathfrak{g} : [X,Y] = \langle \lambda, X\rangle Y \ \text{ for all } \ X \in \mathfrak{h}\}\,.$$

In particular, $\mathfrak{g}_0 = \{Y \in \mathfrak{g} : [X,Y] = 0 \ \text{ for all } \ X \in \mathfrak{h}\}$ is the *centralizer* of $\mathfrak{h}$ in $\mathfrak{g}$. Let $\Phi \subset \mathfrak{g}^* \setminus \{0\}$ be the set of λ such that $\mathfrak{g}_\lambda \neq 0$. We call Φ the set of *roots* of $\mathfrak{h}$ on $\mathfrak{g}$. Since the mutually commuting linear transformations $\operatorname{ad}X$ are semisimple (for $X \in \mathfrak{h}$), there is a *root space decomposition*

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\lambda \in \Phi} \mathfrak{g}_\lambda\,.$$

Let B denote the Killing form of $\mathfrak{g}$. By the same arguments used for the classical groups in Sections 2.4.1 and 2.4.2 (but now using B instead of the trace form on the defining representation of a classical group), it follows that

1. $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subset \mathfrak{g}_{\lambda+\mu}$;
2. $B(\mathfrak{g}_\lambda, \mathfrak{g}_\mu) = 0$ if $\lambda + \mu \neq 0$;
3. the restriction of B to $\mathfrak{g}_0 \times \mathfrak{g}_0$ is nondegenerate;
4. if $\lambda \in \Phi$ then $-\lambda \in \Phi$ and the restriction of B to $\mathfrak{g}_\lambda \times \mathfrak{g}_{-\lambda}$ is nondegenerate.

New arguments are needed to prove the following key result:

Proposition 2.5.18. *A Cartan algebra is its own centralizer in* $\mathfrak{g}$*; thus* $\mathfrak{h} = \mathfrak{g}_0$.

Proof. Since $\mathfrak{h}$ is abelian, we have $\mathfrak{h} \subset \mathfrak{g}_0$. Let $X \in \mathfrak{g}_0$ and let $X = X_s + X_n$ be the Jordan decomposition of X given by Corollary 2.5.13.

(i) X_s and X_n are in $\mathfrak{g}_0$.

Indeed, since $[X, \mathfrak{h}] = 0$ and the adjoint representation of $\mathfrak{g}$ is faithful, we have $[X_s, \mathfrak{h}] = 0$. Hence $X_s \in \mathfrak{h}$ by the maximality of $\mathfrak{h}$, which implies that $X_n = X - X_s$ is also in $\mathfrak{h}$.

(ii) The restriction of B to $\mathfrak{h} \times \mathfrak{h}$ is nondegenerate.

To prove this, let $0 \neq h \in \mathfrak{h}$. Then by property (3) there exists $X \in \mathfrak{g}_0$ such that $B(h, X) \neq 0$. Since $X_n \in \mathfrak{g}_0$ by (i), we have $[h, X_n] = 0$ and hence $\mathrm{ad}\, h\, \mathrm{ad}\, X_n$ is nilpotent on $\mathfrak{g}$. Thus $B(h, X_n) = 0$ and so $B(h, X_s) \neq 0$. Since $X_s \in \mathfrak{h}$, this proves (ii).

(iii) $[\mathfrak{g}_0, \mathfrak{g}_0] = 0$.

For the proof of (iii), we observe that if $X \in \mathfrak{g}_0$, then $\mathrm{ad}\, X_s$ acts by zero on $\mathfrak{g}_0$, since $X_s \in \mathfrak{h}$. Hence $\mathrm{ad}\, X|_{\mathfrak{g}_0} = \mathrm{ad}\, X_n|_{\mathfrak{g}_0}$ is nilpotent. Suppose for the sake of contradiction that $[\mathfrak{g}_0, \mathfrak{g}_0] \neq 0$ and consider the adjoint action of $\mathfrak{g}_0$ on the invariant subspace $[\mathfrak{g}_0, \mathfrak{g}_0]$. By Theorem 2.5.14 there would exist $0 \neq Z \in [\mathfrak{g}_0, \mathfrak{g}_0]$ such that $[\mathfrak{g}_0, Z] = 0$. Then $[\mathfrak{g}_0, Z_n] = 0$ and hence $\mathrm{ad}\, Y\, \mathrm{ad}\, Z_n$ is nilpotent for all $Y \in \mathfrak{g}_0$. This implies that $B(Y, Z_n) = 0$ for all $Y \in \mathfrak{g}_0$, so we conclude from (3) that $Z_n = 0$. Thus $Z = Z_s$ must be in $\mathfrak{h}$. Now

$$B(h, [X, Y]) = B([h, X], Y) = 0 \quad \text{for all } \ h \in \mathfrak{h} \ \text{ and } \ X, Y \in \mathfrak{g}_0 \ .$$

Hence $\mathfrak{h} \cap [\mathfrak{g}_0, \mathfrak{g}_0] = 0$ by (ii), and so $Z = 0$, giving a contradiction.

It is now easy to complete the proof of the proposition. If $X, Y \in \mathfrak{g}_0$ then $\mathrm{ad}\, X_n\, \mathrm{ad}\, Y$ is nilpotent, since $\mathfrak{g}_0$ is abelian by (iii). Hence $B(X_n, Y) = 0$, and so $X_n = 0$ by (3). Thus $X = X_s \in \mathfrak{h}$. □

Corollary 2.5.19. *Let* $\mathfrak{g}$ *be a semisimple Lie algebra and* $\mathfrak{h}$ *a Cartan subalgebra. Then*

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\lambda \in \Phi} \mathfrak{g}_\lambda \ . \tag{2.47}$$

Hence if $Y \in \mathfrak{g}$ *and* $[Y, \mathfrak{h}] \subset \mathfrak{h}$*, then* $Y \in \mathfrak{h}$*. In particular,* $\mathfrak{h}$ *is a maximal abelian subalgebra of* $\mathfrak{g}$.

