Contents

Math 531 Notes, Fall 2007

David E. Radford
University of Illinois at Chicago
Chicago, IL

November 12, 2007

Chapter 1

Basic Concepts

Below are comments relevant to various sections in the text. They are meant to clarify, amplify, or generalize material in the text. Exercises are optional.

1.1 Definitions and first examples

1.1.1 The notion of a Lie algebra

We begin with a discussion of algebras in general. An *algebra over F* is a vector space A together with a bilinear map $m : A{\times}A \longrightarrow A$ referred to as *multiplication*. Usual notation for m is $m((a, b)) = ab$ for all $a, b \in A$. Bilinearity translates to

(BL.1) $(a + a')b = ab + a'b$ and $(\alpha a)b = \alpha(ab)$,

(BL.2) $a(b + b') = ab + ab'$ and $a(\alpha b) = \alpha(ab)$

for all $a, a', b, b' \in A$ and $\alpha \in F$. Sometimes it is very useful to think of multiplication as a linear map $\mathsf{m} : A{\otimes}_F A \longrightarrow A$ which is determined by $\mathsf{m}(a{\otimes}b) = m((a, b)) = ab$ for all $a, b \in A$.

Let A be an algebra over F. Since A is a ring

$$a0 = 0 = 0a \tag{1.1}$$

or all $a \in A$ and

$$-(ab) = (-a)b = a(-b) \tag{1.2}$$

for all $a, b \in A$. A *subalgebra of* A is a subspace B of A such that $ab \in B$ for all $a, b \in B$. Thus a subalgebra is an algebra in its own right with the multiplication of A. Among the subalgebras of A are A, and by (1.1), the singleton set $\{0\}$, usually denoted (0).

A *map of algebras*, or *homomorphism of algebras*, is a linear map $f : A \longrightarrow A'$, where A and A' are algebras over F, such that $f(ab) = f(a)f(b)$ for all $a, b \in A$. Suppose that $f : A \longrightarrow A'$ is a map of algebras. If B is a subalgebra of A then its image $f(B)$ is a subalgebra of A'. If B' is a subalgebra of A' then its preimage $f^{-1}(B')$ is a subalgebra of A.

A *Lie algebra over* F is an algebra L over F such that

(L.1) $a^2 = 0$ and

(L.2) $a(bc) + b(ca) + c(ab) = 0$

for all $a, b, c \in L$. Axiom (L.2) is referred to as the *Jacobi identity*.

Suppose that L is a Lie algebra over F. Then

$$ba = -ab \tag{1.3}$$

for all $b, a \in L$ as

$$0 = (a + b)^2 = (a + b)(a + b) = a^2 + ab + ba + b^2 = ab + ba.$$

When the characteristic of F is not 2, axiom (L.1) and (1.3) are equivalent. Observe that (1.3) holds for rings where the squares of all elements are zero.

In light of (1.2) and (1.3) we see that axiom (L.2) is equivalent to

$$a(bc) = b(ac) + (ab)c \tag{1.4}$$

for all $a, b, c \in L$.

Notation: The product of $a, b \in L$ is usually denoted $[a\,b]$.

Thus the axioms for a Lie algebras are usually written

$$[a\,a] = 0 \qquad \text{and} \qquad [a\,[b\,c]] + [b\,[c\,a]] + [c\,[a\,b]] = 0$$

for all $a, b, c \in L$.

Let $a \in L$. The linear span Fa is a subalgebra of A since $[a\,a] = 0$. Thus every one-dimensional subspace of L is a subalgebra of L. When L and L' are Lie algebras, the condition that a linear map $f : L \longrightarrow L'$ is a map of Lie algebras is expressed $f([a\,b]) = [f(a)\,f(b)]$ for all $a, b \in L$.

Exercise 1.1.1 You might call this variations on the Jacobi identity. Let A be an algebra over F. Show that Axioms (L.1) and (L.2) are equivalent to (L.1) and any of

(a) $(ab)c + (bc)a + (ca)b = 0$ for all $a, b, c \in A$;

(b) $a(bc) + c(ab) + b(ca) = 0$ for all $a, b, c \in A$;

(c) $(ab)c + (ca)b + (bc)a = 0$ for all $a, b, c \in A$.

1.1.2 Linear Lie algebras

An associative algebra gives rise to a Lie algebra. Let A be any associative algebra over F. Then the "bracket" defined by

$$[a\,b] = ab - ba$$

for all $a, b \in A$ gives the vector space A a Lie algebra structure. Note that $[a\,b]$ is the commutator of a and b. Observe that:

(a) a and b commute if and only if $[a\,b] = 0$;

(b) A is commutative if and only if $[a\,b] = 0$ for all $a, b \in A$;

(c) c is in the center of A if and only if $[a\,c] = 0$ for all $a \in A$.

Now let L be a Lie algebra over F. The last two statements motivate the following definitions:

L is *abelian* if $[a\,b] = 0$ for all $a, b \in L$;

The *center of* L is the set $Z(L) = \{c \in L \mid [a\,c] = 0\ \forall a \in L\}$.

For an associative algebra A over F let $\mathcal{L}(A)$ denote the Lie algebra with underlying vector space A and the bracket product. A map of associative algebras $f : A \longrightarrow A'$ is also a map of Lie algebras $f : \mathcal{L}(A) \longrightarrow \mathcal{L}(A')$. The categorically minded reader will notice that the associations $A \mapsto \mathcal{L}(A)$ and $f \mapsto f$ describes a functor from associative algebras to Lie algebras.

There is much to be said about the classical families $A_\ell - D_\ell$. The exercises below deal with important details. We comment here that Lie algebras of the types $B_\ell - D_\ell$ have the form $\mathcal{L}_s$ as described below.

Let $\mathrm{M}(n, F)$ be the (associative) algebra of $n \times n$ matrices over F and $s \in \mathrm{M}(n, F)$. Set

$$\mathcal{L}_s = \{x \in \mathrm{M}(n, F) \,|\, x^t s = -sx\}.$$

Then an easy calculation shows that $\mathcal{L}_s$ is a Lie subalgebra of $gl(n, F) = \mathcal{L}(\mathrm{M}(n, F))$.

Exercise 1.1.2 Using the fact that the trace function $\mathrm{Tr} : \mathrm{M}(n, F) \longrightarrow F$ is linear, use the Rank–Nullity Theorem to show that $sl(n, F) = \mathrm{Ker}\,\mathrm{Tr}$ has dimension $n^2 - 1$.

Exercise 1.1.3 Suppose that $s \in \mathrm{M}(n, F)$ is invertible and the characteristic of F is not 2. Show that $\mathcal{L}_s$ is a Lie subalgebra of $sl(n, F)$. (Thus the classical Lie algebras are subalgebras of $sl(n, F)$ for some n.)

Exercise 1.1.4 Let $s \in \mathrm{M}(n, F)$ and $b : F^n \times F^n \longrightarrow F$ be defined by $b(u, v) = u^t sv$ for all $u, v \in F^n$.

(a) Show that b is a bilinear form.

(b) Show that b is symmetric (meaning that $b(u, v) = b(v, u)$ for all $u, v \in F^n$) if and only if s is symmetric.

(c) Show that b is skew symmetric (meaning that $b(u, v) = -b(v, u)$ for all $u, v \in F^n$) if and only if s is skew symmetric.

(d) Show that

$$\mathcal{L}_s = \{x \in \mathrm{M}(n, F) \,|\, b(xu, v) = -b(u, xv) \text{ for all } u, v \in F^n\}.$$

Thus $\mathcal{L}_s$ can be described in terms of a bilinear form.

All bilinear forms on F^n are described by part (a) of Exercise 1.1.4. This is basic linear algebra included here for the record.

Exercise 1.1.5 Let $b : F^n \times F^n \longrightarrow F$ be a bilinear form and $s \in \mathrm{M}(n, F)$ be the matrix given by $s_{ij} = b(e_i, e_j)$, where $\{e_1, \ldots, e_n\}$ is the standard basis for F^n. Show that $b(u, v) = u^t sv$ for all $u, v \in F^n$.

There is a vector space analog of $\mathcal{L}_s$ which is suggested by part d) of Exercise 1.1.4.

Exercise 1.1.6 Let V be any vector space over F and suppose that $\beta : V \times V \longrightarrow F$ is a bilinear form. Show that $\mathcal{L}_\beta = \{x \in gl(V) \mid \beta(x(u), v) = -\beta(u, x(v)) \text{ for all } u, v \in V\}$ is a subalgebra of $gl(V)$.

The matrices s which describe the families $B_\ell - D_\ell$ are presented in "block" form and are symmetric or skew-symmetric. There is a very nice algebra of matrices presented in block form which makes the analysis of these classical families as vector spaces rather easy. First an exercise on symmetric and skew-symmetric matrices.

Exercise 1.1.7 Let $n \geq 1$.

(a) Show that the subspace of symmetric matrices of $\mathrm{M}(n, F)$ has dimension $\dfrac{n(n+1)}{2}$.

Suppose that the characteristic of F is not 2.

(b) Show that subspace of skew symmetric matrices of $\mathrm{M}(n, F)$ has dimension $\dfrac{(n-1)n}{2}$.

(c) Show that every matrix in $\mathrm{M}(n, F)$ can be written as a sum of a symmetric matrix and a skew symmetric matrix in $\mathrm{M}(n, F)$ in a unique way. (This accounts for the fact that the dimensions of parts a) and b) add to n^2.)

Observe that if F has characteristic 2 then symmetric matrices and skew symmetric matrices of $\mathrm{M}(n, F)$ are the same.

For computational purposes, we will view matrices in terms of blocks of entries and conceptualize these blocks as entries of a matrix. This process is illustrated by

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 7 & 8 & 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 & 17 & 18 \\ 19 & 20 & 21 & 22 & 23 & 24 \\ 25 & 26 & 27 & 28 & 29 & 30 \\ 31 & 32 & 33 & 34 & 35 & 36 \end{pmatrix} = \left(\begin{array}{cc|cccc} 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 7 & 8 & 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 & 17 & 18 \\ 19 & 20 & 21 & 22 & 23 & 24 \\ \hline 25 & 26 & 27 & 28 & 29 & 30 \\ 31 & 32 & 33 & 34 & 35 & 36 \end{array}\right) = \begin{pmatrix} \mathsf{A}_{11} & \mathsf{A}_{12} \\ \mathsf{A}_{21} & \mathsf{A}_{22} \\ \mathsf{A}_{31} & \mathsf{A}_{32} \end{pmatrix},$$

where

$$\mathsf{A}_{1\,1} = \begin{pmatrix} 1 & 2 \end{pmatrix} \qquad \mathsf{A}_{1\,2} = \begin{pmatrix} 3 & 4 & 5 & 6 \end{pmatrix}$$

$$\mathsf{A}_{2\,1} = \begin{pmatrix} 7 & 8 \\ 13 & 14 \\ 19 & 20 \end{pmatrix} \quad \mathsf{A}_{2\,2} = \begin{pmatrix} 9 & 10 & 11 & 12 \\ 15 & 16 & 17 & 18 \\ 21 & 22 & 23 & 24 \end{pmatrix}$$

$$\mathsf{A}_{3\,1} = \begin{pmatrix} 25 & 26 \\ 30 & 31 \end{pmatrix} \quad \mathsf{A}_{3\,2} = \begin{pmatrix} 27 & 28 & 29 & 30 \\ 33 & 34 & 35 & 36 \end{pmatrix}$$

We are regarding the 6×6 matrix A as an $\mathbf{m}\times\mathbf{n}$ matrix, where

$$\mathbf{m} = (m_1, m_2, m_3) = (1, 3, 2) \qquad \text{and} \qquad \mathbf{n} = (n_1, n_2) = (2, 4)$$

describe how the rows and columns respectively are partitioned.

Suppose that A is an $m\times n$ matrix with coefficients in F,

$$\mathbf{m} = (m_1, m_2, \ldots, m_r) \qquad \text{and} \qquad \mathbf{n} = (n_1, n_2, \ldots, n_s)$$

have positive integer entries such that

$$m_1 + \cdots + m_r = m \qquad \text{and} \qquad n_1 + \cdots + n_s = n.$$

Then A can be regarded as the $\mathbf{m}\times\mathbf{n}$ matrix

$$A = \begin{pmatrix} \mathsf{A}_{1\,1} & \cdots & \mathsf{A}_{1\,s} \\ \vdots & & \vdots \\ \mathsf{A}_{r\,1} & \cdots & \mathsf{A}_{r\,s} \end{pmatrix},$$

where $\mathsf{A}_{i\,j}$ is the $m_i\times n_j$ matrix whose entries is given by

$$(\mathsf{A}_{i\,j})_{k\,\ell} = A_{m_1+\cdots+m_{i-1}+k \quad n_1+\cdots+n_{j-1}+\ell}$$

for all $1 \leq i \leq r$, $1 \leq j \leq s$, $1 \leq k \leq m_i$, and $1 \leq k \leq n_j$. By convention $m_0 = n_0 = 0$.

Observe that A^t is the $\mathbf{n}\times\mathbf{m}$ matrix described by

$$A^t = \begin{pmatrix} (\mathsf{A}_{1\,1})^t & \cdots & (\mathsf{A}_{r\,1})^t \\ \vdots & & \vdots \\ (\mathsf{A}_{1\,s})^t & \cdots & (\mathsf{A}_{r\,s})^t \end{pmatrix} = \begin{pmatrix} (\mathsf{A}_{1\,1})^t & \cdots & (\mathsf{A}_{1\,s})^t \\ \vdots & & \vdots \\ (\mathsf{A}_{r\,1})^t & \cdots & (\mathsf{A}_{r\,s})^t \end{pmatrix}^t,$$

where the latter is a formal expression.

Exercise 1.1.8 Suppose that A is an $m{\times}n$ matrix and B is an $n{\times}p$ matrix with coefficients in F and let $C = AB$ be the $m{\times}p$ matrix which is their product. Let

$$\mathbf{m} = (m_1, m_2, \dots, m_r), \qquad \mathbf{n} = (n_1, n_2, \dots, n_s), \qquad \text{and} \qquad \mathbf{p} = (p_1, p_2, \dots, p_t)$$

have positive integer entries and satisfy

$$m_1 + \cdots + m_r = m, \qquad n_1 + \cdots + n_s = n, \qquad \text{and} \qquad p_1 + \cdots + p_t = p.$$

Write

$$A = \begin{pmatrix} \mathsf{A}_{1\,1} & \cdots & \mathsf{A}_{1\,s} \\ \vdots & & \vdots \\ \mathsf{A}_{r\,1} & \cdots & \mathsf{A}_{r\,s} \end{pmatrix}, \qquad B = \begin{pmatrix} \mathsf{B}_{1\,1} & \cdots & \mathsf{B}_{1\,t} \\ \vdots & & \vdots \\ \mathsf{B}_{s\,1} & \cdots & \mathsf{B}_{s\,t} \end{pmatrix}, \qquad \text{and} \qquad C = \begin{pmatrix} \mathsf{C}_{1\,1} & \cdots & \mathsf{C}_{1\,t} \\ \vdots & & \vdots \\ \mathsf{C}_{r\,1} & \cdots & \mathsf{C}_{r\,t} \end{pmatrix}$$

as above. Show that

$$\mathsf{C}_{i\,j} = \sum_{\ell=1}^{s} \mathsf{A}_{i\,\ell}\mathsf{B}_{\ell\,j}$$

for all $1 \le i \le r$ and $1 \le j \le t$.

Exercise 1.1.9 Let $\ell \ge 1$, suppose that $A \in \mathrm{M}(\ell, F)$ is symmetric and invertible, and suppose that the characteristic of F is not 2.

(a) *Generalization of type* D_ℓ. Let $s \in \mathrm{M}(2\ell, F)$ have block form $s = \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}$.
Show that $\mathrm{Dim}\mathcal{L}_s = 2\ell^2 - \ell$.

(b) *Generalization of type* C_ℓ. Let $s \in \mathrm{M}(2\ell, F)$ have block form $s = \begin{pmatrix} 0 & A \\ -A & 0 \end{pmatrix}$.
Show that $\mathrm{Dim}\mathcal{L}_s = 2\ell^2 + \ell$.

(c) *Generalization of type* B_ℓ. Let $s \in \mathrm{M}(2\ell{+}1, F)$ have block form $s = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & A \\ 0 & -A & 0 \end{pmatrix}$.
Show that $\mathrm{Dim}\mathcal{L}_s = 2\ell^2 + \ell$.

1.1.3 Lie algebras of derivations

Let A be an algebra over F. A *derivation of* A is a linear endomorphism $D : A \longrightarrow A$ which satisfies

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

for all $a, b \in A$. Important examples of derivations for us include the formal partial derivatives $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_r}$ of the commutative polynomial algebra $F[x_1, \ldots, x_r]$. The notion of derivation can be thought of as a generalization of the product rule of Differential Calculus $(fg)' = f'g + fg'$ for differentiation. Observe that the kernel of a derivation is a subalgebra of A by (1.1). The set $\mathrm{Der}(A)$ of derivations of A is a subspace of $\mathrm{End}(A)$.

Suppose that D, D' are derivations of A. For $a, b \in A$ the calculation

$$\begin{aligned}(D\circ D')(ab) &= D(D'(ab)) \\ &= D(aD'(b) + D'(a)b) \\ &= \big(a(D(D'(b))) + D(a)D'(b)\big) + \big(D'(a)D(b) + (D(D'(a)))b\big)\end{aligned}$$

shows that

$$(D\circ D')(ab) = a(D(D'(b))) + D(a)D'(b) + D'(a)D(b) + (D(D'(a)))b,$$

and hence

$$(D'\circ D)(ab) = a(D'(D(b))) + D'(a)D(b) + D(a)D'(b) + (D'(D(a)))b.$$

Thus

$$(D\circ D' - D'\circ D)(ab) = a((D\circ D' - D'\circ D)(b)) + ((D\circ D' - D'\circ D)(a))b$$

for all $a, b \in A$ which means that $[D\, D'] = D\circ D' - D'\circ D$ is a derivation of A. Therefore

$$\mathrm{Der}(A) \text{ is a Lie subalgebra of } gl(A) \tag{1.5}$$

By induction on n the Leibnitz rule

$$D^n(ab) = \sum_{\ell=0}^{n} \binom{n}{\ell} D^\ell(a) D^{n-\ell}(b) \tag{1.6}$$

holds for all $n \geq 0$ and $a, b \in A$. By convention $D^0 = 1_{\mathrm{End}(A)} = \mathrm{Id}_A$ is the identity map of A. Note that when $n = 1$ the formula is the definition of derivation.

For $a \in A$ the functions $\ell_a, r_a : A \longrightarrow A$ defined by

$$\ell_a(x) = ax \qquad \text{and} \qquad r_a(x) = xa$$

for all $x \in A$ are called *left* (respectively *right*) *multiplication by* a. Note that ℓ_a is linear by (BL.2) and r_a is linear by (BL.1). The map

$$\pi : A \longrightarrow \mathrm{End}(A)$$

defined by $\pi(a) = \ell_a$ for all $a \in A$ is linear by (BL.1).

Now suppose that A is associative. Let $a, b \in A$. The calculation

$$(\ell_a \circ r_b)(x) = \ell_a(r_b(x)) = a(xb) = (ax)b = r_b(\ell_a(x)) = (r_b \circ \ell_a)(x)$$

for all $x \in A$ shows that ℓ_a and r_b commute. Also

$$\ell_{ab}(x) = (ab)x = a(bx) = (\ell_a(\ell_b(x)) = (\ell_a \circ \ell_b)(x)$$

for all $x \in A$ shows that $\ell_{ab} = \ell_a \circ \ell_b$. Since A is an associative algebra the map π is an algebra map, called the *left regular representation of* A. When A has a unity π is one-one; in this case A can be regarded as a subalgebra of the algebra of endomorphisms of a vector space.

Now suppose that L is a Lie algebra over F. Then (1.4) can be expressed

$$\ell_a(bc) = b\ell_a(c) + \ell_a(b)c$$

for all $a, b, c \in L$; equivalently ℓ_a is a derivation of L for all $a \in L$. Thus $\pi(a) = \ell_a \in \mathrm{Der}(L)$ and we may think of

$$\pi : L \longrightarrow \mathrm{Der}(L)$$

which is linear. Let $a, b \in L$. The calculation

$$[\ell_a\, \ell_b](x) = (\ell_a \circ \ell_b - \ell_b \circ \ell_a)(x) = \ell_a(bx) - b\ell_a(x) = b\ell_a(x) + \ell_a(b)x - b\ell_a(x) = \ell_a(b)x = (ab)x$$

for all $x \in L$ shows that $\ell_{ab} = [\ell_a\, \ell_b]$. Therefore $\pi : L \longrightarrow \mathrm{Der}(L)$ is a Lie algebra map, called the *regular representation of* L.

Notation: $\mathrm{ad}\, a = \ell_a$. Thus $\mathrm{ad}\, a(x) = [a\, x]$ for all $x \in L$.

The fact that $\mathrm{ad}\, a$ is a derivation of L is expressed

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

for all $x, y \in L$ and the fact that π is multiplicative is expressed

$$\mathrm{ad}\, [a\, b] = [\mathrm{ad}\, a \;\; \mathrm{ad}\, b]$$

for all $a, b \in L$.

Let A be an algebra over F. The first two exercises below give a more theoretical way of deducing (1.6) and the fact the $\mathrm{Der}(A)$ is a subalgebra of $\mathrm{End}(A)$. We think of multiplication as the linear map $\mathsf{m} : A \otimes_F A \longrightarrow A$ given by $\mathsf{m}(a \otimes b) = ab$ for all $a, b \in A$.

Exercise 1.1.10 Let $D : A \longrightarrow A$ be linear.

(a) Show that D is a derivation of A if and only if $D \circ \mathsf{m} = \mathsf{m} \circ (\mathrm{Id}_A \otimes D + D \otimes \mathrm{Id}_A)$.

(b) Show that $D^n \circ \mathsf{m} = \mathsf{m} \circ (\mathrm{Id}_A \otimes D + D \otimes \mathrm{Id}_A)^n$ for all $n \geq 0$.

(c) Show that $\mathrm{Id}_A \otimes D$ and $D \otimes \mathrm{Id}_A$ commute; that is $(\mathrm{Id}_A \otimes D) \circ (D \otimes \mathrm{Id}_A) = (D \otimes \mathrm{Id}_A) \circ (\mathrm{Id}_A \otimes D)$.

(d) Use part (c) and the Binomial Theorem to deduce that $D^n \circ \mathsf{m} = \mathsf{m} \circ \left(\sum_{\ell=0}^{n} \binom{n}{\ell} D^{\ell} \otimes D^{n-\ell} \right)$ for all $n \geq 0$.

(e) Deduce (1.6) from part (d).

Exercise 1.1.11 Use part (a) of Exercise 1.1.10 to show that $\mathrm{Der}(A)$ is a Lie subalgebra of $gl(A)$.

Exercise 1.1.12 Show that A is a Lie algebra over F if and only if (L.1) and either of the following hold:

(a) ℓ_a is a derivation of A for all $a \in A$;

(b) r_a is a derivation of A for all $a \in A$.

Exercise 1.1.13 Now assume that A is associative and $a \in A$. Here ℓ_a and r_a are defined for the associative structure on A and $\mathrm{ad}\, a$ is defined for $gl(A)$.

(a) Show that $\mathrm{ad}\, a = \ell_a - r_a$.

(b) Show that $(\mathrm{ad}\, a)^n = \sum_{\ell=0}^{n} (-1)^{\ell} \binom{n}{\ell} \ell_{a^{n-\ell}} \circ r_{a^{\ell}}$ for all $n \geq 0$. [Hint: Recall that ℓ_a and r_b commute for all $a, b \in A$. See part (d) of Exercise 1.1.10.]

(c) Suppose that $a^n = 0$ for some $n > 0$. Show that $(\mathrm{ad}\, a)^{2n-1} = 0$. (Thus if a is a nilpotent element of A then $\mathrm{ad}\, a$ is a nilpotent element of the associative algebra $\mathrm{end}(A)$.)

1.1.4 Abstract Lie Algebras

We start by constructing finite-dimensional algebras in a very concrete way. Let A be a vector space over F with basis $\{a_1, \ldots, a_r\}$ and let $\{a_{\imath,\jmath}^{\ell}\}_{1\leq \imath,\jmath,\ell\leq r}$ be *any* set of scalars. We define a multiplication of basis vectors by

$$a_\imath a_\jmath = \sum_{\ell=1}^{r} a_{\imath,\jmath}^{\ell} a_\ell$$

for all $1 \leq \imath, \jmath \leq r$. There is at most one way to extend this multiplication of basis vectors to an algebra structure on A. To see this, let $a, b \in A$ and write $a = x_1a_1 + \cdots + x_ra_r$ and $b = y_1a_1 + \cdots + y_ra_r$ as linear combinations of basis elements. Then bilinearity forces

$$ab = (\sum_{\imath=1}^{r} x_\imath a_\imath)(\sum_{\jmath=1}^{r} y_\jmath a_\jmath) = \sum_{\imath,\jmath=1}^{r} x_\imath y_\jmath a_\imath a_\jmath = \sum_{\imath,\jmath,\ell=1}^{r} x_\imath y_\jmath a_{\imath,\jmath}^{\ell} a_\ell.$$

One can check that the rule

$$ab = \sum_{\imath,\jmath,\ell=1}^{r} x_\imath y_\jmath a_{\imath,\jmath}^{\ell} a_\ell$$

does indeed define an algebra structure on A. The scalars $a_{\imath,\jmath}^{\ell}$ are called *structure constants*. All finite-dimensional algebras can be described in terms of structure constants.

Now let A be any algebra over F, not necessarily finite-dimensional. We examine what it takes to verify that certain axioms hold.

Define

$$f : A{\times}A{\times}A \longrightarrow A$$

by

$$f(a, b, c) = a(bc) + b(ca) + c(ab)$$

for all $a, b, c \in A$. The Jacobi Identity is the same as the identity $f(a, b, c) = 0$ for all $a, b, c \in A$. What is sufficient to imply the latter?

Observe that f is linear in each variable. (This should be remind you of the determinant function on $n{\times}n$ matrices Det : $F^n \times \cdots \times F^n \longrightarrow F$ thought of a function on the columns of $n{\times}n$ matrices.) That f is linear in each variable means that for all $a, b \in A$ the functions

$$f_{a,b,*}, f_{a,*,b}, f_{*,a,b} : A \longrightarrow A$$

defined by

$$f_{a,b,*}(x) = f(a,b,x), \qquad f_{a,*,b}(x) = f(a,x,b), \text{ and } \qquad f_{*,a,b}(x) = f(x,a,b)$$

for all $x \in A$ respectively are linear. Check that the axioms for an algebra imply that these functions are linear. Note that $f_{a,b,*}(x) = f(a,b,x) = 0$ for all $x \in A$, or equivalently the subspace $\mathrm{Ker} f_{a,b,*} = A$, if $\mathrm{Ker} f_{a,b,*}$ contains a spanning set for A. Thus $f(a,b,x) = 0$ for all $x \in A$ if $f(a,b,x) = 0$ for all x in some spanning set for A.

At this point it is not hard to construct an argument to show that $f(a,b,c) = 0$ for all $a,b,c \in A$ if this equation holds for all a,b,c in some spanning set for A.

If the characteristic of F is not 2 then the axiom $a^2 = 0$ for all $a \in A$ is equivalent to $ab = -ba$, or $ab + ba = 0$, for all $a, b \in A$. The latter holds if and only if $g(a,b) = 0$ for all $a, b \in A$, where

$$g(a,b) = ab + ba.$$

Note that g is linear in each variable; thus $g(a,b) = 0$ for all $a, b \in A$ if and only if $g(a,b) = 0$ for all a, b in some spanning set for A.

1.2 Ideals and homomorphisms

1.2.1 Ideals

Most of what is done in this section applies to algebras in general, so we will start there. Let A be an algebra over F. An *left ideal* (respectively *right*) *ideal of* A is a subspace I of A such that $ab \in I$ (respectively $ba \in I$) for all $a \in A$ and $b \in I$. An *ideal*, or *two-sided ideal, of* A is a subspace of A which is both a left ideal and a right ideal of A. Thus left or right ideals, and ideals, of A are subalgebras of A.

Left, right, and two-sided ideals can be defined in terms of left and right multiplication maps $\ell_a, r_a : A \longrightarrow A$. Suppose that I is a subspace of A. Then I is a left (respectively right, or two-sided) ideal of A if and only if $\ell_a(I) \subseteq I$ (respectively $r_a(I) \subseteq I$, or $\ell_a(I), r_a(I) \subseteq I$) for all $a \in A$. In particular

$$I \text{ is a left ideal of } A \text{ if and only if } \ell_a(I) \subseteq I \text{ for all } a \in A; \tag{1.7}$$

that is I is invariant under the endomorphism ℓ_a for all $a \in A$.

Suppose that $f : A \longrightarrow A'$ is an algebra homomorphism. We have noted that $f(A)$ is a subalgebra of A'. If I is a left (respectively right, two-sided) ideal of A then the image $f(I)$ is a left (respectively right, two-sided) ideal of $f(A)$. If I' is a left (respectively right, two-sided) ideal of A' then the pre-image $f^{-1}(I')$ is a left (respectively right, two-sided) ideal of A.

Note that A is always an ideal of A as is the zero dimensional subspace $(0) = \{0\}$ of A by (1.1). Consequently, if $f : A \longrightarrow B$ is an algebra homomorphism then $\mathrm{Ker} f = f^{-1}((0))$ is an ideal of A.

Let I be an ideal of A. Then the quotient vector space A/I is an algebra over F, where

$$(a + I)(b + I) = ab + I$$

for all $a, b \in A$. The main aspect of this assertion is whether or not multiplication is well-defined. The fact that I is a two-sided ideal is used for this. For suppose that $a, b, a', b' \in A$ and $a + I = a' + I$, $b + I = b' + I$. Then $a' = a + x$ and $b' = b + y$ for some $x, y, \in I$. The calculation

$$a'b' = (a + x)(b + y) = ab + ay + xb + xy = ab + z,$$

where $z = ay + xb + xy \in I$, shows that $ab + I = a'b' + I$. Observe that the linear projection

$$\pi : A \longrightarrow A/I,$$

which is defined by $\pi(a) = a + I$ for all $a \in A$, is an algebra homomorphism.

Note that $\mathrm{Ker}\pi = I$. We have shown that ideals of algebras and kernels of algebra homomorphisms are the same.

Now let $\{I_\imath\}_{\imath \in \mathcal{I}}$ be an indexed set of ideals of A. Then it is easy to see that

$$\bigcap_{\imath \in \mathcal{I}} I_\imath \qquad \text{and the vector space sum} \qquad \sum_{\imath \in \mathcal{I}} I_\imath \qquad \text{are ideals of } A. \tag{1.8}$$

A concrete description of the sum is

$$\sum_{\imath \in \mathcal{I}} I_\imath = \{a_{\imath_1} + \cdots + a_{\imath_r} \,|\, r \geq 1, \imath_1, \ldots, \imath_r \in \mathcal{I}, a_{\imath_\jmath} \in I_{\imath_\jmath} \, \forall 1 \leq \jmath \leq r\}.$$

Let S be a subset of A. The ideal A of A contains S. By the intersection part of (1.8), among all ideals of A which contain S there is a unique minimal one, namely the intersection of all ideals of A which contain S. This ideal is usually denoted by (S).

Suppose that U, V are subspaces of A. Then UV denotes the *span* of the set of products uv, where $u \in U$ and $v \in V$.

If A is associative then A is said to be *simple* if A has exactly two ideals; in other words A is not (0) and the only ideals of A are A and (0). If A is a Lie algebra then A is said to be *simple* if A has exactly two ideals *and* A is not abelian. Thus 1-dimensional Lie algebras, even though they have exactly two ideals, are not simple.

Now let us turn the special case of a Lie algebra L. Since $[y\,x] = -[x\,y]$ for all $x, y \in L$, there is no distinction between left, right, and two-sided ideal of L. Thus a subspace I of L is an ideal of L if and only if $[L\,I] \subseteq I$. In terms of the adjoint action:

$$I \text{ is an ideal of } L \text{ if and if } \operatorname{ad} x(I) \subseteq I \text{ for all } x \in L \tag{1.9}$$

by (1.7). Another way of stating (1.9) is to say that *the ideals of* L *are the subspaces of* L *invariant under* $\operatorname{ad} x$ *for all* $x \in L$.

Suppose that I, J are ideals of L. Let $x \in L$. Since $\operatorname{ad} x$ is a derivation of L, the calculation

$$\operatorname{ad} x([I\,J]) \subseteq [\operatorname{ad} x(I)\,J] + [I\,\operatorname{ad} x(J)] \subseteq [I\,J] + [I\,J] = [I\,J]$$

shows that $[I\,J]$ is an ideal of L by (1.9). Also prove this assertion directly from the axioms for a Lie algebra!

A comment about principal ideals. Let $x \in L$. Then $[L\,x]$ is contained in the ideal generated by x (we should say more formally by $\{x\}$). Observe that

$$[L\,x] = [x\,L] = \operatorname{Im} \operatorname{ad} x.$$

Now $[x\,x] = 0$ means that $\operatorname{Ker} \operatorname{ad} x$ is not (0). Thus when L is finite-dimensional, $[L\,x]$ is a *proper* subspace of L by the Rank–Nullity Theorem applied to the linear endomorphism $\operatorname{ad} x : L \longrightarrow L$. When L is finite-dimensional and simple, $[L\,x]$ is an ideal of L if and only if $x = 0$. See Exercises 1.2.3 and 1.2.4.

Concerning centralizers and normalizers: let U, V be subspaces of L. Then

$$\mathrm{C}_U(V) = \{u \in U \mid [u\,V] = (0)\}$$

is the *centralizer of* V *in* U and

$$\mathrm{N}_U(V) = \{u \in U \mid [u\,V] \subseteq V\}$$

is the *normalizer of* V *in* U. Observe that $\mathrm{C}_U(V) \subseteq \mathrm{N}_U(V)$.

