Lie Algebras

Brooks Roberts

University of Idaho

Course notes from 2018–2019

Contents

Chapter 1

Basic concepts

1.1 References

The main reference for this course is the book *Introduction to Lie Algebras*, by Karin Erdmann and Mark J. Wildon; this is reference [4]. Another important reference is the book [6], *Introduction to Lie Algebras and Representation Theory*, by James E. Humphreys. The best references for Lie theory are the three volumes [1], *Lie Groups and Lie Algebras, Chapters 1-3*, [2], *Lie Groups and Lie Algebras, Chapters 4-6*, and [3], *Lie Groups and Lie Algebras, Chapters 7-9*, all by Nicolas Bourbaki.

1.2 Motivation

Briefly, Lie algebras have to do with the algebra of derivatives in settings where there is a lot of symmetry. As a consequence, Lie algebras appear in various parts of advanced mathematics. The nexus of these applications is the theory of symmetric spaces. Symmetric spaces are rich objects whose theory has components from geometry, analysis, algebra, and number theory. With these short remarks in mind, in this course we will begin without any more motivation, and start with the definition of a Lie algebra. For now, rather than be concerned about advanced applications, the student should instead exercise critical thinking as basic concepts are introduced.

1.3 The definition

Lie algebras are defined as follows. Throughout this chapter F be an arbitrary field. A **Lie algebra over** F is an F-vector space L and an F-bilinear map

$$[\cdot,\cdot] : L \times L \longrightarrow L$$

that has the following two properties:

1. $[x, x] = 0$ for all $x \in L$;

2. $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in L$.

The map $[\cdot, \cdot]$ is called the **Lie bracket** of L. The second property is called the **Jacobi identity**.

Proposition 1.3.1. *Let L be a Lie algebra over F. If $x, y \in L$, then $[x, y] = -[y, x]$.*

Proof. Let $x, y \in L$. Then

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

so that $[x, y] = -[y, x]$. □

If L_1 and L_2 are Lie algebras over F, then a **homomorphism** $T : L_1 \to L_2$ is an F-linear map that satisfies $T([x, y]) = [T(x), T(y)]$ for all $x, y \in L_1$. If L is a Lie algebra over F, then a **subalgebra** of L is an F-vector subspace K of L such that $[x, y] \in K$ for all $x, y \in K$; evidently, a subalgebra is a Lie algebra over F using the same Lie bracket. If L is a Lie algebra over F, then an **ideal** I of L is an F-vector subspace of L such that $[x, y] \in I$ for all $x \in L$ and $y \in I$; evidently, an ideal of L is also a subalgebra of A. Also, because of Proposition 1.3.1, it is not necessary to introduce the concepts of left or right ideals. If L is a Lie algebra over F, then the **center** of L is defined to be

$$Z(L) = \{x \in L : [x, y] = 0 \text{ for all } y \in L\}.$$

Clearly, the center of L is an F-subspace of L.

Proposition 1.3.2. *Let L be a Lie algebra over F. The center $Z(L)$ of L is an ideal of L.*

Proof. Let $y \in L$ and $x \in Z(L)$. If $z \in L$, then $[[y, x], z] = -[[x, y], z] = 0$. This implies that $[y, x] \in Z(L)$. □

If L is a Lie algebra over F, then we say that L is **abelian** if $Z(L) = L$, i.e., if $[x, y] = 0$ for all $x, y \in L$.

Proposition 1.3.3. *Let L_1 and L_2 be Lie algebras over F, and let $T : L_1 \to L_2$ be a homomorphism. The kernel of T is an ideal of L_1.*

Proof. Let $y \in \ker(T)$ and $x \in L_1$. Then $T([x, y]) = [T(x), T(y)] = [T(x), 0] = 0$, so that $[x, y] \in \ker(T)$. □

1.4 Some important examples

Proposition 1.4.1. *Let A be an associative F-algebra. For $x, y \in A$ define*

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

so that $[x, y]$ is just the commutator of x and y. With this definition of a Lie bracket, the F-vector space A is a Lie algebra.

Proof. It is easy to verify that $[\cdot, \cdot]$ is F-bilinear and that property 1 of the definition of a Lie algebra is satisfied. We need to prove that the Jacobi identity is satisfied. Let $x, y, z \in A$. Then

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

This completes the proof. ☐

Note that in the last proof we indeed used that the algebra was associative.

If V is an F-vector space, then the F-vector space $\mathrm{gl}(V)$ of all F-linear operators from V to V is an associative algebra over F under composition, and thus defines a corresponding Lie algebra over F, also denoted by $\mathrm{gl}(V)$, with Lie bracket as defined in Proposition 1.4.1. Similarly, if n is a non-negative integer, then F-vector space $\mathrm{gl}(n, F)$ of all $n \times n$ matrices is an associative algebra under multiplication of matrices, and thus defines a corresponding Lie algebra, also denoted by $\mathrm{gl}(n, F)$.

The example $\mathrm{gl}(n, F)$ shows that in general the Lie bracket is not associative, i.e., it is not in general true that $[x, [y, z]] = [[x, y], z]$ for all $x, y, z \in \mathrm{gl}(n, F)$. For example, if $n = 2$, and

$$x = \begin{bmatrix} & \\ & 1 \end{bmatrix}, \quad y = \begin{bmatrix} & 1 \\ & \end{bmatrix}, \quad z = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}$$

then

$$[x, [y, z]] = x(yz - zy) = xyz - xzy = xyz - \begin{bmatrix} & \\ & 1 \end{bmatrix} \begin{bmatrix} & 1 \\ & \end{bmatrix} = xyz$$

and

$$[[x, y], z] = (xy - yx)z = xyz - yxz = xyz - \begin{bmatrix} & 1 \\ & \end{bmatrix} \begin{bmatrix} & \\ & 1 \end{bmatrix} = xyz - \begin{bmatrix} & 1 \\ & \end{bmatrix}.$$

We describe some more important examples of Lie algebras.

Proposition 1.4.2. *Let n be a non-negative integer, and let $\mathrm{sl}(n, F)$ be the subspace of $\mathrm{gl}(n, F)$ consisting of elements x such that $\mathrm{tr}(x) = 0$. Then $\mathrm{sl}(n, F)$ is a Lie subalgebra of $\mathrm{gl}(n, F)$.*

Proof. It will suffice to prove that $\mathrm{tr}([x, y]) = 0$ for $x, y \in \mathrm{sl}(n, F)$. Let $x, y \in \mathrm{sl}(n, F)$. Then $\mathrm{tr}([x, y]) = \mathrm{tr}(xy - yx) = \mathrm{tr}(xy) - \mathrm{tr}(yx) = \mathrm{tr}(xy) - \mathrm{tr}(xy) = 0$. □

The example $\mathrm{sl}(2, F)$ is especially important. We have

$$\mathrm{sl}(2, F) = \{\begin{bmatrix} a & b \\ c & -a \end{bmatrix} : a, b, c \in F\}.$$

An important basis for $\mathrm{sl}(2, F)$ is

$$e = \begin{bmatrix} & 1 \\ & \end{bmatrix}, \qquad f = \begin{bmatrix} & \\ 1 & \end{bmatrix}, \qquad h = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}.$$

We have:

$$[e, f] = h, \qquad [e, h] = -2e, \qquad [f, h] = 2f.$$

Proposition 1.4.3. *Let n be a non-negative integer, and let $S \in \mathrm{gl}(n, F)$. Let*

$$\mathrm{gl}_S(n, F) = \{x \in \mathrm{gl}(n, F) : {}^t x S = -Sx\}.$$

Then $\mathrm{gl}_S(n, F)$ is a Lie subalgebra of $\mathrm{gl}(n, F)$.

Proof. Let $x, y \in \mathrm{gl}_S(n, F)$. We need to prove $[x, y] \in \mathrm{gl}_S(n, F)$. We have

$$\begin{aligned}
{}^t([x, y])S &= {}^t(xy - yx)S \\
&= ({}^t y {}^t x - {}^t x {}^t y)S \\
&= {}^t y {}^t x S - {}^t x {}^t y S \\
&= -{}^t y S x + {}^t x S y \\
&= Syx - Sxy \\
&= S[y, x] \\
&= -S[x, y].
\end{aligned}$$

This completes the proof. □

If $n = 2\ell$ is even, and

$$S = \begin{bmatrix} & 1_\ell \\ 1_\ell & \end{bmatrix},$$

then we write

$$\mathrm{so}(n, F) = \mathrm{so}(2\ell, F) = \mathrm{gl}_S(n, F).$$

If $n = 2\ell + 1$ is odd, and

$$S = \begin{bmatrix} 1 & & \\ & & 1_\ell \\ & 1_\ell & \end{bmatrix},$$

then we write

$$\mathrm{so}(n, F) = \mathrm{so}(2\ell + 1, F) = \mathrm{gl}_S(n, F).$$

Also, if $n = 2\ell$ is even and

$$S = \begin{bmatrix} & 1_\ell \\ -1_\ell & \end{bmatrix},$$

then we write

$$\mathrm{sp}(n, F) = \mathrm{sp}(2\ell, F) = \mathrm{gl}_S(n, F).$$

If the F-vector space V is actually an algebra R over F, then the Lie algebra $\mathrm{gl}(R)$ admits a natural subalgebra. Note that in the next proposition we *do not* assume that R is associative.

Proposition 1.4.4. *Let R be an F-algebra. Let $\mathrm{Der}(R)$ be the subspace of $\mathrm{gl}(R)$ consisting of* **derivations***, i.e., $D \in \mathrm{gl}(R)$ such that*

$$D(ab) = aD(b) + D(b)a$$

for all $a, b \in R$. Then $\mathrm{Der}(R)$ is a Lie subalgebra of $\mathrm{gl}(R)$.

Proof. Let $D_1, D_2 \in \mathrm{Der}(R)$ and $a, b \in R$. Then

$$\begin{aligned}
[D_1, D_2](ab) &= (D_1 \circ D_2 - D_2 \circ D_1)(ab) \\
&= (D_1 \circ D_2)(ab) - (D_2 \circ D_1)(ab) \\
&= D_1(D_2(ab)) - D_2(D_1(ab)) \\
&= D_1(aD_2(b) + D_2(a)b) - D_2(aD_1(b) + D_1(a)b) \\
&= aD_1(D_2(b)) + D_1(a)D_2(b) + D_2(a)D_1(b) + D_1(D_2(a))b \\
&\qquad - aD_2(D_1(b)) - D_2(a)D_1(b) - D_1(a)D_2(b) - D_2(D_1(a))b \\
&= a([D_1, D_2](b)) + ([D_1, D_2](a))b.
\end{aligned}$$

This proves that $[D_1, D_2]$ is in $\mathrm{Der}(R)$. □

1.5 The adjoint homomorphism

The proof of the next proposition uses the Jacobi identity.

Proposition 1.5.1. *Let L be a Lie algebra over F. Define*

$$\mathrm{ad} : L \longrightarrow \mathrm{gl}(L)$$

by

$$\big(\mathrm{ad}(x)\big)(y) = [x, y]$$

for $x, y \in L$. Then ad is a Lie algebra homomorphism. Moreover, the kernel of ad is $Z(L)$, and the image of ad lies in $\mathrm{Der}(L)$. We refer to ad as the **adjoint homomorphism**.

Proof. Let $x_1, x_2, y \in L$. Then

$$\big(\mathrm{ad}([x_1,x_2])\big)(y) = [[x_1,x_2],y].$$

Also,

$$\begin{aligned}\big([\mathrm{ad}(x_1),\mathrm{ad}(x_2)]\big)(y) &= \big(\mathrm{ad}(x_1)\circ\mathrm{ad}(x_2)\big)(y) - \big(\mathrm{ad}(x_2)\circ\mathrm{ad}(x_1)\big)(y)\\ &= \mathrm{ad}(x_1)\big([x_2,y]\big) - \mathrm{ad}(x_2)\big([x_1,y]\big)\\ &= [x_1,[x_2,y]] - [x_2,[x_1,y]].\end{aligned}$$

It follows that

$$\begin{aligned}&\big(\mathrm{ad}([x_1,x_2])\big)(y) - \big([\mathrm{ad}(x_1),\mathrm{ad}(x_2)]\big)(y)\\ &= [[x_1,x_2],y] - [x_1,[x_2,y]] + [x_2,[x_1,y]]\\ &= -[y,[x_1,x_2]] - [x_1,[x_2,y]] - [x_2,[y,x_1]]\\ &= 0\end{aligned}$$

by the Jacobi identity. This proves that ad is a Lie algebra homomorphism. It is clear that the kernel of the adjoint homomorphism is $Z(L)$. We also have

$$\mathrm{ad}(x)([y_1,y_2]) = [x,[y_1,y_2]]$$

and

$$[y_1,\mathrm{ad}(x)(y_2)] + [\mathrm{ad}(x)(y_1),y_2] = [y_1,[x,y_2]] + [[x,y_1],y_2].$$

Therefore,

$$\begin{aligned}&\mathrm{ad}(x)([y_1,y_2]) - [y_1,\mathrm{ad}(x)(y_2)] - [\mathrm{ad}(x)(y_1),y_2]\\ &= [x,[y_1,y_2]] - [y_1,[x,y_2]] - [[x,y_1],y_2]\\ &= [x,[y_1,y_2]] + [y_1,[y_2,x]] + [y_2,[x,y_1]]\\ &= 0,\end{aligned}$$

again by the Jacobi identity. This proves that the image of ad lies in $\mathrm{Der}(L)$. $\square$

The previous proposition shows that elements of a Lie algebra can always be thought of as derivations of an algebra. It turns out that if L is a finite-dimensional semi-simple Lie algebra over the complex numbers $\mathbb{C}$, then the image of the adjoint homomorphism is $\mathrm{Der}(L)$.

Chapter 2

Solvable and nilpotent Lie algebras

In this chapter F is an arbitrary field.

2.1 Solvability

Proposition 2.1.1. *Let L be a Lie algebra over F, and let I and J be ideals of L. Define $[I, J]$ to be the F-linear span of all the brackets $[x, y]$ for $x \in I$ and $y \in J$. The F-vector subspace $[I, J]$ of L is an ideal of L.*

Proof. Let $x \in L$, $y \in I$ and $z \in J$. We need to prove that $[x, [y, z]] \in [I, J]$. We have

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

by the Jacobi identity. We have $[z, x] \in J$ because J is an ideal, and $[x, y] \in I$ because I is an ideal. It follows that $[x, [y, z]] \in [I, J]$. Note that we also use Proposition 1.3.1. □

By Proposition 1.3.1, if L is a Lie algebra over F, and I and J are ideals of L, then $[I, J] = [J, I]$.

If L is a Lie algebra over F, then the **derived algebra** of L is defined to be $L' = [L, L]$.

Proposition 2.1.2. *The derived algebra of $\mathrm{sl}(2, F)$ is $\mathrm{sl}(2, F)$.*

Proof. This follows immediately from $[e, f] = h, [e, h] = -2e, [f, h] = 2f$. □

Proposition 2.1.3. *Let L be a Lie algebra over F. The quotient algebra L/L' is abelian.*

Proof. This follows immediately from the definition of the derived algebra. □

Let L be a Lie algebra over F. We can consider the following descending sequence of ideals:

$$L \supset L' = [L, L] \supset (L')' = [L', L'] \supset ((L')')' = [(L')', (L')'] \cdots$$

Each term of the sequence is actually an ideal of L; also, the successive quotients are abelian. To improve the notation, we will write

$$\begin{aligned} L^{(0)} &= L, \\ L^{(1)} &= L', \\ L^{(2)} &= (L')', \\ &\cdots \\ L^{(k+1)} &= (L^k)' \\ &\cdots. \end{aligned}$$

We have then

$$L = L^{(0)} \supset L^{(1)} \supset L^{(2)} \supset \cdots.$$

This is called the **derived series** of L. We say that L is **solvable** if $L^{(k)} = 0$ for some non-negative integer k.

Proposition 2.1.4. *Let L be a Lie algebra over F. Then L is solvable if and only if there exists a sequence $I_0, I_1, I_2, \ldots, I_m$ of ideals of L such that*

$$L = I_0 \supset I_1 \supset I_2 \supset \cdots \supset I_{m-1} \supset I_m = 0$$

and I_{k-1}/I_k is abelian for $k \in \{1, \ldots, m\}$.

Proof. Assume that a sequence exists as in the statement of the proposition; we need to prove that L is solvable. To prove this it will suffice to prove that $L^{(k)} \subset I_k$ for $k \in \{0, 1, \ldots, m\}$. We will prove this by induction on k. The induction claim is true if $k = 0$ because $L^{(0)} = L = I_0$. Assume that $k \in \{1, \ldots, m\}$ and that $L^{(j)} \subset I_j$ for all $j \in \{0, 1, \ldots, k-1\}$; we will prove that $L^{(k)} \subset I_k$. By hypothesis, I_{k-1}/I_k is abelian. This implies that $[I_{k-1}, I_{k-1}] \subset I_k$. We have:

$$L^{(k)} = [L^{(k-1)}, L^{(k-1)}] \subset [I_{k-1}, I_{k-1}] \subset I_k.$$

This completes the argument. □

Lemma 2.1.5. *Let L_1 and L_2 be Lie algebras over F. Let $T : L_1 \to L_2$ be a surjective Lie algebra homomorphism. If k is a non-negative integer, then $T(L_1^{(k)}) = L_2^{(k)}$. Consequently, if L_1 is solvable, then so is $L_2 = T(L_1)$.*

Proof. We will prove that $T(L_1^{(k)}) = L_2^{(k)}$ by induction on k. This is clear if $k = 0$. Assume that the statement holds for k; we will prove that it holds for $k + 1$. Now

$$T(L_1^{(k+1)}) = T([L_1^{(k)}, L_1^{(k)}])$$

$$\begin{aligned} &= [T(L_1^{(k)}), T(L_1^{(k)})] \\ &= [L_2^{(k)}, L_2^{(k)}] \\ &= L_2^{(k+1)}. \end{aligned}$$

This completes the proof. □

Lemma 2.1.6. *Let L be a Lie algebra over F. We have $L^{(k+j)} = (L^{(k)})^{(j)}$ for all non-negative integers k and j.*

Proof. Fix a non-negative integer k. We will prove that $L^{(k+j)} = (L^{(k)})^{(j)}$ by induction on j. If $j = 0$, then $L^{(k+j)} = L^{(k)} = (L^{(k)})^{(0)} = (L^{(k)})^{(j)}$. Assume that the statement holds for j; we will prove that it holds for $j + 1$. By the induction hypothesis,

$$\begin{aligned} L^{(k+j+1)} &= [L^{(k+j)}, L^{(k+j)}] \\ &= [(L^{(k)})^{(j)}, (L^{(k)})^{(j)}]. \end{aligned}$$

Also,

$$(L^{(k)})^{(j+1)} = [(L^{(k)})^{(j)}, (L^{(k)})^{(j)}].$$

The lemma is proven. □

Lemma 2.1.7. *Let L be a Lie algebra over F. Let I be an ideal of L. The Lie algebra L is solvable if and only if I and L/I are solvable.*

Proof. If L is solvable then I is solvable because $I^{(k)} \subset L^{(k)}$ for all non-negative integers; also, L/I is solvable by Lemma 2.1.5. Assume that I and L/I are solvable. Since L/I is solvable, there exists a non-negative integer k such that $(L/I)^{(k)} = 0$. This implies that $L^{(k)} + I = I$, so that $L^{(k)} \subset I$. Since I is solvable, there exists an non-negative integer j such that $I^{(j)} = 0$. It follows that $(L^{(k)})^{(j)} \subset I^{(j)} = 0$. Since $L^{(k+j)} = (L^{(k)})^{(j)}$ by Lemma 2.1.6, we conclude that L is solvable. □

Lemma 2.1.8. *Let L be a Lie algebra over F, and let I and J be solvable ideals of L. Then $I + J$ is solvable.*

Proof. We consider the sequence

$$I + J \supset J \supset 0.$$

We have $(I + J)/J \cong I/(I \cap J)$ as Lie algebras. Since I is solvable, these isomorphic Lie algebras are solvable by Lemma 2.1.5. The Lie algebra $I + J$ is now solvable by Lemma 2.1.7. □

Proposition 2.1.9. *Let L be a finite-dimensional Lie algebra over F. Then there exists a solvable ideal I of L such that every solvable ideal of L is contained in I.*

Proof. Since L is finite-dimensional, there exists a solvable ideal I of L of maximal dimension. Let J be a solvable ideal of L. The ideal $I+J$ is solvable by Lemma 2.1.8. Since I has maximum dimension we must have $I+J=I$, so that $J \subset I$. □

If L is a finite-dimensional Lie algebra over F, then the ideal from Proposition 2.1.9 is clearly unique; we refer to it as the **radical** of L, and denote it by $\mathrm{rad}(L)$. We say that finite-dimensional Lie algebra L over F is **semi-simple** if $L \neq 0$ and the radical of L is zero, i.e., $\mathrm{rad}(L) = 0$. Because the center $Z(L)$ of a Lie algebra L is abelian, the center $Z(L)$ is a solvable ideal of L. Hence, $\mathrm{rad}(L)$ contains $Z(L)$. If L is a semi-simple Lie algebra, then $Z(L) = 0$.

Proposition 2.1.10. *Let L be a finite-dimensional Lie algebra over F. The Lie algebra $L/\mathrm{rad}(L)$ is semi-simple.*

Proof. Let I be a solvable ideal in $L/\mathrm{rad}(L)$; we need to prove that $I = 0$. Let $p : L \to L/\mathrm{rad}(L)$ be the projection map; this is a Lie algebra homomorphism. Define $J = p^{-1}(I)$. Evidently, J is an ideal of L containing $\mathrm{rad}(L)$. Let k be a non-negative integer. By Lemma 2.1.5 we have $p(J^{(k)}) = p(J)^{(k)} = I^{(k)}$. There exists a positive integer k such that $I^{(k)} = 0$. It follows that $p(J^{(k)}) = 0$. This implies that $J^{(k)} \subset \mathrm{rad}(L)$. Since $\mathrm{rad}(L)$ is solvable, it follows for some positive integer j we have $(J^{(k)})^j = 0$. Consequently, by Lemma 2.1.6, the ideal J is solvable. This implies that $J \subset \mathrm{rad}(L)$, which in turn implies that $I = 0$. □

The following theorem will not be proven now, but is an important reduction in the structure of Lie algebras.

Theorem 2.1.11 (Levi decomposition)**.** *Assume that the characteristic of F is zero. Let L be a finite dimensional Lie algebra over F. Then there exists a subalgebra S of L such that $L = \mathrm{rad}(L) \oplus S$ as vector spaces.*

Proposition 2.1.12. *Assume that the characteristic of F is not two. The Lie algebra $\mathrm{sl}(2,F)$ is semi-simple. In fact, $\mathrm{sl}(2,F)$ has no ideals except 0 and $\mathrm{sl}(2,F)$.*

Proof. Let I be an ideal of $\mathrm{sl}(2,F)$. Let $x = ae + bh + cf$ be an element of I, with $a, b, c \in F$. Assume that $a \neq 0$. We have

$$[h, x] = 2ae - 2cf,$$
$$[f, x] = -ah + 2bf,$$

so that

$$[f, [h, x]] = -2ah,$$
$$[f, [f, x]] = -2af.$$

It follows that h and f are contained in I. This implies that e is contained in I, so that $I = \mathrm{sl}(2,F)$. The argument is similar if $b \neq 0$ or $c \neq 0$. □

We say that a Lie algebra L over F is **reductive** if $\mathrm{rad}(L) = Z(L)$.

Proposition 2.1.13. *Assume that the characteristic of F is not two. The Lie algebra $\mathrm{gl}(2, F)$ is reductive.*

Proof. Since $\mathrm{tr}([x, y]) = 0$ for any $x, y \in \mathrm{gl}(2, F)$, it follows that $\mathrm{sl}(2, F)$ is an ideal of $\mathrm{gl}(2, F)$. Let $I = \mathrm{rad}(\mathrm{gl}(2, F))$. Then $I \cap \mathrm{sl}(2, F)$ is an ideal of $\mathrm{gl}(2, F)$ and an ideal of $\mathrm{sl}(2, F)$. By Proposition 2.1.12, we must have $I \cap \mathrm{sl}(2, F) = \mathrm{sl}(2, F)$ or $I \cap \mathrm{sl}(2, F) = 0$. Assume that $I \cap \mathrm{sl}(2, F) = \mathrm{sl}(2, F)$, so that $\mathrm{sl}(2, F) \subset I$. By Lemma 2.1.7, $\mathrm{sl}(2, F)$ is solvable. This contradicts the fact that $\mathrm{sl}(2, F)$ is semi-simple by Proposition 2.1.12. We thus have $I \cap \mathrm{sl}(2, F) = 0$. Let

$$x = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be in I. We have

$$[e, x] = \begin{bmatrix} c & d - a \\ & -c \end{bmatrix} \in I \cap \mathrm{sl}(2, F) = 0,$$

$$[f, x] = \begin{bmatrix} -b & \\ a - d & b \end{bmatrix} \in I \cap \mathrm{sl}(2, F) = 0.$$

It follows that

$$x \in Z(\mathrm{gl}(2, F)) = \{\begin{bmatrix} a & \\ & a \end{bmatrix} : a \in F\},$$

so that $I \subset Z(\mathrm{gl}(2, F))$. Since $Z(\mathrm{gl}(2, F)) \subset I = \mathrm{rad}(\mathrm{gl}(2, F))$, the proposition is proven. □

Proposition 2.1.14. *Let $\mathrm{b}(2, F)$ be the F-subspace of $\mathrm{gl}(2, F)$ consisting of upper triangular matrices. Then $\mathrm{b}(2, F)$ is a Lie subalgebra of $\mathrm{gl}(2, F)$, and $\mathrm{b}(2, F)$ is solvable.*

Proof. Let

$$x_1 = \begin{bmatrix} a_1 & b_1 \\ & d_1 \end{bmatrix}, \qquad x_2 = \begin{bmatrix} a_2 & b_2 \\ & d_2 \end{bmatrix}$$

be in $\mathrm{b}(2, F)$. Then

$$[x_1, x_2] = \begin{bmatrix} & b_1 d_2 - b_2 d_1 + a_1 b_2 - a_2 b_1 \\ & \end{bmatrix} \in \begin{bmatrix} & * \\ & \end{bmatrix}.$$

From this formula it follows that $\mathrm{b}(2, F)$ is a Lie subalgebra of $\mathrm{gl}(2, F)$. Moreover, it is clear that

$$\mathrm{b}(2, F)^{(1)} = \begin{bmatrix} & * \\ & \end{bmatrix},$$

$$\mathrm{b}(2, F)^{(2)} = 0,$$

so that $\mathrm{b}(2, F)$ is solvable. □

The following corollary is a consequence of Proposition 2.1.14.

Corollary 2.1.15. *The F-subspace of* $\mathrm{sl}(2, F)$ *consisting of upper triangular matrices is a Lie subalgebra of* $\mathrm{sl}(2, F)$ *and is solvable.*

More generally, one has the following theorem, the proof of which will be omitted:

Theorem 2.1.16. *Let* $\mathrm{b}(n, F)$ *be the Lie algebra over F consisting of all upper triangular $n \times n$ matrices with entries from F. Then* $\mathrm{b}(n, F)$ *is solvable.*

2.2 Nilpotency

There is a stronger property than solvability. Let L be a Lie algebra over F. We define the **lower central series** of L to the sequence of ideals:

$$L^0 = L, \qquad L^1 = L', \qquad , L^k = [L, L^{k-1}], \quad k \geq 2.$$

Evidently, every element of the sequence $L^0, L^1, L^2, \dots$ is an ideal of L. Also, we have that

$$L = L^0 \supset L^1 \supset L^2 \supset \cdots$$

and $L^{(k)} \subset L^k$. The significant difference between the derived series and lower central series is that while $L^{(k)}/L^{(k+1)}$ and L^k/L^{k+1} are both abelian, the quotient L^k/L^{k+1} is in the center of L/L^{k+1}. We say that L is **nilpotent** if $L^k = 0$ for some non-negative integer k. It is clear that if L is nilpotent, then L is solvable.

It is not true that if a Lie algebra is solvable, then it is nilpotent. Consider $\mathrm{b}(2, F)$, the upper triangular 2×2 matrices over F. We have

$$\mathrm{b}(2, F)^1 = \begin{bmatrix} & * \\ & \end{bmatrix},$$

$$\mathrm{b}(2, F)^2 = \begin{bmatrix} & * \\ & \end{bmatrix},$$

$$\cdots$$

$$\mathrm{b}(2, F)^k = \begin{bmatrix} & * \\ & \end{bmatrix}, \quad k \geq 1.$$

On the other hand, the Lie algebra $\mathrm{n}(2, F)$ of strictly upper triangular 2×2 over F is nilpotent:

$$\mathrm{n}(2, F)^k = 0, \quad k \geq 1.$$

Proposition 2.2.1. *Let L be a Lie algebra over F. If L is nilpotent, then any Lie subalgebra of L is nilpotent. If $L/Z(L)$ is nilpotent, then L is nilpotent.*

Proof. The first assertion is clear. Assume that $L/Z(L)$ is nilpotent. We claim that $(L/Z(L))^k = (L^k + Z(L))/Z(L)$ for all non-negative integers k. This statement is clear if $k = 0$. Assume that the statement holds for k; we will prove that it holds for $k+1$. Now

$$\begin{aligned}(L/Z(L))^{k+1} &= [L/Z(L), (L/Z(L))^k] \\ &= [L/Z(L), (L^k + Z(L))/Z(L)] \\ &= (L^{k+1} + Z(L))/Z(L).\end{aligned}$$

This proves the statement by induction. Since $L/Z(L)$ is nilpotent, there exists a non-negative integer k such that $(L/Z(L))^k = 0$. It follows that $(L^k + Z(L))/Z(L) = 0$; this means that $L^k \subset Z(L)$. Therefore, $L^{k+1} = 0$. □

Theorem 2.2.2. *Let* $\mathrm{n}(n, F)$ *be the Lie algebra over* F *consisting of all strictly upper triangular* $n \times n$ *matrices with entries from* F*. Then* $\mathrm{n}(n, F)$ *is nilpotent.*

Chapter 3

The theorems of Engel and Lie

3.1 The theorems

In this chapter we will prove the following theorems:

Theorem 3.1.1 (Engel's Theorem)**.** *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional vector space over F. Suppose that L is a Lie subalgebra of* $\mathrm{gl}(V)$*, and that every element of L is a nilpotent linear transformation. Then there exists a basis for V such that in this basis every element of L is a strictly upper triangular matrix.*

Theorem 3.1.2 (Lie's Theorem)**.** *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional vector space over F. Suppose that L is a solvable Lie subalgebra of* $\mathrm{gl}(V)$*. Then there exists a basis for V such that in this basis every element of L is an upper triangular matrix.*

3.2 Weight spaces

Let V be a vector space over F, and let A be a Lie subalgebra of $\mathrm{gl}(V)$. Let $\lambda : A \to F$ be a linear map; we refer λ as a **weight** of L. We define

$$V_\lambda = \{v \in V : av = \lambda(a)v \text{ for all } a \in A\},$$

and refer to V_λ as the **weight space** for λ.

Lemma 3.2.1 (Invariance Lemma)**.** *Assume that F has characteristic zero. V be a finite-dimensional vector space over F, and let L be a Lie subalgebra of* $\mathrm{gl}(V)$*, and let A be an ideal of L. Let $\lambda : A \to F$ be a weight for A. The weight space V_λ is invariant under L.*

Proof. Let $w \in V_\lambda$ and $y \in L$. We must prove that yw is in V_λ, i.e., that $a(yw) = \lambda(a)yw$ for all $a \in A$. If $w = 0$, then this is clear; assume that $w \neq 0$. Let $a \in A$. Let $a \in A$. We have

$$\begin{aligned} a(yw) &= (ay)w \\ &= ([a, y] + ya)w \\ &= [a, y]w + yaw \\ &= \lambda([a, y])w + \lambda(a)(yw). \end{aligned}$$

Since $w \neq 0$, this calculation shows that we must prove that $\lambda([a, y]) = 0$.

To prove this, we consider the subspace U of V spanned by the vectors

$$w, yw, y^2w, \ldots$$

The subspace U is non-zero (because $w \neq 0$) and finite-dimensional (because V is finite-dimensional). Let m be the largest non-negative integer such that

$$w, yw, y^2w, \ldots, y^m w$$

are linearly independent. This set is a basis for U. We claim that for all $z \in A$ we have $zU \subset U$, and that moreover the matrix of z with respect to the basis $w, yw, y^2w, \ldots, y^m w$ has the form

$$\begin{bmatrix} \lambda(z) & * & \ldots & * \\ & \lambda(z) & \ldots & * \\ & & \ddots & \vdots \\ & & & \lambda(z) \end{bmatrix}.$$

We will prove this claim by induction on the columns. First of all, if $z \in A$, then $zw = \lambda(z)w$; this proves that the first column has the claimed form for all $z \in A$. For the second column, if $z \in A$, then

$$\begin{aligned} z(yw) &= [z, y]w + yzw \\ &= \lambda([z, y])w + \lambda(z)yw. \end{aligned}$$

This proves the claim for the second column. Assume that the claim has been proven for the first k columns with $k \geq 2$; we will prove it for the $k + 1$ column. Let $z \in A$. Then

$$\begin{aligned} z(y^k w) &= zyy^{k-1}w \\ &= [z, y](y^{k-1}w) + yz(y^{k-1}w). \end{aligned}$$

By the induction hypothesis, since $[z, y] \in A$, the vector $u_1 = [z, y](y^{k-1}w)$ is in the span of $w, yw, y^2w, \ldots, y^{k-1}w$. Also, by the induction hypothesis, there exists u_2 in the span of $w, yw, y^2w, \ldots, y^{k-2}w$ such that

$$z(y^{k-1}w) = \lambda(z)y^{k-1}w + u_2.$$

It follows that

$$\begin{aligned} z(y^k w) &= u_1 + y(\lambda(z)y^{k-1w} + u_2) \\ &= \lambda(z)y^k w + u_1 + yu_2. \end{aligned}$$

Since the vector $u_1 + yu_2$ is in the span of $w, yw, y^2w, \ldots, y^{k-1}w$, our claim follows.

Now we can complete the proof. We recall that we are trying to prove that $\lambda([a, y]) = 0$. Let $z = [a, y]$; then $z \in A$. By the last paragraph, z acts on U, and we have that the trace of the action of z on U is $(m+1)\lambda(z) = (m+1)\lambda([a, y])$. On the other hand, $z = [a, y] = ay - ya$, and a and y both act on U. This implies that trace of the action of z on U is zero. We conclude that $\lambda([a, y]) = 0$. □

Corollary 3.2.2. *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional vector space over $\mathbb{C}$. Let $x, y \in \mathrm{gl}(V)$. If x and y commute with $[x, y]$, then $[x, y]$ is nilpotent.*

Proof. Since our field is algebraically closed, it will suffice to prove that the only eigenvalue of $[x, y]$ is zero. Let c be an eigenvalue of $[x, y]$.

Let

$$L = Fx + Fy + F[x, y].$$

Since $[x, [x, y]] = [y, [x, y]] = 0$, the vector space L is a Lie subalgebra of $\mathrm{gl}(V)$. Let

$$A = F[x, y].$$

Evidently, A is an ideal of L; in fact $[z, a] = 0$ for all $z \in L$. Let $\lambda : A \to F$ be the linear functional such that $\lambda([x, y]) = c$. Then the weight space V_λ is

$$\begin{aligned} V_\lambda &= \{v \in V : av = \lambda(a)v \text{ for all } a \in A\} \\ &= \{v \in V : [x, y]v = cv\}. \end{aligned}$$

By the Lemma 3.2.1, the Invariance Lemma, V_λ is mapped by L into itself. Pick a basis for V_λ, and write the action of x and y on V_λ in this basis as matrices X and Y, respectively. On the one hand, we have $\mathrm{tr}[X, Y] = 0$, as usual. On the other hand, $[X, Y]$ acts by c on V_λ, which implies that $\mathrm{tr}[X, Y] = (\dim V_\lambda)c$. It follows that $c = 0$. □

3.3 Proof of Engel's Theorem

Lemma 3.3.1. *Let V be a finite-dimensional vector space over F, and let L be a Lie subalgebra of $\mathrm{gl}(V)$. Let $x \in L$. If x is nilpotent as a linear operator on V, then $\mathrm{ad}(x)$ is nilpotent as an element of $\mathrm{gl}(L)$.*

Proof. Let $y \in L$. By definition,

$$\mathrm{ad}(x)(y) = [x, y] = xy - yx,$$

$$\begin{aligned}
\operatorname{ad}(x)^2(y) &= \operatorname{ad}(x)(\operatorname{ad}(x)(y)) \\
&= \operatorname{ad}(x)(xy - yx) \\
&= [x, xy - yx] \\
&= x(xy - yx) - (xy - yx)x \\
&= x^2y - 2xyx + yx^2, \\
\operatorname{ad}(x)^3(y) &= \operatorname{ad}(x)(\operatorname{ad}(x)^2(y)) \\
&= [x, x^2y - 2xyx + yx^2] \\
&= x(x^2y - 2xyx + yx^2) - (x^2y - 2xyx + yx^2)x \\
&= x^3y - 2x^2yx + xyx^2 - x^2yx + 2xyx^2 - yx^3 \\
&= x^3y - 3x^2yx + 3xyx^2 - yx^3.
\end{aligned}$$

We claim that for all positive integers n,

$$\operatorname{ad}(x)^n(y) = \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^{n-k} y x^k.$$

We will prove this by induction on n. This claim is true if $n = 1$. Assume it holds for n; we will prove that it holds for $n + 1$. Now

$$\begin{aligned}
&\operatorname{ad}(x)^{n+1}(y) \\
&= \operatorname{ad}(x)(\operatorname{ad}(x)^n(y)) \\
&= [x, \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^{n-k} y x^k] \\
&= \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^{n-k+1} y x^k - \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^{n-k} y x^{k+1} \\
&= \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^{n-k+1} y x^k + \sum_{k=1}^{n+1} \binom{n}{k-1} (-1)^k x^{n-k+1} y x^k \\
&= x^{n+1}y + (-1)^{n+1} y x^{n+1} + \sum_{k=1}^{n} (\binom{n}{k} + \binom{n}{k-1})(-1)^k x^{n-k+1} y x^k \\
&= x^{n+1}y + (-1)^{n+1} y x^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} (-1)^k x^{n-k+1} y x^k \\
&\sum_{k=0}^{n+1} \binom{n+1}{k} (-1)^k x^{n+1-k} y x^k.
\end{aligned}$$

This proves our claim by induction.

From the formula we see that if m is positive integer such that $x^m = 0$, then $\operatorname{ad}(x)2m = 0$.

□

Lemma 3.3.2. *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional vector space over F, and let L be a Lie subalgebra of $\mathrm{gl}(V)$. Assume that L is non-zero, and that every element is a nilpotent linear transformation. Then there exists an non-zero vector v in V such that $xv = 0$ for all $x \in L$.*

Proof. We will prove this lemma by induction on $\dim L$. We cannot have $\dim L = 0$ because $L \neq 0$ by assumption. Assume first that $\dim L = 1$. Then $L = Fx$ for some $x \in L$. By assumption, x is a non-zero nilpotent linear transformation. This implies that there exists a positive integer such that $x^k \neq 0$ and $x^{k+1} = 0$. Since $x^k \neq 0$, there exists $w \in V$ such that $v = x^k w \neq 0$. Since $x^{k+1} = 0$, we have $xv = 0$. This proves the lemma in the case $\dim L = 1$.

Assume now that $\dim L > 1$ and that the lemma holds for all Lie algebras as in the statement of the lemma with dimension strictly less than $\dim L$. We need to prove that the statement of the lemma holds for L.

To begin, let A be a maximal proper Lie algebra of L; we will prove that A is an ideal of L and that $\dim A = \dim L - 1$. Set $\bar{L} = L/A$; this is vector space over F. Define

$$\varphi : A \longrightarrow \mathrm{gl}(\bar{L})$$

by

$$\varphi(a)\big(x + A\big) = [a, x] + A$$

for $a \in A$ and $x \in L$. The map φ is well-defined because A is a Lie subalgebra of L. We claim that φ is a Lie algebra homomorphism. Let $a, b \in A$ and $x \in L$. Then

$$\begin{aligned}
[\varphi(a), \varphi(b)](x + A) &= \varphi(a)\big([b, x] + A\big) - \varphi(b)\big([a, x] + A\big) \\
&= [a, [b, x]] - [b, [a, x]] + A \\
&= [a, [b, x]] + [b, [x, a]] + A \\
&= -[x, [a, b]] + A \\
&= [[a, b], x] + A \\
&= \varphi([a, b])(x + A).
\end{aligned}$$

This proves that φ is a Lie algebra homomorphism. Since φ is a Lie algebra homomorphism, it follows that $\varphi(A)$ is a Lie subalgebra of $\mathrm{gl}(\bar{L})$. We claim that the elements of $\varphi(A)$ are nilpotent as linear transformations in $\mathrm{gl}(\bar{L})$. Let $a \in A$. By Lemma 3.3.1, $\mathrm{ad}(a)$ is a nilpotent element of $\mathrm{gl}(L)$, i.e., there exists a positive integer k such that map $\mathrm{ad}(a)^k : L \to L$, defined by $x \mapsto \mathrm{ad}(a)^k(x) = [a, [a, [a, \cdots [a, x] \cdots]$, is zero, i.e., $[a, [a, [a, \cdots [a, x] \cdots] = 0$ for $x \in L$. The definition of φ implies that $\varphi(a)^k = 0$, as desired. We now may apply the induction hypothesis to $\varphi(A)$ and $\bar{L}$. By the induction hypothesis, there exists a non-zero vector $y + A \in \bar{L}$ such that $\varphi(a)(y + A) = 0$ for all $a \in A$. This means that $[a, y] \in A$ for all $a \in A$. Now define the vector subspace $A' = A + Fy$ of L Since $y + A$ is non-zero in $\bar{L}$, this is actually a direct sum, so that $A' = A \oplus Fy$. Moreover, because $[a, y] \in A$ for all $a \in A$, it follows that A' is a Lie subalgebra of L, and also that A is an ideal in A'. By the maximality of A, we must have $L = A \oplus Fy$. This proves that A is an ideal of L and $\dim A = \dim L - 1$.

We now use the induction hypothesis again. Evidently, $\dim A < \dim L$ and also the elements of the Lie algebra $A \subset \mathrm{gl}(V)$ are nilpotent linear transformations. By the induction hypothesis, there exists a non-zero vector $w \in V$ such that $aw = 0$ for all $a \in A$. Define

$$V_0 = \{v \in V : av = 0 \text{ for all } a \in A\}.$$

We have just noted that V_0 is non-zero. By the Invariance Lemma, Lemma 3.2.1, the vector subspace V_0 of V is mapped to itself under the elements of L. Recall the element y from above such that $L = A \oplus Fy$. We have $yV_0 \subset V_0$. Since y is a nilpotent linear transformation of V, the restriction of y to V_0 is also nilpotent. This implies that there exists a non-zero vector $v \in V_0$ such that $yv = 0$. We claim that $xv = 0$ for all $x \in L$. Let $x \in L$. Write $x = a + cy$ for some $a \in A$ and $c \in F$. Then

$$xv = (a + cy)v = av + cyv = 0 + 0 = 0.$$

This proves that the assertion of the lemma holds for L. By induction, the lemma is proven. □

Proof of Theorem 3.1.1, Engel's Theorem. We prove this theorem by induction on $\dim V$. If $\dim V = 0$, then there is nothing to prove. Assume that $\dim V \geq 1$, and that the theorem holds for all Lie algebras satisfying the hypothesis of the theorem that have dimension strictly less than $\dim V$.

By Lemma 3.3.2, there exists a non-zero vector $v \in V$ such that $xv = 0$ for all $x \in L$. Let $U = Fv$. Define $\bar{V} = V/U$. We consider the natural map

$$\varphi : L \longrightarrow \mathrm{gl}(\bar{V})$$

that sends x to the element of $\mathrm{gl}(\bar{V})$ defined by $w + U \mapsto xw + U$. This map is a Lie algebra homomorphism. Consider $\varphi(L)$. This is a Lie subalgebra of $\mathrm{gl}(\bar{V})$, and as linear transformations from $\bar{V}$ to $\bar{V}$, the elements of $\varphi(L)$ are nilpotent. By the induction hypothesis, there exists a ordered basis

$$v_1 + U, \dots, v_{n-1} + U$$

of $\bar{V}$ such that the elements of $\varphi(L)$ are strictly upper triangular in this basis. The vectors

$$v, v_1, \dots, v_{n-1}$$

form an ordered basis for V. It is evident that the elements of L are strictly upper triangular in this basis. □

3.4 Proof of Lie's Theorem

Lemma 3.4.1. *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional vector space over F, and let L be a Lie* *subalgebra of* gl(V)*. Assume that* L *is solvable. Then there exists a non-zero vector* $v \in V$ *such that* v *is an eigenvector for every element of* L*.*

Proof. We will prove this by induction on $\dim L$. If $\dim L = 0$, then there is nothing to prove. If $\dim L = 1$ then this follows from the assumption that F is algebraically closed. Assume that $\dim L > 1$, and that the assertion holds for all Lie algebras as in the statement with dimension strictly less than $\dim L$. Since L is solvable, the derived algebra L', which is actually an ideal of L, is a proper subspace of L. Choose a vector subspace A of L that contains L' such that $\dim A = \dim L - 1$. We claim that A is an ideal of L. Let $x \in L$ and $a \in A$. Then $[x, a] \in L' \subset A$, so that A is an ideal of L. Since A is an ideal of a solvable Lie algebra, A is also solvable; see Lemma 2.1.7. By the induction hypothesis, there exists a non-zero vector v and a weight $\lambda : A \to F$ such that $av = \lambda(a)v$ for $a \in A$. Thus, the weight space

$$V_\lambda = \{w \in V : aw = \lambda(a)w \text{ for } a \in A\}$$

is non-zero. By the Invariance Lemma, Lemma 3.2.1, the Lie algebra L maps the weight space V_λ to itself. Since $\dim A = \dim L - 1$, there exists $z \in L$ such that $L = A + Fz$. Consider the action of z on V_λ. Since F is algebraically closed, there exists a non-zero vector $w \in V_\lambda$ that is eigenvector for z; let $d \in F$ be the eigenvalue. We claim that w is an eigenvector for every element of L. Let $x \in L$, and write $x = a + cz$ for some $a \in A$ and $c \in F$. Then

$$xw = (a + cz)w = aw + czw = \lambda(a)w + cdw = (\lambda(a) + cd)w,$$

proving our claim. □

Proof of Theorem 3.1.2, Lie's Theorem. The proof of this theorem uses the last lemma, Lemma 3.4.1, and is almost identical to the proof of Engel's Theorem. The details will be omitted. □

Chapter 4

Some representation theory

4.1 Representations

Let L be a Lie algebra over F. A **representation** consists of a pair (φ, V), where V is a vector space over F and $\varphi : L \to \mathrm{gl}(V)$ is a Lie algebra homomorphism. Evidently, if V is a vector space over F, and $\varphi : L \to \mathrm{gl}(V)$ is a linear map, then the pair (φ, V) is a representation of L if and only if

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

for $x, y \in L$ and $v \in V$. Let (φ, V) be a representation of L. We will sometimes refer to a representation (φ, V) of L as an L-**module** and omit mention of φ by writing $x \cdot v = \varphi(x)v$ for $x \in L$ and $v \in V$. Note that with this convention we have

$$[x, y] \cdot v = x \cdot (y \cdot v) - y \cdot (x \cdot v)$$

for $x, y \in V$ and $v \in V$. If (φ, V) is a representation of L, and W is an F-vector subspace of V such that $\varphi(x)w \in W$ for $x \in L$ and $w \in W$, then we can define another representation of L with F-vector space W and homomorphism $L \to \mathrm{gl}(W)$ defined by $x \mapsto \varphi(x)|_W$ for $x \in L$. Such a representation is a called a **subrepresentation** of the representation (φ, V). We will also refer to W as an L-submodule of V. We say that the representation (φ, V) is **irreducible** if $V \neq 0$ and the only L-submodules of V are 0 and V. Let (φ_1, V_1) and (φ_2, V_2) be representations of L. An F-linear map $T : V_1 \to V_2$ is a **homomorphism of representations of** L, or an L-**map**, if $T(\varphi_1(x)v) = \varphi_2(x)T(v)$ for $x \in L$ and $v \in V$.

Let L be a Lie algebra over F. An important example of a representation of L is the **adjoint representation** of L, which has as F-vector space L and homomorphism $\mathrm{ad} : L \to \mathrm{gl}(L)$ given by

$$\mathrm{ad}(x)y = [x, y]$$

for $x, y \in L$.

We have also encountered another fundamental example. Assume that V is an F-vector space and L is Lie subalgebra of $\mathrm{gl}(V)$. This situation naturally defines a representation of L with F-vector space V and homomorphism $L \to \mathrm{gl}(V)$ given by inclusion. This representation is often referred to as the **natural representation**.

4.2 Basic results

Theorem 4.2.1. *Assume that F has characteristic zero and is algebraically closed. Let L be a solvable Lie algebra over F. If (φ, V) is an irreducible representation of L, then V is one-dimensional.*

Proof. Assume that (φ, V) is irreducible. We are given a Lie algebra homomorphism $\varphi : L \to \mathrm{gl}(V)$. Consider the image $\varphi(L)$. By Lemma 2.1.5, the Lie algebra $\varphi(L)$ is solvable. The solvable Lie algebra $\varphi(L)$ is a subalgebra of $\mathrm{gl}(V)$. By Lemma 3.4.1 there exists a non-zero vector $v \in V$ that is an eigenvector of every element of L. It follows that Fv is an L-subspace of V. Since (φ, V) is irreducible, it follows that $V = Fv$, so that V is one-dimensional. □

Theorem 4.2.2 (Schur's Lemma)**.** *Assume that F has characteristic zero and is algebraically closed. Let L be a Lie algebra over F. Let (φ, V) be a finite-dimensional irreducible representation of L. If $T : V \to V$ is an homomorphism of representations of L, then there exists a unique $c \in F$ such that $Tv = cv$ for $v \in V$.*

Proof. Since T is an F-linear map, and F is algebraically closed, T has a eigenvector, i.e., there exists a non-zero vector $v \in V$ and $c \in F$ such that $Tv = cv$. Set $R = T - c1_V$. Then R is a homomorphism of representations of L. Consider the kernel $\ker(T)$ of T; this is a nonzero L-submodule of V. Since V is irreducible, we must have $\ker(T) = V$, so that $T = c1_V$. □

Corollary 4.2.3. *Assume that F has characteristic zero and is algebraically closed. Let L be a Lie algebra over F. Let (φ, V) be a finite-dimensional irreducible representation of L. There exists a linear functional $\lambda : Z(L) \to F$ such that $\varphi(z)v = \lambda(z)v$ for $z \in Z(L)$ and $v \in V$.*

Proof. To define $\lambda : Z(L) \to F$ let $z \in Z(L)$. Consider the F-linear map $\varphi(z) : V \to V$. We claim that this is a homomorphism of representations of L. Let $x \in L$ and $v \in V$. Then

$$\begin{aligned}\varphi(x)(\varphi(z)v) &= \varphi([x,z])v + \varphi(z)(\varphi(x)v)\\ &= 0 + \varphi(z)(\varphi(x)v)\\ &= \varphi(z)(\varphi(x)v).\end{aligned}$$

This proves our claim. Applying Theorem 4.2.2, Schur's Lemma, to $\varphi(z)$, we see that there exists a unique $c \in F$ such that $\varphi(z)v = cv$ for $v \in V$. We now define $\lambda(z) = c$. It is straightforward to verify that λ is a linear map. □

4.3 Representations of sl(2)

In this section we will determine all the irreducible representations of $\mathrm{sl}(2,F)$ when F has characteristic zero and is algebraically closed.

We recall that

$$\mathrm{sl}(2,F) = Fe + Fh + Ff$$

where

$$e = \begin{bmatrix} & 1 \\ & \end{bmatrix}, \qquad h = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}, \qquad f = \begin{bmatrix} & \\ 1 & \end{bmatrix}.$$

We have

$$[e,f] = h, \qquad [e,h] = -2e, \qquad [f,h] = 2f.$$

Lemma 4.3.1. *Let V be a vector space over F, and let $\varphi : \mathrm{sl}(2,F) \to \mathrm{gl}(V)$ be an F-linear map. Define*

$$E = \varphi(e), \qquad H = \varphi(h), \qquad F = \varphi(f).$$

The map φ is a representation of $\mathrm{sl}(2,F)$ if and only if

$$[E,F] = H, \qquad [E,H] = -2E, \qquad [F,H] = 2F.$$

Proof. Assume that φ is a representation. Then, by definition, φ is a Lie algebra homomorphism. Applying φ to $[e,f] = h, [e,h] = -2e$, and $[f,h] = 2f$ yields $[E,F] = H, [E,H] = -2E$, and $[F,H] = 2F$.

Now suppose that the relations $[E,F] = H, [E,H] = -2E$, and $[F,H] = 2F$ hold. By linearity, to prove that φ is a Lie algebra homomorphism, it suffices to prove that $\varphi([e,f]) = [\varphi(e), \varphi(f)]$, $\varphi([e,h]) = [\varphi(e), \varphi(h)]$, and $\varphi([f,h]) = [\varphi(f), \varphi(h)]$; this follows from the assumed relations and the definitions of E, F, and H. □

Let d be a non-negative inteber. Let V_d be F-vector space of homogeneous polynomials in the variables X and Y of degree d with coefficients from F. The F-vector space V_d has dimension $d+1$, with basis

$$X^d, \quad X^{d-1}Y, \quad X^{d-2}Y^2, \quad \dots, \quad Y^d.$$

We define linear maps

$$E, H, F : V_d \longrightarrow V_d$$

by

$$\begin{aligned} Ep &= X\frac{\partial p}{\partial Y}, \\ Fp &= Y\frac{\partial p}{\partial X}, \\ Hp &= X\frac{\partial p}{\partial X} - Y\frac{\partial p}{\partial Y}. \end{aligned}$$

Lemma 4.3.2. *Let d be a non-negative integer. The F-linear operators E, F and H act on V_d and satisfy the relations $[E,F] = H, [E,H] = -2E$, and $[F,H] = 2F$.*

Proof. Since E, F, and H are linear operators, it suffices to prove that the claimed identities hold on the above basis for V_d. For k and integer we define

$$p_k = X^{d-k}Y^k.$$

Let $k \in \{0, 1, 2, \ldots, d\}$. We calculate:

$$\begin{aligned}
Ep_k &= E(X^{d-k}Y^k) \\
&= kX^{d-(k-1)}Y^{k-1} \\
&= kp_{k-1}, \\
Fp_k &= F(X^{d-k}Y^k) \\
&= (d-k)X^{d-(k+1)}Y^{k+1} \\
&= (d-k)p_{k+1}, \\
Hp_k &= H(X^{d-k}Y^k) \\
&= (d-k)X^{d-k}Y^k - kX^{d-k}Y^k \\
&= (d-2k)p_k.
\end{aligned}$$

To summarize:

$$Ep_k = k \cdot p_{k-1}, \quad Fp_k = (d-k) \cdot p_{k+1}, \quad Hp_k = (d-2k) \cdot p_k.$$

These formulas show that E, F and H act on V_d. We now have:

$$\begin{aligned}
[E,F]p_k &= EFp_k - FEp_k \\
&= (d-k)Ep_{k+1} - kFp_{k-1} \\
&= (d-k)(k+1)p_k - k(d-k+1)p_k \\
&= (d-2k)p_k \\
&= Hp_k.
\end{aligned}$$

This proves that $[E,F] = H$. Next,

$$\begin{aligned}
[E,H]p_k &= EHp_k - HEp_k \\
&= (d-2k)kp_{k-1} - k(d-2k+2)p_{k-1} \\
&= -2kp_{k-1} \\
&= -2Ep_k.
\end{aligned}$$

This proves that $[E,H] = -2E$. Finally,

$$\begin{aligned}
[F,H]p_k &= FHp_k - HFp_k \\
&= (d-2k)Fp_k - (d-k)Hp_{k+1}
\end{aligned}$$

$$\begin{aligned} &= (d-2k)(d-k)p_{k+1} - (d-k)(d-2k-2)p_{k1} \\ &= 2(d-k)p_{k-1} \\ &= 2Fp_k. \end{aligned}$$

This proves that $[F, H] = F$, and completes the proof. □

Proposition 4.3.3. *Let the notation be as in Lemma 4.3.2. The linear map $\varphi : \mathrm{sl}(2, F) \to \mathrm{gl}(V_d)$ determined by setting $\varphi(e) = E$, $\varphi(f) = F$, and $\varphi(h) = H$ is a Lie algebra homomorphism, so that (φ, V_d) is a representation of $\mathrm{sl}(2, F)$.*

Proof. This follows from Lemma 4.3.2 and Lemma 4.3.1. □

Let d be a non-negative integer. We note from the proof of Lemma 4.3.1 that the basis p_k, $k \in \{0, \ldots, d\}$, of V_d is such that

$$H \cdot p_k = (d-2k)p_k.$$

In other words, V_d has a basis of eigenvectors for H with one-dimensional eigenspaces. Moreover, we see that the matrices of E, F, and H are:

$$\text{matrix of } E = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & d \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix},$$

$$\text{matrix of } F = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ d & 0 & 0 & \cdots & 0 & 0 \\ 0 & d-1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & d-2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix},$$

$$\text{matrix of } H = \begin{bmatrix} d & 0 & 0 & \cdots & 0 \\ 0 & d-2 & 0 & \cdots & 0 \\ 0 & 0 & d-4 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & -d \end{bmatrix}.$$

Proposition 4.3.4. *Let d be a non-negative integer. The representation of $\mathrm{sl}(2, F)$ on V_d is irreducible.*

Proof. Let W be a non-zero $\mathrm{sl}(2, F)$-subspace of V_d. Since W is an $\mathrm{sl}(2, F)$-subspace, the characteristic polynomial of $H|_W$ divides the characteristic polynomial of H. The characteristic polynomial of H splits over F with distinct roots. It follows that the characteristic polynomial of $H|_W$ also splits over F

with distinct roots. In particular, $H|_W$ has an eigenvector. This implies that for some $k \in \{0, \ldots, d\}$ we have $p_k \in W$. By applying powers of E and F we find that all the vectors $v_0, \ldots, v_d$ are contained in W. Hence, $W = V_d$ and V_d is irreducible. ☐

Lemma 4.3.5. *Let V be a representation of $\mathrm{sl}(2, F)$. Assume that v is an eigenvector for h with eigenvalue $\lambda \in F$. Either $ev = 0$, or ev is non-zero and ev is an eigenvector for h such that*

$$h(ev) = (\lambda + 2)ev.$$

Similarly, either $fv = 0$, or fv is non-zero and ev is an eigenvector for h such that

$$h(fv) = (\lambda - 2)fv.$$

Proof. Assume that ev is non-zero. We have

$$\begin{aligned} h(ev) &= (eh + [h, e])v \\ &= (eh + 2e)v \\ &= e(hv) + 2ev \\ &= \lambda ev + 2ev \\ &= (\lambda + 2)ev. \end{aligned}$$

Assume that fv is non-zero. We have

$$\begin{aligned} h(fv) &= (fh + [h, f])v \\ &= (fh - 2f)v \\ &= f(hv) - 2fv \\ &= \lambda fv - 2fv \\ &= (\lambda - 2)fv. \end{aligned}$$

This completes the proof. ☐

Lemma 4.3.6. *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional representation of $\mathrm{sl}(2, F)$. Then there exists an eigenvector $v \in V$ for h such that $ev = 0$.*

Proof. Since F is algebraically closed, h has an eigenvector u with eigenvalue λ. Consider the sequence of eigenvectors

$$u, \quad eu, \quad e^2u, \quad \ldots.$$

By Lemma 4.3.5, because the numbers $\lambda, \lambda+2, \lambda+4, \ldots$ are mutually distinct, if infinitely many of these vectors are non-zero, then V is infinite-dimensional. Since V is finite-dimensional, all but finitely many of these vectors are non-zero. In particular, there exists a non-negative integer k such that $e^k u \neq 0$ but $e^{k+1}u = 0$. Set $v = e^k u$. Then $v \neq 0$, and by Lemma 4.3.5, v is an eigenvector for h and $ev = 0$. ☐

Theorem 4.3.7. *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional irreducible representation of* $\mathrm{sl}(2, F)$*. Then V is isomorphic to V_d where* $\dim V = d + 1$.

Proof. Since V is irreducible, we have $\dim V > 0$ by definition. By Lemma 4.3.6, there exists an eigenvector $v \in V$ for h with eigenvalue λ such that $Ev = 0$. Consider the sequence of vectors

$$v, \quad fv, \quad f^2v, \quad \dots.$$

By Lemma 4.3.5, because the numbers $\lambda, \lambda - 2, \lambda - 4, \dots$ are mutually distinct, if infinitely many of these vectors are non-zero, then V is infinite-dimensional. Since V is finite-dimensional, all but finitely many of these vectors are non-zero. In particular, there exists a non-negative integer d such that $f^d v \neq 0$ but $f^{d+1}v = 0$. We claim that the F-subspace W spanned by the vectors

$$v, \quad fv, \quad f^2v, \quad \dots, \quad f^d v$$

is an $\mathrm{sl}(2, F)$-subspace. Since $f^{d+1}v = 0$ it follows that W is invariant under f. The subspace W is invariant under h by Lemma 4.3.5. To complete the argument that W is invariant under $\mathrm{sl}(2, F)$ it will suffice to prove that W is invariant under e. We will prove that $e(f^j v) \in W$ by induction on j for $j \in \{0, \dots, d\}$. We have $ev = 0 \in W$. If $d = 0$, then we are done; assume that $d > 0$. Assume that j is a positive integer such that $1 \leq j < d$, and that $ev, e(fv), \dots, e(f^{j-1}v) \in W$; we will prove that $e(f^j v) \in W$. We have

$$\begin{aligned} e(f^j v) &= ef(f^{j-1}v) \\ &= (fe + [e, f])(f^{j-1}v) \\ &= (fe + h)(f^{j-1}v) \\ &= f\big(e(f^{j-1}v)\big) + h(f^{j-1}v). \end{aligned}$$

The vector $f\big(e(f^{j-1}v)\big)$ is in W by the induction hypothesis, and the vector $h(f^{j-1}v)$ is in W because W is invariant under h. This proves our claim by induction, so that W is an $\mathrm{sl}(2, F)$-subspace of V. Since V is irreducible and W is non-zero, we obtain $V = W$. In particular, we see that $\dim V = \dim W = d + 1$.

Next, we will prove that $\lambda = d$. To prove this, consider the matrix of h with respect to the basis

$$v, \quad fv, \quad f^2v, \quad \dots, \quad f^d v$$

of $V = W$. The matrix of h with respect to this basis is:

$$\begin{bmatrix} \lambda & & & & \\ & \lambda - 2 & & & \\ & & \lambda - 4 & & \\ & & & \ddots & \\ & & & & \lambda - 2d \end{bmatrix}.$$

It follows that

$$\begin{aligned} \operatorname{trace}(h) &= (d+1)\lambda - 2(1+2+\cdots+d) \\ &= (d+1)\lambda - d(d+1) \\ &= (d+1)(\lambda - d). \end{aligned}$$

On the other hand,

$$\begin{aligned} \operatorname{trace}(h) &= \operatorname{trace}([e,f]) \\ &= \operatorname{trace}(ef - fe) \\ &= \operatorname{trace}(ef) - \operatorname{trace}(fe) \\ &= \operatorname{trace}(ef) - \operatorname{trace}(ef) \\ &= 0. \end{aligned}$$

Since F has characteristic zero we conclude that $\lambda = d$.

Now we define an F-linear map $T : V \to V_d$ by setting

$$T(f^k v) = F^k X^d$$

for $k \in \{0, \ldots, d\}$. This map is evidently an isomorphism. To complete the proof we need to prove that T is an $\mathrm{sl}(2, F)$-map. First we prove that $T(fw) = FT(w)$ for $w \in V$. To prove this it suffices to prove that this holds for $w = f^k v$ for $k \in \{0, \ldots, d\}$. If $k \in \{0, \ldots, d-1\}$, then

$$\begin{aligned} T(f(f^k v)) &= T(f^{k+1} v) \\ &= F^{k+1} X^d \\ &= FT(f^k v). \end{aligned}$$

If $k = d$, then

$$\begin{aligned} T(f(f^d v)) &= T(0) \\ &= 0 \\ &= F^{d+1} X^d \\ &= FT(f^d v). \end{aligned}$$

Next we prove that $T(hw) = HT(w)$ for $w \in V$. Again, it suffices to prove that this holds for $w = f^k v$ for $k \in \{0, \ldots, d\}$. Let $k \in \{0, \ldots, d\}$. Then

$$\begin{aligned} T(h(f^k v)) &= T((d-2k)(f^k v)) \\ &= (d-2k)T(f^k v) \\ &= (d-2k)F^k X^d \\ &= H(f^k X^d) \\ &= H(T(f^k v)). \end{aligned}$$

Finally, we need to prove that $T(ew) = ET(w)$ for $w = f^k v$ for $k \in \{0, \ldots, d\}$. We will prove this by induction on k. If $k = 0$, this clear because $T(ef^0 v) = T(0) = 0 = EX^d = ET(f^0 v)$. Assume that $k \in \{1, \ldots, d\}$ and $T(e(f^j v)) = ET(f^j v)$ for $j \in \{0, \ldots, k-1\}$; we will prove that $T(e(f^k v)) = ET(f^k v)$. Now

$$\begin{aligned}
T(e(f^k v)) &= T(eff^{k-1}v) \\
&= T((fe + [e, f])f^{k-1}v) \\
&= T(fef^{k-1}v) + T(hf^{k-1}v) \\
&= FT(ef^{k-1}v) + HT(f^{k-1}v) \\
&= FET(f^{k-1}v) + HT(f^{k-1}v) \\
&= (FE + H)T(f^{k-1}v) \\
&= EFT(f^{k-1}v) \\
&= ET(f^k v).
\end{aligned}$$

By induction, this completes the proof. □

Chapter 5

Cartan's criteria

5.1 The Jordan-Chevalley decomposition

Theorem 5.1.1. (Jordan-Chevalley decomposition) *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional F-vector space. Let $x \in \mathrm{gl}(V)$. There exist unique elements $x_s, x_n \in \mathrm{gl}(V)$ such that $x = x_s + x_n$, x_s is semi-simple (i.e., diagonalizable), x_n is nilpotent, and x_s and x_n commute. Moreover, there exist polynomials $s_x(X), n_x(X) \in F[X]$ such that $s_x(X)$ and $n_x(X)$ do not have constant terms and $x_s = s_x(x)$ and $n = n_x(x)$.*

Lemma 5.1.2. *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional F-vector space. Let $x, y \in \mathrm{gl}(V)$.*

1. *If x and y commute, then x, y, x_s, x_n, y_s, and y_n pairwise commute.*

2. *If x and y commute, then $(x+y)_s = x_s + y_s$ and $(x+y)_n = x_n + y_n$.*

Proof. Proof of 1. Assume that x and y commute. We have

$$\begin{aligned} xy_s &= xs_y(y) \\ &= s_y(y)x \\ &= y_s x. \end{aligned}$$

Similarly, x commutes with y_n, y commutes with x_s, and y commutes with x_n. Also, we now have

$$\begin{aligned} x_s y_s &= x_s s_y(y) \\ &= s_y(y)x_s \\ &= y_s x_s. \end{aligned}$$

Similarly, x_s commutes with y_n, x_n commutes with y_s, and x_n commutes with y_n.

Proof of 2. Assume that x and y commute. Evidently, $x + y = (x_s + y_s) + (x_n + y_n)$. Since x_s and y_s commute, x_s and y_s can be simultaneously diagonalized; this implies that $x_s + y_s$ is semi-simple. Similarly, since x_n and y_n commute and are nilpotent, $x_n + y_n$ is also nilpotent. Since $x_s + x_n$ and $y_s + y_n$ commute, by uniqueness we have $(x+y)_s = x_s + y_s$ and $(x+y)_n = x_n + y_n$. □

Lemma 5.1.3. *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional F-vector space. Let $x \in \mathrm{gl}(V)$, and consider $\mathrm{ad}(x) : \mathrm{gl}(V) \to \mathrm{gl}(V)$. We have $\mathrm{ad}(x)_s = \mathrm{ad}(x_s)$ and $\mathrm{ad}(x)_n = \mathrm{ad}(x_n)$.*

Proof. Because $x = x_s + x_n$, we have $\mathrm{ad}(x) = \mathrm{ad}(x_s) + \mathrm{ad}(x_n)$. To complete the proof we need to show that $\mathrm{ad}(x_s)$ is simi-simple, $\mathrm{ad}(x_n)$ is nilpotent, and $\mathrm{ad}(x_s)$ and $\mathrm{ad}(x_n)$ commute. By Lemma 3.3.1 the operator $\mathrm{ad}(x_n)$ is nilpotent. To see that $\mathrm{ad}(x_s)$ is diagonalizable, let $v_1, \dots, v_n$ be an ordered basis for V such that x_s is diagonal in this basis. Let $\lambda_1, \dots, \lambda_n \in F$ be such that $x_s(v_i) = \lambda_i v_i$ for $i \in \{1, \dots, n\}$. For $i, j \in \{1, \dots, n\}$ let $e_{ij} \in \mathrm{gl}(V)$ be the standard basis for $\mathrm{gl}(V)$ with respect to the basis $v_1, \dots, v_n$, so that the matrix of e_{ij} has i, j-th entry 1 and all other entries 0. Let $i, j \in \{1, \dots, n\}$. We have

$$\begin{aligned}
\mathrm{ad}(x_s)(e_{ij}) &= [x_s, e_{ij}] \\
&= x_s e_{ij} - e_{ij} x_s \\
&= \lambda_i e_{ij} - \lambda_j e_{ij} \\
&= (\lambda_i - \lambda_j) e_{ij}.
\end{aligned}$$

It follows that $\mathrm{ad}(x_s)$ is diagonalizable. To see that $\mathrm{ad}(x_s)$ and $\mathrm{ad}(x_n)$ commute, let $y \in \mathrm{gl}(V)$. Then

$$\begin{aligned}
\big(\mathrm{ad}(x_s)\mathrm{ad}(x_n)\big)(y) &= \mathrm{ad}(x_s)\big(\mathrm{ad}(x_n)(y)\big) \\
&= \mathrm{ad}(x_s)\big([x_n, y]\big) \\
&= [x_s, [x_n, y]] \\
&= [x_s, x_n y - y x_n] \\
&= x_s(x_n y - y x_n) - (x_n y - y x_n) x_s \\
&= x_s x_n y - x_s y x_n - x_n y x_s + y x_n x_s \\
&= x_n x_s y - x_s y x_n - x_n y x_s + y x_s x_n \\
&= x_n(x_s y - y x_s) - (x_s y - y x_s) x_n \\
&= [x_n, x_s y - y x_s] \\
&= [x_n, [x_s, y]] \\
&= \big(\mathrm{ad}(x_n)\mathrm{ad}(x_s)\big)(y).
\end{aligned}$$

It follows that $\mathrm{ad}(x_s)$ and $\mathrm{ad}(x_n)$ commute. □

5.2 Cartan's first criterion: solvability

Lemma 5.2.1. *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional F-vector space. Let A and B be F-vector* *subspaces of* $\mathrm{gl}(V)$ *such that* $A \subset B$*. Define*

$$M = \{x \in \mathrm{gl}(V) : [x, B] \subset A\} = \{x \in \mathrm{gl}(V) : \mathrm{ad}(x)(B) \subset A\}.$$

Let $x \in M$*. If* $\mathrm{tr}(xy) = 0$ *for all* $y \in M$*, then* x *is nilpotent.*

Proof. Assume that $x \in M$ and $\mathrm{tr}(xy) = 0$ for all $y \in M$. Set $s = x_s$ and $n = x_n$. We need to prove that $s = 0$. Since s is diagonalizable, there exists an ordered basis $v_1, \dots, v_n$ such that the matrix of s with respect to this basis is diagonal, i.e., there exist $\lambda_1, \dots, \lambda_n \in F$ such that the matrix of s in this basis is:

$$\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

We need to prove that this matrix is zero. Since F has characteristic zero, F contains $\mathbb{Q}$. Let W be the $\mathbb{Q}$-vector subspace of F spanned by $\lambda_1, \dots, \lambda_n$, so that

$$W = \mathbb{Q}\lambda_1 + \cdots + \mathbb{Q}\lambda_n.$$

We need to prove that $W = 0$. To prove this we will prove that every $\mathbb{Q}$ linear functional on W is zero.