Since the form B is nondegenerate on $\mathfrak{h} \times \mathfrak{h}$, it defines a bilinear form on $\mathfrak{h}^*$ that we denote by (α, β).

Theorem 2.5.20. *The roots and root spaces satisfy the following properties:*

1. Φ *spans* $\mathfrak{h}^*$.
2. *If* $\alpha \in \Phi$ *then* $\dim[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = 1$ *and there is a unique element* $h_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ *such that* $\langle \alpha, h_\alpha \rangle = 2$ *(call* h_α *the* coroot *to* α*).*
3. *If* $\alpha \in \Phi$ *and* $c \in \mathbb{C}$ *then* $c\alpha \in \Phi$ *if and only if* $c = \pm 1$. *Also* $\dim \mathfrak{g}_\alpha = 1$.
4. *Let* $\alpha, \beta \in \Phi$ *with* $\beta \neq \pm\alpha$. *Let* p *be the largest integer* $j \geq 0$ *with* $\beta + j\alpha \in \Phi$ *and let* q *be the largest integer* $k \geq 0$ *with* $\beta - k\alpha \in \Phi$. *Then*
$$\langle \beta, h_\alpha \rangle = q - p \in \mathbb{Z} \tag{2.48}$$
and $\beta + r\alpha \in \Phi$ *for all integers* r *with* $-q \leq r \leq p$. *Hence* $\beta - \langle \beta, h_\alpha \rangle \alpha \in \Phi$.
5. *If* $\alpha, \beta \in \Phi$ *and* $\alpha + \beta \in \Phi$, *then* $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$.

Proof. (1): If $h \in \mathfrak{h}$ and $\langle \alpha, h \rangle = 0$ for all $\alpha \in \Phi$, then $[h, \mathfrak{g}_\alpha] = 0$ and hence $[h, \mathfrak{g}] = 0$. The center of $\mathfrak{g}$ is trivial, since $\mathfrak{g}$ has no abelian ideals, so $h = 0$. Thus Φ spans $\mathfrak{h}^*$.

(2): Let $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$. Then $[X,Y] \in \mathfrak{g}_0 = \mathfrak{h}$ and for $h \in \mathfrak{h}$ we have
$$B(h, [X,Y]) = B([h,X], Y) = \langle \alpha, h \rangle B(X,Y) .$$
Thus $[X,Y]$ corresponds to $B(X,Y)\alpha$ under the isomorphism $\mathfrak{h} \cong \mathfrak{h}^*$ given by the form B. Since B is nondegenerate on $\mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha}$, it follows that $\dim[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = 1$.

Suppose $B(X,Y) \neq 0$ and set $H = [X,Y]$. Then $0 \neq H \in \mathfrak{h}$. If $\langle \alpha, H \rangle = 0$ then H would commute with X and Y, and hence $\operatorname{ad} H$ would be nilpotent by Lemma 2.5.1, which is a contradiction. Hence $\langle \alpha, H \rangle \neq 0$ and we can rescale X and Y to obtain elements $e_\alpha \in \mathfrak{g}_\alpha$ and $f_\alpha \in \mathfrak{g}_{-\alpha}$ such that $\langle \alpha, h_\alpha \rangle = 2$, where $h_\alpha = [e_\alpha, f_\alpha]$.

(3): Let $\mathfrak{s}(\alpha) = \operatorname{Span}\{e_\alpha, f_\alpha, h_\alpha\} \cong \mathfrak{sl}(2,\mathbb{C})$ and set
$$M_\alpha = \mathbb{C}h_\alpha + \sum_{c \neq 0} \mathfrak{g}_{c\alpha} .$$
Since $[e_\alpha, \mathfrak{g}_{c\alpha}] \subset \mathfrak{g}_{(c+1)\alpha}$, $[f_\alpha, \mathfrak{g}_{c\alpha}] \subset \mathfrak{g}_{(c-1)\alpha}$, and $[e_\alpha, \mathfrak{g}_{-\alpha}] = [f_\alpha, \mathfrak{g}_\alpha] = \mathbb{C}\mathfrak{h}_\alpha$, we see that M_α is invariant under the adjoint action of $\mathfrak{s}(\alpha)$.

The eigenvalues of $\operatorname{ad} h_\alpha$ on M_α are $2c$ with multiplicity $\dim \mathfrak{g}_{c\alpha}$ and 0 with multiplicity one. By the complete reducibility of representations of $\mathfrak{sl}(2,\mathbb{C})$ (Theorem 2.3.6) and the classification of irreducible representations (Proposition 2.3.3) these eigenvalues must be integers. Hence $c\alpha \in \Phi$ implies that $2c$ is an integer. The eigenvalues in any irreducible representation are all even or all odd. Hence $c\alpha$ is not a root for any integer c with $|c| > 1$, since $\mathfrak{s}(\alpha)$ contains the zero eigenspace in M_α. This also proves that the only irreducible component of M_α with even eigenvalues is $\mathfrak{s}(\alpha)$, and it occurs with multiplicity one.

Suppose $(p + 1/2)\alpha \in \Phi$ for some positive integer p. Then $\operatorname{ad} h_\alpha$ would have eigenvalues $2p+1, 2p-1, \ldots, 3, 1$ on M_α, and hence $(1/2)\alpha$ would be a root. But then α could not be a root, by the argument just given, which is a contradiction. Thus we conclude that $M_\alpha = \mathbb{C}h_\alpha + \mathbb{C}e_\alpha + \mathbb{C}f_\alpha$. Hence $\dim \mathfrak{g}_\alpha = 1$.

(4): The notion of α root string through β from Section 2.4.2 carries over verbatim, as does Lemma 2.4.3. Hence the argument in Corollary 2.4.5 applies.