The following assertions follow directly from the axioms for a Lie algebra. We are interested in understanding these assertions in terms of the endomorphisms $\mathrm{ad}\, x$.

First of all $\mathrm{C}_L(V) = \bigcap_{x \in V} \mathrm{Ker}\,\mathrm{ad}\, x$ and is thus a subalgebra of L. The intersection characterization is clear. The kernel of a derivation is a subalgebra, and the intersection of subalgebras of an algebra is a subalgebra. Since $\mathrm{C}_U(V) = U \cap \mathrm{C}_L(V)$ it follows that:

$$\text{If } U \text{ is a sublagebra of } L \text{ then } \mathrm{C}_U(V) \text{ is a subalgebra of } L. \tag{1.10}$$

Observe that $N_L(V) = \{x \in L \,|\, \mathrm{ad}\, x(V) \subseteq V\}$ is a subalgebra of L. To see this note that $\mathcal{L} = \{T \in \mathrm{End}(L) \,|\, T(V) \subseteq V\}$ is a subalgebra of the associative algebra $\mathrm{End}(L)$; thus $\mathcal{L}$ a Lie subalgebra of $gl(L)$. Let $\pi : L \longrightarrow gl(L)$ be the regular representation of L; thus $\pi(x) = \mathrm{ad}\, x$ for all $x \in L$. Since $\mathrm{N}_L(V) = \pi^{-1}(\mathcal{L})$ is the pre-image of a Lie subalgebra under a Lie algebra map, it follows that $\mathrm{N}_L(V)$ is a subalgebras of L. Since $\mathrm{N}_U(V) = U \cap \mathrm{N}_L(V)$ we have shown:

$$\text{If } U \text{ is a sublagebra of } L \text{ then } \mathrm{N}_U(V) \text{ is a subalgebra of } L. \tag{1.11}$$

We next observe that $\mathrm{C}_L(V)$ is an ideal of $\mathrm{N}_L(V)$. The reason for this is that $\pi' : \mathrm{N}_L(V) \longrightarrow gl(V)$ given by $\pi'(x) = \mathrm{ad}\, x|_V$ is a well-defined map of Lie algebras and $\mathrm{C}_L(V) = \mathrm{Ker}\,\pi'$. As $\mathrm{N}_U(V) = U \cap \mathrm{N}_L(V)$ and $\mathrm{C}_U(V) = U \cap \mathrm{C}_L(V)$:

$$\text{If } U \text{ is a sublagebra of } L \text{ then } \mathrm{C}_U(V) \text{ is an ideal of } \mathrm{N}_U(V). \tag{1.12}$$

In the following exercises A is an algebra over F and L is a Lie algebra over F.

Exercise 1.2.1 Suppose that A is associative and U, V are subspaces of A. Show that:

(a) If U is a left ideal of A then UV is a left ideal of A.

(b) If V is a right ideal of A then UV is a right ideal of A.

(c) If U is a left ideal of A and V is a right ideal of A then UV is an ideal of A.

Exercise 1.2.2 Suppose that $D : A \longrightarrow A$ is a derivation of A. Show that:

(a) $\operatorname{Ker} D$ is a subalgebra of A.

(b) If U, V are ideals of A and $D(U) \subseteq U$, $D(V) \subseteq V$ then $D(UV) \subseteq UV$. See (1.9) and the discussion of the subsequent paragraph.

Exercise 1.2.3 Let $\pi : L \longrightarrow \operatorname{Der}(L)$ be the principal representation of L. Recall that $\pi(x) = \operatorname{ad} x$ for all $x \in L$.

(a) Show that $\operatorname{Ker} \pi = \mathrm{Z}(L)$; thus the center of L is an ideal of L.

(b) Suppose that L is simple. Show that L is isomorphic to a subalgebra of $gl(L)$; hence L is linear.

Exercise 1.2.4 Suppose that L is finite-dimensional and simple. Show that $[L\,x]$ only if $x = 0$.

Exercise 1.2.5 Show that L is simple if and only if L has exactly 2 ideals and $\operatorname{Dim} L > 1$.

Exercise 1.2.6 This exercise outlines a proof of the simplicity of $sl(2, F)$ when the characteristic of F is not 2. First of all let F be any field.

(a) Suppose that $\operatorname{Dim} L \leq 2$. Show that $[L\,L] = L$ implies $L = (0)$.

(b) Suppose that $\operatorname{Dim} L = 3$. Show that L is simple if and only if $[L\,L] = L$. [Hint: Let I be an ideal of L and consider the projection $\pi : L \longrightarrow L/I$. Observe that $\pi([L\,L]) = [\pi(L)\,\pi(L)]$.]

(c) Now suppose that the characteristic of F is not 2. Use part (b) to show that $sl(2, F)$ is simple.

1.2.2 Homomorphisms and Representations

It should be no surprise that the proposition of this section holds for all algebras over F. To establish this recall that you have the homomorphism theorems for abelian groups at your disposal.

Let L be a Lie algebra. We merely summarize some important details.

(a) Regarding L as a vector space over F, recall that the associative algebra $\operatorname{End}(L)$ has a Lie algebra structure defined by

$$[S\,T] = S \circ T - T \circ S$$

for all $S, T \in \operatorname{End}(L)$. The vector space $\operatorname{End}(L)$ with this Lie product is referred to as $gl(L)$.

(b) The subset Der A of derivations of the (Lie) algebra L is a (Lie) subalgebra of $gl(L)$.

(c) $\operatorname{ad} x \in \operatorname{Der}(L)$ for all $x \in L$.

(d) The map $\pi : L \longrightarrow gl(L)$ defined by $\pi(x) = \operatorname{ad} x$ for all $x \in L$ is map of Lie algebras. The algebra map π is called the *adjoint representation of* L.

Observe that

$$\operatorname{Ker} \pi = \{x \in L \mid \operatorname{ad} x = 0\} = \{x \in L \mid [x\, L] = (0)\} = \mathrm{Z}(L).$$

Thus π is one-one if and only if the center $\mathrm{Z}(L) = (0)$, as is the case when L is simple.

1.2.3 Automorphisms

This is a rather compact section. Here is what seems to be the main point.

Let L be a Lie subalgebra of $gl(V) = \mathcal{L}(\operatorname{End}(V))$ for some vector space V over F. Any automorphism f of the associative algebra $\operatorname{End}(V)$ is a Lie algebra automorphism of $gl(V)$. Thus if $f(L) = L$ the restriction $f|_L$ is a Lie algebra automorphism of L.

Typical automorphisms of the associative algebra $\operatorname{End}(V)$ are f_u defined by $f_u(x) = u{\circ}x{\circ}u^{-1}$ for all $x \in \operatorname{End}(V)$, where u in a linear automorphism of V. (When V is finite-dimensional and F is algebraically closed these are the only automorphisms of $\operatorname{End}(V)$.)

Suppose that $x \in L$ is a nilpotent endomorphism of V and that the characteristic of F is zero. Then the formal exponential $u = \exp(x)$ is a linear automorphism of V and $f_u(L) = L$. Thus the restriction $f_u|_L$ is a Lie algebra automorphism of L. Now for the details.

Let A be an algebra over F. An *automorphism of* A is a bijective algebra endomorphism of A. The set $\operatorname{Aut}(A)$ of algebra automorphisms of A is a group under function composition. There are interesting connections between $\operatorname{Aut}(A)$ and $\operatorname{Der}(A)$.

Let $D \in \mathrm{Der}(A)$ and $f \in \mathrm{Aut}(A)$. Then $f^{-1} \in \mathrm{Aut}(A)$ and the calculation

$$\begin{aligned}
&(f \circ D \circ f^{-1})(ab) \\
&\qquad = f(D(f^{-1}(ab))) \\
&\qquad = f(D(f^{-1}(a)f^{-1}(b))) \\
&\qquad = f(f^{-1}(a)D(f^{-1}(b)) + D(f^{-1}(a))f^{-1}(b)) \\
&\qquad = a(f(D(f^{-1}(b)))) + (f(D(f^{-1}(a))))b
\end{aligned}$$

for all $a, b \in A$ shows that:

$$\text{If } f \in \mathrm{Aut}(A) \text{ and } D \in \mathrm{Der}(A) \text{ then } f \circ D \circ f^{-1} \in \mathrm{Der}(A). \tag{1.13}$$

By virtue of (1.13) the group $\mathrm{Aut}(A)$ acts on $\mathrm{Der}(A)$ by conjugation.

To continue we will need to discuss nilpotent and unipotent elements, and the exponential and logarithm functions in an algebraic setting. Let A be an associative algebra with unity, for example $\mathrm{End}(V)$ under composition. An element $a \in A$ is *nilpotent* if $a^n = 0$ for some $n > 0$ and a is *unipotent* if $a = 1 + b$ for some nilpotent element $b \in A$.

Lemma 1.2.7 *Let A be an associative algebra with unity over a field of characteristic 0, let $\mathcal{N}$ be the set of nilpotent elements of A, and let $\mathcal{U}$ be the set of unipotent elements of A.*

(a) *$0 \in \mathcal{N}$, and if $a, b \in \mathcal{N}$ commute then $a + \alpha b \in \mathcal{N}$ for all $\alpha \in F$.*

(b) *If $a \in \mathcal{N}$ and $b \in A$ commute then $ab \in \mathcal{N}$.*

(c) *$1 \in \mathcal{U}$, and if $a, b \in \mathcal{U}$ commute then $ab \in \mathcal{U}$.*

(d) *If $a \in \mathcal{U}$ then a has a multiplicative inverse and $a^{-1} \in \mathcal{U}$.*

Proof: Suppose that $a, b \in A$ commute. Then the binomial theorem holds for a, b; that is

$$(a + b)^\ell = \sum_{\imath=0}^{\ell} \binom{\ell}{\imath} a^{\ell-\imath} b^\imath$$

for all $\ell \geq 0$. Suppose $a^m = 0 = b^n$, where m, n are positive integers. Then $(a + b)^{m+n-1} = 0$ as for each $0 \leq i \leq m + n$ one of $a^{m+n-1-i} = 0$ or $b^i = 0$ since either $n \leq m + n - 1 - i$ or $m \leq i$. Note that $b^n = 0$ implies that $(\alpha b)^n = 0$ for all $\alpha \in F$. Part (a) now follows. Part (b) follows from the formula $(ab)^n = a^n b^n$, where $n \geq 0$ and $a, b \in A$ commute, and (1.1). Part (c) follows from parts (a) and (b).

Part (d) follows from parts (a) and (b) as well. Let $b \in \mathcal{N}$. Then $b^n = 0$ for some $n \geq 1$ and

$$(1-b)(1+b+\cdots+b^{n-1}) = 1 - b^n = 1.$$

Replacing b with $-b$ we deduce $(1+b)(1+c) = (1+c)(1+b) = 1$ for some $c \in \mathcal{N}$. $\square$

For the exponential function to be meaningful we will need for reciprocals of factorials to be defined in F. *Thus for the remainder of this section we will assume that the characteristic of F is* 0. Recall that the exponential function of Calculus can be represented by the power series

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

for all real numbers x, and the natural logarithm function is represented by the power series

$$\ln(x) = -\sum_{n=1}^{\infty} \frac{(1-x)^n}{n}$$

for all $0 < x < 2$. Regard $\exp(x)$ and $\ln(x)$ as formal power series in indeterminate x.

Suppose that $a \in \mathcal{N}$. Then $a^n = 0$ for some $n > 0$. Therefore the formal sum

$$\exp(a) = \sum_{n=0}^{\infty} \frac{a^n}{n!}$$

is meaningful since $a^n = a^{n+1} = \cdots = 0$ means that it can be regarded as a *finite* sum. Note that $\exp(a) \in \mathcal{U}$ by parts (a) and (b) of Lemma 1.2.7. Observe that

$$\exp(0) = 1 \tag{1.14}$$

and:

$$\text{If } a, b \in \mathcal{N} \text{ commute then } \exp(a+b) = \exp(a)\exp(b). \tag{1.15}$$

A consequence of the two preceding equations is that:

$$\text{If } a \in \mathcal{N} \text{ then } \exp(a) \text{ is invertible and } \exp(-a) = \exp(a)^{-1}. \tag{1.16}$$

Now suppose that $a \in \mathcal{U}$. Then $1-a$ is nilpotent. Therefore the formal sum

$$\ln(a) = -\sum_{n=1}^{\infty} \frac{(1-a)^n}{n}$$

is meaningful since $(1-a)^n = (1-a)^{n+1} = \cdots = 0$ for some $n > 0$ means that it can be regarded as a *finite* sum. Note that $\ln(a) \in \mathcal{N}$ by parts (a) and (b) of Lemma 1.2.7. It is a nice exercise[1] to show that

$$\exp : \mathcal{N} \longrightarrow \mathcal{U} \text{ and } \ln : \mathcal{U} \longrightarrow \mathcal{N} \text{ are inverses.} \tag{1.17}$$

Now let A be any algebra over F and suppose that $D \in \mathrm{Der}(A)$. Then D belongs to the associative algebra with unity $\mathrm{End}(A)$. Suppose that D is nilpotent. Then $\exp(D)$ is a unipotent element of $\mathrm{End}(A)$. Using the Leibnitz Rule (1.6) we calculate

$$\begin{aligned}
\exp(D)(ab) &= \sum_{n=0}^{\infty} \frac{D^n(ab)}{n!} \\
&= \sum_{n=0}^{\infty} \left(\sum_{\imath=0}^{n} \binom{n}{\imath} \frac{D^{n-\imath}(a)D^{\imath}(b)}{n!} \right) \\
&= \sum_{n=0}^{\infty} \left(\sum_{\imath=0}^{n} \frac{D^{n-\imath}(a)D^{\imath}(b)}{(n-\imath!)\imath!} \right) \\
&= \sum_{\jmath,\imath=0}^{\infty} \frac{D^{\jmath}(a)D^{\imath}(b)}{\jmath!\imath!} \\
&= \left(\sum_{\jmath=0}^{\infty} \frac{D^{\jmath}(a)}{\jmath!} \right) \left(\sum_{\imath=0}^{\infty} \frac{D^{\imath}(b)}{\imath!} \right) \\
&= (\exp(D)(a))(\exp(D)(b))
\end{aligned}$$

for all $a, b \in A$ which shows that $\exp(D)$ is an algebra automorphism of A. As an exercise the reader is challenged (see Exercise 1.2.9 below) to show that if ϕ is a unipotent algebra automorphism of A then $\ln(\phi)$ is a nilpotent derivation of A. Thus exp and ln induce bijective correspondences

[1] Exercise §8 ¶¶2. Godement, *Algebra*, Herman, Houghton Mifflin Company, Boston, 1963. The symbol ¶¶ means an exercise for the very brave.

between the set of nilpotent derivations of A and the set of unipotent algebra automorphisms of A.

Let us finally turn to Lie algebras. Let $L = gl(V) = \mathcal{L}(\mathrm{End}(V))$, where V is a vector space over F, and let $x \in \mathrm{End}(V)$. Recall that $\ell_x, r_x : \mathrm{End}(V) \longrightarrow \mathrm{End}(V)$ defined by $\ell_x(y) = x \circ y$ and $r_x(y) = y \circ x$ for all $y \in \mathrm{End}(V)$ are commuting endomorphisms of $\mathrm{End}(V)$ and $\mathrm{ad}\, x = \ell_x - r_x$.

Now suppose that x is nilpotent. Then ℓ_x, r_x, and $\mathrm{ad}\, x$ are nilpotent; see Exercise 1.1.13. Thus

$$\exp(\mathrm{ad}\, x) = \exp(\ell_x - r_x) = \exp(\ell_x) \circ \exp(-r_x) = \ell_{\exp(x)} \circ r_{\exp(-x)} = \ell_{\exp(x)} \circ r_{(\exp(x))^{-1}}$$

which means

$$(\exp(x)) \circ y \circ (\exp(x))^{-1} = (\exp(\mathrm{ad}\, x))(y) \tag{1.18}$$

for all $y \in gl(V)$.

Let L be *any* Lie algebra over F. Then $\exp(\mathrm{ad}\, x)$ is an algebra automorphism of L whenever the derivation $\mathrm{ad}\, x$ is nilpotent. The subgroup of $\mathrm{Aut}(L)$ generated by these automorphisms is the subgroup $\mathrm{Aut}_{inner}(L)$ of *inner automorphisms of* L. For any $x \in L$ and any algebra homomorphism $\phi : L \longrightarrow L$ observe that

$$\mathrm{ad}\, \phi(x) \circ \phi = \phi \circ \mathrm{ad}\, x.$$

Therefore $\phi \circ (\mathrm{ad}\, x) \circ \phi^{-1} = \mathrm{ad}\,(\phi(x))$ for all $x \in L$ and $\phi \in \mathrm{Aut}(L)$ which shows that $\mathrm{Aut}_{inner}(L)$ is a normal subgroup of $\mathrm{Aut}(L)$.

Now let us assume that L is a subalgebra of $gl(V)$ for some vector space V over F and suppose $x \in L$ is a nilpotent endomorphism of V. Then $\mathrm{ad}\, x$ is nilpotent, and therefore $\mathrm{ad}_L\, x = (\mathrm{ad}\, x)|_L$ is as well. By (1.18) we calculate for $y \in L$ that

$$\exp(\mathrm{ad}_L\, x)(y) = \exp(\mathrm{ad}\, x)|_L(y) = (\exp(\mathrm{ad}\, x))(y) = (\exp(x)) \circ y \circ (\exp(x))^{-1}.$$

Thus and $(\exp(x)) \circ L \circ (\exp(x))^{-1} = L$ and $\exp(\mathrm{ad}_L\, x)$ is the associative algebra (and hence Lie) algebra inner automorphism determined by the restriction $u = \exp(x)|_L$.

In the following exercises we will regard the multiplication of an algebra A over F as a linear map $\mathsf{m} : A \otimes_F A \longrightarrow A$.

Exercise 1.2.8 Show that a linear endomorphism $f : A \longrightarrow A$ is an algebra map if and only if $f \circ \mathsf{m} = \mathsf{m} \circ (f \otimes f)$. (Compare with part (a) of Exercise 1.1.10).

Exercise 1.2.9 Suppose F has characteristic 0 and f is a unipotent algebra automorphism of A.

(a) Show that $f \otimes f$ is a unipotent endomorphism of $A \otimes_F A$. [Hint: Write $f \otimes f - \mathrm{Id}_A \otimes \mathrm{Id}_A$ as the sum of two commuting endomorphisms of $A \otimes_F A$.]

(b) Show that $(\mathrm{Id}_A - f) \circ \mathsf{m} = \mathsf{m} \circ (\mathrm{Id}_A \otimes \mathrm{Id}_A - f \otimes f)$.

(c) Show that $\ln(f) \circ \mathsf{m} = \mathsf{m} \circ \ln(f \otimes f)$.

(d) Show that $\mathrm{Id}_A \otimes \ln(f) + \ln(f) \otimes \mathrm{Id}_A = \ln(f \otimes f)$. [Hint: Show that left and right hand sides of the equation are nilpotent endomorphisms of $A \otimes_F A$. Recall that $\exp : \mathcal{N} \longrightarrow \mathcal{U}$ is one-one.]

(e) Show that $\ln(f)$ is a derivation of A. [Hint: See part (a) of Exercise 1.1.10.]

1.3 Solvable and Nilpotent Lie algebras

1.3.1 Solvability

Let L be a Lie algebra over F and suppose that S is a subspace of L. Define

$$S^{(0)} = S \qquad \text{and} \qquad S^{(n+1)} = [S^{(n)}\, S^{(n)}]$$

for all $n \geq 0$. The following are easily established by induction:

(a) $(S^{(m)})^{(n)} = S^{(m+n)}$ for all $m, n \geq 0$.

(b) If T is a subset of S then $T^{(n)} \subseteq S^{(n)}$ for all $n \geq 0$.

(c) If $f : L \longrightarrow L'$ is a map of Lie algebras then $f(S^{(n)}) = f(S)^{(n)}$ for all $n \geq 0$.

(d) If S is a subalgebra of L then $S^{(n)}$ is an ideal of S for all $n \geq 0$ and $S^{(0)} \supseteq S^{(1)} \supseteq S^{(2)} \supseteq \cdots$.

Comments on the proposition.

Remark 1.3.1 *Reformulate part (b) to read "If $f : L \longrightarrow L'$ is a map of Lie algebras then L is solvable if and only if* $\operatorname{Im} f$ *and* $\operatorname{Ker} f$ *are solvable.".*

Remark 1.3.2 *The proof of part (c) is not so involved in light of the reformulation of part (b).*

Suppose that I, J are solvable ideals of L. Then the (projection) map of Lie algebras $\pi : L \longrightarrow L/J$ induces a map of Lie algebras $f : I + J \longrightarrow (I + J)/J$ given by $f(x) = x + J$ for all $x \in I + J$. Since $f(I) = \operatorname{Im} f$ and $J = \operatorname{Ker} f$ are solvable, $I + J$ is solvable.

1.3.2 Nilpotency

Let L be a Lie algebra over F and suppose that S is a subspace of L. Define

$$S^0 = S \qquad \text{and} \qquad S^{n+1} = [S\, S^n]$$

for all $n \geq 0$. The following are easily established by induction. Note the parallels with (a) – (d) of the preceding section.

(b') If T is a subset of S then $T^n \subseteq S^n$ for all $n \geq 0$.

(c') If $f : L \longrightarrow L'$ is a map of Lie algebras then $f(S^n) = f(S)^n$ for all $n \geq 0$.

Suppose that S is a subalgebra of L. Then:

(d') For all $n \geq 0$ the subspace S^n is an ideal of S, $S^{(n)} \subseteq S^n$, and $S^0 \supseteq S^1 \supseteq S^2 \supseteq \cdots$.

(a') $(S^m)^n \subseteq S^{m+n}$ for all $m, n \geq 0$.

Comments on the proposition.

Remark 1.3.3 *Concerning part (b): If I is a nilpotent ideal of L and L/I is also nilpotent, it does not necessarily follow that L is nilpotent.*

For example, let L the non-abelian two-dimensional Lie algebra over F. Then L has basis $\{x, y\}$ and the product is determined by $[x\, y] = y$. The span $I = Fy$ is a one-dimensional ideal of L. Note that $L^1 = L^2 = \cdots = I$ and thus L is not nilpotent; however L/I and I are. Apropos of item a') above, note that $(L^1)^1 = (0)$ is a proper subspace of $L^{1+1} = I$.

1.3.3 Proof of Engel's Theorem

Remark 1.3.4 *Let L be any Lie algebra and K be a subalgebra of L. If $x \in \mathrm{N}_L(K)$ then $K + Fx$ is a subalgebra of L.*

Remark 1.3.5 *The proof of the Theorem holds verbatum with L finite-dimensional and V* **any** *vector space over F.*

Concerning the notation in paragraph 3, page 13. Let L be a (Lie) subalgebra of $gl(V)$ and let K be an ideal of L. Note that

$$W = \{v \in V \mid K(v) = (0)\}.$$

For a subset S of L define $S(v) = \{s(v) \mid s \in S\}$.

We wish to show that W is L-invariant, that is $y(W) \subseteq W$, or equivalently $y(v) \in W$ for all $y \in L$ and $v \in W$. Let $y \in L$ and $v \in W$. Then $y(v) \in W$ if and only if $x(y(v)) = 0$ for all $x \in K$. For $x \in K$ the product $[x\,y] \in K$ since K is an ideal of L. The last equation follows from the calculation

$$0 = [x\,y](v) = x(y(v)) - y(x(v)) = x(y(v)) - y(0) = x(y(v)).$$

Thus W is L-invariant.

Concerning paragraph 1 on page 13 and the proof of the Corollary, let T be a linear endomorphism of a vector space V over F and suppose that W is a T-invariant subspace of V. Then $\overline{T}$ is a well-defined linear endomorphism of the quotient space V/W, where

$$\overline{T}(v + W) = T(v) + W$$

for all $v \in V$.

Chapter 2

Semisimple Lie Algebras

Below are comments relevant to various sections in the text. They are meant to clarify, amplify, or generalize material in the text. Exercises are optional.

2.1 Theorems pf Lie and Cartan

2.1.1 Lie's Theorem

There are several ideas in the proof of the Theorem of the section which should be highlighted. First of all let A be an associative algebra over the field F and let $a \in A$. Then $f = \operatorname{ad} a$ is a derivation of the Lie algebra $\mathcal{L}(A)$, that is

$$f([x\,y]) = [x\,f(y)] + [f(x)\,y]$$

for all $x, y \in A$. Note that f is also a derivation of the *associative algebra* A, that is

$$f(xy) = xf(y) + f(x)y$$

for all $x, y \in A$. The last equation is equivalent to

$$[a\,xy] = x[a\,y] + [a\,x]y,$$

or $a(xy) - (xy)a = x(ay - ya) + (ax - xa)y$, for all $x, y \in A$. Since f is a derivation of A the formula

$$f(a_1 \cdots a_n) = \sum_{i=1}^{n} a_1 \cdots f(a_i) \cdots a_n$$

holds for all $a_1, \cdots, a_n \in A$. In particular

$$[a\,x^m] = \sum_{i=0}^{m-1} x^i [a\,x] x^{m-1-i} \tag{2.1}$$

for all $m \geq 1$ and $x \in A$.

Now let V be a vector space over F. Let K be a non-empty subset of $\mathrm{End}(V)$, and let $\lambda : K \longrightarrow F$ be a function. Set

$$V_\lambda = \{v \in V \,|\, y(v) = \lambda(y)v \text{ for all } y \in K\}.$$

Since $V_\lambda = \bigcap_{y \in K} \mathrm{Ker}(y - \lambda(y)\mathrm{I}_V)$ it follows that V_λ is a subspace of V.

Suppose that $x, y \in gl(V)$. Since $yx = [y\,x] + xy$ we have

$$y(x(v)) = [y\,x](v) + x(y(v)) \tag{2.2}$$

for all $v \in V$.

Suppose that $x \in gl(V)$ and K is invariant under $\mathrm{ad}\,x$, that is $[y\,x] = -\mathrm{ad}\,x(y) \in K$ for all $y \in K$. Then V_λ is invariant under x if and only if $y(x(v)) = \lambda(y)x(v)$ for all $y \in K$ and $v \in V_\lambda$. Thus V_λ is invariant under x if and only if $\lambda([y\,x]) = 0$ by (2.2).

A major portion of the proof of the Theorem boils down to the following lemma. Note that the definition of W_m therein differs from the one on page 16.

Lemma 2.1.1 *Let V be a vector space over the field F, let $x \in gl(V)$, and let K be a non-empty subset of $gl(V)$ invariant under* $\mathrm{ad}\,x$. *Suppose that $\lambda : K \longrightarrow F$ is a function and $v \in V_\lambda$. Set $W_{-1} = (0)$ and let W_m be the span of $\{v, x(v), \ldots, x^m(v)\}$ for $m \geq 0$. Then $y(x^m(v)) - \lambda(y)x^m(v) \in W_{m-1}$ for all $y \in K$ and $m \geq 0$.*

Proof: The conclusion of the lemma is true for $m = 0$ by the definition of V_λ. Suppose that $m > 0$ and the conclusion of the lemma is true for non-negative integers less than or equal to $m - 1$. Then $y(W_{m-1}) \subseteq W_{m-1}$ for all $y \in K$. Let $y \in K$. By (2.2) we have

$$y(x^m(v)) = [y\,x^m](v) + x^m(y(v)) = [y\,x^m](v) + \lambda(y)x^m(v)$$

and thus

$$y(x^m(v)) - \lambda(y)x^m(v) = [y\,x^m](v).$$

Since $[y\,x] \in K$ by assumption, we can use (2.1) and our induction hypothesis to calculate

$$\begin{aligned}
[y\,x^m](v) &= \sum_{i=0}^{m-1} x^i([y\,x](x^{m-1-i}(v))) \\
&\in \sum_{i=0}^{m-1} x^i([y\,x](W_{m-1-i})) \\
&\subseteq \sum_{i=0}^{m-1} x^i(W_{m-1-i}) \\
&\subseteq W_{m-1}.
\end{aligned}$$

□

Here is a proof of Corollary C which does not involve Engel's Theorem.

PROOF: Let $L_0 \subseteq L_1 \subseteq \cdots \subseteq L_n = L$ be a flag of ideals for L. Choose a basis $\{x_1, \ldots, x_n\}$ for L such that $\{x_1, \ldots, x_i\}$ is a basis for L_i for $1 \leq i \leq n$. Use this basis construct an isomorphism of associative algebras $\mathrm{End}(L) \simeq \mathrm{M}_n(F)$ in the usual way.

Let $f : L \longrightarrow gl(n, F)$ be the composite Lie algebra maps $L \xrightarrow{\pi} gl(L) \simeq gl(n, F)$, where π is the adjoint representation of L. Observe that $\mathrm{Im}\, f \subseteq t(n, F)$ and $\mathrm{Ker}\, f = Z(L)$. Since

$$\mathcal{L} = [L\,L]/(Z(L) \cap [L\,L]) \simeq f([L\,L]) = [f(L)\, f(L)] \subseteq n(n, F)$$

it follows that $\mathcal{L}$ is nilpotent. Since $Z(L) \cap [L\,L] \subseteq Z([L\,L])$ the quotient $[L\,L]/Z([L\,L])$ is a quotient of $\mathcal{L}$ and is therefore nilpotent. Thus $[L\,L]$ is nilpotent. □

Exercise 2.1.2 Assume the hypothesis of the preceding lemma. Although an exact expression for $y(x^m(v))$ is not needed for the proof of Lie's Theorem, it is not too difficult to compute it.

Let $\delta = -\mathrm{ad}\, x$; thus $\delta(z) = [z\,x]$ for all $z \in gl(V)$. We will show that

$$y(x^m(v)) = \sum_{i=0}^{m} \left[\begin{array}{c} m \\ i \end{array} \right] \lambda(\delta^{m-i}(y)) x^i(v)$$

for all $m \geq 0$.

(a) Suppose that $\left[\begin{array}{c} m \\ i \end{array}\right]$ are non-negative integers for all $m \geq i \geq 0$ which satisfy

$$\left[\begin{array}{c} m \\ m \end{array}\right] = 1 \qquad \text{for } m \geq 0$$

and

$$\left[\begin{array}{c} m \\ i \end{array}\right] = \sum_{j=0}^{i} \left[\begin{array}{c} m-1-i-j \\ j \end{array}\right] \qquad \text{for } 0 \leq i < m.$$

Show that

$$y(x^m(v)) = \sum_{i=0}^{m} \left[\begin{array}{c} m \\ i \end{array}\right] \lambda(\delta^{m-i}(y))x^i(v)$$

for all $m \geq 0$.

(b) Show that $\left[\begin{array}{c} m \\ i \end{array}\right] = \binom{m}{i}$ for all $0 \leq i \leq m$. [Hint: Recall that the binomial coefficients can be defined recursively.]

Exercise 2.1.3 Suppose that the field F has characteristic zero and is not algebraically closed. Show that there is a finite-dimensional vector space V over F and a solvable subalgebra L of $gl(V)$ which does not satisfy the conclusion of Lie's Theorem. [Hint: Consider one-dimensional Lie algebras.]

Exercise 2.1.4 Here we examine Lie's Theorem in positive characteristic. Suppose that F is an algebraically closed field of characteristic $p > 0$.

(a) Show that the conclusion of Lie's Theorem is true if $\operatorname{Dim} V < p$.

(b) When $\operatorname{Dim} V = p$ find an example of a solvable Lie subalgebra of $gl(V)$ such that the conclusion of Lie's Theorem is false. [Hint: Let $\{v_i\}_{i \in \boldsymbol{Z}_p}$ be a basis for V and let L be the subalgebra of $gl(V)$ generated by x, y, where

$$x(v_i) = v_{i-1} \qquad \text{and} \qquad y(v_i) = iv_i$$

for all $i \in \boldsymbol{Z}_p$. Note that $[x\,y] = x$ and that x, y have no common eigenvector.]

2.1.2 Jordan-Chevalley decomposition

See "Decomposition of Operators" which is available on the course home page.

2.1.3 Cartan's Criterion

There is some interesting mathematics involved in the proof of the Lemma of the section. Assume that F has *any* characteristic and that the characteristic polynomial of x splits over F.

Let $\mathcal{V}$ be a vector space over F and let $\mathcal{U}$, $\mathcal{W}$ be subspaces of $\mathcal{V}$ such that $\mathcal{U} \subseteq \mathcal{W}$. Let

$$\mathrm{M}(\mathcal{W},\mathcal{U}) = \{T \in \mathrm{End}(\mathcal{V}) \,|\, T(\mathcal{W}) \subseteq \mathcal{U}\}.$$

Then $\mathrm{M}(\mathcal{W},\mathcal{U})$ is a subspace of $\mathrm{End}(\mathcal{V})$ closed under composition. Thus if $T \in \mathrm{M}(\mathcal{W},\mathcal{U})$ and $f(x) \in F[x]$ is a polynomial with *zero constant term* then $f(T) \in \mathrm{M}(\mathcal{W},\mathcal{U})$. Note that $\mathrm{I}_{\mathcal{V}} \in \mathrm{M}(\mathcal{W},\mathcal{U})$ if and only if $\mathcal{U} = \mathcal{W}$.

Suppose that $T \in \mathrm{M}(\mathcal{W},\mathcal{U})$ and the characteristic polynomial of T splits over F (as is the case when F is algebraically closed). Then the nilpotent and semisimple parts $T_n, T_s \in \mathrm{M}(\mathcal{W},\mathcal{U})$ as well since they are polynomials in T with no zero constant term.