Let $f : W \to \mathbb{Q}$ be a $\mathbb{Q}$ linear map. To prove that $f = 0$ it will suffice to prove that $f(\lambda_1) = \cdots = f(\lambda_n) = 0$. Define $y \in \mathrm{gl}(V)$ to be the element with matrix

$$\begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix}$$

with respect to the ordered basis $v_1, \dots, v_n$. Let E_{ij}, $i, j \in \{1, \dots, n\}$ be the standard basis for $\mathrm{gl}(V)$ with respect to the ordered basis $v_1, \dots, v_n$ for V. Calculations show that

$$\begin{aligned} \mathrm{ad}(s)\big(E_{ij}\big) &= (\lambda_i - \lambda_j)E_{ij}, \\ \mathrm{ad}(y)\big(E_{ij}\big) &= (f(\lambda_i) - f(\lambda_j))E_{ij} = f(\lambda_i - \lambda_j)E_{ij} \end{aligned}$$

for $i, j \in \{1, \dots, n\}$. Consider the set

$$\{(\lambda_i - \lambda_j, f(\lambda_i - \lambda_j)) : i, j \in \{1, \dots, n\}\} \cup (0, 0).$$

Let $r(X) \in F[X]$ be the Langrange interpolation polynomial for this set. Then $r(X)$ does not have a contant term because $r(0) = 0$. Also,

$$r(\lambda_i - \lambda_j) = f(\lambda_i - \lambda_j)$$

for $i, j \in \{1, \dots, n\}$. It follows that

$$r(\mathrm{ad}(s)) = \mathrm{ad}(y).$$

By Lemma 5.1.3 we have $\mathrm{ad}(s) = \mathrm{ad}(x)_s$. Hence, by Theorem 5.1.1, there exists a polynomial $p(X) \in F[X]$ with no constant term such that

$$\mathrm{ad}(s) = p(\mathrm{ad}(x)).$$

We now have

$$\mathrm{ad}(y) = p(r(\mathrm{ad}(x)).$$

Now, because $x \in M$, we have $\mathrm{ad}(x)(B) \subset A$. We claim that this implies that $\mathrm{ad}(x)^k(B) \subset A$ for all positive integers k. We prove this claim by induction on k. The claim holds for $k = 1$. Assume it holds for k. Then

$$\begin{aligned}
\mathrm{ad}(x)^{k+1}(B) &= \mathrm{ad}(x)(\mathrm{ad}(x)^k(B)) \\
&\subset \mathrm{ad}(x)(A) \\
&\subset \mathrm{ad}(x)(B) \\
&\subset A.
\end{aligned}$$

This proves the claim. Since $\mathrm{ad}(y)$ is a polynomial in $\mathrm{ad}(x)$ with constant term we conclude that $\mathrm{ad}(y)(B) \subset A$. This implies that $y \in M$, by definition. By our assumption on x we have $\mathrm{tr}(xy) = 0$. This means that:

$$0 = \mathrm{tr}(xy) = f(\lambda_1)\lambda_1 + \cdots + f(\lambda_n)\lambda_n.$$

Applying f to this equation, we get, because $f(\lambda_1), \ldots, f(\lambda_n) \in \mathbb{Q}$,

$$\begin{aligned}
0 &= f\big(f(\lambda_1)\lambda_1 + \cdots + f(\lambda_n)\lambda_n\big) \\
&= f(\lambda_1)^2 + \cdots + f(\lambda_n)^2.
\end{aligned}$$

Since $f(\lambda_1), \ldots, f(\lambda_n) \in \mathbb{Q}$ we obtain $f(\lambda_1) = \cdots = f(\lambda_n) = 0$. This implies that $f = 0$, as desired. □

Lemma 5.2.2. *Let L be a Lie algebra over F. Let K be an extension of F. Define $L_K = K \otimes_F L$. Then L_K is a K-vector space. There exists a unique K-bilinear form*

$$[\cdot,\cdot] : L_K \times L_K \to L_K$$

such that

$$[a \otimes x, b \otimes y] = ab \otimes [x, y]$$

for $a, b \in K$ and $x, y \in L$. With $[\cdot,\cdot]$, L_K is a Lie algebra over K. The F-Lie algebra L is solvable if and only if the K-Lie algebra L_K is solvable. The F-Lie algebra L is nilpotent if and only if the K-Lie algebra L_K is nilpotent.

Proof. It is clear that the K-bilinear form mentioned in the statement of the lemma is unique if it exists. To prove existence, we note first that the abelian group $\mathrm{Hom}_K(L_K, L_K)$ is naturally an K-vector space. For each $(a, x) \in K \times L$, let $T_{(a,x)} : L_K \to L_K$ be the K-linear map such that $T_{(a,x)}(b \otimes y) = ab \otimes [x, y]$ for $b \in K$ and $y \in L$. The map $T_{(a,x)}$ is exists because the function $K \times L \to L_K$ defined by $(b, y) \mapsto ab \otimes [x, y]$ for $b \in K$ and $y \in L$ is F-bilinear; a calculation shows that it is K-linear. The map $K \times L \to \mathrm{Hom}_K(L_K, L_K)$ defined by $(a, x) \mapsto T_{(a,x)}$ for $a \in K$ and $x \in L$ is an F-bilinear map. It follows that there exists a unique F-linear map $B : L_K = K \otimes_F L \to \mathrm{Hom}_F(L_K, L_K)$ sending $a \otimes x$ to $T_{(a,x)}$ for $a \in K$ and $x \in L$. Now define $L_K \times L_K \to L_K$ by $(z_1, z_2) \mapsto B(z_1)(z_2)$. Let $a, b \in K$ and $x, y \in L$. Then

$$\begin{aligned} B(a \otimes x)(b \otimes y) &= T_{(a,x)}(b \otimes y) \\ &= ab \otimes [x, y]. \end{aligned}$$

It is easy to verify that the map $L_K \times L_K \to L_K$ is K-bilinear. It follows that the desired K-bilinear form exists.

Next, a calculation shows that $[\cdot, \cdot] : L_K \times L_K \to L_K$ is a Lie bracket, so that L_K is a Lie algebra over K with this Lie bracket.

Let k be a non-negative integer. We will prove by induction on k that $K \otimes_F L^{(k)} = L_K^{(k)}$. This is clear if $k = 0$. Assume it holds for k. We have

$$\begin{aligned} K \otimes_F L^{(k+1)} &= K \otimes_F [L^{(k)}, L^{(k)}] \\ &= [K \otimes_F L^{(k)}, K \otimes_F L^{(k)}] \\ &= [L_K^{(k)}, L_K^{(k)}] \\ &= L_K^{(k+1)}. \end{aligned}$$

This completes the proof by induction. It follows that $L^{(k)} = 0$ if and only if $L_K^{(k)} = 0$. Hence, L is solvable if and only if L_K is solvable.

Similarly, L is nilpotent if and only if L_K is nilpotent. □

Lemma 5.2.3. *Assume that F has characteristic zero. Let V be a finite-dimensional F-vector space. Let L be a Lie subalgebra of* $\mathrm{gl}(V)$*. If* $\mathrm{tr}(xy) = 0$ *for all $x \in L'$ and $y \in L$, then L is solvable.*

Proof. Assume that $\mathrm{tr}(xy) = 0$ for all $x \in L'$ and $y \in L$. We need to prove that L is solvable.

We will first prove that we may assume that F is algebraically closed. Let $K = \bar{F}$, the algebraic closure of F. Define $V_K = K \otimes_F V$. Then V_K is a K-vector space, and $\dim_K V_K = \dim_F V$. There is a natural inclusion

$$K \otimes \mathrm{Hom}_F(V, V) \hookrightarrow \mathrm{Hom}_K(V_K, V_K)$$

of K-algebras. As both of these K-algebras have the same dimension over K, this map is an isomorphism. Moreover, the diagram

$$\begin{array}{ccc} K \otimes_F \mathrm{Hom}_F(V, V) & \xrightarrow{\sim} & \mathrm{Hom}_K(V_K, V_K) \\ {\scriptstyle \mathrm{id}\otimes\mathrm{tr}}\downarrow & & {\scriptstyle \mathrm{tr}}\downarrow \\ K & \underset{\sim}{\xrightarrow{\mathrm{id}}} & K \end{array}$$

commutes. Define $L_K = K \otimes_F L$; by Lemma 5.2.2, L_K is a Lie algebra over K with Lie bracket as defined in this lemma. Also, by this lemma, to prove that L is solvable it will suffice to prove that L_K is solvable. In addition, the proof of Lemma 5.2.2 shows that $L'_K = K \otimes_F L = K \otimes_F L'$. Let $a, b \in K$, $x \in L'$ and $y \in L$. Then by the commutativity of the diagram,

$$\begin{aligned}\operatorname{tr}((a \otimes x)(b \otimes y)) &= \operatorname{tr}(ab \otimes xy)\\ &= ab \otimes \operatorname{tr}(xy)\\ &= 0.\end{aligned}$$

It follows that $\operatorname{tr}(wz) = 0$ for all $w \in L'_K$ and $z \in L_K$. Consequently, we may assume that F is algebraically closed.

We have the following sequence of ideals of L:

$$0 \subset L' \subset L.$$

The quotient L/L' is abelian. Thus, by Proposition 2.1.4, to prove that L is solvable it will suffice to prove that L' is solvable; and to prove that L' is solvable, it will suffice to prove that L' is nilpotent. By Engel's Theorem, Theorem 3.1.1, to prove that L' is nilpotent it will suffice to prove that every element of L' is a nilpotent linear transformation (because any subalgebra of $\mathrm{gl}(n, F)$ consisting of strictly upper triangular matrices is nilpotent). Let $x \in L'$. Define $A = L'$ and $B = L$. Evidently, $A \subset B \subset \mathrm{gl}(V)$. If M is as in the statement of Lemma 5.2.1, then we have

$$M = \{x \in \mathrm{gl}(V) : [x, L] \subset L'\}.$$

Evidently, $L \subset M$; in particular, $x \in M$. Let $y \in M$. We claim that $\operatorname{tr}(xy) = 0$. Since $x \in L'$, there exist a positive integer m and $x_i, z_i \in L$ for $i \in \{1, \dots, m\}$ such that

$$x = [x_1, z_1] + \dots + [x_m, z_m].$$

Now

$$\begin{aligned}\operatorname{tr}(xy) &= \sum_{i=1}^{m} \operatorname{tr}([x_i, z_i]y)\\ &= \sum_{i=1}^{m} \operatorname{tr}((x_i z_i - z_i x_i)y)\\ &= \sum_{i=1}^{m} \big(\operatorname{tr}(x_i z_i y) - \operatorname{tr}(z_i x_i y))\big)\\ &= \sum_{i=1}^{m} \big(\operatorname{tr}(x_i z_i y) - \operatorname{tr}(x_i y z_i))\big)\\ &= \sum_{i=1}^{m} \operatorname{tr}(x_i [z_i, y])\end{aligned}$$

$$= -\sum_{i=1}^{m} \operatorname{tr}([y, z_i]x_i).$$

If $i \in \{1, \ldots, m\}$, then since $y \in M$, we have $[y, z_i] \in L'$. By our assumption we now have $\operatorname{tr}([y, z_i]x_i) = 0$ for $i \in \{1, \ldots, m\}$. This implies that $\operatorname{tr}(xy) = 0$, proving our claim. From Lemma 5.2.1 we now conclude that x is nilpotent. □

Theorem 5.2.4 (Cartan's First Criterion)**.** *Assume that F has characteristic zero. Let L be a finite-dimensional Lie algebra over F. The Lie algebra L is solvable if and only if* $\operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(y)) = 0$ *for all $x \in L'$ and $y \in L$.*

Proof. Assume that L is solvable; we need to prove that $\operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(y)) = 0$ for all $x \in L'$ and $y \in L$. We will first prove that we may assume that F is algebraically closed. Let $K = \bar{F}$ be the algebraic closure of F. Define $L_K = K \otimes_F L$. Then L_K is a Lie algebra over K, with Lie bracket as defined in Lemma 5.2.2. Moreover, by Lemma 5.2.2 and its proof, we also have that L_K is solvable, and that $L'_K = K \otimes_F L'$. The natural inclusion

$$K \otimes \mathrm{gl}(L) \hookrightarrow \mathrm{gl}(L_K)$$

is an isomorphism of K-algebras. Let $a, b, c \in K$ and $x, y, z \in L$. Then

$$\begin{aligned}
\big(ab \otimes \operatorname{ad}(x)\operatorname{ad}(y)\big)(c \otimes z) &= abc \otimes \big(\operatorname{ad}(x)\operatorname{ad}(y)\big)(z)\\
&= abc \otimes \operatorname{ad}(x)(\operatorname{ad}(y)z))\\
&= abc \otimes \operatorname{ad}(x)([y, z])\\
&= abc \otimes [x, [y, z]].
\end{aligned}$$

And

$$\begin{aligned}
\big(\operatorname{ad}(a \otimes x)\operatorname{ad}(b \otimes y)\big)(c \otimes z) &= \operatorname{ad}(a \otimes x)\big(\operatorname{ad}(b \otimes y)(c \otimes z)\big)\\
&= \operatorname{ad}(a \otimes x)\big([b \otimes y, c \otimes z]\big)\\
&= [a \otimes x, [b \otimes y, c \otimes z]]\\
&= [a \otimes x, bc \otimes [y, z]]\\
&= abc \otimes [x, [y, z]].
\end{aligned}$$

It follows that

$$ab \otimes \operatorname{ad}(x)\operatorname{ad}(y) = \operatorname{ad}(a \otimes x)\operatorname{ad}(b \otimes y).$$

The diagram

$$\begin{array}{ccc}
K \otimes \mathrm{gl}(L) & \xrightarrow{\sim} & \mathrm{gl}(L_K)\\
{\scriptstyle \mathrm{id}\otimes\mathrm{tr}}\big\downarrow & & {\scriptstyle \mathrm{tr}}\big\downarrow\\
K & \xrightarrow{\mathrm{id}} & K
\end{array}$$

commutes. Hence, we obtain

$$ab \cdot \operatorname{tr}(\operatorname{ad}(x)\operatorname{ad}(y)) = \operatorname{tr}\big(\operatorname{ad}(a \otimes x)\operatorname{ad}(b \otimes y)\big).$$

It follows that if $\mathrm{tr}(\mathrm{ad}(w)\mathrm{ad}(z)) = 0$ for all $w \in L'_K$ and $z \in L_K$, then $\mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y)) = 0$ for all $x \in L'$ and $y \in L$. Thus, we may assume that F is algebraically closed.

Next, by Lemma 2.1.5, the Lie algebra $\mathrm{ad}(L) \subset \mathrm{gl}(L)$ is solvable. By Lie's Theorem, Theorem 3.1.2, there exists a basis for L so that in this basis all the elements of $\mathrm{ad}(L)$ are upper triangular; fix such a basis for L, and write the elements of $\mathrm{gl}(L)$ as matrices with respect to this basis. Let $x_1, x_2 \in L$. Then

$$\mathrm{ad}([x_1, x_2]) = [\mathrm{ad}(x_1), \mathrm{ad}(x_2)].$$

Since $\mathrm{ad}(x_1)$ and $\mathrm{ad}(x_2)$ are upper triangular, a calculation shows that the upper triangular matrix $[\mathrm{ad}(x_1), \mathrm{ad}(x_2)]$ is strictly upper triangular. This implies that all the elements of $\mathrm{ad}(L')$ are strictly upper triangular matrices. Another calculation now shows that $\mathrm{ad}(x)\mathrm{ad}(y)$ is strictly upper triangular for $x \in L'$ and $y \in L$; therefore, $\mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y)) = 0$ for $x \in L'$ and $y \in L$.

Now assume that $\mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y)) = 0$ for $x \in L'$ and $y \in L$. Consider $\mathrm{ad}(L)$. By Lemma 2.1.5, $\mathrm{ad}(L') = \mathrm{ad}(L)'$. Therefore, our hypothesis and Lemma 5.2.3 imply that $\mathrm{ad}(L)$ is solvable. Now $\mathrm{ad}(L) \cong L/Z(L)$ as Lie algebras. Hence, $L/Z(L)$ is solvable. Since $Z(L)$ is solvable, we conclude from Lemma 2.1.7 that L is solvable. □

5.3 Cartan's second criterion: semi-simplicity

Let L be a finite-dimensional Lie algebra over F. Define

$$\kappa : L \times L \longrightarrow F$$

by

$$\kappa(x, y) = \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y))$$

for $x, y \in L$. We refer to κ as the **Killing form** on L.

Proposition 5.3.1. *Let L be a finite-dimensional Lie algebra over F. The Killing form on L is a symmetric bilinear form. Moreover, we have*

$$\kappa([x, y], z) = \kappa(x, [y, z])$$

for $x, y, z \in L$.

Proof. The linearity of ad and tr imply that *kappa* is bilinear. The Killing form is symmetric because in general $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ for A and B linear operators on a finite-dimensional vector space. Finally, let $x, y, z \in L$. Then

$$\begin{aligned}
\kappa([x, y], z) &= \mathrm{tr}(\mathrm{ad}([x, y])\mathrm{ad}(z)) \\
&= \mathrm{tr}([\mathrm{ad}(x), \mathrm{ad}(y)]\mathrm{ad}(z)) \\
&= \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y)\mathrm{ad}(z)) - \mathrm{tr}(\mathrm{ad}(y)\mathrm{ad}(x)\mathrm{ad}(z)) \\
&= \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y)\mathrm{ad}(z)) - \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(z)\mathrm{ad}(y))
\end{aligned}$$

$$= \mathrm{tr}(\mathrm{ad}(x)[\mathrm{ad}(y), \mathrm{ad}(z)])$$
$$= \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}([y, z]))$$
$$= \kappa(x, [y, z]).$$

This completes the proof. □

Lemma 5.3.2. *Let L be a finite-dimensional Lie algebra over F. Let I be an ideal of L. Consider I as a Lie algebra over F, and let κ_I be the Killing form for I. We have $\kappa(x, y) = \kappa_I(x, y)$ for $x, y \in I$.*

Proof. Fix a F-vector space basis for I, and extend this to a basis for L. Let $x \in I$. Then because I is an ideal, we have $\mathrm{ad}(x)L \subset I$. It follows that the matrix of $\mathrm{ad}(x)$ in our basis for L has the form

$$\mathrm{ad}(x) = \begin{bmatrix} M(x) & * \\ 0 & 0 \end{bmatrix}$$

where $M(x)$ is the matrix of $\mathrm{ad}(x)|_I$ in our chosen basis for I. Let $y \in I$. Then

$$\begin{aligned} \kappa_I(x, y) &= \mathrm{tr}(\mathrm{ad}(x)|_I \mathrm{ad}(y)|_I) \\ &= \mathrm{tr}(M(x)M(y)) \\ &= \mathrm{tr}(\begin{bmatrix} M(x) & * \\ 0 & 0 \end{bmatrix} \begin{bmatrix} M(y) & * \\ 0 & 0 \end{bmatrix}) \\ &= \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y)) \\ &= \kappa(x, y). \end{aligned}$$

This completes the proof. □

Lemma 5.3.3. *Let L be a finite-dimensional Lie algebra over F. Let I be an ideal of L. Define*

$$I^\perp = \{x \in L : \kappa(x, I) = 0\}.$$

Then $I^\perp$ is an ideal of L.

Proof. It is evident that $I^\perp$ is an F-subspace of L. Let $x \in L$, $y \in I^\perp$ and $z \in I$. Then

$$\kappa([x, y], z) = \kappa(x, [y, z]) = \kappa(x, 0) = 0.$$

It follows that $[x, y] \in I^\perp$, as required. □

Let V be an F-vector space and let $b : V \times V \to F$ be a symmetric bilinear form. We say that b is **non-degenerate** if, for all $x \in V$, if $b(x, y) = 0$ for all $y \in V$, then $x = 0$. Let L be a finite-dimensional Lie algebra over F. Evidently, $L^\perp = 0$ if and only if the Killing form on L is non-degenerate.

Theorem 5.3.4 (Cartan's Second Criterion)**.** *Assume that F has characteristic zero. Let L be a finite-dimensional Lie algebra over F. The Lie algebra L is semi-simple if and only if the Killing form on L is non-degenerate.*

Proof. Assume that L is semi-simple. We need to prove that $L^\perp = 0$. Set $I = L^\perp$. By the definition of I, we have $\kappa(I, L) = 0$. This implies that $\kappa(I, I') = 0$. By Lemma 5.3.2 we get $\kappa_I(I, I') = 0$. By Theorem 5.2.4, Cartan's first criterion, the Lie algebra I is solvable. Since L is semi-simple by assumption, we must have $I = 0$, as required.

Now assume that the Killing form on L is non-degenerate. Assume that L is not semi-simple; we will obtain a contradiction. By definition, since L is not semi-simple, L contains a non-zero solvable ideal I. Consider the sequence $I^{(k)}$ for $k = 0, 1, 2, \ldots$. Each element of the sequence is an ideal of L; also, since I is solvable, there exists a non-negative integer such that $I^{(k)} \neq 0$ and $I^{(k+1)} = 0$. Set $A = I^{(k)}$. Then A is a non-zero ideal of L, and A is abelian. Let $x \in L$ and $a \in A$. Let $y \in L$. Then

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

Since A is an ideal of L we have $[a, y] \in A$, and hence also $[x, [a, y]] \in A$. Since A is abelian, this implies that $[a, [x, [a, y]]] = 0$. It follows that $\mathrm{ad}(a)\mathrm{ad}(x)\mathrm{ad}(a) = 0$ and thus $(\mathrm{ad}(x)\mathrm{ad}(a))^2 = 0$. Since nilpotent operators have trivial traces, we obtain $\mathrm{tr}(\mathrm{ad}(a)\mathrm{ad}(x)) = 0$. Thus, $\kappa(a, x) = 0$. Because $x \in L$ was arbitrary, we have $a \in L^\perp = 0$. Thus, $A = 0$, a contradiction. $\square$

5.4 Simple Lie algebras

Lemma 5.4.1. *Let V be a finite-dimensional F-vector space and let b be a symmetric bilinear form on V. Let W be a subspace of V. Then*

$$\dim W + \dim W^\perp \geq \dim V.$$

If b is non-degenerate, then

$$\dim W + \dim W^\perp = \dim V.$$

Proof. Let $V^\vee$ be the dual space of V, i.e., $V^\vee = \mathrm{Hom}_F(V, F)$. Define

$$V \longrightarrow V^\vee$$

by $v \mapsto \lambda_v$, where λ_v is defined by $\lambda_v(x) = b(x, v)$ for $x \in V$. Let $V^\vee \to W^\vee$ be the restriction map, i.e., defined by $\lambda \mapsto \lambda|_W$ for $\lambda \in V^\vee$. This restriction map is surjective. Consider the composition

$$V \xrightarrow{\sim} V^\vee \longrightarrow W^\vee.$$

The kernel of this linear map is $W^\perp$. It follows that $\dim V - \dim W^\perp \leq \dim W^\vee = \dim W$, i.e., $\dim V \leq \dim W + \dim W^\perp$.

Assume that b is non-degenerate. Then the map $V \to V^\vee$ is injective; since V and $V^\vee$ have the same finite dimension, this map is an isomorphism. It follows that the above composition is surjective. Hence $\dim W + \dim W^\perp = \dim V$. □

Let L be a Lie algebra over F. Let $L_1, \dots, L_t$ be Lie subalgebras of L. We say that L is the **direct sum** of $L_1, \dots, L_t$ if $L = L_1 \oplus \cdots \oplus L_t$ as vector spaces and

$$[x_1 + \cdots + x_t, y_1 + \cdots + y_t] = [x_1, y_1] + \cdots + [x_t, y_t]$$

for $x_i, y_i \in L_i$, $i \in \{1, \dots, t\}$.

Lemma 5.4.2. *Let L be a Lie algebra over F. Let $I_1, \dots, I_t$ be ideals of L. If L is the direct sum of $I_1, \dots, I_t$ as vector spaces, then L is the direct sum of $I_1, \dots, I_t$ as Lie algebras.*

Proof. Assume L is the direct sum of $I_1, \dots, I_t$ as vector spaces. To prove that L is the direct sum of $I_1, \dots, I_t$ as Lie algebras, it will suffice to prove that $[x, y] = 0$ for $x \in I_i$ and $y \in I_j$ for $i, j \in \{1, \dots, t\}$. Let $i, j \in \{1, \dots, t\}$, $x \in I_i$ and $y \in I_j$. Then $[x, y] \in I_i \cap I_j$ because I_i and I_j are ideals. Since $I_i \cap I_j = 0$ we have $[x, y] = 0$. □

Lemma 5.4.3. *Assume that F has characteristic zero. Let L be a semi-simple finite-dimensional Lie algebra over F. Let I be a non-zero proper ideal of L. Then $L = I \oplus I^\perp$ and I is a semi-simple Lie algebra over F.*

Proof. By Lemma 5.4.1 and Lemma 5.4.2, to prove that $L = I \oplus I^\perp$ it will suffice to prove that $I \cap I^\perp = 0$. Let $J = I \cap I^\perp$. Then J is an ideal of L. By Lemma 5.3.2, we have $\kappa_J(J, J) = 0$. In particular, $\kappa_J(J, J') = 0$. By Theorem 5.2.4, Cartan's first criterion, the Lie algebra J is solvable. Since L is semi-simple, we get $J = 0$, as desired.

By Theorem 5.3.4, Cartan's second criterion, to prove that I is semi-simple, it will suffice to prove that if $x \in I$ and $\kappa_I(x, y) = 0$ for all $y \in I$, then $x = 0$. Assume that $x \in I$ is such that $\kappa_I(x, y) = 0$ for all $y \in I$. By Lemma 5.3.2, $\kappa(x, y) = 0$ for all $y \in I$. Let $z \in L$. By the first paragraph, we may write $z = z_1 + z_2$ with $z_1 \in I$ and $z_2 \in I^\perp$. We have $\kappa(x, z) = \kappa(x, z_1) + \kappa(x, z_2)$. Now $\kappa(x, z_1) = 0$ because $z_1 \in I$ and the assumption on x, and $\kappa(x, z_2) = 0$ because $x \in I$ and $z_2 \in I^\perp$. It follows that $\kappa(x, z) = 0$. Since $z \in L$ was arbitrary, we obtain $x \in L^\perp$. By Theorem 5.3.4, Cartan's second criterion, $L^\perp = 0$. Hence, $x = 0$. □

Let L be a Lie algebra over F. We say that L is **simple** if L is not abelian and the only ideals of L are 0 and L. From the definition, we see that a simple Lie algebra is non-zero.

Lemma 5.4.4. *Let L be a Lie algebra over F. If L is simple, then L is semi-simple.*

Proof. Assume that L is simple. Since L is simple we must have $\mathrm{rad}(L) = 0$ or $\mathrm{rad}(L) = L$. If $\mathrm{rad}(L) = 0$, then L is semi-simple by definition. Assume that $\mathrm{rad}(L) = L$; we will obtain a contradiction. Then L is solvable. By the definition of solvability, and since $L \neq 0$, there exists a non-negative integer k such that $L^{(k)} \neq 0$ and $L^{(k+1)} = 0$. Since $L^{(k)}$ is a non-zero ideal of L we must have $L^{(k)} = L$. Since $L^{(k)}$ is abelian, L is abelian, a contradiction. □

Let L be a Lie algebra over F. Let I be an F-subspace of L. We say that I is a **simple ideal** of L if I is an ideal of L and I is simple as a Lie algebra over F.

Theorem 5.4.5. *Assume that F has characteristic zero. Let L be a finite-dimensional Lie algebra over F. The Lie algebra L is semi-simple if and only if there exist simple ideals $I_, \dots, I_t$ of L such that*

$$I = I_1 \oplus \cdots \oplus I_t.$$

Proof. Via induction on $\dim L$, we will prove the assertion that if L is semi-simple, then there exist simple ideals of L as in the theorem. The assertion is trivially true when $\dim L = 0$, because in this case L cannot be semi-simple. Assume that the assertion holds for all Lie algebras over F with dimension less than $\dim L$; we will prove the assertion for L. Assume that L is semi-simple. Let I be an ideal of L with the smallest possible non-zero dimension. Assume that $\dim I = \dim L$, i.e., $I = L$. Then certainly L has no ideals other than 0 and L. Moreover, L is not abelian because $\mathrm{rad}(L) = 0$. It follows that L is simple. Assume that $\dim I < \dim L$. By Lemma 5.4.3 we have $L = I \oplus I^\perp$, and I and $I^\perp$ are semi-simple Lie algebras over F with $\dim I < \dim L$ and $\dim I^\perp < \dim L$. By induction, there exist simple ideals $I_1, \dots, I_r$ of I and simple ideals $J_1, \dots, J_s$ of $I^\perp$ such that

$$I = I_1 \oplus \cdots \oplus I_r \quad \text{and} \quad I^\perp = J_1 \oplus \cdots \oplus J_s.$$

We have

$$L = I_1 \oplus \cdots \oplus I_r \oplus J_1 \oplus \cdots \oplus J_s$$

as F-vector spaces. It is easy to check that $I_1, \dots, I_r, J_1, \dots, J_s$ are ideals of L. The assertion follows now by induction.

Next, assume that there exist simple ideals of L as in the statement of the theorem. Let $x, y, z \in L$. Write $x = x_1 + \cdots + x_t$, $y = y_1 + \cdots + y_t$, and $z = z_1 + \cdots + z_t$ with $x_i, y_i, z_i \in I_i$ for $i \in \{1, \dots, t\}$. We have

$$\begin{aligned}(\mathrm{ad}(x)\mathrm{ad}(y))(z) &= [x, [y, z]] \\ &= \sum_{i=1}^{t} \sum_{j=1}^{t} \sum_{k=1}^{t} [x_i, [y_j, z_k]] \\ &= \sum_{i=1}^{t} [x_i, [y_i, z_i]]\end{aligned}$$

$$= \sum_{i=1}^{t} (\mathrm{ad}(x_i)\mathrm{ad}(y_i))(z_i).$$

It follows that

$$\mathrm{ad}(x)\mathrm{ad}(y) = \begin{bmatrix} \mathrm{ad}(x_1)\mathrm{ad}(y_1) & & \\ & \ddots & \\ & & \mathrm{ad}(x_t)\mathrm{ad}(y_t) \end{bmatrix}.$$

Hence, using Lemma 5.3.2,

$$\kappa(x,y) = \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y)) = \sum_{i=1}^{t} \mathrm{tr}(\mathrm{ad}(x_i)\mathrm{ad}(y_i)) = \sum_{i=1}^{t} \kappa_{I_i}(x_i, y_i).$$

By Theorem 5.3.4, Cartan's second criterion, to prove that L is semi-simple it suffices to prove that $L^\perp = 0$. Let $x \in L^\perp$. Let $i \in \{1, \dots, t\}$ and $y \in I_i$. Write $x = x_1 + \cdots + x_t$ with $x_j \in I_j$ for $j \in \{1, \dots, t\}$. By the above general calculation we have $0 = \kappa(x, y) = \kappa_{I_i}(x_i, y_i)$. Since I_i is semi-simple by Lemma 5.4.4, by Theorem 5.3.4, Cartan's second criterion applied to I_i, we must have $x_i = 0$. It follows that $x = 0$. □

5.5 Jordan decomposition

Let R be an F-algebra; we *do not* assume that R is associative. We recall from Proposition 1.4.4 the Lie algebra $\mathrm{Der}(R)$ of derivations on R, i.e., the Lie subalgebra of $\mathrm{gl}(R)$ consisting of the linear maps $D : R \to R$ such that

$$D(ab) = aD(b) + D(a)b$$

for $a, b \in R$.

Proposition 5.5.1. *Let F be a field of characteristic zero. Let L be a semi-simple finite-dimensional Lie algebra over F. Then the* ad *homomorphism is an isomorphism of L onto* $\mathrm{Der}(L)$*:*

$$\mathrm{ad} : L \xrightarrow{\sim} \mathrm{Der}(L).$$

Proof. By Proposition 1.4.4, the kernel of ad is $Z(L)$. Since L is semi-simple, we have $Z(L) = 0$, so that ad is injective. Set $K = \mathrm{ad}(L)$. Because ad is injective, K is isomorphic to L, and is hence also semi-simple.

By Proposition 1.4.4 we have $K \subset \mathrm{Der}(L)$; we need to prove that $K = \mathrm{Der}(L)$. We first prove that K is an ideal of $\mathrm{Der}(L)$. Let $x \in K$ and $D \in \mathrm{Der}(L)$. Let $y \in L$. Then

$$\begin{aligned} \big([D, \mathrm{ad}(x)]\big)(y) &= \big(D\mathrm{ad}(x) - \mathrm{ad}(x)D\big)(y) \\ &= D\big(\mathrm{ad}(x)(y)\big) - \mathrm{ad}(x)\big(D(y)\big) \end{aligned}$$

$$\begin{aligned}
&= D([x, y]) - [x, D(y)] \\
&= [x, D(y)] + [D(x), y] - [x, D(y)] \\
&= [D(x), y] \\
&= \operatorname{ad}\big(D(x)\big)(y).
\end{aligned}$$

This implies that

$$[D, \operatorname{ad}(x)] = \operatorname{ad}\big(D(x)\big),$$

so that $[D, \operatorname{ad}(x)] \in K$. Next, using the Killing form on $\operatorname{Der}(L)$, define as usual

$$K^\perp = \{D \in \operatorname{Der}(L) : \kappa_{\operatorname{Der}(L)}(D, K) = 0\}.$$

By Lemma 5.3.3, $K^\perp$ is also an ideal of $\operatorname{Der}(L)$. Let $x \in K \cap K^\perp$. Then

$$0 = \kappa_{\operatorname{Der}(L)}(x, K) = \kappa_K(x, K),$$

where the last equality follows from Lemma 5.3.2. Since K is semi-simple we must have $x = 0$ by Theorem 5.3.4, Cartan's second criterion. Therefore, $K \cap K^\perp = 0$. Now since K and $K^\perp$ are both ideals of $\operatorname{Der}(L)$ we have $[K, K^\perp] \subset K$ and $[K, K^\perp] \subset K^\perp$, so that $[K, K^\perp] \subset K \cap K^\perp$. Thus, $[K, K^\perp] = 0$. Let $D \in K^\perp$ and $x \in L$. Then $[D, \operatorname{ad}(x)] = 0$. From above, we also have $[D, \operatorname{ad}(x)] = \operatorname{ad}\big(D(x)\big)$. Therefore, $\operatorname{ad}\big(D(x)\big) = 0$. Since ad is injective, we get $D(x) = 0$. Since $x \in L$ was arbitrary, we obtain $D = 0$. Thus, $K^\perp = 0$. Now by Lemma 5.4.1 we have $\dim K + \dim K^\perp \geq \dim \operatorname{Der}(L)$; therefore, $\dim K = \dim \operatorname{Der}(L)$ so that $K = \operatorname{Der}(L)$. □

We recall the following theorem from linear algebra.

Theorem 5.5.2 (Generalized eigenvalue decomposition)**.** *Assume that F has characteristic zero and is algebraically closed. Let V be a finite-dimensional vector space and let $T \in \operatorname{gl}(V)$. If $\lambda \in F$, then define $V_\lambda(T)$ to be the subset of $v \in V$ such that there exists a non-negative integer such that $(T - \lambda 1_V)^k v = 0$. For $\lambda \in F$, $V_\lambda(T)$ is an F-subspace of V that is mapped to itself by T. We have*

$$V = \bigoplus_{\lambda \in F} V_\lambda(T).$$

Factor the characteristic polynomial of T as

$$(X - \lambda_1)^{n_1} \cdots (X - \lambda_t)^{n_t}$$

where the $\lambda_i \in F$ are pairwise distinct for $i \in \{1, \dots, t\}$, and $n_1, \dots, n_t$ are positive integers such that $n_1 + \cdots + n_t = \dim V$. Define $E(T) = \{\lambda_1, \dots, \lambda_t\}$, the set of eigenvalues of T. For $\lambda \in F$ we have $V_\lambda(T) \neq 0$ if and only if $\lambda \in E(T)$, and $\dim V_{\lambda_i} = n_i$ for $i \in \{1, \dots, t\}$. Let $T = s + n$ be the Jordan-Chevalley decomposition of T, with s diagonalizable and n nilpotent. The set of eigenvalues for T is the same as the set of eigenvalues for s, and $V_\lambda(s) = V_\lambda(T)$ for $\lambda \in E(T) = E(s)$. Moreover, for every $\lambda \in E(T) = E(s)$, $V_\lambda(s)$ is the usual λ-eigenspace for s.

Lemma 5.5.3. *Let L be a Lie algebra over F. Let $D \in \mathrm{Der}(L)$. Let n be a non-negative integer. Let $\lambda, \mu \in F$ and $x, y \in L$. Then*

$$\big(D - (\lambda+\mu)1_L\big)^n([x,y]) = \sum_{k=0}^{n} \binom{n}{k} \big[(D-\lambda 1_L)^k x, (D-\mu 1_L)^{n-k} y\big].$$

Proof. We prove this by induction on n. The claim holds if $n = 0$. Assume it holds for n for all $x, y \in L$; we will prove that it holds for $n+1$ for all $x, y \in L$. Now

$$\begin{aligned}
&(D - (\lambda+\mu)1_L)^{n+1}[x,y] \\
&= (D - (\lambda+\mu)1_L)^n\big((D - (\lambda+\mu)1_L)[x,y]\big) \\
&= (D - (\lambda+\mu)1_L)^n\big(D[x,y] - (\lambda+\mu)[x,y]\big) \\
&= (D - (\lambda+\mu)1_L)^n\big([Dx,y] + [x,Dy] - (\lambda+\mu)[x,y]\big) \\
&= (D - (\lambda+\mu)1_L)^n\big([(D-\lambda 1_L)x,y] + [x,(D-\mu 1_L)y]\big) \\
&= \sum_{k=0}^{n} \binom{n}{k} \big[(D-\lambda 1_L)^{k+1} x, (D-\mu 1_L)^{n-k} y\big] \\
&\quad + \sum_{k=0}^{n} \binom{n}{k} \big[(D-\lambda 1_L)^{k} x, (D-\mu 1_L)^{n-k+1} y\big] \\
&= \sum_{k=0}^{n} \binom{n}{k} \big[(D-\lambda 1_L)^{k+1} x, (D-\mu 1_L)^{n+1-(k+1)} y\big] \\
&\quad + \sum_{k=0}^{n} \binom{n}{k} \big[(D-\lambda 1_L)^{k} x, (D-\mu 1_L)^{n+1-k} y\big] \\
&= \sum_{k=1}^{n+1} \binom{n}{k-1} \big[(D-\lambda 1_L)^{k} x, (D-\mu 1_L)^{n+1-k} y\big] \\
&\quad + \sum_{k=0}^{n} \binom{n}{k} \big[(D-\lambda 1_L)^{k} x, (D-\mu 1_L)^{n+1-k} y\big] \\
&= \sum_{k=1}^{n} (\binom{n}{k-1} + \binom{n}{k}) \big[(D-\lambda 1_L)^{k} x, (D-\mu 1_L)^{n+1-k} y\big] \\
&\quad + \big[(D-\lambda 1_L)^{n+1} x, (D-\mu 1_L)^{0} y\big] + \big[(D-\lambda 1_L)^{0} x, (D-\mu 1_L)^{n+1} y\big] \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} \big[(D-\lambda 1_L)^{k} x, (D-\mu 1_L)^{n+1-k} y\big].
\end{aligned}$$

This completes the proof. □

Lemma 5.5.4. *Assume that F has characteristic zero and is algebraically closed. Let L be a finite-dimensional Lie algebra over F. Let $D \in \mathrm{Der}(L)$, and let $D = S + N$ be the Jordan-Chevalley decomposition of D, with $S \in \mathrm{gl}(L)$ diagonalizable and $N \in \mathrm{gl}(L)$ nilpotent. Then S and N are contained in* $\mathrm{Der}(L)$.

Proof. Using the notation of Theorem 5.5.2, we have

$$L = \bigoplus_{\lambda \in F} L_\lambda(D).$$

Let $\lambda, \mu \in F$. We will first prove that

$$[L_\lambda(D), L_\mu(D)] \subset L_{\lambda+\mu}(D).$$

To prove this, let $x \in L_\lambda(D)$ and $y \in L_\mu(D)$. Let n be a positive even integer such that $(D - \lambda 1_L)^{n/2} x = 0$ and $(D - \mu 1_L)^{n/2} y = 0$. By Lemma 5.5.3 we have

$$\big(D - (\lambda + \mu) 1_L\big)^n ([x, y]) = \sum_{k=0}^{n} \binom{n}{k} \big[(D - \lambda 1_L)^k x, (D - \mu 1_L)^{n-k} y\big].$$

If $k \in \{0, \dots, n\}$, then $k \geq n/2$ or $n - k \geq n/2$. It follows that

$$\big(D - (\lambda + \mu) 1_L\big)^n ([x, y]) = 0$$

so that $[x, y] \in L_{\lambda+\mu}(D)$.

Now we prove that s is a derivation. We need to prove that $S([x, y]) = [S(x), y] + [x, S(y)]$ for $x, y \in L$. By linearity, it suffices to prove this for every $x \in L_\lambda(D)$ and $y \in L_\mu(D)$ for all $\lambda, \mu \in F$. Let $\lambda, \mu \in F$ and $x \in L_\lambda(D)$ and $y \in L_\mu(D)$. From Theorem 5.5.2, $L_\lambda(s) = L_\lambda(D)$, $L_\mu(D) = L_\mu(S)$, $L_{\lambda+\mu}(D) = L_{\lambda+\mu}(S)$ and on these three F-subspaces of L the operator σ acts by λ, μ, and $\lambda + \mu$, respectively. We have $[x, y] \in L_{\lambda+\mu}(D) = L_{\lambda+\mu}(S)$. Hence,

$$\begin{aligned} S([x, y]) &= (\lambda + \mu)[x, y] \\ &= [\lambda x, y] + [x, \mu y] \\ &= [S(x), y] + [x, S(y)]. \end{aligned}$$

It follows that S is a derivation. Since $N = D - S$, N is also a derivation. □

Theorem 5.5.5. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let $x \in L$. Then there exist unique elements $s, n \in L$ such that $x = s + n$, $\mathrm{ad}(s)$ is diagonalizable, $\mathrm{ad}(n)$ is nilpotent, and $[s, n] = 0$. Moreover, if $y \in L$ is such that $[x, y] = 0$, then $[s, y] = [n, y] = 0$.*

Proof. First we prove the existence of s and n. By Proposition 1.5.1 we have $\mathrm{ad}(x) \in \mathrm{Der}(L)$. Let $\mathrm{ad}(x) = S + N$ be the Jordan-Chevalley decomposition of $\mathrm{ad}(x)$ with S diagonalizable and N nilpotent. By Lemma 5.5.4, S and N are derivations. By Proposition 5.5.1, since L is semi-simple, there exist $s, n \in L$ such that $\mathrm{ad}(s) = S$ and $\mathrm{ad}(n) = N$. We have $\mathrm{ad}(x) = \mathrm{ad}(s + n)$. Since L is semi-simple, ad is injective; hence, $x = s + n$. Also, $\mathrm{ad}([s, n]) = [\mathrm{ad}(s), ad(n)] = [S, N] = 0$ because the operators S and N commute. Since ad is injective, we get $[s, n] = 0$. This proves the existence of s and n.

To prove uniqueness, assume that $s', n' \in L$ are such that $x = s' + n'$, $\mathrm{ad}(s')$ is diagonalizable, $\mathrm{ad}(n')$ is nilpotent, and $[s', n'] = 0$. Set $S' = \mathrm{ad}(s')$ and $N' = \mathrm{ad}(n')$. Then $\mathrm{ad}(x) = S' + N'$, S' is diagonalizable, N' is nilpotent, and S' and N' commute. By the uniqueness of the Jordan-Chevalley decomposition for $\mathrm{ad}(x)$ we get $\mathrm{ad}(s) = S = S' = \mathrm{ad}(s')$ and $\mathrm{ad}(n) = N = N' = \mathrm{ad}(n')$. Since ad is injective, $s = s'$ and $n = n'$.

Finally, assume that $y \in L$ is such that $[x, y] = 0$. Then $[\mathrm{ad}(x), \mathrm{ad}(y)] = 0$, i.e., $\mathrm{ad}(y)$ commutes with $\mathrm{ad}(x)$. By Theorem 5.1.1, there exists a polynomial $P(X) \in F[X]$ such that $S = P(\mathrm{ad}(x))$. Since $\mathrm{ad}(y)$ commutes with $\mathrm{ad}(x)$, we get $\mathrm{ad}(y)P(\mathrm{ad}(x)) = P(\mathrm{ad}(x))\mathrm{ad}(y)$. Hence, $\mathrm{ad}(y)$ commutes with S. Thus, $0 = [S, \mathrm{ad}(y)] = [\mathrm{ad}(s), \mathrm{ad}(y)] = \mathrm{ad}([s, y])$. By the injectivity of ad, we obtain $[s, y] = 0$. Similarly, $[n, y] = 0$. □

We refer to the decomposition $x = s+n$ from Theorem 5.5.5 as the **abstract Jordan decomposition** of x. We refer to s as the **semi-simple component** of x, and n as the **nilpotent component** of x.

Chapter 6

Weyl's theorem

6.1 The Casmir operator

Let L be a Lie algebra over F, let V be a finite-dimensional F-vector space, and let $\varphi : L \to \mathrm{gl}(V)$ be a representation. Define

$$\beta_V : L \times L \longrightarrow F$$

by

$$\beta_V(x, y) = \mathrm{tr}(\varphi(x)\varphi(y))$$

for $x, y \in L$.

Lemma 6.1.1. *Assume that F has characteristic zero. Let L be a semi-simple finite-dimensional Lie algebra over F, let V be a finite-dimensional F-vector space, and let $\varphi : L \to \mathrm{gl}(V)$ be a faithful representation. Then β_V is an associative and non-degenerate symmetric bilinear form on L.*

Proof. It is clear that β_V is a symmetric bilinear form. To see that β_V is associative, let $x, y, z \in L$. Then

$$\begin{aligned}
\beta_V([x, y], z) &= \mathrm{tr}(\varphi([x, y])\varphi(z)) \\
&= \mathrm{tr}([\varphi(x), \varphi(y)]\varphi(z)) \\
&= \mathrm{tr}(\varphi(x)\varphi(y)\varphi(z)) - \mathrm{tr}(\varphi(y)\varphi(x)\varphi(z)) \\
&= \mathrm{tr}(\varphi(x)\varphi(y)\varphi(z)) - \mathrm{tr}(\varphi(x)\varphi(z)\varphi(y)) \\
&= \mathrm{tr}(\varphi(x)[\varphi(y), \varphi(z)]) \\
&= \mathrm{tr}(\varphi(x)\varphi([y, z])) \\
&= \beta_V(, x, [y, z]).
\end{aligned}$$

Next, let

$$I = \{x \in L : \beta_V(x, L) = 0\}.$$

To prove that β_V is non-degenerate it will suffice to prove that $I = 0$. We claim that I is an ideal of L. Let $x \in I$ and $y, z \in L$. Then $\beta_V([x, y], z) =$

$\beta_V(x, [y, z]) = 0$. This proves that $[x, y] \in I$, so that I is an ideal of L. Since L is semi-simple, to prove that $I = 0$ it will now suffice to prove that I is solvable. Consider $J = \varphi(I)$. Since φ is faithful, $I \cong J$; thus, it suffices to prove that J is solvable. Now by the definition of I we have $\mathrm{tr}(xy) = 0$ for all $x \in J$ and $y \in \varphi(L)$; in particular, we have $\mathrm{tr}(xy) = 0$ for all $x, y \in J$. By Lemma 5.2.3, the Lie algebra J is solvable. □

Let the notation be as in the statement of Lemma 6.1.1. Since the symmetric bilinear form β_V is non-degenerate, if $x_1, \ldots, x_n$ is an ordered basis for L, then there exists a unique ordered basis $x'_1, \ldots, x'_n$ for L such that

$$\beta_V(x_i, x'_j) = \delta_{ij}$$

for $i, j \in \{1, \ldots, n\}$. We refer to $x'_1, \ldots, x'_n$ as the basis dual to $x_1, \ldots, x_n$ with respect to β_V.

Lemma 6.1.2. *Assume that F has characteristic zero. Let L be a semi-simple finite-dimensional Lie algebra over F, let V be a finite-dimensional F-vector space, and let $\varphi : L \to \mathrm{gl}(V)$ be a faithful representation. Let $x_1, \ldots, x_n$ be an ordered basis for L, with dual basis $x'_1, \ldots, x'_n$ defined with respect to β_V. Define*

$$C = \sum_{i=1}^{n} \varphi(x_i)\varphi(x'_i).$$

Then $C \in \mathrm{gl}(V)$, the definition of C does not depend on the choice of ordered basis for L, and $C\varphi(x) = \varphi(x)C$ for $x \in L$. Moroever, $\mathrm{tr}(C) = \dim L$. We refer to C as the ***Casmir operator*** *for φ.*

Proof. To show that the definition of C does not depend on the choice of basis, let $y_1, \ldots, y_n$ be another ordered basis for L. Let $(m_{ij}) \in \mathrm{GL}(n, F)$ be the matrix such that

$$y_i = \sum_{j=1}^{n} m_{ij} x_j$$

and let $(n_{ij}) \in \mathrm{GL}(n, F)$ be the matrix such that

$$x_i = \sum_{j=1}^{n} n_{ij} y_j$$

for $i \in \{1, \ldots, n\}$. We have

$$\delta_{ij} = \sum_{l=1}^{n} m_{il} n_{lj}, \qquad \delta_{ij} = \sum_{l=1}^{n} n_{il} m_{lj}$$

for $i, j \in \{1, \ldots, n\}$. We have, for $i, j \in \{1, \ldots, n\}$,

$$\beta_V(y_i, \sum_{l=1}^{n} n_{lj} x'_l) = \sum_{l=1}^{n} n_{lj} \beta_V(y_i, x'_l)$$

$$
\begin{aligned}
&= \sum_{l=1}^{n} n_{lj}\beta_V(\sum_{k=1}^{n} m_{ik}x_k, x_l') \\
&= \sum_{l=1}^{n}\sum_{k=1}^{n} n_{lj}m_{ik}\beta_V(x_k, x_l') \\
&= \sum_{l=1}^{n}\sum_{k=1}^{n} n_{lj}m_{ik}\delta_{kl} \\
&= \sum_{l=1}^{n}\sum_{k=1}^{n} n_{lj}m_{il} \\
&= \delta_{ij}.
\end{aligned}
$$

It follows that

$$y_j' = \sum_{l=1}^{n} n_{lj}x_l'$$

for $j \in \{1, \dots, n\}$. Therefore,

$$
\begin{aligned}
\sum_{i=1}^{n} \varphi(y_i)\varphi(y_i') &= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n} m_{ij}n_{li}\varphi(x_j) \\
&= \sum_{j=1}^{n}\sum_{l=1}^{n}(\sum_{i=1}^{n} m_{ij}n_{li})\varphi(x_j)\varphi(x_l') \\
&= \sum_{j=1}^{n}\sum_{l=1}^{n} \delta_{lj}\varphi(x_j)\varphi(x_l') \\
&= \sum_{l=1}^{n} \varphi(x_l)\varphi(x_l').
\end{aligned}
$$

This proves that the definition of C does not depend on the choice of ordered basis for L.

Next, let $x \in L$. We need to prove that $C\varphi(x) = \varphi(x)C$. Let $(a_{jk}) \in \mathrm{M}(n, F)$ be such that

$$[x_j, x] = \sum_{k=1}^{n} a_{jk}x_k$$

for $j \in \{1, \dots, n\}$. We claim that

$$[x_j', x] = -\sum_{k=1}^{n} a_{kj}x_k'.$$

To see this, let $i \in \{1, \dots, n\}$. Then

$$\beta_V([x_j', x] + \sum_{k=1}^{n} a_{kj}x_k', x_i) = \beta_V([x_j', x], x_i) + \sum_{k=1}^{n} a_{kj}\beta_V(x_k', x_i)$$

$$
\begin{aligned}
&= \beta_V(x'_j, [x, x_i]) + a_{ij} \\
&= \beta_V(x'_j, -\sum_{l=1}^{n} a_{il} x_l) + a_{ij} \\
&= -\sum_{l=1}^{n} a_{il} \beta_V(x'_j, x_l) + a_{ij} \\
&= -a_{ij} + a_{ij} \\
&= 0.
\end{aligned}
$$

Since β_V is non-degenerate, we must have $[x'_j, x] = -\sum_{k=1}^{n} a_{kj} x'_k$. We now calculate:

$$
\begin{aligned}
&C\varphi(x) - \varphi(x)C = \sum_{j=1}^{n} \varphi(x_j)\varphi(x'_j)\varphi(x) - \varphi(x)\varphi(x_j)\varphi(x'_j) \\
&= \sum_{j=1}^{n} \varphi(x_j)\varphi(x'_j)\varphi(x) - \varphi(x_j)\varphi(x)\varphi(x'_j) + \varphi(x_j)\varphi(x)\varphi(x'_j) - \varphi(x)\varphi(x_j)\varphi(x'_j) \\
&= \sum_{j=1}^{n} \varphi(x_j)[\varphi(x'_j), \varphi(x)] + [\varphi(x_j), \varphi(x)]\varphi(x'_j) \\
&= \sum_{j=1}^{n} \varphi(x_j)\varphi([x'_j, x]) + \varphi([x_j, x])\varphi(x'_j) \\
&= -\sum_{j=1}^{n}\sum_{k=1}^{n} \Big(a_{kj}\varphi(x_j)\varphi(x'_k) + a_{jk}\varphi(x_k)\varphi(x'_j)\Big) \\
&= -\sum_{j=1}^{n}\sum_{k=1}^{n} a_{kj}\varphi(x_j)\varphi(x'_k) + \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}\varphi(x_k)\varphi(x'_j) \\
&= 0.
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
\mathrm{tr}(C) &= \mathrm{tr}(\sum_{i=1}^{n} \varphi(x_i)\varphi(x'_i)) \\
&= \sum_{i=1}^{n} \mathrm{tr}(\varphi(x_i)\varphi(x'_i)) \\
&= \sum_{i=1}^{n} \beta_V(x_i, x'_i) \\
&= \sum_{i=1}^{n} 1 \\
&= \dim L.
\end{aligned}
$$

This completes the proof. □

6.2 Proof of Weyl's theorem

Lemma 6.2.1. *Let L be a finite-dimensional semi-simple Lie algebra over F, and let I be an ideal of L. Then L/I is semi-simple.*

Proof. By Lemma 5.4.3, $I^\perp$ is also a Lie algebra over F, I and $I^\perp$ are semi-simple as Lie algebras over F, and $L = I \oplus I^\perp$ as Lie algebras. We have $L/I \cong I^\perp$ as Lie algebras; it follows that L/I is semi-simple. □

Lemma 6.2.2. *Let L be a finite-dimensional semi-simple Lie algebra over F. Then $L = L' = [L, L]$.*

Proof. By Theorem 5.4.5, there exist simple ideals $I_1, \dots, I_t$ of L such that $L = I_1 \oplus \cdots \oplus I_t$ as Lie algebras. We have $[L, L] = [I_1, I_1] \oplus \cdots \oplus [I_t, I_t]$. For each $i \in \{1, \dots, t\}$, I_i is not abelian so that $[I_i, I_i]$ is non-zero; this implies that $[I_i, I_i] = I_i$. Hence, $[L, L] = L$. □

Lemma 6.2.3. *Let L be a Lie algebra over F, and let V and W be L-modules. Let*

$$M = \mathrm{Hom}(V, W)$$

be the F-vector space of all F-linear maps from V to W. For $x \in L$ and $T \in M$ define $x \cdot T : V \to W$ by

$$(x \cdot T)(v) = x \cdot T(v) - T(x \cdot v)$$

for $v \in V$. With this definition, M is an L-module. Moreover, the following statements hold:

1. *The F-subspace of $T \in M$ such that $x \cdot T = 0$ for all $x \in L$ is $\mathrm{Hom}_L(V, W)$, the F-vector space of all L-maps from V to W.*

2. *If W is an L-submodule of V, then the F-subspaces*

$$M_1 = \{T \in \mathrm{Hom}(V, W) : f|_W \text{ is a constant}\}$$

and

$$M_0 = \{T \in \mathrm{Hom}(V, W) : f|_W = 0\}$$

are L subspaces of M with $M_0 \subset M_1$ and the action of L maps M_1 into M_0.

Proof. Let $x, y \in L$, $T \in M$, and $v \in V$. Then

$$\begin{aligned}([x, y] \cdot T)(v) &= [x, y] \cdot T(v) - T([x, y] \cdot v) \\ &= x(yT(v)) - y(xT(v)) - T(x(yv)) + T(y(xv))\end{aligned}$$

and

$$\big(x(yT) - y(xT)\big)(v)$$

$$\begin{aligned}
&= \big(x(yT)\big)(v) - \big(y(xT)\big)(v) \\
&= x((yT)(v)) - (yT)(xv) - y((xT)(v)) + (xT)(yv) \\
&= x(yT(v) - T(yv)) - y(T(xv)) + T(y(xv)) \\
&\qquad - y(xT(v) - T(xv)) + xT(yv) - T(x(yv)) \\
&= x(yT(v)) - xT(yv) - y(T(xv)) + T(y(xv)) \\
&\qquad - y(xT(v)) - yT(xv)) + xT(yv) - T(x(yv)) \\
&= x(yT(v)) + T(y(xv)) - y(xT(v)) - T(x(yv)).
\end{aligned}$$

It follows that

$$[x, y] \cdot T = x(yT) - y(xT)$$

so that with the above definition $\mathrm{Hom}(V, W)$ is an L-module.

The assertion 1 of the lemma is clear.

To prove the assertion 2, let $T \in M_1$ and let $a \in F$ be such that $T(w) = aw$ for $w \in W$. Let $x \in L$. Let $w \in W$. Then

$$\begin{aligned}
(xT)(w) &= xT(w) - T(xw) \\
&= axw - axw \\
&= 0.
\end{aligned}$$

The assertion 2 follows.

□

Theorem 6.2.4 (Weyl's Theorem)**.** *Let F be algebraically closed and have characteristic zero. Let L be a finite-dimensional semi-simple Lie algebra over F. If (φ, V) is a finite-dimensional representation of L, then V is a direct sum of irreducible representations of L.*

Proof. By induction, to prove the theorem it will suffice to prove that if W is a proper, non-zero L-subspace of V, then W has a complement, i.e., there exists an L-subspace W' of V such that $V = W \oplus W'$. Let W be a proper, non-zero L-subspace of V.

We first claim that W has a complement in the case that $\dim W = \dim V - 1$. Assume that $\dim W = \dim V - 1$.

We will first prove our claim when W is irreducible; assume that W is irreducible. The kernel $\ker(\varphi)$ of $\varphi : L \to \mathrm{gl}(V)$ is an ideal of L. By Lemma 6.2.1 the Lie algebra $L/\ker(\varphi)$ is semi-simple. By replacing $\varphi : L \to \mathrm{gl}(V)$ by the representation $\varphi : L/\ker(\varphi) \to \mathrm{gl}(V)$, we may assume that φ is faithful. Consider the quotient V/W. By assumption, this is a one-dimensional L-module. Since $[L, L]$ acts by zero on any one-dimensional L-module, and since $L = [L, L]$ by Lemma 6.2.2, it follows that L acts by zero on V/W. This implies that $\varphi(L)V \subset W$. In particular, if C is the Casmir operator for φ, then $CV \subset W$. By Lemma 6.1.2, C is an L-map. Hence, $\ker(C)$ is an L-submodule of V; we will prove that $V = W \oplus \ker(C)$, so that $\ker(C)$ is a complement to W. To prove that $\ker(C)$ is a complement to W it will suffice to prove that $W \cap \ker(C) = 0$ and $\dim \ker(C) = 1$. Consider the restriction $C|_W$ of C to W. This is an L-map from W to W. By Schur's Lemma, Theorem 4.2.2, since W is irreducible, there exists a constant $a \in F$ such that $C(w) = aw$ for $w \in W$. Fix an ordered basis $w_1, \dots, w_t$ for W, and let $v \notin V$. Then $w_1, \dots, w_t, v$ is an ordered basis for V, and the matrix of C in this basis has the form

$$\begin{bmatrix} a & & & * \\ & \ddots & & * \\ & & a & * \\ & & & 0 \end{bmatrix}.$$

It follows that $\mathrm{tr}(C) = (\dim W)a$. On the other hand, by Lemma 6.1.2, we have $\mathrm{tr}(C) = \dim L$. It follows that $(\dim W)a = \dim L$, and in particular, $a \neq 0$. Thus, C is injective on W and maps onto W. Therefore, $W \cap \ker(C) = 0$, and $\dim \ker(C) = \dim V - \dim \mathrm{im}(C) = \dim V - \dim W = 1$. This proves our claim in the case that W is irreducible.

We will now prove our claim by induction on $\dim V$. We cannot have $\dim V = 0$ or 1 because W is non-zero and proper by assumption. Suppose that $\dim V = 2$. Then $\dim W = 1$, so that W is irreducible, and the claim follows from the previous paragraph. Assume now that $\dim V \geq 3$, and that for all L-modules A with $\dim A < \dim V$, if B is an L-submodule of A of co-dimension one, then B has a complement. If W is irreducible, then W has a complement by the previous paragraph. Assume that W is not irreducible, and let W_1 be a L-submodule of W such that $0 < \dim W_1 < \dim W$. Consider the L-submodule W/W_1 of V/W_1. This L-submodule has co-dimension one in V/W_1, and $\dim V/W_1 < \dim V$. By the induction hypothesis, there exists an L-submodule U of V/W_1 such that

$$V/W_1 = U \oplus W/W_1.$$

We have $\dim U = 1$. Let $p : V \to V/W_1$ be the quotient map, and set $M = p^{-1}(U)$. Then M is an L-submodule of V, $W_1 \subset M$, and $M/W_1 = U$. We have

$$\dim M = \dim W_1 + \dim U = 1 + \dim W_1.$$

Since $\dim M = 1 + \dim W_1 < 1 + \dim W \leq \dim V$, we can apply the induction hypothesis again: let W_2 be an L-submodule of M that is a complement to W_1 in M, i.e.,

$$M = W_1 \oplus W_2.$$

We assert that W_2 is a complement to W in V, i.e., $V = W \oplus W_2$. Since $\dim W_2 = 1$, to prove this it suffices to prove that $W \cap W_2 = 0$. Assume that $w \in W \cap W_2$. Then

$$w + W_1 \in \big(W/W_1\big) \cap \big(M/W_1\big) = 0.$$

This implies that $w \in W_1$. Since now $w \in W_2 \cap W_1$, we have $w = 0$, as desired. The proof of our claim is complete.

Using the claim, we will now prove that W has a complement. Set

$$\begin{aligned} M &= \mathrm{Hom}(V,W), \\ M_1 &= \{T \in \mathrm{Hom}(V,W) : f|_W \text{ is multiplication by some constant}\}, \\ M_0 &= \{T \in \mathrm{Hom}(V,W) : f|_W = 0\}. \end{aligned}$$

By Lemma 6.2.3, M, M_1, M_0 are L-modules; clearly, $M_0 \subset M_1$. We claim that $\dim M_1/M_0 = 1$. To prove this, let $w \in W$ be non-zero. Define

$$M_1 \longrightarrow Fw$$

by $T \mapsto T(w)$. This is a well-defined F-linear map. Clearly, since $1_V \in M_1$, this map is surjective; also, the kernel of this map is M_0. It follows that $\dim M_1/M_0 = 1$. By the above claim, the L-submodule M_0 of M_1 has a complement M_1' in M_0, so that

$$M_1 = M_0 \oplus M_0'.$$

Since M_0' is one-dimensional, M_0' is spanned by a single element $T \in M_1$; we may assume that in fact $T(w) = w$ for $w \in W$. Moreover, since M_0' is one-dimensional the action of L on M_0' is trivial (see earlier in the proof for another example of this), so that $xT = 0$ for $x \in L$. The definition of the action of L on M implies that T is an L map. We now claim that

$$V = W \oplus \ker(T).$$

To see this, let $v \in V$. Then $v = T(v) + (v - T(v))$. Evidently, $T(v) \in W$. Also, $T(v - T(v)) = T(v) - T(T(v)) = T(v) - T(v) = 0$ because $T(v) \in W$, and the restriction of T to W is the identity. Thus, $V = W + \ker(T)$. Finally, suppose that $w \in W \cap \ker(T)$. Then $w = T(w)$ and $T(w) = 0$, so that $w = 0$. □

6.3 An application to the Jordan decomposition

Lemma 6.3.1. *Assume that F is algebraically closed and has characteristic zero. Let V be a finite-dimensional F-vector space. Let L be a Lie subalgebra of $\mathrm{gl}(V)$, and assume that L is semi-simple. If $x \in L$, and $x = x_s + x_n$ is the Jordan-Chevalley decomposition of x as an element of $\mathrm{gl}(V)$, then $x_s, x_n \in L$.*

Proof. We will first prove that $[x_s, L] \subset L$ and $[x_n, L] \subset L$. To see this, consider $\mathrm{ad}_{\mathrm{gl}(V)}(x) : \mathrm{gl}(V) \to \mathrm{gl}(V)$. This linear map has a Jordan-Chevalley decomposition $\mathrm{ad}_{\mathrm{gl}(V)}(x) = \mathrm{ad}_{\mathrm{gl}(V)}(x)_s + \mathrm{ad}_{\mathrm{gl}(V)}(x)_n$. Because $x \in L$, the linear map $\mathrm{ad}_{\mathrm{gl}(V)}(x)$ maps L into L (i.e., $[x, L] \subset L$). Because $\mathrm{ad}_{\mathrm{gl}(V)}(x)_s$ and $\mathrm{ad}_{\mathrm{gl}(V)}(x)_n$ are polynomials in $\mathrm{ad}_{\mathrm{gl}(V)}(x)$, these linear maps also map L into L. Now by Lemma 5.1.3 we have $\mathrm{ad}_{\mathrm{gl}(V)}(x)_s = \mathrm{ad}_{\mathrm{gl}(V)}(x_s)$ and $\mathrm{ad}_{\mathrm{gl}(V)}(x)_n = \mathrm{ad}_{\mathrm{gl}(V)}(x_n)$. It follows that $\mathrm{ad}_{\mathrm{gl}(V)}(x_s)$ and $\mathrm{ad}_{\mathrm{gl}(V)}(x_n)$ map L into L, i.e., $[x_s, L] \subset L$ and $[x_n, L] \subset L$.

Define

$$N = \{y \in \mathrm{gl}(V) : [y, L] \subset L\}.$$

Evidently, $L \subset N$; also, we just proved that $x_s, x_n \in N$. Moreover, we claim that N is a Lie subalgebra of $\mathrm{gl}(V)$, and that L is an ideal of N. To see that N is a Lie subalgebra of $\mathrm{gl}(V)$, let $y_1, y_2 \in N$. Let $z \in L$. Then

$$\begin{aligned}[][y_1, y_2], z] &= -[z, [y_1, y_2]] \\ &= [y_1, [y_2, z]] + [y_2, [z, y_1]].\end{aligned}$$

This is contained in L. Hence, $[z_1, z_2] \in N$. To see that L is an ideal of N, let $y \in N$ and $z \in L$; then $[y, z] \in L$ by the definition of N, which implies that L is an ideal of N.

