# *Lie algebras*

Course Notes

**Alberto Elduque**

Departamento de Matemáticas
Universidad de Zaragoza
50009 Zaragoza, Spain

These notes are intended to provide an introduction to the basic theory of finite dimensional Lie algebras over an algebraically closed field of characteristic 0 and their representations. They are aimed at beginning graduate students in either Mathematics or Physics.

The basic references that have been used in preparing the notes are the books in the following list. By no means these notes should be considered as an alternative to the reading of these books.


- N. Jacobson: *Lie algebras*, Dover, New York 1979. Republication of the 1962 original (Interscience, New York).
- J.E. Humphreys: *Introduction to Lie algebras and Representation Theory*, GTM **9**, Springer-Verlag, New York 1972.
- W. Fulton and J. Harris: *Representation Theory. A First Course*, GTM **129**, Springer-Verlag, New York 1991.
- W.A. De Graaf: *Lie algebras: Theory and Algorithms*, North Holland Mathematical Library, Elsevier, Amsterdan 2000.

# Contents

# Chapter 1

# A short introduction to Lie groups and Lie algebras

This chapter is devoted to give a brief introduction to the relationship between Lie groups and Lie algebras. This will be done in a concrete way, avoiding the general theory of Lie groups.

It is based on the very nice article by R. Howe: "Very basic Lie Theory", Amer. Math. Monthly **90** (1983), 600–623.

Lie groups are important since they are the basic objects to describe the symmetry. This makes them an unavoidable tool in Geometry (think of Klein's Erlangen Program) and in Theoretical Physics.

A *Lie group* is a group endowed with a structure of smooth manifold, in such a way that both the algebraic group structure and the smooth structure are compatible, in the sense that both the multiplication ($(g,h) \mapsto gh$) and the inverse map ($g \mapsto g^{-1}$) are smooth maps.

To each Lie group a simpler object may be attached: its Lie algebra, which almost determines the group.

**Definition.** A *Lie algebra* over a field $k$ is a vector space $\mathfrak{g}$, endowed with a bilinear multiplication

$$\begin{aligned} [.,.] : \mathfrak{g} \times \mathfrak{g} &\longrightarrow \mathfrak{g} \\ (x,y) &\mapsto [x,y], \end{aligned}$$

satisfying the following properties:

$$\begin{aligned} &[x,x] = 0 \quad \text{(anticommutativity)} \\ &[[x,y],z] + [[y,z],x] + [[z,x],y] = 0 \quad \text{(Jacobi identity)} \end{aligned}$$

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

**Example.** Let $A$ be any associative algebra, with multiplication denoted by juxtaposition. Consider the new multiplication on $A$ given by

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

for any $x,y \in A$. It is an easy exercise to check that $A$, with this multiplication, is a Lie algebra, which will be denoted by $A^-$.

As for any algebraic structure, one can immediately define in a natural way the concepts of subalgebra, ideal, homomorphism, isomorphism, ..., for Lie algebras.

The most usual Lie groups and Lie algebras are "groups of matrices" and their Lie algebras. These concrete groups and algebras are the ones that will be considered in this chapter, thus avoiding the general theory.

## § 1. One-parameter groups and the exponential map

Let $V$ be a real finite dimensional normed vector space with norm $\|.\|$. (So that $V$ is isomorphic to $\mathbb{R}^n$.)

Then $\mathrm{End}_{\mathbb{R}}(V)$ is a normed space with

$$\begin{aligned}\|A\| &= \sup\left\{\frac{\|Av\|}{\|v\|} : 0 \neq v \in V\right\}\\ &= \sup\left\{\|Av\| : v \in V \text{ and } \|v\| = 1\right\}\end{aligned}$$

The determinant provides a continuous (even polynomial) map $\det : \mathrm{End}_{\mathbb{R}}(V) \to \mathbb{R}$. Therefore

$$GL(V) = \det^{-1}\big(\mathbb{R} \setminus \{0\}\big)$$

is an open set of $\mathrm{End}_{\mathbb{R}}(V)$, and it is a group. Moreover, the maps

$$\begin{array}{rcl} GL(V) \times GL(V) & \to & GL(V) \\ (A, B) & \mapsto & AB \end{array} \qquad \begin{array}{rcl} GL(V) & \to & GL(V) \\ A & \mapsto & A^{-1} \end{array}$$

are continuous. (Actually, the first map is polynomial, and the second one rational, so they are smooth and even analytical maps. Thus, $GL(V)$ is a Lie group.)

### One-parameter groups

A *one-parameter group* of transformations of $V$ is a continuous group homomorphism

$$\phi : (\mathbb{R}, +) \longrightarrow GL(V).$$

Any such one-parameter group $\phi$ satisfies the following properties:

**1.1 Properties.**

(i) *$\phi$ is differentiable.*

*Proof.* Let $F(t) = \int_0^t \phi(u)du$. Then $F'(t) = \phi(t)$ for any $t$ and for any $t, s$:

$$\begin{aligned}F(t+s) &= \int_0^{t+s} \phi(u)du\\ &= \int_0^t \phi(u)du + \int_t^{t+s} \phi(u)du\\ &= \int_0^t \phi(u)du + \int_t^{t+s} \phi(t)\phi(u-t)du\\ &= F(t) + \phi(t)\int_0^s \phi(u)du\\ &= F(t) + \phi(t)F(s).\end{aligned}$$

But
$$F'(0) = \lim_{s\to 0} \frac{F(s)}{s} = \phi(0) = I$$
(the identity map on $V$), and the determinant is continuous, so
$$\lim_{s\to 0} \det\left(\frac{F(s)}{s}\right) = \lim_{s\to 0} \frac{\det F(s)}{s^n} = 1 \neq 0,$$
and hence a small $s_0$ can be chosen with invertible $F(s_0)$. Therefore
$$\phi(t) = \big(F(t+s_0) - F(t)\big)F(s_0)^{-1}$$
is differentiable, since so is $F$. □

(ii) *There is a unique $A \in \mathrm{End}_{\mathbb{R}}(V)$ such that*
$$\phi(t) = e^{tA} \quad \left(= \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}\right).$$
*(Note that the series* $\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}$ *converges absolutely, since* $\|A^n\| \leq \|A\|^n$*, and uniformly on each bounded neighborhood of* $0$*, in particular on* $\mathcal{B}_s(0) = \{A \in \mathrm{End}_{\mathbb{R}}(V) : \|A\| < s\}$*, for any* $0 < s \in \mathbb{R}$*, and hence it defines a smooth, in fact analytic, map from* $\mathrm{End}_{\mathbb{R}}(V)$ *to itself.) Besides,* $A = \phi'(0)$*.*

*Proof.* For any $0 \neq v \in V$, let $v(t) = \phi(t)v$. In this way, we have defined a map $\mathbb{R} \to V$, $t \mapsto v(t)$, which is differentiable and satisfies
$$v(t+s) = \phi(s)v(t)$$
for any $s, t \in \mathbb{R}$. Differentiate with respect to $s$ for $s = 0$ to get
$$\begin{cases} v'(t) = \phi'(0)v(t), \\ v(0) = v, \end{cases}$$
which is a linear system of differential equations with constant coefficients. By elementary linear algebra(!), it follows that
$$v(t) = e^{t\phi'(0)}v$$
for any $t$. Moreover,
$$\left(\phi(t) - e^{t\phi'(0)}\right)v = 0$$
for any $v \in V$, and hence $\phi(t) = e^{t\phi'(0)}$ for any $t$. □

(iii) *Conversely, for any $A \in \mathrm{End}_{\mathbb{R}}(V)$, the map $t \mapsto e^{tA}$ is a one-parameter group.*

*Proof.* If $A$ and $B$ are two commuting elements in $\mathrm{End}_{\mathbb{R}}(V)$, then
$$e^A e^B = \lim_{n\to\infty} \Big(\sum_{p=0}^{n} \frac{A^p}{p!}\Big)\Big(\sum_{q=0}^{n} \frac{B^q}{q!}\Big) = \lim_{n\to\infty} \Big(\sum_{r=0}^{n} \frac{(A+B)^r}{r!} + R_n(A,B)\Big),$$

with

$$R_n(A,B) = \sum_{\substack{1\le p,q\le n\\ p+q>n}} \frac{A^p}{p!}\frac{B^q}{q!},$$

so

$$\|R_n(A,B)\| \le \sum_{\substack{1\le p,q\le n\\ p+q>n}} \frac{\|A\|^p}{p!}\frac{\|B\|^q}{q!} \le \sum_{r=n+1}^{2n} \frac{\big(\|A\|+\|B\|\big)^r}{r!},$$

whose limit is 0. Hence, $e^A e^B = e^{A+B}$.

Now, given any $A \in \mathrm{End}_{\mathbb{R}}(V)$ and any scalars $t, s \in \mathbb{R}$, $tA$ commutes with $sA$, so $\phi(t+s) = e^{tA+sA} = e^{tA}e^{sA} = \phi(t)\phi(s)$, thus proving that $\phi$ is a group homomorphism. The continuity is clear. □

(iv) *There exists a positive real number $r$ and an open set $\mathcal{U}$ in $GL(V)$ contained in $\mathcal{B}_s(I)$, with $s = e^r - 1$, such that the "exponential map":*

$$\begin{aligned} \exp : \mathcal{B}_r(0) &\longrightarrow \mathcal{U} \\ A &\mapsto \exp(A) = e^A \end{aligned}$$

*is a homeomorphism.*

*Proof.* exp is differentiable because of its uniform convergence. Moreover, its differential at 0 satisfies:

$$d\exp(0)(A) = \lim_{t\to 0} \frac{e^{tA} - e^0}{t} = A,$$

so that

$$d\exp(0) = id \quad \text{(the identity map on } \mathrm{End}_{\mathbb{R}}(V))$$

and the Inverse Function Theorem applies.

Moreover, $e^A - I = \sum_{n=1}^{\infty} \frac{A^n}{n!}$, so $\|e^A - I\| \le \sum_{n=1}^{\infty} \frac{\|A\|^n}{n!} = e^{\|A\|} - 1$. Thus $\mathcal{U} \subseteq \mathcal{B}_s(I)$. □

Note that for $V = \mathbb{R}$ ($\dim V = 1$), $GL(V) = \mathbb{R} \setminus \{0\}$ and $\exp : \mathbb{R} \to \mathbb{R} \setminus \{0\}$ is not onto, since it does not take negative values.

Also, for $V = \mathbb{R}^2$, identify $\mathrm{End}_{\mathbb{R}}(V)$ with $\mathrm{Mat}_2(\mathbb{R})$. Then, with $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, it follows that $A^2 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $A^3 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $A^4 = I$. It follows that $e^{tA} = \left(\begin{smallmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{smallmatrix}\right)$. In particular, $e^{tA} = e^{(t+2\pi)A}$ and, therefore, exp is not one-to-one.

### Adjoint maps

1. For any $g \in GL(V)$, the linear map $\mathrm{Ad}\, g : \mathrm{End}_{\mathbb{R}}(V) \to \mathrm{End}_{\mathbb{R}}(V)$, $A \mapsto gAg^{-1}$, is an inner automorphism of the associative algebra $\mathrm{End}_{\mathbb{R}}(V)$.
   The continuous group homomorphism

$$\begin{aligned} \mathrm{Ad} : GL(V) &\longrightarrow GL(\mathrm{End}_{\mathbb{R}}(V)) \\ g &\mapsto \mathrm{Ad}\, g, \end{aligned}$$

is called the *adjoint map* of $GL(V)$.

2. For any $A \in \mathrm{End}_{\mathbb{R}}(V)$, the linear map $\mathrm{ad}_A$ (or $\mathrm{ad}\, A$) : $\mathrm{End}_{\mathbb{R}}(V) \to \mathrm{End}_{\mathbb{R}}(V)$, $B \mapsto [A, B] = AB - BA$, is an inner derivation of the associative algebra $\mathrm{End}_{\mathbb{R}}(V)$. The linear map

$$\begin{aligned} \mathrm{ad} : \mathrm{End}_{\mathbb{R}}(V) &\longrightarrow \mathrm{End}_{\mathbb{R}}(\mathrm{End}_{\mathbb{R}}(V)) \\ A \quad &\mapsto \quad \mathrm{ad}_A \quad (\textit{or}\ \mathrm{ad}\, A), \end{aligned}$$

is called the *adjoint map* of $\mathrm{End}_{\mathbb{R}}(V)$.

We will denote by $\mathfrak{gl}(V)$ the Lie algebra $\mathrm{End}_{\mathbb{R}}(V)^-$. Then ad is a homomorphism of Lie algebras $\mathrm{ad} : \mathfrak{gl}(V) \to \mathfrak{gl}(\mathrm{End}_{\mathbb{R}}(V))$.

**1.2 Theorem.** *The following diagram is commutative:*

$$\begin{array}{ccc} \mathfrak{gl}(V) & \xrightarrow{\mathrm{ad}} & \mathfrak{gl}(\mathrm{End}_{\mathbb{R}}(V)) \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ GL(V) & \xrightarrow{\mathrm{Ad}} & GL(\mathrm{End}_{\mathbb{R}}(V)) \end{array} \tag{1.1}$$

*Proof.* The map $\phi : t \mapsto \mathrm{Ad}\exp(tA)$ is a one-parameter group of transformations of $\mathrm{End}_{\mathbb{R}}(V)$ and, therefore,

$$\mathrm{Ad}\exp(tA) = \exp(t\mathcal{A})$$

with $\mathcal{A} = \phi'(0) \in \mathfrak{gl}(\mathrm{End}_{\mathbb{R}}(V))$. Hence,

$$\mathcal{A} = \lim_{t \to 0} \frac{\mathrm{Ad}\bigl(\exp(tA)\bigr) - I}{t}$$

and for any $B \in \mathrm{End}_{\mathbb{R}}(V)$,

$$\begin{aligned} \mathcal{A}(B) &= \lim_{t \to 0} \frac{\exp(tA)B\exp(-tA) - B}{t} \\ &= \frac{d}{dt}\Bigl(\exp(tA)B\exp(-tA)\Bigr)|_{t=0} \\ &= ABI - IBA = \mathrm{ad}_A(B). \end{aligned}$$

Therefore, $\mathcal{A} = \mathrm{ad}_A$ and $\mathrm{Ad}\bigl(\exp(tA)\bigr) = \exp(t\,\mathrm{ad}_A)$ for any $t \in \mathbb{R}$. □

# § 2. Matrix groups

**2.1 Definition.** Given a real vector space $V$, a *matrix group* on $V$ is a closed subgroup of $GL(V)$.

Any matrix group inherits the topology of $GL(V)$, which is an open subset of the normed vector space $\mathrm{End}_{\mathbb{R}}(V)$.

**2.2 Examples.** 1. $GL(V)$ is a matrix group, called the *general linear group.* For $V = \mathbb{R}^n$, we denote it by $GL_n(\mathbb{R})$.

2. $SL(V) = \{A \in GL(V) : \det A = 1\}$ is called the *special linear group*.

3. Given a nondegenerate symmetric bilinear map $b : V \times V \to \mathbb{R}$, the matrix group

$$O(V, b) = \{A \in GL(V) : b(Au, Av) = b(u, v)\ \forall u, v \in V\}$$

is called the *orthogonal group* relative to $b$.

4. Similarly, give a nondegenerate alternating form $b_a : V \times V \to \mathbb{R}$, the matrix group

$$Sp(V, b_a) = \{A \in GL(V) : b_a(Au, Av) = b_a(u, v)\ \forall u, v \in V\}$$

is called the *symplectic group* relative to $b_a$.

5. For any subspace $U$ of $V$, $P(U) = \{A \in GL(V) : A(U) \subseteq U\}$ is a matrix group. By taking a basis of $U$ and completing it to a basis of $V$, it consists of the endomorphisms whose associated matrix is in upper block triangular form.

6. Any intersection of matrix groups is again a matrix group.

7. Let $T_1, \dots, T_n$ be elements in $\mathrm{End}_{\mathbb{R}}(V)$, then $G = \{A \in GL(V) : [T_i, A] = 0\ \forall i = 1, \dots, n\}$ is a matrix group.

   In particular, consider $\mathbb{C}^n$ as a real vector space, by restriction of scalars. There is the natural linear isomorphism

$$\begin{gathered} \mathbb{C}^n \longrightarrow \mathbb{R}^{2n} \\ (x_1 + iy_1, \dots, x_n + iy_n) \mapsto (x_1, \dots, x_n, y_1, \dots, y_n). \end{gathered}$$

   The multiplication by $i$ in $\mathbb{C}^n$ becomes, through this isomorphism, the linear map $J : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, $(x_1, \dots, x_n, y_1, \dots, y_n) \mapsto (-y_1, \dots, -y_n, x_1, \dots, x_n)$. Then we may identify the group of invertible complex $n \times n$ matrices $GL_n(\mathbb{C})$ with the matrix group $\{A \in GL_{2n}(\mathbb{R}) : [J, A] = 0\}$.

8. If $G_i$ is a matrix group on $V_i$, $i = 1, 2$, then $G_1 \times G_2$ is naturally isomorphic to a matrix group on $V_1 \times V_2$.

9. Let $G$ be a matrix group on $V$, and let $G^o$ be its connected component of $I$. Then $G^o$ is a matrix group too.

   *Proof.* For any $x \in G^o$, $xG^o$ is connected (homeomorphic to $G^o$) and $xG^o \cap G^o \neq \emptyset$ (as $x$ belongs to this intersection). Hence $xG^o \cup G^o$ is connected and, by maximality, we conclude that $xG^o \subseteq G^o$. Hence $G^oG^o \subseteq G^o$. Similarly, $(G^o)^{-1}$ is connected, $(G^o)^{-1} \cap G^o \neq \emptyset$, so that $(G^o)^{-1} \subseteq G^o$. Therefore, $G^o$ is a subgroup of $G$. Moreover, $G^o$ is closed, because the closure of a connected set is connected. Hence $G^o$ is a closed subgroup of $GL(V)$. □

10. Given any matrix group on $V$, its *normalizer* $N(G) = \{g \in GL(V) : gGg^{-1} = G\}$ is again a matrix group.

# § 3. The Lie algebra of a matrix group

Let $G$ be a matrix group on the vector space $V$. Consider the set

$$\mathfrak{g} = \{A \in \mathfrak{gl}(V) : \exp(tA) \in G \ \forall t \in \mathbb{R}\}.$$

Our purpose is to prove that $\mathfrak{g}$ is a Lie algebra, called *the Lie algebra of $G$*.

**3.1 Technical Lemma.** (i) *Let $A, B, C \in \mathfrak{gl}(V)$ such that $\|A\|, \|B\|, \|C\| \leq \frac{1}{2}$ and $\exp(A)\exp(B) = \exp(C)$. Then*

$$C = A + B + \frac{1}{2}[A, B] + S$$

*with $\|S\| \leq 65\bigl(\|A\| + \|B\|\bigr)^3$.*

(ii) *For any $A, B \in \mathfrak{gl}(V)$,*

$$\exp(A + B) = \lim_{n\to\infty} \left(\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right)\right)^n \qquad \textit{(Trotter's Formula).}$$

(iii) *For any $A, B \in \mathfrak{gl}(V)$,*

$$\exp([A, B]) = \lim_{n\to\infty} \left[\exp\left(\frac{A}{n}\right) : \exp\left(\frac{B}{n}\right)\right]^{n^2},$$

*where, as usual, $[g : h] = ghg^{-1}h^{-1}$ denotes the commutator of two elements in a group.*

*Proof.* Note that, by continuity, there are real numbers $0 < r, r_1 \leq \frac{1}{2}$, such that $\exp\bigl(\mathcal{B}_{r_1}(0)\bigr)\exp\bigl(\mathcal{B}_{r_1}(0)\bigr) \subseteq \exp\bigl(\mathcal{B}_r(0)\bigr)$. Therefore, item (i) makes sense.

For (i) several steps will be followed. Assume $A, B, C$ satisfy the hypotheses there.

- Write $\exp(C) = I + C + R_1(C)$, with $R_1(C) = \sum_{n=2}^{\infty} \frac{C^n}{n!}$. Hence

$$\|R_1(C)\| \leq \|C\|^2 \sum_{n=2}^{\infty} \frac{\|C\|^{n-2}}{n!} \leq \|C\|^2 \sum_{n=2}^{\infty} \frac{1}{n!} \leq \|C\|^2, \tag{3.2}$$

  because $\|C\| < 1$ and $e - 2 < 1$.

- Also $\exp(A)\exp(B) = I + A + B + R_1(A, B)$, with

$$R_1(A, B) = \sum_{n=2}^{\infty} \frac{1}{n!}\left(\sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}\right).$$

  Hence,

$$\|R_1(A, B)\| \leq \sum_{n=2}^{\infty} \frac{(\|A\| + \|B\|)^n}{n!} \leq (\|A\| + \|B\|)^2, \tag{3.3}$$

  because $\|A\| + \|B\| \leq 1$.

- Therefore, $C = A+B+R_1(A,B)-R_1(C)$ and, since $\|C\| \leq \frac{1}{2}$ and $\|A\|+\|B\| \leq 1$, equations (3.2) and (3.3) give

$$\|C\| \leq \|A\| + \|B\| + \big(\|A\| + \|B\|\big)^2 + \|C\|^2 \leq 2\big(\|A\| + \|B\|\big) + \frac{1}{2}\|C\|,$$

and thus

$$\|C\| \leq 4\big(\|A\| + \|B\|\big). \tag{3.4}$$

Moreover,

$$\begin{aligned} \|C - (A+B)\| &\leq \|R_1(A,B)\| + \|R_1(C)\| \\ &\leq \big(\|A\| + \|B\|\big)^2 + \Big(4\big(\|A\| + \|B\|\big)\Big)^2 \\ &\leq 17\big(\|A\| + \|B\|\big)^2. \end{aligned} \tag{3.5}$$

- Let us take one more term now, thus $\exp(C) = I + C + \dfrac{C^2}{2} + R_2(C)$. The arguments in (3.2) give, since $e - 2 - \frac{1}{2} < \frac{1}{3}$,

$$\|R_2(C)\| \leq \frac{1}{3}\|C\|^3. \tag{3.6}$$

On the other hand,

$$\begin{aligned} \exp(A)\exp(B) &= I + A + B + \frac{1}{2}(A^2 + 2AB + B^2) + R_2(A,B), \\ &= I + A + B + \frac{1}{2}[A,B] + \frac{1}{2}(A+B)^2 + R_2(A,B), \end{aligned} \tag{3.7}$$

with

$$\|R_2(A,B)\| \leq \frac{1}{3}(\|A\| + \|B\|)^3. \tag{3.8}$$

But $\exp(C) = \exp(A)\exp(B)$, so if $S = C - \big(A + B + \frac{1}{2}[A,B]\big)$, by (3.7) we get

$$S = R_2(A,B) + \frac{1}{2}\big((A+B)^2 - C^2\big) - R_2(C)$$

and, because of (3.4), (3.5), (3.6) and (3.8),

$$\begin{aligned} \|S\| &\leq \|R_2(A,B)\| + \frac{1}{2}\|(A+B)(A+B-C) + (A+B-C)C\| + \|R_2(C)\| \\ &\leq \frac{1}{3}\big(\|A\| + \|B\|\big)^3 + \frac{1}{2}\big(\|A\| + \|B\| + \|C\|\big)\|A+B-C\| + \frac{1}{3}\|C\|^3 \\ &\leq \frac{1}{3}\big(\|A\| + \|B\|\big)^3 + \frac{5}{2}\big(\|A\| + \|B\|\big)\cdot 17\big(\|A\| + \|B\|\big)^2 + \frac{1}{3}4^3\big(\|A\| + \|B\|\big)^3 \\ &= \Big(\frac{65}{3} + \frac{85}{2}\Big)\big(\|A\| + \|B\|\big)^3 \leq 65\big(\|A\| + \|B\|\big)^3. \end{aligned}$$

To prove (ii) it is enough to realize that for large enough $n$,

$$\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right) = \exp(C_n),$$

with (because of (3.5)),

$$\|C_n - \frac{A+B}{n}\| \leq 17\left(\frac{\|A\|+\|B\|}{n}\right)^2.$$

In other words,

$$\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right) = \exp\left(\frac{A+B}{n} + O\big(\frac{1}{n^2}\big)\right).$$

Therefore,

$$\left(\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right)\right)^n = \exp(C_n)^n = \exp(nC_n) \xrightarrow[n\to\infty]{} \exp(A+B),$$

since exp is continuous.

Finally, for (iii) use that for large enough $n$,

$$\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right) = \exp\left(\frac{A+B}{n} + \frac{1}{2n^2}[A,B] + S_n\right),$$

with $\|S_n\| \leq 65\frac{\left(\|A\|+\|B\|\right)^3}{n^3}$, because of the first part of the Lemma. Similarly,

$$\exp\left(\frac{A}{n}\right)^{-1}\exp\left(\frac{B}{n}\right)^{-1} = \exp\left(\frac{-A}{n}\right)\exp\left(\frac{-B}{n}\right) = \exp\left(-\frac{A+B}{n} + \frac{1}{2n^2}[A,B] + S'_n\right)$$

with $\|S'_n\| \leq 65\frac{\left(\|A\|+\|B\|\right)^3}{n^3}$. Again by the first part of the Lemma we obtain

$$\left[\exp\left(\frac{A}{n}\right) : \exp\left(\frac{B}{n}\right)\right] = \exp\left(\frac{1}{n^2}[A,B] + O\big(\frac{1}{n^3}\big)\right), \tag{3.9}$$

since $\frac{1}{2}\left[\frac{A+B}{n} + \frac{1}{2n^2}[A,B] + S_n, -\frac{A+B}{n} + \frac{1}{2n^2}[A,B] + S'_n\right] = O\big(\frac{1}{n^3}\big)$, and one can proceed as before. □

**3.2 Theorem.** *Let $G$ be a matrix group on the vector space $V$ and let $\mathfrak{g} = \{A \in \mathfrak{gl}(V) : \exp(tA) \in G\ \forall t \in \mathbb{R}\}$. Then:*

(i) *$\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{gl}(V)$. ($\mathfrak{g}$ is called the Lie algebra of $G$.)*

(ii) *The map $\exp : \mathfrak{g} \to G$ maps a neighborhood of $0$ in $\mathfrak{g}$ bijectively onto a neighborhood of $1$ in $G$. (Here $\mathfrak{g}$ is a real vector space endowed with the topology coming from the norm of $\operatorname{End}_{\mathbb{R}}(V)$ induced by the norm of $V$.)*

*Proof.* By its own definition, $\mathfrak{g}$ is closed under multiplication by real numbers. Now, given any $A, B \in \mathfrak{g}$ and $t \in \mathbb{R}$, since $G$ is closed, the Technical Lemma shows us that

$$\exp\bigl(t(A+B)\bigr) = \lim_{n\to\infty}\left(\exp\left(\frac{tA}{n}\right)\exp\left(\frac{tB}{n}\right)\right)^n \in G,$$
$$\exp\bigl(t[A,B]\bigr) = \lim_{n\to\infty}\left[\exp\left(\frac{tA}{n}\right) : \exp\left(\frac{B}{n}\right)\right]^{n^2} \in G.$$

Hence $\mathfrak{g}$ is closed too under addition and Lie brackets, and so it is a Lie subalgebra of $\mathfrak{gl}(V)$.