Now we assume the hypothesis of the Lemma. Let $\mathcal{V} = gl(V)$, $\mathcal{U} = A$, and $\mathcal{W} = B$. Then

$$M = \{x \in gl(V) \,|\, \mathrm{ad}\, x \in \mathrm{M}(\mathcal{U}, \mathcal{W})\}.$$

Let $x \in M$. Suppose $y \in \mathcal{V}$ and x, y commute. Then x_n and y commute which means $x_n \circ y$ is nilpotent. Therefore

$$0 = \mathsf{Tr}(x \circ y) = \mathsf{Tr}(x_s \circ y) + \mathsf{Tr}(x_n \circ y) = \mathsf{Tr}(x_s \circ y). \tag{2.3}$$

Let $\{v_1, \ldots, v_n\}$ be a basis of eigenvectors for x_s and let $\lambda_1, \ldots, \lambda_n \in F$ satisfy $x_s(v_i) = \lambda_i v_i$ for all $1 \leq i \leq n$. For each $1 \leq i, j \leq n$ define $e_{i\,j} \in \mathrm{End}(V)$ by

$$e_{i\,j}(v_k) = \delta_{j,k} v_i h$$

for all $1 \leq k \leq n$. Note that $\{e_{i\,j}\}_{1 \leq i,j \leq n}$ is a basis for $\mathrm{End}(V)$ and

$$e_{i\,j} \circ e_{k\,\ell} = \delta_{j,k} e_{i\,\ell}$$

for all $1 \leq i, j, k, \ell \leq n$. Also note that

$$x_s = \sum_{i=1}^{n} \lambda_i e_{i\,i} = \lambda_1 e_{1\,1} + \cdots + \lambda_n e_{n\,n}$$

and

$$\mathrm{ad}\, x_s(e_{i\,j}) = (\lambda_i - \lambda_j) e_{i\,j}$$

for all $1 \leq i, j \leq n$.

Now suppose $\alpha_1, \ldots, \alpha_n \in F$ satisfy

$$\lambda_i - \lambda_j = \lambda_{i'} - \lambda_{j'} \qquad \text{implies} \qquad \alpha_i - \alpha_j = \alpha_{i'} - \alpha_{j'}. \tag{2.4}$$

Observe that (2.4) implies

$$\lambda_i = \lambda_j \qquad \text{implies} \qquad \alpha_i = \alpha_j. \tag{2.5}$$

Let

$$y = \sum_{i=1}^{n} \alpha_i e_{i\,i} = \alpha_1 e_{1\,1} + \cdots + \alpha_n e_{n\,n}.$$

By virtue of (2.5) there exists a polynomial $f(x) \in F[x]$ such that $f(\lambda_i) = \alpha_i$ for all $1 \leq i \leq n$ by LaGrange interpolation. Thus $y = f(x_s)$ which means that y commutes with x. Notice that

$$\operatorname{ad} y(e_{i\,j}) = (\alpha_i - \alpha_j)e_{i\,j}$$

for all $1 \leq i, j \leq n$. Using LaGrange interpolation again, by (2.4) there exists a polynomial $g(x) \in F[x]$ such that $g(\lambda_i - \lambda_j) = \alpha_i - \alpha_j$ for all $1 \leq i, j \leq n$. As $g(\operatorname{ad} y)(e_{i\,j}) = g(\lambda_i - \lambda_j)e_{i\,j}$ for all $1 \leq, j \leq n$ it follows that $\operatorname{ad} y = g(\operatorname{ad} x_s)$. Now $g(x)$ has zero constant term since $g(0) = g(\lambda_1 - \lambda_1) = \alpha_1 - \alpha_1 = 0$ by (2.5). Therefore $y \in M$ and hence

$$0 = \mathsf{Tr}(x_s \circ y) = \sum_{i=1} \lambda_i \alpha_i. \tag{2.6}$$

To complete the proof, regard F as a vector space over its prime field which is the field of rational numbers $\mathbf{Q}$. Let E be the span of $\lambda_1, \cdots, \lambda_n$ over $\mathbf{Q}$ and suppose $\mathsf{f} : F \longrightarrow \mathbf{Q}$ be any $\mathbf{Q}$-linear functional. Then (2.4) holds for $\mathsf{f}(\lambda_1)\ \ldots, \mathsf{f}(\lambda_n)$. By the preceding equation

$$0 = \sum_{i=1} \lambda_i \mathsf{f}(\lambda_i)$$

from which $0 = \sum_{i=1} \mathsf{f}(\lambda_i)^2$ follows, since f is linear, and therefore $\mathsf{f}(\lambda_i) = 0$ for all $1 \leq i \leq n$. Thus $\lambda_1 = \cdots = \lambda_n$ which means that $x_s = 0$. We have shown $x = x_n$ and is nilpotent.

2.2 Killing form

2.2.1 Criterion for semisimplicity

Basic to this section are bilinear forms. The most important one for our purposes the Killing form. We begin by discussing a context for the Killing form and then derive some elementary properties of non-degenerate bilinear forms on finite-dimensional vector spaces in general.

Let V be any vector space over F and suppose that $\beta : V{\times}V \longrightarrow F$ is a bilinear form. Then β is *symmetric* if $\beta(u, v) = \beta(v, u)$ for all $u, v \in V$. If V is an algebra over F then β is *associative* if $\beta(uv, w) = \beta(u, vw)$ for all $u, v, w \in V$.

Suppose that β is symmetric. Then the *radical* $\mathrm{Rad}\,\beta$ *of* β is the set

$$\mathrm{Rad}\,\beta = \{u \in V \mid \beta(u, v) = 0 \ \ \forall \ \ v \in V\} = \{v \in V \mid \beta(u, v) = 0 \ \ \forall \ \ u \in V\}.$$

Evidently $\mathrm{Rad}\,\beta$ is a subspace of V. The bilinear form β is *non-degenerate* if $\mathrm{Rad}\,\beta = (0)$.

Now suppose that V is finite-dimensional. Then

$$\boldsymbol{\beta}(x, y) = \mathsf{Tr}(x{\circ}y)$$

for all $x, y \in \mathrm{End}\,(V)$ defines a symmetric associative bilinear form on the (associative) algebra $\mathrm{End}\,(V)$ in terms of the trace function. We have noted that $[x\,y{\circ}z] = y{\circ}[x\,z] + [x\,y]{\circ}z$ for all $x, y, z \in \mathrm{End}\,(V)$; that is the derivation $\mathrm{ad}\,x$ of $gl(V)$ is also a derivation of the associative algebra $\mathrm{End}\,(V)$. Since the trace function vanishes on commutators, we deduce from the preceding equation (with x and y interchanged) that

$$\mathsf{Tr}([x\,y]{\circ}z) = \mathsf{Tr}(x{\circ}[y\,z]) \tag{2.7}$$

for all $x, y, z \in \mathrm{End}(V)$. Therefore

$$\boldsymbol{\beta}([x\,y], z) = \boldsymbol{\beta}(x, [y\,z]) \tag{2.8}$$

for all $x, y, z \in gl(V)$ which means that $\boldsymbol{\beta}$ is also a symmetric associative bilinear form on the Lie algebra $gl(V)$.

Now let L be a Lie algebra over F and suppose that $\pi : L \longrightarrow gl(V)$ is a representation of L. Since π is a map of Lie algebras, it follows by (2.8) that $\beta_\pi : L{\times}L \longrightarrow F$ defined by

$$\beta_\pi(x, y) = \boldsymbol{\beta}(\pi(x), \pi(y))$$

for all $x, y \in L$ is a symmetric associative bilinear form. The associativity of β_π means that $\operatorname{Rad}\beta_\pi$ is an ideal of L. Observe that:

Remark 2.2.1 $\operatorname{Ker}\pi \subseteq \operatorname{Rad}\beta_\pi$.

Suppose further that L is finite-dimensional. When π is the adjoint representation of L then $\beta_\pi = \kappa$ is the Killing form of L. Thus

$$\kappa(x, y) = \mathsf{Tr}(\operatorname{ad} x \circ \operatorname{ad} y)$$

for all $x, y \in L$.

Now we continue with a discussion of symmetric bilinear forms $\beta : V \times V \longrightarrow F$, where V is finite-dimensional. Note that the function

$$\beta_\ell : V \longrightarrow V^*$$

given by $\beta_\ell(u)(v) = \beta(u, v)$ for all $u, v \in V$ is linear. Non-singularity of β can be expressed in terms of β_ℓ.

Lemma 2.2.2 *Let V be a finite-dimensional vector space over the field F and let $\beta : V \times V \longrightarrow F$ be a symmetric bilinear form. Then the following are equivalent:*

(a) *β is non-singular.*

(b) *β_ℓ is one-one.*

(c) *β_ℓ is onto.*

(d) *β_ℓ is an isomorphism.*

(e) *There exists a basis $\{v_1, \dots, v_n\}$ for V such that the matrix $(\beta(v_i, v_j))$ is invertible.*

(f) *The matrix of part (e) is invertible for all bases for V.*

PROOF: By the Rank–Nullity Theorem $\operatorname{Dim}\operatorname{Ker}\beta_\ell + \operatorname{Dim}\operatorname{Im}\beta_\ell = \operatorname{Dim} V$. Since $\operatorname{Dim} V = \operatorname{Dim} V^*$ and both are finite, parts (b)–(d) are equivalent.

Suppose that $\operatorname{Dim} V = n$ and let $b : F^n \times F^n \longrightarrow F$ be the "standard inner product" defined by

$$b(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}) = x_1 y_1 + \cdots + x_n y_n$$

for all $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in F^n$. Observe that b is non-singular.

Now let $\{v_1, \ldots, v_n\}$ be a basis for V and consider $S \in \mathrm{M}_n(F)$ defined by

$$S = (\beta(v_{i,j})).$$

Suppose that $u = x_1v_1 + \cdots + x_nv_n, v = y_1v_1 + \cdots + y_nv_n \in V$. A straightforward calculation shows that

$$\beta(u, v) = b(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, S\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}).$$

Thus since b is non-singular, $v \in \mathrm{Rad}\,\beta$ if and only if $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ is in the null-space of S. Since S is invertible if and only if its null space is (0), parts (a), (e), and (f) are equivalent. $\Box$

A few more technical details about the symmetric bilinear form $\beta : V \times V \longrightarrow F$. For a subspace U of V set

$$U^\perp = \{v \in V \,|\, \beta(v, U) = (0)\}.$$

Since β is symmetric $U^\perp = \{v \in V \,|\, \beta(U, v) = (0)\}$ as well. Observe that $U^\perp = \cap_{u \in U} \mathrm{Ker}\,\beta_\ell(u)$ and is therefore a subspace of V.

Suppose that $V = L$ is a Lie algebra and β is also associative (as is the case when L is finite-dimensional and β is the Killing form of L). Then if I is an ideal of L, the calculation

$$\beta([L\, I^\perp], I) = \beta([I^\perp\, L], I) = \beta(I^\perp, [L\, I]) \subseteq \beta(I^\perp, I) = (0)$$

shows that $I^\perp$ is an ideal of L as well.

Here are few basic facts about the Killing form which will make the proof of the main theorems of this section and the next fairly straightforward.

Lemma 2.2.3 *Let L be a finite-dimensional Lie algebra over the field F and suppose that I, J are ideals of L. Then:*

(a) *If I is abelian then $I \subseteq \operatorname{Rad}\kappa$. In particular $\mathrm{Z}(L) \subseteq \operatorname{Rad}\kappa$.*

(b) *If $[I\,J] = (0)$ then $\kappa(I, J) = (0)$.*

(c) *$I^{\perp}$ is an ideal of L. Thus $\operatorname{Rad}\kappa = L^{\perp}$ is an ideal of L.*

(d) *Suppose that F is algebraically closed, has characteristic zero, and K is a subalgebra of L. If $\kappa(K, K) = (0)$ then K is solvable. Thus $\operatorname{Rad}\kappa$ is a solvable ideal of L.*

PROOF: Part (a). Suppose $x \in I$ and $y \in L$. Then the calculation

$$(\operatorname{ad} x \circ \operatorname{ad} y)^2(L) = [x\,[y\,[x\,[y\,L]]]] \subseteq [x\,[y\,[x\,L]]] \subseteq [x\,[y\,I]] \subseteq [x\,I] \subseteq [I\,I] = (0)$$

shows that $(\operatorname{ad} x \circ \operatorname{ad} y)^2 = 0$. Therefore $\operatorname{ad} x \circ \operatorname{ad} y$ is nilpotent which means that $\kappa(x, y) = \mathsf{Tr}(\operatorname{ad} x \circ \operatorname{ad} y) = 0$.

Part (b). Let $x \in I$ and $y \in J$. Then

$$(\operatorname{ad} x \circ \operatorname{ad} y)(L) = [x\,[y\,L]] \subseteq [x\,J] \subseteq [I\,J] = (0).$$

Therefore $\operatorname{ad} x \circ \operatorname{ad} y = 0$ which means that $\kappa(x, y) = \mathsf{Tr}(\operatorname{ad} x \circ \operatorname{ad} y) = 0$.

Part (c) was noted more generally above. To see part (d) note the subalgebra $\operatorname{ad} K$ of $gl(L)$ satisfies $\mathsf{Tr}(\mathsf{x} \circ \mathsf{y}) = 0$ for all $\mathsf{x}, \mathsf{y} \in \operatorname{ad} K$. Thus $K/K' \simeq \operatorname{ad} K$ is solvable, where $K' = Z(L) \cap K$. As K' is solvable so is K. $\square$

Exercise 2.2.4 Let $L = gl(n, F)$. The matrix versions for L of the two symmetric associative bilinear forms studied above are the Killing form and the bilinear form $\boldsymbol{\beta}$ defined by $\boldsymbol{\beta}(x, y) = \mathsf{Tr}(x \circ y)$ for all $x, y \in L$. Here we calculate and compare them.

Let $\{e_{i\,j}\}_{1 \le i,j \le n}$ be the standard basis for the underlying vector space $\mathrm{M}_n(F)$ for L. Thus

$$e_{i\,j} e_{k\,\ell} = \delta_{j,k} e_{i,\ell}$$

for all $1 \le i, j, k, \ell \le n$.

(a) Show that

$$(\operatorname{ad} e_{i\,j} \operatorname{ad} e_{k\,\ell})(e_{u\,v}) = \delta_{\ell,u}\delta_{j,k} e_{i\,v} - \delta_{\ell,u}\delta_{i,v} e_{k\,j} - \delta_{k,v}\delta_{j,u} e_{i\,\ell} + \delta_{k,v}\delta_{i,\ell} e_{u\,j}$$

for all $1 \le i, j, k, \ell, u, v \le n$.

(b) Show that the coefficient of $e_{u\,v}$ in the expression in part (a) is

$$\delta_{\ell,u}\delta_{j,k}\delta_{i,u} - \delta_{\ell,u}\delta_{i,v}\delta_{k,u}\delta_{j,v} - \delta_{k,v}\delta_{j,u}\delta_{i,u}\delta_{\ell,v} + \delta_{k,v}\delta_{i,\ell}\delta_{j,v}.$$

(c) Show that

$$\kappa(e_{i\,j}, e_{k\,\ell}) = \mathsf{Tr}(\operatorname{ad} e_{i\,j} \operatorname{ad} e_{k\,\ell}) = 2n\delta_{i,\ell}\delta_{j,k} - 2\delta_{i,j}\delta_{k,\ell}$$

for all $1 \le i, j, k, \ell \le n$.

(d) Show that

$$\boldsymbol{\beta}(e_{i\,j}, e_{k,\ell}) = \mathsf{Tr}(e_{i\,j}e_{k,\ell}) = \delta_{j,k}\delta_{i,\ell}$$

for all $1 \le i, j, k, \ell \le n$.

Let $n = 2$.

(e) Show that

$$\begin{pmatrix} \kappa(e_{1\,1}, e_{1\,1}) & \kappa(e_{1\,1}, e_{1\,2}) & \kappa(e_{1\,1}, e_{2\,1}) & \kappa(e_{1\,1}, e_{2\,2}) \\ \kappa(e_{1\,2}, e_{1\,1}) & \kappa(e_{1\,2}, e_{1\,2}) & \kappa(e_{1\,2}, e_{2\,1}) & \kappa(e_{1\,2}, e_{2\,2}) \\ \kappa(e_{2\,1}, e_{1\,1}) & \kappa(e_{2\,1}, e_{1\,2}) & \kappa(e_{2\,1}, e_{2\,1}) & \kappa(e_{2\,1}, e_{2\,2}) \\ \kappa(e_{2\,2}, e_{1\,1}) & \kappa(e_{2\,2}, e_{1\,2}) & \kappa(e_{2\,2}, e_{2\,1}) & \kappa(e_{2\,2}, e_{2\,2}) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 & -2 \\ 0 & 0 & 4 & 0 \\ 0 & 4 & 0 & 0 \\ -2 & 0 & 0 & 2 \end{pmatrix}.$$

(f) Show that

$$\begin{pmatrix} \boldsymbol{\beta}(e_{1\,1}, e_{1\,1}) & \boldsymbol{\beta}(e_{1\,1}, e_{1\,2}) & \boldsymbol{\beta}(e_{1\,1}, e_{2\,1}) & \boldsymbol{\beta}(e_{1\,1}, e_{2\,2}) \\ \boldsymbol{\beta}(e_{1\,2}, e_{1\,1}) & \boldsymbol{\beta}(e_{1\,2}, e_{1\,2}) & \boldsymbol{\beta}(e_{1\,2}, e_{2\,1}) & \boldsymbol{\beta}(e_{1\,2}, e_{2\,2}) \\ \boldsymbol{\beta}(e_{2\,1}, e_{1\,1}) & \boldsymbol{\beta}(e_{2\,1}, e_{1\,2}) & \boldsymbol{\beta}(e_{2\,1}, e_{2\,1}) & \boldsymbol{\beta}(e_{2\,1}, e_{2\,2}) \\ \boldsymbol{\beta}(e_{2\,2}, e_{1\,1}) & \boldsymbol{\beta}(e_{2\,2}, e_{1\,2}) & \boldsymbol{\beta}(e_{2\,2}, e_{2\,1}) & \boldsymbol{\beta}(e_{2\,2}, e_{2\,2}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(g) Let $x = e_{1\,2}$, $h = e_{1\,1} - e_{2\,2}$, and $y = e_{2\,1}$ be the elements of the standard basis for $sl(2, F)$. Use part (e) to show that

$$\begin{pmatrix} \kappa(x,x) & \kappa(x,h) & \kappa(x,y) \\ \kappa(h,x) & \kappa(h,h) & \kappa(h,y) \\ \kappa(y,x) & \kappa(y,h) & \kappa(y,y) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix}.$$

(h) Continuing with part (g), use part (f) to show that

$$\begin{pmatrix} \boldsymbol{\beta}(x,x) & \boldsymbol{\beta}(x,h) & \boldsymbol{\beta}(x,y) \\ \boldsymbol{\beta}(h,x) & \boldsymbol{\beta}(h,h) & \boldsymbol{\beta}(h,y) \\ \boldsymbol{\beta}(y,x) & \boldsymbol{\beta}(y,h) & \boldsymbol{\beta}(y,y) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

2.2.2 Simple ideals of L

Part (b) of the following lemma is an essential detail used in the proof of the Theorem of this section. We assume a few basic facts about transpose maps.

Let $f : V \longrightarrow W$ be a linear map. Then the linear map $f^* : W^* \longrightarrow V^*$ defined by $f^*(w^*) = w^* \circ f$ for all $w^* \in W^*$ is called the *transpose of* f. Note that f one-one implies f^* is onto and f onto implies that f^* is one-one. If $g : U \longrightarrow V$ is also linear then $(f \circ g)^* = g^* \circ f^*$.

Lemma 2.2.5 *Let V be a finite-dimensional vector space over F, let $\beta : V \times V \longrightarrow F$ is a symmetric bilinear form, and suppose that U is a subspace of V. Then:*

(a) $\mathrm{Dim}\, U + \mathrm{Dim}\, U^{\perp} = \mathrm{Dim}\, V + \mathrm{Dim}\,(U \cap \mathrm{Rad}\,\beta)$.

(b) *Suppose that β is non-singular. Then* $\mathrm{Dim}\, U + \mathrm{Dim}\, U^{\perp} = \mathrm{Dim}\, V$.

PROOF: Part (b) is a direct consequence of part a). To show part (a), we first define a linear map $f : V \longrightarrow U^*$ by $f(v)(u) = \beta(u, v)$ for all $u \in U$ and $v \in V$. Since $\mathrm{Ker}\, f = U^{\perp}$, by the Rank–Nullity Theorem we have

$$\mathrm{Dim}\, U^{\perp} + \mathrm{Dim}\,\mathrm{Im}\, f = \mathrm{Dim}\, V. \tag{2.9}$$

It remains to calculate $\mathrm{Dim}\,\mathrm{Im}\, f$ in terms of dimensions mentioned in the formula of part (a).

Recall that $\mathrm{Ker}\,\beta_\ell = \mathrm{Rad}\,\beta$. Therefore the linear map $\beta_\ell : V \longrightarrow V^*$ induces a one-one linear map $\overline{\beta_\ell} : V/\mathrm{Rad}\,\beta \longrightarrow V^*$ which is defined by $\overline{\beta_\ell}(v + \mathrm{Rad}\,\beta) = \beta_\ell(v)$ for all $v \in V$. Let $\mathsf{c} : V \longrightarrow V^{**}$ be the linear isomorphism defined by $\mathsf{c}(v)(v^*) = v^*(v)$ for all $v \in V$ and $v^* \in V^*$.

The commutative diagram

$$\begin{tikzcd}
V \arrow[r, "\pi"] & V/\mathrm{Rad}\,\beta \\
U \arrow[u, "i"] \arrow[r, "\rho"] & U/(\mathrm{Rad}\,\beta \cap U) \arrow[u, "j"']
\end{tikzcd}$$

where π, ρ are the projections, i is the projection, and $j(u + (U \cap \mathrm{Rad}\,\beta)) = u + \mathrm{Rad}\,\beta$ for all $u \in U$, gives rise to the commutative diagram

$$\begin{tikzcd}
{} \arrow[d] & (V/\mathrm{Rad}\,\beta)^* \arrow[l, "\pi^*"'] \arrow[d, "j^*"] \\
55)V^*25)i^*. & (U/(\mathrm{Rad}\,\beta \cap U))^* \arrow[l, "\rho^*"']
\end{tikzcd}$$

Since j is one-one j^* is onto. Therefore $\mathrm{Im}\,(i^* \circ \pi^*) = \mathrm{Im}\,\rho^*$. Thus $\mathrm{Im}\,\rho^*$ is the image of the composite $i^* \circ \pi^* \circ (\overline{\beta_\ell})^* \circ \mathsf{c}$ since the third and fourth maps are isomorphisms. The calculation

$$\begin{aligned}
\left((i^* \circ \pi^* \circ (\overline{\beta_\ell})^* \circ \mathsf{c})(v)\right)(u) &= \left(((\overline{\beta_\ell} \circ \pi \circ i)^* \circ \mathsf{c})(v)\right)(u) \\
&= \left((\overline{\beta_\ell} \circ \pi \circ i)^* (\mathsf{c}(v))\right)(u) \\
&= (\mathsf{c}(v) \circ \overline{\beta_\ell} \circ \pi \circ i)(u) \\
&= \mathsf{c}(v)\left((\overline{\beta_\ell} \circ \pi \circ i)(u)\right) \\
&= \left(\overline{\beta_\ell} \circ \pi \circ i)(u)\right)(v) \\
&= \beta_\ell(u)(v) \\
&= \beta(u, v) \\
&= f(u)(v)
\end{aligned}$$

for all $u \in U$ and $v \in V$ shows that $f = i^* \circ \pi^* \circ (\overline{\beta_\ell})^* \circ \mathsf{c}$. Since ρ is onto ρ^* is one-one. Thus

$$\begin{aligned}
\mathrm{Dim}\,\mathrm{Im}\,f &= \mathrm{Dim}\,\mathrm{Im}\,\rho^* \\
&= \mathrm{Dim}\,(U/(\mathrm{Rad}\,\beta \cap U))^* \\
&= \mathrm{Dim}\,(U/(\mathrm{Rad}\,\beta \cap U)) \\
&= \mathrm{Dim}\,U - \mathrm{Dim}\,(\mathrm{Rad}\,\beta \cap U);
\end{aligned}$$

the last equation follows by the Rank–Nullity Theorem. Part (a) now follows from this last calculation and (2.9). $\square$

Here is a slight reformulation of the theorem of the section with a proof which allows for an easy induction.

Theorem 2.2.6 *Let L be a finite-dimensional semisimple Lie algebra over an algebraically closed field of characteristic zero. Then:*

(a) *$L = L_1 \oplus \cdots \oplus L_r$ is the direct sum of its simple ideals.*

(b) *Let I be an ideal of L. Then $I = L'_1 \oplus \cdots \oplus L'_r$, where $L_i = L_i$ or $L'_i = (0)$ for all $1 \le i \le r$.*

(c) *Each L_i is a simple Lie algebra.*

(d) *Write $x, y \in L$ as $x = x_1 \oplus \cdots \oplus x_r, y = y_1 \oplus \cdots \oplus y_r$, where $x_i, y_i \in L_i$ for all $1 \le i \le r$. Then*

$$[x\,y] = [x_1\,y_1] \oplus \cdots \oplus [x_r\,y_r]$$

and

$$\kappa_L(x, y) = \kappa_{L_1}(x_1, y_1) + \cdots + \kappa_{L_r}(x_r, y_r).$$

PROOF: If L is simple then there is nothing to prove. Suppose that L is not simple, and let I be any non-zero proper ideal of L.

Let us suppose that there is an ideal J of L such that $I \oplus J = L$. Then J is a non-zero proper ideal of L also. Now $[I\,J] \subseteq I \cap J = (0)$ implies that $[I\,J] = (0)$ and therefore $\kappa(I, J) = (0)$ by part (b) of Lemma 2.2.3. Let $x, y \in L$ and write $x = x' \oplus x'', y = y' \oplus y''$ where $x', y' \in I$ and $x'', y'' \in J$. Then

$$[x\,y] = [x'\,y'] \oplus [x''\,y'']$$

and

$$\kappa_L(x, y) = \kappa_L(x', y') + \kappa_L(x', y'') = \kappa_I(x', y') + \kappa_J(x', y'').$$

Thus an ideal of I or J is an ideal of L, and since κ_L is non-degenerate it follows that κ_I and κ_J are as well. In particular I and J are semisimple Lie algebras.

Suppose that K is a simple ideal of L. Then $[L\,K] \neq (0)$; otherwise $K \subseteq \mathrm{Rad}\kappa = (0)$ by part (b) of Lemma 2.2.3 again. Therefore

$$K = [L\,K] = [I\,K] \oplus [J\,K]$$

is the direct sum of ideals of L. Since K is simple either $[J\,K] = (0)$ or $[I\,K] = (0)$. Our conclusion: $K \subseteq I$ or $K \subseteq J$. At this point the theorem follows by induction on $\mathrm{Dim}\, L$.

It remains to find J. Consider the ideal $J = I^{\perp}$ of L. Since $\kappa(I \cap I^{\perp}, I \cap I^{\perp}) = (0)$, we use part d) of Lemma 2.2.3 to conclude that $I \cap I^{\perp} \subseteq \operatorname{Rad} L = (0)$. We have shown$I \cap I^{\perp} = (0)$. Now $\operatorname{Dim} I + \operatorname{Dim} I^{\perp} = \operatorname{Dim} L$ by part b) of Lemma 2.2.5. Consequently $L = I \oplus I^{\perp} = I \oplus J$. $\square$

Remark 2.2.7 *By virtue of the preceding theorem L is the direct product of simple Lie algebras.*

2.2.3 Inner derivations

Let L be any Lie algebra over the field F and let $\pi : L \longrightarrow gl(L)$ be the adjoint representation of L. Then $\operatorname{Im} \pi = \{\operatorname{ad} x \mid x \in L\}$ is a Lie subalgebra of $gl(L)$; indeed it is a subalgebra of the subalgebra $\operatorname{Der}(L)$ of $gl(L)$.

Let $\delta \in \operatorname{Der}(L)$ and $x \in L$. Then the calculation

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

for all $x, y \in L$ shows that

$$[\delta \operatorname{ad} x] = \operatorname{ad} \delta(x) \tag{2.10}$$

for all $x \in L$. By virtue of (2.10) we have that $\operatorname{Im} \pi = \operatorname{ad} L$ is an ideal of $\operatorname{Der}(L)$.

There seems to be a gap in the proof of the theorem of this section which can easily be fixed. The flow of the proof suggests that $M + I = M + M^{\perp} = D$ is used. As $I \cap M = (0)$ is established, noting that $\operatorname{Rad} \kappa_D \cap M \subseteq \operatorname{Rad} \kappa_M = (0)$ would give $M + I = D$ by part a) of Lemma 2.2.5. Here is a slightly different proof.

Proof: Since L is semisimple, $\mathrm{Z}(L) = (0)$ by part a) of Lemma 2.2.3, and thus the adjoint representation $\pi : L \longrightarrow gl(L)$ is one-one. Two consequences: $M = \operatorname{Im} \pi = \operatorname{ad} L$ is semisimple and, by (2.10), if $\delta \in D$ satisfies $[\delta\, M] = (0)$ then $\delta = 0$.

We next observe that $\operatorname{Rad} \kappa_D \cap M = (0)$. For since M is an ideal of D we have $\kappa_M = \kappa_D|_{M \times M}$. Thus $\operatorname{Rad} \kappa_D \cap M \subseteq \operatorname{Rad} \kappa_M$ and the latter is (0) since M is semisimple.

Let I be an ideal of D which satisfies $I \cap M = (0)$. Since $[I\, M] \subseteq I \cap M = (0)$ it follows that $I = (0)$. Our conclusion: $\mathrm{Rad}\, \kappa_D = (0)$, or equivalently D is semisimple, and consequently $M = D$ by part b) of Theorem 2.2.6. $\square$

2.2.4 Abstract Jordan decomposition

No particular comments.

2.3 Complete reducibility of representations

2.3.1 Modules

There is a lot of background material needed for this section. The material below is an easy adaptation of parts of the theory of modules for associative algebras. A quick read of the following would be a good idea.

Let L be a Lie algebra. Then representations of L and (left) modules for L can be thought of as two different ways of expressing the same idea. A *left L-module* is a vector space V over F together with a map $L \times V \longrightarrow V$ $((x, v) \mapsto x{\cdot}v)$ such that

(M.1) $x{\cdot}(u + v) = x{\cdot}u + x{\cdot}v$ and $x{\cdot}(\alpha u) = \alpha(x{\cdot}u)$,

(M.2) $(x + y){\cdot}v = x{\cdot}v + y{\cdot}v$ and $(\alpha x){\cdot}u = \alpha(x{\cdot}u)$, and

(M.3) $[x\, y]{\cdot}v = x{\cdot}(y{\cdot}v) - y{\cdot}(x{\cdot}v)$

for all $x, y \in L$, $u, v \in V$, and $\alpha \in F$.

Example 2.3.1 *Let $\pi : L \longrightarrow gl(V)$ be a representation of L. Then V is a left L-module with $x{\cdot}v = \pi(x)(v)$ for all $x \in L$ and $v \in V$.*

Suppose that L is a subalgebra of $gl(V)$. Then the inclusion map $i : L \longrightarrow gl(V)$ is a representation of L. Thus:

Example 2.3.2 *Let L be a subalgebra of $gl(V)$. Then V is a left L-module with $x{\cdot}v = x(v)$ for all $x \in L$ and $v \in V$.*

Example 2.3.3 *Let $\pi : L \longrightarrow gl(L)$ be the adjoint representation of L. The left L-module structure on L described in Example 2.3.1 is given by $x{\cdot}v = [x\, v]$ for all $x, v \in L$.*

Example 2.3.4 *Let V be a vector space over F. The rule $x{\cdot}v = 0$ for all $x \in L$ and $v \in V$ gives V a left L-module structure, called the* trivial module structure.

We shall always assume that F has the trivial left L-module structure unless otherwise stated.

Suppose that V is a vector space over F and $L{\times}V \longrightarrow V$ $((x, v) \mapsto x{\cdot}v)$ is a function. Define a function $\pi : L \longrightarrow \mathcal{F}(V)$ from L to the set of all functions $\mathcal{F}(V)$ from V to itself by $\pi(v)(v) = x{\cdot}v$ for all $x \in L$ and $v \in V$. Note that $\mathcal{F}(V)$ is a vector space over F with the usual rules for function addition and scalar product.

Observe that (M.1) is equivalent to saying that $\pi(x)$ is a linear endomorphism of V for all $x \in L$; that is $\pi(x) \in \mathrm{End}(V)$ for all $x \in L$. Note that (M.2) is equivalent to saying that $\pi : L \longrightarrow \mathcal{F}(V)$ is linear, and (M.3) is the same as saying that $\pi([x\,y]) = \pi(x){\circ}\pi(y) - \pi(y){\circ}\pi(x)$ for all $x, y \in L$. We have established, in our careful analysis of the meaning of each module axiom for π, that if (M.1)–(M.3) are satisfied then $\pi(x) \in \mathrm{End}(V)$ for all $x \in L$ and $\pi : L \longrightarrow gl(V)$ is a representation of L.

One comment on the concept of left A-module for any algebra A over F. One would want (M.1) and (M.2) to hold. The formulation of the third axiom evidently would depend on the nature of the multiplication of A; for example if A is associative one would want $(ab){\cdot}v = a{\cdot}(b{\cdot}v)$ for all $a, b \in A$ and $v \in V$.

Let V be a left L-module. A *submodule* of V is a subspace W of V such that $x{\cdot}w \in W$ for all $x \in L$ and $w \in W$. Thus a submodule is a left L-module in its own right with the module structure of V. Note that V and (0) are submodules of V. The module V is *simple*, or *irreducible*, if V has exactly two submodules.