Next, the Lie algebra L acts on V (since L consists of elements of $\mathrm{gl}(V)$). Let W be any L-submodule of V. Define

$$L_W = \{y \in \mathrm{gl}(V) : yW \subset W \text{ and } \mathrm{tr}(y|_W) = 0\}.$$

Evidently, L_W is a Lie subalgebra of $\mathrm{gl}(V)$. We claim that $L \subset L_W$, L is an ideal of L_W, and $x_s, x_n \in L_W$. Since L is semi-simple, we have by Lemma 6.2.2 that $L = [L, L]$. Thus, to prove that $L \subset L_W$, it will suffice to prove that $[a, b] \in L_W$ for $a, b \in L$. Let $a, b \in L$. Since W is an L-submodule of V, we have $[a, b]W \subset W$. Also,

$$\mathrm{tr}([a, b]|_W) = \mathrm{tr}(a|_W b|_W - b|_W a|_W) = \mathrm{tr}(a|_W b_W) - \mathrm{tr}(b|_W a|_W) = 0.$$

It follows that $L \subset L_W$. The argument that L is an ideal of L_W is similar. Next, since x maps W to W, x_s and x_n also map W to W. Since x_n is nilpotent, $x_n|_W$ is also nilpotent. Since $x_n|_W$ is nilpotent, $\mathrm{tr}(x_n|_W) = 0$. We have already proven that $\mathrm{tr}(x|_W) = 0$. Since $x|_W = x_s|_W + x_n|_W$, it follows that $\mathrm{tr}(x_s|_W) = 0$. Hence, $x_s, x_n \in L_W$.

Now define

$$A = \{y \in \mathrm{gl}(V) : [y, L] \subset L\} \cap \bigcap_{W \text{ is an } L\text{-submodule of } V} L_W.$$

By the last two paragraphs, A is a Lie subalgebra of $\mathrm{gl}(V)$, $L \subset A$, L is an ideal of A, and $x_s, x_n \in A$. We will prove that $A = L$, which will complete the proof since this implies that $x_s, x_n \in L$. We may regard A as an L-module via the action defined by $x \cdot a = \mathrm{ad}(x)a = [x, a]$ for $x \in L$ and $a \in A$. Evidently, with this action, L is an L-submodule of A. By Weyl's Theorem, Theorem 6.2.4, L admits a complement L_1 in A so that $A = L \oplus L_1$. We need to prove that the L-module L_1 is zero. We claim that $[L, L_1] = 0$, i.e., the action of L on L_1 is trivial. To see this we first note that $[L, L_1] \subset L_1$ because L_1 is an L-submodule. On the other hand, since L is an ideal of A, we have $[L, A] \subset L$; in particular, $[L, L_1] \subset L$. We now have $[L, L_1] \subset L \cap L_1 = 0$, proving that $[L, L_1] = 0$. Next, consider the action of L on V; by again Weyl's Theorem, Theorem 6.2.4, we can write

$$V = W_1 \oplus \cdots \oplus W_t$$

where W_i is an irreducible L-submodule of V for $i \in \{1, \dots, t\}$. Let $i \in \{1, \dots, t\}$. Let $y \in L_1$. Because $y \in A$ we have $y \in L_{W_i}$. Thus, $yW_i \subset W_i$. Moreover, since $[L, L_1] = 0$, the map $y|_{W_i}$ commutes with the action of L on W_i. By Schur's Lemma, Theorem 4.2.2, y acts by a scalar on W_i. Since we also have $\mathrm{tr}(y|_{W_i}) = 0$ because $y \in L_{W_i}$, it follows that $y|_{W_i} = 0$. We now conclude that $y = 0$, as desired. □

Theorem 6.3.2. *Assume that F is algebraically closed and has characteristic zero. Let V be a finite-dimensional F-vector space. Let L be a Lie subalgebra of $\mathrm{gl}(V)$, and assume that L is semi-simple. If $x \in L$, $x = x_s + x_n$ is the Jordan-Chevalley decomposition of x as an element of $\mathrm{gl}(V)$, and $x = s + n$ is the abstract Jordan decomposition of x, then $x_s = s$ and $x_n = n$.*

Proof. By Lemma 6.3.1 we have $x_s, x_n \in L$. By the uniqueness of the Jordan-Chevalley decomposition of elements of $\mathrm{gl}(L)$, to prove the theorem it will suffice to prove that $\mathrm{ad}_L(x) = \mathrm{ad}_L(x_s) + \mathrm{ad}_L(x_n)$, $\mathrm{ad}_L(x_s)$ is diagonalizable, $\mathrm{ad}_L(x_n)$ is nilpotent, and $[\mathrm{ad}_L(x_s), \mathrm{ad}_L(x_n)] = 0$, i.e., $\mathrm{ad}_L(x_s)$ and $\mathrm{ad}_L(x_n)$ commute. Since $x = x_s + x_n$ we have $\mathrm{ad}_L(x) = \mathrm{ad}_L(x_s) + \mathrm{ad}_L(x_n)$. From the involved definitions, is clear that $\mathrm{ad}_{\mathrm{gl}(V)}(x_s)|_L = \mathrm{ad}_L(x_s)$ and $\mathrm{ad}_{\mathrm{gl}(V)}(x_n)|_L = \mathrm{ad}_L(x_n)$. By Lemma 5.1.3, $\mathrm{ad}_{\mathrm{gl}(V)}(x_s)$ is diagonalizable and $\mathrm{ad}_{\mathrm{gl}(V)}(x_n)$ is nilpotent. This implies that $\mathrm{ad}_{\mathrm{gl}(V)}(x_s)|_L = \mathrm{ad}_L(x_s)$ is diagonalizable and $\mathrm{ad}_{\mathrm{gl}(V)}(x_n)|_L = \mathrm{ad}_L(x_n)$ is nilpotent. Finally, since $\mathrm{ad}_{\mathrm{gl}(V)}(x_s)$ and $\mathrm{ad}_{\mathrm{gl}(V)}(x_n)$ commute, $\mathrm{ad}_{\mathrm{gl}(V)}(x_s)|_L = \mathrm{ad}_L(x_s)$ and $\mathrm{ad}_{\mathrm{gl}(V)}(x_n)|_L = \mathrm{ad}_L(x_n)$ commute. □

Lemma 6.3.3. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let I be an ideal of L. The Lie algebra L/I is semi-simple. Let $x \in L$, and let $x = s+n$ be the abstract Jordan decomposition of x, as in Theorem 5.5.5. Then $x+I = (s+I)+(n+I)$ is the abstract Jordan decomposition of $x+I$, with $s+I$ and $n+I$ being the semi-simple and nilpotent components of $x+I$, respectively.*

Proof. By Lemma 6.2.1 L/I is semi-simple. Since $x = s + n$, we have $x + I = (s + I) + (n + I)$. Let $z \in L$. Let $y \in L$. We have

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

Similarly, if $P(X) \in F[X]$ is a polynomial, then

$$P(\mathrm{ad}(z+I))(y+I) = P(\mathrm{ad}(z))(y) + I.$$

Let $M(X)$ be the minimal polynomial of $\mathrm{ad}(s)$. Then

$$M(\mathrm{ad}(s+I))(y+I) = M(\mathrm{ad}(s))(y) + I = 0 + I = I.$$

Hence, $M(\mathrm{ad}(s+I)) = 0$, so that the minimal polynomial of $\mathrm{ad}(s+I)$ divides $M(X)$. Since s is diagonalizable, $M(X)$ has no repeated roots. Hence, the minimal polynomial of $\mathrm{ad}(s+I)$ has no repeated roots; this implies that $\mathrm{ad}(s+I)$ is diagonalizable. Similarly, since $\mathrm{ad}(n)$ is nilpotent, we see that $\mathrm{ad}(n+I)$ is nilpotent. Finally, we have $[s+I, n+I] = [s,n] + I = 0 + I = I$. □

Theorem 6.3.4. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let V be a finite-dimensional F-vector space, and let $\theta : L \to \mathrm{gl}(V)$ be a homomorphism. Let $x \in L$. Let $x = s + n$ be the abstract Jordan decomposition of x as in Theorem 5.5.5. Then the Jordan-Chevalley decomposition of $\theta(x) \in \mathrm{gl}(V)$ is given by $\theta(x) = \theta(s) + \theta(n)$, with $\theta(s)$ diagonalizable and $\theta(n)$ nilpotent.*

Proof. Set $J = \theta(L)$; this is a Lie subalgebra of $\mathrm{gl}(V)$. Since we have an isomorphism of Lie algebras

$$\theta : L/\ker(\theta) \xrightarrow{\sim} J$$

and since $L/\ker(\theta)$ is semi-simple by Lemma 6.2.1, it follows that J is semi-simple. Moreover, $x+\ker(\theta) = (s+\ker(\theta)) + (n+\ker(\theta))$ is the abstract Jordan decomposition of $x+\ker(\theta)$ by Lemma 6.3.3. Applying the above isomorphism, it follows that $\theta(x) = \theta(s) + \theta(n)$ is the abstract Jordan decomposition of $\theta(x)$ inside J. By Theorem 6.3.2, this is the Jordan-Chevalley decomposition of $\theta(x)$ inside $\mathrm{gl}(V)$. □

Chapter 7

The root space decomposition

Let F have characteristic zero and be algebraically closed. Let L be a finite-dimensional Lie algebra over F. Let H be a Lie subalgebra of L. We say that H is a **Cartan subalgebra** of L if H is non-zero; H is abelian; all the elements of H are semi-simple; and H is not properly contained in another abelian subalgebra of L, the elements of which are all semi-simple.

Theorem 7.0.1 (Second version of Engel's Theorem)**.** *Let L be a Lie algebra over F. Then L is nilpotent if and only if for all $x \in L$, the linear map $\mathrm{ad}(x) \in \mathrm{gl}(L)$ is nilpotent.*

Proof. Assume that L is nilpotent. By definition, this means that there exists a positive integer m such that $L^m = 0$. The definition of L^m implies that, in particular,

$$\underbrace{[x, [x, [x, \cdots, [x}_{m \text{ x's}}, y] \cdots]]]$$

for $x, y \in L$. This means that $\mathrm{ad}(x)^m = 0$. Thus, for every $x \in L$, the linear map $\mathrm{ad}(x)$ is nilpotent. Conversely, assume that for every $x \in L$, the linear map $\mathrm{ad}(x) \in \mathrm{gl}(L)$ is nilpotent. Consider the Lie subalgebra $\mathrm{ad}(L)$ of $\mathrm{gl}(L)$. By Theorem 3.1.1, the original version of Engel's Theorem, there exists a basis for L in which all the elements of $\mathrm{ad}(L)$ are strictly upper triangular; this implies that $\mathrm{ad}(L)$ is a nilpotent Lie algebra. By Proposition 2.2.1, since $\mathrm{ad}(L) \cong L/Z(L)$ is nilpotent, the Lie algebra L is also nilpotent.

□

Lemma 7.0.2. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Then L has a Cartan subalgebra.*

Proof. It will suffice to prove that L contains a non-zero abelian subalgebra consisting of semi-simple elements; to prove this, it will suffice to prove that L

contains a non-zero semi-simple element x (because the subalgebra Fx is non-zero, abelian and contains only semi-simple elements). Assume that L contains only nilpotent elements. Then by Theorem 7.0.1, the second version of Engel's Theorem, L is nilpotent, and hence solvable. This is a contradiction. □

Proposition 7.0.3. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let $H^\vee$ be $\mathrm{Hom}_F(H, F)$, the F-vector space of all F-linear maps from H to F. For $\alpha \in H^\vee$, define*

$$L_\alpha = \{x \in L : \mathrm{ad}(h)x = \alpha(h)x \text{ for all } h \in H\}.$$

Let Φ be the set of all $\alpha \in H^\vee$ such that $\alpha \neq 0$ and $L_\alpha \neq 0$. There is a direct sum decomposition

$$L = L_0 \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

Moreover:

1. *If $\alpha, \beta \in H^\vee$, then*
$$[L_\alpha, L_\beta] \subset L_{\alpha+\beta}.$$
2. *If $\alpha, \beta \in H^\vee$ and $\alpha + \beta \neq 0$, then*
$$\kappa(L_\alpha, L_\beta) = 0,$$
where κ is the Killing form on L.
3. *The restriction of the Killing form κ to L_0 is non-degenerate.*

Proof. Consider the F-vector space $\mathrm{ad}(H)$ of linear operators on L. Since every element of H is semi-simple, the elements of $\mathrm{ad}(H)$ are diagonalizable (recall the definition of the abstract Jordan decomposition, and in particular, the definition of semi-simple). Also, the linear operators in $\mathrm{ad}(H)$ mutually commute because H is abelian. It follows that the elements of $\mathrm{ad}(H)$ can be simultaneously diagonalized, i.e., the above decomposition holds.

To prove 1, let $\alpha, \beta \in H^\vee$. Let $x \in L_\alpha$ and $y \in L_\beta$. Let $h \in H$. Then

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

It follows that $[x, y] \in L_{\alpha+\beta}$.

To prove 2, let $\alpha, \beta \in H^\vee$ and assume that $\alpha + \beta \neq 0$. Let $x \in L_\alpha$, $y \in L_\beta$, and $h \in H$. Then

$$\alpha(h)\kappa(x,y) = \kappa(\alpha(h)x, y)$$

$$\begin{aligned}&= \kappa([h,x],y)\\ &= -\kappa([x,h],y)\\ &= -\kappa(x,[h,y])\\ &= -\kappa(x,\beta(h)y)\\ &= -\beta(h)\kappa(x,y).\end{aligned}$$

It follows that $(\alpha+\beta)(h)\kappa(x,y) = 0$. Since this holds for all $h \in H$ and $\alpha+\beta \neq 0$, it follows that $\kappa(x,y) = 0$. That is, $\kappa(L_\alpha, L_\beta) = 0$.

To prove 3, let $x \in L_0$. Assume that $\kappa(x,y) = 0$ for all $y \in L_0$. By 2, we have then $\kappa(x,L) = 0$. Since κ is non-degenerate, we must have $x = 0$. □

We refer to the decomposition of L in Proposition 7.0.3 as the **root space decomposition** of L with respect to H; an element of Φ is called a **root**.

Lemma 7.0.4. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L. Let $h \in H$ be such that $\dim C_L(h)$ is minimal. Then $C_L(h) = C_L(H)$.*

Proof. We first claim that for all $s \in H$, we have $C_L(h) \subset C_L(s)$. Let $s \in H$. There are filtrations of F-vector spaces:

$$\begin{aligned}&0 \subset C_L(h) \cap C_L(s) \subset C_L(s) \subset C_L(h) + C_L(s) \subset L,\\ &0 \subset C_L(h) \cap C_L(s) \subset C_L(h) \subset C_L(h) + C_L(s) \subset L.\end{aligned}$$

Consider the operators $\mathrm{ad}(h)$ and $\mathrm{ad}(s)$ on L. Since H is a Cartan subalgebra of L, $\mathrm{ad}(h)$ and $\mathrm{ad}(s)$ commute with each other, and both operators are diagonalizable. The restrictions of $\mathrm{ad}(h)$ and $\mathrm{ad}(s)$ to $C_L(h) \cap C_L(s)$ are zero because $[h, C_L(h)] = 0$ and $[s, C_L(s)] = 0$. Let

$$x_1, \dots, x_k$$

be any basis for $C_L(h) \cap C_L(s)$. Next, consider the restrictions of $\mathrm{ad}(h)$ and $\mathrm{ad}(s)$ to $C_L(s)$. Since $[s, C_L(s)] = 0$, the restriction of $\mathrm{ad}(s)$ to $C_L(s)$ is zero. We claim that $\mathrm{ad}(h)$ maps $C_L(s)$ to itself. To see this, let $x \in C_L(s)$. We calculate:

$$\begin{aligned}[\mathrm{ad}(h)x, s] &= [[h,x],s]\\ &= -[s,[h,x]]\\ &= [h,[x,s]] + [x,[s,h]]\\ &= [h,0] + [x,0]\\ &= 0\end{aligned}$$

because $[x,s] = 0$ (since $x \in C_L(s)$) and $[s,h] = 0$ (since H is abelian). It follows that $\mathrm{ad}(h)x \in C_L(s)$, as claimed. Since both $\mathrm{ad}(s)$ and $\mathrm{ad}(h)$ map $C_L(s)$ to itself, since $\mathrm{ad}(s)$ and $\mathrm{ad}(h)$ commute, and since both $\mathrm{ad}(s)$ and $\mathrm{ad}(h)$ are diagonalizable, the restrictions of $\mathrm{ad}(s)$ and $\mathrm{ad}(h)$ to $C_L(s)$ can be simultaneously diagonalized, so that there exist elements $y_1, \dots, y_\ell$ in $C_L(s)$ so that

$$x_1, \dots, x_k, y_1, \dots, y_\ell$$

is a basis for $C_L(s)$, and each element is an eigenvector for $\mathrm{ad}(s)$ and $\mathrm{ad}(h)$ (the elements $x_1, \dots, x_k$ are already in the 0-eigenspaces for the restrictions of $\mathrm{ad}(h)$ and $\mathrm{ad}(s)$ to $C_L(s)$). Since $\mathrm{ad}(s)$ is zero on $C_L(s)$, the elements $y_1, \dots, y_\ell$ are in the 0-eigenspace for $\mathrm{ad}(s)$. Similarly, there exist elements $z_1, \dots, z_m$ in $C_L(h)$ such that

$$x_1, \dots, x_k, z_1, \dots, z_m$$

is a basis for $C_L(h)$ and each element is an eigenvector for $\mathrm{ad}(s)$ and $\mathrm{ad}(h)$; note that since $\mathrm{ad}(h)$ is zero on $C_L(h)$, the elements $z_1, \dots, z_m$ are in the 0 eigenspace for $\mathrm{ad}(h)$. We claim that

$$x_1, \dots, x_k, y_1, \dots, y_\ell, z_1, \dots, z_m$$

form a basis for $C_L(h) + C_L(s)$. It is evident that these vectors span $C_L(h) + C_L(s)$. Now

$$\begin{aligned}
&\dim(C_L(h) + C_L(s)) \\
&= \dim C_L(s) + \dim(C_L(s) + C_L(h))/C_L(s) \\
&= \dim C_L(s) + \dim C_L(h)/(C_L(s) \cap C_L(h)) \\
&= \dim C_L(s) + \dim C_L(h) - \dim(C_L(s) \cap C_L(h)) \\
&= \dim(C_L(s) \cap C_L(h)) + \dim C_L(s) - \dim(C_L(s) \cap C_L(h)) \\
&\qquad + \dim C_L(h) - \dim(C_L(s) \cap C_L(h)) \\
&= k + \ell + m.
\end{aligned}$$

It follows that this is a basis for $C_L(s) + C_L(h)$. Finally, there exist elements $w_1, \dots, w_n$ in L such that

$$x_1, \dots, x_k, y_1, \dots, y_\ell, z_1, \dots, z_m, w_1, \dots, w_n$$

is a basis for L and $w_1, \dots, w_n$ are eigenvectors for $\mathrm{ad}(s)$ and $\mathrm{ad}(h)$. Since $w_1, \dots, w_n$ are not in $C_L(s)$, it follows that the eigenvalues of $\mathrm{ad}(s)$ on these elements do not include zero; similarly, the eigenvalues of $\mathrm{ad}(h)$ on $w_1, \dots, w_n$ do not include zero. Let $\alpha_, \dots, \alpha_n$ in F and $\beta_1, \dots, \beta_n$ be such that

$$\mathrm{ad}(s)w_i = \alpha_i w_i, \qquad \mathrm{ad}(h)w_i = \beta_i w_i$$

for $i \in \{1, \dots, n\}$. Now let c be any element of F such that

$$c \neq 0, \quad \alpha_1 + c\beta_1 \neq 0, \quad \dots, \quad \alpha_n + c\beta_n \neq 0.$$

We have:

$$\mathrm{ad}(s + c \cdot h)x_i = \mathrm{ad}(s)x_i + c \cdot \mathrm{ad}(h)x_i = 0,$$

$$\mathrm{ad}(s+c\cdot h)y_i = \mathrm{ad}(s)y_i + c\cdot \mathrm{ad}(h)y_i = c\cdot \mathrm{ad}(h)y_i = \text{non-zero multiple of } y_i,$$
$$\mathrm{ad}(s+c\cdot h)z_i = \mathrm{ad}(s)z_i + c\cdot \mathrm{ad}(h)z_i = \mathrm{ad}(s)z_i = \text{non-zero multiple of } z_i,$$
$$\mathrm{ad}(s+c\cdot h)w_i = (\alpha_i + c\beta_i)w_i = \text{non-zero multiple of } w_i.$$

Here $c\cdot \mathrm{ad}(h)y_i$ is a multiple of y_i because y_i is an $\mathrm{ad}(h)$ eigenvector, and this multiple is non-zero because otherwise $[h, y_i] = 0$, contradicting $y_i \notin C_L(s) \cap C_L(h)$. Similary, $\mathrm{ad}(s)z_i$ is a non-zero multiple of z_i. Because

$$x_1, \ldots, x_k, y_1, \ldots, y_\ell, z_1, \ldots, z_m, w_1, \ldots, w_n$$

is a basis for L we conclude that if $x \in L$ is such that $[s + c\cdot h, x] = 0$, then x is in the span of $x_1, \ldots, x_k$; this means that

$$C_L(s + c\cdot h) \subset C_L(s) \cap C_L(h).$$

Since $C_L(s) \cap C_L(h) \subset C_L(s + c\cdot h)$ we get

$$C_L(s + c\cdot h) = C_L(s) \cap C_L(h).$$

By the definition of h, we must have $C_L(h) \subset C_L(s + c\cdot h)$; hence

$$C_L(h) \subset C_L(s) \cap C_L(h).$$

This means that $C_L(h) \subset C_L(s)$.

Finally, to see that $C_L(h) = C_L(H)$, we note first that $C_L(H) \subset C_L(h)$. For the converse inclusion, we have by the first part of the proof:

$$C_L(h) \subset \bigcap_{s\in H} C_L(s) = C_L(H).$$

Hence, $C_L(h) = C_L(H)$. □

Proposition 7.0.5. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L. Then $C_L(H) = H$.*

Proof. Clearly, $H \subset C_L(H)$. To prove the other inclusion, let $x \in C_L(H)$; we need to prove that $x \in H$. By Lemma 7.0.4, there exists $h \in H$ such that $C_L(H) = C_L(h)$. Hence, $x \in C_L(h)$. Let $x = s + n$ be the abstract Jordan decomposition of x. We have $[x, h] = 0$. By Theorem 5.5.5, we obtain $[s, h] = 0$ and $[n, h] = 0$. It follows that $s, n \in C_L(h) = C_L(H)$. Consider the subalgebra $H' = H + Fs$ of L. This subalgebra is abelian, and all the elements of it are semi-simple. By the maximality property of H, we have $H' = H$; this implies that $s \in H$. To prove that $x \in H$ it will now suffice to prove that $n = 0$.

We first show that $C_L(h)$ is a nilpotent Lie algebra. By the second version of Engel's Theorem, Theorem 7.0.1, to prove this it will suffice to prove that $\mathrm{ad}_{C_L(h)}(y)$ is nilpotent for all $y \in C_L(h)$. Let $y \in C_L(h)$, and let $y = r + m$ be the abstract Jordan decomposition of y as a element of L, with r semi-simple and m nilpotent. As in the previous paragraph, $r \in C_L(h)$. Let $z \in C_L(h)$. Then

$$\begin{aligned} \mathrm{ad}_{C_L(h)}(y)z &= [y,z] \\ &= [r,z] + [m,z] \\ &= 0 + [m,z] \\ &= \mathrm{ad}(m)z. \end{aligned}$$

The operator $\mathrm{ad}(m) : L \to L$ is nilpotent; it follows that $\mathrm{ad}_{C_L(h)}(y)$ is also nilpotent. Hence, $C_L(h)$ is a nilpotent Lie algebra.

Now we prove that the n from the first paragraph is zero. Since $C_L(h)$ is a nilpotent Lie algebra, it is a solvable Lie algebra. Consider the Lie subalgebra $\mathrm{ad}(C_L(h))$ of $\mathrm{gl}(L)$. Since L is semi-simple, ad is injective (see Proposition 5.5.1). It follows that $\mathrm{ad}(C_L(h))$ is a solvable Lie subalgebra of $\mathrm{gl}(L)$. By Lie's Theorem, Theorem 3.1.2, there exists a basis for L in which all the elements of $\mathrm{ad}(C_L(h))$ are upper-triangular. The element $\mathrm{ad}(n)$ is a nilpotent element of $\mathrm{gl}(L)$, and is hence strictly upper triangular. Let $z \in C_L(h)$. Then

$$\kappa(n,z) = \mathrm{tr}(\mathrm{ad}(n)\mathrm{ad}(z)) = 0$$

because $\mathrm{ad}(n)\mathrm{ad}(z)$ is also strictly upper triangular. Now $C_L(h) = C_L(H) = L_0$ for the choice H of Cartan subalgebra, and by Proposition 7.0.3, the restriction of the Killing form to L_0 is non-degenerate. This implies that $n = 0$. □

Corollary 7.0.6. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L. Then $L_0 = H$.*

Proof. By definition, and by Proposition 7.0.5,

$$\begin{aligned} L_0 &= \{x \in L : [h,x] = 0 \text{ for all } h \in H\} \\ &= \{x \in L : x \in C_L(H)\} \\ &= H. \end{aligned}$$

This completes the proof. □

Lemma 7.0.7. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3. If $\alpha \in \Phi$, then $-\alpha \in \Phi$. Let $\alpha \in \Phi$, and let $x \in L_\alpha$ be non-zero. There exists $y \in L_{-\alpha}$ such that $Fx + Fy + F[x,y]$ is a Lie subalgebra of L isomorphic to $\mathrm{sl}(2,F)$.*

Proof. Let $x \in L_\alpha$ be non-zero. By 3 of Proposition 7.0.3, the Killing form κ of L is non-degenerate; hence, there exists $z \in L$ such that $\kappa(x,z) \neq 0$. Write

$$z = z_0 + \sum_{\beta \in \Phi} z_\beta$$

for some $z_0 \in H = L_0$ and $z_\beta \in L_\beta$, $\beta \in \Phi$. By 2 of Proposition 7.0.3 we have $\kappa(x, L_\beta) = 0$ for all $\beta \in H^\vee$ such that $\beta + \alpha \neq 0$. Therefore,

$$\begin{aligned} \kappa(x,z) &= \kappa(x,z_0) + \sum_{\beta\in\Phi} \kappa(x,z_\beta) \\ &= \sum_{\substack{\beta\in\Phi \\ \alpha+\beta=0}} \kappa(x,z_\beta). \end{aligned}$$

Since $\kappa(x,z) \neq 0$, this implies that there exists $\beta \in \Phi$ such that $\alpha + \beta = 0$, i.e, $-\alpha \in \Phi$. Also, we have proven that there exists $y \in L_{-\alpha}$ such that $\kappa(x,y) \neq 0$. By 1 of Proposition 7.0.3 and Corollary 7.0.6 we have $[x,y] \in L_0 = H$.

Let $c \in F^\times$. We claim that $S(cy) = Fx + Fy + F[x,y]$ is a Lie subalgebra of L. To prove this it suffices to check that $[[x,y],x], [[x,y],y] \in S(cy)$. Now since $[x,y] \in H$, we have by the definition of L_α,

$$[[x,y],x] = \alpha([x,y])x;$$

also, by the definition of $L_{-\alpha}$,

$$[[x,y],y] = -\alpha([x,y])y.$$

This proves that $S(cy)$ is a Lie subalgebra of L.

To complete the proof we will prove that there exists $c \in F^\times$ such that $S(cy)$ is isomorphic to $\mathrm{sl}(2,F)$. Let $c \in F^\times$, and set

$$e = x, \qquad f = cy, \qquad h = [e,f].$$

To prove that there exists a $c \in F^\times$ such that $S(cy)$ is isomorphic to $\mathrm{sl}(2,F)$ it will suffice to prove that there exists a $c \in F^\times$ such that

$$h \neq 0, \qquad [e,h] = -2e, \qquad [f,h] = 2f.$$

We first claim that h is non-zero for all $c \in F^\times$. We will prove the stronger statement that $\alpha([x,y]) \neq 0$. Assume that $\alpha([x,y]) = 0$; we will obtain a contradiction. From above, we have that $[x,y]$ commutes with x and y. This implies that $\mathrm{ad}([x,y]) = [\mathrm{ad}(x), \mathrm{ad}(y)]$ commutes with $\mathrm{ad}(x)$ and $\mathrm{ad}(y)$; these are elements of $\mathrm{gl}(L)$. By Corollary 3.2.2, the element $\mathrm{ad}([x,y])$ is a nilpotent element of $\mathrm{gl}(L)$. However, by the definition of a Cartan subalgebra, $\mathrm{ad}([x,y])$ is semi-simple. It follows that $[x,y] = 0$. Since $\alpha \neq 0$, there exists $t \in H$ such that $\alpha(t) \neq 0$. Now

$$\begin{aligned} 0 &= \kappa(t,[x,y]) \\ &= \kappa(t,[x,y]) \\ &= \kappa([t,x],y) \\ &= \kappa(\alpha(t)x,y) \\ &= \alpha(t)\kappa(x,y). \end{aligned}$$

This is non-zero, a contradiction. Hence, $\alpha([x,y]) \neq 0$ and consequently $h \neq 0$ for any $c \in F^\times$.

Finally, for any $c \in F^\times$ we have

$$\begin{aligned}[e,h] &= -[h,x] \\ &= -\alpha(h)x \\ &= -\alpha([x,cy])x \\ &= -c\alpha([x,y])x \\ &= -c\alpha([x,y])e\end{aligned}$$

and

$$\begin{aligned}[f,h] &= -[h,f] \\ &= -[[x,cy],cy] \\ &= -c[[x,y],cy] \\ &= -c(-\alpha([x,y]))f \\ &= c\alpha([x,y])f\end{aligned}$$

Setting $c = 2/\alpha([x,y])$ now completes the proof. □

Let the notation be as in Lemma 7.0.7 and its proof. We will write

$$e_\alpha = x, \qquad f_\alpha = (2/\alpha([x,y]))y, \qquad h_\alpha = [e_\alpha, f_\alpha].$$

We have $e_\alpha \in L_\alpha$, $f_\alpha \in L_{-\alpha}$ and $h_\alpha \in H$. The subalgebra $Fe_\alpha + Ff_\alpha + Fh_\alpha$ is isomorphic to $\mathrm{sl}(2,F)$. We will write

$$\mathrm{sl}(\alpha) = Fe_\alpha + Ff_\alpha + Fh_\alpha.$$

We note that

$$\alpha(h_\alpha) = \alpha((2/\alpha([x,y]))[x,y]) = 2.$$

Consider the action of $\mathrm{sl}(\alpha)$ on L. By Weyl's Theorem, Theorem 6.2.4, L can be written as a direct sum of irreducible $\mathrm{sl}(\alpha)$ representations. By Theorem 4.3.7 every one of these irreducible representations is of the form V_d for some integer $d \geq 0$. Moreover, the explicit description of the representations V_d shows that V_d is a direct sum of h_α eigenspaces, and each eigenvalue is an integer. It follows that L is a direct sum of h_α eigenspaces, and that each eigenvalue is an integer. As every subspace L_β for $\beta \in \Phi$ is obviously contained in the $\beta(h_\alpha)$-eigenspace for h_α, this implies that for all $\beta \in \Phi$ we have that $\beta(h_\alpha)$ is an integer.

Proposition 7.0.8. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3. Let $\beta \in \Phi$. The space L_β is one-dimensional, and the only F-multiples of β contained in Φ are β and $-\beta$.*

Proof. Consider the set

$$X(\beta) = \{c \in F : c\beta \in \Phi\}.$$

We have $1 \in X(\beta)$. By the definition of Φ, we have $0 \notin X(\beta)$. Let $c \in X(\beta)$. Let $x \in L_{c\beta}$ be non-zero. Then

$$\begin{aligned}[h_\beta, x] &= (c\beta)(h_\beta)x \\ &= c\beta(h_\beta)x \\ &= 2cx.\end{aligned}$$

By the remark preceding the proposition, $2c$ must be an integer; in particular, we may say that c is positive or negative. Define

$$X_+(\beta) = \{c \in F : c\beta \in \Phi \text{ and } c > 0\}$$

and

$$X_-(\beta) = \{c \in F : c\beta \in \Phi \text{ and } c < 0\}.$$

We have

$$X(\beta) = X_-(\beta) \sqcup X_+(\beta).$$

To prove the proposition it will suffice to prove that

$$\#X_+(\beta) = 1 \qquad \text{and} \qquad \dim L_\beta = 1.$$

Let $c_0 \in X_+(\beta)$ be minimal, and define

$$\alpha = c_0\beta.$$

By definition, $\alpha \in \Phi$. The map

$$X_+(\beta) \xrightarrow{\sim} X_+(\alpha), \qquad c \mapsto c/c_0$$

is a well-defined bijection. Evidently, 1 is the minimal element of $X_+(\alpha)$; in particular, $1/2 \notin X_+(\alpha)$.

Now define

$$M = H \oplus \bigoplus_{c \in X(\alpha)} L_{c\alpha}.$$

We claim that M is an $\mathrm{sl}(\alpha)$ module. Let $h \in H$. Then

$$\begin{aligned}&[e_\alpha, h] = -[h, e_\alpha] = -\alpha(h)e_\alpha \in L_\alpha, \\ &[f_\alpha, h] = -[h, f_\alpha] = \alpha(h)f_\alpha \in L_{-\alpha}, \\ &[h_\alpha, h] = 0.\end{aligned}$$

It follows that $[\mathrm{sl}(\alpha), H] \subset M$. Let $c \in X(\alpha)$. Let $x \in L_{c\alpha}$. Then

$$[e_\alpha, x] \in [L_\alpha, L_{c\alpha}] \subset L_{\alpha + c\alpha} = L_{(c+1)\alpha},$$

$$[f_\alpha, x] \in [L_{-\alpha}, L_{c\alpha}] \subset L_{-\alpha+c\alpha} = L_{(c-1)\alpha},$$
$$[h_\alpha, x] = (c\alpha)(h_\alpha)x \in L_{c\alpha};$$

here, we have used 1 of Proposition 7.0.3. This implies that $[\mathrm{sl}(\alpha), L_{c\alpha}] \subset M$. Thus, $\mathrm{sl}(\alpha)$ acts on M. The subspace M contains several subspaces. Evidently,

$$\mathrm{sl}(\alpha) \subset H \oplus L_\alpha \oplus L_{-\alpha} \subset M.$$

It is clear that $\mathrm{sl}(\alpha)$ is an $\mathrm{sl}(\alpha)$ subspace of M. Also, let

$$K = \ker(\alpha) \subset H.$$

We claim that

$$K \cap \mathrm{sl}(\alpha) = 0.$$

To see this, let $k \in K \cap \mathrm{sl}(\alpha)$. Since $K \subset H$, we have $k \in H \cap \mathrm{sl}(\alpha) = F h_\alpha$; write $k = a h_\alpha$ for some $a \in F$. By the definition of K, $\alpha(k) = 0$. Since $\alpha(h_\alpha) = 2$, we get $a = 0$ so that $k = 0$. Now let

$$N = K \oplus \mathrm{sl}(\alpha).$$

We claim that N is an $\mathrm{sl}(\alpha)$ subspace of M. To prove this it will certainly suffice to prove that $[\mathrm{sl}(\alpha), K] = 0$. Let $k \in K$; since $K \subset H$, we have:

$$[e_\alpha, k] = -[k, e_\alpha] = -\alpha(k)e_\alpha = 0,$$
$$[f_\alpha, k] = -[k, f_\alpha] = \alpha(k)f_\alpha = 0,$$
$$[h_\alpha, k] = 0.$$

It follows that N is an $\mathrm{sl}(\alpha)$-subspace of M. Since K is the kernel of the non-zero linear functional α on H, it follows that $\dim K = \dim H - 1$. Since $h_\alpha \in H$ but $h_\alpha \notin K$, we have $H = K \oplus F h_\alpha$. In particular,

$$H \subset N.$$

By Weyl's Theorem, Theorem 6.2.4, there exists an $\mathrm{sl}(\alpha)$-subspace W of M such that

$$M = N \oplus W.$$

We claim that W is zero. Assume that $W \neq 0$; we will obtain a contradiction.

By Weyl's Theorem, Theorem 6.2.4, we may write W as the direct sum of irreducible representations of $\mathrm{sl}(\alpha)$; by Theorem 4.3.7, each of these representations is of the form V_d for some integer $d \geq 0$.

Assume first that W contains a representation V_d with d even. By the explicit description of V_d, there exists a non-zero vector v in V_d such that $h_\alpha v = 0$, i.e., $[h_\alpha, v] = 0$. Write

$$v = h \oplus \bigoplus_{c \in X(\alpha)} v_{c\alpha}$$

with $h \in H$ and $v_{c\alpha} \in L_{c\alpha}$ for $c \in \{c \in F : c\alpha \in \Phi\}$. We have

$$\begin{aligned} 0 &= [h_\alpha, v] \\ &= [h_\alpha, h] + \sum_{c \in X(\alpha)} [h_\alpha, v_{c\alpha}] \\ &= 0 + \sum_{c \in X(\alpha)} c\alpha(h_\alpha) v_{c\alpha} \\ &= \sum_{c \in X(\alpha)} 2c v_{c\alpha}. \end{aligned}$$

Since the vectors $v_{c\alpha}$ lie in the summands of

$$\bigoplus_{c \in X(\alpha)} L_{c\alpha}$$

and this sum is direct, we must have $v_{c\alpha} = 0$ for all $c \in X(\alpha)$. Hence, $v = h \in H \subset N$. On the other hand, $v \in W$. Therefore, $v \in N \cap W = 0$, so that $v = 0$; this is a contradiction. It follows that the V_d that occur in the decomposition of W are such that d is odd.

Let d be an odd integer with $d \geq 1$ and such that V_d occurs in W. By the explicit description of V_d, there exists a vector v in V_d such that $h_\alpha v = v$, i.e, $[h_\alpha, v] = v$. Again write

$$v = h \oplus \bigoplus_{c \in X(\alpha)} v_{c\alpha}$$

with $h \in H$ and $v_{c\alpha} \in L_{c\alpha}$ for $c \in X(\alpha)$. Then

$$\begin{aligned} v &= [h_\alpha, v] \\ &= [h_\alpha, h] + \sum_{c \in X(\alpha)} [h_\alpha, v_{c\alpha}] \\ &= 0 + \sum_{c \in X(\alpha)} c\alpha(h_\alpha) v_{c\alpha} \\ &= \sum_{c \in \{c \in X(\alpha)} 2c v_{c\alpha}. \end{aligned}$$

Therefore,

$$h \oplus \bigoplus_{c \in X(\alpha)} v_{c\alpha} = \bigoplus_{c \in X(\alpha)} 2c v_{c\alpha}$$

Since $v \neq 0$, this implies that for some $c \in X(\alpha)$ we have $2c = 1$, i.e, $c = 1/2 \in X(\alpha)$. This contradicts the fact that $1/2 \notin X(\alpha)$. It follows that $W = 0$.

Since $W = 0$, we have $N = M$. This implies that $\#X_+(\alpha) = 1$ and $\dim L_\alpha = 1$. Hence, $\#X_+(\beta) = 1$. Since $1 \in X_+(\beta)$, we obtain $X_+(\beta) = \{1\}$, so that $c_0 = 1$. This implies that in fact $\beta = \alpha$, so that $\dim L_\beta = 1$. The proof is complete. $\square$

Proposition 7.0.9. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3. Let $\alpha, \beta \in \Phi$ with $\beta \neq \pm\alpha$.*

1. *We have $\beta(h_\alpha) \in \mathbb{Z}$.*
2. *There exist non-negative integers r and q such that*
$$\{k \in \mathbb{Z} : \beta + k\alpha \in \Phi\} = \{k \in \mathbb{Z} : -r \leq k \leq q\}.$$
Moreover, $r - q = \beta(h_\alpha)$.
3. *If $\alpha + \beta \in \Phi$, then $[e_\alpha, e_\beta]$ is a non-zero multiple of $e_{\alpha+\beta}$.*
4. *We have $\beta - \beta(h_\alpha)\alpha \in \Phi$.*

Proof. Proof of 1. Consider the action of $\mathrm{sl}(\alpha)$ on L. By Weyl's Theorem, Theorem 6.2.4, L is a direct sum of irreducible representations of $\mathrm{sl}(\alpha)$. By Theorem 4.3.7, each of these representations is of the form V_d for some integer $d \geq 0$. Each V_d is a direct sum of eigenspaces for h_α, and each eigenvalue for h_α is an integer. It follows that L is a direct sum of eigenspaces for h_α, with each eigenvalue being an integer. Let $x \in L_\beta$ be non-zero. Then $[h_\alpha, x] = \beta(h_\alpha)x$, so that $\beta(h_\alpha)$ is an eigenvalue for h_α. It follows that $\beta(h_\alpha)$ is an integer.

Proof of 2. Let
$$M = \bigoplus_{k \in \mathbb{Z}} L_{\beta+k\alpha}.$$
We claim that there does not exist a $k \in \mathbb{Z}$ such that $\beta + k\alpha = 0$. For suppose such a k exists; we will obtain a contradiction. We have $\beta = -k\alpha$. Hence, $-k\alpha \in \Phi$. By Proposition 7.0.8 we must have $-k = \pm 1$. Thus, $\beta = \pm\alpha$; this contradicts our hypothesis that $\beta \neq \pm\alpha$ and proving our claim. It follows that for every $k \in \mathbb{Z}$ either $\beta + k\alpha \in \Phi$ or $L_{\beta+k\alpha} = 0$. Next, we assert that M is an $\mathrm{sl}(\alpha)$ module. Let $k \in \mathbb{Z}$ and $x \in L_{\beta+k\alpha}$. Then

$$\begin{aligned}
&[e_\alpha, x] \in [L_\alpha, L_{\beta+k\alpha}] \subset L_{\beta+(k+1)\alpha},\\
&[f_\alpha, x] \in [L_{-\alpha}, L_{\beta+k\alpha}] \subset L_{\beta+(k-1)\alpha},\\
&[h_\alpha, x] = (\beta + k\alpha)(h_\alpha)x = (\beta(h_\alpha) + k\alpha(h_\alpha))x = (\beta(h_\alpha) + 2k)x.
\end{aligned}$$

Here we have used 1 of Proposition 7.0.3 and the fact that $\alpha(h_\alpha) = 2$. These formulas show that M is an $\mathrm{sl}(\alpha)$ module. We also see from the last formula that M is the direct sum of h_α eigenspaces because h_α acts on the zero or one-dimensional F-subspace $L_{\beta+k\alpha}$ by $\beta(h_\alpha) + 2k$ for $k \in \mathbb{Z}$; moreover, every eigenvalue for h_α is an integer, and all the eigenvalues for h_α have the same parity. As in the proof of 1, M is a direct sum of irreducible representations of the form V_d for d a non-negative integer. The explicit description of the representations of the form V_d for d a non-negative integer implies that if more than one such representation V_d occurs in the decomposition of M, then either some h_α eigenspace is at least two-dimensional, or the h_α eigenvalues do not all have the same parity. It follows that M is irreducible, and there exists a non-negative integer such that $M \cong V_d$. The explicit description of V_d implies that

$$M = \bigoplus_{n=0}^{d} M(d-2n)$$

where

$$M(n) = \{x \in M : h_\alpha x = nx\}$$

for $n \in \{0, \ldots, d\}$, and that each of the h_α eigenspaces $M(d-2n)$ for $n \in \{0, \ldots, d\}$ is one-dimensional. Now consider the set

$$\{k \in \mathbb{Z} : \beta + k\alpha \in \Phi\}.$$

This set is non-empty since it contains 0. Let

$$k \in \{k \in \mathbb{Z} : \beta + k\alpha \in \Phi\}$$

Then $L_{\beta+k\alpha} \neq 0$, and from above $\beta(h_\alpha) + 2k$ is an eigenvalue for h_α. This implies that there exists $n \in \{0, \ldots, d\}$ such that $d - 2n = \beta(h_\alpha) + 2k$. Solving for k, we obtain $k = (d - \beta(h_\alpha))/2 - n$. It follows that

$$q = (d - \beta(h_\alpha))/2$$

is an integer; since k may assume the value 0, we also see that q is non-negative. Continuing, we have

$$\begin{aligned} d \geq n &\geq 0, \\ -d \leq -n &\leq 0, \\ q - d \leq q - n &\leq q \\ -(d-q) \leq k &\leq q, \\ -r \leq k &\leq q, \end{aligned}$$

where $r = d - q$. Since k may assume the value 0, r is a non-negative integer. We have proven that

$$\{k \in \mathbb{Z} : \beta + k\alpha \in \Phi\} \subset \{k \in \mathbb{Z} : -r \leq k \leq q\}.$$

Now

$$\#\{k \in \mathbb{Z} : \beta + k\alpha \in \Phi\} = \dim M = \dim V_d = d + 1.$$

Also,

$$\begin{aligned} \#\{k \in \mathbb{Z} : -r \leq k \leq q\} &= q - (-r) + 1 \\ &= q + r + 1 \\ &= q + d - q + 1 \\ &= d + 1. \end{aligned}$$

It follows that

$$\{k \in \mathbb{Z} : \beta + k\alpha \in \Phi\} = \{k \in \mathbb{Z} : -r \leq k \leq q\},$$

as desired. Finally,

$$r - q = d - q - q = d - 2q = d - (d - \beta(h_\alpha)) = \beta(h_\alpha).$$

This completes the proof of 2.

Proof of 3. Assume that $\alpha + \beta \in \Phi$. We have that $\alpha + \beta \neq 0$, $L_{\alpha+\beta}$ is non-zero, and $L_{\alpha+\beta}$ is spanned by $e_{\alpha+\beta}$. To prove 3, it will suffice to prove that $[e_\alpha, e_\beta]$ is non-zero because by 1 of Proposition 7.0.3 we have $[e_\alpha, e_\beta] \in L_{\alpha+\beta}$. Assume that $[e_\alpha, e_\beta] = 0$; we will obtain a contradiction. Let M be as in the proof of 2. Now $e_\beta \in L_\beta \subset M$; also, it was proven that $M \cong V_d$. Since $[e_\alpha, e_\beta] = 0$, by the structure of V_d, we have $[h_\alpha, e_\beta] = de_\beta$. On the other hand, since $e_\beta \in L_\beta$, we have $[h_\alpha, e_\beta] = \beta(h_\alpha)e_\beta$. It follows that $d = \beta(h_\alpha)$. This implies that $q = 0$. By 2, we therefore have

$$1 \notin \{k \in \mathbb{Z} : \beta + k\alpha \in \Phi\}.$$

This contradicts the assumption that $\alpha + \beta \in \Phi$.

Proof of 4. We have

$$\begin{aligned} -r &\leq q - r \leq q, \\ -r &\leq -(r - q) \leq q, \\ -r &\leq -\beta(h_\alpha) \leq q. \end{aligned}$$

Here, $r - q = \beta(h_\alpha)$ by 2. It now follows from 2 that $\beta - \beta(h_\alpha)\alpha \in \Phi$. □

Proposition 7.0.10. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3.*

1. *If $h \in H$ is non-zero, then there exists $\alpha \in \Phi$ such that $\alpha(h) \neq 0$.*

2. *The elements of Φ span $H^\vee$.*

Proof. Proof of 1. Let $h \in H$ be non-zero. Assume that $\alpha(h) = 0$ for all $\alpha \in \Phi$. Let $\alpha \in \Phi$. Then $[h, L_\alpha] \subset \alpha(h)L_\alpha = 0$. It follows that $[h, x] = 0$ for all $x \in L$. Hence, $h \in Z(L) = 0$; this is a contradiction.

Proof of 2. Let W be the span in $H^\vee$ of the elements of Φ. Assume that $W \neq H^\vee$; we will obtain a contradiction. Since W is a proper subspace of $H^\vee$, there exists a non-zero linear functional $f : H^\vee \to F$ such that $f(W) = 0$. Since the natural map $H \to (H^\vee)^\vee$ is an isomorphism, there exists $h \in H$ such that $f(\lambda) = \lambda(h)$ for all $\lambda \in H^\vee$. Now $h \neq 0$ because f is non-zero. If $\lambda \in W$, then $\lambda(h) = f(\lambda) = 0$. This contradicts 1. □

Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3. Consider the F-linear map

$$H \longrightarrow H^\vee \tag{7.1}$$

defined by $h \mapsto \kappa(\cdot, h)$. By 3 of Proposition 7.0.3 and Corollary 7.0.6, this map is injective, i.e., the restriction of the Killing form to H is non-degenerate; since both F-vector spaces have the same dimension, it is an isomorphism. There is thus a natural isomorphism between H and $H^\vee$. In particular, for every root $\alpha \in \Phi$ there exists $t_\alpha \in H$ such that

$$\alpha(x) = \kappa(x, t_\alpha)$$

for $x \in H$.

Lemma 7.0.11. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3. Let $\alpha \in \Phi$.*

1. *For $x \in L_\alpha$ and $y \in L_{-\alpha}$ we have*

$$[x, y] = \kappa(x, y)t_\alpha.$$

In particular,

$$h_\alpha = [e_\alpha, f_\alpha] = \kappa(e_\alpha, f_\alpha)t_\alpha.$$

2. *We have*

$$h_\alpha = \frac{2}{\kappa(t_\alpha, t_\alpha)} t_\alpha.$$

and

$$\kappa(t_\alpha, t_\alpha)\kappa(h_\alpha, h_\alpha) = 4.$$

3. *If $\beta \in \Phi$, then*

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

Proof. 1. Let $h \in H$, $x \in L_\alpha$ and $y \in L_{-\alpha}$. We need to prove that $[x, y] - \kappa(x, y)t_\alpha = 0$. Now by 1 of Proposition 7.0.3 we have $[x, y] \in L_0$, and $H = L_0$ by Corollary 7.0.6. Thus, $[x, y] \in H$. It follows that $[x, y] - \kappa(x, y)t_\alpha$ is in H. Let $h \in H$. Then

$$\begin{aligned}
\kappa(h, [x, y] - \kappa(x, y)t_\alpha) &= \kappa(h, [x, y]) - \kappa(h, \kappa(x, y)t_\alpha) \\
&= \kappa([h, x], y) - \kappa(x, y)\kappa(h, t_\alpha) \\
&= \kappa(\alpha(h)x, y) - \kappa(x, y)\alpha(h) \\
&= \alpha(h)\kappa(x, y) - \kappa(x, y)\alpha(h) \\
&= 0.
\end{aligned}$$

Since this holds for all $h \in H$, and since the restriction of the Killing form to H is non-degenerate, we obtain $[x, y] - \kappa(x, y)t_\alpha = 0$. This proves the first and second assertions.

2. We first note that

$$\begin{aligned} 2 &= \alpha(h_\alpha) \\ &= \kappa(h_\alpha, t_\alpha) \\ &= \kappa(\kappa(e_\alpha, f_\alpha)t_\alpha, t_\alpha) \\ 2 &= \kappa(e_\alpha, f_\alpha)\kappa(t_\alpha, t_\alpha) \\ \frac{2}{\kappa(t_\alpha, t_\alpha)} &= \kappa(e_\alpha, f_\alpha). \end{aligned}$$

The first claim of 2 now follows now from 1 by substitution. Next, we have:

$$\begin{aligned} \kappa(h_\alpha, h_\alpha) &= \kappa(\frac{2}{\kappa(t_\alpha, t_\alpha)}t_\alpha, \frac{2}{\kappa(t_\alpha, t_\alpha)}t_\alpha) \\ &= \frac{2^2}{\kappa(t_\alpha, t_\alpha)^2}\kappa(t_\alpha, t_\alpha) \\ &= \frac{4}{\kappa(t_\alpha, t_\alpha)}. \end{aligned}$$

3. Using the definition of $(\cdot, \cdot)$ and t_α and t_β, we have

$$\begin{aligned} \frac{2(\alpha, \beta)}{(\alpha, \alpha)} &= \frac{2\kappa(t_\alpha, t_\beta)}{\kappa(t_\alpha, t_\alpha)} \\ &= \frac{2}{\kappa(t_\alpha, t_\alpha)} \cdot \kappa(t_\alpha, t_\beta) \\ &= \kappa(e_\alpha, f_\alpha) \cdot \kappa(t_\alpha, t_\beta) \\ &= \kappa(\kappa(e_\alpha, f_\alpha) \cdot t_\alpha, t_\beta) \\ &= \kappa(h_\alpha, t_\beta) \\ &= \beta(h_\alpha). \end{aligned}$$

This completes the proof. □

We note that by 2 of Lemma 7.0.11 the element h_α is determined soley by t_α, which in turn is canonically determined by the Killing form.

Proposition 7.0.12. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3. If $\alpha, \beta \in \Phi$, then $\kappa(h_\alpha, h_\beta) \in \mathbb{Z}$ and $\kappa(t_\alpha, t_\beta) \in \mathbb{Q}$.*

Proof. We begin by considering the matrix of the linear operator $\text{ad}(h_\alpha) = [h_\alpha, \cdot]$ with respect to the decomposition

$$L = H \oplus \bigoplus_{\gamma \in \Phi} L_\gamma.$$

Since H is abelian, $\mathrm{ad}(h_\alpha)$ acts by zero on H. If $\gamma \in \Phi$, then $\mathrm{ad}(h_\alpha)$ acts by multiplication by $\gamma(h_\alpha)$ on L_γ (by the definition of L_γ). It follows that the matrix of $\mathrm{ad}(h_\alpha)$, with respect to the above decomposition, is:

$$\begin{bmatrix} 0 & & & \\ & \ddots & & \\ & & \gamma(h_\alpha) & \\ & & & \ddots \end{bmatrix}.$$

Therefore, the matrix of $\mathrm{ad}(h_\alpha) \circ \mathrm{ad}(h_\beta)$ is

$$\begin{bmatrix} 0 & & & \\ & \ddots & & \\ & & \gamma(h_\alpha)\gamma(h_\beta) & \\ & & & \ddots \end{bmatrix}.$$

This implies that

$$\kappa(h_\alpha, h_\beta) = \mathrm{tr}(\mathrm{ad}(h_\alpha) \circ \mathrm{ad}(h_\beta)) = \sum_{\gamma \in \Phi} \gamma(h_\alpha)\gamma(h_\beta).$$

By 1 of Proposition 7.0.9 the product $\gamma(h_\alpha)\gamma(h_\beta)$ is in $\mathbb{Z}$ for all $\gamma \in \Phi$. This implies that $\kappa(h_\alpha, h_\beta) \in \mathbb{Z}$. Next, using Lemma 7.0.11,

$$\begin{aligned} \kappa(t_\alpha, t_\beta) &= \kappa(2^{-1}\kappa(t_\alpha, t_\alpha)h_\alpha, 2^{-1}\kappa(t_\beta, t_\beta)h_\beta) \\ &= 4^{-1}\kappa(t_\alpha, t_\alpha)\kappa(t_\beta, t_\beta)\kappa(h_\alpha, h_\beta) \\ &= 4^{-1}\frac{4}{\kappa(h_\alpha, h_\alpha)}\frac{4}{\kappa(h_\beta, h_\beta)}\kappa(h_\alpha, h_\beta) \\ &= \frac{4\kappa(h_\alpha, h_\beta)}{\kappa(h_\alpha, h_\alpha)\kappa(h_\beta, h_\beta)}. \end{aligned}$$

This completes the proof. □

Let F have characteristic zero and be algebraically closed. Let L be a semisimple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3. We introduce a non-degenerate F-symmetric bilinear form $(\cdot, \cdot)$ on $H^\vee$ via the isomorphism

$$H \xrightarrow{\sim} H^\vee$$

from (7.1). If $\alpha, \beta \in \Phi$, then we have

$$(\alpha, \beta) = \kappa(t_\alpha, t_\beta),$$

and by Proposition 7.0.12,

$$(\alpha, \beta) \in \mathbb{Q}.$$

Let K be a subfield of F. Evidently, $\mathbb{Q} \subset K$. We define V_K to be the K-subspace of $H^\vee$ generated by Φ.

Proposition 7.0.13. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3. Let $\{\alpha_1, \dots, \alpha_\ell\}$ be an F-basis for $H^\vee$ with $\alpha_1, \dots, \alpha_\ell \in \Phi$; such a basis exists by 2 of Proposition 7.0.10. Let $\beta \in \Phi$, and write*

$$\beta = c_1\alpha_1 + \cdots + c_\ell\alpha_\ell$$

for $c_1, \dots, c_\ell \in F$. Then $c_1, \dots, c_\ell \in \mathbb{Q}$.

Proof. Let $i \in \{1, \dots, \ell\}$. Then

$$(\alpha_i, \beta) = c_1(\alpha_i, \alpha_1) + \cdots + c_\ell(\alpha_i, \alpha_\ell).$$

It follows that

$$\begin{bmatrix}(\alpha_1, \beta)\\ \vdots \\ (\alpha_\ell, \beta)\end{bmatrix} = S \begin{bmatrix}c_1\\ \vdots \\ c_\ell\end{bmatrix}$$

where

$$S = \begin{bmatrix}(\alpha_1, \alpha_1) & \cdots & (\alpha_1, \alpha_\ell)\\ \vdots & & \vdots \\ (\alpha_\ell, \alpha_1) & \cdots & (\alpha_\ell, \alpha_\ell)\end{bmatrix}.$$

Since $(\cdot, \cdot)$ is a non-degenerate symmetric bilinear form the matrix S is invertible. Therefore,

$$S^{-1}\begin{bmatrix}(\alpha_1, \beta)\\ \vdots \\ (\alpha_\ell, \beta)\end{bmatrix} = \begin{bmatrix}c_1\\ \vdots \\ c_\ell\end{bmatrix}.$$

By the remark preceding the proposition the entries of all the matrices on the left are in $\mathbb{Q}$; hence, $c_1, \dots, c_\ell \in \mathbb{Q}$. □

Proposition 7.0.14. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3. As a Lie algebra, L is generated by the root spaces L_α for $\alpha \in \Phi$.*

Proof. By the decomposition

$$L = H \oplus \bigoplus_{\alpha\in\Phi} L_\alpha$$

that follows from Proposition 7.0.3 and Corollary 7.0.6 it suffices to prove that H is contained in the F-span of the F-subspaces $[L_\alpha, L_{-\alpha}]$ for $\alpha \in \Phi$. By the discussion preceding Proposition 7.0.8, the elements h_α for $\alpha \in \Phi$ are contained in this F-span. By Lemma 7.0.11, this F-span therefore contains the elements t_α for $\alpha \in \Phi$. By Lemma 7.0.10, the linear forms $\alpha \in \Phi$ span $H^\vee$; this implies that the elements t_α for $\alpha \in \Phi$ span H. The F-span of the F-subspaces $[L_\alpha, L_{-\alpha}]$ for $\alpha \in \Phi$ therefore contains H. □

7.1 An associated inner product space

Let F be algebraically closed and have characteristic zero. Then $\mathbb{Q} \subset F$.

Lemma 7.1.1. *Let V_0 be a finite-dimensional vector space over $\mathbb{Q}$, and assume that $(\cdot,\cdot)_0 : V_0 \times V_0 \to \mathbb{Q}$ is a positive-definite, symmetric bilinear form. Let $V = \mathbb{R} \otimes_{\mathbb{Q}} V_0$, so that V is an $\mathbb{R}$ vector space. Let $(\cdot,\cdot) : V \times V \to \mathbb{R}$ be the symmetric bilinear form determined by the condition that*

$$(a \otimes v, b \otimes w) = ab(v, w)_0$$

for $a, b \in \mathbb{R}$ and $v, w \in V_0$. The symmetric bilinear form $(\cdot,\cdot)$ is positive-definite.

Proof. Let $v_1, \ldots, v_n$ be an orthogonal basis for the $\mathbb{Q}$ vector space V_0; then $1 \otimes v_1, \ldots, 1 \otimes v_n$ is an orthogonal basis for the real vector space V. Let $x \in V$. There exist $a_1, \ldots, a_n \in \mathbb{R}$ such that

$$x = a_1(1 \otimes v_1) + \cdots + a_n(1 \otimes v_n) = a_1 \otimes v_1 + \cdots + a_n \otimes v_n.$$

We have

$$\begin{aligned}(x, x) &= \sum_{i,j=1}^{n} (a_i \otimes v_i, a_j \otimes v_j)\\ &= \sum_{i,j=1}^{n} a_i a_j (v_i, v_j)_0\\ &= \sum_{i=1}^{n} a_i^2 (v_i, v_i)_0.\end{aligned}$$

Since $(\cdot,\cdot)_0$ is positive-definite, $(v_i, v_i)_0 > 0$ for $i \in \{1, \ldots, n\}$. It follows that if $(x, x) = 0$, then $a_1 = \cdots = a_n = 0$, so that $x = 0$. □

Proposition 7.1.2. *Let F be algebraically closed and have characteristic zero. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let the notation be as in Proposition 7.0.3. Let V_0 be the $\mathbb{Q}$ subspace of $H^\vee = \mathrm{Hom}_F(H, F)$ spanned by the elements of Φ. We have $\dim_{\mathbb{Q}} V_0 = \dim_F H^\vee$. The restriction $(\cdot,\cdot)_0$ of the symmetric bilinear form on $H^\vee$ (which corresponds to the Killing form) to $V_0 \times V_0$ takes values in $\mathbb{Q}$ and is positive-definite.*

Proof. Let $\{\alpha_1, \ldots, \alpha_\ell\} \subset \Phi$ be as in the statement of Proposition 7.0.13. Then by Proposition 7.0.13 the set $\{\alpha_1, \ldots, \alpha_\ell\}$ is a basis for the $\mathbb{Q}$ vector space V_0, and is also a basis for the F vector space $H^\vee$. Hence, $\dim_{\mathbb{Q}} V_0 = \dim_F H^\vee = \dim_F H$.

To see that $(\cdot,\cdot)_0$ takes values in $\mathbb{Q}$ it suffices to see that $(\alpha, \beta) \in \mathbb{Q}$ for $\alpha, \beta \in \Phi$. This follows from Proposition 7.0.12.

Let $y \in V_0$. Regard y as an element of $H^\vee$. Let h be the element of H corresponding to y under the isomorphism $H \xrightarrow{\sim} H^\vee$. By Corollary 7.0.6 and Proposition 7.0.8 we have

$$L = H \oplus \bigoplus_{\alpha\in\Phi} L_\alpha$$

and each of the subspaces L_α is one-dimensional. Moreover, $\mathrm{ad}(h)$ acts by 0 on H and by $\alpha(h)$ on L_α for $\alpha \in \Phi$. It follows that

$$\begin{aligned}
(y, y) &= \kappa(h, h) \\
&= \mathrm{tr}(\mathrm{ad}(h) \circ \mathrm{ad}(h)) \\
&= \sum_{\alpha\in\Phi} \alpha(h)^2 \\
&= \sum_{\alpha\in\Phi} \kappa(t_\alpha, h)^2 \\
&= \sum_{\alpha\in\Phi} (\alpha, y)^2.
\end{aligned}$$

Since $(\alpha, y) \in \mathbb{R}$ for $\alpha \in \Phi$, we have $(y, y) \geq 0$. Assume that $(y, y) = 0$. By the above formula for (y, y) we have that $\alpha(h) = \kappa(t_\alpha, h) = (\alpha, y) = 0$ for all $\alpha \in \Phi$, or equivalently, $\alpha(h) = 0$ for all $\alpha \in \Phi$. By Proposition 7.0.10, this implies that $h = 0$, so that $y = 0$. □

Chapter 8

Root systems

8.1 The definition

Let V be a finite-dimensional vector space over $\mathbb{R}$, and fix an inner product $(\cdot,\cdot)$ on V. By definition, $(\cdot,\cdot) : V \times V \to \mathbb{R}$ is a symmetric bilinear form such that $(x,x) > 0$ for all non-zero $x \in V$. Let $v \in V$ be non-zero. We define the **reflection** determined by v to be the unique $\mathbb{R}$ linear map $s_v : V \to V$ such that $s_v(v) = -v$ and $s_v(w) = w$ for all $w \in (\mathbb{R}v)^{\perp}$. A calculation shows that s_v is given by the formula

$$s_v(x) = x - \frac{2(x,v)}{(v,v)}v$$

for $x \in V$. Another calculation also shows that s_v preserves the inner product $(\cdot,\cdot)$, i.e.,

$$(s_v(x), s_v(y)) = (x,y)$$

for $x, y \in V$; that is, s_v is in the orthogonal group $\mathrm{O}(V)$. Evidently,

$$\det(s_v) = -1.$$

We will write

$$\langle x, y\rangle = \frac{2(x,y)}{(y,y)}$$

for $x, y \in V$. We note that the function $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ is linear in the first variable; however, this function is not linear in the second variable. We have

$$s_v(x) = x - \langle x, v\rangle v$$

for $x \in V$.

Let R be a subset of V. We say that R is a **root system** if R satisfies the following axioms:

(R1) The set R is finite, does not contain 0, and spans V.

(R2) If $\alpha \in R$, then α and $-\alpha$ are the only scalar multiples of α that are contained in R.

(R3) If $\alpha \in R$, then $s_\alpha(R) = R$, so that s_α permutes the elements of R.

(R4) If $\alpha, \beta \in R$, then $\langle \alpha, \beta \rangle \in \mathbb{Z}$.

8.2 Root systems from Lie algebras

Let F be algebraically closed and have characteristic zero. Let L be a semi-simple Lie algebra over F. Let H be a Cartan subalgebra of L, and let

$$L = L_0 \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$$

be the root space decomposition of L with respect to L. Here, for a F linear functional $f : H \to F$,

$$L_f = \{x \in L : [h, x] = f(h)x \text{ for all } h \in H\}.$$

In particular,

$$L_0 = \{x \in L : [h, x] = 0 \text{ for all } h \in H\}.$$

Here, Φ is the subset of α in

$$H^\vee = \mathrm{Hom}_F(H, F)$$

such that $L_\alpha \neq 0$. The elements of Φ are called the roots of L with respect to H. By Corollary 7.0.6 we have $L_0 = H$ so that in fact

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

Previously, we proved that the F subspaces L_α for $\alpha \in \Phi$ are one-dimensional (Proposition 7.0.8). We also proved that the restriction of the Killing form κ to H is non-degenerate (Proposition 7.0.3 and Corollary 7.0.6). Thus, there is an induced F linear isomorphism

$$H \xrightarrow{\sim} H^\vee.$$

Via this isomorphism, we defined an F symmetric bilinear form on $H^\vee$ (by transferring over the Killing form via the isomorphism). Let

$$V_0 = \mathbb{Q} \text{ span of } \Phi \text{ in } H^\vee.$$

By Proposition 7.1.2, we have

$$\dim_{\mathbb{Q}} V_0 = \dim_F H^\vee = \dim_F H,$$

and the restriction $(\cdot,\cdot)_0$ of the symmetric bilinear form on $H^\vee$ to V_0 is an inner product, i.e., is positive definite, and is $\mathbb{Q}$ valued. Let

$$V = \mathbb{R} \otimes_{\mathbb{Q}} V_0,$$

so that V is an $\mathbb{R}$ vector space, and define an $\mathbb{R}$ symmetric bilinear form $(\cdot,\cdot)$ on V by declaring $(a \otimes v, b \otimes w) = ab(v,w)_0$ for $a, b \in \mathbb{R}$ and $v, w \in V_0$. By Lemma 7.1.1, we have that $(\cdot,\cdot)$ is positive-definite.

Proposition 8.2.1. *Let the notation be as in the discussion preceding the proposition. The subset Φ of the inner product space V is a root system.*

Proof. It is clear that (R1) is satisfied. (R2) is satisfied by Proposition 7.0.8. To see that (R3) is satisfied, let $\alpha, \beta \in \Phi$. Then by 3 of Lemma 7.0.11,

$$s_\alpha(\beta) = \beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha = \beta - \beta(h_\alpha)\alpha.$$

By 4 of Proposition 7.0.9 we have $\beta - \beta(h_\alpha)\alpha \in \Phi$. It follows that $s_\alpha(\beta) \in \Phi$, so that (R3) is satisfied. To prove that (R4) holds, again let $\alpha, \beta \in \Phi$. We have

$$\langle \alpha, \beta \rangle = \frac{2(\alpha,\beta)}{(\beta,\beta)}.$$

By 3 of Lemma 7.0.11 we have

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

Finally, by 1 of Proposition 7.0.9, this quantity is an integer. This proves (R4). ☐

8.3 Basic theory of root systems

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$. The Cauchy-Schwartz inequality asserts that

$$|(x,y)| \le \|x\|\|y\|$$

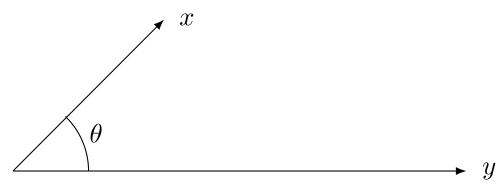
for $x, y \in V$. It follows that if $x, y \in V$ are nonzero, then

$$-1 \le \frac{(x,y)}{\|x\|\|y\|} \le 1.$$

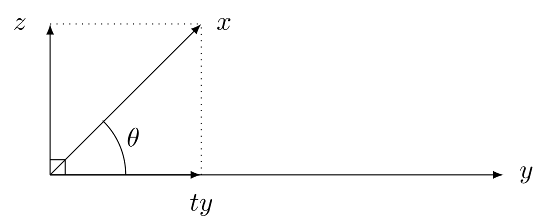
If $x, y \in V$ are nonzero, then we define the **angle** between x and y to be the unique number $0 \le \theta \le \pi$ such that

$$(x,y) = \|x\|\|y\| \cos\theta.$$

The inner product measures the angle between two vectors, though it is a bit more complicated in that the lengths of x and y are also involved. The term "angle" does make sense geometrically. For example, suppose that $V = \mathbb{R}^2$ and we have:



Project x onto y, to obtain ty:



Then we have

$$x = z + ty.$$

Taking the inner product with y, we get

$$\begin{aligned}
(x, y) &= (z, y) + (ty, y) \\
(x, y) &= 0 + t(y, y) \\
(x, y) &= t\|y\|^2 \\
t &= \frac{(x, y)}{\|y\|^2}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\cos\theta &= \frac{\|ty\|}{\|x\|} \\
\cos\theta &= t\frac{\|y\|}{\|x\|} \\
t &= \frac{\|x\|}{\|y\|}\cos\theta.
\end{aligned}$$

If we equate the two formulas for t we get $(x, y) = \|x\|\|y\|\cos\theta$. We say that two vectors are **orthogonal** if $(x, y) = 0$; this is equivalent to the angle between x and y being $\pi/2$. If $(x, y) > 0$, then we will say that x and y form an **acute** angle; this is equivalent to $0 < \theta < \pi/2$. If $(x, y) < 0$, then we will say that x and y form an **obtuse** angle; this is equivalent to $\pi/2 < \theta \leq \pi$.