(5): This follows from the same argument as Corollary 2.4.4. □

2.5.3 Geometry of Root Systems

Let $\mathfrak{g}$ be a semisimple Lie algebra. Fix a Cartan subalgebra $\mathfrak{h}$ and let Φ be the set of roots of $\mathfrak{h}$ on $\mathfrak{g}$. For $\alpha \in \Phi$ there is a TDS triple $\{e_\alpha, f_\alpha, h_\alpha\}$ with $\langle \alpha, h_\alpha \rangle = 2$. Define

$$\check{\alpha} = n_\alpha \alpha\,, \quad \text{where } n_\alpha = B(e_\alpha, f_\alpha) \in \mathbb{Z} \setminus \{0\}\,. \tag{2.49}$$

Then $h_\alpha \longleftrightarrow \check{\alpha}$ under the isomorphism $\mathfrak{h} \cong \mathfrak{h}^*$ given by the Killing form B (see the proof of Theorem 2.5.20 (2)), and we shall call $\check{\alpha}$ the *coroot* to α.

By complete reducibility of representations of $\mathfrak{sl}(2,\mathbb{C})$ we know that $\mathfrak{g}$ decomposes into the direct sum of irreducible representations under the adjoint action of $\mathfrak{s}(\alpha) = \mathrm{Span}\{e_\alpha, f_\alpha, h_\alpha\}$. From Proposition 2.3.3 and Theorem 2.3.6 we see that e_α and f_α act by integer matrices relative to a suitable basis for any finite-dimensional representation of $\mathfrak{sl}(2,\mathbb{C})$. Hence the trace of $\mathrm{ad}(e_\alpha)\,\mathrm{ad}(f_\alpha)$ is an integer.

Since $\mathrm{Span}\,\Phi = \mathfrak{h}^*$ we can choose a basis $\{\alpha_1, \ldots, \alpha_l\}$ for $\mathfrak{h}^*$ consisting of roots. Setting $H_i = h_{\alpha_i}$, we see from (2.49) that $\{H_1, \ldots, H_l\}$ is a basis for $\mathfrak{h}$. Let

$$\mathfrak{h}_\mathbb{Q} = \mathrm{Span}_\mathbb{Q}\{H_1, \ldots, H_l\}\,, \quad \mathfrak{h}^*_\mathbb{Q} = \mathrm{Span}_\mathbb{Q}\{\alpha_1, \ldots, \alpha_l\}\,,$$

where $\mathbb{Q}$ denotes the field of rational numbers.

Lemma 2.5.21. *Each root $\alpha \in \Phi$ is in $\mathfrak{h}^*_\mathbb{Q}$, and the element h_α is in $\mathfrak{h}_\mathbb{Q}$. Let $a, b \in \mathfrak{h}_\mathbb{Q}$. Then $B(a,b) \in \mathbb{Q}$ and $B(a,a) > 0$ if $a \neq 0$.*

Proof. Set $a_{ij} = \langle \alpha_j, H_i \rangle$ and let $A = [a_{ij}]$ be the corresponding $l \times l$ matrix. The entries of A are integers by Theorem 2.5.20 (4), and the columns of A are linearly independent. Hence A is invertible. For $\alpha \in \Phi$ we can write $\alpha = \sum_i c_i \alpha_i$ for unique coefficients $c_i \in \mathbb{C}$. These coefficients satisfy the system of equations

$$\sum_j a_{ij} c_j = \langle \alpha, H_i \rangle \quad \text{for } i = 1, \ldots, l\,.$$

Since the right side of this system consists of integers, it follows that $c_j \in \mathbb{Q}$ and hence $\alpha \in \mathfrak{h}^*_\mathbb{Q}$. From (2.49) we then see that $h_\alpha \in \mathfrak{h}_\mathbb{Q}$ also.

Given $a, b \in \mathfrak{h}_\mathbb{Q}$, we can write $a = \sum_i c_i H_i$ and $b = \sum_j d_j H_j$ with $c_i, d_j \in \mathbb{Q}$. Thus

$$B(a,b) = \mathrm{tr}(\mathrm{ad}(a)\,\mathrm{ad}(b)) = \sum_{i,j} c_i\, d_j\, \mathrm{tr}(\mathrm{ad}(H_i)\,\mathrm{ad}(H_j))\,.$$

By Theorem 2.5.20 (3) we have

$$\operatorname{tr}(\operatorname{ad}(H_i)\operatorname{ad}(H_j)) = \sum_{\alpha\in\Phi} \langle\alpha, H_i\rangle\langle\alpha, H_j\rangle\ .$$

This is an integer by (2.48), so $B(a,b) \in \mathbb{Q}$. Likewise,

$$B(a,a) = \operatorname{tr}(\operatorname{ad}(a)^2) = \sum_{\alpha\in\Phi} \langle\alpha, a\rangle^2\ ,$$

and we have just proved that $\langle\alpha, a\rangle \in \mathbb{Q}$. If $a \neq 0$ then there exists $\alpha \in \Phi$ such that $\langle\alpha, a\rangle \neq 0$, because the center of $\mathfrak{g}$ is trivial. Hence $B(a,a) > 0$. □

Corollary 2.5.22. *Let $\mathfrak{h}_{\mathbb{R}}$ be the real span of $\{h_\alpha : \alpha \in \Phi\}$ and let $\mathfrak{h}^*_{\mathbb{R}}$ be the real span of the roots. Then the Killing form is real-valued and positive definite on $\mathfrak{h}_{\mathbb{R}}$. Furthermore, $\mathfrak{h}_{\mathbb{R}} \cong \mathfrak{h}^*_{\mathbb{R}}$ under the Killing-form duality.*

Proof. This follows immediately from (2.49) and Lemma 2.5.21. □

Let $E = \mathfrak{h}^*_{\mathbb{R}}$ with the bilinear form $(\cdot,\cdot)$ defined by the dual of the Killing form. By Corollary 2.5.22, E is an l-dimensional real Euclidean vector space. We have $\Phi \subset E$, and the coroots are related to the roots by

$$\check{\alpha} = \frac{2}{(\alpha,\alpha)}\alpha \quad \text{for } \alpha \in \Phi$$

by (2.49). Let $\check{\Phi} = \{\check{\alpha} : \alpha \in \Phi\}$ be the set of coroots. Then $(\beta, \check{\alpha}) \in \mathbb{Z}$ for all $\alpha, \beta \in \Phi$ by (2.48).