Let $\{V_i\}_{i\in\mathcal{I}}$ by an indexed family of submodules of V. Then $\sum_{i\in\mathcal{I}} V_i$ and $\cap_{i\in\mathcal{I}} V_i$ are submodules of V. Since V is a submodule of V, the intersection property implies that any subset S of V is contained in a unique minimal submodule of V, called the *submodule of V generated by S*. A submodule W of V is *finitely generated* if it W generated by a finite subset of V.

A *map of left L-modules V and W* is a linear map $f : V \longrightarrow W$ such that $f(x{\cdot}v) = x{\cdot}f(v)$ for all $x \in L$ and $v \in V$. Suppose that W is a submodule of V. Then the quotient vector space V/W has a left L-module structure determined by the requirement that the linear projection $\pi : V \longrightarrow V/W$ is a module map. There are the isomorphism theorems for modules to formulate and prove.

If V, W are left L-modules then the tensor product $V \otimes W$ and the vector space $\mathrm{Hom}(V, W)$ have left L-module structures.

Lemma 2.3.5 *Suppose that L is a Lie algebras over the field F and V, W are left L-modules. Then:*

(a) *$V \otimes W$ is a left L-module where $x{\cdot}(v \otimes w) = x{\cdot}v \otimes w + v \otimes x{\cdot}w$ for all $x \in L$, $v \in V$, and $w \in W$.*

(b) *$\mathrm{Hom}(V, W)$ is a left L-module where $(x{\cdot}f)(v) = -f(x{\cdot}v) + x{\cdot}(f(v))$ for all $x \in L$, $f \in \mathrm{Hom}(V, W)$, and $v \in V$.*

Proof: The proof is a nice exercise in definitions and is left to the reader. Apropos of part (b), be sure to show that if $x \in L$ and $f \in \mathrm{Hom}(V, W)$ then $x{\cdot}f \in \mathrm{Hom}(V, W)$. $\Box$

Remark 2.3.6 $f \in \mathrm{Hom}(V, W)$*is a map of left L-modules if and only if* $x{\cdot}f = 0$ *for all $x \in L$; that is if and only if f spans a trivial left L-module.*

Now suppose that V is a left L-module. Then $V^* = \mathrm{Hom}(V, F)$ has a left L-module structure by part b) of the preceding proposition. Note that

$$(x{\cdot}v^*)(v) = -v^*(x{\cdot}v) \tag{2.11}$$

for all $x \in L$, $v^* \in V^*$, and $v \in V$. The module structure defined by (2.11) is called the *contragredient action*. We will assume that V^* is a left L-module with this action unless otherwise stated.

Corollary 2.3.7 *Let L be a Lie algebra over the field F and suppose that V, W are left L-modules. With the module structures of parts (b) and (a) of Proposition 2.3.5 the one-one linear map $\pi : V^* \otimes W \longrightarrow \mathrm{Hom}(V, W)$ defined by $\pi(v^* \otimes w)(v) = v^*(v)w$ for all $v^* \in V^*$, $w \in W$, and $v \in V$ is a map of left L-modules.* $\Box$

A non-zero left L-module V is *completely reducible* if V is the sum of simple submodules. There are several useful equivalent formulations of completely reducible.

Suppose that V is a vector space over F and W is a subspace of V. Then a *projection of V onto W* is a linear map $f : V \longrightarrow W$ such that $f|_W = I_W$.

In this case $V = W \oplus \mathrm{Ker}\, f$. Conversely, if $V = W \oplus W'$ is the direct sum of subspaces then $f : V \longrightarrow W$ defined by $f(w \oplus w') = w$ is a projection of V onto W and $\mathrm{Ker}\, f = W'$.

Proposition 2.3.8 *Let L be a Lie algebra over the field F and suppose that V is a left L-module. Then the following are equivalent:*

(a) *V is completely reducible.*

(b) *V is the direct sum of simple submodules.*

(c) *V is not zero and if W is a submodule of V then $V = W \oplus W'$ for some submodule W' of V.*

(d) *V is not zero and if W is a submodule of V then there exists a projection $f : V \longrightarrow W$ of V onto W which is a module map.*

Proof: The conclusion of the proposition holds for associative algebras. We make cosmetic changes to the proof in the associative case to prove our version for Lie algebras.

That part (b) implies part (a) is clear. Assume the hypothesis of part (a) and let W be a submodule of V. By Zorn's Lemma there is a submodule W' of V maximal with respect property that $W \cap W' = (0)$, or equivalently $W + W' = W \oplus W'$.

Suppose that $W + W'$ is a proper submodule of V. Then some simple submodule S of V is not contained in $W + W'$. Thus $(W + W') \cap S = (0)$. Hence $W + W' + S = (W \oplus W') \oplus S = W \oplus (W' \oplus S)$ which means that $W \cap (W' \oplus S) = (0)$. This contradiction shows that $W \oplus W' = W + W' = V$ after all. We have shown that part (a) implies part (c).

Assume the hypothesis of part (c) and let W be the sum of all the simple submodules of V. (We will take $W = (0)$ if V has no simple submodules.) By assumption $W \oplus W' = V$ for some submodule W' of V. If $W' = (0)$ part (b) follows.

Suppose that W' is not zero. Then W' contains a non-zero finitely generated submodule W''. By Zorn's Lemma W'' has a maximal proper submodule M. By assumption $M \oplus M' = V$ for some submodule M' of V. The reader is left to show that $M \oplus (M' \cap W'') = W''$ and that $S = M' \cap W''$ is a simple submodule of V. But this means $S \subseteq W \cap W' = (0)$, a contradiction. Therefore $W' = (0)$ after all. We shown that part (c) implies part (b).

We have shown that parts (a)–(c) are equivalent. By our comments preceding the statement of the proposition parts (c) and (d) are equivalent. $\Box$

I think of Shur's Lemma in terms of modules over associative algebras. There seem to be several formulations of this extremely useful lemma. Here, I would venture to say, is the more typical formulation of Shur's Lemma.

Lemma 2.3.9 *Let A be an associative algebra over an algebraically closed field F, let V be a finite-dimensional irreducible left A-module, and suppose that $f : V \longrightarrow V$ is a module map. Then $f = \alpha I$ for some $\alpha \in F$.*

PROOF: Let D be the set of all linear endomorphisms of V which are left A-module maps. It is easy to see that D is a subalgebra of the associative algebra $\mathrm{End}(V)$; in particular D is finite-dimensional. Let $f \in D$. Since $\mathrm{Ker}\, f$, $\mathrm{Im}\, f$ are submodules of V, and V is simple, either $f = 0$ or f is an isomorphism. Therefore D is a division algebra over F. Now we may regard F as a subalgebra of D via the identification of $\alpha \in F$ with $\alpha I \in D$. Since F is algebraically closed necessarily $D = F$. $\Box$

Corollary 2.3.10 *Let V be a finite-dimensional vector space over an algebraically closed field F, let $\mathcal{S}$ be a subset of $\mathrm{End}(V)$ such that (0) and V are the only subspaces of V invariant under all $T \in \mathcal{S}$, and suppose that $f \in \mathrm{End}(V)$ commutes with all $T \in \mathcal{S}$. Then $f = \alpha I$ for some $\alpha \in F$.*

PROOF: We may assume that $V \neq (0)$. Let A be the set of all $T \in \mathrm{End}(V)$ which commute with f. Then A is a subalgebra of the associative algebra $\mathrm{End}(V)$. Now V is a left $\mathrm{End}(V)$-module where $T{\cdot}v = T(v)$ for all $T \in \mathrm{End}(V)$ and $v \in V$. Thus V is a left A-module under the same action. By assumption A contains $\mathcal{S}$. The invariance assumption for $\mathcal{S}$ means that V is a simple left A-module. Since f commutes with all $T \in A$ it follows that f is a left A-module map. Therefore $f = \alpha I$ for some $\alpha \in F$ by Lemma 2.3.9 $\Box$

Here is the connection between Shur's Lemma as formulated on page 26 of the text and Lemma 2.3.9 above, an associative version of Shur's Lemma. Let V be a finite-dimensional vector space over F (which is assumed to be algebraically closed) and suppose that $\pi : L \longrightarrow gl(V)$ is an irreducible representation; that is suppose that V is an irreducible module under the action $x{\cdot}v = \pi(x)(v)$ for all $x \in L$ and $v \in V$. Irreducibility means that V and (0) are the only subspaces of V invariant under $\pi(x)$ for all $x \in L$.

Let f be a linear endomorphism of V which commutes with all $\pi(x)$, $x \in L$. Then V, $\mathcal{S} = \operatorname{Im}\pi$, and f satisfy the hypothesis of Corollary 2.3.10; thus $f = \alpha I$ for some $\alpha \in F$.

In the following exercises L is a Lie algebra over the field F.

Exercise 2.3.11 Let V be a completely reducible left L-module.

(a) Show that a non-zero submodule of V contains a simple submodule of V.

(b) Show that non-zero submodules and quotients of V are completely reducible.

Exercise 2.3.12 Let V be a left L-module and suppose that $\beta : V{\times}V \longrightarrow F$ is a bilinear form. Show that the following are equivalent:

(a) $\beta_\ell : V \longrightarrow V^*$ is a map of left L-modules.

(b) $\beta(x{\cdot}u, v) = -\beta(u, x{\cdot}v)$ for all $x \in L$ and $u, v \in V$.

Exercise 2.3.13 Regard L as a left L-module under $x{\cdot}v = [x\,v]$ for all $x, v \in L$; this is the module action arising form the adjoint representation of L. Suppose that $\beta : L{\times}L \longrightarrow F$ is a bilinear form.

(a) Show that $\beta_\ell : L \longrightarrow L^*$ is a left L-module map if and only if β is associative.

(b) Let $\kappa : L{\times}L \longrightarrow F$ be the Killing form of L. Show that $\kappa_\ell : L \longrightarrow L^*$ is a map of left L-modules.

Exercise 2.3.14 Show that Corollary 2.3.10 implies the following:

Lemma 2.3.15 *Let L be a Lie algebra over an algebraically closed field F, let V be a finite-dimensional simple left L-module, and suppose that $f : V \longrightarrow V$ is a map of left L-modules. Then $f = \alpha I$ for some $\alpha \in F$.*

Exercise 2.3.16 Prove the following corollary:

Corollary 2.3.17 *Let L be a Lie algebra over an algebraically closed field F, let V, W be isomorphic finite-dimensional simple left L-modules, and suppose that $f, g : V \longrightarrow W$ are non-zero maps of left L-modules. Then $g = \alpha f$ for some $\alpha \in F$.*

[Hint: Show that f, g are module isomorphisms and thus $f^{-1}\circ g : V \longrightarrow V$ is a module isomorphism.]

Exercise 2.3.18 Suppose that L is finite-dimensional, simple, and the field F is algebraically closed of characteristic zero.

(a) Let $\beta : L\times L \longrightarrow F$ be an associative bilinear form. Show that $\beta = \alpha\kappa$ for some $\alpha \in F$. [Hint: Note that the Killing form κ is non-degenerate. See Exercises 2.3.13 and 2.3.16.]

(b) Suppose that L is a subalgebra of $gl(V)$ for some finite-dimensional vector space V over F. Show that there is a non-zero $\alpha \in F$ such that

$$\kappa(x, y) = \alpha \mathsf{Tr}(x\circ y)$$

for all $x, y \in L$.

In regard to part (b), see Exercise 2.2.4.

2.3.2 Casimir element of a representation

Let V be a finite-dimensional vector space over the field F and suppose that $\{v_1, \ldots, v_n\}$ is a basis for V. For each $1 \leq i \leq n$ define $v^i \in V^*$ by $v^i(v_j) = \delta_{i,j}$ for all $1 \leq j \leq n$. It is easy to see that $\{v^1, \ldots, v^n\}$ is a basis for V^*, called the *basis for V^* dual to the basis* $\{v_1, \ldots, v_n\}$, or more informally the *dual basis for V^**.

Now let L be a finite-dimensional Lie algebra over the field F and suppose that $\beta : L\times L \longrightarrow F$ is a non-degenerate symmetric associative bilinear form. Case in point: the Killing form β when L semisimple and F an algebra field of characteristic zero. Since β is non-degenerate $\beta_\ell : L \longrightarrow L^*$ is a linear isomorphism. Let $\{x_1, \ldots, x_n\}$ be a basis for L and let $\{x^1, \ldots, x^n\}$ be the dual basis for L^*. Since β_ℓ is an isomorphism there are unique $y_1, \ldots, y_n \in L$ which satisfy $\beta_\ell(y_i) = x^i$ for all $1 \leq i \leq n$. Observe that $\{y_1, \ldots, y_n\}$ is also a basis for L and, since β is symmetric, that

$$\beta(y_i, x_j) = \delta_{i,j} = \beta(x_i, y_j) \tag{2.12}$$

for all $1 \leq i, j \leq n$. By Exercise 2.3.22 we have

$$x = \sum_{i=1}^{n} \beta(x, x_i)y_i = \sum_{i=1}^{n} \beta(x, y_i)x_i \tag{2.13}$$

for all $x \in L$.

The tensor $\sum_{i=1}^{n} x_i \otimes y_i$ is of importance in connection with the Casimir element. For given β it does not depend on the particular choice of basis. See Exercise 2.3.21. A very important relation for us is

$$\sum_{i=1}^{n} [x\, x_i] \otimes y_i = \sum_{i=1}^{n} x_i \otimes [y_i\, x] \tag{2.14}$$

for all $x \in L$.

PROOF: By virtue of Exercise 2.3.21 we need only show that for each $z \in L$ applying $\beta_\ell(z) \otimes I$ to both sides of the equation results in the same element of L. Applying $\beta_\ell(z) \otimes I$ to the left hand side of the equation yields

$$\sum_{i=1}^{n} \beta_\ell(z)([x\, x_i]) y_i = \sum_{i=1}^{n} \beta(z, [x\, x_i]) y_i = \sum_{i=1}^{n} \beta([z\, x], x_i) y_i = [z\, x]$$

by (2.13) since β is associative. Applying $\beta_\ell(z) \otimes I$ to the right hand side yields

$$\sum_{i=1}^{n} \beta_\ell(z)(x_i)[y_i\, x] = \sum_{i=1}^{n} \beta_\ell(z, x_i)[y_i\, x] = [\sum_{i=1}^{n} \beta_\ell(z, x_i) y_i\, x] = [z\, x]$$

by (2.13) again. □

Lemma 2.3.19 *Let $\phi : L \longrightarrow gl(V)$ be a representation of a Lie algebra L over F and suppose that $\sum_{i=1}^{n} x_i \otimes y_i \in L \otimes L$ satisfies (2.14). Then the endomorphism of V defined by*

$$\mathsf{c}_\beta(\phi) = \sum_{i=1}^{n} \phi(x_i) \circ \phi(y_i)$$

is a map of left L-modules.

PROOF: Let $x \in L$. Since ϕ is a Lie algebra map we deduce form (2.14) that

$$\sum_{i=1}^{n} [\phi(x)\ \phi(x_i)] \circ \phi(y_i) = \sum_{i=1}^{n} \phi(x_i) \circ [\phi(y_i)\ \phi(x)]. \tag{2.15}$$

Since the Lie algebra derivation $\operatorname{ad}\phi(x)$ of $gl(V)$ is also an associative algebra derivation of $\operatorname{End}(V)$, we use (2.15) to compute

$$\begin{aligned}[\phi(x)\ \mathsf{c}_\beta(\phi)] &= [\phi(x)\ \sum_{i=1}^{n}\phi(x_i)\circ\phi(y_i)] \\ &= \sum_{i=1}^{n}[\phi(x)\ \phi(x_i)]\circ\phi(y_i) + \sum_{i=1}^{n}\phi(x_i)\circ[\phi(x)\ \phi(y_i)] \\ &= \sum_{i=1}^{n}\phi(x_i)\circ[\phi(y_i)\ \phi(x)] - \sum_{i=1}^{n}\phi(x_i)\circ[\phi(y_i)\ \phi(x)] \\ &= 0.\end{aligned}$$

by (2.15). Therefore $\phi(x)$ and $\mathsf{c}_\beta(\phi)$ commute for all $x \in L$; that is $\mathsf{c}_\beta(\phi)$ is a map of left L-modules. □

In more concrete terms

$$\mathsf{c}_\beta(\phi)(v) = \sum_{i=1}^{n} x_i\cdot(y_i\cdot v) \tag{2.16}$$

for all $v \in V$.

We end with a very important application to semisimple Lie algebras.

Proposition 2.3.20 *Let L be a finite-dimensional semisimple Lie algebra over an algebraically closed field F of characteristic 0, and suppose that $\phi : L \longrightarrow gl(V)$ is a non-trivial finite-dimensional representation of L. Then there is a module map $\mathsf{c}_\beta(\phi) : V \longrightarrow V$ described by (2.16) which satisfies $\mathsf{Tr}(\mathsf{c}_\beta(\phi)) \neq 0$.*

Proof: First of all suppose that $\operatorname{Ker}\phi = (0)$; that is ϕ is a faithful representation. Let $\beta : L \longrightarrow F$ be the symmetric associative bilinear form defined by $\beta(x, y) = \mathsf{Tr}(\phi(x)\circ\phi(y))$ for all $x, y \in L$.

We claim first of all that β is non-singular. The ideal $I = \operatorname{Rad}\beta$ of L is a semisimple Lie algebra. Thus $\phi(I)$ is a semisimple subalgebra of $gl(V)$. Since $0 = \beta(x, y) = \mathsf{Tr}(\phi(x)\circ\phi(y))$ for all $x, y \in I$ it follows that the semisimple subalgebra $\phi(I)$ is also solvable. Therefore $\phi(I) = (0)$. Thus $I \subseteq \operatorname{Ker}\phi = (0)$ which means that $\operatorname{Rad}\beta = (0)$. Therefore β is non-singular.

By our discussion above there are bases $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ for L which satisfy (2.12) and (2.14). Let $\mathsf{c}_\beta(\phi)$ be the module endomorphism of V of Lemma 2.3.19 defined for $\sum_{i=1}^n x_i \otimes y_i$. Observe that

$$\mathsf{Tr}(\mathsf{c}_\beta(\phi)) = \sum_{i=1}^n \mathsf{Tr}(\phi(x_i) \circ \phi(y_i)) = \sum_{i=1}^n \beta(x_i, y_i) = (\mathrm{Dim}\, L)1.$$

Since ϕ is not trivial, $L \neq (0)$, and thus $(\mathrm{Dim}\, L)1 \neq 0$ since the characteristic of F is zero.

Now we pass to the general case. Since L is semisimple $L = \mathrm{Ker}\, \phi \oplus \mathcal{L}$ is the direct sum of ideals of L. Observe there is only one possibility for $\mathcal{L}$. Now $\mathcal{L}$ is a semisimple Lie algebra and the restriction $\phi|_{\mathcal{L}} : \mathcal{L} \longrightarrow gl(V)$ is a non-trivial faithful representation of $\mathcal{L}$. Construct $\mathsf{c}_\beta(\phi) = \mathsf{c}_{\beta|_{\mathcal{L} \times \mathcal{L}}}(\phi|_{\mathcal{L}})$ as above with bases $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ for $\mathcal{L}$. Note that (2.14) is satisfied for all $x \in L$ as $[\mathrm{Rad}\, \beta\ \mathcal{L}] = (0)$. Therefore $\mathsf{c}_\beta(\phi)$ is an endomorphism of L-modules by Lemma 2.3.19. $\square$

The module map $\mathsf{c}_\beta(\phi)$ described in the proof of the preceding proposition is called the *Casimir element of* ϕ. Observe that the construction does not depend on the choice of basis $\{x_1, \ldots, x_n\}$.

Exercise 2.3.21 Let V, W be vector spaces over the field F and suppose that $\mathsf{v} \in V \otimes W$ is not zero. Write $\mathsf{v} = \sum_{i=1}^r v_i \otimes w_i$ as a sum of tensors where r is as small as possible.

(a) Show that $\{v_1, \ldots, v_r\}$ and $\{w_1, \ldots, w_r\}$ are linearly independent sets. [Hint: What happens if $r > 1$ and v_r is a linear combination of $v_1, \ldots, v_{r-1}$?]

(b) Show that there is a $f \in V^*$ such that $0 \neq (f \otimes I)(\mathsf{v}) = \sum_{i=1}^r f(v_i) w_i$.

c) Suppose that $\mathsf{u} \in V \otimes W$. Show that $(f \otimes I)(\mathsf{u}) = 0$ for all $f \in V^*$ implies that $\mathsf{u} = 0$.

Exercise 2.3.22 Let $\beta : V \otimes V \longrightarrow F$ be a non-degenerate symmetric bilinear form on a finite-dimensional vector space V over F.

(a) Show that there are bases $\{x_n, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ for V such that (2.12) holds.

Suppose $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ are bases for V such that (2.12) holds.

(b) Show that $\sum_{i=1}^n \beta(x, x_i) y_i = x = \sum_{i=1}^n \beta(x, y_i) x_i$ for all $x \in V$. [Hint: Possibly too helpful of a hint. Write $x = \sum_{j=1}^n \alpha_j x_j$ and calculate $\beta(x, y_i)$.]

(c) Suppose $\{x'_1, \ldots, x'_n\}$ and $\{y'_1, \ldots, y'_n\}$ are also bases for V such that (2.12) holds. Show that $\sum_{i=1}^n x_i \otimes y_i = \sum_{i=1}^n x'_i \otimes y'_i$. [Hint: Note that $\beta_\ell : V \longrightarrow V^*$ is an isomorphism. Apply $\beta_\ell(x) \otimes I$ to the difference of the two sides of this equation. See Exercise 2.3.21.]

2.3.3 Weyl's Theorem

Concerning the Lemma.

Remark 2.3.23 *Let L be any Lie algebra and suppose that V is a one-dimensional left L-module. Then $[L\,L]\cdot v = (0)$.*

To see this, write $V = Fv$ and let $x, y \in L$. Then $x\cdot v = \alpha v$ and $y\cdot v = \beta v$ for some $\alpha, \beta \in F$. Therefore

$$[x\,y]\cdot v = x\cdot(y\cdot v) - y\cdot(x\cdot v) = x\cdot(\beta v) - y\cdot(\alpha v) = \beta\alpha v - \alpha\beta v = 0.$$

Suppose that L is a Lie algebra over F and let V be a non-zero left L-module. Then V is completely reducible if and only if for every submodule W of V there is a projection $f : V \longrightarrow W$ from V onto W which is a module map by Proposition 2.3.8.

Let W be a submodule of V. Recall that $\mathrm{Hom}(V, W)$ is a left L-module where $x\cdot f(v) = -f(x\cdot v) + x\cdot f(v)$ for all $x \in L$, $f \in \mathrm{Hom}(V, W)$, and $v \in V$. Observe that

$$\mathrm{M}(W, W) = \{f \in \mathrm{Hom}(V, W) \mid f(W) \subseteq W\}$$

is a submodule of $\mathrm{Hom}(V, W)$ and that $\pi : \mathrm{M}(V, W) \longrightarrow \mathrm{End}(W)$ defined by $\pi(f) = f|_W$ is a map of left L-modules. Since the identity map $\mathrm{Id}_W : W \longrightarrow W$ is a module map the linear span $F\mathrm{Id}_W$ is a one-dimensional trivial left L-module by Remark 2.3.6. Note that $\mathcal{W} = \mathrm{Ker}\,\pi$ is a codimension one submodule of $\mathcal{V} = \pi^{-1}(FI_W)$. Observe that any projection $f : V \longrightarrow W$ lies in $\mathcal{V}\backslash\mathcal{W}$. Conversely, if $f \in \mathcal{V}\backslash\mathcal{W}$ then some scalar multiple of f is a projection of V onto W. Therefore there exists a projection of V onto W which is a module map if and only if there exists a one-dimensional trivial submodule $\mathcal{W}'$ of $\mathcal{V}$ such that $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}'$.

We have a characterization of Lie algebras whose finite-dimensional representations are completely reducible in terms of the existence of certain one-dimensional modules.

Theorem 2.3.24 *Let L be a Lie algebra over the field F. Then the following are equivalent:*

a) *All non-zero finite-dimensional left L-modules are completely reducible.*

b) *$[L\,L] = L$ and all finite-dimensional left L-modules V which contain a codimension one simple submodule W contain a one-dimensional submodule W' such that $V = W \oplus W'$.*

PROOF: Suppose that all non-zero left L-modules are completely reducible. Suppose that V is finite-dimensional left L-module and W is a submodule of V. Then $V = W \oplus W'$ for some submodule W' of V. Since $\operatorname{Dim} V = \operatorname{Dim} W + \operatorname{Dim} W'$, if W has codimension one then W' has dimension one. That $L = [L\,L]$ follows by Exercise 2.3.30. We have shown that part a) implies part b).

Conversely, assume the hypothesis of part b). Since $L = [L\,L]$ any one-dimensional left L-module is trivial by Remark 2.3.23. Let V be a finite-dimensional left L-module which contains a submodule W of codimension one. Since any left L-module is trivial, by our discussion preceding the statement of the Theorem to show part b) implies part a) we need only show that V contains a one-dimensional submodule W' such that $V = W \oplus W'$. This we do by induction on $\operatorname{Dim} V$.

If W is simple we are done by assumption. Suppose that W is not simple. We may assume that $W \neq (0)$. Then W contains a non-zero proper submodule W''. Since W/W'' is a codimension one submodule of V/W'', by induction on $\operatorname{Dim} V$ we have $V/W'' = W/W'' \oplus W'/W''$, where W' is a submodule of V containing W''. It is a small exercise to show that W'' is a codimension one submodule of W' and that W' is a proper submodule of V. Thus by induction on $\operatorname{Dim} V$ again, $W' = W'' \oplus W'''$ where W''' is a one-dimensional submodule of W''. Since $W''' \cap W \subseteq W' \cap W \subseteq W''$ it follows that $W''' \cap W \subseteq W''' \cap W'' = (0)$. Therefore $V = W \oplus W'''$. $\Box$

Remark 2.3.25 *Weyl's Theorem follows form Theorem 2.3.24 and the lemma below:*

Lemma 2.3.26 *Let L be a finite-dimensional semisimple Lie algebra over an algebraically closed field of characteristic zero and suppose that V is a finite-dimensional left L module which contains a codimension one simple submodule W. Then $V = W \oplus W'$ for some a one-dimensional submodule (which must be trivial).*

Proof: By Proposition 2.3.20 there are $x_1, \ldots, x_n, y_1, \ldots, y_n \in L$ such that $f : V \longrightarrow V$ defined by $f(v) = \sum_{i=1}^{n} x_i{\cdot}(y_i{\cdot}v)$ for all $v \in V$ is a left L-module map and $\mathsf{Tr}(f) \neq 0$. By the Lemma of the section $x{\cdot}V \subseteq W$; therefore $f(V) \subseteq W$.

Suppose that $f(W) = (0)$. Then $f^2 = 0$ which means that $\mathsf{Tr}(f) = 0$, a contradiction. Therefore $f(W) \neq (0)$. Since W is simple it follows that $f(W) = W$. Therefore $\mathrm{Im}\, f = W$, so $\mathrm{Ker}\, f$ is a one-dimensional submodule of V by the Rank–Nullity Theorem. Since $W \cap \mathrm{Ker}\, f$ is a submodule of W, it follows that $W \cap \mathrm{Ker}\, f = (0)$. Our conclusion: $V = W \oplus \mathrm{Ker}\, f$. $\Box$

The following theorem is very important for associative algebras[1]:

Theorem 2.3.27 *Let A be a finite-dimensional associative algebra over the field F. Then the following are equivalent:*

(a) *All left A-modules are completely reducible.*

(b) *A is completely reducible as a left A-module under multiplication.*

Exercise 2.3.28 The preceding theorem is not the case for Lie algebras as the following example shows. Let $L = Fx$ be a one-dimensional Lie algebra over F, suppose that V is a vector space over F and let $\pi : L \longrightarrow \mathrm{End}(V)$ be *any* linear map, and let $T = \pi(x)$. Note that L is a completely reducible left L-module no matter what the structure.

(a) Show that π is a representation of L.

(b) Show that the submodules of V are the T-invariant subspaces of V.

(c) Show that V is completely reducible if and only if the minimal polynomial of T over F factors into distinct irreducible factors.

(d) Show that there are non-zero left L-modules which are not completely reducible.

Exercise 2.3.29 Let $f : L \longrightarrow L'$ be a map of Lie algebras and suppose that V is a left L'-module.

(a) Show that V is a left L-module where $x{\cdot}v = f(x){\cdot}v$ for all $x \in L$ and $v \in V$.

Regard V as a left L-module according to part (a).

---

[1] See Theorem 4, page 820 of *Abstract Algebra*, Dummit and Foote, Prentice Hall, N.J., 1999

(b) Show that the L'-submodules of V are the L-submodules of V.

Suppose that f is onto.

(c) Show that L-submodules of F are L'-submodules of V.

(d) Suppose that $V \neq (0)$. Show that V is a completely reducible left L-module if and only if V is a completely reducible left L'-module.

Exercise 2.3.30 Let L be a Lie algebra over the field F and suppose that all non-zero left L-modules are completely reducible. Show that $L = [L\,L]$. [Hint: If not, show there is an onto Lie algebra map $f : L \longrightarrow L/I$ where I is a codimension one ideal of L. See Exercises 2.3.28 and 2.3.29.]

2.3.4 Preservation of Jordan decomposition

Let V be any vector space over F and let W be a subspace of V. Recall that $M(W, W) = \{f \in \mathrm{End}(V) \,|\, f(W) \subseteq W\}$ is a subalgebra of the associative algebra $\mathrm{End}(V)$ and is therefore a Lie subalgebra of $gl(V)$. The map $\Pi : M(W, W) \longrightarrow \mathrm{End}(W)$ defined by $\Pi(x) = x|_W$ is a map associative algebras and therefore can be regarded as a map of Lie algebras $\Pi : M(W, W) \longrightarrow gl(W)$.

Let L be a subspace of $gl(V)$ and set

$$\begin{aligned} L_W &= \{x \in gl(V) \,|\, [x\,L] \subseteq L,\ x(W) \subseteq W,\ \mathsf{Tr}(x|_W) = 0\} \\ &= \{x \in gl(V) \,|\, \mathrm{ad}\,x(L) \subseteq L,\ x(W) \subseteq W,\ \mathsf{Tr}(x|_W) = 0\}. \end{aligned}$$

It is easy enough to check directly that L_W is a Lie subalgebra of $gl(V)$. Let $\mathcal{V} = gl(V)$ and $\pi : gl(V) \longrightarrow gl(\mathcal{V})$ be the adjoint representation. Then L is a subspace of $\mathcal{V}$ and

$$L_W = \pi^{-1}(M(L, L)) \cap \Pi^{-1}(sl(W)) \tag{2.17}$$

which shows that L_W is a Lie subalgebra of $gl(V)$. Suppose that $x \in L_W$. Since x_s, x_n are polynomials in x it follows that W is invariant under x_n and x_s. For this same reason L is invariant under $(\mathrm{ad}\,x)_s = \mathrm{ad}\,x_s$ and $(\mathrm{ad}\,x)_n = \mathrm{ad}\,x_n$. Since W is invariant under x, x_s, and x_n we have $x|_W = x_s|_W + x_n|_W$. Now $x_n|_W$ is nilpotent. Therefore $\mathsf{Tr} x_n|_W = 0$ and consequently $\mathsf{Tr}(x_s|_W) = \mathsf{Tr}(x_s|_W) + \mathsf{Tr}(x_n|_W) = \mathsf{Tr}(x|_W) = 0$. We have shown that

$$x \in L_W \ \text{ implies } \ x_s, x_n \in L_W. \tag{2.18}$$

Note that $\Pi([M(W,W)\, M(W,W)]) = [\Pi(M(W,W))\, \Pi(M(W,W))] \subseteq sl(W)$; thus

$$\text{If } L \text{ is a Lie subalgebra and } L(W) \subseteq W \text{ then } [L\, L] \subseteq L_W. \tag{2.19}$$

The technical point of the section is embodied in the following:

Lemma 2.3.31 *Suppose that V is a finite-dimensional vector space over an algebraically closed field F of characteristic zero and L is a subalgebra of $gl(V)$ such that*

(a) $[L\, L] = L$,

(b) *V is a completely reducible left L-module under $x{\cdot}v = x(v)$ for all $x \in L$ and $v \in V$,*

(c) *$gl(V)$ is a completely reducible left L-module under the adjoint action, that is $x{\cdot}\mathsf{v} = \operatorname{ad} x(\mathsf{v}) = [x\, \mathsf{v}]$ for all $\mathsf{v} \in gl(V)$.*

Then $x_s, x_n \in L$ for all $x \in L$.

Proof: Write $V = W_1 + \cdots + W_r$ as the sum of simple submodules and set $L' = L_{W_1} \cap \cdots \cap L_{W_r}$. By (2.19) we conclude that $L \subseteq L_W$. Since $[L_W\, L] \subseteq L$ in any case it follows that $[L\, L_W] \subseteq L_W$. Thus L_W is an L-submodule of $gl(V)$. Now non-zero submodules of completely reducible modules are completely reducible by Exercise 2.3.11. At this point the proof in the text goes through verbatim. □

In the following exercises we explore the implications of a finite-dimensional Lie algebra L being completely reducible under the adjoint action; that is $x{\cdot}v = \operatorname{ad} x(v) = [x\, v]$ for all $x, v \in L$.

Exercise 2.3.32 Let L be a non-zero finite-dimensional Lie algebra over the field F and suppose that L is a completely reducible left L-module under the adjoint action. Then

$$L = L_1 \oplus \cdots \oplus L_s$$

is the direct sum of simple left L-modules, that is simple ideals of L.

(a) Let I be a simple ideal of L.

  (i) Show that $I \subseteq Z(L)$ of $I = [L\, I] = [L_i\, L_i] = L_i$ for some $1 \leq i \leq s$.