Non-zero vectors also define some useful geometric objects. Let $v \in V$ be non-zero. We may consider three sets that partition V:

$$\{x \in V : (x, v) > 0\}, \qquad P = \{x \in V : (x, v) = 0\}, \qquad \{y \in V : (x, v) < 0\}.$$

The first set consists of the vectors that form an acute angle with v, the middle set is the hyperplane P orthogonal to $\mathbb{R}v$, and the last set consists of the vectors that form an obtuse angle with v. We refer to the first and last sets as the **half-spaces** defined by P. Of course, v lies in the first half-space. The formula for the reflection s_v shows that

$$(s_v(x), v) = -(x, v)$$

for x in V, so that S sends one half-space into the other half-space. Also, S acts by the identity on P. Multiplication by -1 also sends one half-space into the other half-space; however, while multiplication by -1 preserves P, it is not the identity on P.

Lemma 8.3.1. *Let V be a vector space over $\mathbb{R}$ with an inner product $(\cdot,\cdot)$. Let $x, y \in V$ and assume that x and y are both non-zero. The following are equivalent:*

1. *The vectors x and y are linearly dependent.*
2. *We have $(x,y)^2 = (x,x)(y,y) = \|x\|^2\|y\|^2$.*
3. *The angle between x and y is 0 or π.*

Proof. $1 \implies 2$. This clear.

$2 \implies 3$. Let θ be the angle between x and y. We have

$$(x,y)^2 = \|x\|^2\|y\|^2\cos^2\theta$$

Assume that $(x,y)^2 = (x,x)(y,y) = \|x\|^2\|y\|^2$. Then $(x,y)^2 = \|x\|^2\|y\|^2 \neq 0$, and $\cos^2\theta = 1$, so that $\cos\theta = \pm 1$. This implies that $\theta = 0$ or $\theta = \pi/2$.

$3 \implies 2$. Assume that the angle θ between x and y is 0 or π. Then $\cos^2\theta = 1$. Hence, $(x,y)^2 = \|x\|^2\|y\|^2$.

$2 \implies 1$. Suppose that $(x,y)^2 = (x,x)(y,y)$. We have

$$\begin{aligned}(y - \frac{(x,y)}{(x,x)}x, y - \frac{(x,y)}{(x,x)}x) &= (y,y) - 2\frac{(x,y)}{(x,x)}(x,y) + \frac{(x,y)^2}{(x,x)^2}(x,x) \\ &= (y,y) - 2\frac{(x,y)^2}{(x,x)} + \frac{(x,y)^2}{(x,x)} \\ &= (y,y) - \frac{(x,y)^2}{(x,x)} \\ &= (y,y) - \frac{(x,x)(y,y)}{(x,x)} \\ &= 0.\end{aligned}$$

It follows that $y - \frac{(x,y)}{(x,x)}x = 0$, so that x and y are linearly dependent. ◻

Lemma 8.3.2. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. Let $\alpha,\beta \in R$, and assume that $\alpha \neq \pm\beta$. Then*

$$\langle \alpha,\beta\rangle\langle\beta,\alpha\rangle \in \{0,1,2,3\}.$$

Proof. Let θ be the angle between α and β. We have

$$\begin{aligned}\langle \alpha,\beta\rangle\langle\beta,\alpha\rangle &= \frac{2(\alpha,\beta)}{(\beta,\beta)}\frac{2(\beta,\alpha)}{(\alpha,\alpha)}\\ &= \frac{4(\alpha,\beta)^2}{\|\alpha\|^2\|\beta\|^2}\\ \langle \alpha,\beta\rangle\langle\beta,\alpha\rangle &= 4\cos^2\theta.\end{aligned}$$

Since $\langle \alpha,\beta\rangle\langle\beta,\alpha\rangle$ is an integer, the above equality implies that $4\cos^2\theta$ is an integer. Since $0 \leq \cos^2\theta \leq 1$, we must have

$$\langle \alpha,\beta\rangle\langle\beta,\alpha\rangle = 4\cos^2\theta \in \{0,1,2,3,4\}.$$

We claim that $4\cos^2\theta = 4$ is impossible. Assume that $4\cos^2\theta = 4$; then $\cos^2\theta = 1$. This implies that $\theta = 0$ or $\theta = \pi$. By Lemma 8.3.1 it follows that α and β are linearly dependent, and consequently that β is a scalar multiple of α. By (R2), we must have $\beta = \pm\alpha$; this is a contradiction. □

Lemma 8.3.3. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. Let $\alpha,\beta \in R$, and assume that $\alpha \neq \pm\beta$ and $\|\beta\| \geq \|\alpha\|$. Let θ be the angle between α and β. Exactly one of the following possibilities holds:*

angle type	θ	$\cos\theta$	$\langle\alpha,\beta\rangle$	$\langle\beta,\alpha\rangle$	$\frac{\|\beta\|}{\|\alpha\|}$
strictly acute	$\pi/6 = 30°$	$\sqrt{3}/2$	1	3	$\sqrt{3}$
	$\pi/4 = 45°$	$\sqrt{2}/2$	1	2	$\sqrt{2}$
	$\pi/3 = 60°$	$1/2$	1	1	1
right	$\pi/2 = 90°$	0	0	0	not determined
strictly obtuse	$2\pi/3 = 120°$	$-1/2$	-1	-1	1
	$3\pi/4 = 135°$	$-\sqrt{2}/2$	-1	-2	$\sqrt{2}$
	$5\pi/6 = 150°$	$-\sqrt{3}/2$	-1	-3	$\sqrt{3}$

Proof. By the assumption $\|\beta\| \geq \|\alpha\|$ we have $(\beta,\beta) = \|\beta\|^2 \geq (\alpha,\alpha) = \|\alpha\|^2$, so that

$$|\langle\beta,\alpha\rangle| = \frac{2|(\beta,\alpha)|}{(\alpha,\alpha)} \geq \frac{2|(\alpha,\beta)|}{(\beta,\beta)} = |\langle\alpha,\beta\rangle|.$$

By (R4) we have that $\langle\alpha,\beta\rangle$ and $\langle\beta,\alpha\rangle$ are integers, and by Lemma 8.3.2 we have $\langle\alpha,\beta\rangle\langle\beta,\alpha\rangle\in\{0,1,2,3\}$. These facts imply that the possibilities for $\langle\alpha,\beta\rangle$ and $\langle\beta,\alpha\rangle$ are as in the table.

Assume first that $\langle\beta,\alpha\rangle=\langle\alpha,\beta\rangle=0$. From above, $\langle\alpha,\beta\rangle\langle\beta,\alpha\rangle=4\cos^2\theta$. It follows that $\cos\theta=0$, so that $\theta=\pi/2=90°$.

Assume next that $\langle\beta,\alpha\rangle\neq 0$. Now

$$\frac{\langle\beta,\alpha\rangle}{\langle\alpha,\beta\rangle}=\frac{2(\beta,\alpha)}{(\alpha,\alpha)}\frac{(\beta,\beta)}{2(\alpha,\beta)}=\frac{(\beta,\beta)}{(\alpha,\alpha)},$$

so that $\frac{\langle\beta,\alpha\rangle}{\langle\alpha,\beta\rangle}$ is positive and

$$\sqrt{\frac{\langle\beta,\alpha\rangle}{\langle\alpha,\beta\rangle}}=\frac{\|\beta\|}{\|\alpha\|},$$

This yields the $\|\beta\|/\|\alpha\|$ column. Finally,

$$\begin{aligned}\langle\alpha,\beta\rangle&=\frac{2(\alpha,\beta)}{(\beta,\beta)}\\&=\frac{2\|\alpha\|\|\beta\|\cos\theta}{\|\beta\|^2}\\ \langle\alpha,\beta\rangle&=2\frac{\|\alpha\|}{\|\beta\|}\cos\theta\end{aligned}$$

so that

$$\cos\theta=\frac{1}{2}\frac{\|\beta\|}{\|\alpha\|}\langle\alpha,\beta\rangle.$$

This gives the $\cos\theta$ column. □

Lemma 8.3.4. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. Let $\alpha,\beta\in R$. Assume that $\alpha\neq\pm\beta$ and $\|\beta\|\geq\|\alpha\|$.*

1. *Assume that the angle θ between α and β is strictly obtuse, so that by Lemma 8.3.3 we have $\theta=2\pi/3=120°$, $\theta=3\pi/4=135°$, or $\theta=5\pi/6=150°$. Then $\alpha+\beta\in R$. Moreover,*

$$\begin{aligned}\theta=3\pi/4=135°&\implies 2\alpha+\beta\in R,\\ \theta=5\pi/6=150°&\implies 3\alpha+\beta\in R.\end{aligned}$$

2. *Assume that the angle between α and β is strictly acute, so that by Lemma 8.3.3 we have $\theta=\pi/6=30°$, $\theta=\pi/4=45°$, or $\theta=\pi/3=60°$. Then $-\alpha+\beta\in R$. Moreover,*

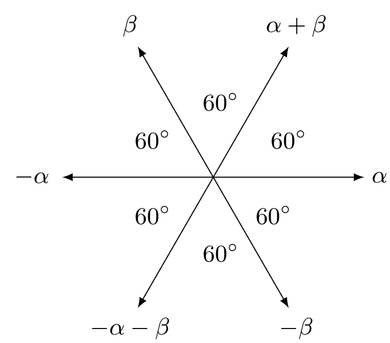
$$\begin{aligned}\theta=\pi/4=45°&\implies -2\alpha+\beta\in R,\\ \theta=\pi/3=60°&\implies -3\alpha+\beta\in R.\end{aligned}$$

Proof. 1. By (R3), we have $s_\beta(\alpha) = \alpha - \langle \alpha, \beta \rangle \beta \in R$. Since the angle between α and β is strictly obtuse, by Lemma 8.3.3 we have that $\langle \alpha, \beta \rangle = -1$. Therefore, $\alpha + \beta \in R$. Assume that $\theta = 3\pi/4 = 135°$. By Lemma 8.3.3 we $\langle \beta, \alpha \rangle = -2$. Hence, $s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha = \beta + 2\alpha \in R$. The case when $\theta = 5\pi/6 = 150°$ is similar.

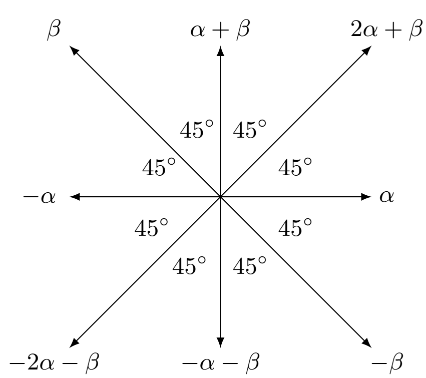
2. By (R3), we have $s_\beta(\alpha) = \alpha - \langle \alpha, \beta \rangle \beta \in R$. Since the angle between α and β is strictly acute, by Lemma 8.3.3 we have that $\langle \alpha, \beta \rangle = 1$. Therefore, $\alpha - \beta \in R$. Hence, $-\alpha + \beta \in R$. Assume that $\theta = \pi/4 = 45°$. By Lemma 8.3.3 we $\langle \beta, \alpha \rangle = 2$. Hence, $s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha = \beta - 2\alpha \in R$. The case $\theta = \pi/3 = 60°$ is similar. □

Proposition 8.3.5. *Let $V = \mathbb{R}^2$ equipt with the usual inner product $(\cdot, \cdot)$, and let R be a root system in V. Let ℓ be the length of the shortest root in R. Let S be the set of pairs (α, β) of non-colinear roots such that $\|\alpha\| = \ell$ and the angle θ between α and β is obtuse, and β is to the left of α. The set S is non-empty. Fix a pair (α, β) in S such that θ is maximal. Then*

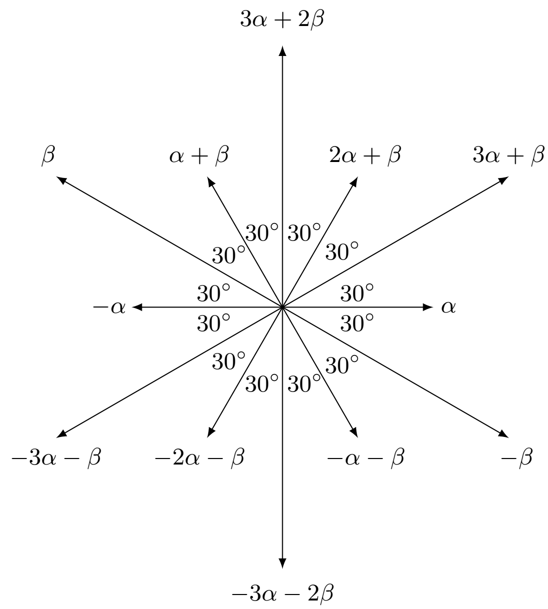
1. *(A_2 root system) If $\theta = 120°$ (so that $\|\alpha\| = \|\beta\|$ by Proposition 8.3.3), then R, α, and β are as follows:*



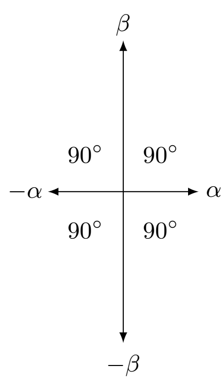
2. *(B_2 root system) If $\theta = 135°$ (so that $\|\beta\| = \sqrt{2}\|\alpha\|$ by Proposition 8.3.3), then R, α, and β are as follows:*



3. *(G_2 root system) If $\theta = 150°$ (so that $\|\beta\| = \sqrt{3}\|\alpha\|$ by Proposition 8.3.3), then R, α, and β are as follows:*



4. *($A_1 \times A_1$ root system) If $\theta = 90°$ (so that the relationship between $\|\beta\|$ and $\|\alpha\|$ is not determined by Proposition 8.3.3), then R, α, and β are as follows:*



Proof. Let (α, β) be a pair of non-colinear roots in R such that $\|\alpha\| = \ell$; such a pair must exist because R contains a basis which includes α. If the angle between α and β is acute, then the angle between α and $-\beta$ is obtuse. Thus, there exists a pair of roots (α, β) in R such that $\|\alpha\| = \ell$ and the angle between α and β is obtuse. If β is the right of α, then $-\beta$ forms an acute angle with

α and is to the left of α; in this case, $s_\alpha(\beta)$ forms an obtuse angle with α and $s_\alpha(\beta)$ is to the left of β. It follows that S is non-empty.

Assume that $\theta = 120°$, so that $\|\alpha\| = \|\beta\|$ by Lemma 8.3.3. By Lemma 8.3.4, $\alpha+\beta \in R$. It follows that $\alpha, \beta, \alpha+\beta, -\alpha, -\beta, -\alpha-\beta \in R$. By geometry, $\|\alpha+\beta\| = \|\alpha\| = \|\beta\|$. It follows that R contains the vectors in 1. Assume that R contains a root γ other than $\alpha, \beta, \alpha+\beta, -\alpha, -\beta, -\alpha-\beta$. By Lemma 8.3.3 we see that γ must lie halfway between two adjacent roots from $\alpha, \beta, \alpha+\beta, -\alpha, -\beta, -\alpha-\beta$. This implies that θ is not maximal, a contradiction.

Assume that $\theta = 135°$, so that $\|\beta\| = \sqrt{2}\|\alpha\|$ by Lemma 8.3.3. By Lemma 8.3.4, we have $\alpha+\beta, 2\alpha+\beta \in R$. It follows that R contains $\alpha, \beta, \alpha+\beta, 2\alpha+\beta, -\alpha, -\beta, -\alpha-\beta, -2\alpha-\beta$, so that R contains the vectors in 2. Assume that R contains a root γ other than $\alpha, \beta, \alpha+\beta, 2\alpha+\beta, -\alpha, -\beta, -\alpha-\beta, -2\alpha-\beta$. Then γ must make an angle strictly less than $30°$ with one of $\alpha, \beta, \alpha+\beta, 2\alpha+\beta, -\alpha, -\beta, -\alpha-\beta, -2\alpha-\beta$. This is impossible by Lemma 8.3.3.

Assume that $\theta = 150°$, so that $\|\beta\| = \sqrt{3}\|\alpha\|$ by Lemma 8.3.3. By Lemma 8.3.4 we have $\alpha+\beta, 3\alpha+\beta \in R$. By geometry, the angle between α and $3\alpha+\beta$ is $30°$. By Lemma 8.3.3, $-\alpha+(3\alpha+\beta) = 2\alpha+\beta \in R$. By geometry, the angle between β and $3\alpha+\beta$ is $120°$. By Lemma 8.3.3, $\beta+3\alpha+\beta = 3\alpha+2\beta \in R$. It now follows that R contains the vectors in 3. Assume that R contains a vector γ other than $\alpha, \beta, \alpha+\beta, 2\alpha+\beta, 3\alpha+\beta, 3\alpha+2\beta, -\alpha, -\beta, -\alpha-\beta, -2\alpha-\beta, -3\alpha-\beta, -3\alpha-2\beta$. Then Then γ must make an angle strictly less than $30°$ with one of $\alpha, \beta, \alpha+\beta, 2\alpha+\beta, 3\alpha+\beta, 3\alpha+2\beta, -\alpha, -\beta, -\alpha-\beta, -2\alpha-\beta, -3\alpha-\beta, -3\alpha-2\beta$. This is impossible by Lemma 8.3.3.

Finally, assume that $\theta = 90°$. Assume that R contains a root γ other than $\alpha, \beta, -\alpha, -\beta$. Arguing as in the first paragraph, one can show that the set S contains a pair with θ larger than $90°$; this is a contradiction. Thus, R is as in 4. □

8.4 Bases

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. Let B be a subset of R. We say that B is a **base** for R if

(B1) B is a basis for the $\mathbb{R}$ vector space V.

(B2) Every element $\alpha \in R$ can be written in the form

$$\alpha = \sum_{\beta \in B} c(\beta)\beta$$

where the coefficients $c(\beta)$ for $\beta \in B$ are all integers of the same sign (i.e., either all greater than or equal to zero, or all less than or equal to zero).

Assume that B is a base for R. We define

$$R^+ = \left\{ \alpha \in R : \begin{array}{l} \alpha \text{ is a linear combination of } \beta \in B \\ \text{with non-negative coefficients} \end{array} \right\},$$

$$R^- = \left\{ \alpha \in R : \begin{array}{c} \alpha \text{ is a linear combination of } \beta \in B \\ \text{with non-positive coefficients} \end{array} \right\}.$$

We have

$$R = R^+ \sqcup R^-.$$

We refer to R^+ as the set of **positive roots** with respect to B and R^- as the set of **negative roots** with respect to B. If $\alpha \in R$ is written as in (B2), then we define the **height** of α to be the integer

$$\text{ht}(\alpha) = \sum_{\beta \in B} c(\beta).$$

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let R be a root system in V. Let $v \in V$ be non-zero. We will say that v is **regular** with respect to R if $(v, \alpha) \neq 0$ for all $\alpha \in R$, i.e., if v does not lie on any of the hyperplanes

$$P_\alpha = \{x \in V : (x, \alpha) = 0\}$$

for $\alpha \in R$. If v is not regular, then we say that v is **singular** with respect to R. Evidently, v is regular with respect to R if and only if

$$v \in V - \cup_{\alpha \in R} P_\alpha.$$

We denote by V_{reg} the set of all vectors in V that are regular with respect to R, so that

$$V_{\text{reg}}(R) = V - \cup_{\alpha \in R} P_\alpha.$$

Evidently, $V_{\text{reg}}(R)$ is an open subset of V; however, it is not entirely obvious that $V_{\text{reg}}(R)$ is non-empty.

Lemma 8.4.1. *Let V be a finite-dimensional vector space over $\mathbb{R}$, and let $U_1, \dots, U_n$ be proper subspaces of V. Define $U = \cup_{i=1}^n U_i$. If U_i is a proper subset of U for all $i \in \{1, \dots, n\}$, then U is not a subspace of V.*

Proof. Assume that U_i is a proper subset of U for all $i \in \{1, \dots, n\}$. Since U_i is a proper subset of U for all $i \in \{1, \dots, n\}$ we must have $n \geq 2$. After replacing the collection of U_i for $i \in \{1, \dots, n\}$ with a subcollection, we may assume that $U_i \not\subseteq U_j$ and $U_j \not\subseteq U_i$ for $i, j \in \{1, \dots, n\}$, $i \neq j$. We will prove that U is not a subspace for collections of proper subspaces $U_1, \dots, U_n$ with $n \geq 2$ and such that that $U_i \not\subseteq U_j$ and $U_j \not\subseteq U_i$ for $i, j \in \{1, \dots, n\}$ by induction on n. Assume that $n = 2$ and that $U = U_1 \cup U_2$ is a subspace; we will obtain a contradiction. Since $U_1 \not\subseteq U_2$ and $U_2 \not\subseteq U_1$, there exist $u_2 \in U_2$ such that $u_2 \notin U_1$ and $u_1 \in U_1$ such that $u_1 \notin U_2$. Since U is a subspace we have $u_1 + u_2 \in U$. Hence, $u_1 + u_2 \in U_1$ or $u_1 + u_2 \in U_2$. If $u_1 + u_2 \in U_1$, then $u_2 \in U_1$, a contradiction; similary, if $u_1 + u_2 \in U_2$, then $u_1 \in U_2$, a contradiction. Thus, the claim holds if $n = 2$.

Suppose that $n \geq 3$ and that the claim holds for $n - 1$; we will prove that the claim holds for n. We argue by contradiction; assume that U is a subspace.

We first note that $U_1 \not\subseteq \cup_{i=1,i\neq 1}^n U_i$; otherwise, $U = \cup_{i=1,i\neq 1}^n U_i$ is a subspace, contradicting the induction hypothesis. Similarly, $U_2 \not\subseteq \cup_{i=1,i\neq 2}^n U_i$. Let $u_1 \in U_1$ be such that $u_1 \notin \cup_{i=1,i\neq 1}^n U_i$ and let $u_2 \in U_2$ be such that $u_2 \notin \cup_{i=1,i\neq 2}^n U_i$. Let $\lambda_1, \ldots, \lambda_{n-1}$ be distinct non-zero elements of $\mathbb{R}$. The $n-1$ vectors

$$u_1 + \lambda_1 u_2, \quad u_1 + \lambda_2 u_2, \quad \ldots, \quad u_1 + \lambda_n u_2$$

are all contained in U, and hence must each lie in some U_i with $i \in \{1, \ldots, n\}$. However, no such vector can be in U_1 because otherwise $u_2 \in U_1$; similarly, no such vector can be in U_2. By the pigeonhole principle, this means that there exist distinct $j, k \in \{2, \ldots, n\}$ and $i \in \{3, \ldots, n\}$ such that $u_1 + \lambda_j u_2, u_1 + \lambda_k u_2 \in U_i$. It follows that $(\lambda_j - \lambda_k)u_2 \in U_i$, so that $u_2 \in U_i$. This is a contradiction. □

Lemma 8.4.2. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. Assume that $\dim V \geq 2$. There exists a $v \in V$ such that v is regular with respect to R, i.e., $V_{\mathrm{reg}}(R)$ is non-empty.*

Proof. Assume that there exists no $v \in V$ such that v is regular with respect to R; we will obtain a contradiction. Since no regular $v \in V$ exists, we have $V = \cup_{\alpha \in R} P_\alpha$. Since $\dim V \geq 2$, and since R contains a basis for V over $\mathbb{R}$, it follows that $\#R \geq 2$. Also, $\dim P_\alpha = \dim V - 1$ for all $\alpha \in R$. We now have a contradiction by Lemma 8.4.1. □

Assume that v is regular with respect to R. As we have mentioned before, v can be used to divide V into three components:

$$\begin{aligned}
&\{x \in V : (x,v) = 0\} : \text{the hyperplane of vectors orthogonal to } v,\\
&\{x \in V : (x,v) > 0\} : \text{the vectors that form a strictly acute angle with } v,\\
&\{x \in V : (x,v) < 0\} : \text{the vectors that form a strictly obtuse angle with } v.
\end{aligned}$$

We will write

$$\begin{aligned}
R^+(v) &= \{\alpha \in R : (\alpha, v) > 0\},\\
R^-(v) &= \{\alpha \in R : (\alpha, v) < 0\}.
\end{aligned}$$

Evidently,

$$R = R^+(v) \sqcup R^-(v).$$

Let $\alpha \in R^+(v)$. We will say that α is **decomposable** if $\alpha = \beta_1 + \beta_2$ for some $\beta_1, \beta_2 \in R^+(v)$. If α is not decomposable we will say that α is **indecomposable**. We define

$$B(v) = \{\alpha \in R^+(v) : \alpha \text{ is indecomposable}\}.$$

Lemma 8.4.3. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. Let $v \in V$ be regular with respect to R (such a v exists by Lemma 8.4.2). The set $B(v)$ is non-empty.*

Proof. Assume that $B(v)$ is empty; we will obtain a contradiction. Let $\alpha \in R^+$ be such that (v, α) is minimal. Since α is decomposable, there exist $\alpha_1, \alpha_2 \in R^+$ such that $\alpha = \alpha_1 + \alpha_2$. Now

$$(v, \alpha) = (v, \alpha_1) + (v, \alpha_2).$$

By the definition of $R^+(v)$, the real numbers (v, α), (v, α_1), and (v, α_2) are all positive. It follows that we must have $(v, \alpha) > (v, \alpha_1)$. This contradicts the definition of α. □

Lemma 8.4.4. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let R be a root system in V. Let $v \in V$ be regular with respect to R (such a v exists by Lemma 8.4.2). If $\alpha, \beta \in B(v)$ with $\alpha \neq \beta$, then the angle between α and β is obtuse, i.e., $(\alpha, \beta) \leq 0$.*

Proof. Assume that the angle between α and β is strictly acute. With out loss of generality, we may assume that $\|\alpha\| \leq \|\beta\|$. Since $(v, \alpha) > 0$ and $(v, \beta) > 0$ we must have $\alpha \neq -\beta$. By Lemma 8.3.4 we have $\gamma = -\alpha + \beta \in R$. Since $\gamma \in R$, we also have $-\gamma \in R$. Since $R = R^+(v) \sqcup R^-(v)$, we have $\gamma \in R^+(v)$ or $-\gamma \in R^+(v)$. Assume that $\gamma \in R^+(v)$. We have $\gamma + \alpha = \beta$ with $\gamma, \alpha \in R^+(v)$. This contradicts the fact that β is indecomposable. Similarly, the assumption that $-\gamma \in R^+(v)$ implies that $\alpha = \gamma + \beta$, contradicting the fact that α is indecomposable. It follows that the angle between α and β is obtuse, i.e., $(\alpha, \beta) \leq 0$. □

Lemma 8.4.5. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$. Let v be a non-zero vector in V, and let $B \subset V$ be a finite set such that $(v, \alpha) > 0$ for all $\alpha \in B$. If $(\alpha, \beta) \leq 0$ for all $\alpha, \beta \in B$, then the set B is linear independent.*

Proof. Assume that $c(\alpha)$ for $\alpha \in B$ are real numbers such that

$$0 = \sum_{\alpha \in B} c(\alpha)\alpha.$$

We need to prove that $c(\alpha) = 0$ for all $\alpha \in B$. Suppose that $c(\alpha) \neq 0$ for some $\alpha \in B$; we will obtain a contradiction. Since $c(\alpha) \neq 0$ for some $\alpha \in B$, we may assume that, after possibly multiplying by -1, that there exists $\alpha \in B$ such that $c(\alpha) > 0$. Define

$$x = \sum_{\alpha \in B,\ c(\alpha) > 0} c(\alpha)\alpha.$$

We also have

$$x = \sum_{\beta \in B,\ c(\beta) < 0} (-c(\beta))\beta.$$

Therefore,

$$(x, x) = (\sum_{\alpha \in B,\ c(\alpha) > 0} c(\alpha)\alpha, \sum_{\beta \in B,\ c(\beta) < 0} (-c(\beta))\beta)$$

$$(x,x) = \sum_{\substack{\alpha\in B,\ c(\alpha)>0 \\ \beta\in B,\ c(\beta)<0}} c(\alpha)\cdot(-c(\beta))(\alpha,\beta).$$

By assumption we have $(\alpha,\beta) \leq 0$ for $\alpha,\beta \in B$. Therefore, $(x,x) \leq 0$. This implies that $x = 0$. Now

$$(v,x) = (v, \sum_{\alpha\in B,\ c(\alpha)>0} c(\alpha)\alpha)$$
$$0 = \sum_{\alpha\in B,\ c(\alpha)>0} c(\alpha)(v,\alpha).$$

By the definition of B we have $(v,\alpha) > 0$ for all $\alpha \in B$. The last displayed equation now yields a contradiction since the set of $\alpha \in B$ such that $c(\alpha) > 0$ is non-empty. □

Proposition 8.4.6. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. Let $v \in V$ be regular with respect to R (such a v exists by Lemma 8.4.2). The set $B(v)$ is a base for R, and the set of positive roots with respect to $B(v)$ is $R^+(v)$.*

Proof. We will begin by proving that (B2) holds. Evidently, since $R^-(v) = -R^+(v)$, to prove that (B2) holds it suffices to prove that every $\beta \in R^+(v)$ can be written as

$$\beta = \sum_{\alpha\in B(v)} c(\alpha)\alpha, \qquad c(\alpha) \in \mathbb{Z}_{\geq 0}. \tag{8.1}$$

Let S be the set of $\beta \in R^+(v)$ for which (8.1) does not hold. We need to prove that S is empty. Suppose that S is not empty; we will obtain a contradiction. Let $\beta \in S$ be such that (v,β) is minimal. Clearly, $\beta \notin B(v)$, i.e., β is decomposable. Let $\beta_1, \beta_2 \in R^+(v)$ be such that $\beta = \beta_1 + \beta_2$. We have

$$(v,\beta) = (v,\beta_1) + (v,\beta_2).$$

By the definition of $R^+(v)$, the real numbers (v,β), (v,β_1), and (v,β_2) are all positive. It follows that we must have $(v,\beta) > (v,\beta_1)$ and $(v,\beta) > (v,\beta_2)$. The definition of β implies that $\beta_1 \notin S$ and $\beta_2 \notin S$. Hence, β_1 and β_2 have expressions as in (8.1). It follows that $\beta = \beta_1 + \beta_2$ has an expression as in (8.1). This contradiction implies that (B2) holds.

Now we prove that $B(v)$ satisfies (B1). Since R spans V, and since every element of R is a linear combination of elements of $B(v)$ because $B(v)$ satisfies (B2), it follows that $B(v)$ spans V. Finally, $B(v)$ is linearly independent by Lemma 8.4.4 and Lemma 8.4.5. □

Lemma 8.4.7. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$. Let $v_1,\dots,v_n$ be a basis for V. There exists a vector $v \in V$ such that $(v,v_1) > 0,\dots,(v,v_n) > 0$.*

Proof. Let $i \in \{1, \dots, n\}$. The subspace V_i of V spanned by $\{v_1, \dots, v_n\} - \{v_i\}$ has dimension $n - 1$. It follows that the orthogonal complement $V_i^{\perp}$ is one-dimensional. Let $w_i \in V$ be such that $V_i^{\perp} = \mathbb{R}w_i$. Evidently, by construction we have $(w_i, v_j) = 0$ for $j \in \{1, \dots, n\}$, $j \neq i$. This implies that $(w_i, v_i) \neq 0$; otherwise, w_i is orthogonal to every element of V, contradicting the fact that $w_i \neq 0$. After possibly replacing w_i with $-w_i$, we may assume that $(w_i, v_i) > 0$. Consider the vector

$$v = w_1 + \cdots + w_n.$$

Let $i \in \{1, \dots, n\}$. Then

$$(v, v_i) = (w_1 + \cdots + w_n, v_i) = (w_i, v_i) > 0.$$

It follows that v is the desired vector. □

Lemma 8.4.8. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, let R be a root system in V, let B be a base for R, and let R^+ be the positive roots in R with respect to B. Let $v \in V$ be regular with respect R, and assume that $R^+(v) = R^+$. Then $B(v) = B$.*

Proof. Let $\beta \in B$. By the assumption $R^+(v) = R^+$ we have $\beta \in R^+(v)$. We claim that β is indecomposable as an element of $R^+(v)$. Suppose not; we will obtain a contradiction. Since β is decomposable there exist $\beta_1, \beta_2 \in R^+(v)$ such that $\beta = \beta_1 + \beta_2$. As $R^+(v) = R^+$ and B is a base for R, we can write

$$\beta_1 = \sum_{\alpha \in B} c_1(\alpha)\alpha,$$
$$\beta_2 = \sum_{\alpha \in B} c_2(\alpha)\alpha$$

for some non-negative integers $c_1(\alpha), c_2(\alpha)$, $\alpha \in B$. This implies that

$$\beta = \sum_{\alpha \in B} \big(c_1(\alpha) + c_2(\alpha)\big)\alpha.$$

Since B is a basis for V and $\beta \in B$, we obtain $c_2(\alpha) = -c_1(\alpha)$ for $\alpha \in B$, $\alpha \neq \beta$, and $c_2(\beta) = 1 - c_1(\alpha)$. As $c_1(\alpha)$ and $c_2(\alpha)$ are both non-negative for $\alpha \in B$, we get $c_1(\alpha) = c_2(\alpha) = 0$ for $\alpha \in B$, $\alpha \neq \beta$. Also, since $c_2(\beta) = 1 - c_1(\beta)$ is a non-negative integer, we must have $1 \geq c_1(\beta)$; since $c_1(\beta)$ is a non-negative integer, this implies that $c_1(\beta) = 0$ or $c_1(\beta) = 1$. If $c_1(\beta) = 0$, then $\beta_1 = 0$, a contraction. If $c_1(\beta) = 1$, then $c_2(\beta) = 0$ so that $\beta_2 = 0$; this is also a contradiction. It follows that β is indecomposable with respect to v. Therefore, $B \subset B(v)$. Since $\#B = \dim V = \#B(v)$, we obtain $B = B(v)$. □

Proposition 8.4.9. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let R be a root system in V. If B is a base for R, then there exists a vector $v \in V$ that is regular with respect to R and such that $B = B(v)$.*

Proof. By Lemma 8.4.7 there exists a vector $v \in V$ such that $(v, \alpha) > 0$ for $\alpha \in B$. We claim that v is regular with respect to R. Let $\beta \in R$, and write

$$\beta = \sum_{\alpha \in B} c(\alpha)\alpha,$$

where the coefficients $c(\alpha)$ for $\alpha \in B$ are integers of the same sign. We have

$$\begin{aligned}(v, \beta) &= (v, \sum_{\alpha \in B} c(\alpha)\alpha) \\ &= \sum_{\alpha \in B} c(\alpha)(v, \alpha).\end{aligned}$$

Since all the coefficients $c(\alpha)$, $\alpha \in B$, have the same sign, and since $(v, \alpha) > 0$ for $\alpha \in B$, it follows that $(v, \beta) > 0$ or $(v, \beta) < 0$. Thus, v is regular with respect to R. Next, since $(v, \alpha) > 0$ for $\alpha \in B$, we have $R^+ \subset R^+(v)$ and $R^- \subset R^-(v)$. Since $R = R^+ \sqcup R^-$ and $R = R^+(v) \sqcup R^-(v)$ we now have $R^+ = R^+(v)$ and $R^- = R^-(v)$. We now have $B = B(v)$ by Lemma 8.4.8. □

8.5 Weyl chambers

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let R be a root system in V. We recall that each root $\alpha \in R$ defines a hyperplane

$$P_\alpha = \{x \in V : (x, \alpha) = 0\}.$$

Also, recall that a vector $v \in V$ is regular with respect to R if and only if

$$v \in V_{\mathrm{reg}}(R) = V - \cup_{\alpha \in R} P_\alpha.$$

Evidently, $V_{\mathrm{reg}}(R)$ is an open subset of V. A path component of the space $V_{\mathrm{reg}}(R)$ is called a **Weyl chamber** of V with respect to R.

Lemma 8.5.1. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let R be a root system in V. Let $v \in V$ be regular with respect R. Let C be the Weyl chamber of V with respect to R that contains the vector v. Then*

$$C = X(v)$$

where

$$X(v) = \{w \in V : (w, \alpha) > 0, \alpha \in B(v)\}.$$

Proof. We need prove that $X(v) \subset V_{\mathrm{reg}}(R)$, $v \in X(v)$, and that $X(v)$ is exactly the set of $w \in V_{\mathrm{reg}}(R)$ that are path connected in $V_{\mathrm{reg}}(R)$ to v.

To see that $X(v) \subset V_{\mathrm{reg}}(R)$ let $w \in X(v)$. To prove that $w \in V_{\mathrm{reg}}(R)$ it suffices to prove that $(w, \beta) > 0$ for all $\beta \in R^+(v)$; this follows from the definition of $X(v)$ and the fact that $B(v)$ is a base for R such that $R^+(v)$ is the set of the positive roots with respect to $B(v)$. Thus, $X(v) \subset V_{\mathrm{reg}}(R)$.

By the definition of $B(v)$ we have $B(v) \subset R^+(v)$. It follows that $v \in X(v)$.

Next, we show that every element of X is path connected in $V_{\mathrm{reg}}(R)$ to v. Let $w \in X(v)$. Define $f : [0,1] \to V_{\mathrm{reg}}(R)$ by $f(t) = (1-t)v + tw$ for $t \in [0,1]$. To see that f is well-defined, let $t \in [0,1]$ and $\beta \in R$. We need to verify that $(f(t), \beta) \neq 0$. We may assume that $\beta \in R^+(v)$. Write

$$\beta = \sum_{\alpha \in B(v)} c(\alpha)\alpha, \quad c(\alpha) \in \mathbb{Z}_{\geq 0}.$$

We have

$$\begin{aligned}(f(t), \beta) &= \big((1-t)v + tw, \sum_{\alpha \in B(v)} c(\alpha)\alpha\big) \\ &= (1-t) \sum_{\alpha \in B(v)} c(\alpha)\big(v, \alpha\big) + t \sum_{\alpha \in B(v)} c(\alpha)\big(w, \alpha\big).\end{aligned}$$

Since $(v, \alpha), (w, \alpha) > 0$ for $\alpha \in B(v)$ it follows that $(f(t), \beta) > 0$; thus, the image of f is indeed in $V_{\mathrm{reg}}(R)$, so that f is well-defined. Evidently, f is continuous, and $f(0) = v$ and $f(1) = w$. It follows that every element of X is path connected in $V_{\mathrm{reg}}(R)$ to v.

Finally, we prove that if $u \in V_{\mathrm{reg}}(R)$ and $u \notin X(v)$, then u is not path connected in $V_{\mathrm{reg}}(R)$ to v. Suppose that $u \in V_{\mathrm{reg}}(R)$, $u \notin X(v)$, and that u is path connected in $V_{\mathrm{reg}}(R)$ to v; we will obtain a contradiction. Since u is path connected in $V_{\mathrm{reg}}(R)$ to v there exists a continuous function $g : [0,1] \to V_{\mathrm{reg}}(R)$ such that $g(0) = v$ and $g(1) = u$. Since $u \notin X(v)$, there exists $\alpha \in B(v)$ such that $(u, \alpha) < 0$. Define $F : [0,1] \to \mathbb{R}$ by $F(t) = (g(t), \alpha)$ for $t \in [0,1]$. We have $F(0) > 0$ and $F(1) < 0$. Since F is continuous, there exists a $t \in (0,1)$ such that $F(t) = 0$. This means that $(g(t), \alpha) = 0$. However, this is a contradiction since $g(t)$ is regular with respect to R. □

Proposition 8.5.2. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let R be a root system in V. The map*

$$\begin{matrix}\text{Weyl chambers in } V \\ \text{with respect to } R\end{matrix} \xrightarrow{\sim} \text{Bases for } R$$

that sends a Weyl chamber C to $B(v)$, where v is any element of C, is a well-defined bijection.

Proof. Let C be a Weyl chamber in V with respect to R, and let $v_1, v_2 \in C$. To prove that the map is well-defined it will suffice to prove that $B(v_1) = B(v_2)$. Let $\alpha \in B(v_1)$. By Lemma 8.5.1, since v_1 and v_2 lie in the same Weyl chamber C, we have $C = X(v_1) = X(v_2)$. This implies that $(v_2, \gamma) > 0$ for $\gamma \in B(v_1)$. In particular, we have $(v_2, \alpha) > 0$. Now let $\beta \in R^+(v_1)$. Write

$$\beta = \sum_{\alpha \in B(v_1)} c(\alpha)\alpha, \quad c(\alpha) \in \mathbb{Z}_{\geq 0}.$$

Then

$$(v_2, \beta) = \sum_{\alpha \in B(v_1)} c(\alpha)(v_2, \alpha).$$

Since $(v_2, \alpha) > 0$ for all $\alpha \in B(v_1)$ we must have $(v_2, \beta) > 0$. Thus, $R^+(v_1) \subset R^+(v_2)$. Similarly, $R^+(v_2) \subset R^+(v_1)$, so that $R^+(v_1) = R^+(v_2)$. We now obtain $B(v_1) = B(v_2)$ by Lemma 8.4.8.

To see that the map is injective, suppose that C_1 and C_2 are Weyl chambers that map to the same base for R. Let $v_1 \in C_1$ and $v_2 \in C_2$. By assumption, we have $B(v_1) = B(v_2)$. Since $B(v_1) = B(v_2)$ we have $X(v_1) = X(v_2)$. By Lemma 8.5.1, this implies that $C_1 = C_2$.

Finally, the map is surjective by Proposition 8.4.9. □

Lemma 8.5.3. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let R be a root system in V. Let C be a Weyl chamber of V with respect to R, and let B be the base of R that corresponds to C, as in Proposition 8.5.2, so that*

$$C = \{w \in V : (w, \alpha) > 0 \text{ for all } \alpha \in B\}.$$

The closure $\bar{C}$ of C is:

$$\bar{C} = \{w \in V : (w, \alpha) \geq 0 \text{ for all } \alpha \in B\}.$$

Every element of V is contained $\bar{C}$ for some Weyl chamber C of V in R.

Proof. The closure of C consists of C and points $w \in V$ with $w \notin C$ such that there exists a sequence $(w_n)_{n=1}^{\infty}$ of elements of C such that $w_n \to w$ as $n \to \infty$. Let w be an element of $\bar{C}$ of the this second type. Assume that there exists $\alpha \in B$ such that $(w, \alpha) < 0$. Since $(w_n, \alpha) \to (w, \alpha)$ as $n \to \infty$, there exists a positive integer n such that $(w_n, \alpha) < 0$. This is a contradiction. It follows that $\bar{C}$ is contained in $\{w \in V : (w, \alpha) \geq 0 \text{ for all } \alpha \in B\}$. Let w be in $\{w \in V : (w, \alpha) \geq 0 \text{ for all } \alpha \in B\}$; we need to prove that $w \in \bar{C}$. Let $w_0 \in C$. Consider the sequence $(w + (1/n)w_0)_{n=1}^{\infty}$. Evidently this sequence converges to w and is contained in C. It follows that w is in $\bar{C}$. This proves the first assertion of the lemma. For the second assertion, let $v \in V$. If $v \in V_{\text{reg}}(R)$, then v is by definition in some Weyl chamber. Assume that $v \notin V_{\text{reg}}(R)$. Then $v \in \cup_{\alpha \in R} P_\alpha$. Define

$$p : V \longrightarrow \mathbb{R} \quad \text{by} \quad p(x) = \prod_{\alpha \in R} (x, \alpha).$$

The function p is a non-zero polynomial function on V, and the set of zeros of p is exactly $\cup_{\alpha \in R} P_\alpha$. Thus, $p(v) = 0$. Since p is a non-zero polynomial function on V, p cannot vanish on an open set. Hence, for each positive integer n, there exists v_n such that $\|v - v_n\| < 1/n$ and $p(v_n) \neq 0$. The sequence $(v_n)_{n=1}^{\infty}$ converges to v and is contained in $V_{\text{reg}}(R)$; in particular every element of the sequence is contained in some Weyl chamber. Since the number of Weyl chambers of V with respect to R is finite by Proposition 8.5.2, it follows that there is a subsequence $(v_{n_k})_{k=1}^{\infty}$ of $(v_n)_{n=1}^{\infty}$ the elements of which are completely contained in one Weyl chamber C. Let C correspond to the base B for R. We have $(v_{n_k}, \alpha) \geq 0$ for all $\alpha \in B$ and positive integers k. Taking limits, we find that $(v, \alpha) \geq 0$ for all $\alpha \in B$, so that $v \in \bar{C}$. □

8.6 More facts about roots

Lemma 8.6.1. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$. Let $\alpha \in V$ be non-zero, let A be an open subset of V, and let $v \in A$ be such that $(v, \alpha) = 0$. Then there exists $w \in A$ such that $(w, \alpha) > 0$.*

Proof. Let $e_1, \dots, e_n$ be the standard basis for V. Write $\alpha = a_1 e_1 + \cdots + a_n e_n$ for some $a_1, \dots, a_n \in \mathbb{R}$, and $v = v_1 e_1 + \cdots + v_n e_n$ for some $v_1, \dots, v_n \in \mathbb{R}$. Since $\alpha \neq 0$, there exists $i \in \{1, \dots, n\}$ such that $a_i \neq 0$. Let $\epsilon \in \mathbb{R}$. Define $w = v + (\epsilon/a_i)e_i$. For sufficiently small ϵ we have $w \in A$ and

$$(w, \alpha) = (v + (\epsilon/a_i)e_i, \alpha) = (v, \alpha) + \epsilon = \epsilon > 0.$$

This completes the proof. □

Lemma 8.6.2. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let R be a root system in V. Let $\alpha \in R$. There exists a base B for R that contains α.*

Proof. We first claim that

$$P_\alpha \not\subseteq \bigcup_{\beta \in R, \beta \neq \pm\alpha} P_\beta.$$

Suppose this is false; we will obtain a contradiction. Since P_α is contained in the union of the sets P_β, $\beta \in R$, $\beta \neq \pm\alpha$, we have

$$P_\alpha = \bigcup_{\beta \in R, \beta \neq \pm\alpha} (P_\alpha \cap P_\beta).$$

By Lemma 8.4.1, as P_α is a subspace of V, there exists $\beta \in R$, $\beta \neq \pm\alpha$, such that $P_\alpha = P_\alpha \cap P_\beta$. This implies that $P_\alpha = P_\beta$; taking orthogonal complements, this implies that $\mathbb{R}\alpha = \mathbb{R}\beta$, a contradiction. Since P_α is not contained in $\bigcup_{\beta \in R, \beta \neq \pm\alpha} P_\beta$, there exists a vector $v \in P_\alpha$ such that $v \notin \bigcup_{\beta \in R, \beta \neq \pm\alpha} P_\beta$. Define a function

$$f : V \to \mathbb{R} \oplus \bigoplus_{\beta \in R, \beta \neq \pm\alpha} \mathbb{R}$$

by

$$f(w) = (w, \alpha) \oplus \bigoplus_{\beta \in R, \beta \neq \pm\alpha} (|(w, \beta)| - |(w, \alpha)|).$$

This function is continuous, and we have

$$f(v) = 0 \oplus \bigoplus_{\beta \in R, \beta \neq \pm\alpha} |(v, \beta)|$$

with $|(v,\beta)| > 0$ for $\beta \in R, \beta \neq \pm\alpha$. Fix $\epsilon > 0$ be such that $|(v,\beta)| > \epsilon > 0$ for $\beta \in R, \beta \neq \pm\alpha$. Since f is continuous, there exists an open set A containing v such that

$$f(A) \subset (-\epsilon, \epsilon) \oplus \bigoplus_{\beta \in R, \beta \neq \pm\alpha} (|(v,\beta)| - \epsilon, |(v,\beta)| + \epsilon).$$

Moreover, by Lemma there exists $w \in A$ such that $(w, \alpha) > 0$. Let $\beta \in R$, $\beta \neq \pm\alpha$. Since $w \in A$, we have

$$0 < |(v,\beta)| - \epsilon < |(w,\beta)| - |(w,\alpha)| = |(w,\beta)| - (w,\alpha)$$

so that

$$(w,\alpha) < |(w,\beta)|.$$

Consider now the base $B(w)$. We claim that $\alpha \in B(w)$. We have $(w,\alpha) > 0$, so that $\alpha \in R^+(w)$. Assume that $\alpha = \beta_1 + \beta_2$ for some $\beta_1, \beta_2 \in R^+(w)$; we obtain a contradiction, proving that $\alpha \in B(w)$. We must have $\beta_1 \neq \pm\alpha_1$ and $\beta_2 \neq \pm\alpha$; otherwise, $0 \in R$ or $2\alpha \in R$, a contradiction. Now

$$(w,\alpha) = (w,\beta_1) + (w,\beta_2).$$

Since $(w,\beta_1) > 0$ and $(w,\beta_2) > 0$ we must have $(w,\alpha) > (w,\beta_1)$. This contradicts $(w,\alpha) < |(w,\beta_1)| = (w,\beta_1)$. □

Lemma 8.6.3. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. Let B be a base for R. Let α be a positive root with respect to B such that $\alpha \notin B$. Then there exists $\beta \in B$ such that $(\alpha,\beta) > 0$ and $\alpha - \beta$ is a positive root.*

Proof. By Proposition 8.4.9 there exists $v \in V_{\mathrm{reg}}(R)$ such that $B = B(v)$. Since α and the elements of B are all in $R^+ = R^+(v)$ (see Proposition 8.4.6) we have $(v,\alpha) > 0$ and $(v,\beta) > 0$ for $\beta \in B$. If $(\alpha,\beta) \leq 0$ for all $\beta \in B$, then by Lemma 8.4.4 Lemma 8.4.5, the set $B \sqcup \{\alpha\}$ is linearly independent, contradicting the fact that B is a basis for the $\mathbb{R}$ vector space V. It follows that there exists $\beta \in B$ such that $(\alpha,\beta) > 0$. By Lemma 8.3.4 we have $\alpha - \beta \in R$. Since α is positive we can write

$$\alpha = c(\beta)\beta + \sum_{\gamma \in B, \gamma \neq \beta} c(\gamma)\gamma$$

with $c(\beta) \geq 0$ and $c(\gamma) \geq 0$ for $\gamma \in B, \gamma \neq \beta$. Since $\alpha \notin B$, we must have $c(\gamma) > 0$ for some $\gamma \in B$ with $\gamma \neq \beta$, or $c(\beta) \geq 2$. Since

$$\alpha - \beta = (c(\beta) - 1)\beta + \sum_{\gamma \in B, \gamma \neq \beta} c(\gamma)\gamma$$

we see that $\alpha - \beta$ is positive. □

Lemma 8.6.4. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let R be a root system in V. Let B be a base for R. If $\alpha \in R^+$, then there exist (not necessarily distinct) $\alpha_1, \dots, \alpha_t \in B$ such that $\alpha = \alpha_1 + \cdots + \alpha_t$, and the partial sums*

$$\begin{aligned}&\alpha_1,\\&\alpha_1 + \alpha_2,\\&\alpha_2 + \alpha_2 + \alpha_3,\\&\cdots\\&\alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_t\end{aligned}$$

are all positive roots.

Proof. We will prove this by induction on $\mathrm{ht}(\alpha)$. If $\mathrm{ht}(\alpha) = 1$ this is clear. Assume that $\mathrm{ht}(\alpha) > 1$ and that the lemma holds for all positive roots γ with $\mathrm{ht}(\gamma) < \mathrm{ht}(\alpha)$. We will prove that the lemma holds for α. If $\alpha \in B$, then $\mathrm{ht}(\alpha) = 1$, contradicting our assumption that $\mathrm{ht}(\alpha) > 1$. Thus, $\alpha \notin B$. By Lemma 8.6.3 there exists $\beta \in B$ such that $\alpha - \beta$ is a positive root. Now $\mathrm{ht}(\alpha - \beta) = \mathrm{ht}(\alpha) - 1$. By the induction hypothesis, the lemma holds for $\alpha - \beta$; let $\alpha_1, \dots, \alpha_t \in B$ be such that $\alpha - \beta = \alpha_1 + \cdots + \alpha_t$, and the partial sums

$$\begin{aligned}&\alpha_1,\\&\alpha_1 + \alpha_2,\\&\alpha_2 + \alpha_2 + \alpha_3,\\&\cdots\\&\alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_t\end{aligned}$$

are all positive roots. Since $\alpha = \alpha_1 + \cdots + \alpha_t + \beta$, the lemma holds for α. □

Lemma 8.6.5. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let R be a root system in V. Let B be a base for R. Let $\alpha \in B$. The reflection s_α maps $R^+ - \{\alpha\}$ onto $R^+ - \{\alpha\}$.*

Proof. Let $\beta \in R^+ - \{\alpha\}$. Write

$$\beta = \sum_{\gamma \in B} c(\gamma)\gamma$$

with $c(\gamma) \in \mathbb{Z}_{\geq 0}$ for $\gamma \in B$. We claim that $c(\gamma_0) > 0$ for some $\gamma_0 \in B$ with $\gamma_0 \neq \alpha$. Suppose this is false, so that $\beta = c(\alpha)\alpha$; we will obtain a contradiction. By (R2), we have $c(\alpha) = \pm 1$. By hypothesis, $\alpha \neq \beta$; hence, $c(\alpha) = -1$, so that $\beta = -\alpha$. This contradicts the fact that β is positive, proving our claim. Now

$$\begin{aligned}s_\alpha(\beta) &= \beta - \langle \alpha, \beta \rangle \alpha\\&= (c(\alpha) - \langle \alpha, \beta \rangle)\alpha + \sum_{\gamma \in B, \gamma \neq \alpha} c(\gamma)\gamma.\end{aligned}$$

This is the expression of the root $s_\alpha(\beta)$ in terms of the base B. Since $c(\gamma_0) > 0$, we see that $s_\alpha(\beta)$ is a positive root and that $s_\alpha(\beta) \neq \alpha$, i.e, $s_\alpha(\beta) \in R^+ - \{\alpha\}$. □

Lemma 8.6.6. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. Let B be a base for R. Set*

$$\delta = \frac{1}{2}\sum_{\beta\in R^+}\beta.$$

If $\alpha \in B$, then

$$s_\alpha(\delta) = \delta - \alpha.$$

Proof. We have

$$\begin{aligned}
s_\alpha(\delta) &= s_\alpha(\frac{1}{2}\alpha) + s_\alpha(\delta - \frac{1}{2}\alpha)\\
&= -\frac{1}{2}\alpha + \frac{1}{2}\sum_{\beta\in R^+-\{\alpha\}} s_\alpha(\beta)\\
&= -\frac{1}{2}\alpha + \frac{1}{2}\sum_{\beta\in R^+-\{\alpha\}} \beta\\
&= -\frac{1}{2}\alpha - \frac{1}{2}\alpha + \frac{1}{2}\sum_{\beta\in R^+} \beta\\
&= -\alpha + \delta.
\end{aligned}$$

This completes the proof. □

8.7 The Weyl group

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. We define the **Weyl group** of R to be the subgroup $\mathcal{W}$ of $\mathrm{O}(V)$ generated by the reflections s_α for $\alpha \in R$.

Lemma 8.7.1. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. The Weyl group $\mathcal{W}$ of R is finite.*

Proof. Define a map

$$\mathcal{W} \longrightarrow \text{The group of permutations of } R$$

by sending w to the permutation that sends $\alpha \in R$ to $w(r)$. By (R3), this map is well-defined. This map is a homomorphism because the group law for both groups is composition of functions. Assume that $w \in \mathcal{W}$ maps to the identity. Then $w(\alpha) = \alpha$ for all $\alpha \in R$. Since R contains a basis for the vector space V, this implies that w is the identity; hence, the map is injective. It follows to that $\mathcal{W}$ is finite. □

Lemma 8.7.2. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$. Let X be a finite subset of V consisting of non-zero vectors that span V. Assume that for every $\alpha \in X$, the reflection s_α maps X into X. Let $s \in \mathrm{GL}(V)$. Assume that $s(X) = X$, that there is a hyperplane P of V that s fixes pointwise, and that for some $\alpha \in X$, $s(\alpha) = -\alpha$. Then $s = s_\alpha$ and $P = P_\alpha$.*

Proof. Let $t = ss_\alpha^{-1} = ss_\alpha$. We have

$$t(\alpha) = s(s_\alpha(\alpha)) = s(-\alpha) = -(-\alpha) = \alpha.$$

We must have $\mathbb{R}\alpha \cap P = 0$; otherwise, $\alpha \in P$, and so $s(\alpha) = \alpha$, a contradiction. Therefore,

$$V = \mathbb{R}\alpha \oplus P.$$

On the other hand, by the definition of $P_\alpha = (\mathbb{R}\alpha)^\perp$, we also have

$$V = \mathbb{R}\alpha \oplus P_\alpha.$$

It follows that the image of P under the projection map $V \to V/\mathbb{R}\alpha$ is all of $V/\mathbb{R}\alpha$; similarly, the image of P_α under $V \to V/\mathbb{R}\alpha$ is all of $V/\mathbb{R}\alpha$. Since s fixes P pointwise, it follows that the endomorphism of $V/\mathbb{R}\alpha$ induced by s is the identity. Similarly, the endomorphism of $V/\mathbb{R}\alpha$ induced by s_α is the identity. Therefore, the endomorphism of $V/\mathbb{R}\alpha$ induced by $t = ss_\alpha$ is also the identity. Let $v \in V$. We then have $t(v) = v + a\alpha$ for some $a \in \mathbb{R}$. Applying t again, we obtain $t^2(v) = t(v)+a\alpha$. Solving this last equation for $a\alpha$ gives $a\alpha = t^2(v)-t(v)$. Substituting into the first equation yields:

$$\begin{aligned} t(v) &= v + t^2(v) - t(v) \\ 0 &= t^2(v) - 2t(v) + v. \end{aligned}$$

That is, $p(t) = 0$ for $p(z) = z^2 - 2z + 1 = (z-1)^2$. It follows that the minimal polynomial of t divides $(z-1)^2$. On the other hand, s and s_α both send X into X, so that t also sends X into X. Let $\beta \in X$, and consider the sequence

$$\beta, \quad t(\beta), \quad t^2(\beta), \quad \dots \quad .$$

These vectors are contained in X. Since X is finite, these vectors cannot be pairwise distinct. This implies that there exists a positive integer $k(\beta)$ such that $t^{k(\beta)}(\beta) = \beta$. Now define

$$k = \prod_{\beta \in X} k(\beta).$$

We then have $t^k(\beta) = \beta$ for all $\beta \in X$. Since X spans V, it follows that $t^k = 1$. This means that the minimal polynomial of t divides $z^k - 1$. The minimal polynomial of t now divides $(z-1)^2$ and $z^k - 1$; this implies that the minimal polynomial of t is $z - 1$, i.e., $t = 1$. $\square$

Lemma 8.7.3. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let R be a root system in V. Let $s \in \mathrm{GL}(V)$. Assume that $s(R) = R$. Then*

$$ss_\alpha s^{-1} = s_{s(\alpha)}$$

for all $\alpha \in R$, and

$$\langle s(\alpha), s(\beta)\rangle = \langle \alpha, \beta\rangle$$

for all $\alpha, \beta \in R$.

Proof. Let $\alpha \in R$. We consider the element $ss_\alpha s^{-1}$ of $\mathrm{GL}(V)$. Let $\beta \in R$. We have

$$(ss_\alpha s^{-1})(s(\beta)) = (ss_\alpha)(\beta) = s\big(s_\alpha(\beta)\big).$$

This vector is contained in R because $s_\alpha(\beta)$ is contained in R, and s maps R into R. Since $s(R) = R$, it follows that $(ss_\alpha s^{-1})(R) = R$. Let $P = s(P_\alpha)$; we claim that $ss_\alpha s^{-1}$ fixes P pointwise. Let $x \in P$. Write $x = s(y)$ for some $y \in P_\alpha$. We have

$$\begin{aligned}(ss_\alpha s^{-1})(x) &= (ss_\alpha s^{-1})(s(y))\\ &= s\big(s_\alpha(y)\big)\\ &= s\big(y\big)\\ &= x.\end{aligned}$$

It follows that $ss_\alpha s^{-1}$ fixes P pointwise. Also, we have:

$$\begin{aligned}(ss_\alpha s^{-1})(s(\alpha)) &= s\big(s_\alpha)(\alpha)\big)\\ &= s(-\alpha)\\ &= -s(\alpha).\end{aligned}$$

By Lemma 8.7.2 we now have that $ss_\alpha s^{-1} = s_{s(\alpha)}$.

Finally, let $\alpha, \beta \in R$. Since $ss_\alpha s^{-1} = s_{s(\alpha)}$, we obtain:

$$\begin{aligned}(ss_\alpha s^{-1})(\beta) &= s_{s(\alpha)}(\beta)\\ &= \beta - \langle \beta, s(\alpha))\rangle s(\alpha).\end{aligned}$$

On the other hand, we also have:

$$\begin{aligned}(ss_\alpha s^{-1})(\beta) &= s\big(s_\alpha(s^{-1}(\beta))\big)\\ &= s\big(s^{-1}(\beta) - \langle s^{-1}(\beta), \alpha\rangle\alpha\\ &= \beta - \langle s^{-1}(\beta), \alpha\rangle s(\alpha).\end{aligned}$$

Equating, we conclude that $\langle \beta, s(\alpha))\rangle = \langle s^{-1}(\beta), \alpha\rangle$. Since this holds for all $\alpha, \beta \in R$, this implies that $\langle s(\alpha), s(\beta)\rangle = \langle \alpha, \beta\rangle$ for all $\alpha, \beta \in R$ (substitute $s(\alpha)$ for β and β for α). □

Lemma 8.7.4. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, let R be a root system in V, and let B be a base for R. Let $t \geq 2$ be an integer, and let $\alpha_1, \dots, \alpha_t$ be elements of B that are not necessarily distinct. For convenience, write $s_1 = s_{\alpha_1}, \dots, s_t = s_{\alpha_t}$. If the root $(s_1 \cdots s_{t-1})(\alpha_t)$ is negative, then for some integer k with $1 \leq k < t$,*

$$s_1 \cdots s_t = s_1 \cdots s_{k-1} s_{k+1} \cdots s_{t-1}.$$

Proof. Consider the roots

$$\begin{aligned}
\beta_0 &= (s_1 \cdots s_{t-1})(\alpha_t),\\
\beta_1 &= (s_2 \cdots s_{t-1})(\alpha_t),\\
\beta_2 &= (s_3 \cdots s_{t-1})(\alpha_t),\\
&\cdots\\
\beta_{t-2} &= s_{t-1}(\alpha_t),\\
\beta_{t-1} &= \alpha_t.
\end{aligned}$$

We have

$$\begin{aligned}
s_1(\beta_1) &= \beta_0,\\
s_2(\beta_2) &= \beta_1,\\
s_3(\beta_3) &= \beta_2,\\
&\cdots\\
s_{t-1}(\beta_{t-1}) &= \beta_{t-2}.
\end{aligned}$$

We also have that β_0 is negative, and β_{t-1} is positive. Let k be the smallest integer in $\{1, \dots, t-1\}$ such that β_k is positive. Consider $s_k(\beta_k) = \beta_{k-1}$. By the choice of k, $s_k(\beta_k) = \beta_{k-1}$ must be negative. Recalling that $s_k = s_{\alpha_k}$, by Lemma 8.6.5 we must have $\beta_k = \alpha_k$. This means that

$$(s_{k+1} \cdots s_{t-1})(\alpha_t) = \alpha_k.$$

By Lemma 8.7.3,

$$\begin{aligned}
(s_{k+1} \cdots s_{t-1}) s_t (s_{k+1} \cdots s_{t-1})^{-1} &= s_{(s_{k+1} \cdots s_{t-1})(\alpha_t)}\\
s_{k+1} \cdots s_{t-1} s_t s_{t-1} \cdots s_{k+1} &= s_{\alpha_k}\\
s_{k+1} \cdots s_{t-1} s_t s_{t-1} \cdots s_{k+1} &= s_k\\
s_{k+1} \cdots s_{t-1} s_t &= s_k s_{k+1} \cdots s_{t-1}.
\end{aligned}$$

Via the last equality, we get:

$$\begin{aligned}
s_1 \cdots s_t &= (s_1 \cdots s_{k-1}) s_k (s_{k+1} \cdots s_t)\\
&= (s_1 \cdots s_{k-1}) s_k (s_k s_{k+1} \cdots s_{t-1})\\
&= s_1 \cdots s_{k-1} s_{k+1} \cdots s_{t-1}.
\end{aligned}$$

This is the desired result. □

Proposition 8.7.5. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, let R be a root system in V, and let B be a base for R. Let $\mathcal{W}$ be the Weyl group of R. Let $s \in \mathcal{W}$ with $s \neq 1$. Assume that s can be written as a product of s_α for $\alpha \in B$. Let*

$$s = s_{\alpha_1} \cdots s_{\alpha_t}$$

with $\alpha_1, \ldots, \alpha_t \in B$ and $t \geq 1$ as small as possible. Then $s(\alpha_t)$ is negative.

Proof. If $t = 1$ then $s = s_{\alpha_1}$, and $s(\alpha_1) = -\alpha_1$ is negative. We may thus assume that $t \geq 2$. Assume that $s(\alpha_t)$ is positive; we will obtain a contradiction. Now

$$\begin{aligned} s(\alpha_t) &= (s_{\alpha_1} \cdots s_{\alpha_t})(\alpha_t) \\ &= (s_{\alpha_1} \cdots s_{\alpha_{t-1}})\big(s_{\alpha_t}(\alpha_t)\big) \\ &= (s_{\alpha_1} \cdots s_{\alpha_{t-1}})\big(-\alpha_t\big) \\ &= -(s_{\alpha_1} \cdots s_{\alpha_{t-1}})\big(\alpha_t\big). \end{aligned}$$

Since this root is positive, the root $(s_{\alpha_1} \cdots s_{\alpha_{t-1}})\big(\alpha_t\big)$ must be negative. By Lemma 8.7.4, this implies that t is not minimal, a contradiction. □

Theorem 8.7.6. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, let R be a root system in V, and let $\mathcal{W}$ be the Weyl group of R. The Weyl group $\mathcal{W}$ acts on the set of bases for R by sending a base B to $s(B)$ for $s \in \mathcal{W}$, and the Weyl group acts on the set of Weyl chambers of V with respect to R by sending a Weyl chamber C to $s(C)$ for $s \in \mathcal{W}$. These actions are compatible with the bijection*

$$i : \begin{matrix}\text{Weyl chambers in } V \\ \text{with respect to } R\end{matrix} \xrightarrow{\sim} \text{Bases for } R$$

from Proposition 8.5.2. These actions are transitive. If B is a base for R, then the Weyl group $\mathcal{W}$ is generated by the reflections s_α for $\alpha \in B$. The stabilizer of any point is trivial.

Proof. Let $s \in \mathcal{W}$. If B is a base for R, then it is clear that $s(B)$ is a base for R. Let C be a Weyl chamber of V with respect to R. Let $v \in C$. By Lemma 8.5.1, we have

$$C = X(v) = \{w \in V : (w, \alpha) > 0 \text{ for } \alpha \in B(v)\}.$$

It follows that

$$\begin{aligned} s(C) &= s(\{w \in V : (w, \alpha) > 0 \text{ for } \alpha \in B(v)\}) \\ &= \{x \in V : (s^{-1}(x), \alpha) > 0 \text{ for } \alpha \in B(v)\} \\ &= \{x \in V : (x, s(\alpha)) > 0 \text{ for } \alpha \in B(v)\} \\ &= \{x \in V : (x, \beta) > 0 \text{ for } \beta \in s(B(v))\}. \end{aligned}$$

Since

$$\begin{aligned} s(B(v)) &= s(\{\alpha \in R : (v, \alpha) > 0\}) \\ &= \{\beta \in R : (v, s^{-1}(\beta)) > 0\} \\ &= \{\beta \in R : (s(v), \beta) > 0\} \\ &= B(s(v)). \end{aligned}$$

Hence,

$$\begin{aligned} s(C) &= \{x \in V : (x, \beta) > 0 \text{ for } \beta \in B(s(v))\} \\ &= X(s(v)). \end{aligned}$$

Thus, $s(C) = X(s(v))$ is another Weyl chamber of V with respect to R. To see that the bijection i respects the actions, again let C be a Weyl chamber of V with respect to R, and let $v \in C$. Then

$$\begin{aligned} i(s(C)) &= i\big(X(s(v))\big) \\ &= B(s(v)) \\ &= s\big(B(v)\big) \\ &= s(i(C)), \end{aligned}$$

proving that the actions are compatible with the bijection i.

To prove that the actions are transitive, fix a base B for R, and define R^+ with respect to B. Let $\mathcal{W}'$ be the subgroup of $\mathcal{W}$ generated by the reflections s_α for $\alpha \in B$. Let $v \in V_{\text{reg}}(R)$ be such that $B = B(v)$; the Weyl chamber of V with respect to R corresponding to $B = B(v)$ under the bijection i is $X(v)$. Let C be another Weyl chamber of V with respect to R, and let $w \in C$. Let

$$\delta = \frac{1}{2} \sum_{\alpha \in R^+} \alpha.$$

Let $s \in \mathcal{W}'$ be such that $(s(w), \delta)$ is maximal. We claim that $(s(w), \alpha) > 0$ for all $\alpha \in B$. To see this, let $\alpha \in B$. Since $s_\alpha s$ is also in $\mathcal{W}'$ we have, by the maximality of $(s(w), \delta)$,

$$\begin{aligned} (s(w), \delta) &\geq ((s_\alpha s)(w), \delta) \\ &= (s(w), s_\alpha(\delta)) \\ &= (s(w), \delta - \alpha) \\ &= (s(w), \delta) - (s(w), \alpha). \end{aligned}$$

That is,

$$(s(w), \delta) \geq (s(w), \delta) - (s(w), \alpha).$$

This implies that $(s(w), \alpha) \geq 0$. If $(s(w), \alpha) = 0$, then $(w, s^{-1}(\alpha)) = 0$; this is impossible since $s^{-1}(\alpha)$ is a root and w is regular. Thus, $(s(w), \alpha) > 0$. Since $(s(w), \alpha) > 0$ for all $\alpha \in B$ it follows that $s(w) \in X(v)$. This implies that $s(C) = X(v)$, so that $\mathcal{W}'$, and hence $\mathcal{W}$, acts transitively the set of Weyl chambers of V with respect to R. Since the bijection i is compatible with the actions, the subgroup $\mathcal{W}'$, and hence $\mathcal{W}$, also acts transitively on the set of bases of R.