An element $h \in E$ is called *regular* if $(\alpha, h) \neq 0$ for all $\alpha \in \Phi$. Since the set

$$\bigcup_{\alpha\in\Phi} \{h \in E : (\alpha, h) = 0\}$$

is a finite union of hyperplanes, regular elements exist. Fix a regular element h_0 and define

$$\Phi^+ = \{\alpha \in \Phi : (\alpha, h_0) > 0\}\ .$$

Then $\Phi = \Phi^+ \cup (-\Phi^+)$. We call the elements of Φ^+ the *positive roots*. A positive root α is called *indecomposable* if there do not exist $\beta, \gamma \in \Phi^+$ such that $\alpha = \beta + \gamma$ (these definitions depend on the choice of h_0, of course).

Proposition 2.5.23. *Let Δ be the set of indecomposable positive roots.*

1. *Δ is a basis for the vector space E.*
2. *Every positive root is a linear combination of the elements of Δ with nonnegative integer coefficients.*
3. *If $\beta \in \Phi^+ \setminus \Delta$ then there exists $\alpha \in \Delta$ such that $\beta - \alpha \in \Phi^+$.*
4. *If $\alpha, \beta \in \Delta$ then the α root string through β is*

$$\beta, \beta+\alpha, \ldots, \beta+p\alpha, \quad \textit{where } p = -(\beta, \check{\alpha}) \geq 0\ . \tag{2.50}$$

Proof. The key to the proof is the following property of root systems:

($\star$) If $\alpha, \beta \in \Phi$ and $(\alpha, \beta) > 0$ then $\beta - \alpha \in \Phi$.

This property holds by Theorem 2.5.20 (4): $\beta - (\beta, \check{\alpha})\alpha \in \Phi$ and $(\beta, \check{\alpha}) \geq 1$, since $(\alpha, \beta) > 0$; hence $\beta - \alpha \in \Phi$. It follows from ($\star$) that

$$(\alpha, \beta) \leq 0 \quad \text{for all } \alpha, \beta \in \Delta \text{ with } \alpha \neq \beta \ . \tag{2.51}$$

Indeed, if $(\alpha, \beta) > 0$ then ($\star$) would imply that $\beta - \alpha \in \Phi$. If $\beta - \alpha \in \Phi^+$ then $\alpha = \beta + (\beta - \alpha)$, contradicting the indecomposability of α. Likewise, $\alpha - \beta \in \Phi^+$ would contradict the indecomposability of β. We now use these results to prove the assertions of the proposition.

(1): Any real linear relation among the elements of Δ can be written as

$$\sum_{\alpha \in \Delta_1} c_\alpha \alpha = \sum_{\beta \in \Delta_2} d_\beta \beta \ , \tag{2.52}$$

where Δ_1 and Δ_2 are disjoint subsets of Δ and the coefficients c_α and d_β are nonnegative. Denote the sum in (2.52) by γ. Then by (2.51) we have

$$0 \leq (\gamma, \gamma) = \sum_{\alpha \in \Delta_1} \sum_{\beta \in \Delta_2} c_\alpha d_\beta \, (\alpha, \beta) \leq 0 \ .$$

Hence $\gamma = 0$, and so we have

$$0 = (\gamma, h_0) = \sum_{\alpha \in \Delta_1} c_\alpha (\alpha, h_0) = \sum_{\beta \in \Delta_2} d_\beta (\beta, h_0) \ .$$

Since $(\alpha, h_0) > 0$ and $(\beta, h_0) > 0$, it follows that $c_\alpha = d_\beta = 0$ for all α, β.

(2): The set $M = \{(\alpha, h_0) : \alpha \in \Phi^+\}$ of positive real numbers is finite and totally ordered. If m_0 is the smallest number in M, then any $\alpha \in \Phi^+$ with $(\alpha, h_0) = m_0$ is indecomposable; hence $\alpha \in \Delta$. Given $\beta \in \Phi^+ \setminus \Delta$, then $m = (\beta, h_0) > m_0$ and $\beta = \gamma + \delta$ for some $\gamma, \delta \in \Phi^+$. Since $(\gamma, h_0) < m$ and $(\delta, h_0) < m$, we may assume by induction on m that γ and δ are nonnegative integral combinations of elements of Δ, and hence so is β.

(3): Let $\beta \in \Phi^+ \setminus \Delta$. There must exist $\alpha \in \Delta$ such that $(\alpha, \beta) > 0$, since otherwise the set $\Delta \cup \{\beta\}$ would be linearly independent by the argument at the beginning of the proof. This is impossible, since Δ is a basis for E by (1) and (2). Thus $\gamma = \beta - \alpha \in \Phi$ by ($\star$). Since $\beta \neq \alpha$, there is some $\delta \in \Delta$ that occurs with positive coefficient in γ. Hence $\gamma \in \Phi^+$.

(4): Since $\beta - \alpha$ cannot be a root, the α-string through β begins at β. Now apply Theorem 2.5.20 (4). □

We call the elements of Δ the *simple roots* (relative to the choice of Φ^+). Fix an enumeration $\alpha_1, \ldots, \alpha_l$ of Δ and write $E_i = e_{\alpha_i}$, $F_i = f_{\alpha_i}$, and $H_i = h_{\alpha_i}$ for the elements of the TDS triple associated with α_i. Define the *Cartan integers*

$C_{ij} = \langle \alpha_j, H_i \rangle$ and the $l \times l$ *Cartan matrix* $C = [C_{ij}]$ as in Section 2.4.3. Note that $C_{ii} = 2$ and $C_{ij} \le 0$ for $i \ne j$.