  (ii) Show that $\operatorname{Dim} I = 1$ if and only if $I \subseteq Z(L)$.

By virtue of part (a) we may write

$$L = Z \oplus L_1 \oplus \cdots \oplus L_r$$

as the direct sum of ideals where $Z \subseteq \mathrm{Z}(L)$ and $\dim L_i > 1$ for all $1 \leq i \leq r$.

(b) Show that L_i is a simple Lie algebra for all $1 \leq i \leq r$.

(c) Show that $Z = \mathrm{Z}(L)$.

(d) Let I be an ideal of L. Show that $I = (I \cap \mathrm{Z}(L)) \oplus L'_1 \oplus \cdots \oplus L'_r$, where $L'_i = (0)$ or $L'_i = L_i$ for all $1 \leq i \leq r$.

(e) Show that $\mathrm{Rad}\, L = \mathrm{Z}(L)$ (and thus L is reductive).

(f) Show that $L = \mathrm{Z}(L) \oplus [L\, L]$, where $\mathcal{L} = [L\, L]$ is semisimple and is a completely reduced left $\mathcal{L}$-module under the adjoint action.

Exercise 2.3.33 Let L be a non-zero finite-dimensional Lie algebra and suppose that all finite-dimensional non-zero left L-modules are completely reducible. Show that L is semisimple. [Hint: See Exercises 2.3.28–2.3.32.]

Exercise 2.3.34 Prove the following theorem. See Exercise 2.3.28 in connection with the theorem.

Theorem 2.3.35 *Let L be an non-zero finite-dimensional Lie algebra over an algebraically closed field of characteristic zero. Then the following are equivalent:*

(a) $\mathrm{Z}(L) = (0)$ *and L is a completely reducible left L-module under the adjoint action.*

(b) *L is semisimple.*

(c) *All non-zero finite-dimensional left L-modules are completely reducible.*

2.4 Representations of $sl(2, F)$

2.4.1 Weights and maximal vectors

For a linear endomorphism $T : V \longrightarrow V$ of a vector space V over F let

$$V_\lambda = \mathrm{Ker}\,(T - \lambda I) = \{v \in V \mid T(v) = \lambda v\}$$

for all $\lambda \in F$. Note that if $V_\lambda \neq (0)$ then V_λ is the subspace of eigenvectors for T belonging to λ.

Remark 2.4.1 *Let $T : V \longrightarrow V$ be a semisimple linear endomorphism of a finite-dimensional vector space V over F. Then $V = \oplus_{\lambda \in F} V_\lambda$.*

Let $L = sl(2, F)$ and suppose that the field F is algebraically closed of characteristic zero. We have seen that L is simple, hence semisimple. Let V be any finite-dimensional left L-module. Then left multiplication

$$h\cdot : V \longrightarrow V$$

by h, defined by $h\cdot(v) = h\cdot v$ for all $v \in V$, is a semisimple endomorphism of V. The effect of the left multiplications determined by x, y, and h on the V_λ's is explained by schematic diagram

$$\begin{tikzcd}
{} \arrow[r, shift right, "x\cdot"'] & V_{\lambda-2} \arrow[l, shift right, "y\cdot"'] \arrow[loop above, "h\cdot"] \arrow[r, shift right, "x\cdot"'] & V_\lambda \arrow[l, shift right, "y\cdot"'] \arrow[loop above, "h\cdot"] \arrow[r, shift right, "x\cdot"'] & V_{\lambda+2} \arrow[l, shift right, "y\cdot"'] \arrow[loop above, "h\cdot"] \arrow[r, shift right, "x\cdot"'] & {} \arrow[l, shift right, "y\cdot"']
\end{tikzcd}.$$

In the following exercises, we expand on the corollary of Section 6.4 of the text which is a very important result for this section.

Exercise 2.4.2 Let $\phi : L \longrightarrow L'$ be a map of finite-dimensional Lie algebras, where L is semisimple. Suppose that F is an algebraically closed field of characteristic zero.

(a) Suppose that L' is semisimple also and ϕ is onto. Show that $\phi(x)_s = \phi(x_s)$ and $\phi(x)_n = \phi(x_n)$ for all $x \in L$. [Hint: Let $x \in L$. Note that $(\mathrm{ad}\,\phi(x))\circ\phi = \phi\circ(\mathrm{ad}\,x)$ and therefore $f(\mathrm{ad}\,\phi(x))\circ\phi = \phi\circ f(\mathrm{ad}\,x)$ for all $f(x) \in F[x]$. Recall that an endomorphism $T : V \longrightarrow V$ of a finite-dimensional vector space over any field F is semisimple if and only if $f(T) = 0$ for some $f(x) \in F[x]$ which splits into distinct linear factors over F, and that T is nilpotent if and only if $f(T) = 0$ where $f(x) = x^m$ for some $m > 0$.]

(b) Suppose that L' is semisimple also and ϕ is one-one. Show that $\phi(x)_s = \phi(x_s)$ and $\phi(x)_n = \phi(x_n)$ for all $x \in L$. [Hint: By part a) we may assume that ϕ is the inclusion; that is $L \subseteq L'$. The adjoint representation $\mathrm{ad}_{L'} : L' \longrightarrow gl(L')$ is one-one. Let $x = x_s + x_n$ be the decomposition of x into semisimple and nilpotent parts *in* L'. Then $\mathrm{ad}_{L'}\, x = \mathrm{ad}_{L'}\, x_s + \mathrm{ad}_{L'}\, x_n$ is the Jordan-Chevalley decomposition of $\mathrm{ad}_L\, x$ in $\mathrm{End}(L')$. Thus $\mathrm{ad}_{L'} x_s, \mathrm{ad}_{L'} x_n \in \mathrm{ad}_{L'}(L)$. Show that $x_s, x_n \in L$ and that $\mathrm{ad}_L\, x = \mathrm{ad}_L\, x_s + \mathrm{ad}_L\, x_n$ is the Jordan-Chevalley decomposition of $\mathrm{ad}_L\, x$ in $\mathrm{End}(L)$; thus $x = x_s + x_n$ is the decomposition of x into semisimple and nilpotent parts in L.]

(c) Show that $\phi(x)_s = \phi(x_s)$ and $\phi(x)_n = \phi(x_n)$ for all $x \in L$.

d) Suppose that $L' = gl(V)$ for some finite-dimensional vector space V over F. Show that $\phi(x) = \phi(x_s) + \phi(x_n)$ is the Jordan-Chevalley decomposition of $\phi(x)$ for all $x \in L$. (That is, go through the details of the proof of the corollary of Section 6.3.)

Exercise 2.4.3 Let F be an algebraically closed field of characteristic zero and suppose that $L \subseteq gl(V)$ is a semisimple Lie subalgebra of $gl(V)$, where V is a finite-dimensional vector space over F. Give a very careful proof of the fact that for $x \in L$ the decomposition $x = x_s + x_n$ into semisimple and nilpotent parts in L is the Jordan–Chevalley decomposition of the endomorphism x of V into commuting semisimple and nilpotent endomorphisms.

2.4.2 Classification of irreducible modules

We continue with the notation and assumptions of the previous section unless otherwise stated. *$L = sl(2, F)$ and the field F is algebraically closed of characteristic zero.*

Let V be any non-zero finite-dimensional left L-module. For $\rho \in F$ we set $V_\rho = \mathrm{Ker}\,(h\cdot - \rho I)$. Since $h\cdot$ is a semisimple endomorphism of V it follows that $V_\rho \neq (0)$ for some $\rho \in F$.

Now the sum

$$V_\rho + V_{\rho+2} + V_{\rho+4} + \cdots$$

is direct since the subscripts are distinct elements of F. Since V is finite-dimensional $V_{\rho+2\ell} = (0)$ for some $\ell \geq 0$. Let ℓ be the least such integer. Then $\ell \geq 1$. Set $\lambda = \rho + 2(\ell - 1)$. Then

$$V_\lambda \neq (0) = V_{\lambda+2}.$$

Choose $u_0 \in V_\lambda \backslash 0$. Note that $x\cdot u_0 = 0$. For all $1 \leq \imath$ define $u_\imath$ by

$$u_\imath = y\cdot u_{\imath-1}.$$

Observe that

$$u_\imath \in V_{\lambda-2\imath}$$

for all $0 \leq \imath$. Now the sum

$$V_\lambda + V_{\lambda-2} + V_{\lambda-4} + \cdots$$

is direct since the subscripts are distinct elements of F. Therefore $u_\ell = 0$ for some $\ell \geq 0$. Let ℓ be the least such integer. Observe that $\ell > 0$. Let $m = \ell - 1$. Then

$$u_m \neq 0 = u_{m+1}.$$

For convenience we set $u_{-1} = 0$.

Let W be linear span of $u_0, \ldots, u_m$. Since these vectors are not zero and belong to different eigenspaces for $h\cdot$ it follows that $\{u_0, \ldots, u_m\}$ is a basis for W. We will show that there are $\alpha_0, \ldots, \alpha_{m+1} \in F$ which satisfy

$$x\cdot u_\imath = \alpha_\imath u_{\imath-1} \quad \text{for all } 0 \leq \imath \leq m+1, \tag{2.20}$$

$\alpha_0 = 0 = \alpha_{m+1}$, and $\alpha_1, \ldots, \alpha_m \neq 0$.

Suppose this is the case. Then W is a submodule of V. Observe that *any* $u_\imath$ generates W as a left L-module. Furthermore W is simple. For let W' be a non-zero submodule of W. Then W' is invariant under $h\cdot$. Since $h\cdot|_{W'}$ is semisimple also, W' contains a non-zero eigenvector for this restriction. Since the eigenspaces of $h\cdot$ are one dimensional and spanned by the $u_\imath$'s, it follows that $u_\imath \in W'$ for some $1 \leq \imath \leq m$. Therefore $W' = W$.

Back to $\alpha_0, \ldots, \alpha_{m+1}$. We show that

$$\alpha_\imath = \imath(\lambda + 1 - \imath) \tag{2.21}$$

for all $0 \leq \imath \leq m+1$. Since $\alpha_0 = 0$ the equation $x\cdot u_0 = 0 = \alpha_0 u_{-1}$ follows. Suppose that $x\cdot u_\jmath = \alpha_\jmath u_{\jmath-1}$ for all $0 \leq \jmath < \imath \leq m+1$. Then $0 < \imath \leq m+1$ and thus

$$\begin{aligned} x\cdot u_\imath = x\cdot(y\cdot u_{\imath-1}) &= [x\,y]\cdot u_{\imath-1} + y\cdot(x\cdot u_{\imath-1}) \\ &= (\lambda - 2(\imath-1))\cdot u_{\imath-1} + y\cdot(\alpha_{\imath-1}u_{\imath-2}) \\ &= \big((\lambda - 2(\imath-1) + (\imath-1)(\lambda+2-\imath)\big)\cdot u_{\imath-1} \\ &= \imath(\lambda+1-\imath)\cdot u_{\imath-1}. \end{aligned}$$

We have established (2.21) by induction on $\imath$.

Since $\alpha_{m+1} = 0$ we conclude that $\lambda = m$. Therefore

$$\alpha_\imath = \imath(m+1-\imath)$$

for all $0 \leq \imath \leq m+1$. Observe that

$$\alpha_\imath = \alpha_{m+1-\imath} \qquad \text{for all } 0 \leq \imath \leq m+1;$$

thus the $\alpha_\imath$'s have a nice symmetry. Note that $\alpha_\imath = 0$ if and only if $\imath = 0$ or $\imath = m + 1$. Thus W is simple.

We will make a change of basis which will give the traditional description of W. Set

$$v_\imath = \frac{u_\imath}{\imath!}$$

for all $0 \leq \imath \leq m$. Then $\{v_0, \ldots, v_m\}$ is a basis for W and

$$h \cdot v_\imath = (m - 2\imath) \cdot v_\imath, \quad x \cdot v_\imath = (m + 1 - \imath) \cdot v_{\imath-1}, \quad y \cdot v_\imath = (\imath + 1) \cdot v_{\imath+1} \tag{2.22}$$

for all $0 \leq \imath \leq m$. Suppose that $V = W$. Identifying the endomorphisms $h\cdot, x\cdot$, and $y\cdot$ of V with their matrices with respect to $\{v_0, \ldots, v_m\}$ we have

$$h\cdot = \begin{pmatrix} m & & & \\ & m-2 & & \\ & & \ddots & \\ & & & -m \end{pmatrix},$$

$$x\cdot = \begin{pmatrix} 0 & m & & & \\ & 0 & m-1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix},$$

and

$$y\cdot = \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 2 & \ddots & & \\ & & \ddots & 0 & \\ & & & m & 0 \end{pmatrix}.$$

We have determined what finite-dimensional simple left L-modules must be. Whether or not they exist (they do) needs to be verified. See Exercise 2.4.8. We close with a result which will be useful in Chapter 8.

Corollary 2.4.4 *Let F be an algebraically closed field of characteristic zero and suppose that V is a finite-dimensional $L = sl(2, F)$-module. Then the eigenvalues of the left multiplication $h\cdot : V \longrightarrow V$ are integers.*

PROOF: We may as well assume that $V \neq (0)$. In this case V is the sum of simple left L-modules which satisfy the conclusion of the corollary. Thus $f(h\cdot) = 0$ for some $f(x) = (x - \lambda_1) \cdots (x - \lambda_r) \in F[x]$ with $\lambda_1, \ldots, \lambda_r \in \boldsymbol{Z}$. Since the eigenvalues of $h\cdot$ must be roots of $f(x)$ the corollary follows. □

Exercise 2.4.5 Let L be a finite-dimensional Lie algebra over a field F with basis $\{x_1, \ldots, x_n\}$ and suppose that V is a vector space over F. Let $T_1, \ldots, T_n \in \mathrm{End}(V)$ and $\phi : L \longrightarrow \mathrm{End}(V)$ be the linear map determined by $\phi(x_\imath) = T_\imath$ for all $1 \leq \imath \leq n$. Then ϕ determines a representation of Lie algebras $\phi : L \longrightarrow gl(V)$ if and only if

$$\phi([x\,y]) = [\phi(x)\ \phi(y)] \tag{2.23}$$

holds for all $x, y \in L$.

Show that (2.23) holds if and only if

$$\phi([x_\imath\, x_\jmath]) = [T_\imath\ T_\jmath]$$

for all $1 \leq \imath < \jmath \leq n$. [Hint: Note that $\beta, \beta' : L \times L \longrightarrow \mathrm{End}(V)$ defined by $\beta(x, y) = \phi([x\,y])$ and $\beta'(x, y) = [\phi(x)\,\phi(y)]$ for all $x, y \in L$ are both bilinear functions $\gamma : L \times L \longrightarrow \mathrm{End}(V)$ which satisfy $\gamma(y, x) = -\gamma(x, y)$ and $\gamma(x, x) = 0$ for all $x, y \in L$.]

Exercise 2.4.6 Suppose that F is a field of characteristic zero, not necessarily algebraically closed, let $L = sl(2, F)$, and let $\{x, y, h\}$ be the standard basis for L. Let $\lambda \in F$ and $\boldsymbol{V}(\lambda)$ be the vector space over F with basis of symbols $\{u_\imath\}_{\imath \in \boldsymbol{Z}}$. Define

$$\alpha_\imath = \imath(\lambda + 1 - \imath)$$

for all $\imath \in \boldsymbol{Z}$.

(a) Let $T_1, T_2, T_3 \in \mathrm{End}(\boldsymbol{V}(\lambda))$ be the endomorphisms determined by

$$T_1(u_\imath) = \alpha_\imath u_{\imath-1}, \quad T_2(u_\imath) = u_{\imath+1}, \quad \text{and} \quad T_3(u_\imath) = (\lambda - 2\imath)u_\imath$$

for all $\imath \in \boldsymbol{Z}$. Show that the linear map $\phi : L \longrightarrow \mathrm{End}(\boldsymbol{V}(\lambda))$ which is defined by $\phi(x) = T_1, \phi(y) = T_2$, and $\phi(h) = T_3$, determines a representation $\phi : L \longrightarrow \mathrm{End}(\boldsymbol{V}(\lambda))$.

(b) Suppose that W is a non-zero submodule of $\boldsymbol{V}(\lambda)$. Show that $u_m \in W$ for some $m \in \boldsymbol{Z}$. [Hint: Let $w \in W \backslash 0$. Then $w \in Fu_k + \cdots + Fu_\ell$ for some integers $k < \ell$. Let $\mathcal{W}$ be the smallest subspace of $\boldsymbol{V}(\lambda)$ containing w and invariant under $h\cdot$. Then $\mathcal{W} \subseteq W, Fu_k + \cdots + Fu_\ell$. The latter inclusion implies that $\mathcal{W}$ is finite-dimensional and $h\cdot|_{\mathcal{W}} : \mathcal{W} \longrightarrow \mathcal{W}$ is a semisimple endomorphism of $\mathcal{W}$.]

(c) Suppose that $\lambda \notin \boldsymbol{Z}$. Show that $\boldsymbol{V}(\lambda)$ is simple.

(d) Suppose that $\lambda \in \boldsymbol{Z}$. Show that $\boldsymbol{V}(\lambda)$ has exactly one proper submodule W, and that W is simple. (In this case $\boldsymbol{V}(\lambda)$ is not completely reducible.)

Exercise 2.4.7 Try your hand at Exercise 7 on page 34 of the text.

Exercise 2.4.8 Suppose that F is a field of characteristic zero, not necessarily algebraically closed, let $L = sl(2, F)$, and let $\{x, y, h\}$ be the standard basis for L. For a non-negative integer $m \geq 0$ show that there is a simple left L-module of dimension $m+1$ with basis $\{v_0, \ldots, v_m\}$ which satisfies (2.22). [Hint: See Exercise 2.4.5.]

Exercise 2.4.9 Suppose that L is a Lie algebra over a field F and let V be a left L-module. Let $\mathrm{End}_L(V)$ be the set of all module maps $f : V \longrightarrow V$.

(a) Show that $\mathrm{End}_L(V)$ is a subalgebra of the associative algebra $\mathrm{End}(V)$.

(b) Determine $\mathrm{End}_L(V)$ where $V = \boldsymbol{V}(\lambda)$ is the $L = sl(2, F)$-module of Exercise 2.4.6.

(c) Determine $\mathrm{End}_L(V)$ where $V = W$ is the simple $L = sl(2, F)$-module of dimension $m + 1$ described in this section. Do not assume that F is algebraically closed.

See Shur's Lemma for Lie algebras in connection with parts (b) and (c).

Exercise 2.4.10 Let $\boldsymbol{V}(\lambda), \boldsymbol{V}(\lambda')$ be the $L = sl(2, F)$-modules of Exercise 2.4.6. Show that $\boldsymbol{V}(\lambda) \simeq \boldsymbol{V}(\lambda')$ as left L-modules if and only if $\lambda = \lambda'$. [Hint: Let $\{u'_\imath\}_{\imath \in \boldsymbol{Z}}$ be the basis for $\boldsymbol{V}(\lambda')$ described in Exercise 2.4.6 and suppose that $f\boldsymbol{V}(\lambda) \longrightarrow \boldsymbol{V}(\lambda')$ is a module isomorphism. Show that $f(u_r) = \alpha u'_0$ for some $r \in \boldsymbol{Z}$ and $\alpha \in F\backslash 0$. Thus, replacing f by $(1/\alpha)f$, we may assume that $f(u_0) = u'_0$. Show that $f(u_{r+\imath}) = u'_\imath$ for all $\imath \geq 0$.]

2.5 Root Space Decompositions

L is a finite-dimensional non-zero semisimple Lie algebra. Recall that the fundamental representation $\phi : L \longrightarrow gl(L), \quad x \mapsto \mathrm{ad}\, x,$ is injective. Let $x, y \in L$. Since $\mathrm{ad}\,[x\, y] = [\mathrm{ad}\, x\ \mathrm{ad}\, y]$ it follows that $[x\, y] = 0$ *if and only if* $\mathrm{ad}\, x$ *and* $\mathrm{ad}\, y$ *are commuting operators.*

Lemma 2.5.1 *Let $x, y \in L$ and suppose that $[x\, y] = 0$. Then:*

(a) $\operatorname{ad} x$ *commutes with* $\operatorname{ad} y$, $\operatorname{ad} y_s$, *and* $\operatorname{ad} y_n$.

(b) $[x\ y_s] = 0 = [x\ y_n]$.

(c) $\kappa(x, y) = \kappa(x_s, y_s)$.

PROOF: Since $[x\, y] = 0$ the operators $\operatorname{ad} x$ and $\operatorname{ad} y$ commute. Now $(\operatorname{ad} y)_s = \operatorname{ad} y_s$ and $(\operatorname{ad} y)_n = \operatorname{ad} y_n$ are polynomials in $\operatorname{ad} y$. Therefore these operators commute with $\operatorname{ad} x$. We have shown parts (a) and (b).

To show part (c) we observe that $\kappa(x, y) = \kappa(x, y_s + y_n) = \kappa(x, y_s) + \kappa(x, y_n)$. Now $\kappa(x, y_n) = \operatorname{tr}(\operatorname{ad} x \circ \operatorname{ad} y_n)$. Since $\operatorname{ad} x$ and $\operatorname{ad} y_n$ commute by part (a), and $\operatorname{ad} y_n$ is nilpotent, for some $m \geq 0$ we have $(\operatorname{ad} x \circ \operatorname{ad} y_n)^m = (\operatorname{ad} x)^m \circ (\operatorname{ad} y_n)^m = 0$. Since the trace of a nilpotent endomorphism is zero, $\kappa(x, y_n) = \operatorname{Tr}(\operatorname{ad} x \circ \operatorname{ad} y_n) = 0$. We have shown that $\kappa(x, y) = \kappa(x, y_s)$. Since κ is symmetric, and $[y_s\, x] = 0$ by part (b), the calculation $\kappa(x, y_s) = \kappa(y_s, x) = \kappa(y_s, x_s)$ completes the proof of part (c). □

2.5.1 Maximal toral subalgebras and roots

Buried in this section is an analog of Engel's Theorem. Let L be a finite-dimensional Lie algebra over a field F. We will say that $x \in L$ is ad-*semisimple* if $\operatorname{ad} x$ is a semisimple endomorphism of L. (This must be a standard definition.) If L is abelian then all $x \in L$ are ad-semisimple as $\operatorname{ad} x = 0$. The converse is true as well.

Lemma 2.5.2 *Let L be a finite-dimensional Lie algebra over the field F. Suppose $x, y \in L \backslash 0$ where y is* ad-*semisimple and* $\operatorname{ad} x(y) = \lambda y$ *for some $\lambda \in F$. Then $\lambda = 0$.*

PROOF: Let $\lambda_1, \ldots, \lambda_r \in F$ be the distinct eigenvalues for $\operatorname{ad} y$ and let $V_1, \ldots, V_r$ be the corresponding subspaces of eigenvectors. Then $L = V_1 \oplus \cdots \oplus V_r$. Recall that 0 is an eigenvalue for $\operatorname{ad} y$ since $0 = [y\, y] = \operatorname{ad} y(y)$. We may take $\lambda_r = 0$.

By assumption $[x\, y] = \operatorname{ad} x(y) = \lambda y$. Write $x = x_1 + \cdots + x_r$ where $x_\imath \in V_\imath$ for all $1 \leq \imath \leq r$. Applying $\operatorname{ad} y$ to both sides of this equation we calculate

$$-\lambda y = [y\, x] = \lambda_1 x_1 + \cdots + \lambda_r x_r = \lambda_1 x_1 + \cdots + \lambda_{r-1} x_{r-1}$$

and thus

$$\lambda_1 x_1 + \cdots + \lambda_{r-1} x_{r-1} + \lambda y = 0.$$

Since $y \in V_r$ we conclude that $\lambda_1 x_1 = \cdots = \lambda y = 0$. $\square$

Theorem 2.5.3 *Let L be a finite-dimensional Lie algebra over a field F which consists of* ad*-semisimple elements. Then L is abelian.*

PROOF: Let $x \in L \backslash 0$. Then the only eigenvalue of $\operatorname{ad} x$ is 0 by Lemma 2.5.2. Therefore $\operatorname{ad} x = 0$ since $\operatorname{ad} x$ is semisimple. $\square$

Now suppose that L is a finite-dimensional semisimple Lie algebra and F is an algebraically closed field of characteristic zero. Suppose that T is a toral subalgebra of L and let $x \in T$. Then $\operatorname{ad}_L x$ is semisimple by definition. Since $\operatorname{ad}_L x(T) \subseteq T$ the restriction $\operatorname{ad}_L x|_T = \operatorname{ad}_T x$ is semisimple. There T consists of ad-semisimple elements and hence T is abelian.

Let V be a finite-dimensional vector space over the field F and suppose that $\mathcal{S}$ is a non-empty subset of endomorphisms of V. Let $\mathbf{\Phi}$ be the set of all functions $\lambda : \mathcal{S} \longrightarrow F$ and set

$$V_\lambda = \{v \in V \mid S(v) = \lambda(S)v \ \text{ for all } S \in \mathcal{S}\}.$$

Observe that $V_\lambda = \cap_{S \in \mathcal{S}} \operatorname{Ker}(S - \lambda(S)I)$ and is therefore a subspace of V. We leave the reader with the important exercise of showing that

$$\sum_{\lambda \in \mathbf{\Phi}} V_\lambda = \oplus_{\lambda \in \mathbf{\Phi}} V_\lambda.$$

Lemma 2.5.4 *Let V be a finite-dimensional vector space over the field F and suppose that $\mathcal{S}$ is a non-empty family of commuting semisimple endomorphisms of V. Then:*

(a) *V has a basis which consists of eigenvectors for all $S \in \mathcal{S}$, or equivalently $V = \oplus_{\lambda \in \mathbf{\Phi}} V_\lambda$.*

(b) *The linear span of $\mathcal{S}$ is a family of commuting semisimple endomorphisms of V.*

PROOF: We first show part (a). Suppose all $S \in \mathcal{S}$ have one eigenvalue. Then $S \in \mathcal{S}$ has the form $S = \alpha I$ for some $\alpha \in F$. Thus any basis for V will do.

We may assume that some $S \in \mathcal{S}$ has at least two eigenvalues. Let $\rho_1, \ldots, \rho_r$ be the distinct eigenvalues of S. Then $V = V_1 \oplus \cdots \oplus V_r$, where $V_\imath = \mathrm{Ker}(S - \rho_\imath I)$ is the subspace of eigenvectors for S belonging to $\rho_\imath$. Let $S' \in \mathcal{S}$. Since S' commutes with S by assumption, it follows that each of the subspaces $V_\imath$ of V is invariant under S'. For each $1 \leq \imath \leq r$ the hypothesis of the lemma applies to the set of restrictions $\mathcal{S}_\imath = \{T|_{V_\imath} \,|\, T \in \mathcal{S}\}$. As $\mathrm{Dim}\, V_\imath < \mathrm{Dim}\, V$ for all $1 \leq \imath \leq r$, part a) follows by induction on $\mathrm{Dim}\, V$.

By part (a) there is a basis $\{v_1, \ldots, v_n\}$ for V such that $S(v_\imath) \in Fv_\imath$ for all $S \in \mathcal{S}$ and $1 \leq \imath \leq n$. The set of all such linear endomorphisms of V which satisfy this property forms a subspace $\mathcal{W}$ of $\mathrm{End}(V)$ which contains $\mathcal{S}$ and any $T, T' \in \mathcal{W}$ commute. Thus part (b) follows. $\square$

Now suppose that H is a toral subalgebra of L, not necessarily maximal. Then H is abelian. By part (a) of Lemma 2.5.1 the set $\mathcal{S} = \{\mathrm{ad}\, h \,|\, h \in H\}$ is a family of commuting diagonalizable operators on L. Thus

$$L = \bigoplus_{\alpha:\mathcal{S} \longrightarrow F} L_\alpha, \tag{2.24}$$

where

$$\begin{aligned} L_\alpha &= \{x \in L \,|\, \mathrm{ad}\, h(x) = \alpha(\mathrm{ad}\, h)x \;\; \forall h \in H\} \\ &= \{x \in L \,|\, [h\; x] = ((\alpha \circ \phi)(h))x \;\; \forall h \in H\}. \end{aligned}$$

Observe that the restriction $(\alpha \circ \phi)|_H : H \longrightarrow F$ is *linear*. By slight abuse of notation we replace $(\alpha \circ \phi)|_H$ by α and thus have

$$L = \bigoplus_{\alpha \in H^*} L_\alpha, \tag{2.25}$$

where

$$L_\alpha = \{x \in L \,|\, [h\; x] = \alpha(h)x \;\; \forall h \in H\}.$$

Observe that $L_0 = \mathrm{C}_L(H)$. Since H is abelian $H \subseteq \mathrm{C}_L(H)$. Since L is finite-dimensional there are only finitely many L_α's such that $L_\alpha \neq (0)$. Let

$$\mathbf{\Phi} = \{\alpha \in H^* \,|\, \alpha \neq 0 \text{ and } L_\alpha \neq (0)\}.$$

Then

$$L = \mathrm{C}_L(H) \oplus \bigoplus_{\alpha \in \mathbf{\Phi}} L_\alpha. \tag{2.26}$$

When H is a maximal toral subalgebra the elements of $\mathbf{\Phi}$ are called *roots of L relative to H*, or more informally *roots*.

We set the stage for the proof of parts (b) and (c) of the Proposition below. Let $0 \neq \beta \in H^*$ and suppose that T is a linear endomorphism of L such that $T(L_\alpha) \subseteq L_{\beta+\alpha}$ for all $\alpha \in H^*$. Fix $\alpha \in H^*$. Then $T^n(L_\alpha) \subseteq L_{n\beta+\alpha}$ for all $n \geq 0$. Since $\beta \neq 0$ the terms of the sequence $\alpha, \beta+\alpha, 2\beta+\alpha, \ldots$ are distinct. Since L is finite-dimensional, we see from (2.25) that $L_{n_\alpha\beta+\alpha} = (0)$ for some $n_\alpha \geq 0$. Thus $T^m(L_\alpha) = (0)$ for all $m \geq n_\alpha$.

Let $\mathbf{\Phi} = \{\alpha_1, \ldots, \alpha_r\}$ and n be the maximum of $n_0, n_{\alpha_1}, \ldots, n_{\alpha_r}$. Then $T^n(L_\alpha) = (0)$ for all $\alpha \in H^*$. Therefore $T^n = 0$; in particular T is nilpotent.

Proposition 2.5.5 *Let $\alpha, \beta \in H^*$. Then:*

(a) $[L_\alpha\, L_\beta] \subseteq L_{\alpha+\beta}$.

(b) *Suppose $\beta \neq 0$ and $x \in L_\beta$. Then* $\operatorname{ad} x$ *is nilpotent.*

(c) *$\kappa(L_\alpha, L_\beta) = (0)$ unless $\alpha + \beta = 0$. Furthermore $\kappa(L_\alpha, L_{-\alpha}) = (0)$ implies $L_\alpha = (0)$.*

Proof: Let $H \in h$. Since $\operatorname{ad} h$ is a derivation of L, for $x \in L_\alpha$ and $y \in L_\beta$ the calculation

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

shows that $[x\,y] \in L_{\alpha+\beta}$. We have shown part (a).

Assume the hypothesis of part (b) and set $T = \operatorname{ad} x$. Then $T(L_\alpha) = [x\,L_\alpha] \subseteq L_{\beta+\alpha}$ by part (a). Thus T is nilpotent.

Let $x \in L_\alpha, y \in L_\beta$ and $T = \operatorname{ad} x \circ \operatorname{ad} y$. Then $T(L_\gamma) = [x\,[y\;L_\gamma]] \subseteq [x\,L_{\beta+\gamma}] \subseteq L_{(\alpha+\beta)+\gamma}$ for all $\gamma \in H^*$. If $\alpha+\beta \neq 0$ then T is nilpotent and thus $\kappa(x, y) = \operatorname{Tr}(T) = 0$.

We have shown that $\kappa(L_\alpha, L_\beta) = (0)$ unless $\beta = -\alpha$. If $\kappa(L_\alpha, L_{-\alpha}) = (0)$ then $\kappa(L_\alpha, L_\beta) = (0)$ for all $\beta \in H^*$ which means $\kappa(L_\alpha, L) = (0)$ by (2.25). In this case $L_\alpha = (0)$ as κ is non-singular. □

Corollary 2.5.6 *Let* $\mathrm{C} = \mathrm{C}_L(H)$. *Then:*

(a) *The restriction $\kappa|_{C\times C}$ is non-degenerate.*

(b) *Let $x \in C$. Then $x_s, x_n \in C$.*

Suppose that H is a maximal toral subalgebra of L. Then:

(c) *All semisimple elements of C belong to H.*

(d) $\mathrm{Z}(C) = H$.

Proof: To show part (a) we let $x \in C$ and suppose that $(0) = \kappa(x, C) = \kappa(x, L_0)$. Let $\beta \in \mathbf{\Phi}$. Since $0 + \beta \neq 0$, by part (c) of Proposition 2.5.5 we conclude $\kappa(x, L_\beta) = (0)$. Therefore $\kappa(x, L) = (0)$ by (2.26). Since κ is non-singular $x = 0$. Part (b) follows by part (b) of Lemma 2.5.1.

Part (c). Let $x \in C$ be semisimple. Now $[H + Fx\ H + Fx] \subseteq [H\ C] + [C\ H] + [Fx\ Fx] = (0)$ means that $H + Fx$ is an abelian subalgebra of L. Note that $\mathrm{ad}\, h$, $\mathrm{ad}\, x$ commute for all $h \in H$ by part (a) of Lemma 2.5.1. Thus for all $\alpha \in F$ the operator $\mathrm{ad}\,(h + \alpha x) = \mathrm{ad}\, h + \alpha \mathrm{ad}\, x$ is diagonalizable; see part (b) of Lemma 2.5.4. Therefore $H + Fx$ is toral. Since $H \subseteq H + Fx$ and the former is a maximal toral subalgebra of L, $H = H + Fx$ which means $x \in H$.