Let B be a base for R, and as above, let $\mathcal{W}'$ be the subgroup of $\mathcal{W}$ generated by the s_α for $\alpha \in B$. To prove that $\mathcal{W} = \mathcal{W}'$ it suffices to prove that if $\alpha \in R$, then $s_\alpha \in \mathcal{W}'$. Let $\alpha \in R$. By Lemma 8.6.2, there exists a base B' for R such that $\alpha \in B'$. By what we have already proven, there exists $s \in \mathcal{W}'$ such that $s(B') = B$. In particular, $s(\alpha) = \beta$ for some $\beta \in B$. Now by Lemma 8.7.3,

$$s_\beta = s_{s(\alpha)} = s s_\alpha s^{-1},$$

which implies that $s_\alpha = s^{-1} s_\beta s$. Since $s^{-1} s_\beta s \in \mathcal{W}'$, we get $s_\alpha \in \mathcal{W}'$, as desired.

Finally, suppose that B is a base for R and that $s \in \mathcal{W}$ is such that $s(B) = B$. Assume that $s \neq 1$; we will obtain a contradiction. Write $s = s_{\alpha_1} \cdots s_{\alpha_t}$ with $\alpha_1, \ldots, \alpha_t \in B$ and $t \geq 1$ minimal. By Proposition 8.7.5, $s(\alpha_t)$ is negative with respect to B. This contradicts $s(\alpha_t) \in B$. □

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, let R be a root system in V, and let $\mathcal{W}$ be the Weyl group of R. Let $s \in W$ with $s \neq 1$, and write

$$s = s_{\alpha_1} \cdots s_{\alpha_t}$$

with $\alpha_1, \ldots, \alpha_t \in B$ and t minimal. We refer to such an expression for s as **reduced**, and define the **length** of s to be the positive integer $\ell(s) = t$. We define $\ell(1) = 0$.

Proposition 8.7.7. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, let R be a root system in V, and let $\mathcal{W}$ be the Weyl group of R. Let $s \in \mathcal{W}$. The length $\ell(s)$ is equal to the number of positive roots α such that $s(\alpha)$ is negative.*

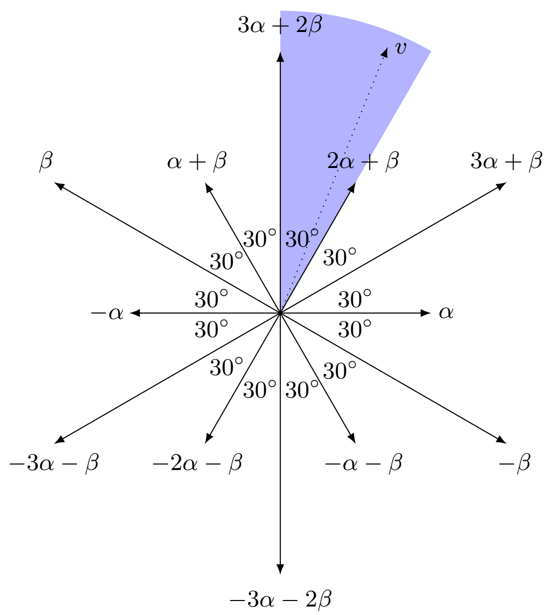
Proof. For $r \in \mathcal{W}$ let $n(r)$ be the number of positive roots α such that $r(\alpha)$ is negative. We need to prove that $\ell(s) = n(s)$. We will prove this by induction on $\ell(s)$. Assume first that $\ell(s) = 0$. Then necessarily $s = 1$. Clearly, $n(1) = 0$. We thus have $\ell(s) = n(s)$. Assume now that $\ell(s) > 0$ and that $\ell(r) = n(r)$ for all $r \in \mathcal{W}$ with $\ell(r) < \ell(s)$. We need to prove that $\ell(s) = n(s)$. Let $s = s_{\alpha_1} \cdots s_{\alpha_t}$ be a reduced expression for s. Set $s' = s s_{\alpha_t}$. Evidently, $\ell(s') = \ell(s) - 1$. By Lemma 8.6.5,

$$\begin{aligned} s(R^+ - \{\alpha_t\}) &= s'(s_{\alpha_t}(R^+ - \{\alpha_t\})) \\ &= s'(R^+ - \{\alpha_t\}). \end{aligned}$$

Also, by Proposition 8.7.5, $s(\alpha_t)$ is negative. Since

$$\begin{aligned} s(\alpha_t) &= s'(s_{\alpha_t}(\alpha_t)) \\ &= -s'(\alpha_t) \end{aligned}$$

we see that $s'(\alpha_t)$ is positive. It follows that $n(s') = n(s) - 1$. By the induction hypothesis, $\ell(s') = n(s')$. This implies now that $\ell(s) = n(s)$, as desired. □



We consider bases, Weyl chambers, and the Weyl group for the root system G_2, which appears in the above diagram. Define the vector v as in the diagram. Then $v \in V_{\text{reg}}(G_2)$. By definition, $R^+(v)$ consists of the roots that form a strictly acute angle with v, i.e.,

$$R^+(v) = \{\alpha, 3\alpha + \beta, 2\alpha + \beta, 3\alpha + 2\beta, \alpha + \beta, \beta\}.$$

By definition, $R^-(v)$ consists of the roots that form a strictly obtuse angle with v, that is:

$$R^-(v) = \{-\alpha, 3\alpha - \beta, -2\alpha - \beta, -3\alpha - 2\beta, -\alpha - \beta, -\beta\}$$

Evidently, $\{\alpha, \beta\}$ is the set of indecomposable roots in $R^+(v)$, so that $B(v) = \{\alpha, \beta\}$ is a base for G_2. The Weyl chambers of V with respect to G_2 consist of the circular sectors with cental angle $30°$ that lie between the roots of G_2. There are 12 such sectors, and hence 12 bases for G_2. The sector containing v is

$$C = X(v) = \{w \in V : (\alpha, v) > 0, (\beta, v) > 0\}.$$

This is the set of vectors that form a strictly acute angle with α and β, and is shaded in blue in the diagram. We know that the Weyl group $\mathcal{W}$ of G_2 acts transitively on both the set of Weyl chambers and bases, with no fixed points. This means that the order of $\mathcal{W}$ is 12. Define:

$$s_1 = s_\alpha, \qquad s_2 = s_\beta.$$

We know that $\mathcal{W}$ is generated by the two elements s_1 and s_2 which each have order two. This means that $\mathcal{W}$ is a dihedral group (the definition of a dihedral group is a group generated by two elements of order two). Consider $s_1 s_2$. This is an element of $\mathrm{SO}(V)$, and hence a rotation. We have

$$\begin{aligned}(s_1 s_2)(\alpha) &= s_\alpha\big(s_\beta(\alpha)\big)\\ &= s_\alpha(\alpha+\beta)\\ &= s_\alpha(\alpha)+s_\alpha(\beta)\\ &= -\alpha+3\alpha+\beta\\ &= 2\alpha+\beta\end{aligned}$$

and

$$\begin{aligned}(s_1 s_2)(\beta) &= s_\alpha\big(s_\beta(\beta)\big)\\ &= -s_\alpha(\beta)\\ &= -3\alpha-\beta.\end{aligned}$$

It follows that $s_1 s_2$ is a rotation in the counterclockwise direction through $60°$. Thus, $s_1 s_2$ has order six. This means that

$$s_1 s_2 s_1 s_2 s_1 s_2 s_1 s_2 s_1 s_2 s_1 s_2 = 1.$$

This implies that:

$$s_1 s_2 s_1 s_2 s_1 s_2 = s_2 s_1 s_2 s_1 s_2 s_1$$

Set

$$r = s_1 s_2.$$

We have

$$\begin{aligned}s_1 r s_1^{-1} &= s_1(s_1 s_2)s_1^{-1}\\ &= s_2 s_1\\ &= s_2^{-1} s_1^{-1}\\ &= (s_1 s_2)^{-1}\\ &= (s_1 s_2)^5\\ &= r^5\\ &= r^{-1}.\end{aligned}$$

We have

$$\mathcal{W} = \langle s_1 s_2\rangle \rtimes \langle s_1\rangle = \langle r\rangle \rtimes \langle s_1\rangle$$

The elements of $\mathcal{W}$ are:

$$\begin{array}{ll} 1, & s_1,\\ r = s_1 s_2, & s_2,\\ r^2 = s_1 s_2 s_1 s_2, & s_2 s_1 s_2,\\ r^3 = s_1 s_2 s_1 s_2 s_1 s_2, & s_2 s_1 s_2 s_1 s_2,\\ r^4 = s_2 s_1 s_2 s_1, & s_1 s_2 s_1 s_2 s_1,\\ r^5 = s_2 s_1, & s_1 s_2 s_1. \end{array}$$

In the ordered basis α, β the linear maps s_1, s_2 and r have the matrices

$$s_1 = \begin{bmatrix} -1 & 3 \\ 0 & 1 \end{bmatrix}, \quad s_2 = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}, \quad r = s_1 s_2 = \begin{bmatrix} -1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ 1 & -1 \end{bmatrix}.$$

Using these matrices, it is easy to calculate that:

$$1: \begin{cases} \alpha & \mapsto \alpha, \\ 3\alpha+\beta & \mapsto 3\alpha+\beta, \\ 2\alpha+\beta & \mapsto 2\alpha+\beta, \\ 3\alpha+2\beta & \mapsto 3\alpha+2\beta, \\ \alpha+\beta & \mapsto \alpha+\beta, \\ \beta & \mapsto \beta, \end{cases} \qquad s_1: \begin{cases} \alpha & \mapsto -\alpha, \\ 3\alpha+\beta & \mapsto \beta, \\ 2\alpha+\beta & \mapsto \alpha+\beta, \\ 3\alpha+2\beta & \mapsto 3\alpha+2\beta, \\ \alpha+\beta & \mapsto 2\alpha+\beta, \\ \beta & \mapsto 3\alpha+\beta, \end{cases}$$

$$r: \begin{cases} \alpha & \mapsto 2\alpha+\beta, \\ 3\alpha+\beta & \mapsto 3\alpha+\beta, \\ 2\alpha+\beta & \mapsto \alpha+\beta, \\ 3\alpha+2\beta & \mapsto \beta, \\ \alpha+\beta & \mapsto -\alpha, \\ \beta & \mapsto -3\alpha-\beta, \end{cases} \qquad s_1 r = s_2: \begin{cases} \alpha & \mapsto \alpha+\beta, \\ 3\alpha+\beta & \mapsto 3\alpha+2\beta, \\ 2\alpha+\beta & \mapsto 2\alpha+\beta, \\ 3\alpha+2\beta & \mapsto 3\alpha+\beta, \\ \alpha+\beta & \mapsto \alpha, \\ \beta & \mapsto -\beta, \end{cases}$$

$$r^2: \begin{cases} \alpha & \mapsto \alpha+\beta, \\ 3\alpha+\beta & \mapsto \beta, \\ 2\alpha+\beta & \mapsto -\alpha, \\ 3\alpha+2\beta & \mapsto -3\alpha-\beta, \\ \alpha+\beta & \mapsto -2\alpha-\beta, \\ \beta & \mapsto -3\alpha-2\beta, \end{cases} \qquad s_1 r^2: \begin{cases} \alpha & \mapsto 2\alpha+\beta, \\ 3\alpha+\beta & \mapsto 3\alpha+\beta, \\ 2\alpha+\beta & \mapsto \alpha, \\ 3\alpha+2\beta & \mapsto -\beta, \\ \alpha+\beta & \mapsto -\alpha-\beta, \\ \beta & \mapsto -3\alpha-2\beta, \end{cases}$$

$$r^3: \begin{cases} \alpha & \mapsto -\alpha, \\ 3\alpha+\beta & \mapsto -3\alpha-\beta, \\ 2\alpha+\beta & \mapsto -2\alpha-\beta, \\ 3\alpha+2\beta & \mapsto -3\alpha-2\beta, \\ \alpha+\beta & \mapsto -\alpha-\beta, \\ \beta & \mapsto -\beta, \end{cases} \qquad s_1 r^3: \begin{cases} \alpha & \mapsto \alpha, \\ 3\alpha+\beta & \mapsto -\beta, \\ 2\alpha+\beta & \mapsto -\alpha-\beta, \\ 3\alpha+2\beta & \mapsto -3\alpha-2\beta, \\ \alpha+\beta & \mapsto -2\alpha-\beta, \\ \beta & \mapsto -3\alpha-\beta, \end{cases}$$

$$r^4: \begin{cases} \alpha & \mapsto -2\alpha-\beta, \\ 3\alpha+\beta & \mapsto -3\alpha-2\beta, \\ 2\alpha+\beta & \mapsto -\alpha-\beta, \\ 3\alpha+2\beta & \mapsto -\beta, \\ \alpha+\beta & \mapsto \alpha, \\ \beta & \mapsto 3\alpha+\beta, \end{cases} \qquad s_1 r^4: \begin{cases} \alpha & \mapsto -\alpha-\beta, \\ 3\alpha+\beta & \mapsto -3\alpha-2\beta, \\ 2\alpha+\beta & \mapsto -2\alpha-\beta, \\ 3\alpha+2\beta & \mapsto -3\alpha-\beta, \\ \alpha+\beta & \mapsto -\alpha, \\ \beta & \mapsto \beta, \end{cases}$$

$$r^5: \begin{cases} \alpha & \mapsto -\alpha-\beta, \\ 3\alpha+\beta & \mapsto -\beta, \\ 2\alpha+\beta & \mapsto \alpha, \\ 3\alpha+2\beta & \mapsto 3\alpha+\beta, \\ \alpha+\beta & \mapsto 2\alpha+\beta, \\ \beta & \mapsto 3\alpha+2\beta, \end{cases} \qquad s_1 r^5: \begin{cases} \alpha & \mapsto -2\alpha-\beta, \\ 3\alpha+\beta & \mapsto -3\alpha-\beta, \\ 2\alpha+\beta & \mapsto -\alpha, \\ 3\alpha+2\beta & \mapsto \beta, \\ \alpha+\beta & \mapsto \alpha+\beta, \\ \beta & \mapsto 3\alpha+2\beta. \end{cases}$$

Using this and that Proposition 8.7.7, we can calculate the length of each element of $\mathcal{W}$. We see that the expressions of the elements of $\mathcal{W}$ in the list from above are in fact reduced, because of Proposition 8.7.7. Thus,

$$\begin{array}{ll}
\ell(1) = 0, & \ell(s_1) = 1, \\
\ell(r = s_1 s_2) = 2, & \ell(s_2) = 1, \\
\ell(r^2 = s_1 s_2 s_1 s_2) = 4, & \ell(s_2 s_1 s_2) = 3, \\
\ell(r^3 = s_1 s_2 s_1 s_2 s_1 s_2) = 6, & \ell(s_2 s_1 s_2 s_1 s_2) = 5, \\
\ell(r^4 = s_2 s_1 s_2 s_1) = 4, & \ell(s_1 s_2 s_1 s_2 s_1) = 5, \\
\ell(r^5 = s_2 s_1) = 2, & \ell(s_1 s_2 s_1) = 3.
\end{array}$$

8.8 Irreducible root systems

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $R \subset V$ be a root system. We say that R is **reducible** if there exist proper subsets $R_1 \subset R$ and $R_2 \subset R$ such that $R = R_1 \cup R_2$ and $(R_1, R_2) = 0$. If R is not reducible we say that R is **irreducible**.

Lemma 8.8.1. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $R \subset V$ be a root system. Assume that R is reducible, so that there exist proper subsets $R_1 \subset R$ and $R_2 \subset R$ such that $R = R_1 \cup R_2$ and $(R_1, R_2) = 0$. Let V_1 and V_2 be the subspaces of V spanned by R_1 and R_2, respectively. Then $V = V_1 \perp V_2$, and R_1 and R_2 are root systems in V_1 and V_2, respectively.*

Proof. Since $(R_1, R_2) = 0$ it is evident that $(V_1, V_2) = 0$. Since $V_1 \oplus V_2 \subset V$ contains R and thus a basis for V, it follows now that V is the orthogonal direct sum of V_1 and V_2. It is easy to see that axioms (R1), (R2), and (R4) for root systems are satisfied by R_1. To see that (R3) is satisfied, let $\alpha, \beta \in R_1$; we need to verify that $s_\alpha(\beta) \in R_1$. Now

$$s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha.$$

This element of R is contained in R_1 or in R_2. Assume that $s_\alpha(\beta) \in R_2$; we will obtain a contradiction. Since $s_\alpha(\beta) \in R_2$, we have

$$\begin{aligned}
0 &= (\alpha, s_\alpha(\beta)) \\
&= (\alpha, \beta) - \langle \beta, \alpha \rangle (\alpha, \alpha) \\
&= (\alpha, \beta) - 2\frac{(\beta, \alpha)}{(\alpha, \alpha)}(\alpha, \alpha) \\
0 &= -(\alpha, \beta),
\end{aligned}$$

so that $(\alpha, \beta) = 0$. Hence, $\langle \alpha, \beta \rangle = 0$. We also have:

$$\begin{aligned}
0 &= (\beta, s_\alpha(\beta)) \\
&= (\beta, \beta) - \langle \beta, \alpha \rangle (\beta, \alpha) \\
0 &= (\beta, \beta).
\end{aligned}$$

This implies that $\beta = 0$, a contradiction. It follows that R_1 is a root system. Similarly, R_2 is a root system. $\square$

Lemma 8.8.2. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $R \subset V$ be a root system. Let B be a base for R. The root system R is reducible if and only if there exist proper subsets $B_1 \subset B$ and $B_2 \subset B$ such that $B = B_1 \cup B_2$ and $(B_1, B_2) = 0$.*

Proof. Assume that R is reducible, so that there exist proper subsets $R_1 \subset R$ and $R_2 \subset R$ such that $R = R_1 \cup R_2$ and $(R_1, R_2) = 0$. Define $B_1 = R_1 \cap B$ and $B_2 = R_2 \cap B$. Evidently, $B = B_1 \cup B_2$. We claim that B_1 and B_2 are proper subsets of B. Assume that $B_1 = B$; we will obtain a contradiction. Since $B_1 = B$ we have $B \subset R_1$. Since B contains a basis for V and since $(R_1, R_2) = 0$, we obtain $(V, R_2) = 0$. This is a contradiction since $(R_2, R_2) \neq 0$. Thus, B_1 is a proper subset of B; similarly, B_2 is a proper subset of B.

Conversely, assume that there exist proper subsets $B_1 \subset B$ and $B_2 \subset B$ such that $B = B_1 \cup B_2$ and $(B_1, B_2) = 0$. Let $\mathcal{W}$ be the Weyl group of R. Define

$$\begin{aligned} R_1 &= \{\alpha \in R : \text{there exists } s \in \mathcal{W} \text{ such that } s(\alpha) \in B_1\}, \\ R_2 &= \{\alpha \in R : \text{there exists } s \in \mathcal{W} \text{ such that } s(\alpha) \in B_2\}. \end{aligned}$$

By Lemma 8.6.2 and Theorem 8.7.6, for every $\alpha \in R$ there exists $s \in \mathcal{W}$ such that $s(\alpha) \in B$. It follows that $R = R_1 \cup R_2$.

To prove $(R_1, R_2) = 0$ we need to introduce some subgroups of $\mathcal{W}$. Let $\mathcal{W}_1$ be the subgroup of $\mathcal{W}$ generated by the s_α with $\alpha \in B_1$, and let $\mathcal{W}_2$ be the subgroup of $\mathcal{W}$ generated by the s_α with $\alpha \in B_2$. We claim that the elements of $\mathcal{W}_1$ commute with the elements of $\mathcal{W}_2$. To prove this, it suffices to verify that $s_{\alpha_1} s_{\alpha_2} = s_{\alpha_2} s_{\alpha_1}$ for $\alpha_1 \in B_1$ and $\alpha_2 \in B_2$. Let $\alpha_1 \in B_1$ and $\alpha_2 \in B_2$. Let $\alpha \in B_1$. Then

$$\begin{aligned} (s_{\alpha_1} s_{\alpha_2})(\alpha) &= s_{\alpha_1}(\alpha - \langle \alpha, \alpha_2 \rangle \alpha_2) \\ &= s_{\alpha_1}(\alpha - 0 \cdot \alpha_2) \\ &= s_{\alpha_1}(\alpha) \\ &= \alpha - \langle \alpha, \alpha_1 \rangle \alpha_1. \end{aligned}$$

And

$$\begin{aligned} (s_{\alpha_2} s_{\alpha_1})(\alpha) &= s_{\alpha_2}(\alpha - \langle \alpha, \alpha_1 \rangle \alpha_1) \\ &= s_{\alpha_2}(\alpha) - \langle \alpha, \alpha_1 \rangle s_{\alpha_2}(\alpha_1) \\ &= \alpha - \langle \alpha, \alpha_2 \rangle \alpha_2 - \langle \alpha, \alpha_1 \rangle \big(\alpha_1 - \langle \alpha_1, \alpha_2 \rangle \alpha_2\big) \\ &= \alpha - \langle \alpha, \alpha_1 \rangle \alpha_1. \end{aligned}$$

Thus, $(s_{\alpha_1} s_{\alpha_2})(\alpha) = (s_{\alpha_2} s_{\alpha_1})(\alpha)$. A similar argument also shows that this equality holds for $\alpha \in B_2$. Since $B = B_1 \cup B_2$ and B is a vector space basis for V, we have $s_{\alpha_1} s_{\alpha_2} = s_{\alpha_2} s_{\alpha_1}$ as claimed. By Theorem 8.7.6 the group $\mathcal{W}$ is generated by the subgroups $\mathcal{W}_1$ and $\mathcal{W}_2$, and by the commutativity property that we have just proven, if $s \in \mathcal{W}$, then there exist $s_1 \in \mathcal{W}_1$ and $s_2 \in \mathcal{W}_2$ such that $s = s_1 s_2 = s_2 s_1$. Now let $\alpha \in R_1$. By definition, there exists $s \in \mathcal{W}$ and $\alpha_1 \in R_1$ such that $\alpha = s(\alpha_1)$. Write $s = s_1 s_2$ with $s_1 \in \mathcal{W}_1$ and $s_2 \in \mathcal{W}_2$.

Since s_2 is a product of elements of the form s_β for $\beta \in B_2$, and each such s_β is the identity on B_1 (use the formula for s_β and $(B_1, B_2) = 0$), it follows that $\alpha = s_1(\alpha_1)$. Writing s_1 as a product of elements of the form s_γ for $\gamma \in B_1$, and using the formula for such s_γ, we see that $\alpha = s(\alpha_1)$ is in the span of B_1. Similarly, if $\alpha \in R_2$, then α is in the span of B_2. Since $(B_1, B_2) = 0$, it now follows that $(R_1, R_2) = 0$.

To see that R_1 and R_2 are proper subsets of R, assume that, say, $R_1 = R$; we will obtain a contradiction. Since $(R_1, R_2) = 0$ we must have $R_2 = 0$. This implies that B_2 is empty (because clearly $B_2 \subset R_2$); this is a contradiction. □

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $R \subset V$ be a root system. Let B be a base for R. Let $v_1, v_2 \in V$, and write

$$v_1 = \sum_{\gamma \in B} c_1(\gamma)\gamma, \qquad v_2 = \sum_{\gamma \in B} c_2(\gamma)\gamma.$$

Here, we use that B is also a vector space basis for V. We define a relation $\succ$ on R by

$$v_1 \succ v_2$$

if and only if

$$c_1(\gamma) \geq c_2(\gamma) \quad \text{for all } \gamma \in B.$$

The relation $\succ$ is a partial order on V. Evidently,

$$R^+ = \{\alpha \in R : \alpha \succ 0\} \quad \text{and} \quad R^- = \{\alpha \in R : \alpha \prec 0\}.$$

We say that α is **maximal** if, for all $\beta \in R$, $\beta \succ \alpha$ implies that $\beta = \alpha$.

Lemma 8.8.3. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $R \subset V$ be a root system. Assume that R is irreducible. Let B be a base for R. With respect to $\succ$, there exists a unique maximal root β in R. Moreover, if β is written as*

$$\beta = \sum_{\alpha \in B} b(\alpha)\alpha,$$

then $b(\alpha) > 0$ for all $\alpha \in B$.

Proof. There exists at least one maximal root in R; let $\beta \in R$ be any maximal root in R. Write

$$\beta = \sum_{\alpha \in B} b(\alpha)\alpha.$$

Since β is maximal, we must have $b(\alpha) \geq 0$ for all $\alpha \in B$. Define

$$B_1 = \{\alpha \in B : b(\alpha) > 0\} \quad \text{and} \quad B_2 = \{\alpha \in B : b(\alpha) = 0\}.$$

We have $B = B_1 \cup B_2$, and B_1 is non-empty. We claim that B_2 is empty. Suppose not; we will obtain a contradiction. Since R is irreducible, by Lemma

8.8.2 we must have $(B_1, B_2) \neq 0$. Proposition 8.4.9 and Lemma 8.4.4 imply that $(\alpha_1, \alpha_2) \leq 0$ for all $\alpha_1 \in B_1$ and $\alpha \in B_2$. For $\alpha_2 \in B_2$ we have

$$(\beta, \alpha_2) = \sum_{\alpha \in B} b(\alpha)(\alpha, \alpha_2) = \sum_{\alpha_1 \in B_1} b(\alpha_1)(\alpha_1, \alpha_2)$$

where each term is less than or equal to zero. Since $(B_1, B_2) \neq 0$, there exist $\alpha_1' \in B_1$ and $\alpha_2' \in B_2$ such that $(\alpha_1', \alpha_2') \neq 0$, so that $(\alpha_1', \alpha_2') < 0$. This implies that $(\beta, \alpha_2') < 0$. By Lemma 8.3.4, either $\beta = \pm\alpha_2'$ or $\beta + \alpha_2'$ is a root. Assume that $\beta = \alpha_2'$. Then $(\beta, \alpha_2') = (\beta, \beta) > 0$, contradicting $(\beta, \alpha_2') < 0$. Assume that $\beta = -\alpha_2'$. Then $b(\alpha_2') = -1 < 0$, a contradiction. It follows that $\beta + \alpha_2'$ is a root. Now $\beta + \alpha_2' \succ \beta$. Since β is maximal, we have $\beta + \alpha_2' = \beta$. This means that $\alpha_2' = 0$, a contradiction. It follows that B_2 is empty, so that $b(\alpha) > 0$ for all $\alpha \in B$. Arguing similarly, we also see that $(\beta, \alpha) \geq 0$ for all $\alpha \in B$ (if $(\beta, \alpha) < 0$ for some $\alpha \in B$, then $\beta + \alpha$ is a root, which contradicts the maximality of β). Since B is a basis for V we cannot have $(\beta, B) = 0$; hence, there exists $\alpha_0 \in B$ such that $(\beta, \alpha_0) > 0$.

Now suppose that β' is another maximal root. Write

$$\beta' = \sum_{\alpha \in B} b'(\alpha)\alpha.$$

As in the last paragraph, $b'(\alpha) > 0$ for all $\alpha \in B$. Now

$$(\beta, \beta') = \sum_{\alpha \in B} b'(\alpha')(\beta, \alpha).$$

As $(\beta, \alpha) \geq 0$ and $b'(\alpha) > 0$ for all $\alpha \in B$, and $(\beta, \alpha_0) > 0$, we see that $(\beta, \beta') > 0$. By Lemma 8.3.4, either $\beta = \beta'$, $\beta = -\beta'$ or $\beta - \beta'$ is a root. Assume that $\beta = -\beta'$. Then $b(\alpha) = -b'(\alpha)$ for $\alpha \in B$; this contradicts the fact that $b(\alpha)$ and $b(\alpha')$ are positive for all $\alpha \in B$. Assume that $\beta - \beta'$ is a root. Then either $\beta - \beta' \succ 0$ or $\beta - \beta' \prec 0$. Assume that $\beta - \beta' \succ 0$. Then $\beta \succ \beta'$, which implies $\beta = \beta'$ by the maximality of β'. Therefore, $\beta - \beta' = 0$; this is not a root, and hence a contradiction. Similarly, the assumption that $\beta - \beta' \prec 0$ leads to a contradiction. We conclude that $\beta = \beta'$. □

Lemma 8.8.4. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $R \subset V$ be a root system. Assume that R is irreducible. Let B be a base for R. Let β be the maximal root of R with respect to B. We have $\beta \succ \alpha$ for all $\alpha \in R$, $\alpha \neq \beta$. Also, if $\alpha \in B$, then $(\beta, \alpha) \geq 0$.*

Proof. Let $\alpha \in R$ with $\alpha \neq \beta$. Since $\alpha \neq \beta$, α is not maximal by Lemma 8.8.3. It follows that there exists $\gamma_1 \in R$ such that $\gamma_1 \succ \alpha$ and $\gamma_1 \neq \alpha$. If $\gamma_1 = \beta$, then $\beta \succ \alpha$. Assume $\gamma_1 \neq \beta$. Since $\gamma_1 \neq \beta$, γ_1 is not maximal by Lemma 8.8.3. It follows that there exists $\gamma_2 \in R$ such that $\gamma_2 \succ \gamma_1$ and $\gamma_2 \neq \gamma_1$. If $\gamma_2 = \beta$, then $\beta \succ \gamma_1 \succ \alpha$, so that $\beta \succ \alpha$. If $\gamma_2 \neq \beta$, we continue to argue in the same fashion. Since R is finite, we eventually conclude that $\beta \succ \alpha$.

Let $\alpha \in B$. Assume that $(\alpha, \beta) < 0$. Then certainly $\alpha \neq \beta$. Also, we cannot have $\alpha = -\beta$ because β is a positive root with respect to B by Lemma 8.8.3. By Lemma 8.3.4, $\alpha + \beta$ is a root. This contradicts the maximality of β. □

Lemma 8.8.5. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $R \subset V$ be a root system. Assume that R is irreducible. The Weyl group $\mathcal{W}$ of R acts irreducibly on V.*

Proof. Assume that U is a $\mathcal{W}$ subspace of V. We need to prove that $U = 0$ or $U = V$. Assume that $U \neq 0$. Since the elements of $\mathcal{W}$ lie in the orthogonal group $\mathrm{O}(V)$ of V, the subspace $U^\perp$ is also a $\mathcal{W}$ subspace. We have $V = U \oplus U^\perp$. Let $\alpha \in R$. We claim that $\alpha \in U$ or $\alpha \in U^\perp$. Write $\alpha = u + u'$ with $u \in U$ and $u' \in U^\perp$. We have

$$\begin{aligned} s_\alpha(\alpha) &= s_\alpha(u) + s_\alpha(u') \\ -\alpha &= s_\alpha(u) + s_\alpha(u') \\ -u - u' &= s_\alpha(u) + s_\alpha(u'). \end{aligned}$$

Since $s_\alpha \in \mathcal{W}$ we have $s_\alpha(u) \in U$ and $s_\alpha(u') \in U^\perp$. It follows that

$$s_\alpha(u) = -u \quad \text{and} \quad s_\alpha(u') = -u'.$$

These equalities imply that $u \in \mathbb{R}\alpha$ and $u' \in \mathbb{R}\alpha$. Since $U \cap U^\perp = 0$, this implies that $u = 0$ or $u' = 0$, as desired. Now define

$$R_1 = \{\alpha \in R : \alpha \in U\} \quad \text{and} \quad R_2 = \{\alpha \in R : \alpha \in U^\perp\}.$$

By we have just proven $R = R_1 \cup R_2$. It is clear that $(R_1, R_2) = 0$. Since R is irreducible, either R_1 is empty or R_2 is empty. If R_1 is empty, then $R \subset U^\perp$, so that $V = U^\perp$ and thus $U = 0$; if R_2 is empty, then $R \subset U$, so that $V = U$. □

Lemma 8.8.6. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $R \subset V$ be a root system. Assume that R is irreducible, and let $\mathcal{W}$ be the Weyl group of R. The function $R \to \mathbb{R}_{>0}$ sending $\alpha \to \|\alpha\|$ takes on at most two values. Moreover, if $\alpha, \beta \in R$ have the same length, then there exists $s \in \mathcal{W}$ such that $s(\alpha) = \beta$.*

Proof. Suppose that there exist $\alpha_1, \alpha_2, \alpha_3 \in R$ such that $\|\alpha_1\| < \|\alpha_2\| < \|\alpha_3\|$; we will obtain a contradiction.

We first assert that there exist roots $\alpha_1', \alpha_2', \alpha_3' \in R$ such that

$$\|\alpha_1'\| = \|\alpha_1\|, \quad \|\alpha_2'\| = \|\alpha_2\|, \quad \|\alpha_3'\| = \|\alpha_3\|$$

and

$$(\alpha_1', \alpha_2') \neq 0, \quad (\alpha_2', \alpha_3') \neq 0, \quad (\alpha_1', \alpha_3') \neq 0.$$

To see this we note that by Lemma 8.8.5, the vectors $s(\alpha_2)$ for $s \in \mathcal{W}$ span V; it follows that there exists $s \in \mathcal{W}$ such that $(\alpha_1, s(\alpha_2)) \neq 0$. Similarly, there exists $r \in \mathcal{W}$ such that $(s(\alpha_2), r(\alpha_3)) \neq 0$. If $(\alpha_1, r(\alpha_3)) \neq 0$, we define

$$\alpha_1' = \alpha_1, \quad \alpha_2' = s(\alpha_2), \quad \alpha_3' = r(\alpha_3)$$

and these vectors have the desired properties. Assume that $(\alpha_1, r(\alpha_3)) = 0$. In this case we define

$$\alpha_1' = \alpha_1, \quad \alpha_2' = s(\alpha_2), \quad \alpha_3' = s_{s(\alpha_2)}(r(\alpha_3)).$$

We have

$$(\alpha_2', \alpha_3') = (s(\alpha_2), s_{s(\alpha_2)}(r(\alpha_3))) = -(s(\alpha_2), r(\alpha_3)) \neq 0.$$

And

$$\begin{aligned}(\alpha_1', \alpha_3') &= (\alpha_1, s_{s(\alpha_2)}(r(\alpha_3)))\\ &= (\alpha_1, r(\alpha_3) - \langle r(\alpha_3), s(\alpha_2)\rangle s(\alpha_2))\\ &= (\alpha_1, r(\alpha_3)) - \langle r(\alpha_3), s(\alpha_2)\rangle(\alpha_1, s(\alpha_2))\\ &= -\langle r(\alpha_3), s(\alpha_2)\rangle(\alpha_1, s(\alpha_2))\\ &= -2\frac{(r(\alpha_3), s(\alpha_2))}{(s(\alpha_2), s(\alpha_2))}(\alpha_1, s(\alpha_2))\\ &\neq 0.\end{aligned}$$

Again, α_1', α_2' and α_3' have the desired properties.

We have $\|\alpha_1'\| < \|\alpha_2'\| < \|\alpha_3'\|$. Thus,

$$1 < \frac{\|\alpha_2'\|}{\|\alpha_1'\|} < \frac{\|\alpha_3'\|}{\|\alpha_1'\|}.$$

Applying Lemma 8.3.3 to the pair α_1' and α_2', and the pair α_1' and α_3', and taking note of the above inequalities, we must have

$$\frac{\|\alpha_2'\|}{\|\alpha_1'\|} = \sqrt{2} \quad \text{and} \quad \frac{\|\alpha_3'\|}{\|\alpha_1'\|} = \sqrt{3}.$$

This implies that

$$\frac{\|\alpha_3'\|}{\|\alpha_2'\|} = \frac{\sqrt{3}}{\sqrt{2}}.$$

However, Lemma 8.3.3 applied to the pair α_2' and α_3' implies that $\sqrt{3}/\sqrt{2}$ is not an allowable value for $\|\alpha_3'\|/\|\alpha_2'\|$. This is a contradiction.

Assume that $\alpha, \beta \in R$ have the same length. Arguing as in the last paragraph, there exists $s \in \mathcal{W}$ such that $(s(\alpha), \beta) \neq 0$. If $s(\alpha) = \beta$, then s is the desired element of $\mathcal{W}$. If $s(\alpha) = -\beta$, then $(s_\beta s)(\alpha) = \beta$, and $s_\beta s$ is the desired element. Assume that $s(\alpha) \neq \pm\beta$. Since $s(\alpha)$ and β have the same length, we have by Lemma 8.3.3 that $\langle s(\alpha), \beta\rangle = \langle \beta, s(\alpha)\rangle = \pm 1$. Assume that $\langle s(\alpha), \beta\rangle = 1$. We have

$$\begin{aligned}(s_\beta s_{s(\alpha)} s_\beta)(s(\alpha)) &= (s_\beta s_{s(\alpha)})(s(\alpha) - \langle s(\alpha), \beta\rangle\beta)\\ &= (s_\beta s_{s(\alpha)})(s(\alpha) - \beta)\\ &= s_\beta(-s(\alpha) - s_{s(\alpha)}(\beta))\end{aligned}$$

$$= s_\beta(-s(\alpha) - \beta + \langle \beta, s(\alpha) \rangle s(\alpha))$$
$$= s_\beta(-\beta)$$
$$= \beta.$$

Assume that $\langle s(\alpha), \beta \rangle = -1$. Then $\langle s(\alpha), \beta' \rangle = 1$ where $\beta' = -\beta = s_\beta(\beta)$. By what we have already proven, there exists $r \in \mathcal{W}$ such that $r(\alpha) = \beta'$. It follows that $(s_\beta r)(\alpha) = \beta$. $\square$

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $R \subset V$ be a root system. Assume that R is irreducible. By Lemma 8.8.6, there are at most two possible lengths for the elements of R. If $\{\|\alpha\| : \alpha \in R\}$ contains two distinct elements ℓ_1 and ℓ_2 with $\ell_1 < \ell_2$, then we refer to the $\alpha \in R$ with $\|\alpha\| = \ell_1$ as **short roots** and the $\alpha \in R$ with $\|\alpha\| = \ell_2$ as **long roots**. If $\{\|\alpha\| : \alpha \in R\}$ contains one element, then we say that all the elements of R are long.

Lemma 8.8.7. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $R \subset V$ be a root system. Assume that R is irreducible. Let B be a base for R, and let $\beta \in R$ be maximal with respect to B. Then β is a long root.*

Proof. Let $\alpha \in R$. We need to prove that $(\beta, \beta) \geq (\alpha, \alpha)$. By Proposition 8.4.9 there exists $v \in V_{\mathrm{reg}}(R)$ such that $B = B(v)$. Let C be the Weyl chamber containing v. By Lemma 8.5.1 we have

$$C = \{w \in V : (w, \gamma) > 0 \text{ for all } \gamma \in B = B(v)\}.$$

By Lemma 8.5.3 there exists a Weyl chamber C' of V with respect to R such that $\alpha \in \bar{C}'$. Let B' be the base corresponding to C', as in Proposition 8.5.2. Now by Lemma 8.5.3 we have

$$\bar{C} = \{w \in V : (w, \alpha) \geq 0 \text{ for all } \alpha \in B\}$$

and

$$\bar{C}' = \{w \in V : (w, \alpha) \geq 0 \text{ for all } \alpha \in B'\}.$$

By Theorem 8.7.6 there exists s in the Weyl group of R such that $s(C') = C$ and $s(B') = B$. It follows that $s(\bar{C}') = \bar{C}$. Replacing α with $s(\alpha)$ (which has the same length as α), we may assume that $\alpha \in \bar{C}$. By Lemma 8.8.4 we also have $\beta \in \bar{C}$. Next, by Lemma 8.8.4, we have $\beta \succ \alpha$. This means that

$$\beta - \alpha = \sum_{\gamma \in B} c(\gamma)\gamma$$

with $c(\gamma) \geq 0$ for all $\gamma \in B$. Let $w \in \bar{C}$. Then

$$(w, \beta - \alpha) = \sum_{\gamma \in B} c(\gamma)(w, \gamma) \geq 0.$$

Applying this observation to $\alpha \in \bar{C}$ and $\beta \in \bar{C}$, we get:

$$(\alpha, \beta - \alpha) \geq 0, \qquad (\beta, \beta - \alpha) \geq 0.$$

This means that

$$(\alpha, \beta) \geq (\alpha, \alpha), \qquad (\beta, \beta) \geq (\beta, \alpha).$$

It follows that $(\beta, \beta) \geq (\alpha, \alpha)$, as desired. $\square$

Chapter 9

Cartan matrices and Dynkin diagrams

9.1 Isomorphisms and automorphisms

Let V_1 and V_2 be a finite-dimensional vector spaces over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)_1$ and $(\cdot,\cdot)_2$, respectively, and let $R_1 \subset V_1$ and $R_2 \subset V_2$ be root systems. We say that R_1 and R_2 are **isomorphic** if there exists an R vector space isomorphism $\phi : V_1 \to V_2$ such that:

1. $\phi(R_1) = R_2$.
2. If $\alpha, \beta \in R_1$, then $\langle \phi(\alpha), \phi(\beta)\rangle = \langle \alpha, \beta\rangle$.

We refer to such a ϕ as an **isomorphism** from R_1 to R_2. Evidently, if ϕ is an isomorphism from R_1 to R_2, then ϕ^{-1} is an isomorphism from R_2 to R_1.

Lemma 9.1.1. *Let V_1 and V_2 be a finite-dimensional vector spaces over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)_1$ and $(\cdot,\cdot)_2$, respectively, and let $R_1 \subset V_1$ and $R_2 \subset V_2$ be root systems. Let $\mathcal{W}_1$ and $\mathcal{W}_2$ be Weyl groups of R_1 and R_2, respectively. Assume that R_1 and R_2 are isomorphic via the R vector space isomorphism $\phi : V_1 \to V_2$. If $\alpha, \beta \in R_1$, then*

$$s_{\phi(\alpha)}(\phi(\beta)) = \phi(s_\alpha(\beta)).$$

The map given by $s \mapsto \phi \circ s \circ \phi^{-1}$ defines an isomorphism of groups

$$\mathcal{W}_1 \xrightarrow{\sim} \mathcal{W}_2.$$

Proof. Let $\alpha, \beta \in R_1$. We have

$$\begin{aligned} s_{\phi(\alpha)}(\phi(\beta)) &= \phi(\beta) - \langle \phi(\beta), \phi(\alpha)\rangle \phi(\alpha) \\ &= \phi(\beta) - \langle \beta, \alpha\rangle \phi(\alpha) \end{aligned}$$

$$= \phi(\beta - \langle \beta, \alpha \rangle \alpha)$$
$$= \phi(s_\alpha(\beta)).$$

Let $s \in \mathcal{W}_1$, $\alpha \in R_1$, and $\alpha' \in R_2$. Then

$$\begin{aligned}(\phi \circ s_\alpha \circ \phi^{-1})(\alpha') &= \phi\big(s_\alpha(\phi^{-1}(\alpha'))\big) \\ &= s_{\phi(\alpha)}(\alpha').\end{aligned}$$

It follows that $\phi \circ s_\alpha \circ \phi^{-1} = s_{\phi(\alpha)}$ is contained in $\mathcal{W}_2$, so that the map $\mathcal{W}_1 \to \mathcal{W}_2$ is well-defined. This map is evidently a homomorphism of groups. The map $\mathcal{W}_2 \to \mathcal{W}_1$ defined by $s' \mapsto \phi^{-1} \circ s' \circ \phi$ is also a well-defined homomorphism and is the inverse of $\mathcal{W}_1 \to \mathcal{W}_2$. □

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $R \subset V$ be a root system. If $\phi : V \to V$ is an isomorpism from R to R then we say that ϕ is an **automorphism** of R.

Lemma 9.1.2. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $R \subset V$ be a root system. A function $\phi : V \to V$ is an automorphism of R if and only if ϕ is an $\mathbb{R}$ vector space isomorphism from V to V, and $\phi(R) = R$. The set of automorphisms of R forms a group* $\mathrm{Aut}(R)$ *under composition of functions. The Weyl group $\mathcal{W}$ of R is a normal subgroup of* $\mathrm{Aut}(R)$.

Proof. Let $\phi : V \to V$ be a function. If ϕ is an automorphism of R, then ϕ is a vector space isomorphism from V to V and $\phi(R) = R$ by definition. Assume that ϕ is a vector space isomorphism from V to V and $\phi(R) = R$. By Lemma 8.7.3 we have $\langle \phi(\alpha), \phi(\beta) \rangle = \langle \alpha, \beta \rangle$ for all $\alpha, \beta \in R$. It follows that ϕ is an automorphism of R. It is clear that $\mathrm{Aut}(R)$ is a group under composition of functions, and that $\mathcal{W}$ is a subgroup of $\mathrm{Aut}(R)$. To see that $\mathcal{W}$ is normal in $\mathrm{Aut}(R)$, let $\alpha, \beta \in R$ and $\phi \in \mathrm{Aut}(R)$. Then

$$\begin{aligned}(\phi \circ s_\alpha \circ \phi^{-1})(\beta) &= \phi\big(s_\alpha(\phi^{-1}(\beta)\big) \\ &= s_{\phi(\alpha)}(\beta).\end{aligned}$$

Since R contains a basis for V this implies that $\phi \circ s_\alpha \circ \phi^{-1} = s_{\phi(\alpha)}$. It follows that $\mathcal{W}$ is normal in $\mathrm{Aut}(R)$. □

9.2 The Cartan matrix

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $R \subset V$ be a root system. Let B be a base for R, and order the elements of B as $\alpha_1, \ldots, \alpha_t$. We define

$$C(\alpha_1, \ldots, \alpha_t) = (\langle \alpha_i, \alpha_j \rangle)_{1 \le i,j \le t} = \begin{bmatrix} \langle \alpha_1, \alpha_1 \rangle & \cdots & \langle \alpha_1, \alpha_t \rangle \\ \vdots & & \vdots \\ \langle \alpha_t, \alpha_1 \rangle & \cdots & \langle \alpha_t, \alpha_t \rangle \end{bmatrix}.$$

Evidently, the entries of $C(\alpha_1, \ldots, \alpha_t)$ are integers.

Lemma 9.2.1. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, let $R \subset V$ be a root system, and let B and B' be bases for R. Order the elements of B as $\alpha_1,\dots,\alpha_t$ and order the elements of B' as $\alpha'_1,\dots,\alpha'_t$. There exists a $t\times t$ permutation matrix P such that*

$$C(\alpha'_1,\dots,\alpha'_t) = P\cdot C(\alpha_1,\dots,\alpha_t)\cdot P^{-1}.$$

Proof. By Theorem 8.7.6 there exists an element s in the Weyl group of R such that $B' = s(B)$. Since $B' = s(B)$, there exists a $t\times t$ permutation matrix P such that $P^{-1}\cdot C(\alpha'_1,\dots,\alpha'_t)\cdot P = C(s(\alpha_1),\dots,s(\alpha_t))$. Now

$$\begin{aligned}
P^{-1}\cdot C(\alpha'_1,\dots,\alpha'_t)\cdot P &= C(s(\alpha_1),\dots,s(\alpha_t))\\
&= (\langle s(\alpha_i), s(\alpha_j)\rangle)_{1\le i,j\le t}\\
&= \Big(\frac{2(s(\alpha_i),s(\alpha_j))}{(s(\alpha_j),s(\alpha_j))}\Big)_{1\le i,j\le t}\\
&= \Big(\frac{2(\alpha_i,\alpha_j)}{(\alpha_j,\alpha_j)}\Big)_{1\le i,j\le t}\\
&= (\langle \alpha_i,\alpha_j\rangle)_{1\le i,j\le t}\\
&= C(\alpha_1,\dots,\alpha_t).
\end{aligned}$$

This is the assertion of the lemma. □

Let t be a positive integer. We will say that two $t\times t$ matrices C and C' with integer entries are equivalent if there exists a permutation matrix P such that $C' = PCP^{-1}$.

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $R\subset V$ be a root system. We define the **Cartan matrix** $C(R)$ of R to be the equivalence class determined by $C(\alpha_1,\dots,\alpha_t)$ where $\alpha_1,\dots,\alpha_t$ are the elements of a base for R. By Lemma 9.2.1, the Cartan matrix of R is well-defined.

Lemma 9.2.2. *Let V and V' be a finite-dimensional vector spaces over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$ and $(\cdot,\cdot)$, respectively, and let $R\subset V$ and $R'\subset V'$ be root systems. The root systems R and R' are isomorphic if and only if R and R' have the same Cartan matrices.*

Proof. Assume that R and R' have the same Cartan matrices. Then V and V' have the same dimension t, and there exists bases $B=\{\alpha_1,\dots,\alpha_t\}$ and $B'=\{\alpha'_1,\dots,\alpha'_t\}$ for R_1 and R_2, respectively, such that $C(\alpha_1,\dots,\alpha_t) = C(\alpha'_1,\dots,\alpha'_t)$. Define $\phi: V_1\to V_2$ by $\phi(\alpha_i)=\alpha'_i$ for $i\in\{1,\dots,t\}$. We need to prove that $\phi(R)=R'$ and that $\langle\phi(\alpha),\phi(\beta)\rangle = \langle\alpha,\beta\rangle$ for $\alpha,\beta\in R$. Let $\alpha,\beta\in B$. Since $C(\alpha_1,\dots,\alpha_t) = C(\alpha'_1,\dots,\alpha'_t)$ we have $\langle\phi(\beta),\phi(\alpha)\rangle = \langle\beta,\alpha\rangle$. Therefore,

$$\begin{aligned}
\phi(s_\alpha(\beta)) &= \phi(\beta - \langle\beta,\alpha\rangle\alpha)\\
&= \phi(\beta) - \langle\beta,\alpha\rangle\phi(\alpha)
\end{aligned}$$

$$= \phi(\beta) - \langle \phi(\beta), \phi(\alpha) \rangle \phi(\alpha)$$
$$= s_{\phi(\alpha)}(\phi(\beta)).$$

Since every element of R is a linear combination of elements of B, it follows that

$$\phi(s_\alpha(\beta)) = s_{\phi(\alpha)}(\phi(\beta))$$

holds for all $\alpha \in B$ and $\beta \in R$. More generally, let s be in the Weyl group of R_1. By Theorem 8.7.6 there exist $\delta_1, \dots, \delta_n \in B$ such that

$$s = s_{\delta_1} \cdots s_{\delta_n}.$$

Let $\beta \in R$. Repeatedly using the identity we have already proved, we find that:

$$\begin{aligned}
\phi(s(\beta)) &= \phi\big((s_{\delta_1} \cdots s_{\delta_n})(\beta)\big) \\
&= s_{\phi(\delta_1)}\Big(\phi\big((s_{\delta_2} \cdots s_{\delta_n})(\beta)\big)\Big) \\
&= s_{\phi(\delta_1)} s_{\phi(\delta_2)}\Big(\phi\big((s_{\delta_3} \cdots s_{\delta_n})(\beta)\big)\Big) \\
&\dots \\
\phi(s(\beta)) &= s_{\phi(\delta_1)} \cdots s_{\phi(\delta_n)}\big(\phi(\beta)\big).
\end{aligned}$$

Again let $\beta \in R$. By Lemma 8.6.2 and Theorem 8.7.6, there exists s in the Weyl group of R such that $s(\beta) \in B$. We have $\phi(s(\beta)) \in B'$. Write s as a product, as above. Then $\phi(s(\beta)) = s_{\phi(\delta_1)} \cdots s_{\phi(\delta_n)}\big(\phi(\beta)\big)$. Since $\phi(s(\beta)) \in B'$, we have $s_{\phi(\delta_1)} \cdots s_{\phi(\delta_n)}\big(\phi(\beta)\big) \in B'$. Applying the inverse of $s_{\phi(\delta_1)} \cdots s_{\phi(\delta_n)}$, we see that $\phi(\beta) \in R'$. Thus, $\phi(R) \subset R'$. A similar argument implies that $\phi(R') \subset R$, so that $\phi(R) = R'$.

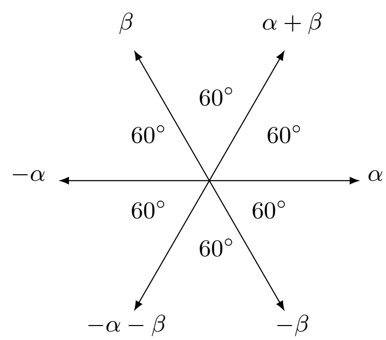
We still need to prove that $\langle \phi(\alpha), \phi(\beta) \rangle = \langle \alpha, \beta \rangle$ for $\alpha, \beta \in R$. By the definition of ϕ, and since $C(\alpha_1, \dots, \alpha_t) = C(\alpha'_1, \dots, \alpha'_t)$, we have $\langle \phi(\alpha), \phi(\beta) \rangle = \langle \alpha, \beta \rangle$ for $\alpha, \beta \in B$. Since this formula is linear in α, the formula holds for all $\alpha \in R$ and $\beta \in B$. Let β be an arbitrary element of R. As before, there exists s in the Weyl group of R such that $s(\beta) \in B$, and $\delta_1, \dots, \delta_n$ such that $\delta_1, \dots, \delta_n \in B$ and $s = s_{\delta_1} \cdots s_{\delta_n}$. Let $\alpha \in R$. Then

$$\begin{aligned}
\langle \alpha, \beta \rangle &= \langle s(\alpha), s(\beta) \rangle \\
&= \langle \phi(s(\alpha)), \phi(s(\beta)) \rangle \\
&= \langle \phi(s(\alpha)), s_{\phi(\delta_1)} \cdots s_{\phi(\delta_n)}\big(\phi(\beta)\big) \rangle \\
&= \langle s^{-1}_{\phi(\delta_n)} \cdots s^{-1}_{\phi(\delta_1)} \phi(s(\alpha)), \phi(\beta) \rangle \\
&= \langle s_{\phi(\delta_n)} \cdots s_{\phi(\delta_1)} \phi(s(\alpha)), \phi(\beta) \rangle \\
&= \langle \phi(s_{\delta_n} \cdots s_{\delta_1} s(\alpha)), \phi(\beta) \rangle \\
&= \langle \phi(\alpha), \phi(\beta) \rangle.
\end{aligned}$$

This completes the proof. □

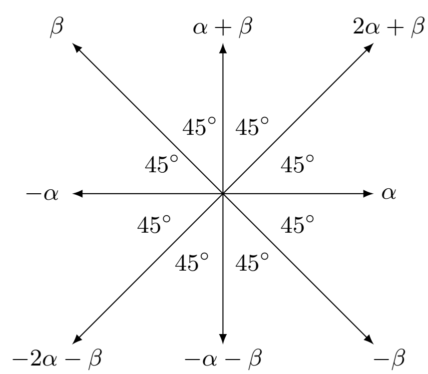
We list the Cartan matrices of the examples from Chapter 8.

1. (A_2 root system)



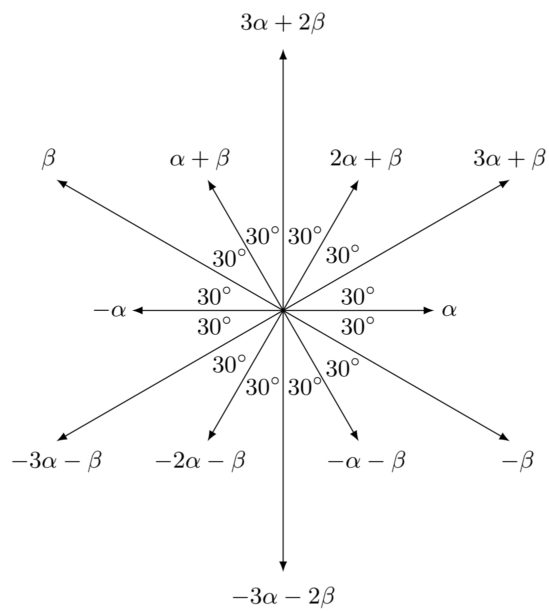
Cartan matrix: $\begin{bmatrix} \langle \alpha,\alpha\rangle & \langle \alpha,\beta\rangle \\ \langle \beta,\alpha\rangle & \langle \beta,\beta\rangle \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$.

2. (B_2 root system)



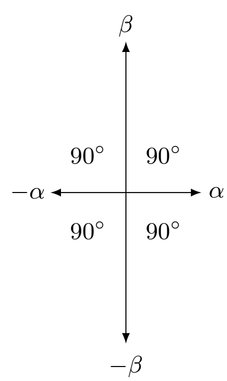
Cartan matrix: $\begin{bmatrix} \langle \alpha,\alpha\rangle & \langle \alpha,\beta\rangle \\ \langle \beta,\alpha\rangle & \langle \beta,\beta\rangle \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -2 & 2 \end{bmatrix}$.

3. (G_2 root system)



$$\text{Cartan matrix: } \begin{bmatrix} \langle\alpha,\alpha\rangle & \langle\alpha,\beta\rangle \\ \langle\beta,\alpha\rangle & \langle\beta,\beta\rangle \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}.$$

4. ($A_1 \times A_1$ root system)



$$\text{Cartan matrix: } \begin{bmatrix} \langle\alpha,\alpha\rangle & \langle\alpha,\beta\rangle \\ \langle\beta,\alpha\rangle & \langle\beta,\beta\rangle \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

9.3 Dynkin diagrams

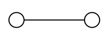
Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $R \subset V$ be a root system. We associate to R a kind of a graph D, called a **Dynkin diagram**, as follows. Let B be a base for R. The vertices of D are labelled with the elements of B. Let $\alpha, \beta \in B$ with $\alpha \neq \beta$. Between the vertices corresponding to α and β we draw

$$d_{\alpha\beta} = \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4\frac{(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)} = 4\frac{(\alpha,\beta)^2}{\|\alpha\|^2\|\beta\|^2}.$$

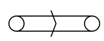
lines; recall that in Lemma 8.3.2 we proved that $d_{\alpha\beta}$ is in $\{0, 1, 2, 3\}$, and that $d_{\alpha\beta}$ was computed in more detail in Lemma 8.3.3. By Lemma 8.3.3, if $d_{\alpha\beta} > 1$, then α and β have different lengths; in this case, we draw an arrow pointing to the shorter root. We will also sometimes consider another graph associated to R. This is called the **Coxeter graph**, and consists of the Dynkin diagram without the arrows pointing to shorter roots.

We have the following of examples of Dynkin diagrams:

1. (A_2 root system)



2. (B_2 root system)



3. (G_2 root system)

4. ($A_1 \times A_1$ root system)

Lemma 9.3.1. *Let V and V' be a finite-dimensional vector spaces over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$ and $(\cdot,\cdot)$, respectively, and let $R \subset V$ and $R' \subset V'$ be root systems. The root systems R and R' are isomorphic if and only if R and R' have the same directed Dynkin diagrams.*

Proof. Assume that R and R' have the same directed Dynkin diagrams. Since R and R' have same directed Dynkin diagrams it follows that R and R' have bases $B = \{\alpha_1, \dots, \alpha_t\}$ and $B' = \{\alpha'_1, \dots, \alpha'_t\}$, respectively, such that for $i, j \in \{1, \dots, t\}$,

$$d_{ij} = \langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle = \langle \alpha'_i, \alpha'_j \rangle \langle \alpha'_j, \alpha'_i \rangle$$

and if $d_{ij} > 1$, then $\|\alpha_j\| > \|\alpha_i\|$ and $\|\alpha'_j\| > \|\alpha'_i\|$ (note that if $i, j \in \{1, \dots, t\}$, then $\langle \alpha_i, \alpha_j \rangle = \langle \alpha_j, \alpha_i \rangle = \langle \alpha'_i, \alpha'_j \rangle = \langle \alpha'_j, \alpha'_i \rangle = 2$). Let $i, j \in \{1, \dots, t\}$. We claim that $\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle$ and $\langle \alpha_j, \alpha_i \rangle = \langle \alpha'_j, \alpha'_i \rangle$. If $i = j$, then this is clear by the previous comment. Assume that $i \neq j$. By Lemma 8.4.4, the angle between α_i and α_j, and the angle between α'_i and α'_j, are obtuse. By Lemma 8.3.2 we have $d_{ij} = 0, 1, 2$ or 3. Assume that $d_{ij} = 0$. By Lemma 8.3.3 we have $\langle \alpha_i, \alpha_j \rangle = \langle \alpha_j, \alpha_i \rangle = \langle \alpha'_i, \alpha'_j \rangle = \langle \alpha'_j, \alpha'_i \rangle = 0$. Assume that $d_{ij} = 1$. By Lemma 8.3.3 we have $\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle = -1$ and $\langle \alpha_j, \alpha_i \rangle = \langle \alpha'_j, \alpha'_i \rangle = -1$. Assume that $d_{ij} = 2$. By Lemma 8.3.3 we have $\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle = -1$ and $\langle \alpha_j, \alpha_i \rangle = \langle \alpha'_j, \alpha'_i \rangle = -2$. Assume that $d_{ij} = 3$. By Lemma 8.3.3 we have $\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle = -1$ and $\langle \alpha_j, \alpha_i \rangle = \langle \alpha'_j, \alpha'_i \rangle = -3$. Our claim follows. We now have an equality of Cartan matrices:

$$C(\alpha_1, \ldots, \alpha_t) = C(\alpha'_1, \ldots, \alpha'_t).$$

By Lemma 9.2.2, R and R' are isomorphic. □

Lemma 9.3.2. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $R \subset V$ be a root system. Let D be the directed Dynkin diagram of R. Then R is irreducible if and only if D is connected.*

Proof. Assume that R is irreducible. Suppose that D is not connected. Let B be a base for R. Since D is not connected there exist proper subsets B_1 and B_2 of B such that $B = B_1 \cup B_2$ and $(B_1, B_2) = 0$. By Lemma 8.8.2 R is reducible, a contradiction. The opposite implication has a similar proof. □

9.4 Admissible systems

We will determine the isomorphism classes of irreducible root systems by introducing a new concept.

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$. Let A be a subset of V. We say that A is an **admissible system** if A satisfies the following conditions:

1. $A = \{v_1, \ldots, v_n\}$ is non-empty and linearly independent.
2. We have $(v_i, v_i) = 1$ and $(v_i, v_j) \leq 0$ for $i, j \in \{1, \ldots, n\}$ with $i \neq j$.
3. If $i, j \in \{1, \ldots, n\}$ with $i \neq j$, then $4(v_i, v_j)^2 \in \{0, 1, 2, 3\}$.

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $A = \{v_1, \ldots, v_n\} \subset V$ be an admissible system. We associate to A a graph Γ_A as follows. The vertices of Γ_A correspond to the elements of A. If $v_i, v_j \in A$ with $i \neq j$, then Γ_A has $d_{ij} = 4(v_i, v_j)^2$ edges between v_i and v_j.

We will classify all the connected Γ_A for A an admissible system. We will use these results to classify all irreducible root systems. For now, we note that there is natural connection between irreducible root systems and admissible systems that have connected graphs. Namely, supose that V is a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $R \subset V$ be an irreducible root system. Let B be a base for R. To B we associate the set A of vectors $v/\sqrt{(v, v)}$ for $v \in B$. Taking note of Lemma 8.4.4, we see that A is an admissible system; by Lemma 9.3.2, Γ_A is connected.

Lemma 9.4.1. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $A = \{v_1, \dots, v_n\} \subset V$ be an admissible system. The number of pairs of vertices $\{v, w\}$, $v \neq w$, of Γ_A that are joined by at least one edge is bounded by $\#A - 1$.*

Proof. Consider the vector $v = \sum_{i=1}^n v_i$. Since A is linearly independent, the vector v is non-zero. This implies that $(v, v) > 0$. Now

$$\begin{aligned}(v,v) &= \sum_{i,j=1}^{n} (v_i, v_j) \\ &= \sum_{i=1}^{n} (v_i, v_i) + \sum_{i,j=1,\ i\neq j}^{n} (v_i, v_j) \\ &= n + 2 \sum_{i,j=1,\ i<j}^{n} (v_i, v_j).\end{aligned}$$

Since $(v, v) > 0$, we obtain

$$n + 2 \sum_{i,j=1,\ i<j}^{n} (v_i, v_j) > 0$$

which implies

$$n > \sum_{i,j=1,\ i<j}^{n} -2(v_i, v_j).$$

Now since $(v_i, v_j) \leq 0$ for $i, j \in \{1, \dots, n\}$ with $i \neq j$, we have

$$\sum_{i,j=1,\ i<j}^{n} -2(v_i, v_j) = \sum_{i,j=1,\ i<j}^{n} \sqrt{4(v_i, v_j)^2} = \sum_{i,j=1,\ i<j}^{n} \sqrt{d_{ij}}.$$

Let N be the number of pairs $\{v_i, v_j\}$, $i, j \in \{1, \dots, n\}$, $i \neq j$, that are joined by at least one edge, i.e., for which $d_{ij} \geq 1$. We have

$$\sum_{i,j=1,\ i<j}^{n} \sqrt{d_{ij}} \geq N.$$

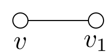
In conclusion, we find that $n > N$. This means that N is bounded by $n - 1 = \#A - 1$. □

Lemma 9.4.2. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $A \subset V$ be an admissible system. The graph Γ_A does not contain a cycle.*

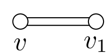
Proof. Assume that Γ_A contains a cycle; we will obtain a contradiction. Let A' be the set of edges involved in the cycle. Evidently, A' is an admissible system. Consider $\Gamma_{A'}$. Since $\Gamma_{A'}$ contains the cycle, the number of pairs of vertices of $\Gamma_{A'}$ that are joined by at least one edge is at least $\#A'$. This contradicts Lemma 9.4.1. $\square$

Lemma 9.4.3. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $A \subset V$ be an admissible system. Let v be a vertex of Γ_A, and let $v_1, \dots, v_k$ be the list of distinct vertices of Γ_A such that $w \in \{v_1, \dots, v_k\}$ if and only if v and w are incident. Then k and all the edges between v and the elements of $\{v_1, \dots, v_k\}$ are as in one of the following:*

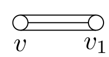
1. *$k = 1$ and*



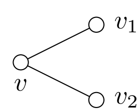
2. *$k = 1$ and*



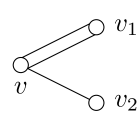
3. *$k = 1$ and*



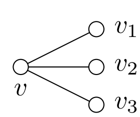
4. *$k = 2$ and*



5. *$k = 2$ and*



6. *$k = 3$ and*



Proof. By Lemma 9.4.2, Γ_A does not contain a cycle; this implies that $(v_i, v_j) = 0$ for $i, j \in \{1, \ldots, k\}$ with $i \neq j$. Consider the subspace U of V spanned by the linearly independent vectors $v_1, \ldots, v_k, v$. There exists a vector $v_0 \in U$ such that $v_0, v_1, \ldots, v_k$ is a basis for U, $(v_0, v_0) = 1$, and $(v_0, v_i) = 0$ for $i \in \{1, \ldots, k\}$. It follows that $v_0, v_1, \ldots, v_k$ is an orthonormal basis for U. Now

$$v = \sum_{i=0}^{k} (v, v_i) v_i.$$

It follows that

$$\begin{aligned}(v, v) &= (\sum_{i=0}^{k} (v, v_i) v_i, \sum_{j=0}^{k} (v, v_j) v_j) \\ &= \sum_{i=0}^{k} \sum_{j=0}^{k} (v, v_i)(v, v_j)(v_i, v_j) \\ &= \sum_{i=0}^{k} (v, v_i)^2.\end{aligned}$$

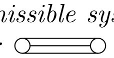
By the definition of an admissible system, $(v, v) = 1$. Therefore,

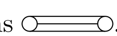
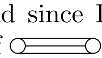
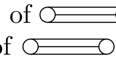
$$1 = \sum_{i=0}^{k} (v, v_i)^2.$$

Now $(v, v_0) \neq 0$ because otherwise $(v_0, U) = 0$. It follows that

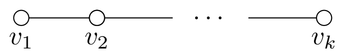
$$4 > \sum_{i=1}^{k} 4(v, v_i)^2.$$

As $4(v, v_i)^2$ is the number of edges between v and v_i, it follows that $4(v, v_i)^2 \geq 1$ for all $i \in \{1, \ldots, k\}$. We conclude that $k \leq 3$; moreover, since $4(v, v_i)^2$ is the number of edges between v and v_i for $i \in \{1, \ldots, k\}$, the possibilities are as listed in the lemma. □

Lemma 9.4.4. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $A = \{v_1, \ldots, v_n\} \subset V$ be an admissible system. Assume that Γ_A is connected and has a triple edge. Then Γ_A is:* 

Proof. By assumption, Γ_A contains . Assume that Γ_A contains another vertex w not this subgraph; we will obtain a contradiction. Since Γ_A is connected, and since Γ_A does not contain a cycle by Lemma 9.4.2, exactly one vertex v of  is on a path to w, and this path does not contain the other vertex of . It now follows that v, the vertices that are incident to v, and the edges between v and these vertices, are not as in one of the possibilities listed in Lemma 9.4.3; this is a contradiction. □

Lemma 9.4.5. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $A \subset V$ be an admissible system. Assume that Γ_A contains the line*



with no other edges between the shown vertices; here $k \geq 2$. Define

$$v = \sum_{i=1}^{k} v_i.$$

Then $v \notin A$. Define

$$A' = \big(A - \{v_1, \dots, v_k\}\big) \cup \{v\}.$$

Then A' is an admissible system, and the graph $\Gamma_{A'}$ is obtained from Γ_A by shrinking the above line to a single vertex.

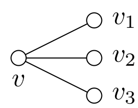
Proof. Since the set A is linearly independent and since $k \geq 2$, we must have $v \notin A$. Similarly, the set A' is linearly independent. To show that property 2 of the definition of an admissible system is satisfied by A' it will suffice to prove that $(v, v) = 1$. Now by assumption we have that $4(v_i, v_{i+1})^2 = 1$ for $i \in \{1, \dots, k-1\}$, or equivalently, $(v_i, v_{i+1}) = -1/2$ for $i \in \{1, \dots, k-1\}$. Also, by assumption, $(v_i, v_j) = 0$ for $i, j \in \{1, \dots, k\}$ $i < j$ and $j \neq i+1$. We obtain:

$$\begin{aligned}
(v, v) &= (\sum_{i=1}^{k} v_i, \sum_{j=1}^{k} v_j) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k} (v_i, v_j) \\
&= \sum_{i=1}^{k} (v_i, v_i) + 2 \sum_{i=1}^{k-1} (v_i, v_{i+1}) \\
&= \sum_{i=1}^{k} 1 + 2 \sum_{i=1}^{k-1} (-1/2) \\
&= k - (k-1) \\
&= 1.
\end{aligned}$$

To prove that property 3 of the definition of an admissible system is satisfied by A' it will suffice to prove that $4(w, v)^2 \in \{0, 1, 2, 3\}$ for $w \in A - \{v_1, \dots, v_k\}$. Let $w \in A - \{v_1, \dots, v_k\}$. If $4(w, v)^2 = 0$ then $4(w, v)^2 \in \{0, 1, 2, 3\}$. Assume that $4(w, v)^2 \neq 0$. Then $(w, v) \neq 0$. This implies that for some $i \in \{1, \dots, k\}$ we have $(w, v_i) \neq 0$, so that $4(w, v_i)^2 \neq 0$. Therefore, there is at least one edge between w and v_i. By Lemma 9.4.2, Γ_A does not contain a cycle. This implies that $(w, v_j) = 0$ for all $j \in \{1, \ldots, k\}$ with $j \neq i$. We now have $(w, v) = (w, v_i)$, so that $4(w, v)^2 = 4(w, v_i)^2 \in \{0, 1, 2, 3\}$, as desired.