To prove the second part of the Theorem, let us first check that, if $(A_n)_{n\in\mathbb{N}}$ is a sequence in $\exp^{-1}(G)$ with $\lim_{n\to\infty}\|A_n\| = 0$, and $(s_n)_{n\in\mathbb{N}}$ is a sequence of real numbers, then any cluster point $B$ of the sequence $(s_nA_n)_{n\in\mathbb{N}}$ lies in $\mathfrak{g}$:

Actually, we may assume that $\lim_{n\to\infty} s_nA_n = B$. Let $t \in \mathbb{R}$. For any $n \in \mathbb{N}$, take $m_n \in \mathbb{Z}$ such that $|m_n - ts_n| \leq 1$. Then,

$$\begin{aligned}\|m_nA_n - tB\| &= \|(m_n - ts_n)A_n + t(s_nA_n - B)\|\\ &\leq |m_n - ts_n|\|A_n\| + |t|\|s_nA_n - B\|.\end{aligned}$$

Since both $\|A_n\|$ and $\|s_nA_n - B\|$ converge to 0, it follows that $\lim_{n\to\infty} m_nA_n = tB$. Also, $A_n \in \exp^{-1}(G)$, so that $\exp(m_nA_n) = \exp(A_n)^{m_n} \in G$. Since exp is continuous and $G$ is closed, $\exp(tB) = \lim_{n\to\infty}\exp(m_nA_n) \in G$ for any $t \in \mathbb{R}$, and hence $B \in \mathfrak{g}$, as required.

Let now $\mathfrak{m}$ be a subspace of $\mathfrak{gl}(V)$ with $\mathfrak{gl}(V) = \mathfrak{g} \oplus \mathfrak{m}$, and let $\pi_{\mathfrak{g}}$ and $\pi_{\mathfrak{m}}$ be the associated projections onto $\mathfrak{g}$ and $\mathfrak{m}$. Consider the analytical function:

$$\begin{aligned}E : \mathfrak{gl}(V) &\longrightarrow GL(V)\\ A &\mapsto \exp\bigl(\pi_{\mathfrak{g}}(A)\bigr)\exp\bigl(\pi_{\mathfrak{m}}(A)\bigr).\end{aligned}$$

Then,

$$\begin{aligned}&\frac{d}{dt}\Bigl(\exp\bigl(\pi_{\mathfrak{g}}(tA)\bigr)\exp\bigl(\pi_{\mathfrak{m}}(tA)\bigr)\Bigr)|_{t=0}\\ &\qquad = \frac{d}{dt}\Bigl(\exp\bigl(\pi_{\mathfrak{g}}(tA)\bigr)\Bigr)|_{t=0}\exp(0) + \exp(0)\frac{d}{dt}\Bigl(\exp\bigl(\pi_{\mathfrak{m}}(tA)\bigr)\Bigr)|_{t=0}\\ &\qquad = \pi_{\mathfrak{g}}(A) + \pi_{\mathfrak{m}}(A) = A.\end{aligned}$$

Hence, the differential of $E$ at 0 is the identity and, thus, $E$ maps homeomorphically a neighborhood of 0 in $\mathfrak{gl}(V)$ onto a neighborhood of 1 in $GL(V)$. Let us take $r > 0$ and a neighborhood $\mathcal{V}$ of 1 in $GL(V)$ such that $E|_{\mathcal{B}_r(0)} : \mathcal{B}_r(0) \to \mathcal{V}$ is a homeomorphism. It is enough to check that $\exp\bigl(\mathcal{B}_r(0)\cap\mathfrak{g}\bigr) = E\bigl(\mathcal{B}_r(0)\cap\mathfrak{g}\bigr)$ contains a neighborhood of 1 in $G$.

Otherwise, there would exist a sequence $(B_n)_{n\in\mathbb{N}} \in \exp^{-1}(G)$ with $B_n \notin \mathcal{B}_r(0)\cap\mathfrak{g}$ and such that $\lim_{n\to\infty} B_n = 0$. For large enough $n$, $\exp(B_n) = E(A_n)$, with $\lim_{n\to\infty} A_n = 0$. Hence $\exp\bigl(\pi_{\mathfrak{m}}(A_n)\bigr) = \exp\bigl(\pi_{\mathfrak{g}}(A_n)\bigr)^{-1}\exp(B_n) \in G$.

Since $\lim_{n\to\infty} A_n = 0$, $\lim_{n\to\infty}\pi_{\mathfrak{m}}(A_n) = 0$ too and, for large enough $m$, $\pi_{\mathfrak{m}}(A_m) \neq 0$, as $A_m \notin \mathfrak{g}$ (note that if $A_m \in \mathfrak{g}$, then $\exp(B_m) = E(A_m) = \exp(A_m)$ and since exp is a bijection on a neighborhood of 0, we would have $B_m = A_m \in \mathfrak{g}$, a contradiction).

The sequence $\left(\frac{1}{\|\pi_{\mathfrak{m}}(A_n)\|}\pi_{\mathfrak{m}}(A_n)\right)_{n\in\mathbb{N}}$ is bounded, and hence has cluster points, which are in $\mathfrak{m}$ (closed in $\mathfrak{gl}(V)$, since it is a subspace). We know that these cluster points are in $\mathfrak{g}$, so in $\mathfrak{g}\cap\mathfrak{m} = 0$. But the norm of all these cluster points is 1, a contradiction. $\square$

**3.3 Remark.** Given any $A \in \mathfrak{gl}(V)$, the set $\{\exp(tA) : t \in \mathbb{R}\}$ is the continuous image of the real line, and hence it is connected. Therefore, if $\mathfrak{g}$ is the Lie algebra of the matrix group $G$, $\exp(\mathfrak{g})$ is contained in the connected component $G^o$ of the identity. Therefore, *the Lie algebra of $G$ equals the Lie algebra of $G^o$.*

Also, $\exp(\mathfrak{g})$ contains an open neighborhood $\mathcal{U}$ of 1 in $G$. Thus, $G^o$ contains the open neighborhood $x\mathcal{U}$ of any $x \in G^o$. Hence $G^o$ is open in $G$ but, as a connected component, it is closed too: $G^o$ is open and closed in $G$.

Let us look at the Lie algebras of some interesting matrix groups.

**3.4 Examples.**

1. The Lie algebra of $GL(V)$ is obviously the whole *general linear Lie algebra* $\mathfrak{gl}(V)$.

2. For any $A \in \mathfrak{gl}(V)$ (or any square matrix $A$), $\det e^A = e^{\operatorname{trace}(A)}$. This is better checked for matrices. Since any real matrix can be considered as a complex matrix, it is well known that given any such matrix there is a regular complex matrix $P$ such that $J = PAP^{-1}$ is upper triangular. Assume that $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $A$ (or $J$), counted according to their multiplicities. Then $Pe^AP^{-1} = e^J$ and $\det e^A = \det e^J = \prod_{i=1}^n e^{\lambda_i} = e^{\sum_{i=1}^n \lambda_i} = e^{\operatorname{trace}(J)} = e^{\operatorname{trace}(A)}$.

   Hence, for any $t \neq 0$, $\det e^{tA} = 1$ if and only if $\operatorname{trace}(A) = 0$. This shows that the Lie algebra of the special linear group $SL(V)$ is the *special linear Lie algebra* $\mathfrak{sl}(V) = \{A \in \mathfrak{gl}(V) : \operatorname{trace}(A) = 0\}$.

3. Let $b : V \times V \to \mathbb{R}$ be a bilinear form. If $A \in \mathfrak{gl}(V)$ satisfies $b(e^{tA}v, e^{tA}w) = b(v, w)$ for any $t \in \mathbb{R}$ and $v, w \in V$, take derivatives at $t = 0$ to get $b(Av, w) + b(v, Aw) = 0$ for any $v, w \in V$. Conversely, if $b(Av, w) = -b(v, Aw)$ for any $v, w \in V$, then $b((tA)^n v, w) = (-1)^n b(v, (tA)^n w)$, so $b(e^{tA}v, e^{tA}w) = b(v, e^{-tA}e^{tA}w) = b(v, w)$ for any $t \in \mathbb{R}$ and $v, w, \in V$.

   Therefore, the Lie algebra of the matrix group $G = \{g \in GL(V) : b(gv, gw) = b(v, w)\ \forall v, w \in V\}$ is $\mathfrak{g} = \{A \in \mathfrak{gl}(V) : b(Av, w) + b(v, Aw) = 0\ \forall v, w \in V\}$.

   In particular, if $b$ is symmetric and nondegenerate, the Lie algebra of the orthogonal group $O(V, b)$ is called the *orthogonal Lie algebra* and denoted by $\mathfrak{o}(V, b)$. Also, for alternating and nondegenerate $b_a$, the Lie algebra of the symplectic group $Sp(V, b_a)$ is called the *symplectic Lie algebra*, and denoted by $\mathfrak{sp}(V, b_a)$.

4. For any subspace $U$ of $V$, consider a complementary subspace $W$, so that $V = U \oplus W$. Let $\pi_U$ and $\pi_W$ be the corresponding projections. For any $A \in \mathfrak{gl}(V)$ and $0 \neq t \in \mathbb{R}$, $e^{tA}(U) \subseteq U$ if and only if $\pi_W(e^{tA}u) = 0$ for any $u \in U$. In this case, by taking derivatives at $t = 0$ we obtain that $\pi_W(Au) = 0$ for any $u \in U$, or $A(U) \subseteq U$. The converse is clear. Hence, the Lie algebra of $P(U) = \{g \in GL(V) : g(U) \subseteq U\}$ is $\mathfrak{p}(U) = \{A \in \mathfrak{gl}(V) : A(U) \subseteq U\}$.

5. The Lie algebra of an intersection of matrix groups is the intersection of the corresponding Lie algebras.

6. The Lie algebra of $G = G_1 \times G_2 \big(\subseteq GL(V_1) \times GL(V_2)\big)$ is the direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ of the corresponding Lie algebras. This follows from the previous items because, inside $GL(V_1 \times V_2)$, $GL(V_1) \times GL(V_2) = P(V_1) \cap P(V_2)$.

7. Given $T_1, \dots T_n \in \mathrm{End}_{\mathbb{R}}(V)$. With similar arguments, one checks that the Lie algebra of $G = \{g \in GL(V) : gT_i = T_i g,\ i = 1, \dots, n\}$ is $\mathfrak{g} = \{A \in \mathfrak{gl}(V) : AT_i = T_i A,\ i = 1, \dots, n\}$. In particular, the Lie algebra of $GL_n(\mathbb{C})$ is $\mathfrak{gl}_n(\mathbb{C})$ (the Lie algebra of complex $n \times n$ matrices).

In the remainder of this chapter, the most interesting properties of the relationship between matrix groups and their Lie algebras will be reviewed.

**3.5 Proposition.** *Let $G$ be a matrix group on a real vector space $V$, and let $G^o$ be its connected component of* $1$. *Let $\mathfrak{g}$ be the Lie algebra of $G$. Then $G^o$ is the group generated by* $\exp(\mathfrak{g})$.

*Proof.* We already know that $\exp(\mathfrak{g}) \subseteq G^o$ and that there exists an open neighborhood $\mathcal{U}$ of $1 \in G$ with $1 \in \mathcal{U} \subseteq \exp(\mathfrak{g})$. Let $\mathcal{V} = \mathcal{U} \cap \mathcal{U}^{-1}$, which is again an open neighborhood of 1 in $G$ contained in $\exp(\mathfrak{g})$. It is enough to prove that $G^o$ is generated, as a group, by $\mathcal{V}$.

Let $H = \cup_{n \in \mathbb{N}} \mathcal{V}^n$, $H$ is closed under multiplication and inverses, so it is a subgroup of $G$ contained in $G^o$. Actually, it is the subgroup of $G$ generated by $\mathcal{V}$. Since $\mathcal{V}$ is open, so is $\mathcal{V}^n = \cup_{v \in \mathcal{V}} v\mathcal{V}^{n-1}$ for any $n$, and hence $H$ is an open subgroup of $G$. But any open subgroup is closed too, as $G \setminus H = \cup_{x \in G \setminus H} xH$ is a union of open sets. Therefore, $H$ is an open and closed subset of $G$ contained in the connected component $G^o$, and hence it fills all of $G^o$. $\square$

**3.6 Theorem.** *Let $G$ and $H$ be matrix groups on the real vector space $V$ with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$.*

(i) *If $H$ is a normal subgroup of $G$, then $\mathfrak{h}$ is an ideal of $\mathfrak{g}$ (that is, $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$).*

(ii) *If both $G$ and $H$ are connected, the converse is valid too.*

*Proof.* Assume that $H$ is a normal subgroup of $G$ and let $A \in \mathfrak{h}$ and $B \in \mathfrak{g}$. Since $H$ is a normal subgroup of $G$, for any $t \in \mathbb{R}$ and $n \in \mathbb{N}$, $\left[e^{t\frac{A}{n}} : e^{\frac{B}{n}}\right] \in H$, and hence, by the Technical Lemma, $e^{t[A,B]} = \lim_{n \to \infty} \left[e^{t\frac{A}{n}} : e^{\frac{B}{n}}\right]^{n^2}$ belongs to $H$ ($H$ is a matrix group, hence closed). Thus, $[A, B] \in \mathfrak{h}$.

Now, assume that both $G$ and $H$ are connected and that $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. Then, for any $B \in \mathfrak{g}$, $\mathrm{ad}_B(\mathfrak{h}) \subseteq \mathfrak{h}$, so $\mathrm{Ad}\, e^B(\mathfrak{h}) = e^{\mathrm{ad}_B}(\mathfrak{h}) \subseteq \mathfrak{h}$. In other words, $e^B \mathfrak{h} e^{-B} \subseteq \mathfrak{h}$. Since $G$ is connected, it is generated by $\{e^B : B \in \mathfrak{g}\}$. Hence, $g\mathfrak{h}g^{-1} \subseteq \mathfrak{h}$ for any $g \in G$. Thus, for any $A \in \mathfrak{h}$ and $g \in G$, $ge^A g^{-1} = e^{gAg^{-1}} \in e^{\mathfrak{h}} \subseteq H$. Since $H$ is connected, it is generated by the $e^A$'s, so we conclude that $gHg^{-1} \subseteq H$ for any $g \in G$, and hence $H$ is a normal subgroup of $G$. $\square$

**3.7 Theorem.** *Let $G$ be a matrix group on the real vector space $V$ with Lie algebra $\mathfrak{g}$, and let $H$ be a matrix group on the real vector space $W$ with Lie algebra $\mathfrak{h}$.*

*If $\varphi : G \to H$ is a continuous homomorphism of groups, then there exists a unique Lie algebra homomorphism $d\varphi : \mathfrak{g} \to \mathfrak{h}$ that makes the following diagram commutative:*

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "d\varphi"] \arrow[d, "\exp"'] & \mathfrak{h} \arrow[d, "\exp"] \\ G \arrow[r, "\varphi"] & H \end{tikzcd}$$

*Proof.* The uniqueness is easy: since exp is bijective on a neighborhood of 0 in $\mathfrak{h}$, $d\varphi$ is determined as $(\exp)^{-1} \circ \varphi \circ \exp$ on a neighborhood of 0 in $\mathfrak{g}$. But $d\varphi$ is linear and any neighborhood of 0 contains a basis. Hence $d\varphi$ is determined by $\varphi$.

Now, to prove the existence of such a linear map, take any $A \in \mathfrak{g}$, then $t \mapsto \varphi(e^{tA})$ is a one-parameter group on $W$ with image in $H$. Thus, there is a unique $B \in \mathfrak{h}$ such that $\varphi(e^{tA}) = e^{tB}$ for any $t \in \mathbb{R}$. Define $d\varphi(A) = B$. Therefore, $\varphi(e^{tA}) = e^{td\varphi(A)}$ for any $t \in \mathbb{R}$ and $A \in \mathfrak{g}$. Now, for any $A_1, A_2 \in \mathfrak{g}$,

$$\begin{aligned}
\varphi\left(e^{t(A_1+A_2)}\right) &= \varphi\left(\lim_{n\to\infty}\left(e^{\frac{tA_1}{n}}e^{\frac{tA_2}{n}}\right)^n\right) \quad \text{(Trotter's formula)}\\
&= \lim_{n\to\infty}\left(\varphi\left(e^{\frac{tA_1}{n}}\right)\varphi\left(e^{\frac{tA_2}{n}}\right)\right)^n \quad \text{(since } \varphi \text{ is continuous)}\\
&= \lim_{n\to\infty}\left(e^{\frac{t}{n}d\varphi(A_1)}e^{\frac{t}{n}d\varphi(A_2)}\right)^n\\
&= e^{t\left(d\varphi(A_1)+d\varphi(A_2)\right)}
\end{aligned}$$

and, hence, $d\varphi$ is linear. In the same spirit, one checks that

$$\varphi\left(e^{t[A_1,A_2]}\right) = \varphi\left(\lim_{n\to\infty}\left[e^{\frac{tA_1}{n}} : e^{\frac{A_2}{n}}\right]^{n^2}\right) = \cdots = e^{t[d\varphi(A_1),d\varphi(A_2)]},$$

thus proving that $d\varphi$ is a Lie algebra homomorphism. □

Several consequences of this Theorem will be drawn in what follows.

**3.8 Corollary.** *Let $G, H, \mathfrak{g}$ and $\mathfrak{h}$ be as in the previous Theorem. If $G$ and $H$ are isomorphic matrix groups, then $\mathfrak{g}$ and $\mathfrak{h}$ are isomorphic Lie algebras.*

**3.9 Remark.** The converse of the Corollary above is false, even if $G$ and $H$ are connected. Take, for instance,

$$G = SO(2) = \left\{\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix} : \theta \in \mathbb{R}\right\}$$

(the special orthogonal group on $\mathbb{R}^2$, which is homeomorphic to the unit circle). Its Lie algebra is

$$\mathfrak{g} = \left\{\begin{pmatrix}0 & -\alpha\\ \alpha & 0\end{pmatrix} : \alpha \in \mathbb{R}\right\}$$

($2 \times 2$ skew-symmetric matrices). Also, take

$$H = \left\{\begin{pmatrix}\alpha & 0\\ 0 & 1\end{pmatrix} : \alpha \in \mathbb{R}_{>0}\right\}$$

which is isomorphic to the multiplicative group of positive real numbers, whose Lie algebra is

$$\mathfrak{h} = \left\{\begin{pmatrix}\alpha & 0\\ 0 & 0\end{pmatrix} : \alpha \in \mathbb{R}\right\}.$$

Both Lie algebras are one-dimensional vector spaces with trivial Lie bracket, and hence they are isomorphic as Lie algebras. However, $G$ is not isomorphic to $H$ (inside $G$ one may find many finite order elements, but the identity is the only such element in $H$). (One can show that $G$ and $H$ are 'locally isomorphic'.)

If $G$ is a matrix group on $V$, and $X \in \mathfrak{g}$, $g \in G$ and $t \in \mathbb{R}$,

$$\exp(t \operatorname{Ad} g(X)) = g\big(\exp(tX)\big)g^{-1} \in G,$$

so $\operatorname{Ad} g(\mathfrak{g}) \subseteq \mathfrak{g}$. Hence, the adjoint map of $GL(V)$ induces an *adjoint map*

$$\operatorname{Ad} : G \longrightarrow GL(\mathfrak{g}),$$

and, by restriction in (1.1), we get the following commutative diagram:

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\text{ad}} & \mathfrak{gl}(\mathfrak{g}) \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow{\text{Ad}} & GL(\mathfrak{g}) \end{array} \tag{3.10}$$

**3.10 Corollary.** *Let $G$ be a matrix group on the real vector space $V$ and let* $\operatorname{Ad} : G \to \operatorname{GL}(\mathfrak{g})$ *be the adjoint map. Then* $d \operatorname{Ad} = \operatorname{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$.

Remember that, given a group $G$, its center $Z(G)$ is the normal subgroup consisting of those elements commuting with every element: $Z(G) = \{g \in G : gh = hg \ \forall h \in G\}$.

**3.11 Corollary.** *Let $G$ be a connected matrix group with Lie algebra $\mathfrak{g}$. Then* $Z(G) = \ker \operatorname{Ad}$, *and this is a closed subgroup of $G$ with Lie algebra the* center *of* $\mathfrak{g}$*:* $Z(\mathfrak{g}) = \{X \in \mathfrak{g} : [X, Y] = 0 \ \forall Y \in \mathfrak{g}\}$ $(= \ker \operatorname{ad})$.

*Proof.* With $g \in Z(G)$ and $X \in \mathfrak{g}$, $\exp(tX) = g\big(\exp(tX)\big)g^{-1} = \exp\big(t \operatorname{Ad} g(X)\big)$ for any $t \in \mathbb{R}$. Taking the derivative at $t = 0$ one gets $\operatorname{Ad} g(X) = X$ for any $X \in \mathfrak{g}$, so that $g \in \ker \operatorname{Ad}$. (Note that we have not used here the fact that $G$ is connected.)

Conversely, take an element $g \in \ker \operatorname{Ad}$, so for any $X \in \mathfrak{g}$ we have $g \exp(X) g^{-1} = \exp\big(\operatorname{Ad} g(X)\big) = \exp(X)$. Since $G$ is connected, it is generated by $\exp(\mathfrak{g})$ and, thus, $ghg^{-1} = h$ for any $h \in G$. That is, $g \in Z(G)$.

Since Ad is continuous, it follows that $Z(G) = \ker \operatorname{Ad} = \operatorname{Ad}^{-1}(I)$ is closed.

Now, the commutativity of the diagram (3.10) shows that $\exp(\ker \operatorname{ad}) \subseteq \ker \operatorname{Ad} = Z(G)$, and hence $\ker \operatorname{ad}$ is contained in the Lie algebra of $Z(G)$. Conversely, if $X \in \mathfrak{g}$ and $\exp(tX) \in Z(G)$ for any $t \in \mathbb{R}$, then $\exp(t \operatorname{ad}_X) = \operatorname{Ad} \exp(tX) = I$ and hence (take the derivative at $t = 0$) $\operatorname{ad}_X = 0$, so $X \in \ker \operatorname{ad}$. Therefore, the Lie algebra of $Z(G) = \ker \operatorname{Ad}$ is $\ker \operatorname{ad}$ which, by its own definition, is the center of $\mathfrak{g}$. □

**3.12 Corollary.** *Let $G$ be a connected matrix group with Lie algebra $\mathfrak{g}$. Then $G$ is commutative if and only if $\mathfrak{g}$ is* abelian*, that is,* $[\mathfrak{g}, \mathfrak{g}] = 0$.

Finally, the main concept of this course will be introduced. Groups are important because they act as symmetries of other structures. The formalization, in our setting, of this leads to the following definition:

**3.13 Definition.** (i) A *representation* of a matrix group $G$ is a continuous homomorphism $\rho : G \to GL(W)$ for a real vector space $W$.

(ii) A *representation* of a Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism $\rho : \mathfrak{g} \to \mathfrak{gl}(W)$, for a vector space $W$.

**3.14 Corollary.** *Let $G$ be a matrix group with Lie algebra $\mathfrak{g}$ and let $\rho : G \to GL(W)$ be a representation of $G$. Then there is a unique representation $d\,\rho : \mathfrak{g} \to \mathfrak{gl}(W)$ such that the following diagram is commutative:*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{d\,\rho} & \mathfrak{gl}(W) \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow{\rho} & GL(W) \end{array}$$

The great advantage of dealing with $d\,\rho$ above is that this is a Lie algebra homomorphism, and it does not involve topology. In this sense, the representation $d\,\rho$ is simpler than the representation of the matrix group, but it contains a lot of information about the latter. The message is that in order to study the representations of the matrix groups, we will study representations of Lie algebras.

# Chapter 2

# Lie algebras

The purpose of this chapter is to present the basic structure of the finite dimensional Lie algebras over fields, culminating in the classification of the simple Lie algebras over algebraically closed fields of characteristic 0.

## § 1. Theorems of Engel and Lie

Let us first recall the definition of representation of a Lie algebra, that has already appeared in the previous chapter.

**1.1 Definition.** Let $L$ be a Lie algebra over a field $k$. A *representation* of $L$ is a Lie algebra homomorphism $\rho : L \to \mathfrak{gl}(V)$, where $V$ is a nonzero vector space over $k$.

We will use the notation $x.v = \rho(x)(v)$ for elements $x \in L$ and $v \in V$. In this case, $V$ is said to be a *module* for $L$.

As for groups, rings or associative algebras, we can immediately define the concepts of submodule, quotient module, irreducible module (or irreducible representation), homomorphism of modules, kernel, image, ...

*In what follows, and unless otherwise stated, all the vector spaces and algebras considered will be assumed to be finite dimensional over a ground field $k$.*

**1.2 Engel's Theorem.** *Let $\rho : L \to \mathfrak{gl}(V)$ be a representation of a Lie algebra $L$ such that $\rho(x)$ is nilpotent for any $x \in L$. Then there is an element $0 \neq v \in V$ such that $x.v = 0$ for any $x \in L$.*

*Proof.* The proof will be done by induction on $n = \dim_k L$, being obvious for $n = 1$.

Hence assume that $\dim_k L = n > 1$ and that the result is true for Lie algebras of smaller dimension. If $\ker \rho \neq 0$, then $\dim_k \rho(L) < \dim_k L = n$, but the inclusion $\rho(L) \hookrightarrow \mathfrak{gl}(V)$ is a representation of the Lie algebra $\rho(L)$ and the induction hypothesis applies.

Therefore, we may assume that $\ker \rho = 0$ and, thus, that $L$ is a subalgebra of $\mathfrak{gl}(V)$. The hypothesis of the Theorem assert then that $x^m = 0$ for any $x \in L \subseteq \mathfrak{gl}(V) = \operatorname{End}_k(V)$, where $m = \dim_k V$. Let $S$ be a proper maximal subalgebra of $L$. For any $x, y \in L$ $\operatorname{ad} x = l_x - r_x$, with $l_x(y) = xy = r_y(x)$, so

$$(\operatorname{ad} x)^{2m-1}(y) = (l_x - r_x)^{2m-1}(y) = \sum_{i=0}^{2m-1} (-1)^i \binom{2m-1}{i} x^{2m-1-i} y x^i.$$

But for any $0 \leq i \leq 2m-1$, either $i$ or $2m-1-i$ is $\geq m$. Hence $(\operatorname{ad} x)^{2m-1} = 0$. In particular, the natural representation of the Lie algebra $S$ on the quotient space $L/S$:

$$\begin{aligned} \varphi : S &\longrightarrow \mathfrak{gl}(L/S) \\ x &\mapsto \varphi(x) : L/S \to L/S \\ &\qquad\qquad y + S \mapsto [x, y] + S \end{aligned}$$

($L$ is a module for $S$ through ad, and $L/S$ is a quotient module) satisfies the hypotheses of the Theorem, but with $\dim_k S < n$. By the induction hypothesis, there exists an element $z \in L \setminus S$ such that $[x, z] \in S$ for any $x \in S$. Therefore, $S \oplus kz$ is a subalgebra of $L$ which, by maximality of $S$, is the whole $L$. In particular $S$ is an ideal of $L$.

Again, by induction, we conclude that the subspace $W = \{v \in V : x.v = 0\ \forall x \in S\}$ is nonzero. But for any $x \in S$, $x.(z.W) \subseteq [x, z].W + z.(x.W) = 0$ ($[x, z] \in S$). Hence $z.W \subseteq W$, and since $z$ is a nilpotent endomorphism, there is a nonzero $v \in W$ such that $z.v = 0$. Hence $x.v = 0$ for any $x \in S$ and for $z$, so $x.v = 0$ for any $x \in L$. □

**1.3 Consequences.** (i) Let $\rho : L \to \mathfrak{gl}(V)$ be an irreducible representation of a Lie algebra $L$ and let $I$ be an ideal of $L$ such that $\rho(x)$ is nilpotent for any $x \in I$. Then $I \subseteq \ker \rho$.