Theorem 2.5.24. *The simple root vectors* $\{E_1, \dots, E_l, F_1, \dots, F_l\}$ *generate* $\mathfrak{g}$. *They satisfy the relations* $[E_i, F_j] = 0$ *for* $i \ne j$ *and* $[H_i, H_j] = 0$, *where* $H_i = [E_i, F_i]$. *They also satisfy the following relations determined by the Cartan matrix:*

$$[H_i, E_j] = C_{ij} E_j \,, \quad [H_i, F_j] = -C_{ij} F_j \,; \tag{2.53}$$

$$\operatorname{ad}(E_i)^{-C_{ij}+1} E_j = 0 \quad \textit{for } i \ne j \,; \tag{2.54}$$

$$\operatorname{ad}(F_i)^{-C_{ij}+1} F_j = 0 \quad \textit{for } i \ne j \,. \tag{2.55}$$

Proof. Let $\mathfrak{g}'$ be the Lie subalgebra generated by the E_i and F_j. Since $\{H_1, \dots, H_l\}$ is a basis for $\mathfrak{h}$, we have $\mathfrak{h} \subset \mathfrak{g}'$. We show that $\mathfrak{g}_\beta \in \mathfrak{g}'$ for all $\beta \in \Phi^+$ by induction on the height of β, exactly as in the proof of Theorem 2.4.11. The same argument with β replaced by $-\beta$ shows that $\mathfrak{g}_{-\beta} \subset \mathfrak{g}'$. Hence $\mathfrak{g}' = \mathfrak{g}$.

The commutation relations in the theorem follow from the definition of the Cartan integers and Proposition 2.5.23 (4). □

The proof of Theorem 2.5.24 also gives the following generalization of Theorem 2.4.11:

Corollary 2.5.25. *Define* $\mathfrak{n}^+ = \sum_{\alpha \in \Phi^+} \mathfrak{g}_\alpha$ *and* $\mathfrak{n}^- = \sum_{\alpha \in \Phi^+} \mathfrak{g}_{-\alpha}$. *Then* $\mathfrak{n}^+$ *and* $\mathfrak{n}^-$ *are Lie subalgebras of* $\mathfrak{g}$ *that are invariant under* $\operatorname{ad} \mathfrak{h}$, *and* $\mathfrak{g} = \mathfrak{n}^- + \mathfrak{h} + \mathfrak{n}^+$. *Furthermore,* $\mathfrak{n}^+$ *is generated by* $\{E_1, \dots, E_l\}$ *and* $\mathfrak{n}^-$ *is generated by* $\{F_1, \dots, F_l\}$.

Remark 2.5.26. We define the *height* of a root (relative to the system of positive roots) just as for the Lie algebras of the classical groups: $\operatorname{ht}\big(\sum_i c_i \alpha_i\big) = \sum_i c_i$ (the coefficients c_i are integers all of the same sign). Then

$$\mathfrak{n}^- = \sum_{\operatorname{ht}(\alpha)<0} \mathfrak{g}_\alpha \quad \text{and} \quad \mathfrak{n}^+ = \sum_{\operatorname{ht}(\alpha)>0} \mathfrak{g}_\alpha \,.$$

Let $\mathfrak{b} = \mathfrak{h} + \mathfrak{n}^+$. Then $\mathfrak{b}$ is a maximal solvable subalgebra of $\mathfrak{g}$ that we call a *Borel subalgebra*.

We call the set Δ of simple roots *decomposable* if it can be partitioned into nonempty disjoint subsets $\Delta_1 \cup \Delta_2$, with $\Delta_1 \perp \Delta_2$ relative to the inner product on E. Otherwise, we call Δ *indecomposable*.

Theorem 2.5.27. *The semisimple Lie algebra* $\mathfrak{g}$ *is simple if and only if* Δ *is indecomposable.*

Proof. Assume that $\Delta = \Delta_1 \cup \Delta_2$ is decomposable. Let $\alpha \in \Delta_1$ and $\beta \in \Delta_2$. Then $p = 0$ in (2.50), since $(\alpha, \beta) = 0$. Hence $\beta + \alpha$ is not a root, and we already know that $\beta - \alpha$ is not a root. Thus

$$[\mathfrak{g}_{\pm\alpha}, \mathfrak{g}_{\pm\beta}] = 0 \quad \text{for all } \alpha \in \Delta_1 \text{ and } \beta \in \Delta_2 \,. \tag{2.56}$$

Let $\mathfrak{m}$ be the Lie subalgebra of $\mathfrak{g}$ generated by the root spaces $\mathfrak{g}_{\pm\alpha}$ with α ranging over Δ_1. It is clear from (2.56) and Theorem 2.5.24 that $\mathfrak{m}$ is a proper ideal in $\mathfrak{g}$. Hence $\mathfrak{g}$ is not simple.

Conversely, suppose $\mathfrak{g}$ is not simple. Then $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$, where each $\mathfrak{g}_i$ is a simple Lie algebra. The Cartan subalgebra $\mathfrak{h}$ must decompose as $\mathfrak{h} = \mathfrak{h}_1 \oplus \cdots \oplus \mathfrak{h}_r$, and by maximality of $\mathfrak{h}$ we see that $\mathfrak{h}_i$ is a Cartan subalgebra in $\mathfrak{g}_i$. It is clear from the definition of the Killing form that the roots of $\mathfrak{g}_i$ are orthogonal to the roots of $\mathfrak{g}_j$ for $i \neq j$. Since Δ is a basis for $\mathfrak{h}^*$, it must contain a basis for each $\mathfrak{h}_i^*$. Hence Δ is decomposable. □

2.5.4 Conjugacy of Cartan Subalgebras

Our results about the semisimple Lie algebra $\mathfrak{g}$ have been based on the choice of a particular Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$. We now show that this choice is irrelevant, generalizing Corollary 2.1.8.