Finally, consider $\Gamma_{A'}$. To see that $\Gamma_{A'}$ is obtained from Γ_A by shrinking the above line to the single vertex v it suffices to see that, for all $i \in \{1, \ldots, k\}$, if there is an edge in Γ_A between v_i and a vertex w with $w \notin \{v_1, \ldots, v_k\}$, then w is not incident to v_j for all $j \in \{1, \ldots, k\}$ with $i \neq j$; this was proven in the last paragraph. □

Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $A = \{v_1, \ldots, v_n\} \subset V$ be an admissible system. We say that a vertex v of Γ_A is a **branch vertex** of Γ_A if v is incident to three distinct vertices of Γ_A by single edges, as in the following picture:



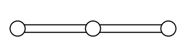
This is possibility 6 from Lemma 9.4.3.

Lemma 9.4.6. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $A = \{v_1, \ldots, v_n\} \subset V$ be an admissible system. Assume that Γ_A is connected. Then:*

1. *Γ_A has at most one double edge.*
2. *Γ_A does not have both a branch vertex and a double edge.*
3. *Γ_A has at most one branch vertex.*

Proof. By Lemma 9.4.4 we may assume that Γ_A does not contain a triple edge.

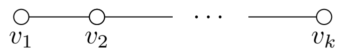
Proof of 1. Assume that Γ_A has at least two double edges; we will obtain a contradiction. Since Γ_A is connected, for every pair of double edges there exists at least one path joining a vertex of one double edge to a vertex of the other double edge; moreover, any such joining path must have at least one edge by Lemma 9.4.3. Chose a pair such that the length of the joining path is the shortest among all joining paths between pairs of double edges. Let $v_1, \ldots, v_k$ be the vertices on this shortest path, with v_1 on the first double edge, v_k on the second double edge, and v_i joined to v_{i+1} for $i \in \{1, \ldots, k-1\}$ by at least one edge. Since this is the shortest path we cannot have v_i and v_j joined by an edge for some $i, j \in \{1, \ldots, k\}$, $i < j$, and $j \neq i + 1$. Also, as this is the shortest choice, it is not the case that v_i is joined to v_{i+1} by a double edge for $i \in \{1, \ldots, k-1\}$. Let A' be as in Lemma 9.4.5; by Lemma 9.4.5, A' is an admissible system. It follows that



is a subgraph of $\Gamma_{A'}$; this contradicts Lemma 9.4.3.

The proof of 2, and then the proof of 3, are similar and will be omitted. □

Lemma 9.4.7. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $A \subset V$ be an admissible system. Assume that Γ_A contains the line*



with no other edges between the shown vertices; here $k \geq 1$. Define

$$v = \sum_{i=1}^{k} i \cdot v_i.$$

Then

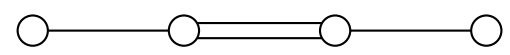
$$(v, v) = \frac{k(k+1)}{2}.$$

Proof. Since the number of edges between v_i and v_{i+1} is one for $i \in \{1, \dots, k-1\}$ it follows that $4(v_i, v_{i+1})^2 = 1$, so that $(v_i, v_{i+1}) = -1/2$ (recall that by the definition of an admissible system we have $(v_i, v_{i+1}) \leq 0$). Also, we have $(v_i, v_j) = 0$ for $i, j \in \{1, \dots, k\}$ with $i < j$ and $j \neq i + 1$. It follows that

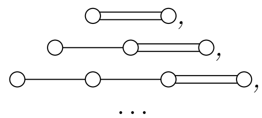
$$\begin{aligned}
(v, v) &= (\sum_{i=1}^{k} i \cdot v_i, \sum_{j=1}^{k} j \cdot v_j) \\
&= \sum_{i=1}^{k} i^2 (v_i, v_i) + 2 \sum_{i=1}^{k-1} i(i+1)(v_i, v_{i+1}) \\
&= \sum_{i=1}^{k} i^2 + 2(-1/2) \sum_{i=1}^{k-1} (i^2 + i) \\
&= k^2 + \sum_{i=1}^{k-1} i^2 - \sum_{i=1}^{k-1} i^2 - \sum_{i=1}^{k-1} i \\
&= k^2 - \sum_{i=1}^{k-1} i \\
&= k^2 - \frac{(k-1)k}{2} \\
&= \frac{2k^2 - k^2 + k}{2} \\
&= \frac{k(k+1)}{2}.
\end{aligned}$$

This completes the calculation. □

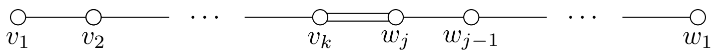
Lemma 9.4.8. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $A = \{v_1, \dots, v_n\} \subset V$ be an admissible system. Assume that Γ_A is connected. If Γ_A contains a double edge, then Γ_A is*



or one of graphs in the following list:



Proof. By Lemma 9.4.6, since Γ_A has a double edge, Γ_A has exactly one double edge, Γ_A has no triple edge, and Γ_A does not contain a branch vertex. It follows that Γ_A has the form



with no other edges between the shown vertices; here $k \geq 1$ and $j \geq 1$. Without loss of generality we may assume that $k \geq j$. Define

$$v = \sum_{i=1}^{k} i \cdot v_i, \qquad w = \sum_{i=1}^{j} i \cdot w_i.$$

By Lemma 9.4.7 we have

$$(v, v) = \frac{k(k+1)}{2}, \qquad (w, w) = \frac{j(j+1)}{2}.$$

We have $4(v_k, w_j)^2 = 2$ since there is a double edge joining v_k and v_j, and $(v_i, w_\ell) = 0$ since no edge joins v_i and w_ℓ for all $i \in \{1, \ldots, k\}$ and $\ell \in \{1, \ldots, j\}$ with $i \neq k$ or $\ell \neq j$. It follows that

$$\begin{aligned}(v, w) &= (\sum_{i=1}^{k} i \cdot v_i, \sum_{\ell=1}^{j} \ell \cdot w_\ell) \\ &= kj(v_k, w_j),\end{aligned}$$

so that

$$(v, w)^2 = k^2 j^2 (v_k, w_j)^2 = \frac{k^2 j^2}{2}.$$

By the Cauchy-Schwarz inequality we have

$$(v, w)^2 < (v, v)(w, w);$$

Note that v and w are linearly independent, so that the inequality is strict. Substituting, we obtain:

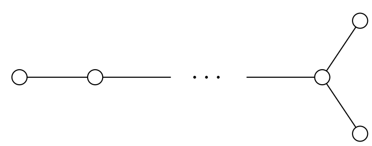
$$\begin{aligned}\frac{k^2 j^2}{2} &< \frac{k(k+1)}{2} \frac{j(j+1)}{2}, \\ 2k^2 j^2 &< k(k+1)j(j+1),\end{aligned}$$

$$\begin{aligned} 2k^2j^2 &< k^2j^2 + jk^2 + j^2k + jk, \\ 2kj &< kj + k + j + 1, \\ kj &< k + j + 1, \\ kj - k - j &< 1, \\ kj - k - j + 1 &< 2, \\ (k-1)(j-1) &< 2. \end{aligned}$$

Recalling that $k \geq j \geq 1$, we find that $k = j = 2$, or k is an arbitrary positive integer and $j = 1$. This proves the lemma. $\square$

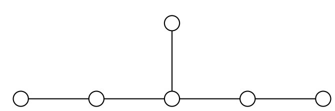
Lemma 9.4.9. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $A = \{v_1, \ldots, v_n\} \subset V$ be an admissible system. Assume that Γ_A is connected, and that Γ_A has a branch vertex. Then Γ_A is either*

D_ℓ, $\ell \geq 4$:



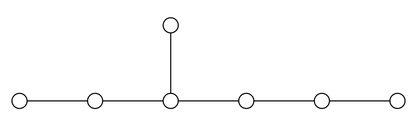
or

E_6 :



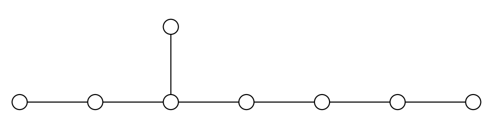
or

E_7 :



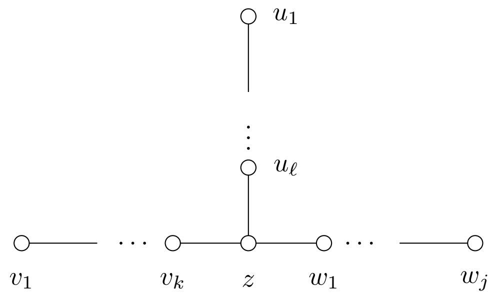
or

E_8 :



.

Proof. By Lemma 9.4.4 and Lemma 9.4.6, since Γ_A is connected and contains a double edge, Γ_A contains exactly one branch vertex, no double edges, and no triple edges. It follows that Γ_A has the form



with $k \geq j \geq \ell$. We define

$$v = \sum_{i=1}^{k} i \cdot v_i, \qquad w = \sum_{i=1}^{j} i \cdot w_i, \qquad u = \sum_{i=1}^{\ell} i \cdot u_i.$$

Since there are no edges between the vertices in $\{v_1, \ldots, v_k\}$ and the vertices in $\{w_1, \ldots, v_j\}$, the vectors v and w are orthogonal. Similarly, v and u are orthogonal, and w and u are orthogonal. Define

$$v' = \frac{v}{\|v\|}, \qquad w' = \frac{w}{\|w\|}, \qquad u' = \frac{u}{\|u\|}.$$

The vectors v', w' and u' are also mutually orthogonal, and have norm one. Let U be the subspace of V spanned by v', w', u' and z. This space is four-dimensional as these vectors are linearly independent. The orthonormal vectors v', w', u' can be extended to an orthonormal basis v', w', u', z' for U. We have

$$z = (z, v')v' + (z, w')w' + (z, u')u' + (z, z')z'$$

so that

$$1 = (z, z) = (z, v')^2 + (z, w')^2 + (z, u')^2 + (z, z')^2.$$

The vector z' cannot be orthogonal to z; otherwise, $(z', U) = 0$, a contradiction. Since $(z, z')^2 > 0$, we obtain

$$(z, v')^2 + (z, w')^2 + (z, u')^2 < 1.$$

Now

$$\begin{aligned} (z, v')^2 &= \frac{(z, v)^2}{(v, v)} \\ &= \frac{2(z, \sum_{i=1}^{k} i v_i)^2}{k(k+1)} \\ &= \frac{2k^2 (z, v_k)^2}{k(k+1)} \\ &= \frac{k}{2(k+1)}. \end{aligned}$$

Similarly,

$$(z, w')^2 = \frac{j}{2(j+1)} \qquad \text{and} \qquad (z, u')^2 = \frac{\ell}{2(\ell+1)}.$$

Substituting, we get:

$$\begin{aligned}
\frac{k}{2(k+1)} + \frac{j}{2(j+1)} + \frac{\ell}{2(\ell+1)} &< 1, \\
\frac{k+1}{2(k+1)} - \frac{1}{2(k+1)} + \frac{j+1}{2(j+1)} - \frac{1}{2(j+1)} + \frac{\ell+1}{2(\ell+1)} - \frac{1}{2(\ell+1)} &< 1, \\
\frac{1}{2} - \frac{1}{2(k+1)} + \frac{1}{2} - \frac{1}{2(j+1)} + \frac{1}{2} - \frac{1}{2(\ell+1)} &< 1, \\
\frac{3}{2} - \frac{1}{2(k+1)} - \frac{1}{2(j+1)} - \frac{1}{2(\ell+1)} &< 1, \\
3 - \frac{1}{k+1} - \frac{1}{j+1} - \frac{1}{\ell+1} &< 2, \\
\frac{1}{k+1} + \frac{1}{j+1} + \frac{1}{\ell+1} &> 1.
\end{aligned}$$

Now $k \geq j \geq \ell \geq 1$. Hence,

$$k+1 \geq j+1 \geq \ell+1 \geq 2$$

and thus

$$\frac{1}{k+1} \leq \frac{1}{j+1} \leq \frac{1}{\ell+1} \leq \frac{1}{2}.$$

It follows that

$$\begin{aligned}
\frac{1}{k+1} + \frac{1}{j+1} + \frac{1}{\ell+1} &> 1, \\
\frac{1}{\ell+1} + \frac{1}{\ell+1} + \frac{1}{\ell+1} &> 1, \\
\frac{3}{\ell+1} &> 1, \\
3 &> \ell+1, \\
2 &> \ell.
\end{aligned}$$

Hence, $\ell = 1$. Substituting $\ell = 1$, we have:

$$\begin{aligned}
\frac{1}{k+1} + \frac{1}{j+1} + \frac{1}{1+1} &> 1, \\
\frac{1}{k+1} + \frac{1}{j+1} &> \frac{1}{2}, \\
\frac{1}{j+1} + \frac{1}{j+1} &> \frac{1}{2}, \\
\frac{2}{j+1} &> \frac{1}{2},
\end{aligned}$$

$$\begin{aligned} \frac{2}{j+1} &> \frac{1}{2}, \\ 3 &> j. \end{aligned}$$

It follows that $j = 1$ or $j = 2$. Assume that $j = 2$. Then the inequality is:

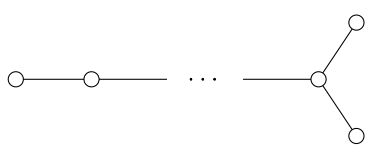
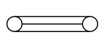
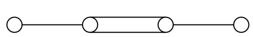
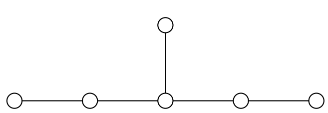
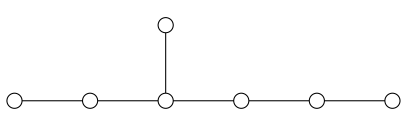
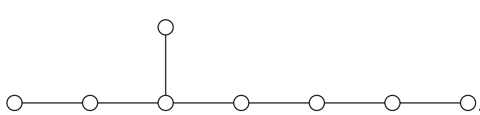
$$\begin{aligned} \frac{1}{k+1} + \frac{1}{2+1} + \frac{1}{1+1} &> 1, \\ \frac{1}{k+1} + \frac{5}{6} &> 1, \\ \frac{1}{k+1} &> \frac{1}{6}, \\ 5 &> k. \end{aligned}$$

This implies that $k = 3$ or $k = 4$. In summary we have found that

$$(k, j, \ell) \in \{(k, 1, 1) : k \in \mathbb{Z}, k \geq 1\} \cup \{(2, 2, 1), (3, 2, 1), (4, 2, 1)\}.$$

This is the assertion of the lemma. □

Theorem 9.4.10. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot, \cdot)$, and let $A = \{v_1, \ldots, v_n\} \subset V$ be an admissible system. Assume that Γ_A is connected. Then Γ_A is one of the following:*

1. *(ℓ vertices, $\ell \geq 1$)* 
2. *(ℓ vertices, $\ell \geq 2$)* 
3. *(ℓ vertices, $\ell \geq 3$)* 
4. 
5. 
6. 
7. 
8. 

Proof. Let ℓ be the number of vertices of Γ_A. If $\ell = 1$, then Γ_A is as in 1 with $\ell = 1$. Assume that $\ell \geq 2$. By Lemma 9.4.3, there exist no two vertices of Γ_A joined by four or more vertices.

Assume that Γ_A has a triple edge. By Lemma 9.4.4, Γ_A is as in 4. Assume for the remainder of the proof that Γ_A does not have a triple edge.

Assume that Γ_A has a double edge. Then by Lemma 9.4.8, Γ_A must be as in 2 or 5. Assume for the remainder of the proof that Γ_A does not have a double edge.

Assume that Γ_A has a branch vertex. By Lemma 9.4.9, Γ_A must be as in 3, 6, 7, or 8. Assume for the remainder of the proof that Γ_A does not have a branch vertex.

Since no two vertices of Γ_A are joined by two or or more vertices, since Γ_A does not have a branch vertex, and since Γ_A does not contain a cycle by Lemma 9.4.2, it follows that Γ_A is as in 1. $\square$

9.5 Possible Dynkin diagrams

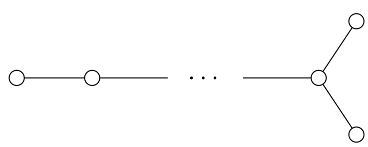
Theorem 9.5.1. *Let V be a finite-dimensional vector space over $\mathbb{R}$ equipt with an inner product $(\cdot,\cdot)$, and let $R \subset V$ be a root system. Assume that R is irreducible. Let D be the Dynkin diagram of R. Then D belongs to one of the following infinite families (each of which has ℓ vertices)*

$A_\ell, \ell \geq 1$:

$B_\ell, \ell \geq 2$:

$C_\ell, \ell \geq 3$:

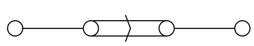
D_ℓ, $\ell \geq 4$:



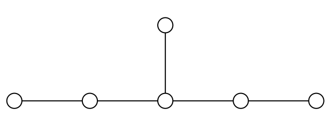
or D is one of the following five diagrams

G_2:

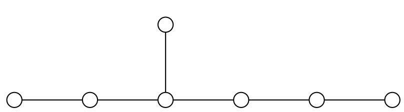
F_4:

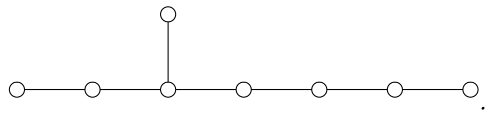


E_6:



E_7:



E_8: 

Proof. Let B a base for R. Let A be the admissible system associated to R and B as at the beginning of Section 9.4. Let C be the Coxeter graph of R; this is the same as Γ_A, the graph associated to A. By Theorem 9.4.10, $\Gamma_A = C$ must be one of the graphs listed in this theorem. This implies the result. □

Chapter 10

The classical Lie algebras

Let F have characteristic zero and be algebraically closed. The classical Lie algebras over F are $\mathrm{sl}(\ell+1,F)$, $\mathrm{so}(2\ell+1,F)$, $\mathrm{sp}(2\ell,F)$, and $\mathrm{so}(2\ell,F)$ for ℓ a positive integer. In this chapter we will prove that these Lie algebras are simple (with the exception of $\mathrm{so}(2\ell,F)$ when $\ell=1$ or $\ell=2$). We will also determine the root systems associated to these classical Lie algebras.

10.1 Definitions

$\mathrm{sl}(\ell+1,F)$

Let F have characteristic zero and be algebraically closed, and let ℓ be a positive integer. We define $\mathrm{sl}(\ell+1,F)$ to be the F-subspace of $g\in\mathrm{gl}(\ell+1,F)$ such that $\mathrm{tr}(g)=0$. The bracket on $\mathrm{sl}(\ell+1,F)$ is inherited from $\mathrm{gl}(\ell+1,F)$, and is defined by $[X,Y]=XY-YX$ for $X,Y\in\mathrm{sl}(\ell+1,F)$. Note that $[X,Y]\in\mathrm{sl}(\ell+1,F)$ for $X,Y\in\mathrm{sl}(\ell+1,F)$ because $\mathrm{tr}([X,Y])=\mathrm{tr}(XY)-\mathrm{tr}(YX)=\mathrm{XY}-\mathrm{XY}=0$. The bracket on $\mathrm{sl}(\ell+1,F)$ satisfies 1 and 2 of the definition of Lie algebra from Section 1.3 because the bracket on $\mathrm{gl}(\ell+1,F)$ satisfies these properties by Proposition 1.4.1.

Lemma 10.1.1. *Let n be a positive integer. Let $S\in\mathrm{gl}(n,F)$. Let L be the F-subspace of $X\in\mathrm{gl}(n,F)$ such that*

$$^{t}XS+SX=0.$$

With the bracket inherited from $\mathrm{gl}(n,F)$, so that $[X,Y]=XY-YX$ for $X,Y\in L$, the subspace L is a Lie subalgebra of $\mathrm{gl}(n,F)$. Moreover, if S is invertible, then $L\subset\mathrm{sl}(n,F)$.

Proof. Let $X,Y\in L$. Then

$$^{t}[X,Y]S+S[X,Y]={}^{t}(XY-YX)S+S(XY-YX)$$

$$\begin{aligned}
&= ({}^tY{}^tX - {}^tX{}^tY)S + SXY - SYX \\
&= {}^tY{}^tXS - {}^tX{}^tYS + SXY - SYX \\
&= -{}^tYSX + {}^tXSY + SXY - SYX \\
&= SYX - SXY + SXY - SYX \\
&= 0.
\end{aligned}$$

It follows that $[X, Y] \in L$. The bracket on L satisfies 1 and 2 of the definition of Lie algebra from Section 1.3 because the bracket on $\mathrm{gl}(n, F)$ satisfies these properties by Proposition 1.4.1. Assume that S is invertible. Let $X \in L$; we need to prove that $\mathrm{tr}(X) = 0$. We have

$$\begin{aligned}
{}^tXS + SX &= 0 \\
{}^tXS &= -SX \\
{}^tX &= -SXS^{-1} \\
\mathrm{tr}({}^tX) &= \mathrm{tr}(-SXS^{-1}) \\
\mathrm{tr}(X) &= -\mathrm{tr}(S^{-1}SX) \\
\mathrm{tr}(X) &= -\mathrm{tr}(X).
\end{aligned}$$

Since F has characteristic zero, this implies that $\mathrm{tr}(X) = 0$. □

$\mathrm{so}(2\ell + 1, F)$

Let F have characteristic zero and be algebraically closed, and let ℓ be a positive integer. Let $S \in \mathrm{gl}(2\ell + 1, F)$ be the matrix

$$S = \begin{bmatrix} 1 & & \\ & & 1_\ell \\ & 1_\ell & \end{bmatrix}.$$

Here, 1_ℓ is the $\ell \times \ell$ identity matrix. We define $\mathrm{so}(2\ell + 1, F)$ to be the Lie subalgebra of $\mathrm{gl}(2\ell + 1, F)$ defined by S as in Lemma 10.1.1. By Lemma 10.1.1, since S is invertible, we have $\mathrm{so}(2\ell + 1, F) \subset \mathrm{sl}(2\ell + 1, F)$.

$\mathrm{sp}(2\ell, F)$

Let F have characteristic zero and be algebraically closed, and let ℓ be a positive integer. Let $S \in \mathrm{gl}(2\ell, F)$ be the matrix

$$S = \begin{bmatrix} & 1_\ell \\ -1_\ell & \end{bmatrix}.$$

Here, 1_ℓ is the $\ell \times \ell$ identity matrix. We define $\mathrm{sp}(2\ell, F)$ to be the Lie subalgebra of $\mathrm{gl}(2\ell, F)$ defined by S as in Lemma 10.1.1. By Lemma 10.1.1, since S is invertible, we have $\mathrm{sp}(2\ell, F) \subset \mathrm{sl}(2\ell, F)$.

$\mathrm{so}(2\ell, F)$

Let F have characteristic zero and be algebraically closed, and let ℓ be a positive integer. Let $S \in \mathrm{gl}(2\ell + 1, F)$ be the matrix

$$S = \begin{bmatrix} & 1_\ell \\ 1_\ell & \end{bmatrix}.$$

Here, 1_ℓ is the $\ell \times \ell$ identity matrix. We define $\mathrm{so}(2\ell, F)$ to be the Lie subalgebra of $\mathrm{gl}(2\ell, F)$ defined by S as in Lemma 10.1.1. By Lemma 10.1.1, since S is invertible, we have $\mathrm{so}(2\ell, F) \subset \mathrm{sl}(2\ell, F)$.

10.2 A criterion for semi-simplicity

Lemma 10.2.1. *Assume that F has characteristic zero and is algebraically closed. Let L be a finite-dimensional Lie algebra over F.*

1. *Assume that L is reductive. Then*

$$L = [L, L] \oplus Z(L)$$

   as Lie algebras, and $[L, L]$ is semi-simple.

2. *Assume that V is a finite-dimensional vector space over F. Let L be a non-zero Lie subalgebra of $\mathrm{gl}(V)$, and assume that L acts irreducibly on V. Then L is reductive and $\dim Z(L) \leq 1$. If L is contained in $\mathrm{sl}(V)$, then L is semi-simple.*

Proof. Proof of 1. Assume that L is reductive. By Lemma 2.1.10, $L/Z(L)$ is semi-simple. Consider the ad action of $L/Z(L)$ on L. By Theorem 6.2.4, Weyl's Theorem, this action is completely reducible; it follows that the ad action of L on L is also completely reducible. Therefore, the L-submodule $Z(L)$ has a complement, i.e., there exists an L-submodule M of L such that $L = M \oplus Z(L)$ as F-vector spaces. Since L acts on L via the ad action, M is an ideal of L. We claim that $M = [L, L]$. Let $x, y \in L$, and write $x = m + u$ and $y = n + v$ with $m, n \in M$ and $u, v \in Z(L)$. Then

$$[x, y] = [m + u, n + v] = [m, n] + [m, v] + [u, n] + [u, v] = [m, n].$$

Therefore, $[x, y] \in [M, M] \subset M$. It follows that $[L, L] \subset M$. Now by Lemma 6.2.2, since $L/Z(L)$ is semi-simple, we have $[L/Z(L), L/Z(L)] = L/Z(L)$. This implies that $([L, L] + Z(L))/Z(L) = L/Z(L)$, so that

$$\dim[L, L] + \dim Z(L) = \dim L.$$

Since now $\dim[L, L] = \dim L - \dim Z(L) = \dim M$, we conclude that $[L, L] = M$. Hence, $L = [L, L] \oplus Z(L)$ as Lie algebras. Since $L = [L, L] \oplus Z(L)$ as Lie algebras we obtain $[L, L] \cong L/Z(L)$ as Lie algebras; since $L/Z(L)$ is semi-simple, we conclude that $[L, L]$ is semi-simple.

Proof of 2. Let $R = \mathrm{Rad}(L)$. By definition, R is a solvable ideal of L. By Lemma 3.4.1, there exists a non-zero vector $v \in V$, and a linear functional $\lambda : R \to F$ such that $rv = \lambda(r)v$ for all $r \in R$. Let $x \in L$ and $r \in R$. Then $[x, r] \in R$ since R is an ideal. Hence,

$$\begin{aligned} [x,r]v &= \lambda([x,r])v \\ xrv - rxv &= \lambda([x,r])v \\ -r(xv) &= -\lambda(r)xv + \lambda([x,r])v \\ r(xv) &= \lambda(r)xv + \lambda([r,x])v. \end{aligned}$$

By assumption, the action of L on V is irreducible. This implies that the vectors xv for $x \in L$ span V. Therefore, there exists vectors $v_1, \dots, v_m$ in V such that $v_1, \dots, v_m, v$ is an ordered basis for V, and constants $c_1, \dots, c_m$ such that

$$rv_i = \lambda(r)v_i + c_i v$$

for $r \in R$ and $i \in \{1, \dots, m\}$. If $r \in R$, then the matrix of r in the basis $v_1, \dots, v_m, v$ is

$$\begin{bmatrix} \lambda(r) & & & c_1 \\ & \ddots & & \vdots \\ & & \lambda(r) & c_m \\ & & & \lambda(r) \end{bmatrix}.$$

In particular, we see that the $\mathrm{tr}(r) = \lambda(r) \cdot \dim V$. Consider $[L, R]$. This ideal of L is contained in R, and we have $\mathrm{tr}([L, R]) = 0$. It follows that $\lambda([L, R]) = 0$. From this, we conclude that in fact

$$r(xv) = \lambda(r)xv$$

for $r \in R$ and $x \in L$. Since the action of L on V is irreducible, it follows that $r \in R$ acts by $\lambda(r)$, i.e., the elements of R are contained in $F \subset \mathrm{gl}(V)$. Thus, $R \subset Z(L)$, so that $R = Z(L)$ and L is hence reductive. Also, $\dim Z(L) = \dim R \leq 1$. Finally, assume that $L \subset \mathrm{sl}(V)$. Then $\mathrm{tr}(x) = 0$ for all $x \in L$. Since $R \subset F \subset \mathrm{gl}(V)$, this implies that $R = 0$; i.e., L is semi-simple. □

10.3 A criterion for simplicity

Lemma 10.3.1. *Let L be a Lie algebra over F, and $S \subset L$ be a subset. Let K be the subalgebra of L generated by S. Let $X \in L$. If $[X, S] = 0$, then $[X, K] = 0$. If $[X, S] \subset K$, then $[X, K] \subset K$.*

Proof. Assume that $[X, S] = 0$. Inductively define subsets $K_1, K_2, K_3, \dots$ by letting $K_1 = S$ and

$$K_k = \bigcup_{i=1}^{k-1} \{[Y, Z] : Y \in K_i, Z \in K_{k-i}\}.$$

Evidently, every element of K is a linear combination of elements from the union $\cup_{k=1}^{\infty} K_k$. Thus, to prove that $[X, K] = 0$ it suffices to prove that $[X, K_k] = 0$ for all positive integers k. We will prove this by induction on k. The case $k = 1$ follows by hypothesis. Let k be a positive integer and that $[X, K_\ell] = 0$ for all positiver integers $\ell \leq k$; we will prove that $[X, K_{k+1}] = 0$. To prove this will suffice to prove that for every pair of positive integers i and j such that $i + j = k + 1$ we have $[X, [Y, Z]] = 0$ for $Y \in K_i$ and $Z \in K_j$. Let i and j be positive integers such that $i + j = k + 1$ and let $Y \in K_i$ and $Z \in K_j$. By the Jacobi identity and the induction hypothesis we have

$$\begin{aligned}
[X, [Y, Z]] &= -[Y, [Z, X]] - [Z, [X, Y]] \\
&= -[Y, 0] - [Z, 0] \\
&= 0.
\end{aligned}$$

We now obtain $[X, K] = 0$ by induction.

To prove the second assertion of the lemma, assume that $[X, S] \subset K$. To prove that $[X, K] \subset K$ it will suffice to prove that $[X, K_k] \subset K$ for all positive integers k. We will prove this by induction on k. The case $k = 1$ is the hypothesis $[X, S] \subset K$. Let k be a positive integer, and assume that $[X, K_\ell] \subset K$ for all positive integers $\ell \leq k$; we will prove that $[X, K_{k+1}] \subset K$. To prove this will suffice to prove that for every pair of positive integers i and j such that $i + j = k + 1$ we have $[X, [Y, Z]] \in K$ for $Y \in K_i$ and $Z \in K_j$. Let i and j be positive integers such that $i + j = k + 1$ and let $Y \in K_i$ and $Z \in K_j$. By the Jacobi identity we have

$$[X, [Y, Z]] = -[Y, [Z, X]] - [Z, [X, Y]].$$

By the induction hypothesis, $[Z, X] = -[X, Z], [X, Y] \in K$. Since $Y, Z \in K$ we obtain $[Y, [Z, X]], [Z, [X, Y]] \in K$. It now follows that $[X, [Y, Z]] \in K$, as desired. We have proven that $[X, K] \subset K$ by induction. □

Proposition 10.3.2. *Let F have characteristic zero and be algebraically closed. Let L be a semi-simple finite-dimensional Lie algebra over F. Let H be a Cartan subalgebra of L, and let Φ be the root system associated to the pair (L, H) as in Section 8.2. Then L is simple if and only if Φ is irreducible.*

Proof. To begin, we recall that as in Section 8.2 we have

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

Assume that L is simple. Assume that Φ is not irreducible; we will obtain a contradiction. Since Φ is not irreducible, there exist non-empty subsets Φ_1 and Φ_2 of Φ such that $\Phi_1 \cap \Phi_2 = \emptyset$ and $(\Phi_1, \Phi_2) = 0$. Let K be the subalgebra generated by the L_α for $\alpha \in \Phi_1$. We claim that K is a non-zero, proper ideal of L; this will contradict the assumption that L is simple. It is clear that K is non-zero because Φ_1 is non-empty.

To prove that K is a proper ideal of L we will first prove that $[L_\beta, K] = 0$ for $\beta \in \Phi_2$. Let $\beta \in \Phi_2$. By Lemma 10.3.1, to prove that $[L_\beta, K] = 0$ it will suffice to prove that $[L_\beta, L_\alpha] = 0$ for $\alpha \in \Phi_1$. Let $\alpha \in \Phi_1$. Now by Proposition 7.0.3, $[L_\beta, L_\alpha] \subset L_{\alpha+\beta}$. Assume that $L_{\alpha+\beta} \neq 0$; we will obtain a contradiction. Consider $\alpha + \beta$. We have $(\alpha + \beta, \alpha) = (\alpha, \alpha) + (\beta, \alpha) = (\alpha, \alpha) + 0 = (\alpha, \alpha) > 0$; this implies that $\alpha + \beta \neq 0$. Since $L_{\alpha+\beta} \neq 0$, and since $\alpha + \beta \neq 0$, we have, by definition, $\alpha + \beta \in \Phi$. Hence, $\alpha + \beta \in \Phi_1$ or $\alpha + \beta \in \Phi_2$. If $\alpha + \beta \in \Phi_1$, then $(\alpha + \beta, \beta) = 0$; since $(\alpha + \beta, \beta) = (\beta, \beta) > 0$, this is a contradiction. Similarly, if $\alpha + \beta \in \Phi_2$, then $(\alpha + \beta, \alpha) = 0$, a contradiction. It follows that $L_{\alpha+\beta} = 0$, implying that $[L_\beta, L_\alpha] = 0$. Hence, $[L_\beta, K] = 0$ for all $\beta \in \Phi_2$.

To see that K is proper, assume that $K = L$. Then $[L_\beta, L] = [L_\beta, K] = 0$ for all $\beta \in \Phi_2$. This means that $L_\beta \subset Z(L)$ for all $\beta \in \Phi_2$; since $Z(L) = 0$ (because L is simple), and since Φ_2 is non-empty, this is a contradiction. Thus, K is proper.

Finally, we need to prove that K is an ideal of L. By Lemma 10.3.1, since $L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$, to prove this it will suffice to prove that $[H, L_\alpha] \subset K, [L_\gamma, L_\alpha] \subset K$ and $[L_\beta, L_\alpha] \subset K$ for all $\alpha \in \Phi_1$, $\gamma \in \Phi_1$, and $\beta \in \Phi_1$. Let $\alpha \in \Phi_1$, $\gamma \in \Phi_1$, and $\beta \in \Phi_1$. Then $[H, L_\alpha] \subset L_\alpha$ by the definition of L_α. Since $L_\alpha \subset K$, we get $[H, L_\alpha] \subset K$. We have $[L_\gamma, L_\alpha] \subset K$ by the definition of K. Finally, we have already proven that $[L_\beta, L_\alpha] = 0$, so that $[L_\beta, L_\alpha] \subset K$. It follows that K is an ideal of K, completing the argument that L is irreducible.

Next, assume that Φ is irreducible, and that L contains a non-zero, proper ideal I; we will obtain a contradiction. Since I is an ideal, the mutually commuting operators $\mathrm{ad}(h) \in \mathrm{gl}(L)$ for $h \in H$ preserve the subspace I. Since every element of H is semi-simple, the elements of $\mathrm{ad}(H) \subset \mathrm{gl}(L)$ are diagonalizable (recall the definition of the abstract Jordan decomposition, and in particular, the definition of semi-simple). The restrictions $\mathrm{ad}(h)|_I$ for $h \in H$ are therefore also diagonalizable. Since the F-subspaces L_α for $\alpha \in \Phi$ are one-dimensional by Proposition 7.0.8, it follows that there exist an F-subspace H_1 of H and a subset Φ_1 of Φ such that

$$I = H_1 \oplus \bigoplus_{\alpha \in \Phi_1} L_\alpha.$$

By Lemma 5.3.3 the subspace $I^\perp$ of L is also an ideal of L. Hence, there also exist an F-subspace H_2 of H and a subset Φ_2 of Φ such that

$$I^\perp = H_2 \oplus \bigoplus_{\beta \in \Phi_2} L_\beta.$$

By Lemma 5.4.3 we have $L = I \oplus I^\perp$. This implies that $H = H_1 \oplus H_2$ and that there is a disjoint decomposition $\Phi = \Phi_1 \sqcup \Phi_2$. Assume that Φ_1 is empty; we will obtain a contradiction. Since Φ_1 is empty, we must have $\Phi_2 = \Phi$, so that $L_\beta \subset I^\perp$ for all $\beta \in \Phi$. By Proposition 7.0.14, $L \subset I^\perp$, implying that $I^\perp = L$ and hence $I = 0$, a contradiction. Thus, Φ_1 is non-empty. Similarly, Φ_2 is non-empty. Let $\alpha \in \Phi_1$ and $\beta \in \Phi_2$; we claim that $(\alpha, \beta) = 0$. We have, by 3 of Lemma 7.0.11,

$$\langle \alpha, \beta \rangle = \frac{2(\alpha, \beta)}{(\beta, \beta)} = \alpha(h_\beta).$$

Also, by the definition of L_α,

$$\alpha(h_\beta)e_\alpha = [h_\beta, e_\alpha].$$

Consider $[h_\beta, e_\alpha]$. On the one hand, since $e_\alpha \in L_\alpha \subset I$, and since I is an ideal of L, we have $[h_\beta, e_\alpha] \in I$. On the other hand, $h_\beta = [e_\beta, f_\beta]$; since $f_\beta \in I^\perp$, and $I^\perp$; we must have $h_\beta \in I^\perp$. Using again that $I^\perp$ is an ideal, we see that $[h_\beta, e_\alpha] \in I^\perp$. Now we have $[h_\beta, e_\alpha] \in I \cap I^\perp = 0$, proving that $[h_\beta, e_\alpha] = 0$. It follows from above that $\alpha(h_\beta) = 0$, and hence that $\langle \alpha, \beta \rangle = 0$, as claimed. This contradicts the irreducibility of Φ. □

10.4 A criterion for Cartan subalgebras

Lemma 10.4.1. *Let F have characteristic zero and be algebraically closed. Let n be a positive integer. Let $h \in \mathrm{gl}(n, F)$ be diagonalizable. Then $\mathrm{ad}(h) : \mathrm{gl}(n, F) \to \mathrm{gl}(n, F)$ is diagonalizable.*

Proof. Since h is diagonalizable, there exists a matrix $A \in \mathrm{GL}(n, F)$ such that AhA^{-1} is diagonal. Let $d = AhA^{-1}$, and let

$$d = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$$

Consider $\mathrm{ad}(d)$. Let $i, j \in \{1, \dots, n\}$. We have

$$\begin{aligned} \mathrm{ad}(d)(e_{ij}) &= [d, e_{ij}] \\ &= de_{ij} - e_{ij}d \\ &= d_i e_{ij} - d_j e_{ij} \\ &= (d_i - d_j)e_{ij}. \end{aligned}$$

Thus, e_{ij} is an eigenvector for d with eigenvalue $d_i - d_j$. Since the set $\{e_{ij} : 1 \le i, j \le n\}$ is a basis for $\mathrm{gl}(n, F)$ it follows that $\mathrm{ad}(d)$ is diagonalizable. Now assume that $x \in \mathrm{gl}(n, F)$ is an eigenvector for $\mathrm{ad}(d)$ with eigenvalue λ. We have

$$\begin{aligned} \mathrm{ad}(h)\big(A^{-1}xA\big) &= hA^{-1}xA - A^{-1}xAh \\ &= A^{-1}(AhA^{-1}x - xAhA^{-1})A \\ &= A^{-1}[d, x]A \\ &= A^{-1}\mathrm{ad}(d)(x)A \\ &= \lambda A^{-1}xA. \end{aligned}$$

It follows that $A^{-1}xA$ is an eigenvector for $\mathrm{ad}(h)$ with eigenvalue λ. Since the vectors $A^{-1}e_{ij}A$ for $i, j \in \{1, \dots, n\}$ are basis for $\mathrm{gl}(n, F)$ and are eigenvectors for $\mathrm{ad}(h)$, it follows that $\mathrm{ad}(h)$ is diagonalizable. □

We remark that the content of the above lemma is already contained in Lemma 5.1.3.

Lemma 10.4.2. *Let F have characteristic zero and be algebraically closed. Let n be a positive integer, and let L be a Lie subalgebra of $\mathrm{gl}(n, F)$. Let H be the abelian subalgebra of L consisting of the diagonal matrices in L; assume that H is non-zero. Let W be the F-subspace of L consisting of elements with zeros on the main diagonal. Assume that no non-zero element of W commutes with all the elements of H, i.e.,*

$$\{x \in W : \mathrm{ad}(h)(x) = [h, x] = 0, h \in H\} = 0.$$

Then H is a Cartan subalgebra of L.

Proof. Evidently, H is abelian. Also, by Lemma 10.4.1, the operators $\mathrm{ad}(h) : \mathrm{gl}(n, F) \to \mathrm{gl}(n, F)$ for $h \in H$ are diagonalizable. To prove that H is a Cartan subalgebra it will suffice to prove that if H' is an abelian subalgebra of L, and $H \subset H'$, then $H = H'$. Assume that H' is an abelian subalgebra of L such that every element of H' and $H \subset H'$. Let $x \in H'$. Now

$$L = H \oplus W.$$

The operators $\mathrm{ad}(h)$ for $h \in H$ leave the subspace W invariant; since $\mathrm{ad}(h)$ is diagonalizable, it follows that $\mathrm{ad}(h)|_W$ is diagonalizable for $h \in H$. For a linear functional $\beta : H \to F$, let

$$W_\beta = \{x \in W : \mathrm{ad}(h)x = \beta(h)x, h \in H\},$$

and let B be the set of linear functionals $\beta : H \to F$ such that $W_\beta \neq 0$. There is a direct sum decomposition

$$W = \bigoplus_{\beta \in B} W_\beta.$$

and hence a direct sum decomposition

$$L = H \oplus \bigoplus_{\beta \in B} W_\beta.$$

The assumption of the lemma is that $0 \notin B$, i.e., $\beta \neq 0$ for all $\beta \in B$. Write

$$x = x_0 + \sum_{\beta \in B} x_\beta$$

where $x_0 \in H$ and $x_\beta \in W_\beta$ for $\beta \in B$. Let $h \in H$. Then $\mathrm{ad}(h)x = [h, x] = 0$ because $h, x \in H'$ and H' is abelian. Applying $\mathrm{ad}(h)$ to the above sum yields

$$\mathrm{ad}(h)x = \mathrm{ad}(h)x_0 + \sum_{\beta \in B} \mathrm{ad}(h)(x_\beta)$$

$$0 = 0 + \sum_{\beta \in B} \beta(h) x_\beta$$
$$0 = \sum_{\beta \in B} \beta(h) x_\beta.$$

Since the subspaces W_β for $\beta \in B$ form a direct sum, we must have $\beta(h)x_\beta = 0$ for all $\beta \in B$ and $h \in H$. Since every $\beta \in B$ is non-zero, we must have $x_\beta = 0$ for all $\beta \in B$. This implies that $x = x_0 \in H$, as desired. □

10.5 The Killing form

Lemma 10.5.1. *Let F have characteristic zero and be algebraically closed. Let n be a positive integer. For $x, y \in \mathrm{gl}(n, F)$ define*

$$t(x, y) = \mathrm{tr}(xy).$$

The function $t : \mathrm{gl}(n, F) \times \mathrm{gl}(n, F) \to F$ is an associative, symmetric bilinear form. If L is a Lie subalgebra of $\mathrm{gl}(n, F)$, L is simple, and the restriction of t to $L \times L$ is non-zero, then L is non-degenerate.

Proof. It is clear that t is F-linear in each variable. Also, t is symmetric because $\mathrm{tr}(xy) = \mathrm{tr}(yx)$ for $x, y \in \mathrm{gl}(n, F)$. To see that t is associative, let $x, y, z \in \mathrm{gl}(n, F)$. Then

$$\begin{aligned} t(x, [y, z]) &= \mathrm{tr}(x(yz - zy)) \\ &= \mathrm{tr}(xyz) - \mathrm{tr}(xzy) \\ &= \mathrm{tr}(xyz) - \mathrm{tr}(yxz) \\ &= \mathrm{tr}((xy - yx)z) \\ &= t([x, y], z). \end{aligned}$$

Assume that L is a subalgebra of $\mathrm{gl}(n, F)$, L is simple, and the restriction of t to $L \times L$ is non-zero. Let $J = \{y \in L : t(x, y) = 0, x \in L\}$. We need to prove that $J = 0$. We claim that J is an ideal of L. Let $y \in L$ and $z \in J$; we need to see that $[y, z] \in J$. Let $x \in L$. Now $t(x, [y, z]) = t([x, y], z) = 0$ because $z \in J$. It follows that J is an ideal. Since L is simple, $J = 0$ or $J = L$. If $J = L$, then the restriction of t to $L \times L$ is zero, a contradiction. Hence, $J = 0$. □

Lemma 10.5.2. *Let L be a Lie algebra over F, and let (π, V) be a representation of L. Let*

$$V^\vee = \mathrm{Hom}_F(V, F),$$

and regard $V^\vee$ as a vector space over F. Define an action $\pi^\vee$ of L on $V^\vee$ by setting

$$(\pi^\vee(x)\lambda)(v) = -\lambda(\pi(x)v)$$

for $x \in L$, $\lambda \in V^\vee$, and $v \in V$. With this definition, $V^\vee$ is a well-defined representation of L.

Proof. We need to prove that the map $\pi^\vee : L \to \mathrm{gl}(V^\vee)$ is a well-defined Lie algebra homomorphism. This map is clearly well-defined and linear. Let $x, y \in L$, $\lambda \in V^\vee$, and $v \in V$. Then

$$\begin{aligned}(\pi^\vee([x,y])\lambda)(v) &= -\lambda(\pi([x,y])v) \\ &= -\lambda(\pi(x)\pi(y)v - \pi(y)\pi(x)v).\end{aligned}$$

And

$$\begin{aligned}&\big(\pi^\vee(x)\pi^\vee(y) - \pi^\vee(y)\pi^\vee(x)\big)\lambda \\ &= \pi^\vee(x)\big(\pi^\vee(y)\lambda\big) - \pi^\vee(y)\big(\pi^\vee(x)\lambda\big),\end{aligned}$$

so that

$$\begin{aligned}&\Big((\pi^\vee(x)\pi^\vee(y) - \pi^\vee(y)\pi^\vee(x))\lambda\Big)(v) \\ &= -\big(\pi^\vee(y)\lambda\big)(\pi(x)v) + \big(\pi^\vee(x)\lambda\big)(\pi(y)v) \\ &= \lambda\big(\pi(y)\pi(x)v\big) - \lambda\big(\pi(x)\pi(y)v\big).\end{aligned}$$

It follows that

$$\pi^\vee([x,y])\lambda = (\pi^\vee(x)\pi^\vee(y) - \pi^\vee(y)\pi^\vee(x))\lambda,$$

proving that $\pi^\vee$ is a Lie algebra homomorphism. □

Lemma 10.5.3. *Let F have characteristic zero and be algebraically closed. Let L be a finite-dimensional simple Lie algebra over F. If $t_1, t_2 : L \times L \to F$ are non-zero, associative, symmetric bilinear forms, then there exists $c \in F^\times$ such that $t_2 = ct_1$.*

Proof. Regard L as a representation π of L via the usual definition $\mathrm{ad}(x)y = [x, y]$ for $x, y \in L$ (see Proposition 1.5.1). Via Lemma 10.5.2 regard $L^\vee$ as a representation of L. For $v \in L$, define $r_1(v) \in L^\vee$ by $(r_1(v))(w) = t_1(v, w)$. We claim that $r_1 : L \to L^\vee$ is a well-defined homomorphism of representations of L. Let $x \in L$ and $v, w \in L$. Then

$$\begin{aligned}r_1(\mathrm{ad}(x)v)(w) &= t_1(\mathrm{ad}(x)v, w) \\ &= t_1([x,v], w) \\ &= t_1(-[v,x], w) \\ &= t_1(v, -[x,w]) \\ &= t_1(v, -\mathrm{ad}(x)w) \\ &= r_1(v)(-\mathrm{ad}(x)w) \\ &= \big(\mathrm{ad}^\vee(x)(r_1(v))\big)(w).\end{aligned}$$

This proves that r_1 is a well-defined homomorphism. Since t_1 is non-zero, r_1 is non-zero. The kernel of r_1 is an L-subspace of L and hence is an ideal of L; since r_1 is non-zero and L is simple, the kernel of r_1 is zero. Since L and $L^\vee$ have the same dimension, r_1 is an isomorphism of representations of L. Similarly, using t_2 we may define another isomorphism $r_2 : LtoL^\vee$ of representations of L. Consider $r_1^{-1} \circ r_2 : L \to L$. This is also an isomorphism of representations of L. By Schur's Lemma, Theorem 4.2.2, there exists $c \in F$ such that $r_1^{-1} \circ r_2 = c\mathrm{id}_L$, or equivalently, $r_2 = cr_1$. Let $v, w \in L$. Then

$$\begin{aligned}(r_2(v))(w) &= c(r_1(v))(w) \\ t_2(v,w) &= ct_1(v,w).\end{aligned}$$

This completes the proof. □

Lemma 10.5.4. *Let F have characteristic zero and be algebraically closed. Let n be a positive integer. Let L be a simple Lie subalgebra of $\mathrm{gl}(n,F)$, and let κ be the Killing form of L. There exists $c \in F^\times$ such that $\kappa = ct$, where $t : L \times L \to F$ is defined by $t(x,y) = \mathrm{tr}(xy)$ for $x, y \in L$*

Proof. This follows from Lemma 10.5.1 and Lemma 10.5.3. □

10.6 Some useful facts

Let n be a positive integer. Let $i, j \in \{1, \dots, n\}$. We let e_{ij} be the element of $\mathrm{gl}(n,F)$ that has 1 as the (i,j)-th entry and zeros elsewhere. Let $i, j, k, \ell \in \{1, \dots, n\}$ and $a \in \mathrm{gl}(n,F)$. Then

$$\begin{aligned}&[e_{ij}, e_{k\ell}] = \delta_{jk} e_{i\ell} - \delta_{\ell i} e_{kj}, \\ &[e_{ij}, e_{ji}] = e_{ii} - e_{jj}, \\ &i \neq \ell \Longrightarrow [e_{ik}, e_{k\ell}] = e_{i\ell}, \\ &j \neq k \Longrightarrow [e_{\ell j}, e_{k\ell}] = -e_{kj}, \\ &i \neq j \Longrightarrow [e_{ij}, [e_{ij}, a]] = -2a_{ji} e_{ij}.\end{aligned}$$

10.7 The Lie algebra $\mathrm{sl}(\ell+1)$

Lemma 10.7.1. *The dimension of the Lie algebra $\mathrm{sl}(\ell+1,F)$ is $(\ell+1)^2 - 1$.*

Proof. A basis for the Lie algebra $\mathrm{sl}(\ell+1,F)$ consists of the elements e_{ij} for $i, j \in \{1, \dots, \ell+1\}$, $i \neq j$, and the elements $e_{ii} - e_{\ell+1,\ell+1}$ for $i \in \{1, \dots, n-1\}$. □

Lemma 10.7.2. *Let F have characteristic zero and be algebraically closed. The natural action of $\mathrm{sl}(\ell+1,F)$ on $V = \mathrm{M}_{\ell+1,1}(F)$ is irreducible, so that $\mathrm{sl}(\ell+1,F)$ is semi-simple.*

Proof. Let $e_1, \dots, e_{\ell+1}$ be the standard basis for V. Let W be a non-zero $\mathrm{sl}(\ell+1,F)$-submodule of V; we need to prove that $W = V$. Let $w \in W$ be non-zero. Write

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_{\ell+1} \end{bmatrix}.$$

Since w is non-zero, there exists $j \in \{1, \dots, \ell+1\}$ such that $w_j \neq 0$. Applying the elements $e_{ij} \in \mathrm{sl}(\ell+1, F)$ for $i \in \{1, \dots, \ell+1\}$, $i \neq j$, to w, we find that the standard basis vectors e_i of V for $i \in \{1, \dots, \ell+1\}$, $i \neq j$ are contained in W. Let $k \in \{1, \dots, \ell+1\}$ with $k \neq j$. Applying the element $e_{jj} - e_{kk}$ to w, we get that $w_j e_j - w_k e_k$ is in W; this implies that e_j is in W. Since W contains the standard basis for V we have $W = V$, as desired. By Lemma 10.2.1, the Lie algebra $\mathrm{sl}(\ell+1, F)$ is semi-simple. □

Lemma 10.7.3. *Let F have characteristic zero and be algebraically closed. The set H of diagonal matrices in $\mathrm{sl}(\ell+1, F)$ is a Cartan subalgebra of $\mathrm{sl}(\ell+1, F)$.*

Proof. Let W be the F subspace of $\mathrm{sl}(\ell+1, F)$ consisting of matrices with zeros on the main diagonal. Let $w \in W$, and assume that w commutes with every element of H. By Lemma 10.4.2, to prove that H is a Cartan subalgebra, it suffices to prove that $w = 0$. Write

$$w = \sum_{\substack{1 \leq i,j \leq \ell+1, \\ i \neq j}} w_{ij} e_{ij}$$

for some $w_{ij} \in F$, $1 \leq i, j \leq \ell+1$, $i \neq j$. Let $h \in H$, with

$$h = \begin{bmatrix} h_{11} & & \\ & \ddots & \\ & & h_{\ell+1,\ell+1} \end{bmatrix}$$

for some $h_{11}, \dots, h_{\ell+1,\ell+1} \in F$. Then

$$\begin{aligned} [h, w] &= \sum_{\substack{1 \leq i,j \leq \ell+1, \\ i \neq j}} w_{ij} [h, e_{ij}] \\ 0 &= \sum_{\substack{1 \leq i,j \leq n, \\ i \neq j}} w_{ij} (h_{ii} - h_{jj}) e_{ij}. \end{aligned}$$

Since the e_{ij} for $i, j \in \{1, \dots, \ell+1\}$ are linearly independent, we get $w_{ij}(h_{ii} - h_{jj}) = 0$ for all $i, j \in \{1, \dots, \ell+1\}$ with $i \neq j$ and all $h \in H$. Let $i, j \in \{1, \dots, \ell+1\}$ with $i \neq j$. Set $h = e_{ii} - e_{jj}$. Then $h \in H$, and we have $w_{ij}(h_{ii} - h_{jj}) = 2w_{ij}$. Since F has characteristic zero, we conclude that $w_{ij} = 0$. Thus, $w = 0$. □

Lemma 10.7.4. *Assume that the characteristic of F is zero and F is algebraically closed. Let H be the Cartan subalgebra of $L = \mathrm{sl}(\ell+1, F)$ consisting of diagonal matrices in $\mathrm{sl}(\ell+1, F)$, as in Lemma 10.7.3. Then Φ consists of the linear forms*

$$\alpha_{ij} : H \longrightarrow F$$

defined by

$$\alpha_{ij}(h) = h_{ii} - h_{jj}$$

for $h \in H$; here, $1 \le i, j \le \ell + 1$ and $i \neq j$. Moreover

$$L_{\alpha_{ij}} = F e_{ij}$$

Proof. Let $1 \le i, j \le \ell + 1$ with $i \neq j$. For $h \in H$ we have

$$[h, e_{ij}] = (h_{ii} - h_{jj}) e_{ij} = \alpha_{ij}(h) e_{ij}.$$

It follows that $\alpha_{ij} \in \Phi$ and $e_{ij} \in L_{\alpha_{ij}}$. Since

$$\mathrm{sl}(\ell+1, F) = H \oplus \sum_{\substack{1 \le i,j \le \ell+1, \\ i \neq j}} F e_{ij} \subset H \oplus \sum_{\substack{1 \le i,j \le \ell+1, \\ i \neq j}} L_{\alpha_{ij}} \subset \mathrm{sl}(\ell+1, F)$$

the inclusion must be an equality. This implies that Φ and $L_{\alpha_{ij}}$ for $1 \le i, j \le \ell + 1$ with $i \neq j$ are as claimed. □

Lemma 10.7.5. *Let F have characteristic zero and be algebraically closed. Let ℓ be a positive integer. Let H be the subalgebra of $\mathrm{sl}(\ell+1, F)$ consisting of diagonal matrices; by Lemma 10.7.3, H is a Cartan subalgebra of $\mathrm{sl}(\ell+1, F)$. Let Φ be the set of roots of $\mathrm{sl}(\ell+1, F)$ defined with respect to H. Let $V = \mathbb{R} \otimes_{\mathbb{Q}} V_0$, where V_0 is the $\mathbb{Q}$ subspace of $H^\vee = \mathrm{Hom}_F(H, F)$ spanned by the elements of Φ; by Proposition 8.2.1, Φ is a root system in V. Let $i \in \{1, \ldots, \ell\}$, and define*

$$\beta_i : H \longrightarrow F$$

by

$$\beta_i(h) = h_{ii} - h_{i+1,i+1}$$

for $h \in H$. The set $B = \{\beta_1, \ldots, \beta_\ell\}$ is a base for Φ. The positive roots in Φ are the α_{ij} with $i < j$, and if $i < j$, then

$$\alpha_{ij} = \beta_i + \beta_{i+1} + \cdots + \beta_{j-1}.$$

Proof. It was proven in Lemma 10.7.4 that the linear functionals $\alpha_{ij} : H \to F$ defined by $\alpha_{ij}(h) = h_{ii} - h_{jj}$ for $h \in H$ and $i, j \in \{1, \ldots, \ell+1\}$, $i \neq j$, constitute the set of roots Φ of $\mathrm{sl}(\ell+1, \mathbb{C})$ with respect to H. Evidently, $B \subset \Phi$. Also, it is clear that B is linearly independent; since B has ℓ elements and the dimension of V is ℓ (by Proposition 7.1.2), it follows that B is a basis for V. Let $i, j \in \{1, \ldots, \ell+1\}$, $i \neq j$. Assume that $i < j$. Then

$$\alpha_{ij} = \beta_i + \beta_{i+1} + \cdots + \beta_{j-1}.$$

Assume that $j < i$. Then

$$\alpha_{ij} = -\alpha_{ji} = -(\beta_j + \beta_{j+1} + \cdots + \beta_{i-1}).$$

It follows that B is a base for Φ and the positive roots in Φ are as described. □

$$\begin{bmatrix} h_{11} & \boxed{\beta_1 = \alpha_{12}} & \boxed{\alpha_{13}} & \boxed{\alpha_{14}} & \boxed{\alpha_{15}} \\ -\alpha_{12} & h_{22} & \boxed{\beta_2 = \alpha_{23}} & \boxed{\alpha_{24}} & \boxed{\alpha_{25}} \\ -\alpha_{13} & -\alpha_{23} & h_{33} & \boxed{\beta_3 = \alpha_{34}} & \boxed{\alpha_{35}} \\ -\alpha_{14} & -\alpha_{24} & -\alpha_{34} & h_{44} & \boxed{\beta_4 = \alpha_{45}} \\ -\alpha_{15} & -\alpha_{25} & -\alpha_{35} & -\alpha_{45} & h_{55} \end{bmatrix}$$

Figure 10.1: The root spaces in $\mathrm{sl}(5, F)$. For this example, $\ell = 3$. The positions are labeled with the corresponding root. Note that the diagonal is our chosen Cartan subalgebra. The positive roots with respect to our chosen base $\{\beta_1, \beta_2, \beta_3, \beta_4\}$ are boxed, while the colored roots form our chosen base. The linear functionals α_{ij} are defined in Proposition 10.7.4.

Lemma 10.7.6. *Assume that the characteristic of F is zero and F is algebraically closed. Let ℓ be a positive integer. The Killing form*

$$\kappa : \mathrm{sl}(\ell+1, F) \times \mathrm{sl}(\ell+1, F) \longrightarrow F$$

is given by

$$\kappa(h, h') = (2\ell + 2) \cdot \mathrm{tr}(hh')$$

for $h, h' \in H$. Here, H is the subalgebra of diagonal matrices in $\mathrm{sl}(\ell+1, F)$; H is a Cartan subalgebra of $\mathrm{sl}(\ell+1, F)$ by Lemma 10.7.3.

Proof. Let $h, h' \in H$. Then:

$$\begin{aligned}
&\kappa(h, h') \\
&= \mathrm{tr}(\mathrm{ad}(h) \circ \mathrm{ad}(h')) \\
&= \sum_{\alpha \in \Phi} \alpha(h)\alpha(h') \\
&= \sum_{\substack{i,j \in \{1,\dots,\ell+1\},\\ i \neq j}} (h_{ii} - h_{jj})(h'_{ii} - h'_{jj}) \\
&= \sum_{\substack{i,j \in \{1,\dots,\ell+1\},\\ i \neq j}} h_{ii}h'_{ii} - \sum_{\substack{i,j \in \{1,\dots,\ell+1\},\\ i \neq j}} h_{ii}h'_{jj} \\
&\qquad - \sum_{\substack{i,j \in \{1,\dots,\ell+1\},\\ i \neq j}} h_{jj}h'_{ii} + \sum_{\substack{i,j \in \{1,\dots,\ell+1\},\\ i \neq j}} h_{jj}h'_{jj} \\
&= 2\ell \sum_{i \in \{1,\dots,\ell+1\}} h_{ii}h'_{ii} - 2 \sum_{\substack{i,j \in \{1,\dots,\ell+1\},\\ i \neq j}} h_{ii}h'_{jj}
\end{aligned}$$

$$\begin{aligned}
&= 2\ell \cdot \mathrm{tr}(hh') - 2 \sum_{i,j\in\{1,\dots,\ell+1\}} h_{ii}h'_{jj} + 2 \sum_{i\in\{1,\dots,\ell+1\}} h_{ii}h'_{ii} \\
&= 2\ell \cdot \mathrm{tr}(hh') - 2\cdot \mathrm{tr}(h)\cdot \mathrm{tr}(h') + 2\cdot \mathrm{tr}(hh') \\
&= (2\ell+2)\cdot \mathrm{tr}(hh') - 2\cdot 0\cdot 0 \\
&= (2\ell+2)\cdot \mathrm{tr}(hh'),
\end{aligned}$$

where we note that $\mathrm{tr}(h) = \mathrm{tr}(h') = 0$ because $h, h' \in \mathrm{sl}(\ell+1, F)$. □

Lemma 10.7.7. *Let the notation as in Lemma 10.7.5. If* $i \in \{1,\dots,\ell\}$*, then*

$$t_{\beta_i} = \frac{1}{2\ell+2}(e_{ii} - e_{i+1,i+1}).$$

Let $i, j \in \{1,\dots,\ell\}$*. Then*

$$(\beta_i,\beta_j) = \begin{cases} \dfrac{2}{2\ell+2} & \text{if } i=j, \\ \dfrac{-1}{2\ell+2} & \text{if } i \text{ and } j \text{ are consecutive}, \\ 0 & \text{if } i\neq j \text{ and } i \text{ and } j \text{ are not consecutive}. \end{cases}$$

Proof. Let $i \in \{1,\dots,\ell\}$, and let $h \in H$. Then

$$\beta_i(h) = h_{ii} - h_{i+1,i+1}.$$

Also,

$$\begin{aligned}
\kappa\big(h, \frac{1}{2\ell+2}(e_{ii} - e_{i+1,i+1})\big) &= \frac{1}{2\ell+2}\kappa(h,e_{ii}) - \frac{1}{2\ell+2}\kappa(h,e_{i+1,i+1}) \\
&= \frac{2\ell+2}{2\ell+2}\cdot \mathrm{tr}(he_{ii}) - \frac{2\ell+2}{2\ell+2}\cdot \mathrm{tr}(he_{i+1,i+1}) \\
&= \mathrm{tr}(he_{ii}) - \mathrm{tr}(he_{i+1,i+1}) \\
&= h_{ii} - h_{i+1,i+1}.
\end{aligned}$$

By definition, t_{β_i} is the unique element of H such that $\beta_i(h) = \kappa(h, t_{\beta_i})$ for all $h \in H$. The last two equalities imply that

$$t_{\beta_i} = \frac{1}{2\ell+2}(e_{ii} - e_{i+1,i+1}).$$

Let $i, j \in \{1,\dots,\ell\}$. By the definition of the inner product on V and Lemma 10.7.6 we have

$$\begin{aligned}
(\beta_i,\beta_j) &= \kappa(t_{\beta_i}, t_{\beta_j}) \\
&= (2\ell+2)\mathrm{tr}(t_{\beta_i}t_{\beta_j}) \\
&= \frac{1}{2\ell+2}\mathrm{tr}\big((e_{ii}-e_{i+1,i+1})(e_{jj}-e_{j+1,j+1})\big)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2\ell+2}\Big(\mathrm{tr}(e_{ii}e_{jj}) - \mathrm{tr}(e_{ii}e_{j+1,j+1}) \\
&\qquad - \mathrm{tr}(e_{i+1,i+1}e_{jj}) + \mathrm{tr}(e_{i+1,i+1}e_{j+1,j+1})\Big) \\
&= \frac{1}{2\ell+2}\Big(\delta_{ij} - \delta_{i,j+1} - \delta_{i+1,j} + \delta_{i+1,j+1}\Big).
\end{aligned}
$$

The formula for (β_i, β_j) follows. □

Lemma 10.7.8. *Let F have characteristic zero and be algebraically closed. The Dynkin diagram of* $\mathrm{sl}(\ell+1, F)$ *is*

A_ℓ: 

and the Cartan matrix of $\mathrm{sl}(\ell+1, F)$ *is*

$$
\begin{bmatrix}
2 & -1 & & & & \\
-1 & 2 & -1 & & & \\
 & -1 & 2 & -1 & & \\
 & & \ddots & \ddots & \ddots & \\
 & & & -1 & 2 & -1 \\
 & & & & -1 & 2
\end{bmatrix}
$$

The Lie algebra $\mathrm{sl}(\ell+1, F)$ *is simple.*

Proof. Let $i, j \in \{1, \ldots, \ell\}$ with $i \neq j$. We have by Lemma 10.7.7,

$$
\begin{aligned}
\langle \beta_i, \beta_j \rangle &= 2\frac{(\beta_i, \beta_j)}{(\beta_j, \beta_j)} \\
&= \begin{cases} -1 & \text{if } i \text{ and } j \text{ are consecutive,} \\ 0 & \text{if } i \text{ and } j \text{ are not consecutive.} \end{cases}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\langle \beta_i, \beta_j \rangle \langle \beta_j, \beta_i \rangle &= 4\frac{(\beta_i, \beta_j)^2}{(\beta_i, \beta_i)(\beta_j, \beta_j)} \\
&= \begin{cases} 1 & \text{if } i \text{ and } j \text{ are consecutive,} \\ 0 & \text{if } i \text{ and } j \text{ are not consecutive.} \end{cases}
\end{aligned}
$$

It follows that the Dynkin diagram of $\mathrm{sl}(\ell+1, F)$ is A_ℓ, and the Cartan matrix of $\mathrm{sl}(\ell+1, F)$ is as stated. Since A_ℓ is connected, $\mathrm{sl}(\ell+1, F)$ is simple by Lemma 9.3.2 and Proposition 10.3.2. □

Lemma 10.7.9. *Assume that the characteristic of F is zero and F is algebraically closed. Let ℓ be a positive integer. The Killing form*

$$\kappa : \mathrm{sl}(\ell+1) \times \mathrm{sl}(\ell+1) \longrightarrow F$$

is given by

$$\kappa(x, y) = (2\ell+2) \cdot \mathrm{tr}(xy).$$

for $x, y \in \mathrm{sl}(\ell+1, F)$.

Proof. By Lemma 10.5.3, there exists $c \in F^\times$ such that $\kappa(x,y) = c\,\mathrm{tr}(xy)$ for $x,y \in \mathrm{sl}(\ell+1,F)$. Let H be the subalgebra of diagonal matrices in $\mathrm{sl}(\ell+1,F)$; H is a Cartan subalgebra of $\mathrm{sl}(\ell+1,F)$ by Lemma 10.7.3. By Lemma 10.7.6 we have $\kappa(h,h') = (2\ell+2)\cdot\mathrm{tr}(hh')$ for $h,h' \in H$. Hence, $c\,\mathrm{tr}(hh') = (2\ell+2)\cdot\mathrm{tr}(hh')$ for $h,h' \in H$. Since there exist $h,h' \in H$ such that $\mathrm{tr}(hh') \neq 0$ we conclude that $c = 2\ell+2$. □

Lemma 10.7.10. *Let the notation as in Lemma 10.7.4 and Lemma 10.7.5. Let* $i,j \in \{1,\dots,\ell+1\}$ *with* $i \neq j$. *The length of every root is* $\frac{1}{\sqrt{\ell+1}}$.

Proof. Let $\alpha \in \Phi^+$. We know that $\alpha_1,\dots,\alpha_\ell$ is an ordered basis for V. By Lemma 10.7.7 the matrix of the inner product $(\cdot,\cdot)$ in this basis is

$$M = \frac{1}{2\ell+2}\begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}.$$

The coordinate vector of α in this basis has the form

$$c = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

A calculation shows that $(\alpha,\alpha) = {}^t cMc = \frac{2}{2\ell+2} = \frac{1}{\ell+1}$; hence the length of α is $\frac{1}{\sqrt{\ell+1}}$. □

10.8 The Lie algebra so$(2\ell+1)$

Lemma 10.8.1. *The Lie algebra* $\mathrm{so}(2\ell+1,F)$ *consists of the* $x \in \mathrm{gl}(2\ell+1,F)$ *of the form*

$$x = \begin{bmatrix} \overbrace{0}^{1} & \overbrace{b}^{\ell} & \overbrace{c}^{\ell} \\ -{}^t c & f & g \\ -{}^t b & G & -{}^t f \end{bmatrix} \begin{matrix} \}\,1 \\ \}\,\ell \\ \}\,\ell \end{matrix}$$

where $g = -{}^t g$ *and* $G = -{}^t G$. *The dimension of* $\mathrm{so}(2\ell+1,F)$ *is* $2\ell^2+\ell$.