*Proof.* Let $W = \{v \in V : x.v = 0\ \forall x \in I\}$, which is not zero by Engel's Theorem. For any $x \in I$, $y \in L$ and $w \in W$, $x.(y.w) = [x, y].w + y.(x.w) = 0$, as $[x, y] \in I$. Hence $W$ is a nonzero submodule of the irreducible module $V$ and, therefore, $W = V$, as required. □

(ii) Let $\rho : L \to \mathfrak{gl}(V)$ be a representation of a Lie algebra $L$. Let $I$ be and ideal of $L$ and let $0 = V_0 \subsetneqq V_1 \subsetneqq \cdots \subsetneqq V_n = V$ be a composition series of $V$. Then $\rho(x)$ is nilpotent for any $x \in I$ if and only if for any $i = 1, \ldots, n$, $I.V_i \subseteq V_{i-1}$.

(iii) The *descending central series* of a Lie algebra $L$ is the chain of ideals $L = L^1 \supseteqq L^2 \supseteqq \cdots \supseteqq L^n \supseteqq \cdots$, where $L^{n+1} = [L^n, L]$ for any $n \in \mathbb{N}$. The Lie algebra is said to be *nilpotent* if there is an $n \in \mathbb{N}$ such that $L^n = 0$. Moreover, if $n = 2$, $L$ is said to be *abelian*. Then

**Theorem.** ***(Engel)*** *A Lie algebra $L$ is nilpotent if and only if* $\operatorname{ad}_x$ *is nilpotent for any $x \in L$.*

*Proof.* It is clear that if $L^n = 0$, then $\operatorname{ad}_x^{n-1} = 0$ for any $x \in L$. Conversely, assume that $\operatorname{ad}_x$ is nilpotent for any $x \in L$, and consider the *adjoint representation* $\operatorname{ad} : L \to \mathfrak{gl}(L)$. Let $0 = L_0 \subsetneqq \cdots \subsetneqq L_{n+1} = L$ be a composition series of this representation. By item (ii) it follows that $L.L_i = [L, L_i] \subseteq L_{i-1}$ for any $i$. Hence $L^i \subseteq L_{n+1-i}$ for any $i$. In particular $L^{n+1} = 0$ and $L$ is nilpotent. □

**1.4 Exercise.** The *ascending central series* of a Lie algebra $L$ is defined as follows: $Z_0(L) = 0$, $Z_1(L) = Z(L) = \{x \in L : [x, L] = 0\}$ (the *center* of $L$) and $Z_{i+1}(L)/Z_i(L) = Z\left(L/Z_i(L)\right)$ for any $i \geq 1$. Prove that this is indeed an ascending chain of ideals and that $L$ is nilpotent if and only if there is an $n \in \mathbb{N}$ such that $Z_n(L) = L$.

Now we arrive to a concept which is weaker than nilpotency.

**1.5 Definition.** Let $L$ be a Lie algebra and consider the descending chain of ideals defined by $L^{(0)} = L$ and $L^{(m+1)} = [L^{(m)}, L^{(m)}]$ for any $m \geq 0$. Then the chain $L = L^{(0)} \supseteq L^{(1)} \supseteq L^{(2)} \supseteq \cdots$ is called the *derived* series of $L$. The Lie algebra $L$ is said to be *solvable* if there is an $n \in \mathbb{N}$ such that $L^{(n)} = 0$.

**1.6 Exercise.** Prove the following properties:

1. Any nilpotent Lie algebra is solvable. However, show that $L = kx + ky$, with $[x, y] = y$, is a solvable but not nilpotent Lie algebra.
2. If $L$ is nilpotent or solvable, so are its subalgebras and quotients.
3. If $I$ and $J$ are nilpotent (or solvable) ideals of $L$, so is $I + J$.
4. Let $I$ be an ideal of $L$ such that both $I$ and $L/I$ are solvable. Then $L$ is solvable. Give an example to show that this is no longer valid with nilpotent instead of solvable.

As a consequence of these properties, the sum of all the nilpotent (respectively solvable) ideals of $L$ is the largest nilpotent (resp. solvable) ideal of $L$. This ideal is denoted by $N(L)$ (resp. $R(L)$) and called the *nilpotent radical* (resp. *solvable radical*) of $L$.

**1.7 Lie's Theorem.** *Let $\rho : L \to \mathfrak{gl}(V)$ be a representation of a solvable Lie algebra $L$ over an algebraically closed field $k$ of characteristic $0$. Then there is a nonzero element $0 \neq v \in V$ such that $x.v \in kv$ for any $x \in L$ (that is, $v$ is a common eigenvector for all the endomorphisms $\rho(x)$, $x \in L$).*

*Proof.* Since $L$ is solvable, $[L, L] \subsetneqq L$ and we may take a codimension 1 subspace of $L$ with $[L, L] \subseteq S$. Then clearly $S$ is an ideal of $L$. Take $z \in L \setminus S$, so $L = S \oplus kz$.

Arguing inductively, we may assume that there is a nonzero common eigenvector $v$ of $\rho(x)$ for any $x \in S$ and, thus, there is a linear form $\lambda : S \to k$, such that $x.v = \lambda(x)v$ for any $x \in S$. Let $W = \{w \in V : x.w = \lambda(x)w \ \forall x \in S\}$. $W$ is a nonzero subspace of $V$. Let $U$ be the linear span of $\{v, z.v, z.(z.v), \ldots\}$, with $v$ as above. The subspace $U$ is invariant under $\rho(z)$, and for any $x \in S$ and $m \in \mathbb{N}$:

$$\rho(x)\rho(z)^m(v) = \rho(x)\rho(z)\rho(z)^{m-1}(v) = \rho([x, z])\rho(z)^{m-1}(v) + \rho(z)\Big(\rho(x)\rho(z)^{m-1}(v)\Big).$$

Now arguing by induction on $m$ we see that

(i) $\rho(x)\rho(z)^m(v) \in U$ for any $m \in \mathbb{N}$, and hence $U$ is a submodule of $V$.

(ii) $\rho(x)\rho(z)^m(v) = \lambda(x)\rho(z)^m(v) + \sum_{i=0}^{m-1} \alpha_i \rho(z)^i(v)$ for suitable scalars $\alpha_i \in k$.

Therefore the action of $\rho(x)$ on $U$ is given by an upper triangular matrix with $\lambda(x)$ on the diagonal and, hence, $\operatorname{trace} \rho(x)|_U = \lambda(x) \dim_k U$ for any $x \in S$. In particular,

$$\operatorname{trace} \rho([x, z])|_U = \begin{cases} \lambda([x, z]) \dim_k U \\ \operatorname{trace}\big[\rho(x)|_U, \rho(z)|_U\big] = 0 \end{cases}$$

(the trace of any commutator is 0), and since the characteristic of $k$ is 0 we conclude that $\lambda([S,L]) = 0$.

But then, for any $0 \neq w \in W$ and $x \in S$,

$$x.(z.w) = [x,z].w + z.(x.w) = \lambda([x,z])w + z.\bigl(\lambda(x)w\bigr) = \lambda(x)z.w,$$

and this shows that $W$ is invariant under $\rho(z)$. Since $k$ is algebraically closed, there is a nonzero eigenvector of $\rho(z)$ in $W$, and this is a common eigenvector for any $x \in S$ and for $z$, and hence for any $y \in L$. □

**1.8 Remark.** Note that the proof above is valid even if $k$ is not algebraically closed, as long as the characteristic polynomial of $\rho(x)$ for any $x \in L$ splits over $k$. In this case $\rho$ is said to be a *split* representation.

**1.9 Consequences.** Assume that the characteristic of the ground field $k$ is 0.

(i) Let $\rho : L \to \mathfrak{gl}(V)$ be an irreducible split representation of a solvable Lie algebra. Then $\dim_k V = 1$.

(ii) Let $\rho : L \to \mathfrak{gl}(V)$ be a split representation of a solvable Lie algebra. Then there is a basis of $V$ such that the coordinate matrix of any $\rho(x)$, $x \in L$, is upper triangular.

(iii) Let $L$ be a solvable Lie algebra such that its adjoint representation $\mathrm{ad} : L \to \mathfrak{gl}(L)$ is split. Then there is a chain of ideal $0 = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_n = L$ with $\dim L_i = i$ for any $i$.

(iv) Let $\rho : L \to \mathfrak{gl}(V)$ be a representation of a Lie algebra $L$. Then $[L, R(L)]$ acts nilpotently on $V$; that is, $\rho(x)$ is nilpotent for any $x \in [L, R(L)]$. The same is true of $[L,L] \cap R(L)$. In particular, with the adjoint representation, we conclude that $[L, R(L)] \subseteq [L,L] \cap R(L) \subseteq N(L)$ and, therefore, $L$ is solvable if and only if $[L,L]$ is nilpotent.

*Proof.* Let $\bar{k}$ be an algebraic closure of $k$. Then $\bar{k} \otimes_k L$ is a Lie algebra over $\bar{k}$ and $\bar{k} \otimes_k R(L)$ is solvable, and hence contained in $R(\bar{k} \otimes_k L)$. Then, by "extending scalars" it is enough to prove the result assuming that $k$ is algebraically closed. Also, by taking a composition series of $V$, it suffices to prove the result assuming that $V$ is irreducible. In this situation, as in the proof of Lie's Theorem, one shows that there is a linear form $\lambda : R(L) \to k$ such that $W = \{v \in V : x.v = \lambda(x)v \ \forall x \in R(L)\}$ is a nonzero submodule of $V$. By irreducibility, we conclude that $x.v = \lambda(x)v$ for any $x \in R(L)$ and any $v \in V$. Moreover, for any $x \in [L,L] \cap R(L)$, $0 = \operatorname{trace} \rho(x) = \lambda(x) \dim_k V$, so $\lambda\bigl([L,L] \cap R(L)\bigr) = 0$ holds, and hence $[L, R(L)].V \subseteq \bigl([L,L] \cap R(L)\bigr).V = 0$.

The last part follows immediately from the adjoint representation. Note that if $[L,L]$ is nilpotent, in particular it is solvable, and since $L/[L,L]$ is abelian (and hence solvable), $L$ is solvable by the exercise above. □

We will prove now a criterion for solvability due to Cartan.

Recall that any endomorphism $f \in \mathrm{End}_k(V)$ over an algebraically closed field decomposes in a unique way as $f = s + n$ with $s, n \in \mathrm{End}_k(V)$, $s$ being semisimple (that is, diagonalizable), $n$ nilpotent and $[s,n]=0$ (*Jordan decomposition*). Moreover, $s(V)\subseteq f(V)$, $n(V)\subseteq f(V)$ and any subspace which is invariant under $f$ is invariant too under $s$ and $n$.

**1.10 Lemma.** *Let $V$ be a vector space over a field $k$ of characteristic 0, and let $M_1\subseteq M_2$ be two subspaces of $\mathfrak{gl}(V)$. Let $A=\{x\in\mathfrak{gl}(V):[x,M_2]\subseteq M_1\}$ and let $z\in A$ be an element such that* $\operatorname{trace}(zy)=0$ *for any $y\in A$. Then $z$ is nilpotent.*

*Proof.* We may extend scalars and assume that $k$ is algebraically closed. Let $m=\dim_k V$. Then the characteristic polynomial of $z$ is $(X-\lambda_1)\cdots(X-\lambda_m)$, for $\lambda_1,\dots,\lambda_m\in k$. We must check that $\lambda_1=\cdots=\lambda_m=0$. Consider the $\mathbb{Q}$ subspace of $k$ spanned by the eigenvalues $\lambda_1,\dots,\lambda_m$: $E=\mathbb{Q}\lambda_1+\cdots+\mathbb{Q}\lambda_m$. Assume that $E\neq 0$ and take $0\neq f:E\to\mathbb{Q}$ a $\mathbb{Q}$-linear form. Let $z=s+n$ be the Jordan decomposition and let $\{v_1,\dots,v_m\}$ be an associated basis of $V$, in which the coordinate matrix of $z$ is triangular and $s(v_i)=\lambda_i v_i$ for any $i$. Consider the corresponding basis $\{E_{ij}:1\leq i,j\leq m\}$ of $\mathfrak{gl}(V)$, where $E_{ij}(v_j)=v_i$ and $E_{ij}(v_l)=0$ for any $l\neq j$. Then $[s,E_{ij}]=(\lambda_i-\lambda_j)E_{ij}$, so that $\mathrm{ad}_s$ is semisimple. Also $\mathrm{ad}_n$ is clearly nilpotent, and $\mathrm{ad}_z=\mathrm{ad}_s+\mathrm{ad}_n$ is the Jordan decomposition of $\mathrm{ad}_z$. This implies that $\mathrm{ad}_z|_{M_2}=\mathrm{ad}_s|_{M_2}+\mathrm{ad}_n|_{M_2}$ is the Jordan decomposition of $\mathrm{ad}_z|_{M_2}$ and $[s,M_2]=\mathrm{ad}_s(M_2)\subseteq\mathrm{ad}_z(M_2)\subseteq M_1$.

Consider the element $y\in\mathfrak{gl}(V)$ defined by means of $y(v_i)=f(\lambda_i)v_i$ for any $i$. Then $[y,E_{ij}]=f(\lambda_i-\lambda_j)E_{ij}$. Let $p(T)$ be the interpolation polynomial such that $p(0)=0$ (trivial constant term) and $p(\lambda_i-\lambda_j)=f(\lambda_i-\lambda_j)$ for any $1\leq i\neq j\leq m$. Then $\mathrm{ad}_y=p(\mathrm{ad}_s)$ and hence $[y,M_2]\subseteq M_1$, so $y\in A$. Thus, $0=\operatorname{trace}(zy)=\sum_{i=1}^m\lambda_i f(\lambda_i)$. Apply $f$ to get $0=\sum_{i=1}^m f(\lambda_i)^2$, which forces, since $f(\lambda_i)\in\mathbb{Q}$ for any $i$, that $f(\lambda_i)=0$ for any $i$. Hence $f=0$, a contradiction. □

**1.11 Proposition.** *Let $V$ be a vector space over a field $k$ of characteristic 0 and let $L$ be a Lie subalgebra of $\mathfrak{gl}(V)$. Then $L$ is solvable if and only if* $\operatorname{trace}(xy)=0$ *for any $x\in[L,L]$ and $y\in L$.*

*Proof.* Assume first that $L$ is solvable and take a composition series of $V$ as a module for $L$: $V=V_0\supseteq V_1\supseteq\cdots\supseteq V_m=0$. Engel's Theorem and Consequences 1.9 show that $[L,L].V_i\subseteq V_{i+1}$ for any $i$. This proves that $\operatorname{trace}\bigl([L,L]L\bigr)=0$.

Conversely, assume that $\operatorname{trace}(xy)=0$ for any $x\in[L,L]$ and $y\in L$, and consider the subspace $A=\{x\in\mathfrak{gl}(V):[x,L]\subseteq[L,L]\}$. For any $u,v\in L$ and $y\in A$,

$$\begin{aligned}\operatorname{trace}\bigl([u,v]y\bigr)&=\operatorname{trace}(uvy-vuy)\\&=\operatorname{trace}(vyu-yvu)\\&=\operatorname{trace}\bigl([v,y]u\bigr)=0\quad(\text{since }[v,y]\in[L,L]).\end{aligned}$$

Hence $\operatorname{trace}(xy)=0$ for any $x\in[L,L]$ and $y\in A$ which, by the previous Lemma, shows that $x$ is nilpotent for any $x\in[L,L]$. By Engel's Theorem, $[L,L]$ is nilpotent, and hence $L$ is solvable. □

**1.12 Theorem.** ***(Cartan's criterion for solvability)***
*Let $L$ be a Lie algebra over a field $k$ of characteristic 0. Then $L$ is solvable if and only if* $\operatorname{trace}(\mathrm{ad}_x\,\mathrm{ad}_y)=0$ *for any $x\in[L,L]$ and any $y\in L$.*

*Proof.* The adjoint representation $\mathrm{ad}:L\to\mathfrak{gl}(L)$ satisfies that $\ker\mathrm{ad}=Z(L)$, which is abelian and hence solvable. Thus $L$ is solvable if and only if so is $L/Z(L)\cong\mathrm{ad}\,L$ and the previous Proposition shows that, since $[\mathrm{ad}\,L,\mathrm{ad}\,L]=\mathrm{ad}[L,L]$, that $\mathrm{ad}\,L$ is solvable if and only if $\operatorname{trace}\bigl(\mathrm{ad}_x\,\mathrm{ad}_y\bigr)=0$ for any $x\in[L,L]$ and $y\in L$. □

The bilinear form $\kappa : L \times L \to k$ given by

$$\kappa(x, y) = \operatorname{trace}(\operatorname{ad}_x \operatorname{ad}_y)$$

for any $x, y \in L$, that appears in Cartan's criterion for solvability, plays a key role in studying Lie algebras over fields of characteristic 0. It is called the *Killing form* of the Lie algebra $L$.

Note that $\kappa$ is symmetric and *invariant* (i.e., $\kappa([x, y], z) = \kappa(x, [y, z])$ for any $x, y, z \in L$).

## § 2. Semisimple Lie algebras

A Lie algebra is said to be *semisimple* if its solvable radical is trivial: $R(L) = 0$. It is called *simple* if it has no proper ideal and it is not abelian.

Any simple Lie algebra is semisimple, and given any Lie algebra $L$, the quotient $L/R(L)$ is semisimple.

**2.1 Theorem.** ***(Cartan's criterion for semisimplicity)***
*Let $L$ be a Lie algebra over a field $k$ of characteristic 0 and let $\kappa(x, y) = \operatorname{trace}(\operatorname{ad}_x \operatorname{ad}_y)$ be its Killing form. Then $L$ is semisimple if and only if $\kappa$ is nondegenerate.*

*Proof.* The invariance of the Killing form $\kappa$ of such a Lie algebra $L$ implies that the subspace $I = \{x \in L : \kappa(x, L) = 0\}$ is an ideal of $L$. By Proposition 1.11, $\operatorname{ad} I$ is a solvable subalgebra of $\mathfrak{gl}(L)$, and this shows that $I$ is solvable. ($\operatorname{ad} I \cong I/Z(L) \cap I$).

Hence, if $L$ is semisimple $I \subseteq R(L) = 0$, and thus $\kappa$ is nondegenerate. (Note that this argument is valid had we started with a Lie subalgebra $L$ of $\mathfrak{gl}(V)$ for some vector space $V$, and had we replaced $\kappa$ by the trace form of $V$: $B : L \times L \to k$, $(x, y) \mapsto B(x, y) = \operatorname{trace}(xy)$.)

Conversely, assume that $\kappa$ is nondegenerate, that is, that $I = 0$. If $J$ were an abelian ideal of $L$, then for any $x \in J$ and $y \in L$, $\operatorname{ad}_x \operatorname{ad}_y(L) \subseteq J$ and $\operatorname{ad}_x \operatorname{ad}_y(J) = 0$. Hence $\bigl(\operatorname{ad}_x \operatorname{ad}_y\bigr)^2 = 0$ and $\kappa(x, y) = \operatorname{trace}\bigl(\operatorname{ad}_x \operatorname{ad}_y\bigr) = 0$. Therefore, $J \subseteq I = 0$. Thus, $L$ does not contain proper abelian ideals, so it does not contain proper solvable ideals and, hence, $R(L) = 0$ and $L$ is semisimple. □

**2.2 Consequences.** Let $L$ be a Lie algebra over a field $k$ of characteristic 0.

(i) $L$ is semisimple if and only if $L$ is a direct sum of simple ideals. In particular, this implies that $L = [L, L]$.

*Proof.* If $L = L_1 \oplus \cdots \oplus L_n$ with $L_i$ a simple ideal of $L$ for any $i$, and $J$ is an abelian ideal of $L$, then $[J, L_i]$ is an abelian ideal of $L_i$, and hence it is 0. Hence $[J, L] = 0$. This shows that the projection of $J$ on each $L_i$ is contained in the center $Z(L_i)$, which is 0 by simplicity. Hence $J = 0$.

Conversely, assume that $L$ is semisimple and let $I$ be a minimal ideal of $L$, take the orthogonal $I^\perp = \{x \in L : \kappa(x, I) = 0\}$, which is an ideal of $L$ by invariance of $\kappa$. Cartan's criterion of solvability (or better Proposition 1.11) shows that $I \cap I^\perp$ is solvable and hence, as $R(L) = 0$, $I \cap I^\perp = 0$ and $L = I \oplus I^\perp$. Now, $I$ is simple, since any ideal $J$ of $I$ satisfies $[J, I^\perp] \subseteq [I, I^\perp] \subseteq I \cap I^\perp = 0$, and hence $[J, L] = [J, I] \subseteq J$. Also, $\kappa = \kappa_I \perp \kappa_{I^\perp}$ is the orthogonal sum of the Killing forms of $I$ and $I^\perp$. So we can proceed with $I^\perp$ as we did for $L$ to complete a decomposition of $L$ into the direct sum of simple ideals. □

(ii) Let $K/k$ be a field extension, then $L$ is semisimple if and only if so is the scalar extension $K \otimes_k L$.

*Proof.* Once a basis of $L$ over $k$ is fixed (which is also a basis of $K \otimes_k L$ over $K$ if we identify $L$ with $1 \otimes L$), the coordinate matrices of the Killing forms of $L$ and $K \otimes_k L$ coincide, whence the result. □

(iii) If $L$ is semisimple and $I$ is a proper ideal of $L$, then both $I$ and $L/I$ are semisimple.

*Proof.* As in (i), $L = I \oplus I^\perp$ and the Killing form of $L$ is the orthogonal sum of the Killing forms of these two ideals: $\kappa_I$ and $\kappa_{I^\perp}$. Hence both Killing forms are nondegenerate and, hence, both $I$ and $I^\perp$ are semisimple. Finally, $L/I \cong I^\perp$. □

(iv) Assume that $L$ is a Lie subalgebra of $\mathfrak{gl}(V)$ and that the trace form $B : L \times L \to k$, $(x, y) \mapsto B(x, y) = \operatorname{trace}(xy)$ is nondegenerate. Then $L = Z(L) \oplus [L, L]$ and $[L, L]$ is semisimple (recall that the center $Z(L)$ is abelian). Moreover, the ideals $Z(L)$ and $[L, L]$ are orthogonal relative to $B$, and hence the restriction of $B$ to both $Z(L)$ and $[L, L]$ are nondegenerate.

*Proof.* Let $V = V_0 \supseteq V_1 \supseteq \cdots \supseteq 0$ be a composition series of $V$ as a module for $L$. Then we know, because of Consequences 1.9 that both $[L, R(L)]$ and $[L, L] \cap R(L)$ act nilpotently on $V$. Therefore, $B\big([L, R(L)], L\big) = 0 = B\big([L, L] \cap R(L), L\big)$ and, as $B$ is nondegenerate, this shows that $[L, R(L)] = 0 = [L, L] \cap R(L)$. In particular, $R(L) = Z(L)$ and, since $L/R(L)$ is semisimple, $L/R(L) = [L/R(L), L/R(L)] = \big([L, L] + R(L)\big)/R(L)$. Hence $L = [L, L] + R(L)$ and $[L, L] \cap R(L) = 0$, whence it follows that $L = Z(L) \oplus [L, L]$. Besides, by invariance of $B$, $B\big(Z(L), [L, L]\big) = B\big([Z(L), L], L\big) = 0$ and the last part follows. □

(v) An endomorphism $d$ of a Lie algebra $L$ is said to be a *derivation* if $d([x, y]) = [d(x), y] + [x, d(y)]$ for any $x, y \in L$. For any $x \in L$, $\operatorname{ad}_x$ is a derivation, called *inner derivation*. Then, if $L$ is semisimple, any derivation is inner.

*Proof.* Let $d$ be any derivation and consider the linear form $L \to k$, $x \mapsto \operatorname{trace}(d \operatorname{ad}_x)$. Since $\kappa$ is nondegenerate, there is a $z \in L$ such that $\kappa(z, x) = \operatorname{trace}(d \operatorname{ad}_x)$ for any $x \in L$. But then, for any $x, y \in L$,

$$
\begin{aligned}
\kappa\big(d(x), y\big) &= \operatorname{trace}\big(\operatorname{ad}_{d(x)} \operatorname{ad}_y\big) \\
&= \operatorname{trace}\big([d, \operatorname{ad}_x] \operatorname{ad}_y\big) \quad \text{(since $d$ is a derivation)} \\
&= \operatorname{trace}\big(d[\operatorname{ad}_x, \operatorname{ad}_y]\big) \\
&= \operatorname{trace}\big(d \operatorname{ad}_{[x,y]}\big) \\
&= \kappa(z, [x, y]) = \kappa([z, x], y).
\end{aligned}
$$

Hence, by nondegeneracy, $d(x) = [z, x]$ for any $x$, so $d = \operatorname{ad}_z$. □

Let $V$ and $W$ be two modules for a Lie algebra $L$. Then both $\mathrm{Hom}_k(V,W)$ and $V \otimes_k W$ are $L$-modules too by means of:

$$(x.f)(v) = x.(f(v)) - f(x.v),$$
$$x.(v \otimes w) = (x.v) \otimes w + v \otimes (x.w),$$

for any $x \in L$, $f \in \mathrm{Hom}_k(V,W)$ and $v \in V$, $w \in W$. (In particular, the dual $V^*$ is a module with $(x.f)(v) = -f(x.v)$ for $x \in L$, $f \in V^*$ and $v \in V$.)

**2.3 Proposition.** *Let $L$ be a Lie algebra over an algebraically closed field $k$ of characteristic 0. Then any irreducible module for $L$ is, up to isomorphism, of the form $V = V_0 \otimes_k Z$, with $V_0$ and $Z$ modules such that $\dim_k Z = 1$ and $V_0$ is irreducible and annihilated by $R(L)$. (Hence, $V_0$ is a module for the semisimple Lie algebra $L/R(L)$.)*

*Proof.* By the proof of Consequence 1.9.(iv), we know that there is a linear form $\lambda : R(L) \to k$ such that $x.v = \lambda(x)v$ for any $x \in R(L)$ and $v \in V$. Moreover, $\lambda\big([L, R(L)]\big) = 0 = \lambda\big([L,L] \cap R(L)\big)$. Thus we may extend $\lambda$ to a form $L \to k$, also denoted by $\lambda$, in such a way that $\lambda\big([L,L]\big) = 0$.

Let $Z = kz$ be a one dimensional vector space, which is a module for $L$ by means of $x.z = \lambda(x)z$ and let $W = V \otimes_k Z^*$ ($Z^*$ is the dual vector space to $Z$), which is also an $L$-module. Then the linear map

$$W \otimes_k Z \longrightarrow V$$
$$(v \otimes f) \otimes z \mapsto f(z)v$$

is easily seen to be an isomorphism of modules. Moreover, since $V$ is irreducible, so is $W$, and for any $x \in R(L)$, $v \in V$ and $f \in Z^*$, $x.(v \otimes f) = (x.v) \otimes f + v \otimes (x.f) = \lambda(x)v \otimes f - \lambda(x)v \otimes f = 0$ (since $(x.f)(z) = -f(x.z) = -\lambda(x)f(z)$). Hence $W$ is annihilated by $R(L)$. □

This Proposition shows the importance of studying the representations of the semisimple Lie algebras.

Recall the following definition.

**2.4 Definition.** A module is said to be *completely reducible* if and only if it is a direct sum of irreducible modules or, equivalently, if any submodule has a complementary submodule.

**2.5 Weyl's Theorem.** *Any representation of a semisimple Lie algebra over a field of characteristic 0 is completely reducible.*

*Proof.* Let $L$ be a semisimple Lie algebra over the field $k$ of characteristic 0, and let $\rho : L \to \mathfrak{gl}(V)$ be a representation and $W$ a submodule of $V$. Does there exist a submodule $W'$ such that $V = W \oplus W'$?

We may extend scalars and assume that $k$ is algebraically closed, because the existence of $W'$ is equivalent to the existence of a solution to a system of linear equations: does there exist $\pi \in \mathrm{End}_L(V)$ such that $\pi(V) = W$ and $\pi|_W = I_W$ (the identity map on $W$)?