If $X \in \mathfrak{g}$ is nilpotent, then $\operatorname{ad} X$ is a nilpotent derivation of $\mathfrak{g}$, and $\exp(\operatorname{ad} X)$ is a Lie algebra automorphism of $\mathfrak{g}$, called an *elementary automorphism*. It satisfies

$$\operatorname{ad}\big(\exp(\operatorname{ad} X)Y\big) = \exp(\operatorname{ad} X)\operatorname{ad} Y \exp(-\operatorname{ad} X) \quad \text{for } Y \in \mathfrak{g} \tag{2.57}$$

by Proposition 1.3.14. Let $\operatorname{Int}(\mathfrak{g})$ be the subgroup of $\operatorname{Aut}(\mathfrak{g})$ generated by the elementary automorphisms.

Theorem 2.5.28. *Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$ and let $\mathfrak{h}_1$ and $\mathfrak{h}_2$ be Cartan subalgebras of $\mathfrak{g}$. Then there exists an automorphism $\varphi \in \operatorname{Int}(\mathfrak{g})$ such that $\varphi(\mathfrak{h}_1) = \mathfrak{h}_2$.*

To prove this theorem, we need some preliminary results. Let $\mathfrak{g} = \mathfrak{n}^- + \mathfrak{h} + \mathfrak{n}^+$ be the triangular decomposition of $\mathfrak{g}$ from Corollary 2.5.25 and let $\mathfrak{b} = \mathfrak{h} + \mathfrak{n}^+$ be the corresponding Borel subalgebra. We shall call an element $H \in \mathfrak{h}$ *regular* if $\alpha(H) \neq 0$ for all roots α. From the root space decomposition of $\mathfrak{g}$ under $\operatorname{ad}\mathfrak{h}$, we see that this condition is the same as $\dim \operatorname{Ker}(\operatorname{ad} H) = \dim \mathfrak{h}$.

Lemma 2.5.29. *Suppose $Z \in \mathfrak{b}$ is semisimple. Write $Z = H + Y$, where $H \in \mathfrak{h}$ and $Y \in \mathfrak{n}^+$. Then $\dim \operatorname{Ker}(\operatorname{ad} Z) = \dim \operatorname{Ker}(\operatorname{ad} H) \geq \dim \mathfrak{h}$, with equality if and only if H is regular.*

Proof. Enumerate the positive roots in order of nondecreasing height as $\{\beta_1, \ldots, \beta_n\}$ and take an ordered basis for $\mathfrak{g}$ as

$$\{X_{-\beta_n}, \ldots, X_{-\beta_1}, H_1, \ldots, H_l, X_{\beta_1}, \ldots, X_{\beta_n}\} .$$

Here $X_\alpha \in \mathfrak{g}_\alpha$ and $\{H_1, \ldots, H_l\}$ is any basis for $\mathfrak{h}$. Then the matrix for $\operatorname{ad} Z$ relative to this basis is upper triangular and has the same diagonal as $\operatorname{ad} H$, namely

$$[-\beta_n(H), \dots, -\beta_1(H), \underbrace{0, \dots, 0}_{l}, \beta_1(H), \dots, \beta_n(H)] .$$

Since $\operatorname{ad} Z$ is semisimple, these diagonal entries are its eigenvalues, repeated according to multiplicity. Hence

$$\dim \operatorname{Ker}(\operatorname{ad} Z) = \dim \mathfrak{h} + 2 \operatorname{Card}\{\alpha \in \Phi^+ : \alpha(H) = 0\} .$$

This implies the statement of the lemma. □

Lemma 2.5.30. *Let $H \in \mathfrak{h}$ be regular. Define $f(X) = \exp(\operatorname{ad} X)H - H$ for $X \in \mathfrak{n}^+$. Then f is a polynomial map of $\mathfrak{n}^+$ onto $\mathfrak{n}^+$.*

Proof. Write the elements of $\mathfrak{n}^+$ as $X = \sum_{\alpha \in \Phi^+} X_\alpha$ with $X_\alpha \in \mathfrak{g}_\alpha$. Then

$$f(X) = \sum_{k \geq 1} \frac{1}{k!} (\operatorname{ad} X)^k H = -\sum_{\alpha \in \Phi^+} \alpha(H) X_\alpha + \sum_{k \geq 2} p_k(X) ,$$

where $p_k(X)$ is a homogeneous polynomial map of degree k on $\mathfrak{h}$. Note that $p_k(X) = 0$ for all sufficiently large k by the nilpotence of $\operatorname{ad} X$. From this formula it is clear that f maps a neighborhood of zero in $\mathfrak{n}^+$ bijectively onto some neighborhood U of zero in $\mathfrak{n}^+$.

To show that f is globally surjective, we introduce a one-parameter group of *grading automorphisms* of $\mathfrak{g}$ as follows: Set

$$\mathfrak{g}_0 = \mathfrak{h} , \quad \mathfrak{g}_n = \sum_{\operatorname{ht}(\beta) = n} \mathfrak{g}_\beta \quad \text{for } n \neq 0 .$$