Part (d). Since $H \subseteq C$ necessarily $H \subseteq \mathrm{Z}(C)$ by definition of C. Conversely, let $x \in \mathrm{Z}(C)$. By part (b) $x_s, x_n \in C$ and thus $x_s, x_n \in \mathrm{Z}(C)$ by part (b) of Lemma 2.5.1 again. Since $[C\ x_n] = (0)$, $\kappa(C, x_n) = (0)$ by part (c) of Lemma 2.5.1. Therefore $x_n = 0$ by part (a). Thus $x = x_s \in H$ by part (c). □

2.5.2 The Centralizer of H

Theorem 2.5.7 *Suppose H is a maximal toral subalgebra of L. Then $H = \mathrm{C}_L(H)$.*

Proof: In light of part (d) of Corollary 2.5.6 we need only show that C is abelian.

First we note that C is nilpotent. Let $x \in C$. Then $\mathrm{ad}_C\, x = \mathrm{ad}_C\, x_s + \mathrm{ad}_C\, x_n = \mathrm{ad}\, x_n|_C$ since $[x_s\ C] = (0)$ by parts (b) and (c) of Corollary 2.5.6. Thus $\mathrm{ad}_C\, x$ is nilpotent for all $x \in C$ which means C is nilpotent by Engel's Theorem.

Next we note that $\mathrm{Z}(C) \cap [C\ C] = (0)$, which is equivalent to $\kappa(C, \mathrm{Z}(C) \cap [C\ C]) = (0)$ by part (a) of Corollary 2.5.6. Using part (c) of Lemma 2.5.1 and part (c) of Corollary 2.5.6 we compute

$$\kappa(C, \mathrm{Z}(C) \cap [C\ C]) \subseteq \kappa(H, \mathrm{Z}(C) \cap [C\ C]) \subseteq \kappa(H, [C\ C]) = \kappa([H\ C], C) = \kappa((0), C) = (0).$$

Since $\mathrm{Z}(C) \cap [C\ C] = (0)$ it follows that $[C\ C] = (0)$ by 3.3 Lemma of Humphreys. $\square$

2.5.3 Orthogonality Properties

Let H be a maximal toral subalgebra of L. Then $\mathrm{C}_L(H) = H$; in particular the restriction $\kappa|_{H \times H}$ is non-degenerate. Thus $\kappa_\ell : H \longrightarrow H^*$,given by $h \mapsto \kappa_\ell(h)$, where $\kappa_\ell(h) : H \longrightarrow F$ is defined by

$$\kappa_\ell(h)(x) = \kappa(h, x)$$

for all $x \in H$, is injective. Since $\mathrm{Dim}\, H = \mathrm{Dim}\, H^*$ it follows that κ_ℓ is bijective. Thus for $\alpha \in H^*$ there is a unique $t_\alpha \in H$ such that $\kappa_\ell(t_\alpha) = \alpha$. Observe that t_α is determined by

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

for all $x \in H$.

We will use the following lemma in the proof of the next proposition.

Lemma 2.5.8 *Let X, Y, T be endomorphisms of a finite-dimensional vector space V over F which satisfy $T = [X\ Y]$ and $[T\ X] = 0 = [T\ Y]$. Then T is nilpotent.*

Proof: We may assume $V \neq (0)$. Let $\lambda \in F$ be an eigenvalue for T and $U = \{v \in V \,|\, T(v) = \lambda v\}$. Since X and Y commute with T it follows that $X(U) \subseteq U$ and $Y(U) \subseteq U$. From

$$\lambda \mathrm{Id}_U = T|_U = [X\ Y]|_U = [X|_U\ Y|_U]$$

we calculate $\lambda \mathrm{Dim}\, U = \mathrm{Tr}(T|_U) = \mathrm{Tr}([X|_U\ Y|_U]) = 0$. Since $\mathrm{Dim}\, U > 0$ and the characteristic of F is zero, $\lambda = 0$. Therefore T is nilpotent. $\square$

Proposition 2.5.9 *Let H be a maximal toral subalgebra of L. Then:*

(a) $\mathbf{\Phi}$ *spans* H^*. *Thus* $|\mathbf{\Phi}| \geq \mathrm{Dim}\, H$ *and the* t_α*'s span* H.

Let $\alpha \in \mathbf{\Phi}$. *Then:*

(b) $-\alpha \in \mathbf{\Phi}$.

(c) *If* $x \in L_\alpha$ *and* $y \in L_{-\alpha}$ *then* $[x\, y] = \kappa(x, y) t_\alpha$.

(d) $[L_\alpha\, L_{-\alpha}] = Ft_\alpha$. *Thus L is generated as a Lie algebra by the root spaces L_α.*

(e) $\alpha(t_\alpha) = \kappa(t_\alpha, t_\alpha) \neq 0$.

(f) *There are $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$ such that the span of x_α, y_α, and $h_\alpha = [x_\alpha\, y_\alpha]$ is a Lie-subalgebra $\mathrm{S}_\alpha \simeq sl(2,F)$; indeed $[h_\alpha\, x_\alpha] = 2x_\alpha$ and $[h_\alpha\, y_\alpha] = -2y_\alpha$.*

(g) *All h_α's arising in part (f) are the same; $h_\alpha = \dfrac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)} = \dfrac{2t_\alpha}{\alpha(t_\alpha)}$. Furthermore $h_{-\alpha} = -h_\alpha$.*

Proof: Let $\mathcal{V}$ be the span of $\mathbf{\Phi} \subseteq H^*$. Suppose that $f \in H^{**}$ vanishes on $\mathcal{V}$. Since H is finite-dimensional, the linear map $H \longrightarrow H^{**}$, given by $h \longrightarrow \widehat{h}$, where $\widehat{h}(\alpha) = \alpha(h)$ for all $\alpha \in H^*$, is an isomorphism. Thus $f = \widehat{h}$ for some $h \in H$. This means $0 = f(\alpha) = \widehat{h}(\alpha) = \alpha(h)$, hence $[h\, L_\alpha] = (0)$, for all $\alpha \in \mathbf{\Phi}$. Since $[h\, L_0] = [h\, H] = (0)$ it follows that $[h\, L] = (0)$ by (2.25). We have shown $h \in \mathrm{Z}(L) = (0)$ which means that the only function $f \in H^{**}$ which vanishes on $\mathcal{V}$ is $f = 0$. Consequently $\mathcal{V} = H^*$ which establishes part (a).

Part (c) of Proposition 2.5.5 implies $\kappa(L_\alpha, L_{-\alpha}) \neq (0)$; in particular $L_{-\alpha} \neq (0)$ which establishes part (b). Suppose $x \in L_\alpha$ and $y \in L_{-\alpha}$. Then $[x\, y] \in L_0 = H$. For all $h \in H$ the calculation

$$\kappa_\ell(\kappa(x,y)t_\alpha)(h) = \kappa(x,y)\alpha(h) = \kappa([h\, x], y) = \kappa(h, [x\, y]) = \kappa([x\, y], h) = \kappa_\ell([x\, y])(h)$$

shows that $\kappa_\ell(\kappa(x,y)t_\alpha) = \kappa_\ell([x\, y]t_\alpha)$. Since κ_ℓ is injective $\kappa(x,y)t_\alpha = [x\, y]$. We have shown part (c). In light of (2.25) part (d) follows from part (c) since $\kappa(L_\alpha, L_{-\alpha}) \neq (0)$ and the t_α's span H.

There are $x \in L_\alpha$ and $y \in L_{-\alpha}$ such that $[x\, y] = t_\alpha$ by part (d). Suppose $\alpha(t_\alpha) = 0$. Since $[t_\alpha\, x] = \alpha(t_\alpha)x = 0$ and $[t_\alpha\, y] = (-\alpha)(t_\alpha)y = 0$, Lemma 2.5.8 applied to $T = \operatorname{ad} t_\alpha$, $X = \operatorname{ad} x$, and $Y = \operatorname{ad} y$ shows that $\operatorname{ad} t_\alpha$ is nilpotent. Since this operator is also diagonalizable, $\operatorname{ad} t_\alpha = 0$ and therefore $t_\alpha = 0$, a contradiction. Thus $\alpha(t_\alpha) \neq 0$ after all. We have shown part (e).

Let $h_\alpha = \dfrac{2t_\alpha}{\alpha(t_\alpha)}$. By part (d) there are $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$ such that $[x_\alpha\, y_\alpha] = h_\alpha$. Since $\alpha(h_\alpha) = 2$, $[h_\alpha\, x_\alpha] = \alpha(h_\alpha)x = 2x_\alpha$ and $[h_\alpha\, y_\alpha] = (-\alpha)(h_\alpha)y = -2y_\alpha$. Since $L_{-\alpha} + L_0 + L_\alpha$ is direct and y_α, h_α, x_α are not zero, the span S_α of these vectors is 3-dimensional. We have established part (f).

Part (g). $h_\alpha = at_\alpha$ for some $a \in F$ by part (c). From $\alpha(h_\alpha)x_\alpha = [h_\alpha\, x_\alpha] = 2x_\alpha$ we deduce $a\alpha(t_\alpha) = \alpha(h_\alpha) = 2$. Since $\kappa_\ell(-t_\alpha) = -\kappa_\ell(t_\alpha) = -\alpha$ it follows that $-t_\alpha = t_{-\alpha}$. $\square$

2.5.4 Integrality Properties

We continue with the notation of the preceding section and will use the results of the two preceding sections without particular reference.

Proposition 2.5.10 *Let H be a maximal toral subalgebra of L and $\mathbf{\Phi}$ the corresponding set of roots and let $\alpha, \beta \in \mathbf{\Phi}$. Then:*

(a) $\operatorname{Dim} L_\alpha = 1$. *(In particular for any non-zero $x \in L_\alpha$ there is a unique $y \in L_{-\alpha}$ such that $[x\, y] = h_\alpha$.)*

(b) *Let $c \in F$. Then $c\alpha \in \mathbf{\Phi}$ if and only if $c = \pm 1$.*

(c) $\beta(h_\alpha) \in \mathbf{Z}$ *and* $\beta - \beta(h_\alpha)\alpha \in \mathbf{\Phi}$.

(d) $\alpha + \beta \in \mathbf{\Phi}$ *then* $[L_\alpha\, L_\beta] = L_{\alpha+\beta}$.

(e) *Suppose $\beta \neq \pm\alpha$ and let $r, q \geq 0$ be the largest non-negative integers such that $\beta - r\alpha, \beta + q\alpha \in \mathbf{\Phi}$. Then $\beta + i\alpha \in \mathbf{\Phi}$ for all $-r \leq i \leq q$ and $\beta(h_\alpha) = r - q$.*

Proof: We first show parts (a) and (b). Note that the parenthetical part of part (a) follows from the preceding sentence and part (f) of Proposition 2.5.9.

Regard L as a left S_α-module under the adjoint action. For a left S_α-submodule V of L and $\lambda \in F$ recall that

$$V_\lambda = \{v \in V \mid [h_\alpha\, v] = \lambda v\}.$$

Let $\alpha \in \mathbf{\Phi}$ and set $M = \bigoplus_{c \in F} L_{c\alpha}$. Note that M is a subalgebra of L since $[L_{c\alpha}\, L_{d\alpha}] \subseteq L_{(c+d)\alpha}$ for all $c, d \in F$. Since $S_\alpha = Fy_\alpha + Fh_\alpha + Fx_\alpha \subseteq L_{-\alpha} + L_0 + L_\alpha$ it follows that S_α is a subalgebra of M. Therefore M is a left S_α-module under the adjoint action.

Note that $L_0 = H = \mathrm{Ker}\,\alpha \oplus Fh_\alpha$. Observe that $S_\alpha \subseteq M' = Fy_\alpha + L_0 + Fx_\alpha = \mathrm{Ker}\,\alpha \oplus S_\alpha$ and the latter is a subalgebra of L as well. Thus M' is a S_α-submodule. That M' is a subalgebra of M follows from $[h\,H] = (0)$, $[h\,x_\alpha] = \alpha(h)x_\alpha = 0$, and $[h\,y_\alpha] = -\alpha(h)y_\alpha = 0$ for $h \in \mathrm{Ker}\,\alpha$. These calculations show that as a direct sum of simple S_α-modules $M' = \left(\bigoplus_{i=1}^n Fh_i\right) \oplus S_\alpha$, where $\{h_1, \ldots, h_n\}$ is a basis for $\mathrm{Ker}\,\alpha$.

Recall that $\alpha(h_\alpha) = 2$. For $x \in L_{c\alpha}$ the calculation $[h_\alpha\, x] = c\alpha(h_\alpha)x = 2cx$ shows that $L_\alpha \subseteq L_{2c}$. Since $\sum_{c\in F} L_{2c}$ is direct,

$$L_{c\alpha} = M_{2c} \text{ for all } c \in F \text{ and thus } M = \bigoplus\nolimits_{c\in F} M_{2c}.$$

Since S_α is semisimple M is the direct sum of simple S_α-modules S. Now S is the direct sum of one-dimensional weight spaces. Thus $S = \bigoplus_{c\in F} S \cap M_{2c}$. Recall that the weights of S are $-m+2\ell$, where $0 \le \ell \le m$, for some $m \ge 0$.

Suppose that $c\alpha \in \mathbf{\Phi}$. Then some simple S_α-submodule S of M intersects $L_{c\alpha} = M_{2c}$. Thus $2c \in \mathbf{Z}$.

Suppose that $2c$ is even. Then S intersects $L_0 = H$ which means $S \subseteq M'$. The weights of M' are $-2, 0, 2$. Therefore $c = \pm 1$. *In particular* $2\alpha \notin \mathbf{\Phi}$.

Suppose that $2c$ is odd. Then S intersects $M_1 = L_{(1/2)\alpha}$. Since twice $(1/2)\alpha$ is a root, this is not possible.

We have shown $M = M'$ which means that $\mathrm{Dim}\, L_\alpha = 1$ and $c\alpha \in \mathbf{\Phi}$ if and only if $c = \pm 1$. Parts (a) and (b) are established.

Now let $\beta \in \mathbf{\Phi}$ where $\beta \neq \pm\alpha$. Observe that $N = \bigoplus_{i\in\mathbf{Z}} L_{\beta+i\alpha}$ is a left S_α-submodule of L,

$$N_{\beta(h_\alpha)+2i} = L_{\beta+i\alpha}, \text{ and thus } N = \bigoplus\nolimits_{i\in\mathbf{Z}} N_{\beta(h_\alpha)+2i}.$$

Suppose that $\beta + i\alpha = 0$. Then $\beta = -i\alpha$ is a root; therefore $i = \pm 1$, a contradiction. Thus $\beta + i\alpha \neq 0$ for all $i \in \mathbf{Z}$.

Since there are no integers i, i' which satisfy $\beta(h_\alpha) + 2i = 0$ and $\beta(h_\alpha) + 2i' = 1$, there can be no $n, n' \in N$ with weights $0, 1$ respectively. Since root spaces are one-dimensional and N is the direct sum of simple S_α-modules, N is simple. For some $m \ge 0$ the weights of N are $-m + 2i$, where $0 \le i \le m$. Write

$$-m = \beta(h_\alpha) + 2(-r) \text{ and } m = \beta(h_\alpha) + 2q.$$

Then $r, q \ge 0$ since $\beta(h_\alpha)$ is a weight as $\beta \in \mathbf{\Phi}$. Therefore

$$N = \bigoplus_{i=0}^{m} M_{-m+2i} = \bigoplus_{i=0}^{m} M_{\beta(h_\alpha)+2(-r+i)} = \bigoplus_{i=-r}^{q} M_{\beta(h_\alpha)+2i} = \bigoplus_{i=-r}^{q} L_{\beta+i\alpha};$$

thus $\beta + i\alpha \in \mathbf{\Phi}$ if and only if $-r \leq i \leq q$. Since $-m = -2r + \beta(h_\alpha)$ and $m = 2q + \beta(h_\alpha)$ it follows that $\beta(h_\alpha) = r - q$. Since $-r \leq -r + q \leq q$, $\beta - \beta(h_\alpha)\alpha = \beta + (-r+q)\alpha \in \mathbf{\Phi}$. As $\alpha(h_\alpha) = 2$, and $\alpha - \alpha(h_\alpha)\alpha = -\alpha \in \mathbf{\Phi}$ by part (b), we have established parts (c)–(f). □

2.5.5 Rationality properties. Summary

We continue with the notation of the previous two sections. First the very important:

$$\kappa(h,k) = \sum_{\alpha \in \mathbf{\Phi}} \alpha(h)\alpha(k) \tag{2.27}$$

for all $h, k \in H$.

PROOF: Let $h, k \in H$. Then for $\alpha \in H^*$ and $x \in L_\alpha$ we have

$$(\operatorname{ad} h \circ \operatorname{ad} k)(x) = [h\,[k\,x]] = \alpha(h)\alpha(k)x.$$

Therefore L_α is invariant under $\operatorname{ad} h \circ \operatorname{ad} k$ and $(\operatorname{ad} h \circ \operatorname{ad} k)|_{L_\alpha} = \alpha(h)\alpha(k)\mathrm{Id}_{L_\alpha}$. As $L = L_0 \oplus \left(\bigoplus_{\alpha \in \mathbf{\Phi}} L_\alpha\right)$ we have

$$\begin{aligned}
\kappa(h,k) &= \operatorname{Tr}(\operatorname{ad} h \circ \operatorname{ad} k) \\
&= \operatorname{Tr}((\operatorname{ad} h \circ \operatorname{ad} k)|_{L_0}) + \sum_{\alpha \in \mathbf{\Phi}} \operatorname{Tr}((\operatorname{ad} h \circ \operatorname{ad} k)|_{L_\alpha}) \\
&= 0(h)0(k)\operatorname{Dim} H + \sum_{\alpha \in \mathbf{\Phi}} \alpha(h)\alpha(k)\operatorname{Dim} L_\alpha \\
&= \sum_{\alpha \in \mathbf{\Phi}} \alpha(h)\alpha(k).
\end{aligned}$$

□

Let $\alpha, \beta \in H^*$. Since κ symmetric

$$\alpha(t_\beta) = \kappa(t_\alpha, t_\beta) = \beta(t_\alpha). \tag{2.28}$$

Now suppose that $\alpha, \beta \in \mathbf{\Phi}$. Then

$$\frac{2\beta(t_\alpha)}{\alpha(t_\alpha)} = \beta(h_\alpha) \in \mathbf{Z} \tag{2.29}$$

by part (c) of Proposition 2.5.10. Using (2.27) and (2.28) we deduce $\alpha(t_\alpha) = \sum_{\gamma \in \mathbf{\Phi}} \gamma(t_\alpha)^2$ from which

$$\frac{4}{\alpha(t_\alpha)} = \sum_{\gamma \in \mathbf{\Phi}} \left(\frac{2\gamma(t_\alpha)}{\alpha(t_\alpha)} \right)^2 \in \mathbf{Z}$$

follows. Thus $\alpha(t_\alpha) \in \mathbf{Q}$ which means

$$\kappa(t_\alpha, t_\beta) = \kappa(t_\beta, t_\alpha) = \beta(t_\alpha) \in \mathbf{Q} \tag{2.30}$$

by (2.29).

Let $\{t_{\alpha_1}, \ldots, t_{\alpha_n}\}$ be a basis for H over F and $A = (a_{i\,j}) \in \mathrm{M}(n, F)$ be the matrix of the bilinear form $\kappa|_{H \times H}$. By definition $a_{i\,j} = \kappa(t_{\alpha_i}, t_{\alpha_j})$ for all $1 \leq i, j \leq n$. Since κ is non-degenerate and symmetric A is invertible and symmetric. Now $A \in \mathrm{M}(n, \mathbf{Q})$ by (2.30). Since $\mathrm{Det}\, A \neq 0$ we conclude that $A^{-1} \in \mathrm{M}(n, \mathbf{Q})$. Note that

$$\kappa(\sum_{i=1}^n c_i t_{\alpha_i}, \sum_{j=1}^n d_j t_{\alpha_j}) = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}^t A \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} \tag{2.31}$$

for all $c_1, \ldots, c_n, d_1, \ldots, d_n \in F$.

Let $\alpha \in \mathbf{\Phi}$. Then $t_\alpha = \sum_{j=1}^n c_j t_{\alpha_j}$ for some $c_1, \ldots, c_n \in F$. For $1 \leq i \leq n$ set $d_i = \kappa(t_{\alpha_i}, t_\alpha)$. Applying $\kappa_\ell(t_{\alpha_i})$ to both sides of the first equation we obtain $d_i = \sum_{j=1}^n c_j a_{i\,j}$; thus

$$A \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} \in \mathbf{Q}^n$$

which means

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = A^{-1} \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} \in \mathbf{Q}^n.$$

Therefore $c_1, \ldots, c_n \in \mathbf{Q}$.

Let $\boldsymbol{E}_{\boldsymbol{Q}}$ be the $\mathbf{Q}$-span of $\mathbf{\Phi}$. We have shown that $\{t_{\alpha_1}, \ldots, t_{\alpha_n}\}$ is a basis for $\boldsymbol{E}_{\boldsymbol{Q}}$ as a vector space over $\boldsymbol{Q}$. Now let $\boldsymbol{E}$ be the vector space over the field of real numbers $\boldsymbol{R}$ with basis of symbols $\{t_{\alpha_1}, \ldots, t_{\alpha_n}\}$. We regard $\boldsymbol{E}_{\boldsymbol{Q}}$ as a $\boldsymbol{Q}$-subspace of $\boldsymbol{E}$ by $t_\alpha \mapsto t_\alpha$.

Let (,) be the $\boldsymbol{R}$-bilinear form on $\boldsymbol{E}$ whose matrix is A. Then (,) is symmetric and non-degenerate since A is symmetric and non-degenerate. Observe that $(\, , \,)|_{\boldsymbol{E}_{\boldsymbol{Q}} \times \boldsymbol{E}_{\boldsymbol{Q}}} = \kappa|_{\boldsymbol{E}_{\boldsymbol{Q}} \times \boldsymbol{E}_{\boldsymbol{Q}}}$.

At this point we identify t_α and α. Thus $(\alpha, \beta) = \kappa(t_\alpha, t_\beta) = \alpha(t_\beta) = \beta(t_\alpha)$. Note that

$$(\mathbf{u}, \mathbf{v}) = \sum_{\gamma \in \boldsymbol{\Phi}} (\gamma, \mathbf{u})(\gamma, \mathbf{v}) \tag{2.32}$$

for all $\mathbf{u}, \mathbf{v} \in \boldsymbol{E}$ by (2.27) as both sides of this equation are $\mathbf{R}$-bilinear forms which agree on pairs $\mathbf{u} = \alpha$ and $\mathbf{v} = \beta$ which come from a spanning set. As a consequence of the last equation (,) is positive definite; that is $(\mathbf{u}, \mathbf{u}) \geq 0$ for all $\mathbf{u} \in \boldsymbol{E}$, with equality if and only if $\mathbf{u} = \mathbf{0}$.

Theorem 2.5.11 *Let $\boldsymbol{E}$, $\boldsymbol{\Phi}$, and (,) be the symmetric positive definite bilinear form above. Then:*

(a) *$\boldsymbol{\Phi}$ spans $\boldsymbol{E}$ and $\mathbf{0} \notin \boldsymbol{\Phi}$.*

Suppose $\alpha, \beta \in \boldsymbol{\Phi}$.

(b) *Let $c \in \boldsymbol{R}$. Then $c\alpha \in \boldsymbol{\Phi}$ f and only if $c = \pm 1$.*

(c) $\beta - \dfrac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha \in \boldsymbol{\Phi}$.

(d) $\dfrac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \boldsymbol{Z}$.

□

We may identify $\boldsymbol{E}$ with $\boldsymbol{R}^n$ and the inner product (,) with the standard inner product on $\boldsymbol{R}^n$. Thus finding the $\boldsymbol{\Phi}$'s of the theorem is a problem in the Euclidean space $\boldsymbol{R}^n$.

Chapter 3

Root Systems

Below are comments relevant to various sections in the text. They are meant to clarify, amplify, or generalize material in the text. Exercises are optional.

3.1 Axiomatics

3.1.1 Reflections in a euclidean space

We begin with some general comments about euclidean spaces, projections, and reflections. See Exercises 3.1.2–3.1.7 for details.

Suppose that V is any vector space over $\boldsymbol{R}$ and that $(\ ,\ ) : V \times V \longrightarrow \boldsymbol{R}$ is a symmetric positive definite bilinear form. Let $u \in V\backslash 0$. Then

$$P = u^{\perp} = \{v \in V \,|\, (u, v) = 0\}$$

is a codimension one subspace, or a hyperspace, of V. When V is finite-dimensional all codimension one subspaces of V have this form.

Define a linear function $\pi_u : V \longrightarrow V$ by

$$\pi_u(v) = v - \frac{(u, v)}{(u, u)} u$$

for all $v \in V$. Then $\pi_u(v) = v$ for all $v \in P$ as $(u, v) = 0$ in this case. Observe that $(\pi_u(v), u) = 0$ for all $v \in V$. Therefore $\operatorname{Im} \pi \subseteq P$. The map π_u regarded as a map from V to P is a linear projection from V onto P.

The projection π_u is singled out from the others by the fact that $\pi_u(v)$ is the vector in U closest to v, meaning that

$$||v - \pi_u(v)|| \leq ||v - u'||$$

for all $u' \in U$. The *reflection* σ_u *of* V *through* P is defined by

$$\sigma_u = 2\pi_u - \mathrm{Id}_V;$$

thus

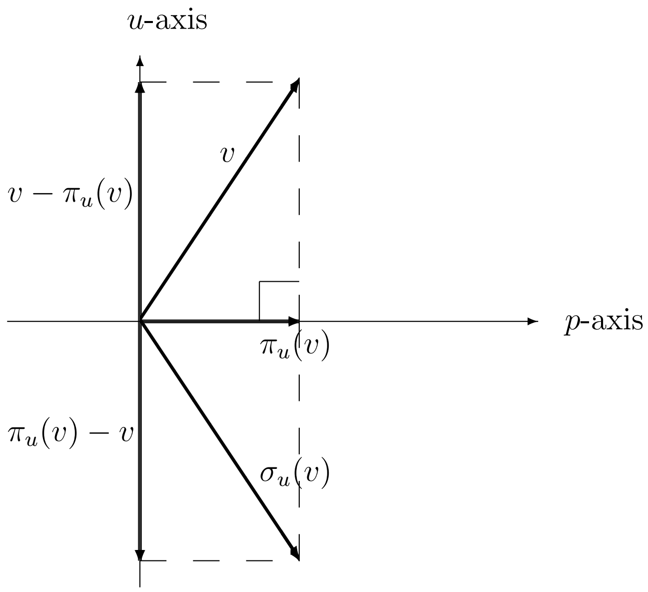
$$\sigma_u(v) = v - \frac{2(v,u)}{(u,u)}u$$

for all $v \in V$. Note that $\sigma_u(u) = -u$ and $\sigma_u(v) = v$ for all $v \in P$. In particular

$$\sigma_u^2 = \mathrm{Id}_V,$$

that is σ_u is an involution of V.

The following diagram indicates the relationship between π_u and σ_u.



$$v - \pi_u(v) = \frac{(v,u)}{(u,u)}u, \quad \pi_u(v) = v - \frac{(v,u)}{(u,u)}u, \quad \sigma_u(v) = \pi_u(v) + (\pi_u(v) - v) = 2\pi_u(v) - v$$

The u-axis is the span of u and the "positive" part of this axis is the set $\{au \,|\, a > 0\}$. The p-axis is the span of $\pi_u(v)$, which is usually one-dimensional. Observe that the p-axis lies in P and, of course, all of P

is perpendicular to u. Also observe that σ_u is an *isometry of* V; that is $||\sigma_u(v)|| = ||v||$ for all $v \in V$.

Comments concerning the lemma of the section. Suppose that F is any field and V is a vector space over F. Then $\mathrm{GL}(V)$ denotes the group of linear automorphisms of V under function composition. For a non-empty subset Φ of V let

$$G_\Phi = \{g \in \mathrm{GL}(V) \,|\, g(\Phi) = \Phi\}.$$

Then G_Φ is a subgroup of $\mathrm{GL}(V)$ and the restriction map $g \mapsto g|_\Phi$ determines a group homomorphism $\pi : G \longrightarrow \mathrm{Sym}(\Phi)$ from G to the group of permutations on Φ. If Φ spans V then π is injective since $\mathrm{Ker}\,\pi = (\mathrm{Id}_V)$ in this case. Thus if Φ spans V and V is finite-dimensional then G_Φ is a finite group.

A codimension one subspace of V is called a *hyperplane of* V. From this point on we will assume that F is of characteristic zero.

Suppose that P is a hyperplane of V, $\sigma \in \mathrm{GL}(V)$ fixes P pointwise, and $\sigma(\alpha) = -\alpha$ for some non-zero $\alpha \in V$. Then $\alpha \notin P$ since $\sigma(\alpha) \neq \alpha$. Thus $V = P \oplus F\alpha$. Since $\sigma(p) = p$ and $\sigma^2(\alpha) = \alpha$ it follows that $\sigma^2 = \mathrm{Id}_V$; that is σ is an involution of V. In particular $\sigma^{-1} = \sigma$. The lemma of 9.1 is a consequence of part (b) of the following.

Lemma 3.1.1 *Suppose that $\sigma, \sigma' \in \mathrm{GL}(V)$ each fix a hyperplane of V pointwise and that $\sigma(\alpha) = -\alpha = \sigma'(\alpha)$ for some non-zero $\alpha \in V$.*

(a) *$\sigma = \sigma'$ or $\tau = \sigma'\sigma$ has infinite order.*

(b) *Suppose that V is finite-dimensional and Φ is a finite spanning set for V. If $\sigma(\Phi) = \Phi = \sigma'(\Phi)$ then $\sigma = \sigma'$.*

PROOF: Part (b) follows from part (a) since σ, σ', hence $\tau = \sigma'\sigma$, belong to the finite group G_Φ in this case.

Part (a). Let P, P' be hyperplanes fixed pointwise by σ, σ' respectively. Then $P \oplus F\alpha = V = P' \oplus F\alpha$.

Consider $\tau = \sigma'\sigma$. Then $\tau(\alpha) = \alpha$. Let $p \in P$. Then $p = p' + a\alpha$ for some $p' \in P'$ and $a \in F$. Therefore

$$\tau(p) = \sigma'(\sigma(p)) = \sigma'(p) = \sigma'(p' + a\alpha) = p' - a\alpha = p - 2a\alpha$$

which shows that $(\tau - \mathrm{Id}_V)(p) = -2a\alpha$. Since $\tau(\alpha) = \alpha = \mathrm{Id}_V(\alpha)$ the operator $(\tau - \mathrm{Id}_V)^2$ vanishes on p and α. Therefore $(\tau - \mathrm{Id}_V)^2 = 0$, or equivalently $\tau^2 = 2\tau - \mathrm{Id}_V$. This equation implies $\tau^n = n\tau - (n-1)\mathrm{Id}_V$ for all $n \geq 1$.

Suppose that $\tau = \sigma'\sigma$ has finite order. Then $\tau^n = \mathrm{Id}_V$ for some $n \geq 1$. Therefore $\mathrm{Id}_V = \tau^n = n\tau - (n-1)\mathrm{Id}_V$ or $n\tau = n\mathrm{Id}_V$. Since the characteristic of F is zero $\sigma'\sigma = \tau = \mathrm{Id}_V$ from which $\sigma' = \sigma^{-1} = \sigma$ follows. $\square$

Exercise 3.1.2 Let V be a finite-dimensional vector space over F.

(a) Show that the hyperspaces of V are the $\mathrm{Ker}\, v^*$'s where $v^* \in V^*\backslash 0$.

(b) Now suppose that $F = \boldsymbol{R}$ and $\beta = (\ ,\ ) : V{\times}V \longrightarrow \boldsymbol{R}$ is a symmetric positive definite bilinear form. Show that the hyperspaces of V are the $u^\perp$'s, where $u \in V\backslash 0$. [Hint: Recall that $\beta_\ell : V \longrightarrow V^*$ given by $\beta_\ell(u,v) = \beta(u,v) = (u,v)$ for all $u, v \in V$ is a linear isomorphism.]

In the next exercise we discuss closest vectors and projections.

Definition 3.1.3 *Let V be a vector space over $\boldsymbol{R}$ and suppose that $(\ ,\ ) : V{\times}V \longrightarrow \boldsymbol{R}$ is a symmetric positive definite bilinear form. Let U be a subspace of V and $v \in V$. Then $u \in U$ is a* closest vector *in U to v if $||v - u|| \leq ||v - u'||$ for all $u' \in U$.*

Exercise 3.1.4 Let V be a vector space over $\boldsymbol{R}$ and let $(\ ,\ ) : V{\times}V \longrightarrow \boldsymbol{R}$ be a symmetric positive definite bilinear form. Suppose that U is a subspace of V, $v \in V$, and $u \in U$ is a closest vector in U to v.

(a) Show that $(v-u, u') = 0$ for all $u' \in U$. [Hint: Let $t \in \boldsymbol{R}$. Note that

$$||v-u||^2 \leq ||v-(u+tu')||^2 = ||v-u||^2 - 2t(v-u, tu') + t^2||u'||^2$$

which implies that $0 \leq -2t(v-u, u') + t^2||u'||^2$ for all $t \in \boldsymbol{R}$. Consider the cases $t > 0$ and $t < 0$ and take limits as $t \mapsto 0$.]

(b) Suppose that $u' \in U$ is also a closest vector in U to v. Show that $u = u'$. [Hint: Compute $||v-u'||^2 = ||(v-u)+(u-u')||^2$. For $x, y \in V$ recall that $||x+y||^2 = ||x||^2 + ||y||^2$ if and only if $(x,y) = 0$.]