Now, assume first that $W$ is irreducible and $V/W$ trivial (that is, $L.V \subseteq W$). Then we may change $L$ by its quotient $\rho(L)$, which is semisimple too (or 0, which is a trivial case), and hence assume that $0 \neq L \leq \mathfrak{gl}(V)$. Consider the trace form $b_V : L \times L \to k$, $(x, y) \mapsto \text{trace}(xy)$. By Cartan's criterion for solvability, $\ker b_V$ is a solvable ideal of $L$, hence 0, and thus $b_V$ is nondegenerate. Take dual bases $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ of $L$ relative to $b_V$ (that is, $b_V(x_i, y_j) = \delta_{ij}$ for any $i, j$).

Then the element $c_V = \sum_{i=1}^n x_i y_i \in \text{End}_k(V)$ is called the *Casimir element* and

$$\text{trace}(c_V) = \sum_{i=1}^n \text{trace}(x_i y_i) = \sum_{i=1}^n b_V(x_i, y_i) = n = \dim_k L.$$

Moreover, for any $x \in L$, there are scalars such that $[x_i, x] = \sum_{j=1}^n \alpha_{ij} x_j$ and $[y_i, x] = \sum_{i=1}^n \beta_{ij} y_j$ for any $i$. Since

$$b_V\big([x_i, x], y_j\big) + b_V\big(x_i, [y_j, x]\big) = 0$$

for any $i, j$, it follows that $\alpha_{ij} + \beta_{ji} = 0$ for any $i, j$, so

$$[c_V, x] = \sum_{i=1}^n \Big([x_i, x]y_i + x_i[y_i, x]\Big) = \sum_{i,j=1}^n (\alpha_{ji} + \beta_{ij}) x_i y_j = 0.$$

We then have that $c_V(V) \subseteq W$ and, by Schur's Lemma ($W$ is assumed here to be irreducible), $c_V|_W \in \text{End}_L(W) = kI_W$. Besides, $\text{trace}(c_V) = \dim_k L$. Therefore,

$$c_V|_W = \frac{\dim_k L}{\dim_k W} I_W$$

and $V = \ker c_V \oplus \text{im}\, c_V = \ker c_V \oplus W$. Hence $W' = \ker c_V$ is a submodule that complements $W$.

Let us show now that the result holds as long as $L.V \subseteq W$.

To do so, we argue by induction on $\dim_k W$, the result being trivial if $\dim_k W = 0$. If $W$ is irreducible, the result holds by the previous arguments. Otherwise, take a maximal submodule $Z$ of $W$. By the induction hypothesis, there is a submodule $\tilde{V}$ such that $V/Z = W/Z \oplus \tilde{V}/Z$, and hence $V = W + \tilde{V}$ and $W \cap \tilde{V} = Z$. Now, $L.\tilde{V} \subseteq \tilde{V} \cap W = Z$ and $\dim_k Z < \dim_k W$, so there exists a submodule $W'$ of $\tilde{V}$ such that $\tilde{V} = Z \oplus W'$. Hence $V = W + W'$ and $W \cap W' \subseteq W \cap \tilde{V} \cap W' = Z \cap W' = 0$, as required.

In general, consider the following submodules of the $L$-module $\text{Hom}_k(V, W)$:

$$M = \{f \in \text{Hom}_k(V, W) : \text{there exists } \lambda_f \in k \text{ such that } f|_W = \lambda_f id\},$$
$$N = \{f \in \text{Hom}_k(V, W) : f|_W = 0\}.$$

For any $x \in L$, $f \in M$, and $w \in W$:

$$(x.f)(w) = x.\big(f(w)\big) - f(x.w) = x.(\lambda_f w) - \lambda_f(x.w) = 0,$$

so $L.M \subseteq N$. Then there exists a submodule $X$ of $\text{Hom}_k(V, W)$ such that $M = N \oplus X$. Since $L.X \subseteq X \cap N = 0$, $X$ is contained in $\text{Hom}_L(V, W)$. Take $f \in X$ with $\lambda_f = 1$, so $f(V) \subseteq W$ and $f|_W = id$. Then $W = \ker f \oplus W$, and $\ker f$ is a submodule of $V$ that complements $W$. $\square$

**2.6 Consequences on Jordan decompositions.** Let $k$ be an algebraically closed field of characteristic 0.

(i) Let $V$ be a vector space over $k$ and let $L$ be a semisimple Lie subalgebra of $\mathfrak{gl}(V)$. For any $x \in L$, consider its Jordan decomposition $x = x_s + x_n$. Then $x_s, x_n \in L$.

*Proof.* We know that $\operatorname{ad} x_s$ is semisimple, $\operatorname{ad} x_n$ nilpotent, and that $\operatorname{ad} x = \operatorname{ad} x_s + \operatorname{ad} x_n$ is the Jordan decomposition of $\operatorname{ad} x$. Let $W$ be any irreducible submodule of $V$ and consider the Lie subalgebra of $\mathfrak{gl}(V)$:

$$L_W = \{z \in \mathfrak{gl}(V) : z(W) \subseteq W \text{ and } \operatorname{trace}(z|_W) = 0\}.$$

Since $L = [L, L]$, $\operatorname{trace}(x|_W) = 0$ for any $x \in L$. Hence $L \subseteq L_W$. Moreover, for any $x \in L$, $x(W) \subseteq W$, so $x_s(W) \subseteq W$, $x_n(W) \subseteq W$ and $x_s, x_n \in L_W$.

Consider also the Lie subalgebra of $\mathfrak{gl}(V)$:

$$N = \{z \in \mathfrak{gl}(V) : [z, L] \subseteq L\} = \{z \in \mathfrak{gl}(V) : \operatorname{ad} z(L) \subseteq L\}.$$

Again, for any $x \in L$, $\operatorname{ad} x(L) \subseteq L$, so $\operatorname{ad} x_s(L) \subseteq L$, $\operatorname{ad} x_n(L) \subseteq L$, and $x_s, x_n \in N$.

Therefore, it is enough to prove that $L = \left(\cap_W L_W\right) \cap N$. If we denote by $\tilde{L}$ the subalgebra $\left(\cap_W L_W\right) \cap N$, then $L$ is an ideal of $\tilde{L}$.

By Weyl's Theorem, there is a subspace $U$ of $\tilde{L}$ such that $\tilde{L} = L \oplus U$ and $[L, U] \subseteq U$. But $[L, U] \subseteq [L, N] \subseteq L$, so $[L, U] = 0$. Then, for any $z \in U$ and irreducible submodule $W$ of $V$, $z|_W \in \operatorname{Hom}_L(W, W) = kI_W$ (by Schur's Lemma) and $\operatorname{trace}(z|_W) = 0$, since $z \in L_W$. Therefore $z|_W = 0$. But Weyl's Theorem asserts that $V$ is a direct sum of irreducible submodules, so $z = 0$. Hence $U = 0$ and $L = \tilde{L}$. $\square$

(ii) Let $L$ be a semisimple Lie algebra. Then $L \cong \operatorname{ad} L$, which is a semisimple subalgebra of $\mathfrak{gl}(L)$. For any $x \in L$, let $\operatorname{ad} x = s + n$ be the Jordan decomposition in $\operatorname{End}_k(L) = \mathfrak{gl}(L)$. By item (i), there are unique elements $x_s, x_n \in L$ such that $s = \operatorname{ad} x_s$, $n = \operatorname{ad} x_n$. Since ad is one-to-one, $x = x_s + x_n$. This is called the *absolute Jordan decomposition* of $x$.

Note that $[x, x_s] = 0 = [x, x_n]$, since $[\operatorname{ad} x, \operatorname{ad} x_s] = 0 = [\operatorname{ad} x, \operatorname{ad} x_n]$.

(iii) Let $L$ be a semisimple Lie algebra and let $\rho : L \to \mathfrak{gl}(V)$ be a representation. Let $x \in L$ and let $x = x_s + x_n$ be its absolute Jordan decomposition. Then $\rho(x) = \rho(x_s) + \rho(x_n)$ is the Jordan decomposition of $\rho(x)$.

*Proof.* Since $\rho(L) \cong L/\ker\rho$ is a quotient of $L$, $\rho(x_s) = \rho(x)_s$ and $\rho(x_n) = \rho(x)_n$ (this is because $\operatorname{ad}_{\rho(L)} \rho(x_s)$ is semisimple and $\operatorname{ad}_{\rho(L)} \rho(x_n)$ is nilpotent). Here $\operatorname{ad}_{\rho(L)}$ denotes the adjoint map in the Lie algebra $\rho(L)$, to distinguish it from the adjoint map of $\mathfrak{gl}(V)$. By item (i), if $\rho(x) = s + n$ is the Jordan decomposition of $\rho(x)$, $s, n \in \rho(L)$ and we obtain two Jordan decompositions in $\mathfrak{gl}\big(\rho(L)\big)$: $\operatorname{ad}_{\rho(L)} \rho(x) = \operatorname{ad}_{\rho(L)} s + \operatorname{ad}_{\rho(L)} n = \operatorname{ad}_{\rho(L)} \rho(x_s) + \operatorname{ad}_{\rho(L)} \rho(x_n)$. By uniqueness, $s = \rho(x_s)$ and $n = \rho(x_n)$. $\square$

There are other important consequences that can be drawn from Weyl's Theorem:

**2.7 More consequences.**

(i) **(Whitehead's Lemma)** Let $L$ be a semisimple Lie algebra over a field $k$ of characteristic 0, let $V$ be a module for $L$, and let $\varphi : L \to V$ be a linear map such that

$$\varphi\bigl([x, y]\bigr) = x.\varphi(y) - y.\varphi(x),$$

for any $x, y \in L$. Then there is an element $v \in V$ such that $\varphi(x) = x.v$ for any $x \in L$.

*Proof.* $\varphi$ belongs to the $L$-module $\mathrm{Hom}_k(L, V)$, and for any $x, y \in L$:

$$(x.\varphi)(y) = x.\varphi(y) - \varphi([x, y]) = y.\varphi(x) = \mu_{\varphi(x)}(y), \tag{2.1}$$

where $\mu_v(x) = x.v$ for any $x \in L$ and $v \in V$. Moreover, for any $x, y \in L$ and $v \in V$,

$$(x.\mu_v)(y) = x.\bigl(\mu_v(y)\bigr) - \mu_v([x, y]) = x.(y.v) - [x, y].v = y.(x.v) = \mu_{x.v}(y).$$

Thus, $\mu_V$ is a submodule of $\mathrm{Hom}_k(L, V)$, which is contained in $W = \{f \in \mathrm{Hom}_k(L, V) : x.f = \mu_{f(x)}\ \forall x \in L\}$, and this satisfies $L.W \subseteq \mu_V$. By Weyl's Theorem there is another submodule $\tilde{W}$ such that $W = \mu_V \oplus \tilde{W}$ and $L.\tilde{W} \subseteq \tilde{W} \cap \mu_V = 0$. But for any $f \in \tilde{W}$ and $x, y \in L$, (2.1) gives

$$\begin{aligned} 0 = (x.f)(y) &= x.f(y) - f([x, y]) = \mu_{f(y)}(x) - f([x, y]) \\ &= (y.f)(x) - f([x, y]) = -f([x, y]). \end{aligned}$$

Therefore, $f(L) = f([L, L]) = 0$. Hence $\tilde{W} = 0$ and $\varphi \in W = \mu_V$, as required. □

(ii) **(Levi-Malcev Theorem)** Let $L$ be a Lie algebra over a field $k$ of characteristic 0, then there exists a subalgebra $S$ of $L$ such that $L = R(L) \oplus S$. If nontrivial, $S$ is semisimple. Moreover, if $T$ is any semisimple subalgebra of $L$, then there is an automorphism $f$ of $L$, in the group of automorphisms generated by $\{\exp \mathrm{ad}_z : z \in N(L)\}$, such that $f(T) \subseteq S$.

*Proof.* In case $S$ is a nontrivial subalgebra of $L$ with $L = R(L) \oplus S$, then $S \cong L/R(L)$ is semisimple.

Let us prove the existence result by induction on $\dim L$, being trivial if $\dim L = 1$ (as $L = R(L)$ in this case). If $I$ is an ideal of $L$ with $0 \subsetneqq I \subsetneqq R(L)$, then by the induction hypothesis, there exists a subalgebra $T$ of $L$, containing $I$, with $L/I = R(L)/I \oplus T/I$. Then $T/I$ is semisimple, so $I = R(T)$ and, by the induction hypothesis again, $T = I \oplus S$ for a subalgebra $S$ of $L$. It follows that $L = R(L) \oplus S$, as required. Therefore, it can be assumed that $R(L)$ is a minimal nonzero ideal of $L$, and hence $[R(L), R(L)] = 0$ and $[L, R(L)]$ is either 0 or $R(L)$.

In case $[L, R(L)] = 0$, $L$ is a module for the semisimple Lie algebra $L/R(L)$, so Weyl's Theorem shows that $L = R(L) \oplus S$ for an ideal $S$.

Otherwise, $[L, R(L)] = R(L)$. Consider then the module $\mathfrak{gl}(L)$ for $L$ ($x.f = [\mathrm{ad}_x, f]$ for any $x \in L$ and $f \in \mathfrak{gl}(L)$). Let $\rho$ be the associated representation. Then the subspaces

$$M = \{f \in \mathfrak{gl}(L) : f(L) \subseteq R(L) \text{ and there exists } \lambda_f \in k \text{ such that } f|_{R(L)} = \lambda_f id\},$$
$$N = \{f \in \mathfrak{gl}(L) : f(L) \subseteq R(L) \text{ and } f\big(R(L)\big) = 0\},$$

are submodules of $\mathfrak{gl}(L)$, with $\rho(L)(M) \subseteq N \subsetneqq M$. Moreover, for any $x \in R(L)$, $f \in M$ and $z \in L$:

$$[\mathrm{ad}_x, f](z) = [x, f(z)] - f\big([x, z]\big) = -\lambda_f \,\mathrm{ad}_x(z), \tag{2.2}$$

since $[x, f(z)] \in [R(L), R(L)] = 0$. Hence, $\rho\big(R(L)\big)(M) \subseteq \{\mathrm{ad}_x : x \in R(L)\} \subseteq N$. Write $R = \{\mathrm{ad}_x : x \in R(L)\}$. Therefore, $M/R$ is a module for the semisimple Lie algebra $L/R(L)$ and, by Weyl's Theorem, there is another submodule $\tilde{N}$ with $R \subsetneqq \tilde{N} \subseteq M$ such that $M/R = N/R \oplus \tilde{N}/R$. Take $g \in \tilde{N} \setminus N$ with $\lambda_g = -1$. Since $\rho(L)(M) \subseteq N$, $\rho(L)(g) \subseteq R$, so for any $y \in L$, there is an element $\alpha(y) \in R(L)$ such that

$$[\mathrm{ad}_y, g] = \mathrm{ad}_{\alpha(y)},$$

and $\alpha : L \to R(L)$ is linear. Equation (2.2) shows that $\alpha|_{R(L)} = id$, so that $L = R(L) \oplus \ker\alpha$ and $\ker\alpha = \{x \in L : \rho(x)(g) = 0\}$ is a subalgebra of $L$.

Moreover, if $T$ is a semisimple subalgebra of $L$, let us prove that there is a suitable automorphism of $L$ that embeds $T$ into $S$. Since $T$ is semisimple, $T = [T, T] \subseteq [L, L] = [L, R(L)] \oplus S \subseteq N(L) \oplus S$. If $N(L) = 0$, the result is clear. Otherwise, let $I$ be a minimal ideal of $L$ contained in $N(L)$ (hence $I$ is abelian). Arguing by induction on $\dim L$, we may assume that there are elements $z_1, \ldots, z_r$ in $N(L)$ such that

$$T' = \exp \mathrm{ad}_{z_1} \cdots \exp \mathrm{ad}_{z_r}(T) \subseteq I \oplus S.$$

Now, it is enough to prove that there is an element $z \in I$ such that $\exp \mathrm{ad}_z(T') \subseteq S$. Therefore, it is enough to prove the result assuming that $L = R \oplus S$, where $R$ is an abelian ideal of $L$. In this case, let $\varphi : T \to R$ and $\psi : T \to S$ be the projections of $T$ on $R$ and $S$ respectively (that is, for any $t \in T$, $t = \varphi(t) + \psi(t)$). For any $t_1, t_2 \in T$,

$$\begin{aligned}
[t_1, t_2] &= [\varphi(t_1) + \psi(t_1), \varphi(t_2) + \psi(t_2)] \\
&= [\varphi(t_1), t_2] + [t_1, \varphi(t_2)] + [\psi(t_1), \psi(t_2)],
\end{aligned}$$

since $[R, R] = 0$. Hence $\varphi([t_1, t_2]) = [\varphi(t_1), t_2] + [t_1, \varphi(t_2)]$. Withehead's Lemma shows the existence of an element $z \in R$ such that $\varphi(t) = [t, z]$ for any $t \in T$. But then, since $(\mathrm{ad}_z)^2 = 0$ because $R$ is abelian,

$$\exp \mathrm{ad}_z(t) = t + [z, t] = t - \varphi(t) = \psi(t) \in S,$$

for any $t \in T$. Therefore, $\exp \mathrm{ad}_z(T) \subseteq S$. □

(iii) Let $L$ be a Lie algebra over a field $k$ of characteristic 0, then $[L, R(L)] = [L, L] \cap R(L)$.

*Proof.* $L = R(L) \oplus S$ for a semisimple (if nonzero) subalgebra $S$, so $[L, L] = [L, R(L)] \oplus [S, S] = [L, R(L)] \oplus S$, and $[L, L] \cap R(L) = [L, R(L)]$. □

# § 3. Representations of $\mathfrak{sl}_2(k)$

Among the simple Lie algebras, the Lie algebra $\mathfrak{sl}_2(k)$ of two by two trace zero matrices plays a distinguished role. In this section we will study its representations over fields of characteristic 0.

First note that $\mathfrak{sl}_2(k) = kh + kx + ky$ with $h = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $x = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ and $y = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$, and that

$$[h, x] = 2x, \quad [h, y] = -2y, \quad [x, y] = h.$$

If the characteristic of the ground field $k$ is $\neq 2$, then $\mathfrak{sl}_2(k)$ is a simple Lie algebra.

Let $V(n)$ be the vector space spanned by the homogeneous degree $n$ polynomials in two indeterminates $X$ and $Y$, and consider the representation given by:

$$\begin{aligned} \rho_n : \mathfrak{sl}_2(k) &\longrightarrow \mathfrak{gl}\big(V(n)\big) \\ h &\mapsto X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y}, \\ x &\mapsto X\frac{\partial}{\partial Y}, \\ y &\mapsto Y\frac{\partial}{\partial X}. \end{aligned}$$

**3.1 Exercise.** Check that this indeed gives a representation of $\mathfrak{sl}_2(k)$.

**3.2 Theorem.** *Let $k$ be a field of characteristic* 0*. Then the irreducible representations of $\mathfrak{sl}_2(k)$ are, up to isomorphism, exactly the $\rho_n$, $n \geq 0$.*

*Proof.* Let us assume first that $k$ is algebraically closed, and let $\rho : \mathfrak{sl}_2(k) \to \mathfrak{gl}(V)$ be an irreducible representation.

Since $\operatorname{ad} x$ is nilpotent, the consequences of Weyl's Theorem assert that $\rho(x)$ is nilpotent too (similarly, $\rho(y)$ is nilpotent and $\rho(h)$ semisimple). Hence $W = \{w \in V : x.w = 0\} \neq 0$. For any $w \in W$,

$$x.(h.w) = [x, h].w + h.(x.w) = -2x.w + h.(x.w) = 0,$$

so $W$ is $h$-invariant and, since $\rho(h)$ is semisimple, there is a nonzero $v \in W$ such that $h.v = \lambda v$ for some $\lambda \in k$.

But $\rho(y)$ is nilpotent, so there is an $n \in \mathbb{Z}_{\geq 0}$ such that $v, \rho(y)(v), \ldots, \rho(y)^n(v) \neq 0$ but $\rho(y)^{n+1}(v) = 0$. Now, for any $i > 0$,

$$\begin{aligned} \rho(h)\rho(y)^i(v) &= \rho\big([h, y]\big)\rho(y)^{i-1}(v) + \rho(y)\rho(h)\rho(y)^{i-1}(v) \\ &= -2\rho(y)^i(v) + \rho(y)\Big(\rho(h)\rho(y)^{i-1}(v)\Big) \end{aligned}$$

which shows, recursively, that

$$h.\Big(\rho(y)^i(v)\Big) = (\lambda - 2i)\rho(y)^i(v),$$

and

$$\begin{aligned} \rho(x)\rho(y)^i(v) &= \rho\big([x, y]\big)\rho(y)^{i-1}(v) + \rho(y)\rho(x)\rho(y)^{i-1}(v) \\ &= \big(\lambda - 2(i-1)\big)\rho(y)^{i-1}(v) + \rho(y)\Big(\rho(x)\rho(y)^{i-1}(v)\Big) \end{aligned}$$

which proves that
$$x.\Big(\rho(y)^i(v)\Big) = i\big(\lambda - (i-1)\big)\rho(y)^{i-1}(v).$$
Therefore, with $v_0 = v$ and $v_i = \rho(y)^i(v)$, for $i > 0$, we have
$$\begin{aligned} h.v_i &= (\lambda - 2i)v_i,\\ y.v_i &= v_{i+1}, \quad (v_{n+1} = 0),\\ x.v_i &= i\big(\lambda - (i-1)\big)v_{i-1}, \quad (v_{-1} = 0). \end{aligned}$$
Hence, $\oplus_{i=0}^n kv_i$ is a submodule of $V$ and, since $V$ is irreducible, we conclude that $V = \oplus_{i=0}^n kv_i$. Besides,
$$0 = \operatorname{trace}\rho(h) = \lambda + (\lambda - 2) + \cdots + (\lambda - 2n) = (n+1)\lambda - (n+1)n.$$
So $\lambda = n$. The conclusion is that there is a unique irreducible module $V$ of dimension $n+1$, which contains a basis $\{v_0, \dots, v_n\}$ with action given by
$$h.v_i = (n-2i)v_i, \quad y.v_i = v_{i+1}, \quad x.v_i = i(n+1-i)v_{i-1}$$
(where $v_{n+1} = v_{-1} = 0$.) Then, a fortiori, $V$ is isomorphic to $V(n)$. (One can check that the assignment $v_0 \mapsto X^n$, $v_i \mapsto n(n-1)\cdots(n-i+1)X^{n-i}Y^i$ gives an isomorphism.)

Finally, assume now that $k$ is not algebraically closed and that $\bar{k}$ is an algebraic closure of $k$. If $V$ is an $\mathfrak{sl}_2(k)$-module, then $\bar{k} \otimes_k V$ is an $\mathfrak{sl}_2(\bar{k})$-module which, by Weyl's Theorem, is completely reducible. Then the previous arguments show that the eigenvalues of $\rho(h)$ are integers (and hence belong to $k$). Now the same arguments above apply, since the algebraic closure was only used to insure the existence of eigenvalues of $\rho(h)$ on the ground field. □

**3.3 Remark.** Actually, the result above can be proven easily without using Weyl's Theorem. For $k$ algebraically closed of characteristic 0, let $0 \neq v \in V$ be an eigenvector for $\rho(h)$: $h.v = \lambda v$. Then, with the same arguments as before, $h.\rho(x)^n(v) = (\lambda + 2n)\rho(x)^n v$ and, since the dimension is finite and the characteristic 0, there is a natural number $n$ such that $\rho(x)^n(v) = 0$. This shows that $W = \{w \in V : x.w = 0\} \neq 0$. In the same vein, for any $w \in W$ there is a natural number $m$ such that $\rho(y)^m(w) = 0$. This is all we need for the proof above.

**3.4 Corollary.** *Let $k$ be a field of characteristic 0 and let $\rho : \mathfrak{sl}_2(k) \to \mathfrak{gl}(V)$ be a representation. Consider the eigenspaces $V_0 = \{v \in V : h.v = 0\}$ and $V_1 = \{v \in V : h.v = v\}$. Then $V$ is a direct sum of $\dim_k V_0 + \dim_k V_1$ irreducible modules.*

*Proof.* By Weyl's Theorem, $V = \oplus_{i=1}^N W^i$, with $W^i$ irreducible for any $i$. Now, for any $i$, there is an $n_i \in \mathbb{Z}_{\geq 0}$ such that $W^i \cong V(n_i)$, and hence $\rho(h)$ has eigenvalues $n_i, n_i - 2, \dots, -n_i$, all with multiplicity 1, on $W^i$. Hence $\dim_k W_0^i + \dim_k W_1^i = 1$ for any $i$, where $W_0^i = W^i \cap V_0$, $W_1^i = W^i \cap V_1$. Since $V_0 = \oplus_{i=1}^N W_0^i$ and $V_1 = \oplus_{i=1}^N W_1^i$, the result follows. □

Actually, the eigenvalues of $\rho(h)$ determine completely, up to isomorphism, the representation, because the number of copies of $V(n)$ that appear in the module $V$ in the Corollary above is exactly $\dim_k V_n - \dim_k V_{n+2}$, where $V_n = \{v \in V : h.v = nv\}$ for any $n$; because $n$ appears as eigenvalue in $V(n)$ and in $V(n+2m)$ $(m \geq 1)$ with multiplicity 1, but $n+2$ is also an eigenvalue of $\rho(h)$ in $V(n+2m)$, again with multiplicity 1.

**3.5 Corollary.** ***(Clebsch-Gordan formula)***
*Let $n, m \in \mathbb{Z}_{\geq 0}$, with $n \geq m$, and let $k$ be a field of characteristic $0$. Then, as modules for $\mathfrak{sl}_2(k)$,*

$$V(n) \otimes_k V(m) \cong V(n+m) \oplus V(n+m-2) \oplus \cdots \oplus V(n-m).$$

*Proof.* The eigenvalues of the action of $h$ on $V(n) \otimes_k V(m)$ are $n - 2i + m - 2j = (n+m) - 2(i+j)$, $(0 \leq i \leq n,\ 0 \leq j \leq m)$. Therefore, for any $0 \leq p \leq n+m$,

$$\dim_k V_{n+m-2p} = \Big| \{(i,j) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} : 0 \leq i \leq n,\ 0 \leq j \leq m,\ i+j = p\} \Big|$$

and $\dim_k V_{n+m-2p} - \dim_k V_{n+m-2(p-1)} = 1$ for any $p = 1, \ldots, m$, while $\dim_k V_{n+m-2p} - \dim_k V_{n+m-2(p-1)} = 0$ for $p = m+1, \ldots, \left[\frac{n+m}{2}\right]$. □

## § 4. Cartan subalgebras

In the previous section, we have seen the importance of the subalgebra $kh$ of $\mathfrak{sl}_2(k)$. We look for similar subalgebras in any semisimple Lie algebra.

**4.1 Definition.** Let $L$ be a Lie algebra over a field $k$. A subalgebra $H$ of $L$ is said to be a *Cartan subalgebra* of $L$ if it is nilpotent and self normalizing ($N_L(H) = H$, where for any subalgebra $S$ of $L$, $N_L(S) = \{y \in L : [y, S] \subseteq S\}$ is the *normalizer* of $S$ in $L$).

**4.2 Example.** $kh$ is a Cartan subalgebra of $\mathfrak{sl}_2(k)$ if the characteristic of $k$ is $\neq 2$.

**4.3 Definition.** Let $L$ be a semisimple Lie algebra over a field $k$ of characteristic $0$. For any $x \in L$, let $x = x_s + x_n$ be its absolute Jordan decomposition in $\bar{k} \otimes_k L$, with $\bar{k}$ an algebraic closure of $k$. The element $x$ will be said to be *semisimple* (respectively, *nilpotent*) if $x = x_s$ (resp., if $x = x_n$); that is, if $\operatorname{ad} x \in \mathfrak{gl}(L)$ is semisimple (resp., nilpotent).