Proof. Let $x \in \mathrm{gl}(2\ell+1, F)$, and write

$$x = \begin{bmatrix} \overbrace{a}^{1} & \overbrace{b}^{\ell} & \overbrace{c}^{\ell} \\ B & f & g \\ C & G & h \end{bmatrix} \begin{matrix} \}\, 1 \\ \}\, \ell \\ \}\, \ell \end{matrix}$$

where $a \in F$, $f \in \mathrm{gl}(\ell, F)$, $h \in \mathrm{gl}(\ell, F)$, and b, c, g, B, C and G are appropriately sized matrices with entries from F. By definition, $s \in \mathrm{so}(2\ell+1, F)$ if and only if ${}^t xS = -Sx$. We have

$${}^t xS = \begin{bmatrix} {}^t a & {}^t B & {}^t C \\ {}^t b & {}^t f & {}^t G \\ {}^t c & {}^t g & {}^t h \end{bmatrix} \begin{bmatrix} 1 & & \\ & & 1_\ell \\ & 1_\ell & \end{bmatrix} = \begin{bmatrix} a & {}^t C & {}^t B \\ {}^t b & {}^t G & {}^t f \\ {}^t c & {}^t h & {}^t g \end{bmatrix}.$$

And:

$$-Sx = -\begin{bmatrix} 1 & & \\ & & 1_\ell \\ & 1_\ell & \end{bmatrix} \begin{bmatrix} a & b & c \\ B & f & g \\ C & G & h \end{bmatrix} = \begin{bmatrix} -a & -b & -c \\ -C & -G & -h \\ -B & -f & -g \end{bmatrix}.$$

It follows that $x \in \mathrm{so}(2\ell+1, F)$ if and only if:

$$\begin{aligned} a &= 0, \\ B &= -{}^t c, \\ C &= -{}^t b, \\ G &= -{}^t G, \\ h &= -{}^t f, \\ g &= -{}^t g. \end{aligned}$$

This completes the proof. □

Lemma 10.8.2. *Assume that the characteristic of* F *is not two. The Lie algebras* $\mathrm{so}(3, F)$ *and* $\mathrm{sl}(2, F)$ *are isomorphic.*

Proof. Recalling the structure of $\mathrm{sl}(2, F)$, it suffices to prove that $\mathrm{so}(3, F)$ has a vector space basis e, f, h such that $[e, f] = h$, $[h, e] = 2e$ and $[h, f] = -2f$. Define the following elements of $\mathrm{so}(3, F)$:

$$e = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad f = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}, \qquad h = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

Evidently, e, f and h form a vector space basis for $\mathrm{so}(3, F)$, and calculations prove that $[e, f] = h$, $[h, e] = 2e$ and $[h, f] = -2f$. □

Lemma 10.8.3. *Let ℓ be an integer with $\ell \geq 2$. Let $x \in \mathrm{M}_{\ell,1}(F)$ be non-zero. There exists $w \in \mathrm{gl}(\ell, F)$ such that $-{}^{\mathrm{t}}w = w$ and $wx \neq 0$.*

Proof. Since $x \neq 0$ there exists $j \in \{1, \ldots, \ell\}$ such that $x_j \neq 0$. Since $\ell \geq 2$, there exists $i \in \{1, \ldots, \ell\}$ such that $i \neq j$. Set $w = e_{ij} - {}^{\mathrm{t}}e_{ij} = e_{ij} - e_{ji}$. Then $wx = x_j e_i - x_i e_j \neq 0$. □

Lemma 10.8.4. *Let ℓ be an integer with $\ell \geq 2$. Assume that the characteristic of F is zero and F is algebraically closed. The natural representation of $\mathrm{so}(2\ell+1, F)$ on $\mathrm{M}_{2\ell+1,1}(F)$ given by multiplication of matrices is irreducible. The Lie algebra $\mathrm{so}(2\ell+1, F)$ is semi-simple.*

Proof. Assume that V is a non-zero $\mathrm{so}(2\ell+1, F)$ subspace of $\mathrm{M}_{2\ell+1,1}(F)$; we need to prove that $V = \mathrm{M}_{2\ell+1,1}(F)$. We will write the elements of $\mathrm{M}_{2\ell+1,1}(F)$ in the form

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \begin{matrix} \}\,1 \\ \}\,\ell \\ \}\,\ell \end{matrix}.$$

We first claim that V contains an element of the form

$$\begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}$$

with $y \neq 0$. To see this, let

$$v = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

be a non-zero element of V. Assume first that $y = z = 0$, so that $x \neq 0$. Let $c \in F^\ell$ be such that ${}^t c x \in \mathrm{M}_{\ell,1}(F)$ is non-zero. Since

$$\begin{bmatrix} 0 & 0 & c \\ -{}^t c & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} v = \begin{bmatrix} 0 & 0 & c \\ -{}^t c & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -{}^t c x \\ 0 \end{bmatrix} \in V,$$

our claim holds in this case. We may thus assume that $y \neq 0$ or $z \neq 0$. Assume that $z \neq 0$. Let $g \in \mathrm{M}_{\ell,\ell}(F)$ be such that $-{}^t g = g$ and $gz \neq 0$; such a g exists by Lemma 10.8.3. Since

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & g \\ 0 & 0 & 0 \end{bmatrix} v = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & g \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ gz \\ 0 \end{bmatrix} \in V,$$

our claim holds in this case. We may now assume that $z = 0$ and $y \neq 0$ so that v has the form

$$v = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}.$$

Let $f \in \mathrm{gl}(\ell, F)$ be such that $fy \neq 0$. Then

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & -{}^t f \end{bmatrix} v = \begin{bmatrix} 0 & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & -{}^t f \end{bmatrix} \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ fy \\ 0 \end{bmatrix} \in V,$$

proving our claim in this final case. Thus, our claim holds; that is, V contains a vector

$$w = \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}$$

with $y \neq 0$. If $f \in \mathrm{gl}(\ell, F)$, then

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & -{}^t f \end{bmatrix} w = \begin{bmatrix} 0 & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & -{}^t f \end{bmatrix} \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ fy \\ 0 \end{bmatrix} \in V.$$

Since the action of $\mathrm{gl}(\ell, F)$ on $\mathrm{M}_{\ell,1}(F)$ is irreducible, it follows that V contains the subspace

$$\begin{bmatrix} 0 \\ \mathrm{M}_{\ell,1}(F) \\ 0 \end{bmatrix}.$$

Let $G \in \mathrm{M}_{\ell,\ell}(F)$ be such that $-{}^t G = G$ and $Gy \neq 0$; such a G exists by Lemma 10.8.3. We have

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & G & 0 \end{bmatrix} w = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & G & 0 \end{bmatrix} \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ Gy \end{bmatrix} \in V.$$

Acting on this vector by elements of $\mathrm{so}(2\ell + 1, F)$ by elements of the form

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & -{}^t f \end{bmatrix}$$

for $f \in \mathrm{gl}(\ell, F)$ we deduce that V contains the subspace

$$\begin{bmatrix} 0 \\ 0 \\ \mathrm{M}_{\ell,1}(F) \end{bmatrix}.$$

Finally, let $b \in \mathrm{M}_{\ell,1}(F)$ and $y \in \mathrm{M}_{1,\ell}(F)$ be such that $by \neq 0$. Then

$$\begin{bmatrix} 0 & b & 0 \\ 0 & 0 & 0 \\ -{}^t b & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} by \\ 0 \\ 0 \end{bmatrix} \in V.$$

It follows that V also contains the one-dimensional space

$$\begin{bmatrix} F \\ 0 \\ 0 \end{bmatrix}.$$

We conclude that $V = \mathrm{M}_{2\ell+1}(F)$, as desired.

Finally, $\mathrm{so}(2\ell+1, F)$ is semi-simple by Lemma 10.2.1 (note that $\mathrm{so}(2\ell+1, F)$ is contained in $\mathrm{sl}(2\ell+1, F)$ by Lemma 10.1.1). □

Lemma 10.8.5. *Let F be a field, and let n be a positive integer. Let $a \in \mathrm{gl}(n, F)$. If $ah = ha$ for all diagonal matrices $h \in \mathrm{gl}(n, F)$, then a is a diagonal matrix. If F does not have characteristic two, and $ah = -ha$ for all diagonal matrices $h \in \mathrm{gl}(n, F)$, then $a = 0$.*

Proof. Assume that $ah = ha$ for all diagonal matrices $h \in \mathrm{gl}(n, F)$. Let $h \in \mathrm{gl}(n, F)$ be a diagonal matrix. Then for all $i, j \in \{1, \dots, n\}$ we have $a_{ij}h_{jj} = h_{ii}a_{ij}$, i.e., $(h_{ii} - h_{jj})a_{ij} = 0$. It follows that $a_{ij} = 0$ for $i, j \in \{1, \dots, n\}$ with $i \neq j$; that is, a is a diagonal matrix.

Assume that F does not have characteristic two. Assume that $ah = -ha$ for all diagonal matrices $h \in \mathrm{gl}(n, F)$. Let $h \in \mathrm{gl}(n, F)$ be a diagonal matrix. Then for all $i, j \in \{1, \dots, n\}$ we have $a_{ij}h_{jj} = -h_{ii}a_{ij}$, i.e., $(h_{ii} + h_{jj})a_{ij} = 0$. This implies that $a = 0$; note that this uses that F does not have characteristic two. □

Lemma 10.8.6. *Let F have characteristic zero and be algebraically closed. The set H of diagonal matrices in* $\mathrm{so}(2\ell+1, F)$ *is a Cartan subalgebra of* $\mathrm{so}(2\ell+1, F)$.

Proof. By Lemma 10.4.2, to prove that H is a Cartan subalgebra of $\mathrm{so}(2\ell+1, F)$, it suffices prove that if $w \in \mathrm{so}(2\ell + 1, F)$ has zero entries on the main diagonal and $wh = hw$ for $h \in H$, then $w = 0$. Let w be such an element of $\mathrm{so}(2\ell+1, F)$, and write, as usual,

$$w = \begin{bmatrix} 0 & b & c \\ -{}^t c & f & g \\ -{}^t b & G & -{}^t f \end{bmatrix}.$$

Let $h \in H$, so that h has the form

$$h = \begin{bmatrix} 0 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & -{}^t d \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & -d \end{bmatrix}$$

with $d \in \mathrm{gl}(\ell, F)$ diagonal. We have

$$wh = \begin{bmatrix} 0 & b & c \\ -{}^t c & f & g \\ -{}^t b & G & -{}^t f \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & -d \end{bmatrix} = \begin{bmatrix} 0 & bd & -cd \\ -{}^t c & fd & -gd \\ -{}^t b & Gd & {}^t fd \end{bmatrix}$$

and

$$hw = \begin{bmatrix} 0 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & -d \end{bmatrix} \begin{bmatrix} 0 & b & c \\ -{}^t c & f & g \\ -{}^t b & G & -{}^t f \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ -d^t c & df & dg \\ d^t b & -dG & d^t f \end{bmatrix}.$$

It follows that

$$\begin{aligned} bd &= 0, \\ cd &= 0, \\ fd &= df, \\ gd &= -dg, \\ Gd &= -dG. \end{aligned}$$

Since these equations hold for all diagonal matrices $d \in \mathrm{gl}(\ell, F)$, it follows that $b = 0$ and $c = 0$. Also, by Lemma 10.8.5, f is a diagonal matrix and $g = 0$ and $G = 0$. Since, by assumption, w has zero entries on the main diagonal, we see that $f = 0$. Thus, $w = 0$. □

Lemma 10.8.7. *Let ℓ be an integer with $\ell \geq 2$. Let F have characteristic zero and be algebraically closed. Let ℓ be a positive integer. Let H be the subalgebra of $\mathrm{so}(2\ell+1, F)$ consisting of diagonal matrices; by Lemma 10.8.6, H is a Cartan subalgebra of $\mathrm{so}(2\ell+1, F)$. Let Φ be the set of roots of $\mathrm{so}(2\ell+1, F)$ defined with respect to H. Let $V = \mathbb{R} \otimes_{\mathbb{Q}} V_0$, where V_0 is the $\mathbb{Q}$ subspace of $H^\vee = \mathrm{Hom}_F(H, F)$ spanned by the elements of Φ; by Proposition 8.2.1, Φ is a root system in V. For $j \in \{1, \dots, \ell\}$, define a linear functional*

$$\alpha_j : H \longrightarrow F$$

by

$$\alpha_j(\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}) = h_{jj}$$

for $h \in \mathrm{gl}(\ell, F)$ and h diagonal. The set Φ consists of the following $2\ell^2$ linear functionals on H:

$$\begin{gathered} \alpha_1, \dots, \alpha_n, \\ -\alpha_1, \dots, -\alpha_n, \\ \alpha_i - \alpha_j, \quad i, j \in \{1, \dots, \ell\}, \quad i \neq j, \\ \alpha_i + \alpha_j, \quad i, j \in \{1, \dots, \ell\}, \quad i < j, \\ -(\alpha_i + \alpha_j), \quad i, j \in \{1, \dots, \ell\}, \quad i < j. \end{gathered}$$

The set

$$B = \{\beta_1 = \alpha_1 - \alpha_2,\ \beta_2 = \alpha_2 - \alpha_3,\ \dots,\ \beta_{\ell-1} = \alpha_{\ell-1} - \alpha_\ell,\ \beta_\ell = \alpha_\ell\}$$

is a base for Φ, and the positive roots with respect to B are

$$\alpha_1, \dots, \alpha_n,$$

$$\alpha_i - \alpha_j, \quad i,j \in \{1,\dots,\ell\}, \quad i<j,$$
$$\alpha_i + \alpha_j, \quad i,j \in \{1,\dots,\ell\}, \quad i<j.$$

The root spaces are:

$$L_{\alpha_j} = F \cdot \begin{bmatrix} 0 & 0 & e_{1j} \\ -{}^t e_{1j} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad j \in \{1,\dots,\ell\},$$

$$L_{-\alpha_j} = F \cdot \begin{bmatrix} 0 & e_{1j} & 0 \\ 0 & 0 & 0 \\ -{}^t e_{1j} & 0 & 0 \end{bmatrix}, \quad j \in \{1,\dots,\ell\},$$

$$L_{\alpha_i-\alpha_j} = F \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & e_{ij} & 0 \\ 0 & 0 & -e_{ji} \end{bmatrix}, \quad i,j \in \{1,\dots,\ell\}, \quad i \neq j,$$

$$L_{\alpha_i+\alpha_j} = F \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & e_{ij}-e_{ji} \\ 0 & 0 & 0 \end{bmatrix}, \quad i,j \in \{1,\dots,\ell\}, \quad i<j,$$

$$L_{-(\alpha_i+\alpha_j)} = F \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & e_{ij}-e_{ji} & 0 \end{bmatrix}, \quad i,j \in \{1,\dots,\ell\}, \quad i<j.$$

Proof. Let $h \in \mathrm{gl}(\ell,F)$ be a diagonal matrix. We have

$$\begin{aligned}
&[\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & e_{1j} & 0 \\ 0 & 0 & 0 \\ -{}^t e_{1j} & 0 & 0 \end{bmatrix}] \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix} \begin{bmatrix} 0 & e_{1j} & 0 \\ 0 & 0 & 0 \\ -{}^t e_{1j} & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & e_{1j} & 0 \\ 0 & 0 & 0 \\ -{}^t e_{1j} & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ h^t e_{1j} & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & e_{1j}h & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&= -h_{jj} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -{}^t e_{1j} & 0 & 0 \end{bmatrix} - h_{jj} \begin{bmatrix} 0 & e_{1j} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&= (-h_{jj}) \cdot \begin{bmatrix} 0 & e_{1j} & 0 \\ 0 & 0 & 0 \\ -{}^t e_{1j} & 0 & 0 \end{bmatrix}.
\end{aligned}$$

That is,

$$[\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & e_{1j} & 0 \\ 0 & 0 & 0 \\ -{}^t e_{1j} & 0 & 0 \end{bmatrix}] = (-h_{jj}) \cdot \begin{bmatrix} 0 & e_{1j} & 0 \\ 0 & 0 & 0 \\ -{}^t e_{1j} & 0 & 0 \end{bmatrix}.$$

Taking transposes of this equation yields:

$$[\begin{bmatrix} 0 & 0 & -e_{1j} \\ {}^te_{1j} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}] = (-h_{jj}) \cdot \begin{bmatrix} 0 & 0 & -e_{1j} \\ {}^te_{1j} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$-[\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & 0 & -e_{1j} \\ {}^te_{1j} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}] = (-h_{jj}) \cdot \begin{bmatrix} 0 & 0 & -e_{1j} \\ {}^te_{1j} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$[\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & 0 & e_{1j} \\ -{}^te_{1j} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}] = h_{jj} \cdot \begin{bmatrix} 0 & 0 & e_{1j} \\ -{}^te_{1j} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

And

$$\begin{aligned}
&[\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & e_{ij}-e_{ji} \\ 0 & 0 & 0 \end{bmatrix}] \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & e_{ij}-e_{ji} \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & e_{ij}-e_{ji} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & h_{ii}e_{ij}-h_{jj}e_{ji} \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -h_{jj}e_{ij}+h_{ii}e_{ji} \\ 0 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & h_{ii}e_{ij}-h_{jj}e_{ji}+h_{jj}e_{ij}-h_{ii}e_{ji} \\ 0 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & (h_{ii}+h_{jj})e_{ij}-(h_{ii}+h_{jj})e_{ji} \\ 0 & 0 & 0 \end{bmatrix} \\
&= (h_{ii}+h_{jj}) \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & e_{ij}-e_{ji} \\ 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

Taking transposes, we obtain:

$$[\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & e_{ji}-e_{ij} & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}] = (h_{ii}+h_{jj}) \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & e_{ji}-e_{ij} & 0 \end{bmatrix}$$

$$-[\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & e_{ji}-e_{ij} & 0 \end{bmatrix}] = (h_{ii}+h_{jj}) \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & e_{ji}-e_{ij} & 0 \end{bmatrix}$$

$$[\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & e_{ij}-e_{ji} & 0 \end{bmatrix}] = -(h_{ii}+h_{jj}) \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & e_{ij}-e_{ji} & 0 \end{bmatrix}.$$

And

$$\begin{aligned}
&[\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix},\begin{bmatrix}0&0&0\\0&e_{ij}&0\\0&0&-e_{ji}\end{bmatrix}]\\
&=\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix}\begin{bmatrix}0&0&0\\0&e_{ij}&0\\0&0&-e_{ji}\end{bmatrix}-\begin{bmatrix}0&0&0\\0&e_{ij}&0\\0&0&-e_{ji}\end{bmatrix}\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix}\\
&=\begin{bmatrix}0&0&0\\0&h_{ii}e_{ij}&0\\0&0&h_{jj}e_{ji}\end{bmatrix}-\begin{bmatrix}0&0&0\\0&h_{jj}e_{ij}&0\\0&0&h_{ii}e_{ji}\end{bmatrix}\\
&=\begin{bmatrix}0&0&0\\0&(h_{ii}-h_{jj})e_{ij}&0\\0&0&(h_{ii}-h_{jj})(-e_{ji})\end{bmatrix}\\
&=(h_{ii}-h_{jj})\cdot\begin{bmatrix}0&0&0\\0&e_{ij}&0\\0&0&-e_{ji}\end{bmatrix}.
\end{aligned}$$

These calculations show that the linear functionals from the statement of the lemma are indeed roots, and that the root spaces of these roots are as stated (recall that any root space is one-dimensional by Proposition 7.0.8). Since the span of H and the stated root spaces is $\mathrm{so}(2\ell+1,F)$ it follows that these roots are all the roots of $\mathrm{so}(2\ell+1,F)$ with respect to H. It is straightforward to verify that B is a base for Φ, and that the positive roots of Φ with respect to B are as stated. Note that the dimension of V is ℓ (by Proposition 7.1.2). □

$$\left[\begin{array}{c|ccc|ccc}
0 & -\alpha_1 & -\alpha_2 & -\alpha_3 & \boxed{\alpha_1} & \boxed{\alpha_2} & \colorbox{cyan}{$\boxed{\beta_3=\alpha_3}$}\\
\hline
* & h_{11} & \colorbox{cyan}{$\boxed{\beta_1=\alpha_1-\alpha_2}$} & \boxed{\alpha_1-\alpha_3} & 0 & \boxed{\alpha_1+\alpha_2} & \boxed{\alpha_1+\alpha_3}\\
* & \alpha_2-\alpha_1 & h_{22} & \colorbox{cyan}{$\boxed{\beta_2=\alpha_2-\alpha_3}$} & * & 0 & \boxed{\alpha_2+\alpha_3}\\
* & \alpha_3-\alpha_1 & \alpha_3-\alpha_2 & h_{33} & * & * & 0\\
\hline
* & 0 & -(\alpha_1+\alpha_2) & -(\alpha_1+\alpha_3) & -h_{11} & * & *\\
* & * & 0 & -(\alpha_2+\alpha_3) & * & -h_{22} & *\\
* & * & * & 0 & * & * & -h_{33}
\end{array}\right]$$

Figure 10.2: The decomposition of $\mathrm{so}(7,F)=\mathrm{so}(2\cdot 3+1,F)$. For this example, $\ell=3$. The positions are labeled with the corresponding root. Note that the diagonal is our chosen Cartan subalgebra. The positive roots with respect to our chosen base $\{\beta_1,\beta_2,\beta_3\}$ are boxed, while the colored roots form our chosen base. Positions labeled with $*$ are determined by other entries. The linear functionals α_1,α_2 and α_3 are defined in Proposition 10.8.7.

Lemma 10.8.8. *Assume that the characteristic of F is zero and F is algebraically closed. Let ℓ be a positive integer. The Killing form*

$$\kappa : \mathrm{so}(2\ell+1, F) \times \mathrm{so}(2\ell+1, F) \longrightarrow F$$

is given by

$$\kappa(h, h') = (2\ell - 1) \cdot \mathrm{tr}(hh')$$

for $h, h' \in H$. Here, H is the subalgebra of diagonal matrices in $\mathrm{so}(2\ell+1, F)$; H is a Cartan subalgebra of $\mathrm{so}(2\ell+1, F)$ by Lemma 10.8.6.

Proof. Let $h, h' \in H$. Then

$$\begin{aligned}
&\kappa(\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & h' & 0 \\ 0 & 0 & -h' \end{bmatrix}) \\
&= \mathrm{tr}(\mathrm{ad}(\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix}) \circ \mathrm{ad}(\begin{bmatrix} 0 & 0 & 0 \\ 0 & h' & 0 \\ 0 & 0 & -h' \end{bmatrix})) \\
&= 2 \sum_{\alpha \in \Phi^+} \alpha(\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix})\alpha(\begin{bmatrix} 0 & 0 & 0 \\ 0 & h' & 0 \\ 0 & 0 & -h' \end{bmatrix}) \\
&= 2 \sum_{i \in \{1, \dots \ell\}} h_i h'_i \\
&\quad + 2 \sum_{\substack{i,j \in \{1,\dots,\ell\}, \\ i<j}} (h_i - h_j)(h'_i - h'_j) \\
&\qquad + 2 \sum_{\substack{i,j \in \{1,\dots,\ell\}, \\ i<j}} (h_i + h_j)(h'_i + h'_j) \\
&= \mathrm{tr}(\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & h' & 0 \\ 0 & 0 & -h' \end{bmatrix}) \\
&\quad + 2 \sum_{\substack{i,j \in \{1,\dots,\ell\}, \\ i<j}} (h_i h'_i - h_i h'_j - h_j h'_i + h_j h'_j + h_i h'_i + h_i h'_j + h_j h'_i + h_j h'_j) \\
&= \mathrm{tr}(\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & h' & 0 \\ 0 & 0 & -h' \end{bmatrix}) + 4 \sum_{\substack{i,j \in \{1,\dots,\ell\}, \\ i<j}} (h_i h'_i + h_j h'_j) \\
&= \mathrm{tr}(\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & h' & 0 \\ 0 & 0 & -h' \end{bmatrix}) + 4 \sum_{\substack{i,j \in \{1,\dots,\ell\}, \\ i<j}} h_i h'_i + 4 \sum_{\substack{i,j \in \{1,\dots,\ell\}, \\ i<j}} h_j h'_j \\
&= \mathrm{tr}(\begin{bmatrix} 0 & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & -h \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & h' & 0 \\ 0 & 0 & -h' \end{bmatrix})
\end{aligned}$$

$$+4\sum_{i\in\{1,\dots,\ell\}}(\ell-i)h_ih'_i+4\sum_{j\in\{1,\dots,\ell\}}(j-1)h_jh'_j$$

$$=\mathrm{tr}(\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix}\begin{bmatrix}0&0&0\\0&h'&0\\0&0&-h'\end{bmatrix})+4\sum_{i\in\{1,\dots,\ell\}}(\ell-i+i-1)h_ih'_i$$

$$=\mathrm{tr}(\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix}\begin{bmatrix}0&0&0\\0&h'&0\\0&0&-h'\end{bmatrix})+4(\ell-1)\sum_{i\in\{1,\dots,\ell\}}h_ih'_i$$

$$=\mathrm{tr}(\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix}\begin{bmatrix}0&0&0\\0&h'&0\\0&0&-h'\end{bmatrix})+2(\ell-1)\mathrm{tr}(\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix}\begin{bmatrix}0&0&0\\0&h'&0\\0&0&-h'\end{bmatrix})$$

$$=(2\ell-1)\mathrm{tr}(\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix}\begin{bmatrix}0&0&0\\0&h'&0\\0&0&-h'\end{bmatrix}).$$

This completes the proof. □

Lemma 10.8.9. *Let the notation as in Lemma 10.8.7. For $i,j\in\{1,\dots,\ell\}$,*

$$(\beta_i,\beta_j)=\begin{cases}\dfrac{2}{4\ell-2} & \text{if } i=j\in\{1,\dots,\ell-1\},\\[2mm] \dfrac{1}{4\ell-2} & \text{if } i=j=\ell,\\[2mm] \dfrac{-1}{4\ell-2} & \text{if } i \text{ and } j \text{ are consecutive},\\[2mm] 0 & \text{if } i \text{ and } j \text{ are not consecutive and } i\neq j.\end{cases}$$

Proof. Let $i\in\{1,\dots,\ell\}$. We have

$$\begin{aligned}&\kappa(\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix},\frac{1}{4\ell-2}\begin{bmatrix}0&0&0\\0&e_{ii}&0\\0&0&-e_{ii}\end{bmatrix})\\&=\frac{2\ell-1}{4\ell-2}\mathrm{tr}(\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix}\begin{bmatrix}0&0&0\\0&e_{ii}&0\\0&0&-e_{ii}\end{bmatrix})\\&=\frac{2\ell-1}{4\ell-2}\cdot 2h_{ii}\\&=h_{ii}\\&=\alpha(\begin{bmatrix}0&0&0\\0&h&0\\0&0&-h\end{bmatrix}).\end{aligned}$$

It follows that

$$t_{\alpha_i}=\frac{1}{4\ell-2}\begin{bmatrix}0&0&0\\0&e_{ii}&0\\0&0&-e_{ii}\end{bmatrix}.$$

Also let $j \in \{1, \dots, \ell\}$. Then

$$\begin{aligned}
(\alpha_i, \alpha_j) &= \kappa(t_{\alpha_i}, t_{\alpha_j}) \\
&= \frac{2\ell - 1}{(4\ell-2)^2} \mathrm{tr}(\begin{bmatrix} 0 & 0 & 0 \\ 0 & e_{ii} & 0 \\ 0 & 0 & -e_{ii} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & e_{jj} & 0 \\ 0 & 0 & -e_{jj} \end{bmatrix}) \\
&= \frac{2\ell - 1}{(4\ell-2)^2} \cdot \begin{cases} 2 & \text{if } i = j, \\ 0 & \text{if } i \neq j \end{cases} \\
&= \begin{cases} \dfrac{1}{4\ell - 2} & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}
\end{aligned}$$

Assume that $i, j \in \{1, \dots, \ell - 1\}$. Then

$$\begin{aligned}
(\beta_i, \beta_j) &= (\alpha_i - \alpha_{i+1}, \alpha_j - \alpha_{j+1}) \\
&= (\alpha_i, \alpha_j) - (\alpha_i, \alpha_{j+1}) - (\alpha_{i+1}, \alpha_j) + (\alpha_{i+1}, \alpha_{j+1}) \\
&= \begin{cases} \dfrac{2}{4\ell - 2} & \text{if } i = j, \\ \dfrac{-1}{4\ell - 2} & \text{if } i \text{ and } j \text{ are consecutive}, \\ 0 & \text{if } i \text{ and } j \text{ are not consecutive and } i \neq j. \end{cases}
\end{aligned}$$

Assume that $i \in \{1, \dots, \ell - 1\}$. Then

$$\begin{aligned}
(\beta_i, \beta_\ell) &= (\alpha_i - \alpha_{i+1}, \alpha_\ell) \\
&= (\alpha_i, \alpha_\ell) - (\alpha_{i+1}, \alpha_\ell) \\
&= -(\alpha_{i+1}, \alpha_\ell) \\
&= \begin{cases} \dfrac{-1}{4\ell - 2} & \text{if } i = \ell - 1, \\ 0 & \text{if } i \neq \ell - 1. \end{cases}
\end{aligned}$$

Finally,

$$\begin{aligned}
(\beta_\ell, \beta_\ell) &= (\alpha_\ell, \alpha_\ell) \\
&= \frac{1}{4\ell - 2}.
\end{aligned}$$

This completes the proof. □

Lemma 10.8.10. *Let ℓ be an integer such that $\ell \geq 2$. Let F have characteristic zero and be algebraically closed. Let ℓ be a positive integer. The Dynkin diagram of* $\mathrm{so}(2\ell+1, F)$ *is*

B_ℓ:

and the Cartan matrix of so$(2\ell+1, F)$ *is*

$$\begin{bmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & -1 & 2 & -1 & \\ & & & & -1 & 2 & -2 \\ & & & & & -1 & 2 \end{bmatrix}$$

The Lie algebra so$(2\ell+1, F)$ *is simple.*

Proof. Let $i, j \in \{1, \ldots, \ell\}$ with $i \neq j$. Then

$$\begin{aligned} \langle \beta_i, \beta_j \rangle &= 2\frac{(\beta_i, \beta_j)}{(\beta_j, \beta_j)} \\ &= \begin{cases} -2 & \text{if } i \text{ and } j \text{ are consecutive and } j = \ell, \\ -1 & \text{if } i \text{ and } j \text{ are consecutive and } j \neq \ell, \\ 0 & \text{if } i \text{ and } j \text{ are not consecutive.} \end{cases} \end{aligned}$$

Hence,

$$\begin{aligned} \langle \beta_i, \beta_j \rangle \langle \beta_j, \beta_i \rangle &= 4\frac{(\beta_i, \beta_j)^2}{(\beta_i, \beta_i)(\beta_j, \beta_j)} \\ &= \begin{cases} 2 & \text{if } i \text{ and } j \text{ are consecutive and } j = \ell \text{ or } i = \ell, \\ 1 & \text{if } i \text{ and } j \text{ are consecutive and } i \neq \ell \text{ and } j \neq \ell, \\ 0 & \text{if } i \text{ and } j \text{ are not consecutive.} \end{cases} \end{aligned}$$

It follows that the Dynkin diagram of so$(2\ell+1, F)$ is B_ℓ, and the Cartan matrix of so$(2\ell+1, F)$ is as stated. Since B_ℓ is connected, so$(2\ell+1, F)$ is simple by Lemma 9.3.2 and Proposition 10.3.2. □

Lemma 10.8.11. *Assume that the characteristic of F is zero and F is algebraically closed. Let ℓ be a positive integer. The Killing form*

$$\kappa : \mathrm{so}(2\ell+1, F) \times \mathrm{so}(2\ell+1, F) \longrightarrow F$$

is given by

$$\kappa(x, y) = (2\ell - 1) \cdot \mathrm{tr}(xy).$$

for $x, y \in$ so$(2\ell+1, F)$.

Proof. By Lemma 10.5.3, there exists $c \in F^\times$ such that $\kappa(x, y) = c\mathrm{tr}(xy)$ for $x, y \in$ so$(2\ell+1, F)$. Let H be the subalgebra of diagonal matrices in so$(2\ell+1, F)$; H is a Cartan subalgebra of so$(2\ell+1, F)$ by Lemma 10.8.6. By Lemma 10.8.8 we have $\kappa(h, h') = (2\ell-1)\cdot\mathrm{tr}(hh')$ for $h, h' \in H$. Hence, $c\mathrm{tr}(hh') = (2\ell-1)\cdot\mathrm{tr}(hh')$ for $h, h' \in H$. Since there exist $h, h' \in H$ such that $\mathrm{tr}(hh') \neq 0$ we conclude that $c = 2\ell - 1$. □

10.9 The Lie algebra $\mathrm{sp}(2\ell)$

Lemma 10.9.1. *Let ℓ be a positive integer. The Lie algebra $\mathrm{sp}(2\ell, F)$ consists of the matrices*

$$\begin{bmatrix} a & b \\ c & -{}^t a \end{bmatrix}$$

for $a, b, c \in \mathrm{gl}(\ell, F)$ with ${}^t b = b$ and ${}^t c = c$. The dimension of $\mathrm{sp}(2\ell, F)$ is $2\ell^2 + \ell$.

Proof. Let

$$x = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

with $a, b, c, d \in \mathrm{gl}(\ell, F)$. Then, by definition, $x \in \mathrm{sp}(2\ell, F)$ if and only if ${}^t x S = -Sx$ where

$$S = \begin{bmatrix} 0 & 1_\ell \\ -1_\ell & 0 \end{bmatrix}.$$

Thus,

$$\begin{aligned}
& x \in \mathrm{sp}(2\ell, F) \\
\iff & {}^t x S = -Sx \\
\iff & {}^t\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1_\ell \\ -1_\ell & 0 \end{bmatrix} = -\begin{bmatrix} 0 & 1_\ell \\ -1_\ell & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \\
\iff & \begin{bmatrix} {}^t a & {}^t c \\ {}^t b & {}^t d \end{bmatrix} \begin{bmatrix} 0 & 1_\ell \\ -1_\ell & 0 \end{bmatrix} = \begin{bmatrix} -c & -d \\ a & b \end{bmatrix} \\
\iff & \begin{bmatrix} -{}^t c & {}^t a \\ -{}^t d & {}^t b \end{bmatrix} = \begin{bmatrix} -c & -d \\ a & b \end{bmatrix}.
\end{aligned}$$

This is the first assertion of the lemma. Using this result it is straightforward to see that $\dim_F \mathrm{sp}(2\ell, F) = 2\ell^2 + \ell$. □

Lemma 10.9.2. *Let ℓ be a positive integer. Let F have characteristic zero and be algebraically closed. The natural action of $\mathrm{sp}(2\ell, F)$ on $V = \mathrm{M}_{2\ell,1}(F)$ is irreducible, so that $\mathrm{sp}(2\ell, F)$ is semi-simple.*

Proof. Let W be a non-zero $\mathrm{sp}(2\ell, F)$ subspace of V; we need to prove that $W = V$. Since W is non-zero, W contains a non-zero vector

$$v = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Assume first that $x \neq 0$ and $y = 0$. Now

$$\begin{bmatrix} a & 0 \\ 0 & -{}^t a \end{bmatrix} w = \begin{bmatrix} a & 0 \\ 0 & -{}^t a \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} ax \\ 0 \end{bmatrix}$$

for $a \in \mathrm{gl}(\ell, F)$. Since $x \neq 0$ and the action of $\mathrm{gl}(\ell, F)$ on $\mathrm{M}_{\ell,1}(F)$ is irreducible, it follows that W contains all vectors of the form

$$\begin{bmatrix} * \\ 0 \end{bmatrix}.$$

Now

$$\begin{bmatrix} 0 & 0 \\ 1_\ell & 0 \end{bmatrix}$$

is contained in $\mathrm{sp}(2\ell, F)$ and

$$\begin{bmatrix} 0 & 0 \\ 1_\ell & 0 \end{bmatrix} \begin{bmatrix} x' \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ x' \end{bmatrix}$$

for $x \in \mathrm{M}_{\ell,1}(F)$. It follows that W contains all the vectors of the form

$$\begin{bmatrix} 0 \\ * \end{bmatrix}.$$

We conclude that, in the current case, $W = V$. If $x = 0$ and $y \neq 0$, then a similar argument shows that $W = V$. Assume that $x \neq 0$ and $y \neq 0$. Since $x \neq 0$ and $y \neq 0$, there exists $a \in \mathrm{GL}(\ell, F)$ such that $ax = y$. Now

$$\begin{bmatrix} a & -1 \\ 0 & -{}^t a \end{bmatrix}$$

is contained in $\mathrm{sp}(2\ell, F)$, and

$$\begin{bmatrix} a & -1 \\ 0 & -{}^t a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ -{}^t ay \end{bmatrix}.$$

Since a is invertible, and $y \neq 0$, we have $-{}^t ay \neq 0$. We are now in the situation of a previous case; it follows that $W = V$.

Finally, $\mathrm{sp}(2\ell, F)$ is semi-simple by Lemma 10.2.1 (note that $\mathrm{sp}(2\ell, F)$ is contained in $\mathrm{sl}(2\ell, F)$ by Lemma 10.1.1). □

Lemma 10.9.3. *Let F have characteristic zero and be algebraically closed. The set H of diagonal matrices in $\mathrm{sp}(2\ell, F)$ is a Cartan subalgebra of $\mathrm{sp}(2\ell, F)$.*

Proof. By Lemma 10.4.2, to prove that H is a Cartan subalgebra of $\mathrm{sp}(2\ell, F)$, it suffices prove that if $w \in \mathrm{sp}(2\ell, F)$ has zero entries on the main diagonal and $wh = hw$ for $h \in H$, then $w = 0$. Let w be such an element of $\mathrm{sp}(2\ell, F)$, and write, as usual,

$$w = \begin{bmatrix} a & b \\ c & -{}^t a \end{bmatrix}.$$

By assumption, a has zero on the main diagonal. Let $h \in H$, so that

$$h = \begin{bmatrix} d & 0 \\ 0 & -d \end{bmatrix}$$

where $d \in \mathrm{gl}(\ell, F)$ is diagonal. We have

$$wh = \begin{bmatrix} a & b \\ c & -{}^{t}a \end{bmatrix} \begin{bmatrix} d & 0 \\ 0 & -d \end{bmatrix} = \begin{bmatrix} ad & -bd \\ cd & {}^{t}ad \end{bmatrix}$$

and

$$hw = \begin{bmatrix} d & 0 \\ 0 & -d \end{bmatrix} \begin{bmatrix} a & b \\ c & -{}^{t}a \end{bmatrix} = \begin{bmatrix} da & db \\ -dc & d\,{}^{t}a \end{bmatrix}.$$

It follows that

$$\begin{aligned} ad &= da, \\ bd &= -db, \\ cd &= -dc, \\ {}^{t}ad &= d\,{}^{t}a. \end{aligned}$$

Lemma 10.8.5 implies that $b = c = 0$ and that a is diagonal. Since a has zeros on the main diagonal by assumption, we also get $a = 0$. Hence, $w = 0$. □

Lemma 10.9.4. *Let ℓ be an integer such that $\ell \geq 2$. Let F have characteristic zero and be algebraically closed. Let ℓ be a positive integer. Let H be the subalgebra of $\mathrm{sp}(2\ell, F)$ consisting of diagonal matrices; by Lemma 10.9.3, H is a Cartan subalgebra of $\mathrm{sp}(2\ell, F)$. Let Φ be the set of roots of $\mathrm{sp}(2\ell, F)$ defined with respect to H. Let $V = \mathbb{R} \otimes_{\mathbb{Q}} V_0$, where V_0 is the $\mathbb{Q}$ subspace of $H^{\vee} = \mathrm{Hom}_F(H, F)$ spanned by the elements of Φ; by Proposition 8.2.1, Φ is a root system in V. For $i \in \{1, \dots, \ell\}$, define a linear functional*

$$\alpha_i : H \longrightarrow F$$

by

$$\alpha_i(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}) = h_{ii}$$

for $h \in \mathrm{gl}(\ell, F)$ and h diagonal. The set Φ consists of the following $2\ell^2$ linear functionals on H:

$$\begin{gathered} \alpha_i - \alpha_j, \quad i, j \in \{1, \dots, \ell\}, \quad i \neq j, \\ 2\alpha_1, \dots, 2\alpha_\ell, \\ -2\alpha_1, \dots, -2\alpha_\ell, \\ \alpha_i + \alpha_j, \quad i, j \in \{1, \dots, \ell\}, \quad i < j, \\ -(\alpha_i + \alpha_j), \quad i, j \in \{1, \dots, \ell\}, \quad i < j. \end{gathered}$$

The set

$$B = \{\beta_1 = \alpha_1 - \alpha_2,\ \beta_2 = \alpha_2 - \alpha_3,\ \dots,\ \beta_{\ell-1} = \alpha_{\ell-1} - \alpha_\ell,\ \beta_\ell = 2\alpha_\ell\}$$

is a base for Φ, and the positive roots with respect to B are the set P, where P consists of the following roots:

$$\alpha_i - \alpha_j, \quad i, j \in \{1, \dots, \ell\}, \quad i < j,$$

$$2\alpha_1, \dots, 2\alpha_n,$$
$$\alpha_i + \alpha_j, \quad i, j \in \{1, \dots, \ell\}, \quad i < j.$$

The root spaces are:

$$\begin{aligned}
L_{\alpha_i - \alpha_j} &= F \cdot \begin{bmatrix} e_{ij} & 0 \\ 0 & -e_{ij} \end{bmatrix}, \quad i, j \in \{1, \dots, \ell\}, \quad i \neq j, \\
L_{2\alpha_i} &= F \cdot \begin{bmatrix} 0 & e_{ii} \\ 0 & 0 \end{bmatrix}, \quad i \in \{1, \dots, \ell\}, \\
L_{-2\alpha_i} &= F \cdot \begin{bmatrix} 0 & 0 \\ e_{ii} & 0 \end{bmatrix}, \quad i \in \{1, \dots, \ell\}, \\
L_{\alpha_i + \alpha_j} &= F \cdot \begin{bmatrix} 0 & e_{ij} + e_{ji} \\ 0 & 0 \end{bmatrix}, \quad i, j \in \{1, \dots, \ell\}, \quad i < j, \\
L_{-(\alpha_i + \alpha_j)} &= F \cdot \begin{bmatrix} 0 & 0 \\ e_{ij} + e_{ji} & 0 \end{bmatrix}, \quad i, j \in \{1, \dots, \ell\}, \quad i < j.
\end{aligned}$$

Proof. Let $h \in \mathrm{gl}(\ell, F)$ be a diagonal matrix. Let $i, j \in \{1, \dots, \ell\}$ with $i \neq j$. Then

$$\begin{aligned}
[\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \begin{bmatrix} e_{ij} & 0 \\ 0 & -e_{ij} \end{bmatrix}] &= \begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix} \begin{bmatrix} e_{ij} & 0 \\ 0 & -e_{ij} \end{bmatrix} - \begin{bmatrix} e_{ij} & 0 \\ 0 & -e_{ij} \end{bmatrix} \begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix} \\
&= \begin{bmatrix} h_{ii}e_{ij} & 0 \\ 0 & -h_{ii}e_{ij} \end{bmatrix} - \begin{bmatrix} h_{jj}e_{ij} & 0 \\ 0 & -h_{jj}e_{ij} \end{bmatrix} \\
&= (h_{ii} - h_{jj}) \begin{bmatrix} e_{ij} & 0 \\ 0 & -e_{ij} \end{bmatrix}.
\end{aligned}$$

This equation proves that $\alpha_i - \alpha_j$ is a root and that $L_{\alpha_i - \alpha_j}$ is as stated. Next, let $h \in \mathrm{gl}(\ell, F)$ be a diagonal matrix, and let $i, j \in \{1, \dots, \ell\}$. Then

$$\begin{aligned}
[\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & e_{ij} + e_{ji} \\ 0 & 0 \end{bmatrix}] &= \begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix} \begin{bmatrix} 0 & e_{ij} + e_{ji} \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & e_{ij} + e_{ji} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix} \\
&= \begin{bmatrix} 0 & h_{ii}e_{ij} + h_{jj}e_{ji} \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -h_{jj}e_{ij} - h_{ii}e_{ji} \\ 0 & 0 \end{bmatrix} \\
&= (h_{ii} + h_{jj}) \begin{bmatrix} 0 & e_{ij} + e_{ji} \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

This proves that $2\alpha_i$ is a root for $i \in \{1, \dots, \ell\}$ and that $\alpha_i + \alpha_j$ is a root for $i, j \in \{1, \dots, \ell\}$ with $i < j$; also the root spaces of these roots are as stated. Again let $h \in \mathrm{gl}(\ell, F)$ be a diagonal matrix, and let $i, j \in \{1, \dots, \ell\}$. Taking tranposes of the last equation, we obtain:

$$\begin{aligned}
{}^{\mathrm{t}}[\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & e_{ij} + e_{ji} \\ 0 & 0 \end{bmatrix}] &= (h_{ii} + h_{jj})\, {}^{\mathrm{t}}\begin{bmatrix} 0 & e_{ij} + e_{ji} \\ 0 & 0 \end{bmatrix} \\
[\begin{bmatrix} 0 & 0 \\ e_{ij} + e_{ji} & 0 \end{bmatrix}, \begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}] &= (h_{ii} + h_{jj}) \begin{bmatrix} 0 & 0 \\ e_{ij} + e_{ji} & 0 \end{bmatrix}
\end{aligned}$$

$$[\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ e_{ij}+e_{ji} & 0 \end{bmatrix}] = -(h_{ii}+h_{jj}) \begin{bmatrix} 0 & 0 \\ e_{ij}+e_{ji} & 0 \end{bmatrix}.$$

This proves that $-2\alpha_i$ is a root for $i \in \{1, \dots, \ell\}$ and that $-(\alpha_i + \alpha_j)$ is a root for $i, j \in \{1, \dots, \ell\}$ with $i < j$; also the root spaces of these roots are as described.

To see that B is a base for Φ we note first that $\dim_F V = \ell$, and that the elements of B are evidently linearly independent; it follows that B is a basis for the F-vector space V. Since B is the disjoint union of P and $\{-\lambda : \lambda \in P\}$, to prove that B is a base for Φ it will now suffice to prove that every element of P is a linear combination of elements from B with non-negative integer coefficients. Let $i, j \in \{1, \dots, \ell\}$ with $i < j$. Then

$$\alpha_i - \alpha_j = \beta_{i+1} + \cdots + \beta_j.$$

Also, we have

$$\begin{aligned} 2\alpha_\ell &= \beta_\ell, \\ 2\alpha_{\ell-1} &= 2(\alpha_{\ell-1} - \alpha_\ell) + 2\alpha_\ell = 2\beta_{\ell-1} + \beta_\ell, \\ 2\alpha_{\ell-2} &= 2(\alpha_{\ell-2} - \alpha_{\ell-1}) + 2\alpha_{\ell-1} = 2\beta_{\ell-2} + 2\beta_{\ell-1} + \beta_\ell, \\ &\cdots \\ 2\alpha_1 &= 2\beta_1 + \cdots 2\beta_{\ell-1} + \beta_\ell. \end{aligned}$$

Finally, let $i, j \in \{1, \dots, \ell\}$ with $i < j$. Then

$$\alpha_i + \alpha_j = (\alpha_i - \alpha_j) + 2\alpha_j = \beta_{i+1} + \cdots + \beta_j + 2\beta_1 + \cdots + 2\beta_{\ell-1} + \beta_\ell.$$

This completes the proof. □

h_{11}	$\boxed{\beta_1 = \alpha_1 - \alpha_2}$	$\boxed{\alpha_1 - \alpha_3}$	$\boxed{2\alpha_1}$	$\boxed{\alpha_1 + \alpha_2}$	$\boxed{\alpha_1 + \alpha_3}$
$\alpha_2 - \alpha_1$	h_{22}	$\boxed{\beta_2 = \alpha_2 - \alpha_3}$	$*$	$\boxed{2\alpha_2}$	$\boxed{\alpha_2 + \alpha_3}$
$\alpha_3 - \alpha_1$	$\alpha_3 - \alpha_2$	h_{33}	$*$	$*$	$\boxed{\beta_3 = 2\alpha_3}$
$-2\alpha_1$	$-(\alpha_1 + \alpha_2)$	$-(\alpha_1 + \alpha_3)$	$-h_{11}$	$*$	$*$
$*$	$-2\alpha_2$	$-(\alpha_2 + \alpha_3)$	$*$	$-h_{22}$	$*$
$*$	$*$	$-2\alpha_3$	$*$	$*$	$-h_{33}$

Figure 10.3: The decomposition of $\mathrm{sp}(6, F)$. For this example, $\ell = 3$. The positions are labeled with the corresponding root. Note that the diagonal is our chosen Cartan subalgebra. The positive roots with respect to our chosen base $\{\beta_1, \beta_2, \beta_3\}$ are boxed, while the colored roots form our chosen base. Positions labeled with $*$ are determined by other entries. The linear functionals α_1, α_2 and α_3 are defined in Proposition 10.9.4.

Lemma 10.9.5. *Assume that the characteristic of F is zero and F is algebraically closed. Let ℓ be a positive integer. The Killing form*

$$\kappa : \mathrm{sp}(2\ell, F) \times \mathrm{sp}(2\ell, F) \longrightarrow F$$

is given by

$$\kappa(h, h') = (2\ell + 2) \cdot \mathrm{tr}(hh')$$

for $h, h' \in H$. Here, H is the subalgebra of diagonal matrices in $\mathrm{sp}(2\ell, F)$; H is a Cartan subalgebra of $\mathrm{sp}(2\ell, F)$ by Lemma 10.9.3.

Proof. Let $h, h' \in \mathrm{gl}(\ell, F)$ be diagonal matrices. Then

$$\begin{aligned}
&\kappa(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \begin{bmatrix} h' & 0 \\ 0 & -h' \end{bmatrix}) \\
&= \mathrm{tr}(\mathrm{ad}(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}) \circ \mathrm{ad}(\begin{bmatrix} h' & 0 \\ 0 & -h' \end{bmatrix})) \\
&= \sum_{i,j \in \{1,\dots,\ell\}, i \neq j} (h_{ii} - h_{jj})(h'_{ii} - h'_{jj}) \\
&\quad + 2 \sum_{i \in \{1,\dots,\ell\}} 4h_{ii}h'_{ii} \\
&\qquad + 2 \sum_{i,j \in \{1,\dots,\ell\}, i<j} (h_{ii} + h_{jj})(h'_{ii} + h'_{jj}) \\
&= \sum_{i,j \in \{1,\dots,\ell\}} (h_{ii} - h_{jj})(h'_{ii} - h'_{jj}) \\
&\quad + 8 \sum_{i \in \{1,\dots,\ell\}} h_{ii}h'_{ii} \\
&\qquad + \sum_{i,j \in \{1,\dots,\ell\}} (h_{ii} + h_{jj})(h'_{ii} + h'_{jj}) \\
&\qquad\quad - \sum_{i \in \{1,\dots,\ell\}} (h_{ii} + h_{ii})(h'_{ii} + h'_{ii}) \\
&= \sum_{i,j \in \{1,\dots,\ell\}} h_{ii}h'_{ii} - h_{ii}h'_{jj} - h_{jj}h'_{ii} + h_{jj}h'_{jj} \\
&\quad + 4 \sum_{i \in \{1,\dots,\ell\}} h_{ii}h'_{ii} \\
&\qquad + \sum_{i,j \in \{1,\dots,\ell\}} h_{ii}h'_{ii} + h_{ii}h'_{jj} + h_{jj}h'_{ii} + h_{jj}h'_{jj} \\
&= 2\ell \sum_{i \in \{1,\dots,\ell\}} h_{ii}h'_{ii} \\
&\quad + 4 \sum_{i \in \{1,\dots,\ell\}} h_{ii}h'_{ii}
\end{aligned}$$

$$
\begin{aligned}
&\quad + 2\ell \sum_{i\in\{1,\dots,\ell\}} h_{ii}h'_{ii} \\
&= (4\ell+4) \sum_{i\in\{1,\dots,\ell\}} h_{ii}h'_{ii} \\
&= (2\ell+2)\cdot \mathrm{tr}(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}\begin{bmatrix} h' & 0 \\ 0 & -h' \end{bmatrix}).
\end{aligned}
$$

This completes the calculation. □

Lemma 10.9.6. *Let ℓ be an integer such that $\ell \geq 2$. Let the notation as in Lemma 10.9.4. For $i, j \in \{1, \dots, \ell\}$,*

$$
(\beta_i, \beta_j) = \begin{cases} \dfrac{2}{4\ell+4} & \text{if } i,j \in \{1,\dots,\ell-1\} \text{ and } i=j \\ \dfrac{-1}{4\ell+4} & \text{if } i,j \in \{1,\dots,\ell-1\} \text{ and } i \text{ and } j \text{ are consecutive,} \\ \dfrac{-2}{4\ell+4} & \text{if } \{i,j\} = \{\ell-1,\ell\}, \\ \dfrac{4}{4\ell+4} & \text{if } i=j=\ell, \\ 0 & \text{if none of the above conditions hold.} \end{cases}
$$

Proof. Let $h \in \mathrm{gl}(\ell, F)$ be a diagonal matrix. Let $i \in \{1, \dots, \ell\}$. Then

$$
\begin{aligned}
\kappa(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \frac{1}{4\ell+4}\begin{bmatrix} e_{ii} & 0 \\ 0 & -e_{ii} \end{bmatrix}) &= \frac{2\ell+2}{4\ell+4}\mathrm{tr}(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}\begin{bmatrix} e_{ii} & 0 \\ 0 & -e_{ii} \end{bmatrix}) \\
&= h_{ii} \\
&= \alpha_i(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}).
\end{aligned}
$$

Since this holds for all diagonal $h \in \mathrm{gl}(\ell, F)$, it follows that

$$
t_{\alpha_i} = \frac{1}{4\ell+4}\begin{bmatrix} e_{ii} & 0 \\ 0 & -e_{ii} \end{bmatrix}.
$$

Also let $j \in \{1, \dots, \ell\}$. Then

$$
\begin{aligned}
(\alpha_i, \alpha_j) &= \kappa(t_{\alpha_i}, t_{\alpha_j}) \\
&= (2\ell+2)\cdot \mathrm{tr}(\frac{1}{4\ell+4}\begin{bmatrix} e_{ii} & 0 \\ 0 & -e_{ii} \end{bmatrix} \cdot \frac{1}{4\ell+4}\begin{bmatrix} e_{jj} & 0 \\ 0 & -e_{jj} \end{bmatrix}) \\
&= \begin{cases} \dfrac{1}{4\ell+4} & \text{if } i=j, \\ 0 & \text{if } i \neq j. \end{cases}
\end{aligned}
$$

Let $i, j \in \{1, \ldots, \ell - 1\}$. Then

$$
\begin{aligned}
(\beta_i, \beta_j) &= (\alpha_i - \alpha_{i+1}, \alpha_j - \alpha_{j+1}) \\
&= (\alpha_i, \alpha_j) - (\alpha_i, \alpha_{j+1}) - (\alpha_{i+1}, \alpha_j) + (\alpha_{i+1}, \alpha_{j+1}) \\
&= \begin{cases} \dfrac{2}{4\ell + 4} & \text{if } i = j, \\ \dfrac{-1}{4\ell + 4} & \text{if } i \text{ and } j \text{ are consecutive}, \\ 0 & \text{if } i \neq j \text{ and } i \text{ and } j \text{ are not consecutive}. \end{cases}
\end{aligned}
$$

Let $i \in \{1, \ldots, \ell - 1\}$. Then

$$
\begin{aligned}
(\beta_i, \beta_\ell) &= (\alpha_i - \alpha_{i+1}, 2\alpha_\ell) \\
&= 2(\alpha_i, \alpha_\ell) - 2(\alpha_{i+1}, \alpha_\ell) \\
&= -2(\alpha_{i+1}, \alpha_\ell) \\
&= \begin{cases} \dfrac{-2}{4\ell + 4} & \text{if } i = \ell - 1, \\ 0 & \text{if } i \neq \ell - 1. \end{cases}
\end{aligned}
$$

Finally,

$$
\begin{aligned}
(\beta_\ell, \beta_\ell) &= 4(\alpha_\ell, \alpha_\ell) \\
&= \frac{4}{4\ell + 4}.
\end{aligned}
$$

This completes the proof. □

Lemma 10.9.7. *Let ℓ be an integer such that $\ell \geq 2$. Let F have characteristic zero and be algebraically closed. Let ℓ be a positive integer. The Dynkin diagram of* $\mathrm{sp}(2\ell, F)$ *is*

C_ℓ:

and the Cartan matrix of $\mathrm{sp}(2\ell, F)$ *is*

$$
\begin{bmatrix}
2 & -1 & & & & & \\
-1 & 2 & -1 & & & & \\
 & -1 & 2 & -1 & & & \\
 & & \ddots & \ddots & \ddots & & \\
 & & & -1 & 2 & -1 & \\
 & & & & -1 & 2 & -1 \\
 & & & & & -2 & 2
\end{bmatrix}
$$

The Lie algebra $\mathrm{sp}(2\ell, F)$ *is simple.*

Proof. Let $i, j \in \{1, \ldots, \ell\}$ with $i \neq j$. Then

$$\langle \beta_i, \beta_j \rangle = 2\frac{(\beta_i, \beta_j)}{(\beta_j, \beta_j)} = \begin{cases} -1 & \text{if } i, j \in \{1, \ldots, \ell - 1\} \text{ and } i \text{ and } j \text{ are consecutive,} \\ -1 & \text{if } i = \ell - 1 \text{ and } j = \ell, \\ -2 & \text{if } i = \ell \text{ and } j = \ell - 1, \\ 0 & \text{if none of the above conditions hold.} \end{cases}$$

Hence,

$$\langle \beta_i, \beta_j \rangle \langle \beta_j, \beta_i \rangle = 4\frac{(\beta_i, \beta_j)^2}{(\beta_i, \beta_i)(\beta_j, \beta_j)} = \begin{cases} 1 & \text{if } i, j \in \{1, \ldots, \ell - 1\} \text{ and } i \text{ and } j \text{ are consecutive,} \\ 2 & \text{if } i = \ell - 1 \text{ and } j = \ell, \\ 0 & \text{if none of the above conditions hold.} \end{cases}$$

It follows that the Dynkin diagram of $\mathrm{sp}(2\ell, F)$ is C_ℓ, and the Cartan matrix of $\mathrm{sp}(2\ell, F)$ is as stated. Since C_ℓ is connected, $\mathrm{sp}(2\ell, F)$ is simple by Lemma 9.3.2 and Proposition 10.3.2. □

10.10 The Lie algebra $\mathrm{so}(2\ell)$

Lemma 10.10.1. *Let ℓ be a positive integer. The Lie algebra $\mathrm{so}(2\ell, F)$ consists of the matrices*

$$\begin{bmatrix} a & b \\ c & -{}^t a \end{bmatrix}$$

for $a, b, c \in \mathrm{gl}(\ell, F)$ with $-{}^t b = b$ and $-{}^t c = c$. The dimension of $\mathrm{so}(2\ell, F)$ is $2\ell^2 - \ell$.

Proof. Let $x \in \mathrm{gl}(2\ell, F)$. Write

$$x = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

with $a, b, c, d \in \mathrm{gl}(\ell, F)$. By definition, $x \in \mathrm{so}(2\ell, F)$ if and only if ${}^t x S + Sx = 0$, where

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

Hence

$$\begin{aligned} &x \in \mathrm{so}(2\ell, F) \\ &\iff {}^{\mathrm{t}}xS = -Sx, \\ &\iff \begin{bmatrix} {}^{\mathrm{t}}a & {}^{\mathrm{t}}c \\ {}^{\mathrm{t}}b & {}^{\mathrm{t}}d \end{bmatrix} \begin{bmatrix} 0 & 1_\ell \\ 1_\ell & 0 \end{bmatrix} = - \begin{bmatrix} 0 & 1_\ell \\ 1_\ell & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \\ &\iff \begin{bmatrix} {}^{\mathrm{t}}c & {}^{\mathrm{t}}a \\ {}^{\mathrm{t}}d & {}^{\mathrm{t}}b \end{bmatrix} = \begin{bmatrix} -c & -d \\ -a & -b \end{bmatrix}. \end{aligned}$$

This is the first assertion of the lemma. □

Lemma 10.10.2. *Let ℓ be an integer such that $\ell \geq 2$. Let F have characteristic zero and be algebraically closed. The natural action of $\mathrm{so}(2\ell, F)$ on $V = \mathrm{M}_{2\ell,1}(F)$ is irreducible, so that $\mathrm{so}(2\ell, F)$ is semi-simple.*

Proof. Let W be a non-zero $\mathrm{so}(2\ell, F)$ subspace of V; we need to prove that $W = V$. Since W is non-zero, W contains a non-zero vector

$$v = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Assume first that $x \neq 0$ and $y = 0$. Now

$$\begin{bmatrix} a & 0 \\ 0 & -{}^{\mathrm{t}}a \end{bmatrix} w = \begin{bmatrix} a & 0 \\ 0 & -{}^{\mathrm{t}}a \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} ax \\ 0 \end{bmatrix}$$

for $a \in \mathrm{gl}(\ell, F)$. Since $x \neq 0$ and the action of $\mathrm{gl}(\ell, F)$ on $\mathrm{M}_{\ell,1}(F)$ is irreducible, it follows that W contains all vectors of the form

$$\begin{bmatrix} * \\ 0 \end{bmatrix}.$$

By Lemma 10.8.3 there exists $c \in \mathrm{gl}(\ell, F)$ such that $-{}^{\mathrm{t}}c = c$ and $cx \neq 0$. The matrix

$$\begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix}$$

is contained in $\mathrm{so}(2\ell, F)$ and

$$\begin{bmatrix} 0 & 0 \\ c & 0 \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ cx \end{bmatrix}$$

This non-zero. An argument as above shows that W contains all the vectors of the form

$$\begin{bmatrix} 0 \\ * \end{bmatrix}.$$

We conclude that, in the current case, $W = V$. If $x = 0$ and $y \neq 0$, then a similar argument shows that $W = V$. Assume that $x \neq 0$ and $y \neq 0$. By Lemma

10.8.3 there exists $b \in \mathrm{gl}(\ell, F)$ such that $-{}^{t}b = b$ and $by \neq 0$. Since $by \neq 0$ and $x \neq 0$, there exists $a \in \mathrm{GL}(\ell, F)$ such that $ax = -by$. Now

$$\begin{bmatrix} a & b \\ 0 & -{}^{t}a \end{bmatrix}$$

is contained in $\mathrm{so}(2\ell, F)$, and

$$\begin{bmatrix} a & b \\ 0 & -{}^{t}a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ -{}^{t}ay \end{bmatrix}.$$

Since a is invertible, and $y \neq 0$, we have $-{}^{t}ay \neq 0$. We are now in the situation of a previous case; it follows that $W = V$.

Finally, $\mathrm{so}(2\ell, F)$ is semi-simple by Lemma 10.2.1 (note that $\mathrm{so}(2\ell, F)$ is contained in $\mathrm{sl}(2\ell, F)$ by Lemma 10.1.1). □

Lemma 10.10.3. *Let ℓ be an integer such that $\ell \geq 2$. Let F have characteristic zero and be algebraically closed. The set H of diagonal matrices in $\mathrm{so}(2\ell, F)$ is a Cartan subalgebra of $\mathrm{so}(2\ell, F)$.*

Proof. By Lemma 10.4.2, to prove that H is a Cartan subalgebra of $\mathrm{so}(2\ell, F)$, it suffices prove that if $w \in \mathrm{so}(2\ell, F)$ has zero entries on the main diagonal and $wh = hw$ for $h \in H$, then $w = 0$. Let w be such an element of $\mathrm{so}(2\ell, F)$, and write, as usual,

$$w = \begin{bmatrix} a & b \\ c & -{}^{t}a \end{bmatrix}.$$

By assumption, a has zeros on the main diagonal. Let $h \in H$, so that

$$h = \begin{bmatrix} d & 0 \\ 0 & -d \end{bmatrix}$$

where $d \in \mathrm{gl}(\ell, F)$ is diagonal. We have

$$wh = \begin{bmatrix} a & b \\ c & -{}^{t}a \end{bmatrix} \begin{bmatrix} d & 0 \\ 0 & -d \end{bmatrix} = \begin{bmatrix} ad & -bd \\ cd & {}^{t}ad \end{bmatrix}$$

and

$$hw = \begin{bmatrix} d & 0 \\ 0 & -d \end{bmatrix} \begin{bmatrix} a & b \\ c & -{}^{t}a \end{bmatrix} = \begin{bmatrix} da & db \\ -dc & d\,{}^{t}a \end{bmatrix}.$$

It follows that

$$\begin{aligned} ad &= da, \\ bd &= -db, \\ cd &= -dc, \\ {}^{t}ad &= d\,{}^{t}a. \end{aligned}$$

Lemma 10.8.5 implies that $b = c = 0$ and that a is diagonal. Since a has zeros on the main diagonal by assumption, we also get $a = 0$. Hence, $w = 0$. □

Lemma 10.10.4. *Let ℓ be an integer such that $\ell \geq 2$. Let F have characteristic zero and be algebraically closed. Let ℓ be a positive integer. Let H be the subalgebra of $\mathrm{so}(2\ell, F)$ consisting of diagonal matrices; by Lemma 10.10.3, H is a Cartan subalgebra of $\mathrm{so}(2\ell, F)$. Let Φ be the set of roots of $\mathrm{so}(2\ell, F)$ defined with respect to H. Let $V = \mathbb{R} \otimes_{\mathbb{Q}} V_0$, where V_0 is the $\mathbb{Q}$ subspace of $H^\vee = \mathrm{Hom}_F(H, F)$ spanned by the elements of Φ; by Proposition 8.2.1, Φ is a root system in V. For $i \in \{1, \dots, \ell\}$, define a linear functional*

$$\alpha_i : H \longrightarrow F$$

by

$$\alpha_i(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}) = h_{ii}$$

for $h \in \mathrm{gl}(\ell, F)$ and h diagonal. The set Φ consists of the following $2\ell^2 - 2\ell$ linear functionals on H:

$$\begin{aligned} \alpha_i - \alpha_j, &\quad i, j \in \{1, \dots, \ell\}, \quad i \neq j, \\ \alpha_i + \alpha_j, &\quad i, j \in \{1, \dots, \ell\}, \quad i < j, \\ -(\alpha_i + \alpha_j), &\quad i, j \in \{1, \dots, \ell\}, \quad i < j. \end{aligned}$$

The set

$$B = \{\beta_1 = \alpha_1 - \alpha_2,\ \beta_2 = \alpha_2 - \alpha_3,\ \dots,\ \beta_{\ell-1} = \alpha_{\ell-1} - \alpha_\ell,\ \beta_\ell = \alpha_{\ell-1} + \alpha_\ell\}$$

is a base for Φ, and the positive roots with respect to B are the set P, where P consists of the following roots:

$$\begin{aligned} \alpha_i - \alpha_j, &\quad i, j \in \{1, \dots, \ell\}, \quad i < j, \\ \alpha_i + \alpha_j, &\quad i, j \in \{1, \dots, \ell\}, \quad i < j. \end{aligned}$$

The root spaces are:

$$\begin{aligned} L_{\alpha_i - \alpha_j} &= F \cdot \begin{bmatrix} e_{ij} & 0 \\ 0 & -e_{ij} \end{bmatrix}, \quad i, j \in \{1, \dots, \ell\}, \quad i \neq j, \\ L_{\alpha_i + \alpha_j} &= F \cdot \begin{bmatrix} 0 & e_{ij} + e_{ji} \\ 0 & 0 \end{bmatrix}, \quad i, j \in \{1, \dots, \ell\}, \quad i < j, \\ L_{-(\alpha_i + \alpha_j)} &= F \cdot \begin{bmatrix} 0 & 0 \\ e_{ij} + e_{ji} & 0 \end{bmatrix}, \quad i, j \in \{1, \dots, \ell\}, \quad i < j. \end{aligned}$$

Proof. Let $h \in \mathrm{gl}(\ell, F)$ be a diagonal matrix. Let $i, j \in \{1, \dots, \ell\}$ with $i \neq j$. Then

$$\begin{aligned} [\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \begin{bmatrix} e_{ij} & 0 \\ 0 & -e_{ij} \end{bmatrix}] &= \begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix} \begin{bmatrix} e_{ij} & 0 \\ 0 & -e_{ij} \end{bmatrix} - \begin{bmatrix} e_{ij} & 0 \\ 0 & -e_{ij} \end{bmatrix} \begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix} \\ &= \begin{bmatrix} h_{ii}e_{ij} & 0 \\ 0 & -h_{ii}e_{ij} \end{bmatrix} - \begin{bmatrix} h_{jj}e_{ij} & 0 \\ 0 & -h_{jj}e_{ij} \end{bmatrix} \end{aligned}$$

$$= (h_{ii} - h_{jj}) \begin{bmatrix} e_{ij} & 0 \\ 0 & -e_{ij} \end{bmatrix}.$$

This equation proves that $\alpha_i - \alpha_j$ is a root and that $L_{\alpha_i - \alpha_j}$ is as stated. Next, let $h \in \mathrm{gl}(\ell, F)$ be a diagonal matrix, and let $i, j \in \{1, \dots, \ell\}$ with $i < j$. Then

$$\begin{aligned}
[\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & e_{ij} - e_{ji} \\ 0 & 0 \end{bmatrix}] &= \begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix} \begin{bmatrix} 0 & e_{ij} - e_{ji} \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & e_{ij} - e_{ji} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix} \\
&= \begin{bmatrix} 0 & h_{ii}e_{ij} - h_{jj}e_{ji} \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -h_{jj}e_{ij} + h_{ii}e_{ji} \\ 0 & 0 \end{bmatrix} \\
&= (h_{ii} + h_{jj}) \begin{bmatrix} 0 & e_{ij} - e_{ji} \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

This proves that that $\alpha_i + \alpha_j$ is a root for $i, j \in \{1, \dots, \ell\}$ with $i < j$; also the root spaces of these roots are as stated. Again let $h \in \mathrm{gl}(\ell, F)$ be a diagonal matrix, and let $i, j \in \{1, \dots, \ell\}$ with $i < j$. Taking tranposes of the last equation, we obtain:

$$\begin{aligned}
{}^t[\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & e_{ij} - e_{ji} \\ 0 & 0 \end{bmatrix}] &= (h_{ii} + h_{jj}) {}^t\begin{bmatrix} 0 & e_{ij} - e_{ji} \\ 0 & 0 \end{bmatrix} \\
[\begin{bmatrix} 0 & 0 \\ e_{ji} - e_{ij} & 0 \end{bmatrix}, \begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}] &= (h_{ii} + h_{jj}) \begin{bmatrix} 0 & 0 \\ e_{ji} - e_{ij} & 0 \end{bmatrix} \\
[\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ e_{ij} - e_{ji} & 0 \end{bmatrix}] &= -(h_{ii} + h_{jj}) \begin{bmatrix} 0 & 0 \\ e_{ij} - e_{ji} & 0 \end{bmatrix}.
\end{aligned}$$

This proves that that $-(\alpha_i + \alpha_j)$ is a root for $i, j \in \{1, \dots, \ell\}$ with $i < j$; also the root spaces of these roots are as described.