This makes $\mathfrak{g}$ a *graded Lie algebra*: $[\mathfrak{g}_k, \mathfrak{g}_n] \subset \mathfrak{g}_{k+n}$ and $\mathfrak{g} = \bigoplus_{n \in \mathbb{Z}} \mathfrak{g}_n$. For $t \in \mathbb{C}^\times$ and $X_n \in \mathfrak{g}_n$ define

$$\delta_t\Big(\sum_n X_n\Big) = \sum_n t^n X_n .$$

The graded commutation relations imply that $\delta_t \in \operatorname{Aut}(\mathfrak{g})$. Thus $t \mapsto \delta_t$ is a regular homomorphism from $\mathbb{C}^\times$ to $\operatorname{Aut}(\mathfrak{g})$ (clearly $\delta_s \delta_t = \delta_{st}$). Since $\delta_t H = H$ for $H \in \mathfrak{h}$, we have $\delta_t f(X) = f(\delta_t X)$. Now let $Y \in \mathfrak{n}^+$. Since $\lim_{t \to 0} \delta_t Y = 0$, we can choose t sufficiently small that $\delta_t Y \in U$. Then there exists $X \in \mathfrak{n}^+$ such that $\delta_t Y = f(X)$, and hence $Y = \delta_{t^{-1}} f(X) = f(\delta_{t^{-1}} X)$. □

Corollary 2.5.31. *Suppose $Z \in \mathfrak{b}$ is semisimple and* $\dim \operatorname{Ker}(\operatorname{ad} Z) = \dim \mathfrak{h}$. *Then there exist $X \in \mathfrak{n}^+$ and a regular element $H \in \mathfrak{h}$ such that* $\exp(\operatorname{ad} X)H = Z$.

Proof. Write $Z = H + Y$ with $H \in \mathfrak{h}$ and $Y \in \mathfrak{n}^+$. By Lemma 2.5.29, H is regular, so by Lemma 2.5.30 there exists $X \in \mathfrak{n}^+$ with $\exp(\operatorname{ad} X)H = H + Y = Z$. □

We now come to the key result relating two Borel subalgebras.

Lemma 2.5.32. *Suppose $\mathfrak{b}_i = \mathfrak{h}_i + \mathfrak{n}_i$ are Borel subalgebras of $\mathfrak{g}$, for $i = 1, 2$. Then*

$$\mathfrak{b}_1 = \mathfrak{b}_1 \cap \mathfrak{b}_2 + \mathfrak{n}_1 . \tag{2.58}$$

Proof. The right side of (2.58) is contained in the left side, so it suffices to show that both sides have the same dimension. For any subspace $V \subset \mathfrak{g}$ let $V^\perp$ be the orthogonal of V relative to the Killing form on $\mathfrak{g}$. Then $\dim V^\perp = \dim \mathfrak{g} - \dim V$, since the Killing form is nondegenerate. It is easy to see from the root space decomposition that $\mathfrak{n}_i \subset (\mathfrak{b}_i)^\perp$. Since $\dim \mathfrak{n}_i = \dim \mathfrak{g} - \dim \mathfrak{b}_i$, it follows that $(\mathfrak{b}_i)^\perp = \mathfrak{n}_i$. Thus we have

$$(\mathfrak{b}_1 + \mathfrak{b}_2)^\perp = (\mathfrak{b}_1)^\perp \cap (\mathfrak{b}_2)^\perp = \mathfrak{n}_1 \cap \mathfrak{n}_2 \,. \tag{2.59}$$

But $\mathfrak{n}_2$ contains all the nilpotent elements of $\mathfrak{b}_2$, so $\mathfrak{n}_1 \cap \mathfrak{n}_2 = \mathfrak{n}_1 \cap \mathfrak{b}_2$. Thus (2.59) implies that

$$\dim(\mathfrak{b}_1 + \mathfrak{b}_2) = \dim \mathfrak{g} - \dim(\mathfrak{n}_1 \cap \mathfrak{b}_2) \,. \tag{2.60}$$

Set $d = \dim(\mathfrak{b}_1 \cap \mathfrak{b}_2 + \mathfrak{n}_1)$. Then by (2.60) we have

$$\begin{aligned} d &= \dim(\mathfrak{b}_1 \cap \mathfrak{b}_2) + \dim \mathfrak{n}_1 - \dim(\mathfrak{n}_1 \cap \mathfrak{b}_2) \\ &= \dim(\mathfrak{b}_1 \cap \mathfrak{b}_2) + \dim(\mathfrak{b}_1 + \mathfrak{b}_2) + \dim \mathfrak{n}_1 - \dim \mathfrak{g} \\ &= \dim \mathfrak{b}_1 + \dim \mathfrak{b}_2 + \dim \mathfrak{n}_1 - \dim \mathfrak{g} \,. \end{aligned}$$

Since $\dim \mathfrak{b}_1 + \dim \mathfrak{n}_1 = \dim \mathfrak{g}$, we have shown that $d = \dim \mathfrak{b}_2$. Clearly $d \le \dim \mathfrak{b}_1$, so this proves that $\dim \mathfrak{b}_2 \le \dim \mathfrak{b}_1$. Reversing the roles of $\mathfrak{b}_1$ and $\mathfrak{b}_2$, we conclude that $\dim \mathfrak{b}_1 = \dim \mathfrak{b}_2 = d$, and hence (2.58) holds. □

Proof of Theorem 2.5.28. We may assume that $\dim \mathfrak{h}_1 \le \dim \mathfrak{h}_2$. Choose systems of positive roots relative to $\mathfrak{h}_1$ and $\mathfrak{h}_2$ and let $\mathfrak{b}_i = \mathfrak{h}_i + \mathfrak{n}_i$ be the corresponding Borel subalgebras, for $i = 1, 2$. Let H_1 be a regular element in $\mathfrak{h}_1$. By Lemma 2.5.32 there exist $Z \in \mathfrak{b}_1 \cap \mathfrak{b}_2$ and $Y_1 \in \mathfrak{n}_1$ such that $H_1 = Z + Y_1$. Then by Lemma 2.5.30 there exists $X_1 \in \mathfrak{n}_1$ with $\exp(\operatorname{ad} X_1) H_1 = Z$. In particular, Z is a semisimple element of $\mathfrak{g}$ and by Lemma 2.5.29 we have

$$\dim \operatorname{Ker}(\operatorname{ad} Z) = \dim \operatorname{Ker}(\operatorname{ad} H_1) = \dim \mathfrak{h}_1 \,.$$