A subalgebra $T$ of $L$ is said to be *toral* if all its elements are semisimple.

**4.4 Lemma.** (i) *Let $f, g$ be two endomorphisms of a nonzero vector space $V$. Let $\mu \in k$ be an eigenvalue of $f$, and let $W = \{v \in V : (f - \mu I)^n(v) = 0$ for some $n\}$ be the corresponding generalized eigenspace. (I denotes the identity map.) If there exists a natural number $m > 0$ such that $(\operatorname{ad} f)^m(g) = 0$, then $W$ is invariant under $g$.*

(ii) *Let $\rho : L \to \mathfrak{gl}(V)$ be a representation of a nilpotent Lie algebra $L$ over an algebraically closed field $k$ of characteristic $0$. Then there exists a finite subset $\Lambda$ of the dual vector space $L^*$ such that $V = \oplus_{\lambda \in \Lambda} V_\lambda$, where $V_\lambda = \{v \in V : (\rho(x) - \lambda(x)I)^n(v) = 0$ for some $n$ and for any $x \in L\}$.*

(iii) *Any toral subalgebra of a semisimple Lie algebra over an algebraically closed field $k$ of characteristic $0$ is abelian.*

*Proof.* For (i) denote by $l_f$ and $r_f$ the left and right multiplication by $f$ in $\mathrm{End}_k(V)$. Then, for any $n > 0$, $f^n g = l_f^n(g) = (\operatorname{ad} f + r_f)^n(g) = \sum_{i=0}^n \binom{n}{i} (\operatorname{ad} f)^i(g) f^{n-i}$, and hence, since $\operatorname{ad}(f - \mu I) = \operatorname{ad} f$, we obtain also $(f - \mu I)^n g = \sum_{i=0}^n \binom{n}{i} (\operatorname{ad} f)^i(g)(f - \mu I)^{n-i}$. Therefore, if $(f - \mu I)^n(v) = 0$, then $(f - \mu I)^{n+m-1}(g(v)) = 0$, so $g(v) \in W$.

For (ii) note that if $z \in L$ satisfies that $\rho(z)$ has more than one eigenvalue, then $V = W_1 \oplus \cdots \oplus W_r$, where the $W_i$'s are the generalized eigenspaces for $\rho(z)$. By item (i), the $W_i$'s are submodules of $V$, so the argument can be repeated to get a decomposition $V = V_1 \oplus \cdots \oplus V_n$, where for each $i$, $\rho(x)$ has a unique eigenvalue on $V_i$ for any $x \in L$. Lie's Theorem (1.7) shows then that for any $i$ there is a linear form $\lambda_i$ on $L$ such that $V_i \subseteq V_{\lambda_i}$, thus completing the proof.

For (iii) note that if $T$ is a toral subalgebra of $L$ and $x \in T$ with $[x, T] \neq 0$ then, since $x$ is semisimple, there is a $y \in T$ and a $0 \neq \alpha \in k$ with $[x, y] = \alpha y$. But then $(\operatorname{ad} y)^2(x) = 0$ and, since $y$ is semisimple, $\operatorname{ad}_y(x) = 0$, a contradiction. Hence $T$ is an abelian subalgebra of $L$. □

**4.5 Theorem.** *Let $L$ be a semisimple Lie algebra over an algebraically closed field $k$ of characteristic $0$, and let $H$ be a subalgebra of $L$. Then $H$ is a Cartan subalgebra of $L$ if and only if it is a maximal toral subalgebra of $L$.*

*Proof.* Assume first that $H$ is a Cartan subalgebra of $L$ so, by the previous lemma, $L = \oplus_{\lambda \in H^*} L_\lambda$, where $L_\lambda = \{x \in L : \forall h \in H\ (\operatorname{ad} h - \lambda(h)I)^n(x) = 0 \text{ for some } n\}$ for any $\lambda$. But then $H$ acts by nilpotent endomorphisms on $L_0$, and hence on $L_0/H$. If $H \neq L_0$, Engel's Theorem shows that there is an element $x \in L_0 \setminus H$ such that $[h, x] \in H$ for any $h \in H$, that is, $x \in N_L(H) \setminus H$, a contradiction with $H$ being self-normalizing. Hence we have $L = H \oplus \bigl(\oplus_{0 \neq \lambda \in H^*} L_\lambda\bigr)$.

One checks immediately that $[L_\lambda, L_\mu] \subseteq L_{\lambda+\mu}$ and, thus, $\kappa\bigl(L_\lambda, L_\mu\bigr) = 0$ if $\lambda \neq -\mu$, where $\kappa$ is the Killing form of $L$. Since $\kappa$ is nondegenerate and $\kappa\bigl(H, L_\lambda\bigr) = 0$ for any $0 \neq \lambda \in H^*$, the restriction of $\kappa$ to $H$ is nondegenerate too.

Now, $H$ is nilpotent, and hence solvable. By Proposition 1.11 applied to $\operatorname{ad} H \subseteq \mathfrak{gl}(L)$, $\kappa\bigl([H, H], H\bigr) = 0$ and, since $\kappa|_H$ is nondegenerate, we conclude that $[H, H] = 0$, that is, $H$ is abelian.

For any $x \in H$, $[x, H] = 0$ implies that $[x_s, H] = 0 = [x_n, H]$. Hence $x_n \in H$ and $\operatorname{ad} x_n$ is nilpotent. Thus, for any $y \in H$, $[x_n, y] = 0$, so $\operatorname{ad}_{x_n} \operatorname{ad}_y$ is a nilpotent endomorphism of $L$. This shows that $\kappa(x_n, H) = 0$ and hence $x_n = 0$. Therefore $H$ is toral. On the other hand, if $H \subseteq S$, for a toral subalgebra $S$ of $L$, then $S$ is abelian, so $[S, H] = 0$ and $S \subseteq N_L(H) = H$. Thus, $H$ is a maximal toral subalgebra of $L$.

Conversely, let $T$ be a maximal toral subalgebra of $L$. Then $T$ is abelian. Let $\{x_1, \ldots, x_m\}$ be a basis of $T$. Then $\operatorname{ad} x_1, \ldots, \operatorname{ad} x_m$ are commuting diagonalizable endomorphisms of $L$, so they are simultaneously diagonalizable. This shows that $L = \oplus_{\lambda \in T^*} L_\lambda(T)$, where $T^*$ is the dual vector space to $T$ and $L_\lambda(T) = \{y \in L : [t, y] = \lambda(t) y\ \forall t \in T\}$. As before, $[L_\lambda(T), L_\mu(T)] \subseteq L_{\lambda+\mu}(T)$ for any $\lambda, \mu \in T^*$ and $L_0(T) = C_L(T)\ (= \{x \in L : [x, T] = 0\})$, the centralizer of $T$.

For any $x = x_s + x_n \in C_L(T)$, both $x_s, x_n \in C_L(T)$. Hence $T + kx_s$ is a toral subalgebra. By maximality, $x_s \in T$. Then $\operatorname{ad} x|_{C_L(T)} = \operatorname{ad} x_n|_{C_L(T)}$ is nilpotent, so by Engel's Theorem, $H = C_L(T)$ is a nilpotent subalgebra. Moreover, for any $x \in N_L(H)$ and $t \in T$, $[x, t] \in [x, H] \subseteq H$, so $[[x, t], t] = 0$ and, since $t$ is semisimple, we get $[x, t] = 0$, so $x \in C_L(T) = H$. Thus $N_L(H) = H$ and $H$ is a Cartan subalgebra of $L$. By the first part of the proof, $H$ is a toral subalgebra which contains $T$ and, by maximality of $T$, $T = H$ is a Cartan subalgebra of $L$. □

**4.6 Corollary.** *Let $L$ be a semisimple Lie algebra over a field $k$ of characteristic 0 and let $H$ be a subalgebra of $L$. Then $H$ is a Cartan subalgebra of $L$ if and only if it is a maximal subalgebra among the subalgebras which are both abelian and toral.*

*Proof.* The properties of being nilpotent and self normalizing are preserved under extension of scalars. Thus, if $\bar{k}$ is an algebraic closure of $k$ and $H$ is nilpotent and self normalizing, so is $\bar{k} \otimes_k H$. Hence $\bar{k} \otimes_k H$ is a Cartan subalgebra of $\bar{k} \otimes_k L$. By the previous proof, it follows that $\bar{k} \otimes_k H$ is abelian, toral and self centralizing, hence so is $H$. But, since $H = C_L(H)$, $H$ is not contained in any bigger abelian subalgebra.

Conversely, if $H$ is a subalgebra which is maximal among the subalgebras which are both abelian and toral, the arguments in the previous proof show that $C_L(H)$ is a Cartan subalgebra of $L$, and hence abelian and toral and containing $H$. Hence $H = C_L(H)$ is a Cartan subalgebra. □

**4.7 Exercises.**

(i) Let $L = \mathfrak{sl}(n)$ be the Lie algebra of $n \times n$ trace zero matrices, and let $H$ be the subalgebra consisting of the diagonal matrices of $L$. Prove that $H$ is a Cartan subalgebra of $L$ and that $L = H \oplus \big(\oplus_{1 \leq i \neq j \leq n} L_{\epsilon_i - \epsilon_j}(H)\big)$, where $\epsilon_i \in H^*$ is the linear form that takes any diagonal matrix to its $i^{\text{th}}$ entry. Also show that $L_{\epsilon_i - \epsilon_j}(H) = kE_{ij}$, where $E_{ij}$ is the matrix with 1 in the $(i, j)$ position and 0's elsewhere.

(ii) Check that $\mathbb{R}^3$ is a Lie algebra under the usual vector cross product. Prove that it is toral but not abelian.

**4.8 Engel subalgebras.** There is another approach to Cartan subalgebras with its own independent interest.

Let $L$ be a Lie algebra over a field $k$, and let $x \in L$, the subalgebra

$$E_L(x) = \{y \in L : \exists n \in \mathbb{N} \text{ such that } (\operatorname{ad} x)^n(y) = 0\}$$

is called an *Engel subalgebra* of $L$ relative to $x$.

$E_L(x)$ is indeed a subalgebra and $\dim_k E_L(x)$ is the multiplicity of 0 as an eigenvalue of $\operatorname{ad} x$.

The main properties of Engel subalgebras are summarized here:

1. Let $S$ be a subalgebra of $L$, and let $x \in L$ such that $E_L(x) \subseteq S$. Then $N_L(S) = S$, where $N_L(S) = \{y \in L : [y, S] \subseteq S\}$ is the *normalizer* of $S$ in $L$. (Note that $N_L(S)$ is always a subalgebra of $L$ and $S$ is an ideal of $N_L(S)$.)

   *Proof.* We have $x \in E_L(x) \subseteq S$ so 0 is not an eigenvalue of the action of $\operatorname{ad} x$ on $N_L(S)/S$. On the other hand $\operatorname{ad} x\big(N_L(S)\big) \subseteq [S, N_L(S)] \subseteq S$. Hence $N_L(S)/S = 0$, or $N_L(S) = S$. □

2. Assume that $k$ is infinite. Let $S$ be a subalgebra of $L$ and let $z \in S$ be an element such that $E_L(z)$ is minimal in the set $\{E_L(x) : x \in S\}$. If $S \subseteq E_L(z)$, then $E_L(z) \subseteq E_L(x)$ for any $x \in S$.

   *Proof.* Put $S_0 = E_L(z)$. Then $S \subseteq S_0 \subseteq L$. For any $x \in S$ and $\alpha \in k$, $z + \alpha x \in S$, so that $\mathrm{ad}(z+\alpha x)$ leaves invariant both $S$ and $S_0$. Hence, the characteristic polynomial of $\mathrm{ad}(z + \alpha x)$ is a product $f_\alpha(X)g_\alpha(X)$, where $f_\alpha(X)$ is the characteristic polynomial of the restriction of $\mathrm{ad}(z + \alpha x)$ to $S_0$ and $g_\alpha(X)$ is the characteristic polynomial of the action of $\mathrm{ad}(z+\alpha x)$ on the quotient $L/S_0$. Let $r = \dim_k S_0$ and $n = \dim_k L$. Thus,

$$\begin{aligned} f_\alpha(X) &= X^r + f_1(\alpha)X^{r-1} + \cdots + f_r(\alpha), \\ g_\alpha(X) &= X^{n-r} + g_1(\alpha)X^{n-r-1} + \cdots + g_{n-r}(\alpha), \end{aligned}$$

   with $f_i(\alpha), g_i(\alpha)$ polynomials in $\alpha$ of degree $\leq i$, for any $i$.

   By hypothesis, $g_{n-r}(0) \neq 0$, and since $k$ is infinite, there are different scalars $\alpha_1, \dots, \alpha_{r+1} \in k$ with $g_{n-r}(\alpha_j) \neq 0$ for any $j = 1, \dots, r + 1$. This shows that $E_L(z + \alpha_j x) \subseteq S_0$ for any $j$. But $S_0 = E_L(z)$ is minimal, so $E_L(z) = E_L(z + \alpha_j x)$ for any $j$. Hence $f_{\alpha_j}(X) = X^r$ for any $j = 1, \dots, r + 1$, and this shows that $f_i(\alpha_j) = 0$ for any $i = 1, \dots, r$ and $j = 1, \dots, r + 1$. Since the degree of each $f_i$ is at most $r$, this proves that $f_i = 0$ for any $i$ and, thus, $\mathrm{ad}(z + \alpha x)$ is shown to act nilpotently on $E_L(z) = S_0$ for any $\alpha \in k$: $E_L(z) \subseteq E_L(z + \alpha x)$ for any $x \in S$ and $\alpha \in k$. Therefore, $E_L(z) \subseteq E_L(x)$ for any $x \in S$. □

3. Let $L$ be a Lie algebra over an infinite field $k$ and let $H$ be a subalgebra of $L$. Then $H$ is a Cartan subalgebra of $L$ if and only if it is a minimal Engel subalgebra of $L$.

   *Proof.* If $H = E_L(z)$ is a minimal Engel subalgebra of $L$, then by Property 1 above, $H$ is self normalizing, while Property 2 shows that $H \subseteq E_L(x)$ for any $x \in H$ which, by Engel's Theorem, proves that $H$ is nilpotent.

   Conversely, let $H$ be a nilpotent self normalizing subalgebra. By nilpotency, $H \subseteq E_L(x)$ for any $x \in H$ and, hence, it is enough to prove that there is an element $z \in H$ with $H = E_L(z)$. Take $z \in H$ with $E_L(z)$ minimal in $\{E_L(x) : x \in H\}$. By Property 2 above, $H \subseteq E_L(z) \subseteq E_L(x)$ for any $x \in H$. This means that $\mathrm{ad}\, x$ acts nilpotently on $E_L(z)/H$ for any $x \in H$ so, if $H \subsetneqq E_L(z)$, Engel's Theorem shows that there is an element $y \in E_L(z) \setminus H$ such that $[x, y] \in H$ for any $x \in H$, but then $y \in N_L(H) \setminus H$, a contradiction. Hence $H = E_L(z)$, as required. □

## § 5. Root space decomposition

Throughout this section, $L$ will denote a semisimple Lie algebra over an algebraically closed field $k$ of characteristic 0, with Killing form $\kappa$. Moreover, $H$ will denote a fixed Cartan subalgebra of $L$.

The arguments in the previous section show that there is a finite set $\Phi \subseteq H^* \setminus \{0\}$ of nonzero linear forms on $H$, whose elements are called *roots*, such that

$$L = H \oplus \big(\oplus_{\alpha\in\Phi} L_\alpha\big), \tag{5.3}$$

where $L_\alpha = \{x \in L : [h, x] = \alpha(h)x \ \forall h \in H\} \neq 0$ for any $\alpha \in \Phi$. Moreover, $H = C_L(H)$ and $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$, where $H = L_0$ and $L_\mu = 0$ if $0 \neq \mu \notin \Phi$.

### 5.1. Properties of the roots

(i) If $\alpha, \beta \in \Phi \cup \{0\}$ and $\alpha + \beta \neq 0$, then $\kappa\big(L_\alpha, L_\beta\big) = 0$.

*Proof.* $\operatorname{ad} x_\alpha \operatorname{ad} x_\beta$ takes each $L_\mu$ to $L_{\mu+(\alpha+\beta)} \neq L_\mu$ so its trace is 0. □

(ii) If $\alpha \in \Phi$, then $-\alpha \in \Phi$. Moreover, the restriction $\kappa : L_\alpha \times L_{-\alpha} \to k$ is nondegenerate.

*Proof.* Otherwise, $\kappa(L_\alpha, L)$ would be 0, a contradiction with the nondegeneracy of $\kappa$. □

(iii) $\Phi$ spans $H^*$.

*Proof.* Otherwise, there would exist a $0 \neq h \in H$ with $\alpha(h) = 0$ for any $\alpha \in \Phi$, so $\operatorname{ad} h = 0$ and $h = 0$, because $Z(L) = 0$ since $L$ is semisimple. □

(iv) For any $\alpha \in \Phi$, $[L_\alpha, L_{-\alpha}] \neq 0$.

*Proof.* It is enough to take into account that $0 \neq \kappa\big(L_\alpha, L_{-\alpha}\big) = \kappa\big([H, L_\alpha], L_{-\alpha}\big) = \kappa\big(H, [L_\alpha, L_{-\alpha}]\big)$. □

(v) For any $\alpha \in H^*$, let $t_\alpha \in H$ such that $\kappa(t_\alpha, \,.\,) = \alpha \in H^*$. Then for any $\alpha \in \Phi$, $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$,

$$[x_\alpha, y_\alpha] = \kappa(x_\alpha, y_\alpha)t_\alpha.$$

*Proof.* For any $h \in H$,

$$\begin{aligned}\kappa(h, [x_\alpha, y_\alpha]) &= \kappa([h, x_\alpha], y_\alpha)\\ &= \alpha(h)\kappa(x_\alpha, y_\alpha) = \kappa(t_\alpha, h)\kappa(x_\alpha, y_\alpha) = \kappa\big(h, \kappa(x_\alpha, y_\alpha)t_\alpha\big)\end{aligned}$$

and the result follows by the nondegeneracy of the restriction of $\kappa$ to $H = L_0$. □

(vi) For any $\alpha \in \Phi$, $\alpha(t_\alpha) \neq 0$.

*Proof.* Take $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$ such that $\kappa(x_\alpha, y_\alpha) = 1$. By the previous item $[x_\alpha, y_\alpha] = t_\alpha$. In case $\alpha(t_\alpha) = 0$, then $[t_\alpha, x_\alpha] = 0 = [t_\alpha, y_\alpha]$, so $S = kx_\alpha + kt_\alpha + ky_\alpha$ is a solvable subalgebra of $L$. By Lie's Theorem $kt_\alpha = [S, S]$ acts nilpotently on $L$ under the adjoint representation. Hence $t_\alpha$ is both semisimple ($H$ is toral) and nilpotent, hence $t_\alpha = 0$, a contradiction since $\alpha \neq 0$. □

(vii) For any $\alpha \in \Phi$, $\dim_k L_\alpha = 1$ and $k\alpha \cap \Phi = \{\pm\alpha\}$.

*Proof.* With $x_\alpha$, $y_\alpha$ and $t_\alpha$ as above, $S = kx_\alpha + kt_\alpha + ky_\alpha$ is isomorphic to $\mathfrak{sl}_2(k)$, under an isomorphism that takes $h$ to $\frac{2}{\alpha(t_\alpha)}t_\alpha$, $x$ to $x_\alpha$, and $y$ to $\frac{2}{\alpha(t_\alpha)}y_\alpha$.

Now, $V = H \oplus \big(\oplus_{0\neq\mu\in k} L_{\mu\alpha}\big)$ is a module for $S$ under the adjoint representation, and hence it is a module for $\mathfrak{sl}_2(k)$ through the isomorphism above. Besides, $V_0 = \{v \in V : [t_\alpha, v] = 0\}$ coincides with $H$. The eigenvalues taken by $h = \frac{2}{\alpha(t_\alpha)}t_\alpha$ are $\mu\alpha(h) = \frac{2\mu\alpha(t_\alpha)}{\alpha(t_\alpha)} = 2\mu$ and, thus, $\mu \in \frac{1}{2}\mathbb{Z}$, since all these eigenvalues are integers. On the other hand, $\ker\alpha$ is a trivial submodule of $V$, and $S$ is another submodule. Hence $\ker\alpha \oplus S$ is a submodule of $V$ which exhausts the eigenspace of $\operatorname{ad} h$ with eigenvalue 0. Hence by Weyl's Theorem, $V$ is the direct sum of $\ker\alpha\oplus S$ and a direct sum of irreducible submodules for $S$ in which 0 is not an eigenvalue for the action of $h$. We conclude that the only even eigenvalues of the action of $h$ are 0, 2 and $-2$, and this shows that $2\alpha \notin \Phi$. That is, the double of a root is never a root. But then $\frac{1}{2}\alpha$ cannot be a root neither, since otherwise $\alpha = 2\frac{1}{2}\alpha$ would not be a root. As a consequence, 1 is not an eigenvalue of the action of $h$ on $V$, and hence $V = \ker\alpha \oplus S$. In particular, $L_\alpha = kx_\alpha$, $L_{-\alpha} = ky_\alpha$ and $k\alpha\cap\Phi = \{\pm\alpha\}$. □

(viii) For any $\alpha \in \Phi$, let $h_\alpha = \frac{2}{\alpha(t_\alpha)}t_\alpha$, which is the unique element $h$ in $[L_\alpha, L_{-\alpha}] = kt_\alpha$ such that $\alpha(h) = 2$, and let $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$ such that $[x_\alpha, y_\alpha] = h_\alpha$. Then, for any $\beta \in \Phi$, $\beta(h_\alpha) \in \mathbb{Z}$.

*Proof.* Consider the subalgebra $S_\alpha = kx_\alpha + kh_\alpha + ky_\alpha$, which is isomorphic to $\mathfrak{sl}_2(k)$. From the representation theory of $\mathfrak{sl}_2(k)$, we know that the set of eigenvalues of the adjoint action of $h_\alpha$ on $L$ are integers. In particular, $\beta(h_\alpha) \in \mathbb{Z}$. □

More precisely, consider the $S_\alpha$-module $V = \oplus_{m\in\mathbb{Z}} L_{\beta+m\alpha}$. The eigenvalues of the adjoint action of $h_\alpha$ on $V$ are $\{\beta(h_\alpha) + 2m : m \in \mathbb{Z}$ such that $L_{\beta+m\alpha} \neq 0\}$, which form a chain of integers:

$$\beta(h_\alpha) + 2q,\ \beta(h_\alpha) + 2(q-1),\ \ldots, \beta(h_\alpha) - 2r,$$

with $r, q \in \mathbb{Z}_{\geq 0}$ and $\beta(h_\alpha) + 2q = -\big(\beta(h_\alpha) - 2r\big)$. Therefore, $\beta(h_\alpha) = r - q \in \mathbb{Z}$. The chain $(\beta + q\alpha, \ldots, \beta - r\alpha)$ is called the *$\alpha$-string* through $\beta$. It is contained in $\Phi \cup \{0\}$.

**5.2 Remark.** Since the restriction of $\kappa$ to $H$ is nondegenerate, it induces a nondegenerate symmetric bilinear form $(.\,|\,.) : H^* \times H^* \to k$, given by $(\alpha|\beta) = \kappa(t_\alpha, t_\beta)$ (where, as before, $t_\alpha$ is determined by $\alpha = \kappa(t_\alpha, .\,)$ for any $\alpha \in H^*$). Then for any $\alpha, \beta \in \Phi$, $\beta(t_\alpha) = \kappa(t_\beta, t_\alpha) = (\beta|\alpha)$. Hence

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

(ix) For any $\alpha \in \Phi$, consider the linear map $\sigma_\alpha : H^* \to H^*$, $\beta \mapsto \beta - 2\frac{(\beta|\alpha)}{(\alpha|\alpha)}\alpha$. (This is the *reflection* through $\alpha$, since $\sigma_\alpha(\alpha) = -\alpha$ and if $\beta$ is orthogonal to $\alpha$, that is, $(\beta|\alpha) = 0$, then $\sigma_\alpha(\beta) = \beta$. Hence $\sigma_\alpha^2 = 1$.)

Then $\sigma_\alpha(\Phi) \subseteq \Phi$. In particular, the group $\mathcal{W}$ generated by $\{\sigma_\alpha : \alpha \in \Phi\}$ is a finite subgroup of $GL(H^*)$, which is called the *Weyl group*.

*Proof.* For any $\alpha, \beta \in \Phi$, $\sigma_\alpha(\beta) = \beta - (r - q)\alpha$ ($r$ and $q$ as before), which is in the $\alpha$-string through $\beta$, and hence belongs to $\Phi$. (Actually, $\sigma_\alpha$ changes the order in the $\alpha$-string, in particular $\sigma_\alpha(\beta + q\alpha) = \beta - r\alpha$.)

Now $\mathcal{W}$ embeds in the symmetric group of $\Phi$, and hence it is finite. □

(x) Let $\{\alpha_1, \dots, \alpha_n\}$ be a basis of $H^*$ contained in $\Phi$. Then $\Phi \subseteq \mathbb{Q}\alpha_1 + \cdots + \mathbb{Q}\alpha_n$.

*Proof.* For any $\alpha \in \Phi$, $\alpha = \mu_1\alpha_1 + \cdots + \mu_n\alpha_n$ with $\mu_1, \dots, \mu_n \in k$. But for $i = 1, \dots, n$,

$$\frac{2(\alpha|\alpha_i)}{(\alpha_i|\alpha_i)} = \sum_{j=1}^{n} \mu_j \frac{2(\alpha_j|\alpha_i)}{(\alpha_i|\alpha_i)}$$

and this gives a system of linear equations on the $\mu_j$'s with a regular integral matrix. Solving by Crammer's rule, one gets that the $\mu_j$'s belong to $\mathbb{Q}$. □

(xi) For any $\alpha, \beta \in \Phi$, $(\alpha|\beta) \in \mathbb{Q}$. Moreover, the restriction $(.\,|\,.) : \mathbb{Q}\Phi \times \mathbb{Q}\Phi \to \mathbb{Q}$ is positive definite.

*Proof.* Since $L = H \oplus \bigl(\oplus_{\alpha\in\Phi} L_\alpha\bigr)$ and $\dim_k L_\alpha = 1$ for any $\alpha \in \Phi$, given any $\beta \in \Phi$,

$$(\beta|\beta) = \kappa(t_\beta, t_\beta) = \operatorname{trace}\bigl((\operatorname{ad} t_\beta)^2\bigr) = \sum_{\alpha\in\Phi} \alpha(t_\beta)^2 = \frac{(\beta|\beta)^2}{4} \sum_{\alpha\in\Phi} \alpha(h_\beta)^2,$$

and, therefore,

$$(\beta|\beta) = \frac{4}{\sum_{\alpha\in\Phi} \alpha(h_\beta)^2} \in \mathbb{Q}_{>0}.$$

Now, for any $\alpha, \beta \in \Phi$, $\frac{2(\alpha|\beta)}{(\beta|\beta)} \in \mathbb{Z}$, so $(\alpha|\beta) = \frac{(\beta|\beta)}{2}\frac{2(\alpha|\beta)}{(\beta|\beta)} \in \mathbb{Q}$. And for any $\beta \in \mathbb{Q}\Phi$, $\beta = \mu_1\alpha_1 + \cdots + \mu_n\alpha_n$ for some $\mu_j$'s in $\mathbb{Q}$, so

$$(\beta|\beta) = \sum_{\alpha\in\Phi} \alpha(t_\beta)^2 = \sum_{\alpha\in\Phi} \bigl(\mu_1\alpha(t_{\alpha_1}) + \cdots + \mu_n\alpha(t_{\alpha_n})\bigr)^2 \geq 0.$$

Besides $(\beta|\beta) = 0$ if and only if $\alpha(t_\beta) = 0$ for any $\alpha \in \Phi$, if and only if $t_\beta = 0$ since $\Phi$ spans $H^*$, if and only if $\beta = 0$. □

Therefore, if the dimension of $H$ is $n$ (this dimension is called the *rank* of $L$, although we do not know yet that it does not depend on the Cartan subalgebra chosen), then $E_{\mathbb{Q}} = \mathbb{Q}\Phi$ is an $n$-dimensional vector space over $\mathbb{Q}$ endowed with a positive definite symmetric bilinear form $(\,|\,)$.