To see that B is a base for Φ we note first that $\dim_F V = \ell$, and that the elements of B are evidently linearly independent; it follows that B is a basis for the F-vector space V. Since B is the disjoint union of P and $\{-\lambda : \lambda \in P\}$, to prove that B is a base for Φ it will now suffice to prove that every element of P is a linear combination of elements from B with non-negative integer coefficients. Let $i, j \in \{1, \dots, \ell\}$ with $i < j$. Then

$$\begin{aligned}
\alpha_i - \alpha_j &= \sum_{k=i}^{j-1} (\alpha_k - \alpha_{k+1}) \\
&= \sum_{k=i}^{j-1} \beta_k.
\end{aligned}$$

Also, we have

$$\begin{aligned}
\alpha_i + \alpha_j &= (\alpha_{\ell-1} + \alpha_\ell) + (\alpha_i - \alpha_{\ell-1}) + (\alpha_j - \alpha_\ell) \\
&= \beta_\ell + (\alpha_i - \alpha_{\ell-1}) + (\alpha_j - \alpha_\ell).
\end{aligned}$$

Since $\alpha_i - \alpha_{\ell-1}$ and $\alpha_j - \alpha_\ell$ are both linear combinations of elements from B with non-negative integer coefficients by what we have already proven, it follows that $\alpha_i + \alpha_j$ is linear combination of elements from B with non-negative integer coefficients. This completes the proof. ☐

h_{11}	$\beta_1 = \alpha_1 - \alpha_2$	$\alpha_1 - \alpha_3$	0	$\alpha_1 + \alpha_2$	$\alpha_1 + \alpha_3$
$\alpha_2 - \alpha_1$	h_{22}	$\beta_2 = \alpha_2 - \alpha_3$	$*$	0	$\beta_3 = \alpha_2 + \alpha_3$
$\alpha_3 - \alpha_1$	$\alpha_3 - \alpha_2$	h_{33}	$*$	$*$	0
0	$-(\alpha_1 + \alpha_2)$	$-(\alpha_1 + \alpha_3)$	$-h_{11}$	$*$	$*$
$*$	0	$-(\alpha_2 + \alpha_3)$	$*$	$-h_{22}$	$*$
$*$	$*$	0	$*$	$*$	$-h_{33}$

Figure 10.4: The decomposition of so$(6, F)$. For this example, $\ell = 3$. The positions are labeled with the corresponding root. Note that the diagonal is our chosen Cartan subalgebra. The positive roots with respect to our chosen base $\{\beta_1, \beta_2, \beta_3\}$ are boxed, while the colored roots form our chosen base. Positions labeled with $*$ are determined by other entries. The linear functionals α_1, α_2 and α_3 are defined in Proposition 10.10.4.

Lemma 10.10.5. *Assume that the characteristic of F is zero and F is algebraically closed. Let ℓ be a positive integer. The Killing form*

$$\kappa : \mathrm{so}(2\ell, F) \times \mathrm{so}(2\ell, F) \longrightarrow F$$

is given by

$$\kappa(h, h') = (2\ell - 2) \cdot \mathrm{tr}(hh')$$

for $h, h' \in H$. Here, H is the subalgebra of diagonal matrices in so$(2\ell, F)$*; H is a Cartan subalgebra of* so$(2\ell, F)$ *by Lemma 10.10.3.*

Proof. Let $h, h' \in \mathrm{gl}(\ell, F)$ be diagonal matrices. Then

$$\begin{aligned}
&\kappa(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \begin{bmatrix} h' & 0 \\ 0 & -h' \end{bmatrix}) \\
&= \mathrm{tr}(\mathrm{ad}(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}) \circ \mathrm{ad}(\begin{bmatrix} h' & 0 \\ 0 & -h' \end{bmatrix})) \\
&= \sum_{i,j \in \{1,\dots,\ell\}, i \neq j} (h_{ii} - h_{jj})(h'_{ii} - h'_{jj}) \\
&\quad + 2 \sum_{i,j \in \{1,\dots,\ell\}, i < j} (h_{ii} + h_{jj})(h'_{ii} + h'_{jj}) \\
&= \sum_{i,j \in \{1,\dots,\ell\}} (h_{ii} - h_{jj})(h'_{ii} - h'_{jj}) \\
&\quad + \sum_{i,j \in \{1,\dots,\ell\}, i \neq j} (h_{ii} + h_{jj})(h'_{ii} + h'_{jj})
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i,j\in\{1,\dots,\ell\}} (h_{ii}-h_{jj})(h'_{ii}-h'_{jj}) \\
&\qquad + \sum_{i,j\in\{1,\dots,\ell\}} (h_{ii}+h_{jj})(h'_{ii}+h'_{jj}) \\
&\qquad\qquad - \sum_{i,j\in\{1,\dots,\ell\},i=j} (h_{ii}+h_{jj})(h'_{ii}+h'_{jj}) \\
&= \sum_{i,j\in\{1,\dots,\ell\}} h_{ii}h'_{ii}-h_{ii}h'_{jj}-h_{jj}h'_{ii}+h_{jj}h'_{jj} \\
&\qquad + \sum_{i,j\in\{1,\dots,\ell\}} h_{ii}h'_{ii}+h_{ii}h'_{jj}+h_{jj}h'_{ii}+h_{jj}h'_{jj} \\
&\qquad\qquad -4\sum_{i\in\{1,\dots,\ell\}} h_{ii}h'_{ii} \\
&= 4\ell \sum_{i\in\{1,\dots,\ell\}} h_{ii}h'_{ii} - 4\sum_{i\in\{1,\dots,\ell\}} h_{ii}h'_{ii} \\
&= (4\ell-4)\sum_{i\in\{1,\dots,\ell\}} h_{ii}h'_{ii} \\
&= (2\ell-2)\cdot \mathrm{tr}(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}\begin{bmatrix} h' & 0 \\ 0 & -h' \end{bmatrix}).
\end{aligned}
$$

This completes the calculation. □

Lemma 10.10.6. *Let ℓ be an integer such that $\ell \geq 2$. Let the notation as in Lemma 10.10.4. Assume first that $\ell \geq 3$. For $i, j \in \{1, \dots, \ell\}$ we have:*

$$
(\beta_i,\beta_j) = \begin{cases} \dfrac{2}{4\ell-4} & \text{if } i=j, \\ \dfrac{-1}{4\ell-4} & \text{if } i,j\in\{1,\dots,\ell-1\} \text{ and } i \text{ and } j \text{ are consecutive}, \\ \dfrac{-1}{4\ell-4} & \text{if } \{i,j\}=\{\ell-2,\ell\}, \\ 0 & \text{if none of the above conditions hold.} \end{cases}
$$

Assume that $\ell = 2$. Then:

$$
\begin{aligned}
(\beta_1,\beta_1) &= \frac{1}{2}, \\
(\beta_2,\beta_2) &= \frac{1}{2}, \\
(\beta_1,\beta_2) &= 0.
\end{aligned}
$$

Proof. Let $h \in \mathrm{gl}(\ell, F)$ be a diagonal matrix. Let $i \in \{1, \dots, \ell\}$. Then

$$
\kappa(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}, \frac{1}{4\ell-4}\begin{bmatrix} e_{ii} & 0 \\ 0 & -e_{ii} \end{bmatrix}) = \frac{2\ell-2}{4\ell-4}\mathrm{tr}(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}\begin{bmatrix} e_{ii} & 0 \\ 0 & -e_{ii} \end{bmatrix})
$$

$$= h_{ii}$$
$$= \alpha_i(\begin{bmatrix} h & 0 \\ 0 & -h \end{bmatrix}).$$

Since this holds for all diagonal $h \in \mathrm{gl}(\ell, F)$, it follows that

$$t_{\alpha_i} = \frac{1}{4\ell - 4} \begin{bmatrix} e_{ii} & 0 \\ 0 & -e_{ii} \end{bmatrix}.$$

Also let $j \in \{1, \dots, \ell\}$. Then

$$\begin{aligned}
(\alpha_i, \alpha_j) &= \kappa(t_{\alpha_i}, t_{\alpha_j}) \\
&= (2\ell - 2) \cdot \mathrm{tr}(\frac{1}{4\ell - 4} \begin{bmatrix} e_{ii} & 0 \\ 0 & -e_{ii} \end{bmatrix} \cdot \frac{1}{4\ell - 4} \begin{bmatrix} e_{jj} & 0 \\ 0 & -e_{jj} \end{bmatrix}) \\
&= \begin{cases} \dfrac{1}{4\ell - 4} & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}
\end{aligned}$$

Let $i, j \in \{1, \dots, \ell - 1\}$. Then

$$\begin{aligned}
(\beta_i, \beta_j) &= (\alpha_i - \alpha_{i+1}, \alpha_j - \alpha_{j+1}) \\
&= (\alpha_i, \alpha_j) - (\alpha_i, \alpha_{j+1}) - (\alpha_{i+1}, \alpha_j) + (\alpha_{i+1}, \alpha_{j+1}) \\
&= \begin{cases} \dfrac{2}{4\ell - 4} & \text{if } i = j, \\ \dfrac{-1}{4\ell - 4} & \text{if } i \text{ and } j \text{ are consecutive}, \\ 0 & \text{if } i \neq j \text{ and } i \text{ and } j \text{ are not consecutive}. \end{cases}
\end{aligned}$$

Let $i \in \{1, \dots, \ell - 1\}$. Assume that $\ell \geq 3$. Then

$$\begin{aligned}
(\beta_i, \beta_\ell) &= (\alpha_i - \alpha_{i+1}, \alpha_{\ell-1} + \alpha_\ell) \\
&= (\alpha_i, \alpha_{\ell-1}) + (\alpha_i, \alpha_\ell) - (\alpha_{i+1}, \alpha_{\ell-1}) - (\alpha_{i+1}, \alpha_\ell) \\
&= \begin{cases} \dfrac{-1}{4\ell - 4} & \text{if } i = \ell - 2, \\ 0 & \text{if } i \neq \ell - 2. \end{cases}
\end{aligned}$$

Assume that $\ell = 2$. Then necessarily $i = 1$, and

$$\begin{aligned}
(\beta_i, \beta_\ell) &= (\beta_1, \beta_2) \\
&= (\alpha_1 - \alpha_2, \alpha_1 + \alpha_2) \\
&= (\alpha_1, \alpha_1) + (\alpha_1, \alpha_2) - (\alpha_2, \alpha_1) - (\alpha_2, \alpha_2) \\
&= 0.
\end{aligned}$$

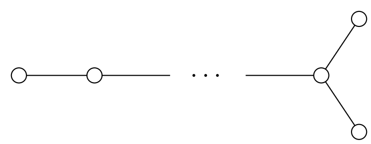
Finally,

$$\begin{aligned}
(\beta_\ell, \beta_\ell) &= (\alpha_{\ell-1} + \alpha_\ell, \alpha_{\ell-1} + \alpha_\ell) \\
&= (\alpha_{\ell-1}, \alpha_{\ell-1}) + (\alpha_{\ell-1}, \alpha_\ell) + (\alpha_\ell, \alpha_{\ell-1}) + (\alpha_\ell, \alpha_\ell) \\
&= \frac{2}{4\ell - 4}.
\end{aligned}$$

This completes the proof. □

Lemma 10.10.7. *Let ℓ be an integer such that $\ell \geq 3$. Let F have characteristic zero and be algebraically closed. The Dynkin diagram of* $\mathrm{so}(2\ell, F)$ *is*

D_ℓ:



and the Cartan matrix of $\mathrm{so}(2\ell, F)$ *is*

$$\begin{bmatrix}
2 & -1 & & & & & \\
-1 & 2 & -1 & & & & \\
 & -1 & 2 & -1 & & & \\
 & & \ddots & \ddots & \ddots & & \\
 & & & -1 & 2 & -1 & -1 \\
 & & & & -1 & 2 & \\
 & & & & -1 & & 2
\end{bmatrix}$$

The Lie algebra $\mathrm{so}(2\ell, F)$ *is simple.*

Proof. Let $i, j \in \{1, \ldots, \ell\}$ with $i \neq j$. Then

$$\begin{aligned}
\langle \beta_i, \beta_j \rangle &= 2\frac{(\beta_i, \beta_j)}{(\beta_j, \beta_j)} \\
&= \begin{cases}
-1 & \text{if } i, j \in \{1, \ldots, \ell - 1\} \text{ and } i \text{ and } j \text{ are consecutive,} \\
-1 & \text{if } \{i, j\} = \{\ell - 2, \ell\}, \\
0 & \text{if none of the above conditions hold.}
\end{cases}
\end{aligned}$$

Hence,

$$\begin{aligned}
\langle \beta_i, \beta_j \rangle \langle \beta_j, \beta_i \rangle &= 4\frac{(\beta_i, \beta_j)^2}{(\beta_i, \beta_i)(\beta_j, \beta_j)} \\
&= \begin{cases}
1 & \text{if } i, j \in \{1, \ldots, \ell - 1\} \text{ and } i \text{ and } j \text{ are consecutive,} \\
1 & \text{if } \{i, j\} = \{\ell - 2, \ell\}, \\
0 & \text{if none of the above conditions hold.}
\end{cases}
\end{aligned}$$

It follows that the Dynkin diagram of sp($2\ell, F$) is C_ℓ, and the Cartan matrix of sp($2\ell, F$) is as stated. Since C_ℓ is connected, sp($2\ell, F$) is simple by Lemma 9.3.2 and Proposition 10.3.2. □

Chapter 11

Representation theory

11.1 Weight spaces again

Let F be algebraically closed and have characteristic zero. Let L be a finite-dimensional, semi-simple Lie algebra over F. Let H be a Cartan subalgebra of L, let

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$$

be the root space decomposition of L with respect to L from Chapter 7. Let (ϕ, V) be a representation of L, so that V is an F-vector space, and $\phi : L \to \mathrm{gl}(V)$ is a homorphism of Lie algebras. If $\lambda : H \to F$ is a linear functional, then we define

$$V_\lambda = \{v \in V : \phi(h)v = \lambda(h)v, h \in H\}.$$

If $\lambda : H \to F$ is a linear functional and $V_\lambda \neq 0$, then we say that λ is a **weight** of H on V, and refer to V_λ as a **weight space**.

Lemma 11.1.1. *Let F be algebraically closed and have characteristic zero. Let L be a finite-dimensional, semi-simple Lie algebra over F. Let H be a Cartan subalgebra of L, let*

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$$

be the root space decomposition of L with respect to L from Chapter 7. Let (ϕ, V) be a representation of L. Let V' be the F-subspace of V generated by the subspaces V_λ for λ a weight of H on V.

1. *Let $\lambda : H \to F$ be a linear functional, and let $\alpha \in \Phi$. If $x \in L_\alpha$, then $\phi(x)V_\lambda \subset V_{\lambda+\alpha}$.*

2. *the F-subspace V' of V is an L-subspace.*

3. *The F-subspace V' of V is the direct sum of the V_λ for λ a weight of H on V, so that*

$$V' = \bigoplus_{\lambda \text{ is a weight of } H \text{ on } V} V_\lambda.$$

4. *If V is finite-dimensional, then $V' = V$.*

Proof. Proof of 1. Let $\lambda : H \to F$ be a linear functional, and let $\alpha \in \Phi$. Let $x \in L_\alpha$ and $v \in V_\lambda$. We have

$$\begin{aligned} \phi([h,x])v &= (\phi(h)\phi(x) - \phi(x)\phi(h))v \\ \phi(\alpha(h)x)v &= \phi(h)\big(\phi(x)v\big) - \phi(x)\big(\phi(h)v\big) \\ \alpha(h)\phi(x)v &= \phi(h)\big(\phi(x)v\big) - \lambda(h)\phi(x)v \\ \phi(h)\big(\phi(x)v\big) &= \big(\lambda(h) + \alpha(h)\big)\phi(x)v. \end{aligned}$$

This implies that $\phi(x)v \in V_{\lambda+\alpha}$.

Proof of 2. Clearly, the operators $\phi(h)$ for $h \in H$ preserve the subspace V'. By 1, the operators $\phi(x)$ for $x \in L_\alpha$, $\alpha \in \Phi$ also preserve V'. Since $L = H \oplus \oplus_{\alpha\in\Phi} L_\alpha$, it follows that L preserves V'.

Proof of 3. Assume that V' is not the direct sum of the subspaces V_λ for $\lambda \in H^\vee$; we will obtain a contradiction. By our assumption, there exist an integer $t \geq 2$ and distinct $\lambda_1, \ldots, \lambda_t \in H^\vee$ such that $V_{\lambda_1} \cap (V_{\lambda_2} + \cdots + V_{\lambda_t}) \neq 0$. We may assume that t is the smallest integer with these properties. Let $v_1 \in V_{\lambda_1} \cap (V_{\lambda_2} + \cdots + V_{\lambda_t})$ be non-zero. Write

$$v_1 = v_2 + \cdots + v_t$$

where $v_i \in V_{\lambda_i}$ for $i \in \{2, \ldots, t\}$. The minimality of t implies that v_i is non-zero for $i \in \{2, \ldots, t\}$. Let $h \in H$. Then

$$\begin{aligned} \phi(h)v_1 &= \phi(h)(v_2 + \cdots + v_t) \\ \lambda_1(h)v_1 &= \lambda_2(h)v_2 + \cdots + \lambda_t(h)v_t, \end{aligned}$$

and, after multiplying $v_1 = v_2 + \cdots + v_t$ by $\lambda_1(h)$,

$$\lambda_1(h)v_1 = \lambda_1(h)v_2 + \cdots + \lambda_1(h)v_t.$$

Subtracting, we obtain:

$$0 = (\lambda_1(h) - \lambda_2(h))v_2 + \cdots + (\lambda_1(h) - \lambda_t(h))v_t.$$

The minimality of t implies that $\lambda_1(h) - \lambda_i(h) = 0$ for all $h \in H$ and $i \in \{2, \ldots, t\}$, i.e, $\lambda_1 = \cdots = \lambda_t$. This is a contradiction.

Proof of 4. Assume that V is finite-dimensional; we need to prove that $V \subset V'$. The operators $\phi(h) \in \mathrm{gl}(V)$ for $h \in H$ are diagonalizable by Theorem 6.3.4 and the definition of a Cartan subalgebra. Since H is abelian, the operators $\phi(h)$ for $h \in H$ mutually commute. It follows that (see Theorem 8 from Section 6.5 of

[5]) that there exists a basis $v_1, \dots, v_n$ for V such that each v_i for $i \in \{1, \dots, n\}$ is an eigenvector for every operator $\phi(h)$ for $h \in H$. Let $i \in \{1, \dots, n\}$. For $h \in H$, let $\lambda(h) \in F$ be such that $\phi(h)v_i = \lambda(h)v_i$. Since the map $H \to \mathrm{gl}(V)$ given by $h \mapsto \phi(h)$ is linear, and v_i is non-zero, the function $\lambda : H \to F$ is also linear. It follows that λ is a weight of H on V and that $v_i \in V_\lambda$. We conclude that $V \subset V'$. □

11.2 Borel subalgebras

Lemma 11.2.1. *Let F be algebraically closed and have characteristic zero. Let L be a finite-dimensional, semi-simple Lie algebra over F. Let H be a Cartan subalgebra of L, let*

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$$

be the root space decomposition of L with respect to L from Chapter 7, and let B be a base for Φ. Let Φ^+ be the positive roots in Φ with respect to B. Define

$$N = \sum_{\alpha \in \Phi^+} L_\alpha$$

and

$$P = H + N = H + \sum_{\alpha \in \Phi^+} L_\alpha.$$

Then N and P are subalgebras of L. Moreover,

$$[P, P] = N,$$

N is nilpotent, and P is solvable.

Proof. Let $\alpha, \beta \in \Phi^+$; we will first prove that $[L_\alpha, L_\beta] \subset N$ and that $[H, L_\alpha] \subset L_\alpha$. Since α and β are both positive roots we must have $\alpha + \beta \neq 0$. By Proposition 7.0.3 we have $[L_\alpha, L_\beta] \subset L_{\alpha+\beta}$. If $\alpha + \beta$ is not a root, then, as $\alpha + \beta \neq 0$, we must have $L_{\alpha+\beta} = 0$ (by definition), so that $[L_\alpha, L_\beta] \subset L_{\alpha+\beta} = 0 \subset N$. Assume that $\alpha + \beta$ is a root. Then $\alpha + \beta$ is a positive root because α and β are positive. It follows that $[L_\alpha, L_\beta] \subset L_{\alpha+\beta} \subset N$. The definition of L_α implies that $[H, L_\alpha] \subset L_\alpha$.

Since $[H, H] = 0$, the previous paragraph implies that N and P are subalgebras of L, and also that $[P, P] \subset N$. To prove that $N \subset [P, P]$ it suffices to prove that $L_\alpha \subset [P, P]$ if α is a positive root. Let α be a positive root. Let $x \in L_\alpha$. Let $h \in H$ be such that $\alpha(h) \neq 0$. We have $[h, x] = \alpha(h)x$. Since $[h, x] \in [P, P]$, it follows that $\alpha(h)x \in [P, P]$. Since $\alpha(h) \neq 0$, we get $x \in [P, P]$. It follows now that $[P, P] = N$.

To see that N is nilpotent, we note that by Proposition 7.0.3, for k a positive integer:

$$N^1 = [N, N] \subset \sum_{\alpha_1, \alpha_2 \in \Phi^+} L_{\alpha_1 + \alpha_2},$$

$$N^2 = [N, N^1] \subset \sum_{\alpha_1,\alpha_2,\alpha_3 \in \Phi^+} L_{\alpha_1+\alpha_2+\alpha_3},$$

$$\dots$$

$$N^{k+1} = [N, N^k] \subset \sum_{\alpha_1,\dots,\alpha_k \in \Phi^+} L_{\alpha_1+\dots+\alpha_k}.$$

For k a positive integer, define

$$S_k = \{\alpha_1 + \dots + \alpha_k : \alpha_1, \dots, \alpha_k \in \Phi^+\}.$$

Evidently, the sets S_k for k a positive integer do not contain the zero linear functional. Recall the height function from page 93. Let $m = \max(\{\mathrm{ht}(\beta) : \beta \in \Phi^+\})$. Since $\mathrm{ht}(\lambda) \geq k$ for all $\lambda \in S_k$, the set S_k for $k \geq m+1$ cannot contain any elements of Φ^+. Also, it is clear that S_k does not contain any elements of the set Φ^- of negative roots (by the basic properties of the base B). Thus, if $k \geq m+1$, then $L_\lambda = 0$ for all $\lambda \in S_k$. It follows that $N^{m+2} = 0$ so that N is nilpotent.

Finally, P is solvable because $[P, P] = N$ and N is nilpotent. □

We refer to P as in Lemma 11.2.1 as a **Borel subalgebra**.

11.3 Maximal vectors

Let F be algebraically closed and have characteristic zero. Let L be a finite-dimensional, semi-simple Lie algebra over F. Let H be a Cartan subalgebra of L, let

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$$

be the root space decomposition of L with respect to L from Chapter 7, and let B be a base for Φ. Let Φ^+ be the positive roots in Φ with respect to B. Define $N = \sum_{\alpha \in \Phi^+} L_\alpha$ as in Lemma 11.2.1. Let (ϕ, V) be a representation of L. Let $v \in V$. We say that v **generates** V if the vectors

$$\phi(x_1) \cdots \phi(x_t) v,$$

for t a positive integer and $x_1, \dots, x_t \in L$, span the F-vector space V. Assume that λ is a weight of H on V, and let $v \in V_\lambda$ be non-zero. We say that v is a **maximal vector of weight** λ if $\phi(x)v = 0$ for all $x \in N$.

Lemma 11.3.1. *Let F be algebraically closed and have characteristic zero. Let L be a finite-dimensional, semi-simple Lie algebra over F. Let H be a Cartan subalgebra of L, let Φ be the roots of L with respect to H, and let B be a base for Φ. Define N and the Borel subalgebra P as as in Lemma 11.2.1. Let (ϕ, V) be a representation of L. If V is finite-dimensional, then V has a maximal vector of weight λ for some weight λ of H on V.*

Proof. Let P be the Borel subalgebra of L defined with respect to our chosen base. By Lemma 11.2.1, P is solvable. Consider $\phi(P) \subset \mathrm{gl}(V)$. Since ϕ is a map of Lie algebras, $\phi(P)$ is a Lie subalgebra of $\mathrm{gl}(V)$. By Lemma 2.1.5, $\phi(P)$ is solvable. By Lemma 3.4.1, a version of Lie's Theorem, there exists a non-zero vector v of V such that v is a common eigenvector for the operators $\phi(x) \in \mathrm{gl}(V)$, $x \in P$. For $x \in P$, let $c(x) \in F$ be such that $\phi(x)v = c(x)v$. It is easy to see that the function $c : P \to F$ is F-linear. We claim that $c(N) = 0$. Let $x, y \in P$. Then

$$\begin{aligned} \phi([x,y])v &= c([x,y])v \\ (\phi(x)\phi(y) - \phi(y)\phi(x))v &= c([x,y])v \\ \phi(x)\phi(y)v - \phi(y)\phi(x)v &= c([x,y])v \\ c(x)c(y)v - c(y)c(x)v &= c([x,y])v \\ 0 &= c([x,y])v. \end{aligned}$$

Since v is non-zero, we see that $c([x,y]) = 0$. Since, by Lemma 11.2.1, $N = [P,P]$, we get that $c(N) = 0$. Define $\lambda : H \to F$ by $\lambda(h) = c(h)$ for $h \in H$. Evidently, v is in the weight space V_λ. Since $c(N) = 0$ we also have $\phi(x)v = 0$ for $x \in N$. It follows that v is a maximal vector for the weight λ of H on V. □

Theorem 11.3.2. *Let F be algebraically closed and have characteristic zero. Let L be a finite-dimensional, semi-simple Lie algebra over F. Let H be a Cartan subalgebra of L, let Φ be the roots of L with respect to H, and let $B = \{\alpha_1, \ldots, \alpha_n\}$ be a base for Φ. Define N and the Borel subalgebra P as as in Lemma 11.2.1. Let (ϕ, V) be a representation of L. Let $v \in V$. Assume that v generates V, and that v is a maximal vector of weight λ. Then*

$$V = \bigoplus_{\mu \text{ is a weight of } H \text{ on } V} V_\mu.$$

Moreover, if μ is a weight of H on V, then

$$\mu = \lambda - (c_1\alpha_1 + \cdots + c_n\alpha_n)$$

for some non-negative integers $c_1, \ldots, c_n$. Thus, if μ is a weight of H on V, then $\mu \prec \lambda$. Here, $\prec$ is the partial order from page 116. For every weight μ of H on V the subspace V_μ is finite-dimensional, and the subspace V_λ is one-dimensional.

Proof. For each $\beta \in \Phi^-$, fix a non-zero element y_β in the one-dimensional space L_β. We first claim that the vector space V is spanned by v and the vectors

$$w = \phi(y_{\beta_1}) \cdots \phi(y_{\beta_k})v$$

for k a positive integer and $\beta_1, \ldots, \beta_k \in \Phi^-$. To see this, we recall that, as a vector space, L is spanned by H, L_α for $\alpha \in \Phi^+$ and $\beta \in \Phi^-$, and that v generates V. This implies that the vector space V is spanned by v and the vectors of the form

$$\phi(z_1) \cdots \phi(z_\ell)v$$

for ℓ a positive integer, and, for $i \in \{1, \ldots, \ell\}$, the element z_i is in H, or in L_α for some $\alpha \in \Phi^+$, or in L_β for some $\beta \in \Phi^-$. Since $N = \oplus_{\alpha \in \Phi^+} L_\alpha$ acts by zero on v (as v is a maximal vector), and since $\phi(h)v = \lambda(h)v$ for $h \in H$, our claim follows.

Next, let $w = \phi(y_{\beta_1}) \cdots \phi(y_{\beta_k})v$ be a vector as above with k a positive integer. By 1 of Lemma 11.1.1, w is contained in $V_{\lambda+\beta_1+\cdots+\beta_k}$. Let M be the set of linear functionals $\mu : H \to F$ such that $\mu = \lambda$, or there exists a positive integer k and $\beta_1, \ldots, \beta_k \in \Phi^-$ such that $\mu = \lambda + \beta_1 + \cdots \beta_k$ and $V_\mu \neq 0$. The result of the previous paragraph imply that the subspaces V_μ for $\mu \in M$ span V. By 3 of Lemma 11.1.1, the span of the subspaces V_μ for $\mu \in M$ is direct, i.e., V is the direct sum of the subspaces V_μ for $\mu \in M$. Let $\nu : H \to F$ be any weight of H on V. Let $u \in V_\nu$ be non-zero. There exist unique elements $\mu_1, \ldots, \mu_t \in M$ and non-zero $v_1 \in V_{\mu_1}, \ldots, v_t \in V_{\mu_t}$ such that $u = v_1 + \cdots + v_t$. Let $h \in H$. Then

$$\begin{aligned} \phi(h)u &= \phi(h)v_1 + \cdots + \phi(h)v_t \\ \nu(h)u &= \mu_1(h)v_1 + \cdots + \mu_t(h)v_t \\ \nu(h)(v_1 + \cdots + v_t) &= \mu_1(h)v_1 + \cdots + \mu_t(h)v_t \\ \nu(h)v_1 + \cdots + \nu(h)v_t &= \mu_1(h)v_1 + \cdots + \mu_t(h)v_t. \end{aligned}$$

Since this equality holds for all $h \in H$, and the sum of $V_{\mu_1}, \ldots, V_{\mu_t}$ is direct, we must have $\nu = \mu_1 = \cdots = \mu_t$. Since $\mu_1, \ldots, \mu_t$ are mutually distinct, we obtain $t = 1$ and $\nu = \mu_1$. Recalling the definition of the set M, and the fact that every element of Φ^- can be uniquely written as a linear combination of the elements of $B = \{\alpha_1, \ldots, \alpha_n\}$ with non-positive integral coefficients, we see that ν has the form as stated in the theorem.

Finally, let μ be a weight of H on V. Let $u \in V_\mu$ be non-zero. By the first paragraph, w can be written as linear combination of v and elements of the form $w = \phi(y_{\beta_1}) \cdots \phi(y_{\beta_k})v$. Hence, there exists a positive integer ℓ, elements $c_0, c_1, \ldots, c_\ell$ of F, and for each $i \in \{1, \ldots, \ell\}$ a positive integer k_i and $\beta_{i,1}, \ldots, \beta_{i,k_i} \in \Phi^-$ such that

$$u = c_0 v + \sum_{i=1}^{\ell} c_i \phi(y_{\beta_{i,1}}) \cdots \phi(y_{\beta_{i,k_i}})v.$$

Since $\phi(y_{\beta_{i,1}}) \cdots \phi(y_{\beta_{i,k_i}})v$ is contained in $V_{\lambda+\beta_{i,1}+\cdots+\beta_{i,k_i}}$, and since the sum of weight spaces is direct by 3 of Lemma 11.1.1, we see that for each $i \in \{1, \ldots, \ell\}$, if

$$c_i \phi(y_{\beta_{i,1}}) \cdots \phi(y_{\beta_{i,k_i}})v$$

is non-zero, then

$$\mu = \lambda + \beta_{i,1} + \cdots + \beta_{i,k_i},$$

or equivalently,

$$\mu - \lambda = \beta_{i,1} + \cdots + \beta_{i,k_i}.$$

It follows that the dimension of V_μ is bounded by N, where N is 1 plus the number of m-tuples $(\beta_1, \dots, \beta_m)$, where m is a positive integer and $\beta_1, \dots, \beta_m \in \Phi^-$, such that

$$\mu - \lambda = \beta_1 + \cdots + \beta_m.$$

If $\mu = \lambda$, then $N = 1$, so that $\dim V_\lambda = 1$. Assume $\mu \neq \lambda$. Recall the height function ht from page 93. If m is a positive integer and $\beta_1, \dots, \beta_m \in \Phi^-$ are such that $\mu - \lambda = \beta_1 + \cdots + \beta_m$, then

$$\mathrm{ht}(\mu - \lambda) = \mathrm{ht}(\beta_1 + \cdots + \beta_m) \leq -m,$$

or equivalently, $-\mathrm{ht}(\mu - \lambda) \geq m$. Since Φ^- is finite, it follows that N is finite, as desired. □

Let the notation be is as in Theorem 11.3.2. We will say that λ is the **highest weight** for V. By Theorem 11.3.2, if μ is a weight of H on V, then $\lambda \succ \mu$. In particular, if λ' is a weight of H on V, and $\lambda' \succ \mu$ for all weights of H on V, then $\lambda' = \lambda$; this fact justifies the uniqueness part of the terminology "*the* highest weight".

Corollary 11.3.3. *Let F be algebraically closed and have characteristic zero. Let L be a finite-dimensional, semi-simple Lie algebra over F. Let H be a Cartan subalgebra of L. Let (ϕ, V) be a representation of L. Assume that V is irreducible. If $v_1 \in V$ and $v_2 \in V$ are maximal vectors of weights λ_1 and λ_2 of H on V, respectively, then $\lambda_1 = \lambda_2$, and there exists $c \in F^\times$ such that $v_2 = cv_1$.*

Proof. Since V is irreducible, the vectors v_1 and v_2 both generate V. By Theorem 11.3.2 we have $\lambda_1 = \lambda_2$. Therefore, $V_{\lambda_1} = V_{\lambda_2}$. Again by Theorem 11.3.2, $\dim V_{\lambda_1} = \dim V_{\lambda_2} = 1$. This implies that v_2 is an $F^\times$ multiple of v_1. □

Corollary 11.3.4. *Let the notation and objects be as in Theorem 11.3.2. If W is an L-subspace of V, then*

$$W = \bigoplus_{\mu \text{ is a weight of } H \text{ on } W} W_\mu.$$

The L-representation V is indecomposable, and has a unique maximal proper L-subspace U. The quotient V/U is irreducible, and if W is any L-subspace of V such that V/W is non-zero and irreducible, then $W = U$.

Proof. Let W be an L-subspace of V; we will first prove that W is the direct sum of its weight spaces. By Theorem 11.3.2, if $w \in W$ and is non-zero, then w has a unique expression as

$$w = w_{\mu_1} + \cdots + w_{\mu_k},$$

where $\mu_1, \dots, \mu_k$ are distinct weights of H on V, and w_{μ_i} is a non-zero element of V_{μ_i} for $i \in \{1, \dots, k\}$; we need to prove that in fact w_{μ_i} is contained in W_{μ_i} for $i \in \{1, \dots, k\}$. If w is a non-zero element of W and $w_{\mu_i} \notin W$ for some

$i \in \{1, \ldots, k\}$, then we will say that w has property P. Suppose that there exists a non-zero $w \in W$ which has property P; we will obtain a contradiction. We may assume that k is minimal. Since k is minimal, we must have $k > 1$: otherwise, $w = w_{\mu_1} \in W \cap V_{\mu_1} = W_{\mu_1}$, a contradiction. Also, we claim that $w_{\mu_i} \notin W$ for $i \in \{1, \ldots, k\}$. To see this, let $X = \{i \in \{1, \ldots, k\} : w_{\mu_i} \in W\}$, and assume that X is non-empty. Since w has property P, the set X is a proper subset of $\{1, \ldots, k\}$. We have

$$w - \sum_{i \in X} w_{\mu_i} = \sum_{j \in \{1, \ldots, k\} - X} w_{\mu_j}.$$

This vector is contained W and has property P; since k is minimal, this is a contradiction. This proves our claim. Next, since μ_1 and μ_2 are distinct, there exists $h \in H$ such that $\mu_1(h) \neq \mu_2(h)$. Now

$$\begin{aligned} w &= w_{\mu_1} + w_{\mu_2} + \cdots + w_{\mu_k} \\ \phi(h)w &= \phi(h)w_{\mu_1} + \phi(h)w_{\mu_2} + \cdots + \phi(h)w_{\mu_k} \\ \phi(h)w &= \mu_1(h)w_{\mu_1} + \mu_2(h)w_{\mu_2} + \cdots + \mu_k(h)w_{\mu_k}. \end{aligned}$$

Also, we have

$$\mu_2(h)w = \mu_2(h)w_{\mu_1} + \mu_2(h)w_{\mu_2} + \cdots + \mu_2(h)w_{\mu_k}.$$

Subtracting yields:

$$\begin{aligned} &\phi(h)w - \mu_2(h)w \\ &\quad = (\mu_1(h) - \mu_2(h))w_{\mu_1} + (\mu_3(h) - \mu_2(h))w_{\mu_3} + \cdots + (\mu_k(h) - \mu_2(h))w_{\mu_k}. \end{aligned}$$

Since W is an L-subspace, this vector is contained in W. Also, $(\mu_1(h) - \mu_2(h))w_{\mu_1} \notin W$. It follows that this vector has property P. This contradicts the minimality of k. Hence, W is the direct sum of its weight spaces, as desired.

To see that V is indecomposable, assume that there exists L-subspaces W_1 and W_2 of W and $V = W_1 \oplus W_2$; we need to prove that $W_1 = V$ or $W_2 = V$. Write $v = w_1 + w_2$ with $w_1 \in W_1$ and $w_2 \in W_2$. By the last paragraph,

$$\begin{aligned} w_1 &= w_{1,\mu_1} + \cdots + w_{1,\mu_k}, \\ w_2 &= w_{2,\nu_1} + \cdots + w_{2,\nu_\ell} \end{aligned}$$

where $\mu_1, \ldots, \mu_k$ are distinct weights of H on W_1, $\nu_1, \ldots, \nu_\ell$ are distinct weights of H on W_2, and $w_{1,\mu_i} \in W_{1,\mu_i}$ and $w_{2,\nu_j} \in W_{2,\nu_j}$ are non-zero for $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, \ell\}$. We have

$$v = w_1 + w_2 = w_{1,\mu_1} + \cdots + w_{1,\mu_k} + w_{2,\nu_1} + \cdots + w_{2,\nu_\ell}.$$

Now v is a vector of weight λ. Since the weight space decomposition is direct, one of $\mu_1, \ldots, \mu_k, \nu_1, \ldots, \nu_\ell$ is λ. Since V_λ is one-dimensional and spanned by v, this implies that $v \in W_1$ or $v \in W_2$. Therefore, $W_1 = V$ or $W_2 = V$.

Let U be the F-subspace spanned by all the proper L-subspaces of V. Clearly, U is an L-subspace. We claim that U is proper. To prove this it suffices to prove that $v \notin U$. Assume $v \in U$; we will obtain a contradiction. Since $v \in U$, there exists proper L-subspaces $U_1, \dots, U_t$ of V and vectors $w_1 \in U_1, \dots, w_t \in U_t$ such that $v = w_1 + \dots + w_t$. An argument as in the last paragraph now implies that for some $i \in \{1, \dots, t\}$ we have $v \in U_i$. This implies that $U_i = V$, contradicting that U_i is a proper subspace. The construction of U implies that U is maximal among proper L-subspaces of V, and that U is the unique proper maximal L-subspace of V.

To see that V/U is irreducible, assume that Q is an L-subspace of V/U. Let $W = \{w \in V : w + U \in Q\}$. Evidently, W is an L-subspace of V. If $W = V$, then $Q = V/U$. If W is a proper subspace of V, then by the definition of U, $W \subset U$, so that $Q = 0$. Thus, V/U is irreducible.

Finally, W be any L-subspace of V such that V/W is non-zero and irreducible. Since V/W is non-zero, W is a proper subspace of V. By the definition of U we get $W \subset U$. Now U/W is an L-subspace of V/W. Since V/W is irreducible, we have $U/W = 0$ or $U/W = V/W$. If $U/W = 0$, then $W = U$, as desired. If $U/W = V/W$, then $V = U$, a contradiction. Thus, $W = U$. □

Corollary 11.3.5. *Let F be algebraically closed and have characteristic zero. Let L be a finite-dimensional, semi-simple Lie algebra over F. Let H be a Cartan subalgebra of L. Let (ϕ_1, V_1) and (ϕ_2, V_2) be irreducible representations of L. Assume that V_1 and V_2 are generated by the maximal vectors $v_1 \in V_1$ and $v_2 \in V_2$ of weights λ_1 and λ_2, respectively. If $\lambda_1 = \lambda_2$, then V_1 and V_2 are isomorphic.*

Proof. Assume that $\lambda_1 = \lambda_2$. Let $\lambda = \lambda_1 = \lambda_2$. Let $V = V_1 \oplus V_2$. The F-vector space V is a representation of L with action ϕ defined by $\phi(x)(v_1 \oplus v_2) = \phi_1(x)w_1 \oplus \phi_2(x)w_2$ for $w_1 \in V_1$, $w_2 \in V_2$ and $x \in L$. Let $v = v_1 \oplus v_2$, and let V be the L-subspace of $V_1 \oplus V_2$ generated by v. The vector v is a maximal vector of V of weight λ. Let $p_1 : V \to V_1$ and $p_2 : V \to V_2$ be the projection maps. The maps p_1 and p_2 are L-maps. Since $p_1(v) = v_1$ and $p_2(v) = v_2$, and since V_1 and V_2 are generated by v_1 and v_2, respectively, it follows that p_1 and p_2 are surjective. Therefore, $V/\ker(p_1) \cong V_1$ and $V/\ker(p_2) \cong V_2$; since V_1 and V_2 are irreducible by assumption, the L-spaces $V/\ker(p_1)$ and $V/\ker(p_2)$ are irreducible. By Corollary 11.3.4, we have $\ker(p_1) = \ker(p_2)$, so that $V_1 \cong V/\ker(p_1) = V/\ker(p_2) \cong V_2$. □

11.4 The Poincaré-Birkhoff-Witt Theorem

Let F be a field, and let L be a Lie algebra over F. Let T be the tensor algebra of the F-vector space L. We have

$$T = T^0 \oplus T^1 \oplus T^2 \oplus \cdots$$

where $T^0 = F = F \cdot 1$, $T^1 = L$, $T^2 = L \otimes L$, and if k is a positive integer, then

$$T^k = \underbrace{L \otimes \cdots \otimes L}_{k}.$$

With tensor multiplication, T is an associative algebra with identity 1.

Let J be the two-sided ideal of T generated by all the elements of the form

$$x \otimes y - y \otimes x - [x, y]$$

for $x, y \in L$. We define

$$U(L) = T/J,$$

and refer to $U(L)$ as the **universal enveloping algebra** of L. We let

$$T \xrightarrow{p} T/L = U(L)$$

be the natural projection map. Evidently, $U(L)$ is an associative algebra over F. If $u, v \in U(L)$, then we will write the product of u and v as uv. We will write $p(1) = 1$. The element $1 \in U(L)$ is an identity for $U(L)$. We have

$$p(T^0) = p(F \cdot 1) = F \cdot 1 \subset U(L).$$

Let

$$T_+ = T^1 \oplus T^2 \oplus T^3 \oplus \cdots .$$

Then

$$T = T^0 \oplus T_+ = F \cdot 1 \oplus T_+.$$

Evidently, T_+ is a two-sided ideal of T. Since $x \otimes y - y \otimes x - [x, y] \in T_+$ for $x, y \in L$, it follows that

$$J \subset T_+.$$

We claim that

$$p(T^0) \cap p(T_+) = 0.$$

To see this, let $a \in F$ and $z \in T_+$ be such that $p(a \cdot 1) = p(z)$. Then $p(a \cdot 1 - z) = 0$. This means that $a \cdot 1 - z \in J$. Since $J \subset T_+$, we get $a \cdot 1 - z \in T_+$, and therefore $a \cdot 1 \in T_+$. As $T^0 \cap T_+ = 0$, this yields $a = 0$, as desired. Letting

$$U_+ = p(T_+),$$

we obtain the direct sum decomposition

$$U(L) = F \cdot 1 \oplus U_+.$$

If $u \in U(L)$, then the component of u in $F \cdot 1$ is called the **constant term** of u.

Let

$$\sigma : L \longrightarrow U(L)$$

be the composition

$$L \longrightarrow T \xrightarrow{p} U(L)$$

where the first map is the inclusion map, and the second map is the projection map p. We refer to σ as the **canonical map** of L into $U(L)$. Let $x, y \in L$. Then

$$\begin{aligned}\sigma(x)\sigma(y) - \sigma(y)\sigma(x) &= (x+J)(y+J) - (y+J)(x+J)\\ &= (x \otimes y + J) - (y \otimes x + J)\\ &= x \otimes y - y \otimes x + J\\ &= [x,y] + x \otimes y - y \otimes x - [x,y] + J\\ &= [x,y] + J\\ &= \sigma([x,y]).\end{aligned}$$

That is,

$$\sigma(x)\sigma(y) - \sigma(y)\sigma(x) = \sigma([x,y])$$

for $x, y \in L$.

Lemma 11.4.1. *Let F be a field, and let L be a Lie algebra over F. Let $\sigma : L \to U(L)$ be the canonical map. Let A be an associative algebra with identity, and assume that*

$$L \xrightarrow{\tau} A$$

is a linear map such that

$$\tau(x)\tau(y) - \tau(y)\tau(x) = \tau([x,y])$$

for $x, y \in L$. There exists a unique F-algebra homomorphism

$$U(L) \xrightarrow{\tau'} A$$

such that $\tau'(1) = 1$ and $\tau' \circ \sigma = \tau$, so that

$$\begin{array}{ccc} L & & \\ \sigma \big\downarrow & \overset{\tau}{\searrow} & \\ U(L) & \xrightarrow{\tau'} & A \end{array}$$

commutes.

Proof. To prove the existence of τ', we note first that by the universal property of T, there exists an algebra homomorphism

$$T \xrightarrow{\varphi} A$$

such that $\varphi(1) = 1$, and $\varphi(x) = \tau(x)$ for $x \in L$. Let $x, y \in L$. Then

$$\begin{aligned}\varphi(x \otimes y - y \otimes x - [x,y]) &= \varphi(x \otimes y) - \varphi(y \otimes x) - \varphi([x,y])\\ &= \varphi(x)\varphi(y) - \varphi(y)\varphi(x) - \tau([x,y])\\ &= \tau(x)\tau(y) - \tau(y)\tau(x) - \tau([x,y])\end{aligned}$$

$$= 0.$$

Since φ is an algebra homomorphism, and since φ is zero on the generators of J, it follows that $\varphi(J) = 0$. Therefore, there exists an algebra homorphism $\tau' : U(L) = T/J \to A$ such that $\tau'(x + J) = \varphi(x)$ for $x \in T$. Evidently, since $\varphi(1) = 1$, we have $\tau'(1) = 1$. Also, if $x \in L$, then

$$\begin{aligned}(\tau' \circ \sigma)(x) &= \tau'(\sigma(x)) \\ &= \tau'(x + J) \\ &= \varphi(x) \\ &= \tau(x).\end{aligned}$$

This proves the existence of τ'. The uniqueness of τ' follows from the fact that $U(L)$ is generated by 1 and $\sigma(L)$, the assumption that τ' is determined on these elements. □

We will consider sequences $(i_1, \ldots, i_p)$ where p is as positive integer, $i_1, \ldots, i_p$ are positive integers, and

$$i_1 \leq \cdots \leq i_p.$$

We let X be the set consisting of all such sequences, along with the empty set $\emptyset$. Let $I \in X$. If $I \neq \emptyset$, so that there exists a positive integer p, and $i_1, \ldots, i_p \in \mathbb{Z}_{>0}$ such that $I = (i_1, \ldots, i_p)$ with $i_1 \leq \cdots \leq i_p$, then we define

$$d(I) = p.$$

If $I = \emptyset$, then we define

$$d(\emptyset) = 0.$$

Let F be a field, and let L be a Lie algebra over F. Assume that L is finite-dimensional and non-zero. We fix an ordered basis

$$x_1, x_2, x_3, \ldots, x_n$$

for L as a vector space over F. We define the images of these vectors in $U(L)$ as

$$y_1 = \sigma(x_1), y_2 = \sigma(x_2), y_3 = \sigma(x_3), \ldots, y_n = \sigma(x_n).$$

Let $I \in X$. If $I \neq \emptyset$, so that there exists a positive integer p, and $i_1, \ldots, i_p \in \mathbb{Z}_{>0}$ such that $I = (i_1, \ldots, i_p)$ with $i_1 \leq \cdots \leq i_p$, then we define

$$y_I = y_{i_1} y_{i_2} y_{i_3} \cdots y_{i_p} \in U(L)$$

If $I = \emptyset$, then we define

$$y_\emptyset = 1 \in U(L).$$

Lemma 11.4.2. *Let F be a field, and let L be a finite-dimensional Lie algebra over F. Fix an ordered basis $x_1, \ldots, x_n$ for L, and define y_I for $I \in X$ as above. Then the elements y_I for $I \in X$ span $U(L)$ as a vector space over F.*

For k a non-negative integer, let X_k be the subset of $I \in X$ such that $d(I) \leq k$.

Let n be a positive integer, and let $z_1, \ldots, z_n$ be indeterminants. We define

$$P = F[z_1, \ldots, z_n].$$

Let $I \in X$. If $I \neq \emptyset$, so that there exists a positive integer p, and $i_1, \ldots, i_p \in \mathbb{Z}_{>0}$ such that $I = (i_1, \ldots, i_p)$ with $i_1 \leq \cdots \leq i_p$, then we define

$$z_I = z_{i_1} z_{i_2} z_{i_3} \cdots z_{i_p} \in F[z_1, \ldots, z_n].$$

If $I = \emptyset$, then we define

$$z_\emptyset = 1 \in F[z_1, \ldots, z_n].$$

Evidently, the elements z_I for $I \in X$ form a basis for $F[z_1, \ldots, z_n]$. For convenience, we define

$$P = F[z_1, \ldots, z_n].$$

Also, if k is a non-negative integer, then we let P_k be the F-subspace of P of polynomials of degree less than or equal to k. Evidently, if k is a non-negative integer, then P_k has as basis the elements z_I for $I \in X_k$.

Let $I \in X$. Let $i \in \{1, \ldots, n\}$. We say that $i \leq I$ if and only if $I = \emptyset$, or, if $I \neq \emptyset$, so that $I = (i_1, \ldots, i_p)$ for some positive integers p and $i_1, \ldots, i_p$ with $i_1 \leq \cdots \leq i_p$, then $i \leq i_1$.

Lemma 11.4.3. *Let L be a finite-dimensional Lie algebra over F, and let $x_1, \ldots, x_n$ be a basis for the F vector space L. Let the notation be as in the discussion preceding the lemma. For every non-negative integer p, there exists a unique linear map*

$$f_p : L \otimes P_p \longrightarrow P$$

such that:

($\mathrm{A_p}$) *If $i \in \{1, \ldots, n\}$ and $I \in X_p$ with $i \leq I$, then*

$$f_p(x_i \otimes z_I) = z_i z_I.$$

($\mathrm{B_p}$) *If $i \in \{1, \ldots, n\}$, q is a non-negative integer such that $q \leq p$, and $I \in X_q$, then*

$$f_p(x_i \otimes z_I) - z_i z_I \in P_q.$$

In particular, $f_p(L \otimes P_q) \subset P_{q+1}$ for non-negative integers q with $q \leq p$.

($\mathrm{C_p}$) *If $p \geq 1$, $i, j \in \{1, \ldots, n\}$ and $J \in X_{p-1}$, then*

$$f_p(x_i \otimes f_p(x_j \otimes z_J)) = f_p(x_j \otimes f_p(x_i \otimes z_J)) + f_p([x_i, x_j] \otimes z_J).$$

Moreover, for every positive integer p, the restriction of f_p to $L \otimes P_{p-1}$ is f_{p-1}.

Proof. We will prove by induction that the following statement holds for all non-negative integers p: ($\mathrm{S_p}$) there exists a unique linear map $f_p : L \otimes P_p \to P$ satisfying ($\mathrm{A_p}$), ($\mathrm{B_p}$) and ($\mathrm{C_p}$) and such that the restriction of f_p to $L \otimes P_{p-1}$ is f_{p-1} when p is positive.

Suppose that $p = 0$. Clearly, there exists a unique linear map $f_0 : L \otimes P_0 \to P$ such that $f_0(x_i \otimes 1) = z_i$ for $i \in \{1, \ldots, n\}$. It is clear that ($\mathrm{A_0}$), ($\mathrm{B_0}$) and ($\mathrm{C_0}$) hold; for this, note that, by definition, $X_0 = \{I \in X : d(I) = 0\} = \{\emptyset\}$, $z_\emptyset = 1$, and $i \leq \emptyset$ for all $i \in \{1, \ldots, n\}$. It follows that ($\mathrm{S_0}$) holds.

Suppose that $p = 1$. To define the linear map $f_p : L \otimes P_1 \to P$ it suffices to define $f_1(x_i \otimes z_I) \in P$ for $i \in \{1, \ldots, n\}$ and $I \in X_0 \sqcup X_1$. If $i \in \{1, \ldots, n\}$ and $I \in X_0$, then $I = \emptyset$, and we define $f_1(x_i \otimes z_I) = f_0(x_i \otimes z_I) = z_i$. Assume that $i \in \{1, \ldots, n\}$ and $I \in X_1$. Write $I = (j)$. There are two cases. Assume first that $i \leq I$, i.e., $i \leq j$. In this case we define $f_1(x_i \otimes z_I) = z_i z_j$. Assume that $i \not\leq I$, so that $i > j$. We define:

$$f_1(x_i \otimes z_I) = z_i z_I + f_0([x_i, x_j] \otimes z_\emptyset),$$

i.e.,

$$f_1(x_i \otimes z_j) = z_i z_j + f_0([x_i, x_j] \otimes 1).$$

It is straightfoward to verify that f_1 satisfies ($\mathrm{A_1}$) and ($\mathrm{B_1}$). To see that f_1 satisfies ($\mathrm{C_1}$), let $i, j \in \{1, \ldots, n\}$ and $J \in X_{1-1} = X_0$. Then $J = \emptyset$. We need to prove that

$$f_1(x_i \otimes f_1(x_j \otimes z_\emptyset)) = f_1(x_j \otimes f_1(x_i \otimes z_\emptyset)) + f_1([x_i, x_j] \otimes z_\emptyset),$$

which is

$$f_1(x_i \otimes z_j) = f_1(x_j \otimes z_i) + f_0([x_i, x_j] \otimes 1).$$

Assume first that $i \leq j$. In this case,

$$f_1(x_i \otimes z_j) = z_i z_j,$$

and

$$\begin{aligned} f_1(x_j \otimes z_i) + f_0([x_i, x_j] \otimes 1) &= z_j z_i + f_0([x_j, x_i] \otimes 1) + f_0([x_i, x_j] \otimes 1) \\ &= z_j z_i - f_0([x_i, x_j] \otimes 1) + f_0([x_i, x_j] \otimes 1) \\ &= z_i z_j. \end{aligned}$$

This proves ($\mathrm{C_1}$) in the case $i \leq j$. Now assume that $i > j$. Then

$$f_1(x_i \otimes z_j) = z_i z_j + f_0([x_i, x_j] \otimes 1),$$

and

$$f_1(x_j \otimes z_i) + f_0([x_i, x_j] \otimes 1) = z_j z_i + f_0([x_i, x_j] \otimes 1).$$

This proves ($\mathrm{C_1}$) in the remaining case $i > j$. It follows that ($\mathrm{S_1}$) holds.

Now suppose that p is a positive integer such that $p \geq 2$ and that ($\mathrm{S_k}$) holds for $k = 0, \ldots, p-1$. To define the linear map $f_p : L \otimes P_p \to P$ it suffices to define $f_p(x_i \otimes z_I) \in P$ for $i \in \{1, \ldots, n\}$ and $I \in X_q$ with q such that $0 \leq q \leq p$. Let $i \in \{1, \ldots, n\}$ and assume that $I \in X_q$ with $0 \leq q < p$. In this case we define $f_p(x_i \otimes z_I) = f_{p-1}(x_i \otimes z_I)$. Assume that $i \in \{1, \ldots, n\}$ and $I \in X_p$. If $i \leq I$, then we define

$$f_p(x_i \otimes z_I) = z_i z_I.$$

Assume that $i \not\leq I$. To see how to define $f_p(x_i \otimes z_I)$ in this case, assume for the moment that f_k exists and satisfies ($\mathrm{A_k}$), ($\mathrm{B_k}$), and ($\mathrm{C_k}$) for non-negative integers k and that f_{k-1} is the restriction of f_k for $k = 1, \ldots, p$; we will find a formula for $f_p(x_i \otimes z_I)$ in terms of f_{p-1}. Let $I = (j, i_2, \ldots, i_p)$. By the definition of X, $j \leq i_2 \leq \cdots \leq i_p$. Since $i \not\leq I$, we must have $i > j$. Define $J = (i_2, \ldots, i_p)$; note that the definition of J is meaningful since $p \geq 2$. Clearly, $J \in X$ with $d(J) = p - 1$. We calculate, using ($\mathrm{A_{p-1}}$) and then ($\mathrm{C_p}$):

$$\begin{aligned}
f_p(x_i \otimes z_I) &= f_p(x_i \otimes z_j z_J) \\
&= f_p(x_i \otimes f_{p-1}(x_j \otimes z_J)) \\
&= f_p(x_i \otimes f_p(x_j \otimes z_J)) \\
&= f_p(x_j \otimes f_p(x_i \otimes z_J)) + f_p([x_i, x_j] \otimes z_J) \\
&= f_p(x_j \otimes f_{p-1}(x_i \otimes z_J)) + f_p([x_i, x_j] \otimes z_J).
\end{aligned}$$

Now since ($\mathrm{S_{p-1}}$) holds, we have by ($\mathrm{B_{p-1}}$),

$$f_{p-1}(x_i \otimes z_J) - z_i z_J \in P_{p-1}.$$

Define

$$w(i, J) = f_{p-1}(x_i \otimes z_J) - z_i z_J.$$

As just indicated, we have that $w(i, J) \in P_{p-1}$. Continuing the calculation, we get:

$$\begin{aligned}
f_p(x_i \otimes z_I) &= f_p(x_j \otimes f_{p-1}(x_i \otimes z_J)) + f_p([x_i, x_j] \otimes z_J) \\
&= f_p(x_j \otimes (z_i z_J + w(i, J))) + f_p([x_i, x_j] \otimes z_J) \\
&= f_p(x_j \otimes z_i z_J) + f_p(x_j \otimes w(i, J)) + f_p([x_i, x_j] \otimes z_J) \\
&= z_j z_i z_J + f_{p-1}(x_j \otimes w(i, J)) + f_p([x_i, x_j] \otimes z_J) \\
&= z_i z_I + f_{p-1}(x_j \otimes w(i, J)) + f_p([x_i, x_j] \otimes z_J).
\end{aligned}$$

Dropping our temporary assumption, we are now motivated to define:

$$f_p(x_i \otimes z_I) = z_i z_I + f_{p-1}\big(x_j \otimes w(i, J)\big) + f_{p-1}([x_i, x_j] \otimes z_J).$$

It is clear that f_p extends f_{p-1}; also, it is straightforward to verify that f_p satisfies ($\mathrm{A_p}$) and ($\mathrm{B_p}$). We need to prove that f_p satisfies ($\mathrm{C_p}$). Assume that $i, j \in \{1, \ldots, n\}$ and $J \in X_{p-1}$. The case $i = j$ holds trivially, so we assume that $i \neq j$. There are 5 possible cases.

Case 1: assume that $j < i$ and $j \leq J$. We have

$$\begin{aligned} f_p(x_i \otimes f_p(x_j \otimes z_J)) &= f_p(x_i \otimes z_j z_J) \\ &= f_p(x_i \otimes z_I) \\ &= z_i z_I + f_{p-1}\big(x_j \otimes w(i,J)\big) + f_{p-1}([x_i, x_j] \otimes z_J) \end{aligned}$$

where $I = (j, j_1, \ldots, j_{p-1})$, $J = (j_1, \ldots, j_{p-1})$ and we have used the definition of $f_p(x_i \otimes z_I)$ from the last paragraph. We note that $j \leq I$. Also,

$$\begin{aligned} f_p(x_j \otimes f_p(x_i \otimes z_J)) &= f_p(x_j \otimes f_{p-1}(x_i \otimes z_J)) \\ &= f_p(x_j \otimes (z_i z_J + w(i,J))) \\ &= f_p(x_j \otimes z_i z_J) + f_p(x_j \otimes w(i,J)) \\ &= z_j z_i z_J + f_{p-1}(x_j \otimes w(i,J)) \\ &= z_i z_I + f_{p-1}(x_j \otimes w(i,J)). \end{aligned}$$

Substituting, we obtain ($\mathrm{C_p}$) in this case.

Case 2: assume that $j < i$ and $i \leq J$. We then have $j < i$ and $j \leq J$. Case 1 now applies to prove ($\mathrm{C_p}$).

Case 3: assume that $i < j$ and $i \leq J$. Then ($\mathrm{C_p}$) follows from Case 1 with i replaced by j, j replaced by i and noting that $[x_j, x_i] = -[x_i, x_j]$.

Case 4: assume that $i < j$ and $j \leq J$. We then have $i < j$ and $i \leq J$. Case 3 now applies to prove ($\mathrm{C_p}$).

Case 5: assume that $i \not\leq J$ and $j \not\leq J$. Write $J = (k, \ldots, j_{p-1})$. By assumption $k < i$ and $k < j$. If $p > 2$, then define $K = (j_2, \ldots, j_{p-1})$; if $p = 2$, then define $K = \emptyset$. We have $k \leq K$. For the remainder of the proof we will write $f_q(x \otimes z)$ as $x \cdot z$ for q a non-negative integer with $q \leq p$, $x \in L$, and $z \in P_q$. Now

$$\begin{aligned} x_j \cdot z_J &= x_j \cdot (z_k z_K) \\ &= x_j \cdot (x_k \cdot z_K) \\ &= x_k \cdot (x_j \cdot z_K) + [x_j, x_k] \cdot z_K \end{aligned}$$

where the last equality follows from ($\mathrm{C_{p-1}}$). Now $x_j \cdot z_K = z_j z_K + w$ for some $w \in P_{p-2}$. Therefore,

$$x_j \cdot z_J = x_k \cdot (z_j z_K) + x_k \cdot w + [x_j, x_k] \cdot z_K.$$

Applying x_i, we get:

$$x_i \cdot (x_j \cdot z_J) = x_i \cdot (x_k \cdot (z_j z_K)) + x_i \cdot (x_k \cdot w) + x_i \cdot ([x_j, x_k] \cdot z_K).$$

Consider $x_i \cdot (x_k \cdot (z_j z_K))$. We may write $z_j z_K = z_M$ where M is formed from j and the entries of K. We have $k < i$ and $k \leq M$. By Case 1,

$$x_i \cdot (x_k \cdot (z_j z_K)) = x_k \cdot (x_i \cdot (z_j z_K)) + [x_i, x_k] \cdot (z_j z_K).$$

Since $w \in P_{p-2}$, we have by $(\mathrm{C}_{\mathrm{p}-2})$,

$$x_i \cdot (x_k \cdot w) = x_k \cdot (x_i \cdot w) + [x_i, x_k] \cdot w.$$

Substituting, this yields

$$\begin{aligned} x_i \cdot (x_j \cdot z_J) = x_k \cdot (x_i \cdot (z_j z_K)) + [x_i, x_k] \cdot (z_j z_K) \\ + x_k \cdot (x_i \cdot w) + [x_i, x_k] \cdot w + x_i \cdot ([x_j, x_k] \cdot z_K) \\ = x_k \cdot (x_i \cdot (x_j \cdot z_K)) + [x_i, x_k] \cdot (x_j \cdot z_K) + x_i \cdot ([x_j, x_k] \cdot z_K) \\ = x_k \cdot (x_i \cdot (x_j \cdot z_K)) + [x_i, x_k] \cdot (x_j \cdot z_K) \\ + [x_j, x_k] \cdot (x_i \cdot z_K) + [x_i, [x_j, x_k]] \cdot z_K, \end{aligned}$$

where we have applied $(\mathrm{C}_{\mathrm{p}-2})$ to $x_i \cdot ([x_j, x_k] \cdot z_K)$. The same argument with i and j interchanged yields

$$\begin{aligned} x_j \cdot (x_i \cdot z_J) = x_k \cdot (x_j \cdot (x_i \cdot z_K)) + [x_j, x_k] \cdot (x_i \cdot z_K) \\ + [x_i, x_k] \cdot (x_j \cdot z_K) + [x_j, [x_i, x_k]] \cdot z_K. \end{aligned}$$

Therefore, the difference is:

$$\begin{aligned} & x_i \cdot (x_j \cdot z_J) - x_j \cdot (x_i \cdot z_J) \\ & = x_k \cdot (x_i \cdot (x_j \cdot z_K)) - x_k \cdot (x_j \cdot (x_i \cdot z_K)) \\ & \quad + [x_i, [x_j, x_k]] \cdot z_K - [x_j, [x_i, x_k]] \cdot z_K \\ & = x_k \cdot \Big(x_i \cdot (x_j \cdot z_K) - x_j \cdot (x_i \cdot z_K) \Big) \\ & \quad + \Big([x_i, [x_j, x_k]] - [x_j, [x_i, x_k]] \Big) \cdot z_K \\ & = x_k \cdot ([x_i, x_j] \cdot z_K) + \big([x_i, [x_j, x_k]] + [x_j, [x_k, x_i]] \big) \cdot z_K \\ & = [x_i, x_j] \cdot (x_k \cdot z_K) + \big([x_k, [x_i, x_j]] + [x_i, [x_j, x_k]] + [x_j, [x_k, x_i]] \big) \cdot z_K \\ & = [x_i, x_j] \cdot z_J. \end{aligned}$$

This is $(\mathrm{C}_{\mathrm{p}})$. □

Lemma 11.4.4. *Let L be a finite-dimensional Lie algebra over F, and let $x_1, \dots, x_n$ be a basis for the F vector space L. Let the notation be as in the discussion preceding Lemma 11.4.3. There exists a representation ρ of L on $F[z_1, \dots, z_n]$ with the property that*

$$\rho(x_i) z_I = z_i z_I$$

for $i \in \{1, \dots, n\}$ and $I \in X$ with $i \leq I$.

Proof. We will use Lemma 11.4.3. For $x \in L$ and $p(z_1, \dots, z_n) \in F[z_1, \dots, z_n]$ define

$$\rho(x)\big(p(z_1, \dots, z_n)\big) = f_k(x \otimes p(z_1, \dots, z_n))$$

where k is any non-negative integer such that $p(z_1, \dots, z_n) \in P_k$. The assertions of Lemma 11.4.3 imply that ρ is a Lie algebra action with the stated property. □

Theorem 11.4.5. *Let F be a field, and let L be a finite-dimensional Lie algebra over F. Fix an ordered basis $x_1, \dots, x_n$ for L, and define y_I for $I \in X$ as preceding Lemma 11.4.2. Then the elements y_I for $I \in X$ are a basis for $U(L)$ as a vector space over F.*

Proof. By Lemma 11.4.2 it suffices to prove that y_I for $I \in X$ are linearly independent. Let ρ be the action of L on $F[z_1, \dots, z_n]$ from Lemma 11.4.4. By Lemma 11.4.1, there exists an action ρ' of $U(L)$ on $F[z_1, \dots, z_n]$ such that $\rho' \circ \sigma = \rho$. Let $I \in X$; we claim that

$$\rho'(y_I) \cdot 1 = z_I.$$

We will prove this by induction on $d(I)$. Assume that $d(I) = 0$. Then $I = \emptyset$. We have $y_I = 1$ so that $\rho'(y_I) = \rho(1) = 1$, and $z_I = 1$. Hence, $\rho'(y_I) \cdot 1 = z_I$. Assume that p is a positive integer and the claim holds for all I with $d(I) < p$. Let $I \in X$ be such that $d(I) = p$. Write $I = (i_1, \dots, i_p)$ for some positive integer p and $i_1, \dots, i_p \in \{1, \dots, n\}$ with $i_1 \leq \cdots \leq i_p$. Let $J = (i_2, \dots, i_p)$ if $p \geq 2$ and $J = \emptyset$ if $p = 1$. We have $i_1 \leq J$. Now

$$\begin{aligned}
\rho'(y_I) \cdot 1 &= \rho'(y_{i_1} y_J) \cdot 1 \\
&= \rho'(y_{i_1})(\rho'(y_J) \cdot 1) \\
&= \rho(x_{i_1})(z_J) \\
&= z_{i_1} z_J \\
&= z_I.
\end{aligned}$$

This proves the claim by induction. It follows now that the y_I for $I \in X$ are linearly independent because the z_I are linearly independent for $I \in X$. □

Index

Bibliography

[1] Nicolas Bourbaki. Lie groups and lie algebras, chapter 1–3, 1989.

[2] Nicolas Bourbaki. Lie groups and lie algebras. chapters 4–6. translated from the 1968 french original by andrew pressley. elements of mathematics, 2002.

[3] Nicolas Bourbaki. Lie groups and lie algebras. chapters 7–9. translated from the 1975 french original by andrew pressley. elements of mathematics, 2008.

[4] Karin Erdmann and Mark J. Wildon. *Introduction to Lie Algebras*. Springer Undergraduate Mathematics Series. Springer, 2006.

[5] Kenneth Hoffman and Ray Kunze. *Linear Algebra*. Prentice-Hall, second edition, 1971.

[6] James E Humphreys. *Introduction to Lie algebras and representation theory*, volume 9. Springer Science & Business Media, 2012.

[7] Jean-Pierre Serre. *Complex Semisimple Lie Algebras*. Springer Monographs in Mathematics. Springer, 1987.