But $Z \in \mathfrak{b}_2$, so Lemma 2.5.29 gives $\dim \operatorname{Ker}(\operatorname{ad} Z) \ge \dim \mathfrak{h}_2$. This proves that $\dim \mathfrak{h}_1 = \dim \mathfrak{h}_2$. Now apply Corollary 2.5.31: there exists $X_2 \in \mathfrak{n}_2$ such that

$$\exp(\operatorname{ad} X_2) Z = H_2 \in \mathfrak{h}_2 \,.$$

Since $\dim \operatorname{Ker}(\operatorname{ad} H_2) = \dim \operatorname{Ker}(\operatorname{ad} Z) = \dim \mathfrak{h}_2$, we see that H_2 is regular. Hence $\mathfrak{h}_2 = \operatorname{Ker}(\operatorname{ad} H_2)$. Thus the automorphism $\varphi = \exp(\operatorname{ad} X_2) \exp(\operatorname{ad} X_1) \in \operatorname{Int} \mathfrak{g}$ maps $\mathfrak{h}_1$ onto $\mathfrak{h}_2$. □

Remark 2.5.33. Let $Z \in \mathfrak{g}$ be a semisimple element. We say that Z is *regular* if $\dim \operatorname{Ker}(\operatorname{ad} Z)$ has the smallest possible value among all elements of $\mathfrak{g}$. From Theorem 2.5.28 we see that this minimal dimension is the *rank* of $\mathfrak{g}$. Furthermore, if Z is regular then $\operatorname{Ker}(\operatorname{ad} Z)$ is a Cartan subalgebra of $\mathfrak{g}$ and all Cartan subalgebras are obtained this way.

2.5.5 Exercises

1. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra and let B be the Killing form of $\mathfrak{g}$. Show that $B([X,Y],Z) = B(X,[Y,Z])$ for all $X,Y,Z \in \mathfrak{g}$.
2. Let $\mathfrak{g} = \mathbb{C}X + \mathbb{C}Y$ be the two-dimensional Lie algebra with commutation relations $[X,Y] = Y$. Calculate the Killing form of $\mathfrak{g}$.
3. Suppose $\mathfrak{g}$ is a simple Lie algebra and $\omega(X,Y)$ is an invariant symmetric bilinear form on $\mathfrak{g}$. Show that ω is a multiple of the Killing form B of $\mathfrak{g}$. (HINT: Use the nondegeneracy of B to write $\omega(X,Y) = B(TX,Y)$ for some $T \in \mathrm{End}(\mathfrak{g})$. Then show that the eigenspaces of T are invariant under $\mathrm{ad}\,\mathfrak{g}$.)
4. Let $\mathfrak{g} = \mathfrak{sl}(n,\mathbb{C})$. Show that the Killing form B of $\mathfrak{g}$ is $2n\,\mathrm{tr}_{\mathbb{C}^n}(XY)$. (HINT: Calculate $B(H,H)$ for $H = \mathrm{diag}[1,-1,0,\ldots,0]$ and then use the previous exercise.)
5. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra and let $\mathfrak{h} \subset \mathfrak{g}$ be an ideal. Prove that the Killing form of $\mathfrak{h}$ is the restriction to $\mathfrak{h}$ of the Killing form of $\mathfrak{g}$.
6. Prove formula (2.45).
7. Let D be a derivation of a finite-dimensional Lie algebra $\mathfrak{g}$. Prove that $\exp(tD)$ is an automorphism of $\mathfrak{g}$ for all scalars t. (HINT: Let $X,Y \in \mathfrak{g}$ and consider the curves $\varphi(t) = \exp(tD)[X,Y]$ and $\psi(t) = [\exp(tD)X, \exp(tD)Y]$ in $\mathfrak{g}$. Show that $\varphi(t)$ and $\psi(t)$ satisfy the same differential equation and $\varphi(0) = \psi(0)$.)

2.6 Notes

Section 2.1.2. The proof of the conjugacy of maximal tori for the classical groups given here takes advantage of a special property of the defining representation of a classical group, namely that it is *multiplicity-free* for the maximal torus. In Chapter 11 we will prove the conjugacy of maximal tori in any connected linear algebraic group using the general structural results developed there.

Section 2.2.2. A linear algebraic group $G \subset \mathbf{GL}(n,\mathbb{C})$ is connected if and only if the defining ideal for G in $\mathbb{C}[G]$ is *prime*. Weyl [164, Chapter X, Supplement B] gives a direct argument for this in the case of the symplectic and orthogonal groups.

Sections 2.4.1 and 2.5.2. The *roots* of a semisimple Lie algebra were introduced by Killing as the roots of the characteristic polynomial $\det(\mathrm{ad}(x) - \lambda I)$, for $x \in \mathfrak{g}$ (by the Jordan decomposition, one may assume that x is semisimple and hence that $x \in \mathfrak{h}$). See the Note Historique in Bourbaki [12] and Hawkins [63] for details.

Section 2.3.3. See Borel [17, Chapter II] for the history of the proof of complete reducibility for representations of $\mathbf{SL}(2,\mathbb{C})$. The proof given here is based on arguments first used by Cartan [26].

Sections 2.4.3 and 2.5.3. Using the set of roots to study the structure of $\mathfrak{g}$ is a fundamental technique going back to Killing and Cartan. The most thorough axiomatic treatment of root systems is in Bourbaki [12]; for recent developments see Humphreys [78] and Kane [83]. The notion of a set of simple roots and the associated Dynkin diagram was introduced in Dynkin [44], which gives a self-contained development of the structure of semisimple Lie algebras.

Section 2.5.1. In this section we follow the exposition in Hochschild [68]. The proof of Theorem 2.5.3 is from Hochschild [68, Theorem XI.1.2], and the proof of Theorem 2.5.7 is from Hochschild [68, Theorem XI.1.6].