Then $E = \mathbb{R} \otimes_{\mathbb{Q}} E_{\mathbb{Q}}$ is an euclidean $n$-dimensional vector space which contains a subset $\Phi$ which satisfies:

(R1) $\Phi$ is a finite subset that spans $E$, and $0 \notin \Phi$.

(R2) For any $\alpha \in \Phi$, $-\alpha \in \Phi$ too and $\mathbb{R}\alpha \cap \Phi = \{\pm\alpha\}$.

(R3) For any $\alpha \in \Phi$, the reflection on the hyperplane $(\mathbb{R}\alpha)^\perp$ leaves $\Phi$ invariant (i.e., for any $\alpha, \beta \in \Phi$, $\sigma_\alpha(\beta) \in \Phi$).

(R4) For any $\alpha, \beta \in \Phi$, $\langle\beta|\alpha\rangle = 2\dfrac{(\beta|\alpha)}{(\alpha|\alpha)} \in \mathbb{Z}$.

A subset $\Phi$ of an euclidean space, satisfying these properties (R1)–(R4), is called a *root system*, and the subgroup $\mathcal{W}$ of the orthogonal group $O(E)$ generated by the reflections $\sigma_\alpha$, $\alpha \in \Phi$, is called the *Weyl group* of the root system. The dimension of the euclidean space is called the *rank* of the root system. Note that $\mathcal{W}$ is naturally embedded in the symmetric group of $\Phi$, and hence it is a finite group.

## § 6. Classification of root systems

Our purpose here is to classify the root systems. Hence we will work in the abstract setting considered at the end of the last section. The arguments in this section follow the ideas in the article by R.W. Carter: *Lie Algebras and Root Systems*, in Lectures on Lie Groups and Lie Algebras (R.W. Carter, G. Segal and I. Macdonal), London Mathematical Society, Student Texts **22**, Cambridge University Press, 1995.

Let $\Phi$ be a root system in a euclidean space $E$. Take $\nu \in E$ such that $(\nu|\alpha) \neq 0$ for any $\alpha \in \Phi$. This is always possible since $\Phi$ is finite. (Here $(\,|\,)$ denotes the inner product on $E$.) Let $\Phi^+ = \{\alpha \in \Phi : (\nu|\alpha) > 0\}$ be the set of *positive roots*, so $\Phi = \Phi^+ \cup \Phi^-$ (disjoint union), where $\Phi^- = -\Phi^+$ (the set of negative roots).

A positive root $\alpha$ is said to be *simple* if it is not the sum of two positive roots. Let $\Delta = \{\alpha \in \Phi^+ : \alpha \text{ is simple}\}$, $\Delta$ is called a *system of simple roots* of $(E, \Phi)$.

**6.1 Proposition.** *Let $\Phi$ be a root system on a euclidean vector space $E$ and let $\Delta = \{\alpha_1, \dots, \alpha_n\}$ be a system of simple roots in $(E, \Phi)$. Then:*

(i) *For any $\alpha \neq \beta$ in $\Delta$, $(\alpha|\beta) \leq 0$.*

(ii) *$\Delta$ is a basis of $E$.*

(iii) $\Phi^+ \subseteq \mathbb{Z}_{\geq 0}\alpha_1 + \cdots + \mathbb{Z}_{\geq 0}\alpha_n$.

(iv) *For any $\alpha \in \Delta$, $\sigma_\alpha\big(\Phi^+ \setminus \{\alpha\}\big) = \Phi^+ \setminus \{\alpha\}$.*

(v) *If $\nu' \in E$ is a vector such that $(\nu'|\alpha) \neq 0$ for any $\alpha \in \Phi$ and $\Delta'$ is the associated system of simple roots, then there is an element $\sigma \in \mathcal{W}$ such that $\sigma(\Delta) = \Delta'$.*

*Proof.* For any $\alpha, \beta \in \Phi$, consider the integer

$$N_{\alpha\beta} = \langle\alpha|\beta\rangle\langle\beta|\alpha\rangle = \frac{4(\alpha|\beta)^2}{(\alpha|\alpha)(\beta|\beta)} \in \mathbb{Z}_{\geq 0}.$$

The Cauchy-Schwarz inequality shows that $0 \leq N_{\alpha\beta} \leq 4$ and that $N_{\alpha\beta} = 4$ if and only if $\beta = \pm\alpha$, since $\mathbb{R}\alpha \cap \Phi = \{\pm\alpha\}$ by (R2).

Assume that $\alpha, \beta \in \Phi^+$ with $\alpha \neq \pm\beta$ and $(\alpha|\beta) \geq 0$. Then $0 \leq N_{\alpha\beta} = \langle\alpha|\beta\rangle\langle\beta|\alpha\rangle \leq 3$, so either $(\alpha|\beta) = 0$ or $\langle\alpha|\beta\rangle = 1$ or $\langle\beta|\alpha\rangle = 1$. If, for instance, $\langle\beta|\alpha\rangle = 1$, then $\sigma_\alpha(\beta) = \beta - \langle\beta|\alpha\rangle\alpha = \beta - \alpha \in \Phi$. If $\beta - \alpha \in \Phi^+$, then $\beta = \alpha + (\beta - \alpha)$ is not simple, while if $\beta - \alpha \in \Phi^-$, then $\alpha = \beta + (\alpha - \beta)$ is not simple. This proves item (i).

Now, for any $\alpha \in \Phi^+$, either $\alpha \in \Delta$ or $\alpha = \beta + \gamma$, with $\beta, \gamma \in \Phi^+$. But in the latter case, $(\nu|\alpha) = (\nu|\beta) + (\nu|\gamma)$, with $0 < (\nu|\alpha), (\nu|\beta), (\nu|\gamma)$, so that both $(\nu|\beta)$ and $(\nu|\gamma)$

are strictly lower than $(\nu|\alpha)$. Now, we proceed in the same way with $\beta$ and $\gamma$. They are either simple or a sum of "smaller" positive roots. Eventually we end up showing that $\alpha$ is a sum of simple roots, which gives (iii).

In particular, this shows that $\Delta$ spans $E$. Assume that $\Delta$ were not a basis, then there would exist disjoint nonempty subsets $I, J \subseteq \{1, \dots, n\}$ and positive scalars $\mu_i$ such that $\sum_{i\in I} \mu_i\alpha_i = \sum_{j\in J} \mu_j\alpha_j$. Let $\gamma = \sum_{i\in I} \mu_i\alpha_i = \sum_{j\in J} \mu_j\alpha_j$. Then $0 \leq (\gamma|\gamma) = \sum_{\substack{i\in I\\ j\in J}} \mu_i\mu_j(\alpha_i|\alpha_j) \leq 0$ (because of (i)). Thus $\gamma = 0$, but this would imply that $0 < \sum_{i\in I} \mu_i(\nu|\alpha_i) = (\nu|\gamma) = 0$, a contradiction that proves (ii).

In order to prove (iv), we may assume that $\alpha = \alpha_1$. Let $\alpha \neq \beta \in \Phi^+$, then (iii) shows that $\beta = \sum_{i=1}^n m_i\alpha_i$, with $m_i \in \mathbb{Z}_{\geq 0}$ for any $i$. Since $\beta \neq \alpha$, there is a $j \geq 2$ such that $m_j > 0$. Then $\sigma_\alpha(\beta) = \beta - \langle\beta|\alpha\rangle\alpha = (m_1 - \langle\beta|\alpha\rangle)\alpha_1 + m_2\alpha_2 + \cdots + m_n\alpha_n \in \Phi$, and one of the coefficients, $m_j$, is $> 0$. Hence $\alpha \neq \sigma_\alpha(\beta) \notin \Phi^-$, so that $\sigma_\alpha(\beta) \in \Phi^+ \setminus \{\alpha\}$.

Finally, let us prove (v). We know that $\Phi = \Phi^+ \cup \Phi^- = \Phi'^+ \cup \Phi'^-$ (with obvious notation). Let $\rho = \frac{1}{2}\sum_{\alpha\in\Phi^+} \alpha$ (which is called the *Weyl vector*), and let $\sigma \in \mathcal{W}$ such that $(\sigma(\nu')|\rho)$ is maximal. Then, for any $\alpha \in \Delta$:

$$\begin{aligned}
(\sigma(\nu')|\rho) &\geq (\sigma_\alpha\sigma(\nu')|\rho)\\
&= \big(\sigma(\nu')|\sigma_\alpha(\rho)\big) \quad \text{(since } \sigma_\alpha^2 = 1 \text{ and } \sigma_\alpha \in O(E))\\
&= \big(\sigma(\nu')|\rho - \alpha\big) \quad \text{(because of (iv))}\\
&= \big(\sigma(\nu')|\rho\big) - \big(\sigma(\nu')|\alpha\big)\\
&= \big(\sigma(\nu')|\rho\big) - \big(\nu'|\sigma^{-1}(\alpha)\big),
\end{aligned}$$

so $\big(\nu'|\sigma^{-1}(\alpha)\big) \geq 0$. This shows that $\sigma^{-1}(\Delta) \subseteq \Phi'^+$, so $\sigma^{-1}(\Phi^\pm) = \Phi'^\pm$ and $\sigma^{-1}(\Delta)$ then coincides with the set of simple roots in $\Phi'^+$, which is $\Delta'$. □

Under the previous conditions, with $\Delta = \{\alpha_1, \dots, \alpha_n\}$, consider

- The square matrix $C = \Big(\langle\alpha_i|\alpha_j\rangle\Big)_{1\leq i,j\leq n}$, which is called the *Cartan matrix* of the root system.

  Note that for any $\alpha \neq \beta$ in $\Phi$ with $(\alpha|\beta) \leq 0$,

$$\langle\alpha|\beta\rangle = \frac{2(\alpha|\beta)}{(\beta|\beta)} = -\sqrt{\frac{(\alpha|\alpha)}{(\beta|\beta)}}\sqrt{\frac{4(\alpha|\beta)(\beta|\alpha)}{(\alpha|\alpha)(\beta|\beta)}} = -\frac{\sqrt{(\alpha|\alpha)}}{\sqrt{(\beta|\beta)}}\sqrt{N_{\alpha\beta}},$$

  so we get a factorization of the Cartan matrix as $C = D_1\hat{C}D_2$, where $D_1$ (respectively $D_2$) is the diagonal matrix with the elements $\sqrt{(\alpha_1|\alpha_1)}, \dots, \sqrt{(\alpha_n|\alpha_n)}$ (resp. $\frac{1}{\sqrt{(\alpha_1|\alpha_1)}}, \dots, \frac{1}{\sqrt{(\alpha_n|\alpha_n)}}$) on the diagonal, and

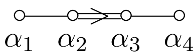
$$\hat{C} = \begin{pmatrix} 2 & -\sqrt{N_{\alpha_1\alpha_2}} & \cdots & -\sqrt{N_{\alpha_1\alpha_n}}\\ -\sqrt{N_{\alpha_2\alpha_1}} & 2 & \cdots & -\sqrt{N_{\alpha_2\alpha_n}}\\ \vdots & \vdots & \ddots & \vdots\\ -\sqrt{N_{\alpha_n\alpha_1}} & -\sqrt{N_{\alpha_n\alpha_2}} & \cdots & 2 \end{pmatrix}$$

  This matrix $\hat{C}$ is symmetric and receives the name of *Coxeter matrix* of the root system. It is nothing else but the coordinate matrix of the inner product $(\,|\,)$ in the basis $\{\hat{\alpha}_1, \dots, \hat{\alpha}_n\}$ with $\hat{\alpha}_i = \frac{\sqrt{2}\alpha_i}{\sqrt{(\alpha_i|\alpha_i)}}$. Note that $\det C = \det\hat{C}$.

**6.2 Exercise.** What are the possible Cartan and Coxeter matrices for $n = 2$? Here $\Delta = \{\alpha, \beta\}$, and you may assume that $(\alpha|\alpha) \leq (\beta|\beta)$.

- The *Dynkin diagram* of $\Phi$, which is the graph which consists of a node for each simple root $\alpha$. The nodes associated to $\alpha \neq \beta \in \Delta$ are connected by $N_{\alpha\beta}$ ($= 0, 1, 2$ or $3$) arcs. Moreover, if $N_{\alpha\beta} = 2$ or $3$, then $\alpha$ and $\beta$ have different length and an arrow is put pointing from the long to the short root. For instance,

$$C = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -2 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} \quad \longmapsto$$




- The *Coxeter graph* is the graph obtained by omitting the arrows in the Dynkin diagram.

  In our previous example it is .

Because of item (v) in Proposition 6.1, these objects depend only on $\Phi$ and not on $\Delta$, up to the same permutation of rows and columns in $C$ and up to the numbering of the vertices in the graphs.

The root system $\Phi$ is said to be *reducible* if $\Phi = \Phi_1 \cup \Phi_2$, with $\emptyset \neq \Phi_i$ ($i = 1, 2$) and $\big(\Phi_1|\Phi_2\big) = 0$. Otherwise, it is called *irreducible.*

**6.3 Theorem.**

(a) *A root system $\Phi$ is irreducible if and only if its Dynkin diagram (or Coxeter graph) is connected.*

(b) *Let $L$ be a semisimple Lie algebra over an algebraically closed field $k$ of characteristic $0$. Let $H$ be a Cartan subalgebra of $L$ and let $\Phi$ be the associated root system. Then $\Phi$ is irreducible if and only if $L$ is simple.*

*Proof.* For (a), if $\Phi$ is reducible with $\Phi = \Phi_1 \cup \Phi_2$ and $\Delta$ is a system of simple roots, then it is clear that $\Delta = \big(\Delta \cap \Phi_1\big) \cup \big(\Delta \cap \Phi_2\big)$ and the nodes associated to the elements in $\Delta \cap \Phi_1$ are not connected with those associated to $\Delta \cap \Phi_2$. Hence the Dynkin diagram is not connected.

Conversely, if $\Delta = \Delta_1 \cup \Delta_2$ (disjoint union) with $\emptyset \neq \Delta_1, \Delta_2$ and $\big(\Delta_1|\Delta_2\big) = 0$, let $E_1 = \mathbb{R}\Delta_1$ and $E_2 = \mathbb{R}\Delta_2$, so that $E$ is the orthogonal sum $E = E_1 \perp E_2$. Then $\Phi_i = \Phi \cap E_i$ is a root system in $E_i$ with system of simple roots $\Delta_i$ ($i = 1, 2$). It has to be checked that $\Phi = \Phi_1 \cup \Phi_2$. For any $\alpha \in \Phi_1$, $\sigma_\alpha|_{E_2}$ is the identity. Hence, item (v) in Proposition 6.1 shows that there exists an element $\sigma \in \mathcal{W}_1$ such that $\sigma(\Delta_1) = -\Delta_1$, where $\mathcal{W}_1$ is the subgroup of the Weyl group $\mathcal{W}$ generated by $\{\sigma_\alpha : \alpha \in \Phi_1\}$. Order the roots so that $\Delta_1 = \{\alpha_1, \ldots, \alpha_r\}$ and $\Delta_2 = \{\alpha_{r+1}, \ldots, \alpha_n\}$. Then any $\beta \in \Phi$ can be written as $\beta = m_1\alpha_1 + \cdots + m_n\alpha_n$, with $m_i \in \mathbb{Z}$ for any $i$, and either $m_i \geq 0$ or $m_i \leq 0$ for any $i$. But $\sigma(\beta) \in \Phi$ and, since $\sigma(\Delta_1) = -\Delta_1$, $\sigma(\beta) = m'_1\alpha_1 + \cdots + m'_r\alpha_r + m_{r+1}\alpha_{r+1} + \cdots + m_n\alpha_n$, where $(m'_1, \ldots, m'_r)$ is a permutation of $(-m_1, \ldots, -m_r)$. Since the coefficients of $\sigma(\beta)$ are also either all nonnegative or all nonpositive, we conclude that either $m_1 = \cdots = m_r = 0$ or $m_{r+1} = \cdots = m_n = 0$, that is, either $\beta \in \Phi_1$ or $\beta \in \Phi_2$.

For (b), assume first that $\Phi$ is reducible, so $\Phi = \Phi_1 \cup \Phi_2$ with $\bigl(\Phi_1|\Phi_2\bigr) = 0$ and $\Phi_1 \neq \emptyset \neq \Phi_2$. Then the subspace $\sum_{\alpha\in\Phi_1^+}\Bigl(L_\alpha + L_{-\alpha} + [L_\alpha, L_{-\alpha}]\Bigr)$ is a proper ideal of $L$, since

$$[L_\alpha, L_\beta] \begin{cases} = 0 & \text{if } \alpha+\beta \notin \Phi\cup\{0\}\text{, in particular if } \alpha\in\Phi_1 \text{ and } \beta\in\Phi_2\\ \subseteq L_{\alpha+\beta} & \text{otherwise.}\end{cases}$$

Hence $L$ is not simple in this case.

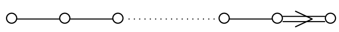
Conversely, if $L$ is not simple, then $L = L_1 \oplus L_2$ with $L_1$ and $L_2$ proper ideals of $L$. Hence $\kappa(L_1, L_2) = 0$ by the definition of the Killing form, and $H = C_L(H) = C_{L_1}(H) \oplus C_{L_2}(H) = (H \cap L_1) \oplus (H \cap L_2)$, because for any $h \in H$ and $x_i \in L_i$ $(i = 1, 2)$, $[h, x_1 + x_2] = [h, x_1] + [h, x_2]$, with the first summand in $L_1$ and the second one in $L_2$. Now, for any $\alpha \in \Phi$, $\alpha(H \cap L_i) \neq 0$ for some $i = 1, 2$. Then $L_\alpha = [H \cap L_i, L_\alpha] \subseteq L_i$, so the element $t_\alpha$ such that $\kappa(t_\alpha, .) = \alpha$ satisfies that $t_\alpha \in [L_\alpha, L_{-\alpha}] \subseteq L_i$. As a consequence, $\Phi = \Phi_1 \cup \Phi_2$ (disjoint union), with $\Phi_i = \{\alpha \in \Phi : \alpha(H \cap L_i) \neq 0\}$, and $(\alpha|\beta) = \kappa(t_\alpha, t_\beta) = 0$ for any $\alpha \in \Phi_1$ and $\beta \in \Phi_2$. Thus, $\Phi$ is reducible. □

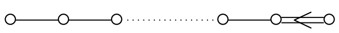
**6.4 Remark.** The proof above shows that the decomposition of the semisimple Lie algebra $L$ into a direct sum of simple ideals gives the decomposition of its root system $\Phi$ into an orthogonal sum of irreducible root systems.

Dynkin diagrams are classified as follows:

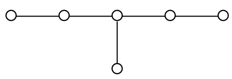
**6.5 Theorem.** *The Dynkin diagrams of the irreducible root systems are precisely the following (where $n$ indicates the number of nodes):*

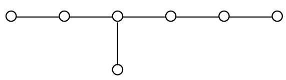
$(A_n)$  , $n \geq 1$.

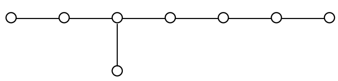
$(B_n)$  , $n \geq 2$.

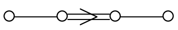
$(C_n)$  , $n \geq 3$.

$(D_n)$  , $n \geq 4$.

$(E_6)$  .

$(E_7)$  .

$(E_8)$  .

$(F_4)$  .

$(G_2)$  .

Most of the remainder of this section will be devoted to the proof of this Theorem.

First, it will be shown that the 'irreducible Coxeter graphs' are the ones corresponding to $(A_n)$, $(B_n = C_n)$, $(D_n)$, $(E_{6,7,8})$, $(F_4)$ and $(G_2)$. Any Coxeter graph determines the symmetric matrix $\bigl(a_{ij}\bigr)$ with $a_{ii} = 2$ and $a_{ij} = -\sqrt{N_{ij}}$ for $i \neq j$, where $N_{ij}$ is the number of lines joining the vertices $i$ and $j$. We know that this matrix is the coordinate matrix of a positive definite quadratic form on a real vector space.

Any graph formed by nodes and lines connecting these nodes will be called a 'Coxeter type graph'. For each such graph we will take the symmetric matrix $\bigl(a_{ij}\bigr)$ defined as before and the associated quadratic form on $\mathbb{R}^n$, which may fail to be positive definite, such that $q(e_i, e_j) = a_{ij}$, where $\{e_1, \dots, e_n\}$ denotes the canonical basis of $\mathbb{R}^n$.

**6.6 Lemma.** *Let $V$ be a real vector space with a basis $\{v_1, \dots, v_n\}$ and a positive definite quadratic form $q : V \to \mathbb{R}$ such that $q(v_i, v_j) \leq 0$ for any $i \neq j$, and $q(v_1, v_2) < 0$. (Here $q(v, w) = \frac{1}{2}\bigl(q(v+w) - q(v) - q(w)\bigr)$ gives the associated symmetric bilinear form.)*

*Let $\tilde{q} : V \to \mathbb{R}$ be a quadratic form such that its associated symmetric bilinear form satisfies $\tilde{q}(v_i, v_j) = q(v_i, v_j)$ for any $(i, j) \neq (1, 2)$, $i \leq j$, and $0 \geq \tilde{q}(v_1, v_2) > q(v_1, v_2)$. Then $\tilde{q}$ is positive definite too and $\det \tilde{q} > \det q$ (where $\det$ denotes the determinant of the quadratic form in any fixed basis).*

*Proof.* We apply a Gram-Schmidt process to obtain a new suitable basis of $\mathbb{R}v_2 + \cdots + \mathbb{R}v_n$ as follows:

$$\begin{aligned} w_n &= v_n \\ w_{n-1} &= v_{n-1} + \lambda_{n-1,n} w_n \\ \vdots\ &\quad \vdots \\ w_2 &= v_2 + \lambda_{2,3} w_3 + \cdots + \lambda_{2,n} w_n \end{aligned}$$

where the $\lambda$'s are determined by imposing that $q(w_i, w_j) = 0$ for any $i > j \geq 2$. Note that $q(w_i, w_j) = \tilde{q}(w_i, w_j)$ for any $i > j \geq 2$, and that this process gives that $\lambda_{i,j} \geq 0$ for any $2 \leq i < j \leq n$ and $q(v_i, w_j) \leq 0$ for any $1 \leq i < j \leq n$. Now take $w_1 = v_1 + \lambda_{1,3} w_3 + \cdots + \lambda_{1,n} w_n$, and determine the coefficients by imposing that $q(w_1, w_i) = 0$ for any $i \geq 3$. Then $q(w_1, w_2) = q(v_1, w_2) \leq q(v_1, v_2) < 0$, $\tilde{q}(w_1, w_2) = \tilde{q}(v_1, w_2) \leq \tilde{q}(v_1, v_2) \leq 0$, and $0 \geq \tilde{q}(w_1, w_2) > q(w_1, w_2)$.

In the basis $\{w_1, \dots, w_n\}$, the coordinate matrices of $q$ and $\tilde{q}$ present the form

$$\begin{pmatrix} \alpha_1 & \beta & 0 & \cdots & 0 \\ \beta & \alpha_2 & 0 & \cdots & 0 \\ 0 & 0 & \alpha_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_n \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \alpha_1 & \tilde{\beta} & 0 & \cdots & 0 \\ \tilde{\beta} & \alpha_2 & 0 & \cdots & 0 \\ 0 & 0 & \alpha_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_n \end{pmatrix}$$

with $0 \geq \tilde{\beta} > \beta$. Since $q$ is positive definite, $\alpha_i \geq 0$ for any $i$ and $\alpha_1\alpha_2 - \beta^2 > 0$. Hence $\alpha_1\alpha_2 - \tilde{\beta}^2 > \alpha_1\alpha_2 - \beta^2 > 0$ and the result follows. $\square$

Note that by suppressing a line connecting nodes $i$ and $j$ in a Coxeter type graph, with associated quadratic form $q$, the quadratic form $\tilde{q}$ associated to the new graph obtained differs only in that $0 > \tilde{q}(e_i, e_j) > q(e_i, e_j)$. Hence the previous Lemma immediately implies the following result:

**6.7 Corollary.** *If some lines connecting two nodes on a Coxeter type graph with positive definite quadratic form are suppressed, then the new graph obtained is a new Coxeter type graph with positive definite quadratic form.*

Let us compute now the matrices associated to some Coxeter type graphs, as well as their determinants.

$A_n$ $(n \geq 1)$ . Here the associated matrix is

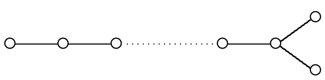
$$M_{A_n} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{pmatrix}$$

whose determinant can be computed recursively by expanding along the first row: $\det M_{A_n} = 2\det M_{A_{n-1}} - \det M_{A_{n-2}}$, obtaining that $\det M_{A_n} = n+1$ for any $n \geq 1$.

$B_n = C_n$ $(n \geq 2)$ . Here

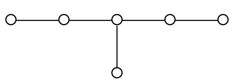
$$M_{B_n} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -\sqrt{2} \\ 0 & 0 & 0 & \cdots & -\sqrt{2} & 2 \end{pmatrix}$$

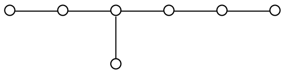
and, by expanding along the last row, $\det M_{B_n} = 2\det M_{A_{n-1}} - 2\det M_{A_{n-2}}$, so that $\det M_{B_n} = 2$.

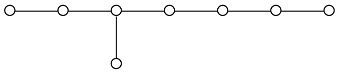
$D_n$ $(n \geq 4)$ . The associated matrix is

$$M_{D_n} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 0 & 2 \end{pmatrix}$$

so that $\det M_{D_4} = 4 = \det M_{D_5}$ and by expanding along the first row, $\det M_{D_n} = 2\det M_{D_{n-1}} - \det M_{D_{n-2}}$. Hence $\det M_{D_n} = 4$ for any $n \geq 4$.

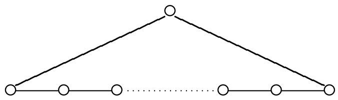
$E_6$ . Here $\det M_{E_6} = 2\det M_{D_5} - \det M_{A_4} = 8 - 5 = 3$ (expansion along the row corresponding to the leftmost node).

$E_7$ . Here $\det M_{E_7} = 2\det M_{E_6} - \det M_{D_5} = 6 - 4 = 2$.

$E_8$ . Here $\det M_{E_8} = 2\det M_{E_7} - \det M_{E_6} = 4 - 3 = 1$.

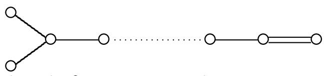
$F_4$ . Here $\det M_{F_4} = \det M_{B_3} - \det M_{A_2} = 4 - 3 = 1$.

$G_2$ . Here $\det M_{G_2} = \det \begin{vmatrix} 2 & -\sqrt{3} \\ -\sqrt{3} & 2 \end{vmatrix} = 1$.