By virtue of part (b) for $v' \in V$ there is at most one closest vector in U to v'.

(c) Suppose that $v' \in V$ and there is a closest vector $u' \in U$ to v'. For $a \in \boldsymbol{R}$ show that $au + u'$ is a closest vector in U to $av + v'$.

Exercise 3.1.5 Let V be a vector space over $\boldsymbol{R}$ and let $(\ ,\ ) : V{\times}V \longrightarrow \boldsymbol{R}$ be a symmetric positive definite bilinear form.

(a) Suppose that U is a subspace of V and every vector in V has a closest vector in U. Define $\pi : V \longrightarrow U$ by $\pi(v) = u$, where $u \in U$ is the vector in U closest to v. Show that π is linear and $\pi(u) = u$ for all $u \in U$. (Thus π is a projection from V onto U.)

(b) Let $u \in V \backslash 0$ and $U = u^{\perp}$. For $v \in V$ show that $v - \dfrac{(v,u)}{(u,u)}u$ is a vector in U closest to v.

Suppose that U is finite-dimensional.

(c) Show that every $v \in V$ has a closest vector in U, building an argument on the following idea: if $v \in V \backslash U$ then U is a hyperplane of the subspace $\mathcal{V} = U + Fv$. [Hint: See part b).]

(d) Show that every $v \in V$ has a closest vector in U, building an argument on the following idea: Let $v \in V \backslash U$ and let $\{u_1, \ldots, u_r\}$ be an orthonormal basis for U (use Gram–Schmidt). Then $v - (v, u_1)u_1 - \cdots - (v, u_r)u_r$ is perpendicular to all $u \in U$.

Definition 3.1.6 *We will provisionally call the projection of part (a) the* geometric projection of V onto U.

Exercise 3.1.7 Let V be a vector space over $\boldsymbol{R}$ and let $(\ ,\ ) : V \times V \longrightarrow \boldsymbol{R}$ be a symmetric positive definite bilinear form and let U be a subspace of V

(a) Suppose that there is a geometric projection $\pi : V \longrightarrow U$. Show that $V = U \oplus U^{\perp}$ and $\operatorname{Ker} \pi = U^{\perp}$.

Suppose that $V = U \oplus U^{\perp}$.

(b) Show that $\pi : V \longrightarrow U$ defined by $\pi(u \oplus u') = u$, where$u \in U$ and $u' \in U^{\perp}$, is a geometric projection of V onto U.

(c) Let $\sigma : V \longrightarrow V$ be defined by $\sigma = 2\pi - I$. Show that $\sigma(u \oplus u') = u - u'$ for all $u \in U$ and $u' \in U^{\perp}$.

(d) Show that $\sigma^2 = I$ and that σ is an isometry.

(e) Suppose that $v \in U \backslash 0$ and set $\mathcal{V} = U + Fv$. Show that $\sigma(\mathcal{V}) \subseteq \mathcal{V}$ and that $\sigma|_{\mathcal{V}}$ is the reflection through the hyperplane U of $\mathcal{V}$.

Definition 3.1.8 *The map σ of part (c) is called the* reflection of V through U.

3.1.2 Root systems

For the reader's convenience we recall the axioms of a root system $\mathbf{\Phi}$ for E.

(R1) $\mathbf{\Phi} \subset E\backslash 0$ and is a finite spanning set of E.

(R2) Let $a \in F$ and $\alpha \in \mathbf{\Phi}$. Then $a\alpha \in \mathbf{\Phi}$ if and only if $a = \pm 1$.

(R3) $\sigma_\alpha(\mathbf{\Phi}) \subseteq \mathbf{\Phi}$ for all $\alpha \in \mathbf{\Phi}$.

(R4) $<\beta, \alpha> = \dfrac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \boldsymbol{Z}$ for all $\alpha, \beta \in \mathbf{\Phi}$.

The structure of a root system is reflected in the structure of rank 2 root systems, which are analyzed in detail in 9.3, since:

Lemma 3.1.9 *Let $\mathbf{\Phi} \subseteq E$ be a root system for E, let $\alpha_1, \dots, \alpha_r \in \mathbf{\Phi}$, let E' be the span of these α_i's. Then $\mathbf{\Phi}' = \mathbf{\Phi} \cap E'$ is a root system for E'.*

PROOF: Let $\alpha \in E'\backslash 0$. Then the reflection through the hyperplane P_α in E' is $\sigma'_\alpha = \sigma_\alpha|_{E'}$. Thus for $\alpha \in \mathbf{\Phi}'$ it follows that $\sigma'_\alpha(\mathbf{\Phi}') \subseteq \mathbf{\Phi} \cap E' = \mathbf{\Phi}'$. $\square$

Let $\mathbf{\Phi}$ be a root system for E. We have noted in the preceding paragraph that

$$G = \{\sigma \in \mathrm{GL}(E) \mid \sigma(\mathbf{\Phi}) = \mathbf{\Phi}\}$$

is a finite subgroup of $\mathrm{GL}(E)$. The *Weyl group $\mathcal{W}$ of $\mathbf{\Phi}$* is the subgroup of G generated by the isometries σ_α where $\alpha \in \mathbf{\Phi}$. Thus $\mathcal{W}$ is a finite subgroup of G and is also a subgroup of the group of isometries of E.

Proposition 3.1.10 *Let E, E' be Euclidean spaces with root systems Φ, Φ' respectively.*

(a) *Suppose that $\tau : E \longrightarrow E'$ is a linear isomorphism such that $\tau(\Phi) = \Phi'$. Then $\tau\sigma_\alpha\tau^{-1} = \sigma_{\tau(\alpha)}$ for all $\alpha \in \Phi$ and $<\tau(\beta), \tau(\alpha)> = <\beta, \alpha>$ for all $\beta, \alpha \in \Phi$.*

(b) *Suppose that $\tau \in G_\Phi$. Then $\tau\sigma_\alpha\tau^{-1} = \sigma_{\tau(\alpha)}$ for all $\alpha \in \Phi$ and $<\tau(\beta), \tau(\alpha)> = <\beta, \alpha>$ for all $\beta, \alpha \in \Phi$.*

(c) *The Weyl group $\mathcal{W}$ of Φ is a normal subgroup of G_Φ.*

Proof: Part (c) follows from part (b) and part (b) follows from part (a) with $E = E'$ and $\Phi = \Phi'$.

Part (a). Let $\alpha \in \Phi$. Then σ_α fixes $P = \alpha^\perp$ pointwise and $\sigma_\alpha(\alpha) = -\alpha$. Therefore $\tau\sigma_\alpha\tau^{-1}$ fixes the hyperplane $\tau(P)$ of E' pointwise and $(\tau\sigma_\alpha\tau^{-1})(\tau(\alpha)) = -\tau(\alpha) = \sigma_{\tau(\alpha)}(\tau(\alpha))$. Since $\tau(\Phi) = \Phi'$ and $\sigma_\alpha(\Phi) = \Phi$ it follows that $\tau\sigma_\alpha\tau^{-1}(\Phi') = \Phi'$. As $\sigma_{\tau(\alpha)}(\Phi') = \Phi'$ as well, we conclude from part (b) of Lemma 3.1.1 that $\tau\sigma_\alpha\tau^{-1} = \sigma_{\tau(\alpha)}$.

Now let $\beta \in \Phi$. Then $\tau\sigma_\alpha\tau^{-1}(\tau(\beta)) = \tau(\sigma_\alpha(\beta)) = \tau(\beta) - <\beta, \alpha>\tau(\alpha)$. On the other hand $\sigma_{\tau(\alpha)}(\tau(\beta)) = \tau(\beta) - <\tau(\beta), \tau(\alpha)>\tau(\alpha)$. Thus $<\beta, \alpha> = <\tau(\beta), \tau(\alpha)>$. $\square$

Let E, E' be Euclidean spaces with root systems Φ, Φ' respectively. An *isomorphism of root systems* Φ, Φ' is an linear isomorphism $\tau : E \longrightarrow E'$ such that $\tau(\Phi) = \Phi'$ and $<\tau(\beta), \tau(\alpha)> = <\beta, \alpha>$ for all $\beta, \alpha \in \Phi$. By part (a) of the preceding proposition an isomorphism of root systems Φ, Φ' is a linear isomorphism $\tau : E \longrightarrow E'$ such that $\tau(\Phi) = \Phi'$.

Suppose that $a_\alpha \in \boldsymbol{R}$ for all $\alpha \in \Phi$ and let $\Phi'' = \{a_\alpha\alpha \,|\, \alpha \in \Phi\}$. It is easy to see that Φ'' is a root system if and only if

(r1) $a_\alpha \neq 0$;

(r2) $a_\alpha = a_{-\alpha}$;

(r3) $a_\beta = a_{\sigma_\alpha(\beta)}$;

(r4) $\left(\dfrac{a_\beta}{a_\alpha}\right) <\beta, \alpha> \in \boldsymbol{Z}$

for all $\alpha, \beta \in \Phi$. Let $c \in \boldsymbol{R}$ be non-zero. Then Φ'' is a root system with $a_\alpha = c$ for all $\alpha \in \Phi$. Φ'' is also a root system with $a_\alpha = \dfrac{1}{||\alpha||^2} = \dfrac{1}{(\alpha, \alpha)}$ for all $\alpha \in \Phi$. For since σ_α is an isometry (r3) holds. The calculation

$$\left(\frac{a_\beta}{a_\alpha}\right) <\beta, \alpha> = \left(\frac{(\alpha, \alpha)}{(\beta, \beta)}\right)\left(\frac{2(\beta, \alpha)}{(\alpha, \alpha)}\right) = \frac{2(\alpha, \beta)}{(\beta, \beta)} = <\alpha, \beta>$$

establishes (r4). Combining these constructions results in a root system

$$\Phi^\vee = \{\alpha^\vee \,|\, \alpha \in \Phi\},$$

called the *dual root system*, where $\alpha^\vee = \dfrac{\alpha}{(\alpha, \alpha)}$ for $\alpha \in \Phi$. Note

$$<\beta^\vee, \alpha^\vee> = <\alpha, \beta>$$

for all $\alpha, \beta \in \Phi$.

Since $\sigma_{au} = \sigma_u$ for all $a \in \boldsymbol{R}\backslash 0$ and $u \in E\backslash 0$ it follows that

$$\sigma_{\alpha^\vee} = \sigma_\alpha$$

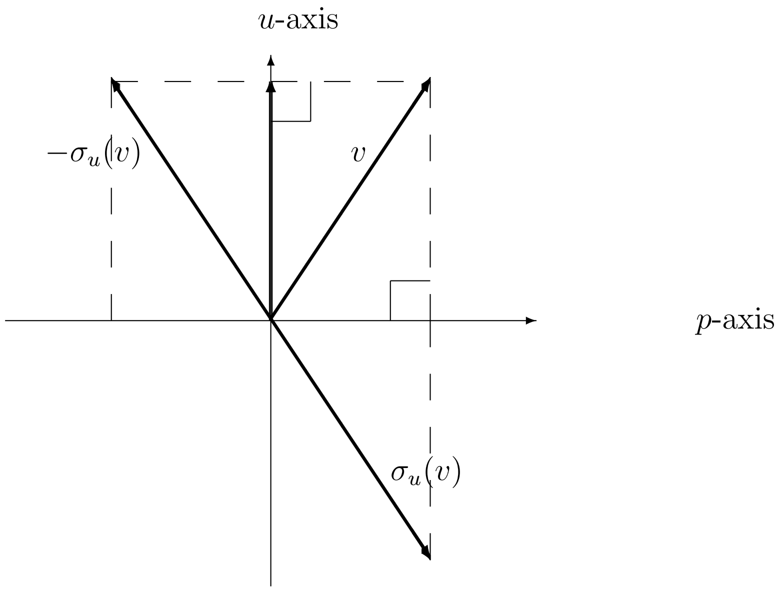
for all $\alpha \in \boldsymbol{\Phi}$. Therefore the Weyl group of $\boldsymbol{\Phi}$ *is* the Weyl group of $\boldsymbol{\Phi}^\vee$.

3.1.3 Examples

Let $\boldsymbol{\Phi} \subseteq E$ satisfy (R1)–(R3). For $\alpha \in \boldsymbol{\Phi}$ observe that

$$\tau_\alpha = -\sigma_\alpha$$

is the reflection of E through the line $\boldsymbol{R}\alpha$. See the figure below, which is derived from the figure of the discussion of 9.2 above, and see Exercise 3.1.7.



Note that (R1)–(R3) are satisfied with the τ_α's replacing the σ_α's. We will use the τ_α's to study root systems in $\boldsymbol{R}^2$. First some general observations.

Among all of the pairs $\alpha', \beta' \in \boldsymbol{\Phi}$, where $\beta' \neq \pm\alpha'$, choose one α, β with smallest angle θ between them. Now $0 \leq \theta \leq \pi$ by definition. (If

$-\alpha \in \mathbf{\Phi}$ then $0 \leq \theta \leq \pi/2$.) We define a sequence in E inductively by setting $\alpha_0 = \alpha, \alpha_1 = \beta$, and for $i \geq 2$ we let α_{i+1} be the reflection of α_{i-1} through α_i. Thus

$$\alpha_{i+1} = \tau_{\alpha_i}(\alpha_{i-1})$$

for all $i \geq 2$. By induction

$$\alpha_0, \alpha_1, \alpha_2, \dots \in \mathbf{\Phi}.$$

Now the angle between α_{i-1} and α_i is the same as the angle between α_i and α_{i+1}. Therefore by induction

$$\text{the angle between } \alpha_i \text{ and } \alpha_{i+1} \text{ is } \theta$$

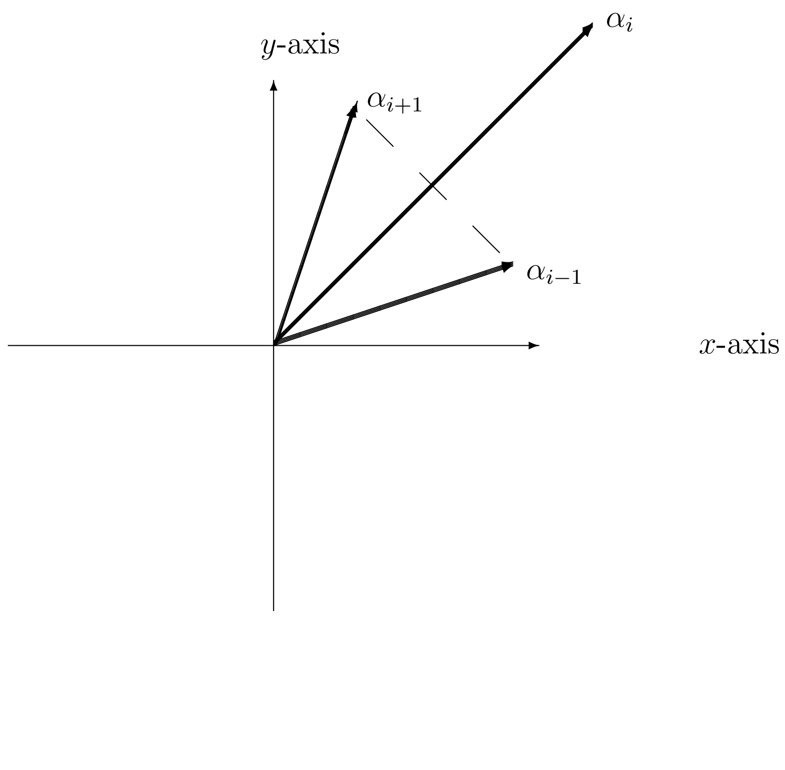
for all $i \geq 0$. Since reflections are isometries if follows that

$$\alpha_0, \alpha_2, \alpha_4, \dots \text{ have the same length}$$

and

$$\alpha_1, \alpha_3, \alpha_5, \dots \text{ have the same length.}$$

The figure below illustrates the construction of $\alpha_0, \alpha_1, \alpha_2, \dots$ when $E = \boldsymbol{R}^2$.



Now suppose that $E = \boldsymbol{R}^2$. Since $0 \leq \theta \leq \pi$ the root α lies between α_n and α_{n+1} for some $n > 0$. Let n be the least such integer. Then $n\theta \leq 2\pi < (n+1)\theta$. Since the angle between α_n and 2π is $0 \leq 2\pi - n\theta < \theta$ it follows that $\pi - n\theta = 0$. Therefore

$$\theta = \frac{2\pi}{n}.$$

If the rank of $\boldsymbol{\Phi}$ is 2 then $n > 2$. Evidently

$$\boldsymbol{\Phi} = \{\alpha_0, \alpha_1, \ldots, \alpha_{n-1}\}.$$

Suppose that $-\alpha \in \boldsymbol{\Phi}$. We have noted that $0 \leq \theta \leq \pi/2$ in this case. Thus $-\alpha$ lies between α_m and α_{m+1} for some $m > 0$. Letting m be the smallest such positive integer we argue as above that

$$\theta = \frac{\pi}{m}.$$

In particular $n = 2m$.

The condition (R4) imposes restrictions on m, namely $m = 2, 3, 4$ or 6. Our analysis at this point accounts for the diagram possibilities listed in Section 9.2. The actual construction requires a bit more work. See Exercises 3.1.11–3.1.12 below. The final touches on the construction of the rank two root systems are encouraged in Exercise 3.1.15.

Exercise 3.1.11 Let E be a finite-dimensional euclidean space and suppose that $\boldsymbol{\Phi} \subset E \backslash 0$ is a finite set.

(a) Show that $\boldsymbol{\Phi}$ satisfies (R1)–(R3) if and only if $\boldsymbol{\Phi}$ satisfies (R1)–(R3) with the reflections σ_α through the hyperplanes $\alpha^\perp$ replaced by the reflections τ_α through the lines $F\alpha$.

(b) Let $\boldsymbol{\Phi}_{\mathsf{n}} = \{\frac{v}{||v||} \mid v \in \boldsymbol{\Phi}\}$. Show that $\boldsymbol{\Phi}$ satisfies (R1) and (R3) if and only if $\boldsymbol{\Phi}_{\mathsf{n}}$ does. (Note that $\boldsymbol{\Phi}_{\mathsf{n}}$ is merely the set of normalizations of $\boldsymbol{\Phi}$.)

Exercise 3.1.12 Suppose that $E = \boldsymbol{R}^2$ with the usual (positive definite) inner product and let $\boldsymbol{\Phi}$ be a rank two system of roots for E. We can assume that $(0, a) \in \boldsymbol{\Phi}$ for some $a \in \boldsymbol{R} \backslash 0$.

(a) Show that $\boldsymbol{\Phi}_{\mathsf{n}} = \{u_1, \ldots, u_{2m-1}\}$ for some $m > 1$, where

$$u_i = \begin{pmatrix} \cos\left(\frac{\pi}{m} i\right) \\ \\ \sin\left(\frac{\pi}{m} i\right) \end{pmatrix}$$

for all $0 \leq i < 2m$.

For $i \in \boldsymbol{Z}$ we let u_i be defined by as above, we let $\tau_i = \tau_{u_i}$ be the reflection of E through the line Fu_i, and we let $\sigma_i = -\tau_i$ be the reflection of E through the hyperplane, or line in this case, $u_i^\perp$.

(b) Show that $\tau_i(u_j) = u_{2i-j}$ for all $i, j \in \boldsymbol{Z}$. [Hint: Since the u_i's have length 1 observe that $\tau_i(v) = 2(v, u_i)u_i - v$ for all $v \in E$. The calculation of $\tau_i(u_j)$ only involves some basic trigonometric formulas.]

(c) Show that $\sigma_i(u_j) = u_{m+2i-j}$ for all $i, j \in \boldsymbol{Z}$.

d) Show that $\boldsymbol{\Phi} = \{\alpha_0, \dots, \alpha_{2m-1}\}$, where

(1) if m is odd then $\alpha_0 = au_0, \alpha_1 = au_1, \dots, \alpha_{2m-1} = au_{2m-1}$ for some $a > 0$, and

(2) if m is even then $\alpha_0 = au_0, \alpha_2 = au_2, \alpha_4 = au_4$ and $\alpha_1 = bu_1, \alpha_3 = bu_3, \alpha_5 = bu_5, \dots$ for some $a, b > 0$.

[Hint: First note that $\alpha_0 = au_0$ and $\alpha_1 = bu_1$ for some $a, b \in \boldsymbol{R}\backslash 0$. Using part (c) we see that $\boldsymbol{\Phi} = \{\pm au_{m+2i}, \pm bu_{m-1+2i} \,|\, i \in \boldsymbol{Z}\}$.]

e) Suppose (1) or (2) is satisfied. Show that $\boldsymbol{\Phi} = \{\alpha_0, \dots, \alpha_{2m-1}\}$ satisfies (R1)–(R3).

Exercise 3.1.13 Let $\mathcal{W}_m$ be the subgroup of isometries of $E = \boldsymbol{R}^2$ generated by the reflections $\sigma_0, \dots, \sigma_{2m-1}$ of Exercise 3.1.12.

(a) Show that $\mathcal{W}_m$ is isomorphic to the subgroup W_m of $\mathrm{Sym}(\boldsymbol{Z}_{2m})$ generated by $\boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_{2m-1}$, where

$$\boldsymbol{\sigma}_i(j) = m + 2i - j$$

for all $0 \le i < 2m$ and $j \in \boldsymbol{Z}_{2m}$.

(b) Show that $\mathcal{W}_n \simeq D_{2m}$. [Hint: Note that $\tau = \boldsymbol{\sigma}_1\boldsymbol{\sigma}_0$ has order m.]

3.1.4 Pairs of roots

Let $\beta, \alpha \in \Phi$. Then

$$<\beta,\alpha><\alpha,\beta> = \left(\frac{2(\beta,\alpha)}{(\alpha,\alpha)}\right)\left(\frac{2(\alpha,\beta)}{(\beta,\beta)}\right) = 4\left(\frac{(\alpha,\beta)}{\|\alpha\|\,\|\beta\|}\right)^2 = 4\cos^2\theta, \quad (3.1)$$

where θ is the angle between α and β. By convention $0 \leq \theta \leq \pi$. Note that $\cos^2\theta = 1$ if and only if β and α are scalar multiples of each other; that is $\beta = \pm\alpha$. In this case $<\beta, \alpha> = \pm 2$.

Suppose $\beta \neq \pm\alpha$. Then $<\beta, \alpha><\alpha, \beta> \in \{0, 1, 2, 3\}$. Now $<\beta, \alpha><\alpha, \beta> = 0$ if and only if $(\alpha, \beta) = 0$ if and only if α and β are perpendicular.

Suppose further that α and β are not perpendicular. Then $<\beta, \alpha><\alpha, \beta> \in \{1, 2, 3\}$. Since $<\beta, \alpha>$ and $<\alpha, \beta>$ are integers, one of $<\beta, \alpha>$ or $<\alpha, \beta>$ is ± 1. *Assume further that* $||\beta|| \geq ||\alpha||$. From

$$\frac{||\beta||^2}{||\alpha||^2} = \frac{<\beta, \alpha>}{<\alpha, \beta>}$$

we conclude that $<\alpha, \beta> = \pm 1$; therefore $\dfrac{||\beta||^2}{||\alpha||^2} = |<\beta, \alpha>|$. Since $<\beta, \alpha>, <\alpha, \beta>$ and (α, β) all have the same sign,

$$\cos\theta = <\alpha, \beta> \left(\frac{\sqrt{|<\beta, \alpha>|}}{2} \right).$$

The values for the quantities described in the table below now fall into place very quickly; again our assumptions are $\beta \neq \pm\alpha$ and $||\beta|| \geq ||\alpha||$.

$<\alpha, \beta>$	$<\beta, \alpha>$	$\cos\theta$	θ	$\|\|\beta\|\|^2/\|\|\alpha\|\|^2$
0	0	0	$\pi/2$	undetermined
1	1	$1/2$	$\pi/3$	1
-1	-1	$-1/2$	$2\pi/3$	1
1	2	$\sqrt{2}/2$	$\pi/4$	2
-1	-2	$-\sqrt{2}/2$	$3\pi/4$	2
1	3	$\sqrt{3}/2$	$\pi/6$	3
-1	-3	$-\sqrt{3}/2$	$5\pi/6$	3

Lemma 3.1.14 *Suppose that* $\beta, \alpha \in \Phi$ *and* $\beta \neq \pm\alpha$.

(a) *If* $(\beta, \alpha) < 0$ *then* $\beta + \alpha \in \Phi$.

(b) $(\beta, \alpha) > 0$ *then* $\beta - \alpha \in \Phi$.

Proof: Part (b) follows from part (a) with the pair β and $-\alpha$. To show part (a). Since $(\beta, \alpha) = (\alpha, \beta)$, we may assume without loss of generality that $||\alpha|| \geq ||\beta||$. From the table $<\beta,\alpha> = -1$ which means $\beta + \alpha = \beta - <\beta,\alpha>\alpha = \sigma_\alpha(\beta) \in \Phi$. $\Box$

Consider the "sequence"

$$\ldots, \beta - 2\alpha, \beta - \alpha, \beta, \beta + \alpha, \beta + \alpha, \ldots \ , \tag{3.2}$$

where $\beta \neq \pm\alpha$. Since β is not a scalar multiple of α it follows that $\beta + i\alpha \neq \pm\alpha$ for all $i \in \boldsymbol{Z}$. For $i, i' \in \boldsymbol{Z}$ note that $\beta + i\alpha = \beta + i'\alpha$ if and only if $i = i'$ since $\alpha \neq 0$. We order the elements of the sequence by $\beta + i\alpha \preceq \beta + i'\alpha$ if and only if $i \leq i'$. Since $\sigma_\alpha(\beta + i\alpha) = \beta - (<\beta,\alpha> + i)\alpha$ the involution σ_α induces an *order reversing involution of the sequence.*

Suppose that $\beta + i\alpha \preceq \beta + i'\alpha$. Then

$$[\beta + i\alpha,\ \beta + i'\alpha] = \{\beta + j\alpha \,|\, \beta + i\alpha \preceq \beta + j\alpha \preceq \beta + i'\alpha\} = \{\beta + j\alpha \,|\, i \leq j \leq i'\}$$

is the *segment with endpoints* $\beta + i\alpha$ *and* $\beta + i'\alpha$. Since σ_α induces an order reversing involution of the sequence

$$\sigma_\alpha([\beta + i\alpha,\ \beta + i'\alpha]) = [\beta - (<\beta,\alpha> + i')\alpha,\ \beta - (<\beta,\alpha> + i)\alpha]. \tag{3.3}$$

Suppose that $\beta + \ell\alpha \notin \Phi$. Assume $\beta + k\alpha, \beta + m\alpha \in \Phi$, where $k \leq \ell$ and $\ell \leq m$. Let k be the largest such integer and m be the smallest. Then $k < \ell$ and $(\beta + k\alpha) + \alpha \notin \Phi$ and $\ell < m$ and $(\beta + m\alpha) - \alpha \notin \Phi$. Therefore

$$(\beta + k\alpha, \alpha) \not< 0 \qquad \text{and} \qquad (\beta + m\alpha, \alpha) \not> 0,$$

by Lemma 3.1.14, or equivalently

$$(\beta, \alpha) + k(\alpha, \alpha) \geq 0 \qquad \text{and} \qquad (\beta, \alpha) + m(\alpha, \alpha) \leq 0$$

from which $(k - m)(\alpha, \alpha) \geq 0$, a contradiction. we have shown

$$\beta + i\alpha \preceq \beta + i'\alpha \in \Phi \qquad \text{implies} \qquad [\beta + i\alpha,\ \beta + i'\alpha] \subseteq \Phi. \tag{3.4}$$

Now $\beta + (-0)\alpha = \beta + 0\alpha = \beta \in \Phi$. Since Φ is finite there are largest non-negative integers r, q such that $\beta + (-r)\alpha, \beta + q\alpha \in \Phi$. In light of (3.4) it follows that the set of all roots in the sequence (3.2) is the segment $[\beta + (-r)\alpha,\ \beta + q\alpha]$, called the α-*string through* β. Since σ_α is injective and $\sigma_\alpha(\Phi) \subseteq \Phi$,

$$[\beta + (-r)\alpha,\ \beta + q\alpha] = [\beta - (<\beta,\alpha> + q)\alpha,\ \beta - (<\beta,\alpha> + (-r))\alpha] \tag{3.5}$$

by (3.3). Therefore

$$<\beta, \alpha> = r - q. \tag{3.6}$$

Using the table we see that $<\beta', \alpha'> \leq 3$ for all $\beta', \alpha' \in \Phi$. Since $\beta' = -(\beta + (-r)\alpha)$ and $\beta'' = \beta + q\alpha$ are roots, the calculation

$$2(r+q) = <(r+q)\alpha, \alpha> = <\beta' + \beta'', \alpha> = <\beta', \alpha> + <\beta'', \alpha> \leq 6$$

shows that $r + q \leq 3$. Therefore roots strings have length at most 4.

Exercise 3.1.15 Use the table on page 45 of the text together with Exercises 3.1.11–3.1.12 to construct the rank 2 root systems.

3.2 Bases and Weyl Chambers

3.2.1 Bases and Weyl chambers

The existence of a regular element is a consequence of the following lemma. We first note that a vector space V of dimension at least 2 over an infinite field has an infinite number of one-dimensional subspaces. Indeed let $\{v_1, v_2\}$ be a linearly independent subset of V. Then the $v_1 + av_2$'s, where $a \in F$, span different one-dimensional subspaces of V.

Lemma 3.2.1 *Let V be a finite-dimensional vector space over and infinite field F and suppose that $V = V_1 \bigcup \cdots \bigcup V_r$ is the union of a finite number of subspaces. Then $V_i = V$ for some $1 \leq i \leq r$.*

PROOF: We may as well assume that $\operatorname{Dim} V \geq 2$. Note there are an infinite number of codimension one subspaces of V. For the codimension one subspaces of V are the $\operatorname{Ker} f$'s, where $f \in V^* \backslash 0$, and these kernels are in one-one correspondence with the one-dimensional subspaces of V^*. By our assumption V^* is at least two-dimensional.

Let $\mathcal{V}$ be a codimension one subspace of V. Since $\mathcal{V} = (\mathcal{V} \bigcap V_1) \bigcup \cdots \bigcup (\mathcal{V} \bigcap V_r)$ it follows by induction on $\operatorname{Dim} V$ that $\mathcal{V} = \mathcal{V} \bigcap V_i$, or equivalently $\mathcal{V} \subseteq V_i$, for some $1 \leq i \leq r$. Thus some V_i contains two different codimension one subspaces $\mathcal{V}, \mathcal{V}'$ from which $V = \mathcal{V} + \mathcal{V}' \subseteq V_i$ follows. $\square$

Corollary 3.2.2 *Let V be a finite-dimensional vector space over F and let $U, U_1, \ldots, U_r$ be subspaces of V. Suppose that $U \subseteq \bigcup_{i=1}^r U_i$. Then $U \subseteq U_i$ for some $1 \leq i \leq r$.*

PROOF: By assumption $U = \bigcup_{i=1}^{r}(U \cap U_i)$ and therefore $U = U \cap U_i$, or equivalently, $U \subseteq U_i$, for some $1 \leq i \leq r$ by Lemma 3.2.1. $\square$

Some comments on the proof of 10.1 Theorem'.

(1) Write $\mathbf{\Phi}^+(\gamma) = \{\beta_1, \ldots, \beta_r\}$ where $(\gamma, \beta_1) \leq \cdots \leq (\gamma, \beta_r)$ and use induction on i.

(5) Let $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ be *any* linear basis for E and suppose that $a_1, \ldots, a_n \in F$. Then there is a $\gamma \in E$ such that $(\gamma, \alpha_i) = a_i$ for all $1 \leq i \leq n$.

To see this observe that there is a linear functional $f \in E^*$ such that $f(\alpha_i) = a_i$ for all $1 \leq i \leq n$. Now $\beta = (\ ,\ )$ is a non-degenerate bilinear form. Therefore $\beta_\ell : E \longrightarrow E^*$ is a linear isomorphism. This means $f = \beta_\ell(\gamma)$ for some $\gamma \in E$. Thus

$$a_i = f(\alpha_i) = \beta_\ell(\gamma)(\alpha_i) = \beta(\gamma, \alpha_i) = (\gamma, \alpha_i)$$

for all $1 \leq i \leq n$.

There is quite a bit to say about the Weyl chambers. Let $\sigma \in \mathrm{GL}(E)$ be an isometry. Then $(\sigma(u), \sigma(v)) = (u, v)$ for all $u, v \in E$. Consequently

$$\sigma(u^\perp) = \sigma(u)^\perp$$

for all $u \in E$. This means that $\sigma \in \mathcal{W}$ permutes the set of codimension one subspaces $\{P_\alpha \,|\, \alpha \in \mathbf{\Phi}\}$. Therefore

$$\sigma(\bigcup\nolimits_{\alpha \in \mathbf{\Phi}} P_\alpha) = \bigcup\nolimits_{\alpha \in \mathbf{\Phi}} P_\alpha$$

which means that

$$\sigma(X) = X$$

where

$$X = E \backslash (\bigcup\nolimits_{\alpha \in \mathbf{\Phi}} P_\alpha).$$

By definition the vectors in X are *regular*. The vectors in $\bigcup_{\alpha \in \mathbf{\Phi}} P_\alpha$ are *singular*.

The distance function

$$d(u, v) = ||u - v|| = \sqrt{(u - v, u - v)}$$

for all $u, v \in E$ gives E the structure of a metric space. We will assume some of the elementary facts about the resulting topology without proof.

All linear endomorphisms of E are continuous. In particular all linear automorphisms of E are homeomorphisms of E. Consequently all $\sigma \in \mathcal{W}$ permute the connected components of X, the *Weyl chambers of E*. For $\gamma \in X$ let $\mathcal{C}(\gamma)$ be the Weyl chamber (connected component of X) containing γ. Then

$$\sigma(\mathcal{C}(\gamma)) = \mathcal{C}(\sigma(\gamma)) \tag{3.7}$$

for all $\sigma \in \mathcal{W}$.