$\tilde{A}_n$  ($n+1$ nodes, $n \geq 2$). Then

$$M_{\tilde{A}_n} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & -1 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ -1 & 0 & 0 & \cdots & -1 & 2 \end{pmatrix}$$

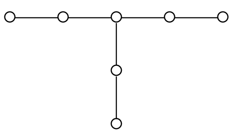
so the sum of the rows is the zero row. Hence $\det M_{\tilde{A}_n} = 0$.

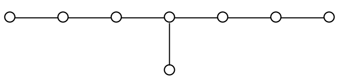
$\tilde{B}_n$  ($n+1$ nodes, $n \geq 3$) Let us number the nodes so that the leftmost nodes are nodes 1 and 2, and node 3 is connected to both of them. Then we may expand $\det M_{\tilde{B}_n} = 2\det M_{B_n} - \det M_{A_1} \det M_{B_{n-1}} = 4 - 4 = 0$. (For $n = 3$, $\det M_{\tilde{B}_3} = 2\det M_{B_2} - \det M_{A_1}^2 = 4 - 4 = 0$.)

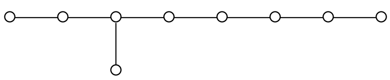
$\tilde{C}_n$  ($n+1$ nodes, $n \geq 2$). Then $\det M_{\tilde{C}_n} = 2\det M_{B_n} - 2\det M_{B_{n-1}} = 0$. (For $n = 2$, $\det M_{\tilde{C}_3} = 2\det M_{B_2} - 2\det M_{A_1} = 0$.)

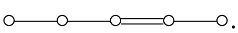
$\tilde{D}_n$  ($n+1$ nodes, $n \geq 4$). Here

$$\det M_{\tilde{D}_n} = \begin{cases} 2\det M_{D_4} - \det M_{A_1}^3 = 8 - 8 = 0, & \text{if } n = 4, \\ 2\det M_{D_5} - \det M_{A_1} \det M_{A_3} = 8 - 8 = 0, & \text{if } n = 5, \\ 2\det M_{D_n} - \det M_{A_1} \det M_{D_{n-2}} = 8 - 8 = 0, & \text{otherwise.} \end{cases}$$

$\tilde{E}_6$ . Here $\det M_{\tilde{E}_6} = 2\det M_{E_6} - \det M_{A_5} = 6 - 6 = 0$.

$\tilde{E}_7$ . Here $\det M_{\tilde{E}_7} = 2\det M_{E_7} - \det M_{D_6} = 4 - 4 = 0$.

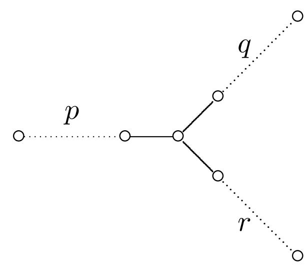
$\tilde{E}_8$ . Here $\det M_{\tilde{E}_8} = 2\det M_{E_8} - \det M_{E_7} = 2 - 2 = 0$.

$\tilde{F}_4$ . Here $\det M_{\tilde{F}_4} = 2\det M_{F_4} - \det M_{B_3} = 2 - 2 = 0$.

$\tilde{G}_2$ . Here $\det M_{\tilde{G}_2} = 2\det M_{G_2} - \det M_{A_1} = 2 - 2 = 0$.

Now, if $\mathcal{G}$ is a connected Coxeter graph and we suppress some of its nodes (and the lines connecting them), a new Coxeter type graph with positive definite associated quadratic form is obtained. The same happens, because of the previous Lemma 6.6, if only some lines are suppressed. The new graphs thus obtained will be called subgraphs.

- If $\mathcal{G}$ contains a cycle, then it has a subgraph (isomorphic to) $\tilde{A}_n$, and this is a contradiction since $\det M_{\tilde{A}_n} = 0$, so its quadratic form is not positive definite.
- If $\mathcal{G}$ contains a node which is connected to four different nodes, then it contains a subgraph of type $\tilde{D}_4$, a contradiction.
- If $\mathcal{G}$ contains a couple of nodes (called 'triple nodes') connected to three other nodes, then it contains a subgraph of type $\tilde{D}_n$, a contradiction again.
- If $\mathcal{G}$ contains two couples of nodes connected by at least two lines, then it contains a subgraph of type $\tilde{C}_n$, which is impossible.
- If $\mathcal{G}$ contains a triple node and two nodes connected by at least two lines, then it contains a subgraph of type $\tilde{B}_n$.
- If $\mathcal{G}$ contains a 'triple link', then either it is isomorphic to $G_2$ or contains a subgraph of type $\tilde{G}_2$, this latter possibility gives a contradiction.
- If $\mathcal{G}$ contains a 'double link' and this double link is not at a extreme of the graph, then either $\mathcal{G}$ is isomorphic to $F_4$ or contains a subgraph of type $\tilde{F}_4$, which is impossible.
- If $\mathcal{G}$ contains a 'double link' at one extreme, then the Coxeter graph is $B_n = C_n$.
- Finally, if $\mathcal{G}$ contains only simple links, then it is either $A_n$ or it contains a unique triple node. Hence it has the form:

  with $1 \leq p \leq q \leq r$. But then either $p = 1$ or it contains a subgraph of type $\tilde{E}_6$, a contradiction. If $p = 1$, then either $q \leq 2$ or it contains a subgraph of type $\tilde{E}_7$, another contradiction. Finally, with $p = 1$ and $q = 2$, either $r \geq 4$ or it contains a subgraph of type $\tilde{E}_8$, a contradiction again. Therefore, either $p = q = 1$ and we get $D_n$, or $p = 1$, $q = 2$ and $r = 2, 3$ or 4, thus obtaining $E_6$, $E_7$ and $E_8$.

Therefore, the only possible connected Coxeter graphs are those in Theorem 6.5. What remains to be proven is to show that for each Dynkin diagram $(A)$–$(G)$, there exists indeed an irreducible root system with this Dynkin diagram.

For types $(A)$–$(D)$ we will prove a stronger statement, since we will show that there are simple Lie algebras, over an algebraically closed field of characteristic 0, such that their Dynkin diagrams of their root systems relative to a Cartan subalgebra and a set of simple roots are precisely the Dynkin diagrams of types $(A)$–$(D)$.

$(A_n)$ Let $L = \mathfrak{sl}_{n+1}(k)$ be the Lie algebra of $n+1$ trace zero square matrices. Let $H$ be the subspace of diagonal matrices in $L$, which is an abelian subalgebra, and let $\epsilon_i : H \to k$ the linear form such that $\epsilon_i\bigl(\operatorname{diag}(\alpha_1, \dots, \alpha_{n+1})\bigr) = \alpha_i$, $i = 1, \dots, n+1$. Then $\epsilon_1 + \cdots + \epsilon_{n+1} = 0$. Moreover,

$$L = H \oplus \Bigl(\oplus_{1\le i\ne j\le n+1} kE_{ij}\Bigr) \tag{6.4}$$

where $E_{ij}$ is the matrix with a 1 in the $(i,j)$ entry, and 0's elsewhere. Since $[h, E_{ij}] = (\epsilon_i - \epsilon_j)(h)E_{ij}$ for any $i \ne j$, it follows that $H$ is toral and a Cartan subalgebra of $L$. It also follows easily that $L$ is simple (using that any ideal is invariant under the adjoint action of $H$) and that the set of roots of $L$ relative to $H$ is

$$\Phi = \{\epsilon_i - \epsilon_j : 1 \le i \ne j \le n+1\}.$$

The restriction of the Killing form to $H$ is determined by

$$\begin{aligned}\kappa(h,h) &= \sum_{1\le i\ne j\le n+1} (\alpha_i - \alpha_j)^2 = 2 \sum_{1\le i<j\le n+1} (\alpha_i^2 + \alpha_j^2 - 2\alpha_i\alpha_j) \\ &= 2(n+1) \sum_{1\le i\le n+1} \alpha_i^2 = 2(n+1)\operatorname{trace}(h^2)\end{aligned} \tag{6.5}$$

for any $h = \operatorname{diag}(\alpha_1, \dots, \alpha_{n+1}) \in H$, since $0 = (\alpha_1 + \cdots + \alpha_{n+1})^2 = \sum_{1\le i\le n+1} \alpha_i^2 + 2\sum_{1\le i<j\le n+1} \alpha_i\alpha_j$. Therefore, for any $i \ne j$, $t_{\epsilon_i - \epsilon_j} = \frac{1}{2(n+1)}(E_{ii} - E_{jj})$ and

$$\bigl(\epsilon_i - \epsilon_j | \epsilon_h - \epsilon_k\bigr) = (\epsilon_i - \epsilon_j)\bigl(t_{\epsilon_h - \epsilon_k}\bigr) = \frac{1}{2(n+1)}\bigl(\delta_{ih} - \delta_{jh} - \delta_{ik} + \delta_{jk}\bigr),$$

where $\delta_{ij}$ is the Kronecker symbol. Thus we get the euclidean vector space $E = \mathbb{R} \otimes_{\mathbb{Q}} \mathbb{Q}\Phi$ and can take the vector $\nu = n\epsilon_1 + (n-2)\epsilon_2 + \cdots + (-n)\epsilon_{n+1} = n(\epsilon_1 - \epsilon_{n+1}) + (n-2)(\epsilon_2 - \epsilon_n) + \cdots \in E$, which satisfies $\bigl(\nu | \epsilon_i - \epsilon_j\bigr) > 0$ if and only if $i < j$. For this $\nu$ we obtain the set of positive roots $\Phi^+ = \{\epsilon_i - \epsilon_j : 1 \le i < j \le n+1\}$ and the system of simple roots $\Delta = \{\epsilon_1 - \epsilon_2, \epsilon_2 - \epsilon_3, \dots, \epsilon_n - \epsilon_{n+1}\}$. The corresponding Dynkin diagram is $(A_n)$.

$(B_n)$ Consider the following 'orthogonal Lie algebra':

$$\begin{aligned} L &= \mathfrak{so}_{2n+1}(k) \\ &= \left\{ X \in \mathfrak{gl}_{2n+1}(k) : X^t \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_n \\ 0 & I_n & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_n \\ 0 & I_n & 0 \end{pmatrix} X = 0 \right\} \\ &= \left\{ \begin{pmatrix} 0 & -b^t & -a^t \\ a & A & B \\ b & C & -A^t \end{pmatrix} : a, b \in \operatorname{Mat}_{n\times 1}(k)\,, \right. \\ &\qquad\qquad \left. A, B, C \in \operatorname{Mat}_n(k),\ B^t = -B,\ C^t = -C \right\} \end{aligned} \tag{6.6}$$

where $I_n$ denotes the identity $n \times n$ matrix. Number the rows and columns of these matrices as $0, 1, \dots, n, \bar{1}, \dots, \bar{n}$ and consider the subalgebra $H$ consisting again of the diagonal matrices on $L$: $H = \{\operatorname{diag}(0, \alpha_1, \dots, \alpha_n, -\alpha_1, \dots, -\alpha_n) :$

$\alpha_i \in k,\, i = 1, \ldots, n\}$. Again we get the linear forms $\epsilon_i : H \to k$, such that $\epsilon_i\big(\operatorname{diag}(0, \alpha_1, \ldots, \alpha_n, -\alpha_1, \ldots, -\alpha_n)\big) = \alpha_i$, $i = 1, \ldots, n$. Then,

$$\begin{aligned} L &= H \oplus \big(\oplus_{i=1}^n k(E_{0i} - E_{\bar{i}0})\big) \oplus \big(\oplus_{i=1}^n k(E_{0\bar{i}} - E_{i0})\big) \oplus \big(\oplus_{1\le i\ne j\le n} k(E_{ij} - E_{\bar{j}\bar{i}})\big) \\ &\qquad \oplus \big(\oplus_{1\le i<j\le n} k(E_{i\bar{j}} - E_{j\bar{i}})\big) \oplus \big(\oplus_{1\le i<j\le n} k(E_{\bar{i}j} - E_{\bar{j}i})\big) \\ &= H \oplus \big(\oplus_{i=1}^n L_{-\epsilon_i}\big) \oplus \big(\oplus_{i=1}^n L_{\epsilon_i}\big) \oplus \big(\oplus_{1\le i\ne j\le n} L_{\epsilon_i - \epsilon_j}\big) \\ &\qquad \oplus \big(\oplus_{1\le i<j\le n} L_{\epsilon_i+\epsilon_j}\big) \oplus \big(\oplus_{1\le i<j\le n} L_{-(\epsilon_i+\epsilon_j)}\big), \end{aligned}$$

where $L_\alpha = \{x \in L : [h, x] = \alpha(h)x \ \forall h \in H\}$. It follows easily from here that $H$ is a Cartan subalgebra of $L$, that $L$ is simple and that the set of roots is

$$\Phi = \{\pm\epsilon_i, \pm\epsilon_i \pm \epsilon_j : 1 \le i < j \le n\}.$$

Also, for any $h \in H$ as above,

$$\begin{aligned} \kappa(h,h) &= 2(\alpha_1^2 + \cdots + \alpha_n^2) + \sum_{1\le i\ne j\le n} (\alpha_i - \alpha_j)^2 + 2\sum_{1\le i<j\le n} (\alpha_i + \alpha_j)^2 \\ &= 2(\alpha_1^2 + \cdots + \alpha_n^2) + 2\sum_{1\le i<j\le n} \Big((\alpha_i - \alpha_j)^2 + (\alpha_i + \alpha_j)^2\Big) \\ &= \big(2 + 4(n-1)\big)(\alpha_1^2 + \cdots + \alpha_n^2) = 2(2n-1)(\alpha_1^2 + \cdots + \alpha_n^2) \\ &= (2n-1)\operatorname{trace}(h^2). \end{aligned} \tag{6.7}$$

Therefore, $t_{\epsilon_i} = \frac{1}{2(2n-1)}(E_{ii} - E_{\bar{i}\bar{i}})$ and $\big(\epsilon_i|\epsilon_j\big) = \epsilon_i(t_{\epsilon_j}) = \frac{1}{2(2n-1)}\delta_{ij}$. We can take the element $\nu = n\epsilon_1 + (n-1)\epsilon_2 + \cdots + \epsilon_n$, whose inner product with any root is never 0 and gives $\Phi^+ = \{\epsilon_i, \epsilon_i \pm \epsilon_j : 1 \le i < j \le n\}$ and system of simple roots $\Delta = \{\epsilon_1 - \epsilon_2, \epsilon_2 - \epsilon_3, \ldots, \epsilon_{n-1} - \epsilon_n, \epsilon_n\}$. The associated Dynkin diagram is $(B_n)$.

**6.8 Exercise.** Prove that $\mathfrak{so}_3(k)$ is isomorphic to $\mathfrak{sl}_2(k)$. ($k$ being algebraically closed.)

$(C_n)$ Consider now the 'symplectic Lie algebra':

$$\begin{aligned} L &= \mathfrak{sp}_{2n}(k) \\ &= \left\{ X \in \mathfrak{gl}_{2n}(k) : X^t \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} + \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} X = 0 \right\} \\ &= \left\{ \begin{pmatrix} A & B \\ C & -A^t \end{pmatrix} : A, B, C \in \operatorname{Mat}_n(k),\ B^t = B,\ C^t = C \right\} \end{aligned} \tag{6.8}$$

where $n \ge 2$ (for $n = 1$ we get $\mathfrak{sp}_2(k) = \mathfrak{sl}_2(k)$). Number the rows and columns as $1, \ldots, n, \bar{1}, \ldots, \bar{n}$. As before, the subspace $H$ of diagonal matrices is a Cartan subalgebra with set of roots

$$\Phi = \{\pm 2\epsilon_i, \pm\epsilon_i \pm \epsilon_j : 1 \le i < j \le n\}$$

where $\epsilon_i(h) = \alpha_i$ for any $i$, with $h = \mathrm{diag}(\alpha_1, \dots, \alpha_n, -\alpha_1, \dots, -\alpha_n)$. Here

$$\begin{aligned}
\kappa(h,h) &= 2\sum_{i=1}^{n} 4\alpha_i^2 + \sum_{1\le i\ne j\le n} (\alpha_i-\alpha_j)^2 + 2\sum_{1\le i<j\le n} (\alpha_i+\alpha_j)^2 \\
&= 8(\alpha_1^2+\cdots+\alpha_n^2) + 2\sum_{1\le i<j\le n}\Big((\alpha_i-\alpha_j)^2+(\alpha_i+\alpha_j)^2\Big) \\
&= \big(8+4(n-1)\big)(\alpha_1^2+\cdots+\alpha_n^2) = 4(n+1)(\alpha_1^2+\cdots+\alpha_n^2) \\
&= 2(n+1)\,\mathrm{trace}(h^2),
\end{aligned} \tag{6.9}$$

$t_{\epsilon_i} = \frac{1}{4n}(E_{ii} - E_{\bar{i}\bar{i}})$, $\big(\epsilon_i|\epsilon_j\big) = \frac{1}{4n}\delta_{ij}$. Besides, we can take $\nu = n\epsilon_1 + (n-1)\epsilon_2 + \cdots + \epsilon_n$, which gives $\Phi^+ = \{2\epsilon_i, \epsilon_i \pm \epsilon_j : 1 \le i < j \le n\}$ and $\Delta = \{\epsilon_1 - \epsilon_2, \dots, \epsilon_{n-1} - \epsilon_n, 2\epsilon_n\}$, whose associated Dynkin diagram is $(C_n)$.

$(D_n)$ Finally, consider the 'orthogonal Lie algebra':

$$\begin{aligned}
L &= \mathfrak{so}_{2n}(k) \\
&= \left\{ X \in \mathfrak{gl}_{2n}(k) : X^t \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix} + \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix} X = 0 \right\} \\
&= \left\{ \begin{pmatrix} A & B \\ C & -A^t \end{pmatrix} : A, B, C \in \mathrm{Mat}_n(k),\ B^t = -B,\ C^t = -C \right\}
\end{aligned} \tag{6.10}$$

with $n \ge 4$. Number the rows and columns as $1, \dots, n, \bar{1}, \dots, \bar{n}$. As it is always the case, the subspace $H$ of diagonal matrices is a Cartan subalgebra with set of roots

$$\Phi = \{\pm\epsilon_i \pm \epsilon_j : 1 \le i < j \le n\}$$

where $\epsilon_i(h) = \alpha_i$ for any $i$, with $h = \mathrm{diag}(\alpha_1, \dots, \alpha_n, -\alpha_1, \dots, -\alpha_n)$. Here

$$\begin{aligned}
\kappa(h,h) &= \sum_{1\le i\ne j\le n} (\alpha_i-\alpha_j)^2 + 2\sum_{1\le i<j\le n} (\alpha_i+\alpha_j)^2 \\
&= 4(n-1)(\alpha_1^2+\cdots+\alpha_n^2) \\
&= 2(n-1)\,\mathrm{trace}(h^2),
\end{aligned} \tag{6.11}$$

$t_{\epsilon_i} = \frac{1}{4(n-1)}(E_{ii} - E_{\bar{i}\bar{i}})$, $\big(\epsilon_i|\epsilon_j\big) = \frac{1}{4(n-1)}\delta_{ij}$. Also, we can take $\nu = n\epsilon_1 + (n-1)\epsilon_2 + \cdots + \epsilon_n$, which gives $\Phi^+ = \{\epsilon_i \pm \epsilon_j : 1 \le i < j \le n\}$ and $\Delta = \{\epsilon_1 - \epsilon_2, \dots, \epsilon_{n-1} - \epsilon_n, \epsilon_{n-1} + \epsilon_n\}$, whose associated Dynkin diagram is $(D_n)$.

The remaining Dynkin diagrams correspond to the so called *exceptional* simple Lie algebras, whose description is more involved. Hence, we will proceed in a different way:

$(E_8)$ Let $E = \mathbb{R}^8$ with the canonical inner product $(\,.\,|\,.\,)$ and canonical orthonormal basis $\{e_1, \dots, e_8\}$. Take $e_0 = \frac{1}{2}(e_1 + \cdots + e_8)$ and $Q = \{m_0 e_0 + \sum_{i=1}^{8} m_i e_i : m_i \in \mathbb{Z}\,\forall i,\ \sum_{i=1}^{8} m_i \in 2\mathbb{Z}\}$, which is an additive subgroup of $\mathbb{R}^8$. Consider the set

$$\Phi = \{v \in Q : (v|v) = 2\}.$$

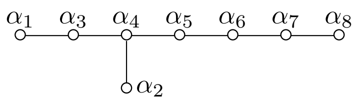
For $v = \sum_{i=0}^{8} m_i e_i \in Q$, $(v|v) = \sum_{i=1}^{8}(m_i + \frac{1}{2}m_0)^2$, so if $m_0$ is even, then $m_i + \frac{1}{2}m_0 \in \mathbb{Z}$ for any $i$ and the only possibilities for $v$ to belong to $\Phi$ are $v = \pm e_i \pm e_j$, $1 \leq i < j \leq 8$. On the other hand, if $m_0$ is odd, then $m_i + \frac{1}{2}m_0 \in \frac{1}{2} + \mathbb{Z}$ for any $i$ and the only possibilities are $v = \frac{1}{2}(\pm e_1 \pm e_2 \pm \cdots \pm e_8)$. Moreover, since $\sum_{i=1}^{8} m_i$ must be even, the number of $+$ signs in the previous expression must be even. In particular, $\Phi$ satisfies the restrictions (R1) and (R2) of the definition of root system.

Besides, for any $v \in \Phi$, $(v|v) = 2$ and for any $v, w \in \Phi$, $\langle v|w\rangle = \dfrac{2(v|w)}{(w|w)} = (v|w)$ is easily shown to be in $\mathbb{Z}$, hence (R4) is satisfied too. The proof that (R3) is satisfied is a straightforward computation. Thus, $\Phi$ is a root system.

Take now $\nu = \sum_{i=1}^{8} 2^i e_i$, then $(\nu|\alpha) \neq 0$ for any $\alpha \in \Phi$. The associated set of positive roots is $\Phi^+ = \{\frac{1}{2}(\pm e_1 \pm e_2 \pm \cdots \pm e_7 + e_8), \pm e_i + e_j : i < j\}$, and the set of simple roots is

$$\begin{aligned}\Delta = \big\{&\alpha_1 = \tfrac{1}{2}(e_1 - e_2 - e_3 - e_4 - e_5 - e_6 - e_7 + e_8), \alpha_2 = e_1 + e_2, \alpha_3 = e_2 - e_1,\\ &\alpha_4 = e_3 - e_2, \alpha_5 = e_4 - e_3, \alpha_6 = e_5 - e_4, \alpha_7 = e_6 - e_5, \alpha_8 = e_7 - e_6\big\}\end{aligned}$$

with associated Dynkin diagram




of type $(E_8)$.

$(E_7)$ **and** $(E_6)$ These are obtained as the 'root subsystems' of $(E_8)$ generated by $\Delta \setminus \{\alpha_8\}$ and $\Delta \setminus \{\alpha_7, \alpha_8\}$ above.

$(F_4)$ Here consider the euclidean vector space $E = \mathbb{R}^4$, $e_0 = \frac{1}{2}(e_1 + e_2 + e_3 + e_4)$, $Q = \{m_0 e_0 + \sum_{i=1}^{4} m_i e_i : m_i \in \mathbb{Z}\}$, and

$$\Phi = \{v \in Q : (v|v) = 1 \text{ or } 2\} = \{\pm e_i, \pm e_i \pm e_j\ (i < j),\ \tfrac{1}{2}(\pm e_1 \pm e_2 \pm e_3 \pm e_4)\}.$$

This is a root system and with $\nu = 8e_1 + 4e_2 + 2e_3 + e_4$ one obtains $\Phi^+ = \{e_i, e_i \pm e_j\ (i < j),\ \frac{1}{2}(e_1 \pm e_2 \pm e_3 \pm e_4)\}$ and

$$\Delta = \{e_2 - e_3, e_3 - e_4, e_4, \tfrac{1}{2}(e_1 - e_2 - e_3 - e_4)\},$$

with associated Dynkin graph $(F_4)$.

$(G_2)$ In the euclidean vector space $E = \{(\alpha, \beta, \gamma) \in \mathbb{R}^3 : \alpha + \beta + \gamma = 0\} = \mathbb{R}(1,1,1)^\perp$, with the restriction of the canonical inner product on $\mathbb{R}^3$, consider the subset $Q = \{m_1 e_1 + m_2 e_2 + m_3 e_3 : m_i \in \mathbb{Z},\ m_1 + m_2 + m_3 = 0\}$, and

$$\begin{aligned}\Phi &= \{v \in Q : (v|v) = 2 \text{ or } 6\}\\ &= \{\pm(e_i - e_j)\,(i < j), \pm(2e_1 - e_2 - e_3), \pm(-e_1 + 2e_2 - e_3), \pm(-e_1 - e_2 + 2e_3)\}.\end{aligned}$$

Again, $\Phi$ is a root system, and with $\nu = -2e_1 - e_2 + 3e_3$, $\Phi^+ = \{e_i - e_j\,(i > j),\ -2e_1 + e_2 + e_3,\ e_1 - 2e_2 + e_3,\ -e_1 - e_2 + 2e_3\}$ and

$$\Delta = \{e_2 - e_1, e_1 - 2e_2 + e_3\},$$

with associated Dynkin diagram of type $(G_2)$.

This finishes the classification of the connected Dynkin diagrams. To obtain from this classification a classification of the root systems, it is enough to check that any root system is determined by its Dynkin diagram.

**6.9 Definition.** Let $\Phi_i$ be a root system in the euclidean space $E_i$, $i = 1, 2$, and let $\varphi : E_1 \to E_2$ be a linear map. Then $\varphi$ is said to be a *root system isomorphism* between $\Phi_1$ and $\Phi_2$ if $\varphi(\Phi_1) = \Phi_2$ and for any $\alpha, \beta \in \Phi_1$, $\langle \varphi(\alpha)|\varphi(\beta)\rangle = \langle \alpha|\beta\rangle$.

**6.10 Exercise.** Prove that if $\varphi$ is a root system isomorphism between the irreducible root systems $\Phi_1$ and $\Phi_2$, then $\varphi$ is a similarity of multiplier $\dfrac{\big(\varphi(\alpha)|\varphi(\alpha)\big)}{(\alpha|\alpha)}$ for a fixed $\alpha \in \Phi_1$.

The next result is already known for roots that appear inside the semisimple Lie algebras over algebraically closed fields of characteristic 0, because of the representation theory of $\mathfrak{sl}_2(k)$.

**6.11 Lemma.** *Let $\Phi$ be a root system, $\alpha, \beta \in \Phi$ two roots such that $\beta \neq \pm\alpha$, let $r = \max\{i \in \mathbb{Z}_{\geq 0} : \beta - i\alpha \in \Phi\}$ and $q = \max\{i \in \mathbb{Z}_{\geq 0} : \beta + i\alpha \in \Phi\}$. Then $\langle \beta|\alpha\rangle = r - q$, $r + q \leq 3$ and all the elements in the chain $\beta - r\alpha, \beta - (r-1)\alpha, \dots, \beta, \dots, \beta + q\alpha$ belong to $\Phi$ (this is called the* $\alpha$-chain *of $\beta$).*

*Proof.* Take $\gamma = \beta + q\alpha \in \Phi$, then $\langle \gamma|\alpha\rangle = \langle \beta|\alpha\rangle + 2q$. Besides, $\gamma + i\alpha \notin \Phi$ for any $i \in \mathbb{Z}_{>0}$, $\gamma - (r+q)\alpha \in \Phi$, and $\gamma - (r+q+i)\alpha \notin \Phi$ for any $i \in \mathbb{Z}_{>0}$.