We will next determine the connected components of X. Let $\alpha \in \mathbf{\Phi}$. Since $\beta_\ell(\alpha) : E \longrightarrow \boldsymbol{R}$ is continuous,

$$P^+_\alpha = \{\gamma \in E \,|\, (\alpha, \gamma) > 0\} = \beta_\ell(\alpha)^{-1}((0, \infty))$$

and

$$P^-_\alpha = \{\gamma \in E \,|\, (\alpha, \gamma) < 0\} = \beta_\ell(\alpha)^{-1}((-\infty, 0))$$

are open subsets of E hence of X. Let $\mathbf{\Phi} = \{\alpha_1, \ldots, \alpha_r\}$ be a listing of the elements of $\mathbf{\Phi}$. Since $E = P^+_\alpha \cup P_\alpha \cup P^-_\alpha$ is a partition of E it follows that

$$\begin{aligned} X &= E \backslash (\textstyle\bigcup_{\alpha \in \mathbf{\Phi}} P_\alpha) \\ &= \textstyle\bigcap_{\alpha \in \mathbf{\Phi}} E \backslash P_\alpha \\ &= \textstyle\bigcap_{\alpha \in \mathbf{\Phi}} (P^+_\alpha \bigcup P^-_\alpha) \\ &= \bigcup_{\substack{(n_1, \ldots, n_r) \\ n_i \in \{+,-\}}} (P^{n_1}_{\alpha_1} \cap \cdots \cap P^{n_r}_{\alpha_r}) \end{aligned}$$

is the union of disjoint open subsets of E. Once we show that the $P^{n_1}_{\alpha_1} \cap \cdots \cap P^{n_r}_{\alpha_r}$'s are connected it will follows that the non-empty ones are the connected components of X.

Let $\gamma, \gamma' \in X$. Then $\varphi : \boldsymbol{R} \longrightarrow E$ defined by

$$\varphi(t) = t\gamma + (1 - t)\gamma'$$

for all $t \in \boldsymbol{R}$ is continuous. Thus

$$\varphi([0, 1]) = \{t\gamma + (1 - t)\gamma' \,|\, 0 \leq t \leq 1\}$$

is connected since it is the continuous image of a connected set. Let $\alpha \in \mathbf{\Phi}$. We may assume without loss of generality that $(\gamma, \alpha) \leq (\gamma', \alpha)$. Thus

$$\{(\alpha, t\gamma + (1 - t)\gamma') \,|\, 0 \leq t \leq 1\} = [(\alpha, \gamma), (\alpha, \gamma')].$$

At this point the reader can see that $\gamma, \gamma' \in P_{\alpha_1}^{n_1} \cap \cdots \cap P_{\alpha_r}^{n_r}$ implies that the line segment $\varphi([0,1]) \subseteq P_{\alpha_1}^{n_1} \cap \cdots \cap P_{\alpha_r}^{n_r}$ also. Therefore $P_{\alpha_1}^{n_1} \cap \cdots \cap P_{\alpha_r}^{n_r}$ is connected.

If the Weyl chambers were defined as the $P_{\alpha_1}^{n_1} \cap \cdots \cap P_{\alpha_r}^{n_r}$ in the first place then one could deduce (3.7) from the fact that $\sigma \in \mathcal{W}$ is an isometry and one could deduce that the chambers are convex subsets of X.

As a matter of convenience here we show that any $\alpha \in \mathbf{\Phi}$ belongs to a base.

Proposition 3.2.3 *Let $\mathbf{\Phi}$ be a root system for E. Then any $\alpha \in \mathbf{\Phi}$ belongs to a base.*

Proof: For $\beta \in \Phi$ observe that $P_\alpha = P_\beta$ if and only if $\beta = \pm\alpha$. Thus

$$P_\alpha \not\subseteq \bigcup_{\beta \neq \pm\alpha} P_\beta$$

by Corollary 3.2.2. Choose $u \in P_\alpha$ such that $u \notin P_\beta$ for all $\beta \neq \pm\alpha$.

For $\beta \neq \pm\alpha$ the function $f_\beta : \boldsymbol{R} \longrightarrow \boldsymbol{R}$ defined by

$$f_\beta(t) = |(\beta, u + \frac{t\alpha}{(\alpha,\alpha)})| - t$$

is continuous. Since $f_\beta(0) = |(\beta, u)| > 0$ it follows that $0 \in \mathcal{O}_\beta = f_\beta^{-1}((0,\infty))$ and thus $0 \in \bigcap_{\beta \neq \pm\alpha} \mathcal{O}_\beta$. This intersection is an open subset of $\boldsymbol{R}$ and thus contains a $t > 0$.

Let $\gamma = u + \dfrac{t\alpha}{(\alpha,\alpha)}$. Then $(\alpha, \gamma) = t$ and $|(\beta, \gamma)| > t$ for all $\beta \neq \pm\alpha$. Therefore $\alpha \in \mathbf{\Phi}^+(\gamma)$ and is indecomposable. $\square$

Let $\omega \in \mathcal{W}$. Since σ is an isometry $(u, v) = (\sigma(u), \sigma(v))$ for all $u, v \in E$. Thus if Δ is a base $\sigma(\Delta)$ is a base and, using (3.7), it follows that

$$\sigma(\Delta(\gamma)) = \Delta(\sigma(\gamma)) \tag{3.8}$$

for all $\gamma \in X$.

Now $\Delta \subseteq \mathcal{C}(\gamma)$ for some $\gamma \in X$. By definition $(\gamma, \alpha) > 0$ for all $\alpha \in \Delta$, and $\mathcal{C}(\Delta) = \mathcal{C}(\gamma)$. Suppose that $\gamma' \in X$ and $(\gamma', \alpha) > 0$ for all $\alpha \in \Delta$. Then $(\gamma', \beta), (\gamma, \beta)$ have the same sign for all $\beta \in \mathbf{\Phi}$ since Δ is a base. Therefore

$$\mathcal{C}(\Delta) = \{\gamma' \in X \mid (\gamma', \alpha) > 0 \text{ for all } \alpha \in \Delta\}. \tag{3.9}$$

Exercise 3.2.4 Show that the conclusion of Lemma 3.2.1 holds when

a) F is an infinite field and $V_1, \ldots, V_r$ is replaced with a family of subspaces $\{V_i\}_{i\in\mathcal{I}}$, where $|\mathcal{I}| < |F|$, or

b) F is a finite field and $r \leq |F|$.

3.2.2 Lemmas on simple roots

No particular comments.

3.2.3 The Weyl group

Comments on the proof of the theorem of the section:

(a) See Equation 3.9.

(c) See Proposition 3.2.2.

We define $\ell(\sigma) = 0$ if $\sigma = 1$ is the identity of the Weyl group. Suppose that $\sigma \in \mathcal{W}\backslash 1$. Then $\sigma = \sigma_{\alpha_1}\cdots\sigma_{\alpha_t}$, where $\alpha_1, \ldots, \alpha_t \in \Delta$. Let $\ell(\sigma)$ be the smallest possible t. In any case $\ell(\sigma)$ is the *length of* σ which is a non-negative integer.

We can associate another non-negative integer to $\sigma \in \mathcal{W}$ as well. Let $n(\sigma)$ be the number of negative roots in the set $\{\sigma(\beta) \,|\, \beta \in \Delta\}$. Observe that $n(1) = 0$ as well.

Suppose that $\ell(\sigma) = t > 0$ and write $\sigma = \sigma_{\alpha_1}\cdots\sigma_{\alpha_t}$ as above. Then

$$\ell(\sigma) = \ell(\sigma_{\alpha_1}\cdots\sigma_{\alpha_{t-1}}) = \ell(\sigma\sigma_{\alpha_t}) + 1.$$

We will show that $n(\sigma) = n(\sigma\sigma_{\alpha_t}) + 1$ as well. Since σ_{α_t} permutes $\mathbf{\Phi}^+\backslash\alpha_t$ by Lemma B of §10.2, it follows that

$$\{\sigma(\beta) \,|\, \beta \in \mathbf{\Phi}^+\backslash\alpha_t\} = \{\sigma\sigma_{\alpha_t}(\sigma_{\alpha_t}(\beta)) \,|\, \beta \in \mathbf{\Phi}^+\backslash\alpha_t\} = \{\sigma\sigma_{\alpha_t}(\beta) \,|\, \beta \in \mathbf{\Phi}^+\backslash\alpha_t\}.$$

Now

$$\sigma(\alpha_t) = \sigma_{\alpha_1}\cdots\sigma_{\alpha_{t-1}}(\alpha_t) \in \mathbf{\Phi}^-$$

by the corollary in §10.2, and thus

$$\sigma\sigma_{\alpha_t}(\alpha_t) = \sigma(-\alpha_t) = -\sigma(\alpha_t) \in \mathbf{\Phi}^+.$$

Therefore $n(\sigma) = n(\sigma\sigma_{\alpha_t}) + 1$. We have proved Lemma A by induction on $\ell(\sigma)$:

Lemma 3.2.5 $\ell(\sigma) = n(\sigma)$ *for all* $\sigma \in \mathcal{W}$. $\Box$

Concerning Lemma B, we write $\mathcal{C}(\Delta) = P_{\alpha_1}^{n_1} \cap \cdots P_{\alpha_r}^{n_r}$ as in the notes on §10.1, and note that

$$\overline{\mathcal{C}(\Delta)} \subseteq \mathcal{C}(\Delta) \cup P_{\alpha_1} \cdots \cup P_{\alpha_r}$$

as $P_{\alpha_i}^{n_i} \cup P_{\alpha_i}$ is closed for all $1 \leq i \leq r$.

3.2.4 Irreducible root systems

Let E be a Euclidean space. A subset $S \subseteq E$ is *reducible* if S is the union of non-empty subsets S' and S'' such that $(s', s'') = 0$ for all $s' \in S'$ and $s'' \in S''$. Note that $S' \cap S'' = \emptyset$ when $0 \notin S$. The set S is *irreducible* if it is not reducible.

By definition the empty set is irreducible. Observe that any singleton subset of E is irreducible. If $0 \in S$ then S is irreducible if and only if $S = \{0\}$. In particular $\{0\}$ is a maximal irreducible subset of E.

By Zorn's Lemma every irreducible subset of S is contained in a maximal irreducible subset of S. The latter are called *irreducible components of* S. At this point it is not hard to establish:

Proposition 3.2.6 *Let* E *be a Euclidean space and suppose that* S *is a subset of* E. *Then:*

(a) S *is the union of its irreducible components.*

(b) *Any irreducible subset of* S *is contained in an component of* S.

(c) *Suppose* S' *and* S'' *are different irreducible components of* S. *Then* $S' \cap S'' = \emptyset$ *and* $(s', s'') = 0$ *for all* $s' \in S'$ *and* $s'' \in S''$.

$\Box$

Now suppose that E is finite-dimensional and that $\mathbf{\Phi}$ is a root system for E. Write $\mathbf{\Phi} = \mathbf{\Phi}_1 \cup \cdots \cup \mathbf{\Phi}_r$ as the disjoint union of its irreducible components and let E_i be the span of $\mathbf{\Phi}_i$ for all $1 \leq i \leq r$. Then $E = E_1 + \cdots + E_r$ is an orthogonal direct sum and $\mathbf{\Phi}_i$ is a root system for E_i for all $1 \leq i \leq r$. Thus understanding root systems boils down to understanding irreducible root systems.

Let U be a subspace of E. Then $E = U \oplus U^\perp$; use the Gram-Schmidt process. If σ is an isometry of E and $\sigma(U) \subseteq U$ then $\sigma(U) = U$; thus $\sigma(U^\perp) \subseteq U^\perp$. In particular, if U is σ-invariant then $U^\perp$ is σ-invariant as well.

Lemma 3.2.7 *Let E be a finite-dimensional Euclidean space, let $v \in E \backslash 0$, and let U be a subspace of E. Then:*

(a) *If $v \in U$ then U and $U^\perp$ are σ_v-invariant.*

(b) *Suppose that U is σ_v-invariant. Then $v \in U$ or $v \in U^\perp$.*

PROOF: Suppose that $v \in U$. The formula $\sigma_v(u) = u - \dfrac{2(u,v)}{(v,v)}v$ for all $u \in E$ shows implies that U is σ_v-invariant. To complete the proof of part a) we note that σ_v is a isometry of E.

Suppose that U is σ_v-invariant. Then $U^\perp$ is as well. Write $v = v' + v''$, where $v' \in U$ and $v'' \in U^\perp$. Since $E = U \oplus U^\perp$, the calculation

$$-(v' + v'') = -v = \sigma_v(v) = \sigma_v(v') + \sigma_v(v'')$$

shows that $-v' = \sigma_v(v')$ and $-v'' = \sigma_v(v'')$. Thus

$$v' = \frac{(v',v)}{(v,v)}v \qquad \text{and} \qquad v'' = \frac{(v'',v)}{(v,v)}v.$$

Since $(v', v'') = 0$, we deduce from the two preceding equations that $(v',v)(v'',v) = 0$. Therefore one of (v',v) or (v'',v) is zero, that is $v = v''$ or $v = v'$. We have established part b). $\square$

Corollary 3.2.8 *Let E be a finite-dimensional Euclidean space, let $\mathbf{\Phi}$ be a root system for E, and suppose that $\Delta \subseteq \mathbf{\Phi}$ is a base. Then $\mathbf{\Phi}$ is reducible if and only if Δ is reducible.*

PROOF: Suppose that $\mathbf{\Phi}$ is reducible and write $\mathbf{\Phi} = \mathbf{\Phi}' \cup \mathbf{\Phi}''$, where $\mathbf{\Phi}', \mathbf{\Phi}''$ are non-empty and $(\alpha', \alpha'') = 0$ for all $\alpha' \in \mathbf{\Phi}$ and $\alpha'' \in \mathbf{\Phi}''$. Let $\Delta' = \Delta \cap \mathbf{\Phi}'$ and $\Delta'' = \Delta \cap \mathbf{\Phi}''$. Then $\Delta = \Delta' \cup \Delta''$ and $(\alpha', \alpha'') = 0$ for all $\alpha' \in \Delta'$ and $\alpha'' \in \Delta''$. If $\Delta' = \emptyset$ then $\Delta \subseteq \mathbf{\Phi}''$ and hence $\mathbf{\Phi}' = \emptyset$ since Δ is a basis for E. Thus $\Delta' \neq \emptyset$. Likewise $\Delta'' \neq \emptyset$ and therefore Δ is reducible.

Conversely, suppose that Δ is reducible and write $\Delta = \Delta' \cup \Delta''$, where Δ' and Δ'' are not empty and $(\alpha', \alpha'') = 0$ for all $\alpha' \in \Delta'$ and $\alpha'' \in \Delta''$. Let

U be the span of Δ'. Then $U^\perp$ is the span of Δ''. Now the reflections σ_α, where $\alpha \in \Delta$, generate the Weyl group $\mathcal{W}$. Therefore both $U = U^{\perp\perp}$ and $U^\perp$ are invariant under all $\sigma \in \mathcal{W}$ by part (a) of Lemma 3.2.7. This means that $\mathbf{\Phi} = (U \cap \mathbf{\Phi}) \cup (U^\perp \cap \mathbf{\Phi})$ by Lemma 3.2.7 again and the theorem of the preceding section. We have shown that $\mathbf{\Phi}$ is reducible. $\Box$

Let Δ be any linear basis for E and for $v \in E$ write $v = \sum_{\alpha \in \Delta} v_\alpha \alpha$, where $v_\alpha \in \boldsymbol{R}$ for all $\alpha \in \Delta$. For $u, v \in E$ the relation $u \prec v$ if and only if $u_\alpha \leq v_\alpha$ for all $\alpha \in \Delta$ defines a partial ordering on E. Roots can be compared in terms of this partial order and also in terms of the height function. Here is a refinement of Lemma A.

Lemma 3.2.9 *Let E be a finite-dimensional Euclidean space, let $\mathbf{\Phi}$ be a root system for E, and suppose that $\Delta \subseteq \mathbf{\Phi}$ is a base.*

(a) *There exists a unique root $\beta \in \mathbf{\Phi}$ maximal with respect to the partial order $\prec$.*

(b) *There exist unique root $\beta' \in \mathbf{\Phi}$ of maximal height.*

(c) $\beta = \beta'$

(d) $(\beta, \alpha) \geq (\beta, \alpha) \geq 0$ *for all positive roots* $\alpha \in \mathbf{\Phi}$ *and* $(\beta, \alpha) > 0$ *for some* $\alpha \in \Delta$.

(e) $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$, *where k_α is a positive integer for all $\alpha \in \Delta$*

Proof: Let $\beta' \in \mathbf{\Phi}$ have maximal height. Then β' is maximal with respect to the order $\prec$. Thus parts b) and c) follow from part a).

Let $\beta \in \mathbf{\Phi}$ be maximal with respect to the order $\prec$. Then $\beta \in \mathbf{\Phi}^+$. We first show part d).

Suppose that $(\beta, \alpha) \geq 0$ does not hold for all $\alpha \in \mathbf{\Phi}^+$. Then there exists a $\alpha \in \mathbf{\Phi}^+$ such that $(\beta, \alpha) < 0$. Therefore $\beta \neq \pm\alpha$ and $\beta + \alpha \in \mathbf{\Phi}$. Since $\alpha \in \mathbf{\Phi}^+$ we have $\beta \prec \beta + \alpha$. As $\beta \neq \beta + \alpha$ we have a contradiction. Thus $(\beta, \alpha) \geq 0$ for all $\alpha \in \mathbf{\Phi}^+$ after all. Since Δ is a basis for E it follows that $(\beta, \alpha) \neq 0$ for some $\alpha \in \Delta$. Since $\alpha \in \mathbf{\Phi}^+$ necessarily $(\beta, \alpha) > 0$.

To complete the proof of part (d), let $\alpha \in \mathbf{\Phi}^+$. If $(\beta, \alpha) = 0$ then $(\beta, \beta) \geq (\beta, \alpha) = 0$. In light of the preceding paragraph we may assume that $(\beta, \alpha) > 0$. Thus $\beta - \alpha \in \mathbf{\Phi}$. If $\alpha - \beta \in \mathbf{\Phi}^+$ then $\alpha = (\alpha - \beta) + \beta$ is the sum of positive roots and the maximality of β is contradicted. Thus

$\beta - \alpha \in \mathbf{\Phi}^+$ from which we deduce that $(\beta, \beta) - (\beta, \alpha) = (\beta, \beta - \alpha) \geq 0$. We have established part d).

To show part (e), write $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$, where k_α is a non-negative integer for all $\alpha \in \Delta$. Let $\Delta' = \{\alpha' \in \Delta \,|\, k_{\alpha'} > 0\}$ and let $\Delta'' = \{\alpha'' \in \Delta \,|\, k_{\alpha''} = 0\}$. Then Δ is the union of Δ' and Δ''. Now $(\alpha, \alpha') \leq 0$ for distinct $\alpha, \alpha' \in \Delta$. For $\alpha'' \in \Delta''$ we calculate, using part (d), that

$$0 \leq (\beta, \alpha'') = \sum_{\alpha' \in \Delta'} k_{\alpha'} (\alpha', \alpha'') \leq 0$$

shows that $(\alpha', \alpha'') = 0$ for all $\alpha' \in \Delta'$. Since Δ is irreducible and Δ' is not empty, Δ'' is empty.

Finally we show part (c). Suppose that β'' is also maximal with respect to the partial order $\prec$. Then $(\beta, \beta'') = \sum_{\alpha \in \Delta} k_\alpha (\alpha, \beta'') > 0$ by parts d) and e). Suppose that $\beta \neq \pm\beta''$. Then $\beta - \beta'' \in \mathbf{\Phi}$. One of $\beta - \beta''$ or $\beta'' - \beta$ is positive. In the first case $\beta = (\beta - \beta'') + \beta''$ is the sum of positive roots which contradicts the maximality of β''. In the second case the maximality of β is contradicted. Thus $\beta = \beta''$. $\square$

Recall that a subgroup G of $\mathrm{GL}(E)$ acts on E by the rule $\sigma{\cdot}v = \sigma(v)$ for all $\sigma \in G$ and $v \in E$. The action is irreducible if the only subspaces U of E for which $\sigma(U) \subseteq U$ for all $\sigma \in G$ are E and (0). A paraphrase of Lemma B is:

Lemma 3.2.10 *Let E be a finite-dimensional Euclidean space and let $\mathbf{\Phi}$ be an irreducible root system for E.*

(a) *The Weyl group acts irreducibly on E.*

(b) *For $\alpha \in \mathbf{\Phi}$ the span of $\{\sigma(\alpha) \,|\, \sigma \in \mathcal{W}\}$ is E.*

Proof: Let U be a subspace of E and suppose that $\sigma(U) \subseteq U$ for all $\sigma \in \mathcal{W}$. Let $\mathbf{\Phi}' = \mathbf{\Phi} \cap U$ and $\mathbf{\Phi}'' = \mathbf{\Phi} \cap U^\perp$. Then $\mathbf{\Phi} = \mathbf{\Phi}' \cup \mathbf{\Phi}''$ by Lemma 3.2.7. Since $\mathbf{\Phi}$ is irreducible, $\mathbf{\Phi}' = \mathbf{\Phi}$, in which case $U = E$, or $\mathbf{\Phi}' = \emptyset$, in which case $\mathbf{\Phi}'' = \mathbf{\Phi}$, $U^\perp = E$ and thus $U = (0)$. We have shown part (a).

As for part (b), if $\alpha \in \mathbf{\Phi}$ then the span of $\{\sigma(\alpha) \,|\, \alpha \in \mathbf{\Phi}\}$ is a $\mathcal{W}$-invariant subspace of E which must be E by part (a). $\square$

Here is a refinement of Lemma C.

Lemma 3.2.11 *Let E be a finite-dimensional Euclidean space and let $\mathbf{\Phi}$ be a root system for E. Then:*

(a) *Suppose* $\alpha, \beta \in \mathbf{\Phi}$ *and* $||\alpha|| \leq ||\beta||$. *Then* $(\alpha, \beta) = 0$ *or* $\dfrac{||\beta||^2}{||\alpha||^2} = 1, 2, 3.$

(b) *Let* $\alpha_1, \ldots, \alpha_r \in \mathbf{\Phi}$. *If* $(\alpha_i, \alpha_j) \neq 0$ *for all* $1 \leq i, j \leq r$ *then at most two root lengths occur among* $\{\alpha_1, \ldots, \alpha_r\}$.

(c) *Suppose that* $\mathbf{\Phi}$ *is irreducible. Then at most two root lengths occur among* $\mathbf{\Phi}$.

Proof: To show part (a) we may assume that $\beta \neq \pm\alpha$ and $(\alpha, \beta) \neq 0$. In this case $|{<}\beta, \alpha{>}{<}\alpha, \beta{>}| \leq 3$ so the absolute value of one of the integers ${<}\beta, \alpha{>}.{<}\alpha, \beta{>}$ is 1 and that of the other is $1, 2$, or 3. Since $\dfrac{||\beta||^2}{||\alpha||^2} = \dfrac{{<}\beta, \alpha{>}}{{<}\alpha, \beta{>}}$, part (a) now follows. Part (b) is a direct consequence of part (a).

To show part (c), let $\alpha, \beta \in \mathbf{\Phi}$ and suppose that $||\alpha|| \leq ||\beta||$. By part (b) of Lemma 3.2.10 there is a $\sigma \in \mathcal{W}$ such that $(\beta, \sigma(\alpha)) \neq 0$. Thus $\dfrac{||\beta||^2}{||\alpha||^2} = \dfrac{||\beta||^2}{||\sigma(\alpha)||^2} = 1, 2, 3$ by part (a). Therefore there cannot be roots α, β, ϵ with $||\alpha|| < ||\beta|| < ||\epsilon||$ as

$$\frac{||\epsilon||^2}{||\alpha||^2} = \left(\frac{||\epsilon||^2}{||\beta||^2}\right)\left(\frac{||\beta||^2}{||\alpha||^2}\right) = 4, 6, 9.$$

□

Lemma 3.2.12 *Let* E *be a finite-dimensional Euclidean space, let* $\mathbf{\Phi}$ *be a root system for* E*, and let* $\Delta \subseteq \mathbf{\Phi}$ *be a base. Then the unique root of maximal height is a root of maximal length.*

Proof:

□

3.3 Classification

3.3.1 Cartan matrix of $\mathbf{\Phi}$

We begin with a brief review and slight generalization of relevant material. Let E be an Euclidean space and suppose that $u \in E \backslash 0$. The reflection $\sigma_u : E \longrightarrow E$ through the hyperplane $P_u = u^\perp$ is given by

$$\sigma_u(v) = v - \frac{2(v, u)}{(u, u)} u$$

for all $v \in E$. Set

$$<v, u> = \frac{2(v, u)}{(u, u)} = \frac{2(v, u)}{||u||^2}.$$

Then by definition

$$\sigma_u(v) = v - <v, u>u$$

for all $v \in E$. Note that $<v, u>$ and $<u, v>$ are both zero or they both have the same sign.

Suppose further that $v \in E\backslash 0$. Using the Cauchy–Schwartz inequality with the calculation

$$<v, u><u, v> = \left(\frac{2(v, u)}{||u||^2}\right)\left(\frac{2(u, v)}{||v||^2}\right) = 4\left(\frac{(u, v)}{||u||\;||v||}\right)^2$$

we conclude that

$$<v, u><u, v> = 4\cos^2\theta,$$

where θ is the angle between the vectors u and v. (Recall that $0 \leq \theta \leq \pi$ so the cosine of θ determines θ.) In particular

$$0 \leq <v, u><u, v> \leq 4,$$

with $<v, u><u, v> = 4$ if and only if u and v are scalar multiples of each other.

Observe that $<u, v> = 0$ if and only if $(u, v) = 0$ if and only if $<v, u> = 0$ which is the case if and only if u and v are at right angles to each other. Suppose that $(u, v) \neq 0$. Then

$$\frac{<v, u>}{<u, v>} = \frac{||v||^2}{||u||^2}.$$

Suppose that $\boldsymbol{\Phi}$ is a root system for E. Then $<\beta, \alpha> \in \boldsymbol{Z}$ for all $\beta, \alpha \in \boldsymbol{\Phi}$. *We shall assume that β and α are non-proportional roots and $||\alpha|| \leq ||\beta||$ in the sequel.* Thus

$$<\beta, \alpha><\alpha, \beta> = 0, 1, 2, \quad \text{or} \quad 3.$$

We analyze the possibilities.

Case 0: $<\beta, \alpha><\alpha, \beta> = 0$. This is the case $(\alpha, \beta) = 0$. Thus $<\beta, \alpha> = 0 = <\alpha, \beta>$.

In the remaining cases $(\alpha, \beta) \neq 0$.

Case 1: $<\beta,\alpha><\alpha,\beta> = 1$. Here

$$<\beta,\alpha> = <\alpha,\beta> = \pm 1;$$

thus

$$||\beta|| = ||\alpha||.$$

Since $4\cos^2\theta = 1$ we observe that

$$\theta = \begin{cases} \frac{\pi}{3} & : \ (\alpha,\beta) > 0; \\ \frac{2\pi}{3} & : \ (\alpha,\beta) < 0. \end{cases}$$

Case 2: $<\beta,\alpha><\alpha,\beta> = 2$. Here

$$<\alpha,\beta> = \pm 1 \quad \text{and} \quad <\beta,\alpha> = 2<\alpha,\beta>$$

and

$$||\beta|| = \sqrt{2}||\alpha||.$$

Since $4\cos^2\theta = 2$ we observe that

$$\theta = \begin{cases} \frac{\pi}{4} & : \ (\alpha,\beta) > 0; \\ \frac{3\pi}{4} & : \ (\alpha,\beta) < 0. \end{cases}$$

Case 3: $<\beta,\alpha><\alpha,\beta> = 3$. Here

$$<\alpha,\beta> = \pm 1 \quad \text{and} \quad <\beta,\alpha> = 3<\alpha,\beta>$$

and

$$||\beta|| = \sqrt{3}||\alpha||.$$

Since $4\cos^2\theta = 3$ we observe that

$$\theta = \begin{cases} \frac{\pi}{6} & : \ (\alpha,\beta) > 0; \\ \frac{5\pi}{6} & : \ (\alpha,\beta) < 0. \end{cases}$$

Now let Δ be a base for $\mathbf{\Phi}$ and let

$$\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$$

be a listing of the distinct elements of Δ. Recall that

$$\text{Rank}\, \mathbf{\Phi} = \ell = \text{Dim}\, E.$$

The *Cartan matrix of* $\mathbf{\Phi}$ is $\mathcal{C} = (<\alpha_i, \alpha_j>)$. Since $<\alpha, \alpha> = 2$ for all $\alpha \in \mathbf{\Phi}$ it follows that the diagonal entries of $\mathcal{C}$ are all 2.

Suppose that $\alpha, \beta \in \Delta$ are different. Then α and β are non-proportional roots and

$$(\alpha, \beta) \leq 0.$$

We revisit the cases above.

Case 0': $<\beta, \alpha><\alpha, \beta> = 0$. This is the case $(\alpha, \beta) = 0$; thus $<\beta, \alpha> = 0 = <\alpha, \beta>$.

In the remaining cases $(\alpha, \beta) \neq 0$.

Case 1': $<\beta, \alpha><\alpha, \beta> = 1$. Here

$$<\beta, \alpha> = <\alpha, \beta> = -1,$$

$$||\beta|| = ||\alpha||,$$

and

$$\theta = \frac{2\pi}{3}.$$

Case 2': $<\beta, \alpha><\alpha, \beta> = 2$. Here

$$<\alpha, \beta> = -1 \quad \text{and} \quad <\beta, \alpha> = -2,$$

$$||\beta|| = \sqrt{2}||\alpha||,$$

and

$$\theta = \frac{3\pi}{4}.$$

Case 3': $<\beta, \alpha><\alpha, \beta> = 3$. Here

$$<\alpha, \beta> = -1 \quad \text{and} \quad <\beta, \alpha> = -3,$$

$$||\beta|| = \sqrt{3}||\alpha||,$$

and

$$\theta = \frac{5\pi}{6}.$$

The off-diagonal entries of $\mathcal{C}$ are thus zero or negative. For distinct $1 \leq i, j \leq \ell$ both of $<\alpha_i, \alpha_j>, <\alpha_j, \alpha_i>$ are zero or one is -1 and the other is -1, -2 or -3. The four possibilities fit unambiguously into the four cases just described.

There is only one possible root system (up to isomorphism) with given Cartan matrix.

Proposition 3.3.1 *Suppose that $\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$ and $\Delta' = \{\alpha'_1, \ldots, \alpha'_\ell\}$ are bases for rank ℓ root systems $\mathbf{\Phi}$ for E and $\mathbf{\Phi}'$ for E' respectively. Suppose further that $<\alpha_i, \alpha_j> = <\alpha'_i, \alpha'_j>$ for all $1 \leq i, j \leq \ell$. Then there is a linear isomorphism $\phi : E \longrightarrow E'$ which is an isomorphism of the root systems $\mathbf{\Phi}$ and $\mathbf{\Phi}'$.*

PROOF: We sketch a proof. Before continuing review the discussion following the Lemma of §9.2 in the text.

Since Δ and Δ' are bases for E and E' respectively, the set bijection $\Delta \longrightarrow \Delta'$ given by $\alpha_i \mapsto \alpha'_i$ determines a linear isomorphism $\Phi : E \longrightarrow E'$. Since $<\alpha_i, \alpha_j> = <\alpha'_i, \alpha'_j> = <\phi(\alpha_i), \phi(\alpha_j)>$ for all $1 \leq i, j \leq \ell$ it follows that

$$\sigma_{\alpha'_i}(\phi(\alpha_j)) = \phi(\sigma_{\alpha_i}(\alpha_j))$$

for all $1 \leq i, j \leq \ell$, or equivalently, since Δ spans E, that

$$\sigma_{\alpha'_i} \circ \phi = \phi \circ \sigma_{\alpha_i}$$

for all $1 \leq i \leq \ell$. Since $\mathcal{W}$ is generated by simple reflections, $\phi \circ \mathcal{W} \circ \phi^{-1} \subseteq \mathcal{W}'$. Replacing ϕ by ϕ^{-1} gives $\phi^{-1} \circ \mathcal{W}' \circ \phi \subseteq \mathcal{W}$. We have shown that $\phi \circ \mathcal{W} \circ \phi^{-1} = \mathcal{W}'$.

To complete the proof we need only show that $\phi(\mathbf{\Phi}) = \mathbf{\Phi}'$. Let $\beta \in \mathbf{\Phi}$. Then $\beta = \tau(\alpha_i)$ for some $\tau \in \mathcal{W}$ and $1 \leq i \leq \ell$; see part c) of the theorem of §10.3 of the text. The calculation

$$\phi(\beta) = (\phi \circ \tau \circ \phi^{-1})(\phi(\alpha_i)) = (\phi \circ \tau \circ \phi^{-1})(\alpha'_i) \in (\phi \circ \mathcal{W} \circ \phi^{-1})(\alpha'_i) \subseteq \mathcal{W}'(\alpha'_i) \subseteq \mathbf{\Phi}'$$

shows that $\phi(\mathbf{\Phi}) \subseteq \mathbf{\Phi}'$. Replacing ϕ by ϕ^{-1} gives $\phi^{-1}(\mathbf{\Phi}') \subseteq \mathbf{\Phi}$. Therefore $\phi(\mathbf{\Phi}) = \mathbf{\Phi}'$ as required. $\square$

3.3.2 Coxeter graphs and Dynkin diagrams

See the case discussion for §11.1

3.3.3 Irreducible components

See the discussion of irreducible components for §10.4.

3.3.4 Classification theorem

This is very nicely written — no further comments.