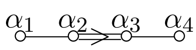
Then $\sigma_\alpha(\gamma) = \gamma - \langle \gamma|\alpha\rangle\alpha \in \Phi$, so $\langle \gamma|\alpha\rangle \leq r + q$; while $\sigma_\alpha\big(\gamma - (r+q)\alpha\big) = \gamma - \langle \gamma|\alpha\rangle\alpha + (r+q)\alpha \in \Phi$, so $r + q - \langle \gamma|\alpha\rangle \leq 0$, or $\langle \gamma|\alpha\rangle \geq r + q$. We conclude that $\langle \gamma|\alpha\rangle = r + q$ and this is $\leq 3$ by the argument in the proof of Proposition 6.1. Besides, $\langle \beta|\alpha\rangle = \langle \gamma|\alpha\rangle - 2q = r - q$.

Thus, $\langle \gamma|\alpha\rangle = 0, 1, 2$ or 3. If $\langle \gamma|\alpha\rangle = 0$, then the $\alpha$-chain of $\beta$ consists only of $\gamma = \beta \in \Phi$. If $\langle \gamma|\alpha\rangle = 1$, then the $\alpha$-chain consists of $\gamma \in \Phi$ and $\gamma - \alpha = \sigma_\alpha(\gamma) \in \Phi$. If $\langle \gamma|\alpha\rangle = 2$, then $\langle \alpha|\gamma\rangle = 1$ and the $\alpha$-chain consists of $\gamma \in \Phi$, $\gamma - \alpha = -\sigma_\gamma(\alpha) \in \Phi$ and $\gamma - 2\alpha = \sigma_\alpha(\gamma) \in \Phi$. Finally, if $\langle \gamma|\alpha\rangle = 3$, then again $\langle \alpha|\gamma\rangle = 1$ and the $\alpha$-chain consists of $\gamma \in \Phi$, $\gamma - \alpha = \sigma_\gamma(\alpha)$, $\gamma - 2\alpha = \sigma_\alpha(\gamma - \alpha) \in \Phi$, and $\gamma - 3\alpha = \sigma_\alpha(\gamma) \in \Phi$. □

**6.12 Theorem.** *Each Dynkin diagram determines a unique (up to isomorphism) root system.*

*Proof.* First note that it is enough to assume that the Dynkin diagram is connected. We will do it.

Let $\Delta$ be the set of nodes of the Dynkin diagram and fix arbitrarily the length of a 'short node'. Then the diagram determines the inner product on $E = \mathbb{R}\Phi = \mathbb{R}\Delta$. This is better seen with an example. Take, for instance the Dynkin diagram $(F_4)$, so we have $\Delta = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, with




Fix, for simplicity, $(\alpha_3|\alpha_3) = 2 = (\alpha_4|\alpha_4)$. Then

- $-1 = \langle \alpha_3|\alpha_4\rangle = \dfrac{2(\alpha_3|\alpha_4)}{(\alpha_4|\alpha_4)}$, so $(\alpha_3|\alpha_4) = -1$,

- $-2 = \langle\alpha_2|\alpha_3\rangle = \dfrac{2(\alpha_2|\alpha_3)}{(\alpha_3|\alpha_3)}$, so $(\alpha_2|\alpha_3) = -2$.
- $-1 = \langle\alpha_3|\alpha_2\rangle = \dfrac{2(\alpha_3|\alpha_2)}{(\alpha_2|\alpha_2)}$, so $(\alpha_2|\alpha_2) = 4 = (\alpha_1|\alpha_1)$.
- $-1 = \langle\alpha_1|\alpha_2\rangle$, so $(\alpha_1|\alpha_2) = -1$.

Since $\Delta$ is a basis of $E$, the inner product is completely determined up to a nonzero positive scalar (the arbitrary length we have imposed on the short roots of $\Delta$). For any other connected Dynkin diagram, the argument is the same.

Now, with $\Delta = \{\alpha_1, \dots, \alpha_n\}$, any $\alpha \in \Phi^+$ appears as $\alpha = \sum_{i=1}^n m_i\alpha_i$ with $m_i \in \mathbb{Z}_{\geq 0}$. Define the *height* of $\alpha$ as $\mathrm{ht}(\alpha) = m_1 + \cdots + m_n$. It is enough to prove that for any $N \in \mathbb{N}$, the subset $\{\alpha \in \Phi^+ : \mathrm{ht}(\alpha) = N\}$ is determined by the Dynkin diagram, and this is done by induction on $N$:

For $N = 1$ this is obvious, since $\mathrm{ht}(\alpha) = 1$ if and only if $\alpha \in \Delta$.

Assume that the result is valid for $1, \dots, N$. Then it is enough to prove that the roots of height $N+1$ are precisely the vectors $\gamma = \beta + \alpha$, with $\mathrm{ht}(\beta) = N$, $\alpha \in \Delta$ and such that $\langle\beta|\alpha\rangle < r$ with $r = \max\{i \in \mathbb{Z}_{\geq 0} : \beta - i\alpha \in \Phi^+\}$. Note that the height of the roots $\beta - i\alpha \in \Phi^+$, with $i \geq 0$, is at most $N$, and hence all these roots are determined by $\Delta$. Actually, if $\beta \in \Phi$ and $\alpha \in \Delta$ satisfy these conditions, then $r > \langle\beta|\alpha\rangle = r - q$ by the Lemma, so $q \geq 1$, and $\beta + \alpha$ is in the $\alpha$-chain of $\beta$, and hence it is a root. Conversely, let $\gamma = \sum_{i=1}^n m_i\alpha_i$ be a root of height $N+1$. Then $0 < (\gamma|\gamma) = \sum_{i=1}^n m_i(\gamma|\alpha_i)$, so there is an $i$ with $(\gamma|\alpha_i) > 0$ and $m_i > 0$. From the previous Lemma we know that $\beta = \gamma - \alpha_i \in \Phi$, and $\mathrm{ht}(\beta) = N$. Besides, $\beta + \alpha_i \in \Phi$, so $q \geq 1$ in the previous Lemma, and hence $r - q = \langle\beta|\alpha_i\rangle < r$, as required. □

**6.13 Remark.** Actually, the proof of this Theorem gives an algorithm to obtain a root system $\Phi$, starting with its Dynkin diagram.

**6.14 Exercise.** Use this algorithm to obtain the root system associated to the Dynkin diagram $(G_2)$.

**6.15 Exercise.** Let $\Phi$ a root system and let $\Delta = \{\alpha_1, \dots, \alpha_n\}$ be a system of simple roots of $\Phi$. Let $\alpha = m_1\alpha_1 + \cdots + m_n\alpha_n$ be a positive root of maximal height and consider $\Delta_1 = \{\alpha_i : m_i \neq 0\}$ and $\Delta_2 = \Delta \setminus \Delta_1$. Prove that $\big(\Delta_1|\Delta_2\big) = 0$.
In particular, if $\Phi$ is irreducible this shows that $\alpha$ “involves” all the simple roots ($\Delta = \Delta_1$).

# § 7. Classification of the semisimple Lie algebras

Throughout this section, the ground field $k$ will be assumed to be algebraically closed of characteristic 0.

The aim here is to show that each root system $\Phi$ determines, up to isomorphism, a unique semisimple Lie algebra over $k$.

Let $L = H \oplus \big(\oplus_{\alpha\in\Phi} L_\alpha\big)$ be the root space decomposition of a semisimple Lie algebra over $k$, relative to a Cartan subalgebra $H$. We want to prove that the multiplication in $L$ is determined by $\Phi$.

For any $\alpha \in \Phi$, there are elements $x_\alpha \in L_\alpha$, $y_\alpha \in L_{-\alpha}$ such that $[x_\alpha, y_\alpha] = h_\alpha$, with $\alpha(h_\alpha) = 2$. Besides, $L_\alpha = kx_\alpha$, $L_{-\alpha} = ky_\alpha$ and $S_\alpha = L_\alpha \oplus L_{-\alpha} \oplus [L_\alpha, L_{-\alpha}] = kx_\alpha \oplus ky_\alpha \oplus kh_\alpha$ is a subalgebra isomorphic to $\mathfrak{sl}_2(k)$. Also, for any $\beta \in \Phi \setminus \{\pm\alpha\}$, recall that the $\alpha$-chain of $\beta$ consists of roots $\beta - r\alpha, \dots, \beta, \dots, \beta + q\alpha$, where $\langle \beta|\alpha\rangle = r - q$.

**7.1 Lemma.** *Under the hypotheses above, let $\alpha, \beta \in \Phi$ with $\alpha + \beta \in \Phi$, then $[L_\alpha, L_\beta] = L_{\alpha+\beta}$. Moreover, for any $x \in L_\beta$,*

$$\begin{cases} [y_\alpha, [x_\alpha, x]] = q(r+1)x, \\ [x_\alpha, [y_\alpha, x]] = r(q+1)x. \end{cases}$$

*Proof.* This is a straightforward consequence of the representation theory of $\mathfrak{sl}_2(k)$, since $\oplus_{i=-r}^{q} L_{\beta+i\alpha}$ is a module for $S_\alpha \cong \mathfrak{sl}_2(k)$. Hence, there are elements $v_i \in L_{\beta+(q-i)\alpha}$, $i = 0, \dots, r+q$, such that $[y_\alpha, v_i] = v_{i+1}$, $[x_\alpha, v_i] = i(r+q+1-i)v_{i-1}$, with $v_{-1} = v_{r+q+1} = 0$ (see the proof of Theorem 3.2); whence the result. $\square$

Let $\Delta = \{\alpha_1, \dots, \alpha_n\}$ be a system of simple roots of $\Phi$. For any $i = 1, \dots, n$, let $x_i = x_{\alpha_i}$, $y_i = y_{\alpha_i}$ and $h_i = h_{\alpha_i}$. For any $\alpha \in \Phi^+$, the proof of Theorem 6.12 shows that $\alpha$ is a sum of simple roots: $\alpha = \alpha_{i_1} + \dots + \alpha_{i_r}$, with $\alpha_{i_1} + \dots + \alpha_{i_j} \in \Phi^+$ for any $j = 1, \dots, r = \mathrm{ht}(\alpha)$. For any $\alpha \in \Phi^+$ we fix one such sequence $I_\alpha = (i_1, \dots, i_r)$ and take $x_\alpha = \operatorname{ad} x_{i_r} \cdots \operatorname{ad} x_{i_2}\big(x_{i_1}\big)$ and $y_\alpha = \operatorname{ad} y_{i_r} \cdots \operatorname{ad} y_{i_2}\big(y_{i_1}\big)$. These elements are nonzero by the previous Lemma, and hence $L_\alpha = kx_\alpha$ and $L_{-\alpha} = ky_\alpha$.

**7.2 Lemma.** *For any $\alpha \in \Phi^+$, let $J = J_\alpha = (j_1, \dots, j_r)$ be another sequence such that $\alpha = \alpha_{j_1} + \dots + \alpha_{j_r}$, and let $x_J = \operatorname{ad} x_{j_r} \cdots \operatorname{ad} x_{j_2}\big(x_{j_1}\big)$ and $y_J = \operatorname{ad} y_{j_r} \cdots \operatorname{ad} y_{j_2}\big(y_{j_1}\big)$. Then there are rational numbers $q, q' \in \mathbb{Q}$, determined by $\Phi$, such that $x_J = qx_\alpha$, $y_J = q'y_\alpha$.*

*Proof.* Since $x_J \in L_\alpha$, the previous Lemma shows that $x_J = q_1[x_{i_r}, [y_{i_r}, x_J]]$, for some $q_1 \in \mathbb{Q}$ which depends on $\Phi$. Let $s$ be the largest integer with $j_s = i_r$, then

$$\begin{aligned} [y_{i_r}, x_J] &= \operatorname{ad} x_{j_r} \cdots \operatorname{ad} x_{j_{s+1}} \operatorname{ad} y_{i_r} \operatorname{ad} x_{j_s}(x_K) \\ &\qquad \text{(where } K = (j_1, \dots, j_{s-1})\text{, since } [y_i, x_j] = 0 \text{ for any } i \neq j) \\ &= q_2 \operatorname{ad} x_{j_r} \cdots \operatorname{ad} x_{j_{s+1}}(x_K) \quad \text{(by the previous Lemma)} \\ &= q_2 q_3 x_{I'} \quad \text{(by induction on } r = \mathrm{ht}(\alpha)\text{)}, \end{aligned}$$

where $q_2, q_3 \in \mathbb{Q}$ depend on $\Phi$ and $I' = (i_1, \dots, i_{r-1})$. Therefore, $x_J = q_1 q_2 q_3 [x_{i_r}, x_{I'}] = q_1 q_2 q_3 x_\alpha$, with $q_1, q_2, q_3 \in \mathbb{Q}$ determined by $\Phi$. The proof for $y_J$ is similar. $\square$

Hence, we may consider the following basis for $L$: $\mathcal{B} = \{h_1, \dots, h_n, x_\alpha, y_\alpha : \alpha \in \Phi^+\}$, with the $x_\alpha$'s and $y_\alpha$'s chosen as above.

**7.3 Proposition.** *The product of any two elements in $\mathcal{B}$ is a rational multiple of another element of $\mathcal{B}$, determined by $\Phi$, with the exception of the products $[x_\alpha, y_\alpha]$, which are linear combinations of the $h_i$'s, with rational coefficients determined by $\Phi$.*

*Proof.* First note that $[h_i, h_j] = 0$, $[h_i, x_\alpha] = \alpha(h_i)x_\alpha = \langle\alpha|\alpha_i\rangle x_\alpha$ and $[h_i, y_\alpha] = -\langle\alpha|\alpha_i\rangle y_\alpha$, are all determined by $\Phi$.

Consider now $\alpha, \beta \in \Phi^+$, and the corresponding fixed sequences $I_\alpha = (i_1, \dots, i_r)$, $I_\beta = (j_1, \dots, j_s)$.

To deal with the product $[x_\alpha, x_\beta]$, let us argue by induction on $r$. If $r = 1$, $[x_\alpha, x_\beta] = 0$ if $\alpha + \beta \notin \Phi$, while $[x_\alpha, x_\beta] = qx_{\alpha+\beta}$ for some $q \in \mathbb{Q}$ determined by $\Phi$ by the previous Lemma. On the other hand, if $r > 1$ and $I'_\alpha = (i_1, \dots, i_{r-1})$, then $[x_\alpha, x_\beta] = [[x_{i_r}, x_{I'_\alpha}], x_\beta] = [x_{i_r}, [x_{I'_\alpha}, x_\beta]] - [x_{I'_\alpha}, [x_{i_r}, x_\beta]]$ and now the induction hypothesis and the previous Lemma yield the result. The same arguments apply to products $[y_\alpha, y_\beta]$.

Finally, we will argue by induction on $r$ too to deal with the product $[x_\alpha, y_\beta]$. If $r = 1$ and $\alpha = \alpha_i$, then $[x_\alpha, y_\beta] = 0$ if $0 \neq \beta - \alpha \notin \Phi$, $[x_\alpha, y_\beta] = h_i$ if $\alpha = \beta$, while if $\beta - \alpha = \gamma \in \Phi$, then $y_\beta = q[y_i, y_\gamma]$ for some $q \in \mathbb{Q}$ determined by $\Phi$, and $[x_\alpha, y_\beta] = q[x_i, [y_i, y_\gamma]] = qq'y_\gamma$, determined by $\Phi$. On the other hand, if $r > 1$ then, as before, $[x_\alpha, y_\beta] = [x_{i_r}, [x_{I'_\alpha}, y_\beta]] - [x_{I'_\alpha}, [x_{i_r}, y_\beta]]$ and the induction hypothesis applies. □

What remains to be done is, on one hand, to show that for each of the irreducible root systems $E_6, E_7, E_8, F_4, G_2$ there is a simple Lie algebra $L$ over $k$ and a Cartan subalgebra $H$ such that the corresponding root system is of this type. Since we have constructed explicitly these root systems, the dimension of such an $L$ must be $|\Phi| + \operatorname{rank}(\Phi)$, so $\dim_k L = 78, 133, 248, 52$ and $14$ respectively. Later on, some explicit constructions of these algebras will be given.

On the other hand, given a simple Lie algebra $L$ over $k$ and two Cartan subalgebras $H_1$ and $H_2$, it must be shown that the corresponding root systems $\Phi_1$ and $\Phi_2$ are isomorphic. The next Theorem solves this question:

**7.4 Theorem.** *Let $L$ be one of the Lie algebras $\mathfrak{sl}_n(k)$ ($n \geq 2$), $\mathfrak{so}_n(k)$ ($n \geq 3$), or $\mathfrak{sp}_{2n}(k)$ ($n \geq 1$), and let $H$ be any Cartan subalgebra of $L$. Then there is an element $g$ of the matrix group $GL_n(k)$, $O_n(k)$ or $Sp_{2n}(k)$ respectively, such that $gHg^{-1}$ is the subspace of diagonal matrices in $L$. In particular, for any two Cartan subalgebras of $L$, there is an automorphism $\varphi \in \operatorname{Aut}(L)$ such that $\varphi(H_1) = H_2$.*

*The last assertion is valid too for the simple Lie algebras containing a Cartan subalgebra such that the associated root system is exceptional.*

*Proof.* For the first part, let $V$ be the 'natural' module for $L$ ($V = k^n$ (column vectors) for $\mathfrak{sl}_n(k)$ or $\mathfrak{so}_n(k)$, and $V = k^{2n}$ for $\mathfrak{sp}_{2n}(k)$). Since $H$ is toral and abelian, the elements of $H$ form a commuting space of diagonalizable endomorphisms of $V$. Therefore there is a simultaneous diagonalization: $V = \oplus_{\lambda \in H^*} V_\lambda$, where $V_\lambda = \{v \in V : h.v = \lambda(h)v \,\forall h \in H\}$.

If $L = \mathfrak{sl}_n(k)$, then this means that there is an element $g \in GL_n(k)$ such that $gHg^{-1} \subseteq \{\text{diagonal matrices}\}$. Now, the map $x \mapsto gxg^{-1}$ is an automorphism of $L$ and hence $gHg^{-1}$ is a Cartan subalgebra too, in particular it is a maximal toral subalgebra. Since the set of diagonal matrices in $L$ is a Cartan subalgebra too, we conclude by maximality that $gHg^{-1}$ coincides with the space of diagonal matrices in $L$.

If $L = \mathfrak{so}_n(k)$ or $L = \mathfrak{sp}_{2n}(k)$, there is a nondegenerate symmetric or skew symmetric bilinear form $b : V \times V \to k$ such that (by its own definition) $L = \{x \in \mathfrak{gl}(V) : b(x.v, w) + b(v, x.w) = 0 \;\forall v, w \in V\}$. But then, for any $h \in H$, $\lambda, \nu \in H^*$ and $v \in V_\lambda$, $w \in V_\mu$, $0 = b(h.v, w) + b(v, h.w) = \big(\lambda(h) + \mu(h)\big)b(v, w)$. Hence we conclude that

$b(V_\lambda, V_\mu) = 0$ unless $\lambda = -\mu$. This implies easily the existence of a basis of $V$ consisting of common eigenvectors for $H$ in which the coordinate matrix of $b$ is either

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_n \\ 0 & I_n & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}$$

according to $L$ being $\mathfrak{so}_{2n+1}(k)$, $\mathfrak{sp}_{2n}(k)$ or $\mathfrak{so}_{2n}(k)$. Therefore, there is a $g \in SO_{2n+1}(k)$, $Sp_{2n}(k)$ or $SO_{2n}(k)$ (respectively) such that $gHg^{-1}$ is contained in the space of diagonal matrices of $L$. As before, we conclude that $gHg^{-1}$ fills this space.

Finally, let $L$ be a simple Lie algebra with a Cartan subalgebra $H$ such that the associated root system $\Phi$ is exceptional. Let $H'$ be another Cartan subalgebra and $\Phi'$ the associated root system. If $\Phi'$ were classical, then Proposition 7.3 would show that $L$ is isomorphic to one of the simple classical Lie algebras, and by the first part of the proof, there would exist an automorphism of $L$ taking $H'$ to $H$, so that $\Phi$ would be classical too, a contradiction. Hence $\Phi'$ is exceptional, and hence the fact that $\dim_k L = |\Phi| + \operatorname{rank}(\Phi)$, and the same for $\Phi'$, shows that $\Phi$ and $\Phi'$ are isomorphic. But by Proposition 7.3 again, we can choose bases $\{h_1, \ldots, h_n, x_\alpha, y_\alpha : \alpha \in \Phi\}$ and $\{h'_1, \ldots, h'_n, x'_\alpha, y'_\alpha : \alpha \in \Phi'\}$ with the same multiplication table. Therefore, there is an automorphism $\varphi$ of $L$ such that $\varphi(h_i) = h'_i$, $\varphi(x_\alpha) = x'_\alpha$ and $\varphi(y_\alpha) = y'_\alpha$, for any $i = 1, \ldots, n$ and $\alpha \in \Phi$. In particular, $\varphi(H) = H'$. $\square$

**7.5 Remark.** There is a more general classical result which asserts that if $H_1$ and $H_2$ are any two Cartan subalgebras of an arbitrary Lie algebra over $k$, then there is an automorphism $\varphi$, in the subgroup of the automorphism group generated by $\{\exp \operatorname{ad} x : x \in L, \operatorname{ad} x \text{ nilpotent}\}$ such that $\varphi(H_1) = H_2$. For an elementary (not easy!) proof, you may consult the article by A.A. George Michael: *On the conjugacy theorem of Cartan subalgebras*, Hiroshima Math. J. **32** (2002), 155-163.

The dimension of any Cartan subalgebra is called the *rank* of the Lie algebra.

Summarizing all the work done so far, and assuming the existence of the exceptional simple Lie algebras, the following result has been proved:

**7.6 Theorem.** *Any simple Lie algebra over $k$ is isomorphic to a unique algebra in the following list:*

$$\mathfrak{sl}_{n+1}(k)\ (n \geq 1,\ A_n), \quad \mathfrak{so}_{2n+1}(k)\ (n \geq 2,\ B_n), \quad \mathfrak{sp}_{2n}(k)\ (n \geq 3,\ C_n),$$
$$\mathfrak{so}_{2n}(k)\ (n \geq 4,\ D_n), \quad E_6,\ E_7,\ E_8,\ F_4,\ G_2.$$

**7.7 Remark.** There are the following isomorphisms among different Lie algebras: $\mathfrak{so}_3(k) \cong \mathfrak{sp}_2(k) = \mathfrak{sl}_2(k)$, $\mathfrak{so}_4(k) \cong \mathfrak{sl}_2(k) \oplus \mathfrak{sl}_2(k)$, $\mathfrak{sp}_4(k) \cong \mathfrak{so}_5(k)$, $\mathfrak{so}_6(k) \cong \mathfrak{sl}_4(k)$.

*Proof.* This can be checked by computing the root systems associated to the natural Cartan subalgebras. If the root systems are isomorphic, then so are the Lie algebras.

Alternatively, note that the Killing form on the three dimensional simple Lie algebra $\mathfrak{sl}_2(k)$ is symmetric and nondegenerate, hence the orthogonal Lie algebra $\mathfrak{so}_3(k) \cong \mathfrak{so}\big(\mathfrak{sl}_2(k), \kappa\big)$, which has dimension 3 and contains the subalgebra $\operatorname{ad} \mathfrak{sl}_2(k) \cong \mathfrak{sl}_2(k)$, which is three dimensional too. Hence $\mathfrak{so}_3(k) \cong \mathfrak{sl}_2(k)$.

Now consider $V = \mathrm{Mat}_2(k)$, which is endowed with the quadratic form det and its associated symmetric bilinear form $b(x, y) = \frac{1}{2}\big(\det(x + y) - \det(x) - \det(y)\big) = -\frac{1}{2}\big(\mathrm{trace}(xy) - \mathrm{trace}(x)\,\mathrm{trace}(y)\big)$. Then we get the one-to-one Lie algebra homomorphism $\mathfrak{sl}_2(k) \oplus \mathfrak{sl}_2(k) \to \mathfrak{so}(V, b) \cong \mathfrak{so}_4(k)$, $(a, b) \mapsto \varphi_{a,b}$, where $\varphi_{a,b}(x) = ax - xb$. By dimension count, this is an isomorphism.

Next, consider the vector space $V = k^4$. The determinant provides a linear isomorphism $\det : \Lambda^4 V \cong k$, which induces a symmetric nondegenerate bilinear map $b : \Lambda^2 V \times \Lambda^2 V \to k$. The Lie algebra $\mathfrak{sl}(V)$ acts on $\Lambda^2(V)$, which gives an embedding $\mathfrak{sl}_4(k) \cong \mathfrak{sl}(V) \hookrightarrow \mathfrak{so}\big(\Lambda^2 V, b\big) \cong \mathfrak{so}_6(k)$. By dimension count, these Lie algebras are isomorphic. Finally, consider a nondegenerate skew-symmetric bilinear form $c$ on $V$. Then $c$ may be considered as a linear map $c : \Lambda^2 V \to k$ and the dimension of $K = \ker c$ is 5. The embedding $\mathfrak{sl}(V) \hookrightarrow \mathfrak{so}(\Lambda^2 V, b)$ restricts to an isomorphism $\mathfrak{sp}_4(k) \cong \mathfrak{sp}(V, c) \cong \mathfrak{so}(K, b) \cong \mathfrak{so}_5(k)$. $\square$

## § 8. Exceptional Lie algebras

In this section a construction of the exceptional simple Lie algebras will be given, thus completing the proof of Theorem 7.6. The hypothesis of the ground field $k$ being algebraically closed of characteristic 0 will be kept here. Many details will be left to the reader.

Let $V = k^3 = \mathrm{Mat}_{3\times 1}(k)$ and let $\times$ denote the usual cross product on $V$. For any $x \in V$, let $l_x$ denote the coordinate matrix, in the canonical basis, of the map $y \mapsto x \times y$. Hence for

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \mapsto \quad l_x = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}.$$

Consider also the map $V^3 \to k$, $(x, y, z) \mapsto (x \times y) \cdot z$ (where $u \cdot v$ denotes the canonical inner product on $V$). Then a simple computation gives that for any $a \in \mathfrak{sl}_3(k)$, $l_{ax} = -(l_x a + a^t l_x)$. Also, the identity of the double cross product: $(x \times y) \times z = (x \cdot z)y - (y \cdot z)x$, shows that $l_{x\times y} = yx^t - xy^t$. Using these properties, the proof of the following result follows at once.

**8.1 Proposition.** *The subspace*

$$L = \left\{ \begin{pmatrix} 0 & -2y^t & -2x^t \\ x & a & l_y \\ y & l_x & -a^t \end{pmatrix} : a \in \mathfrak{sl}_3(k),\ x, y \in k^3 \right\}$$

*is a fourteen dimensional Lie subalgebra of* $\mathfrak{gl}_7(k)$.

For any $a \in \mathfrak{sl}_3(k)$, and $x, y \in k^3$, let $M_{(a,x,y)}$ denote the matrix $\begin{pmatrix} 0 & -2y^t & -2x^t \\ x & a & l_y \\ y & l_x & -a^t \end{pmatrix}$.

In particular we get:

$$[M_{(a,0,0)}, M_{(0,x,0)}] = M_{(0,ax,0)}, \quad [M_{(a,0,0)}, M_{(0,0,y)}] = M_{(0,0,-a^t y)}.$$

Let $H$ be the space of diagonal matrices in $L$, $\dim_k H = 2$ and let $\epsilon_i : H \to k$, the linear map such that $\epsilon_i\bigl(\operatorname{diag}(0, \alpha_1, \alpha_2, \alpha_3, -\alpha_1, -\alpha_2, -\alpha_3)\bigr) = \alpha_i$, $i = 1, 2, 3$. Thus, $\epsilon_1 + \epsilon_2 + \epsilon_3 = 0$. Let $\{e_1, e_2, e_3\}$ be the canonical basis of $V = k^3$. Then we have a root space decomposition

$$L = H \oplus \bigl(\oplus_{\alpha\in\Phi} L_\alpha\bigr),$$

with $\Phi = \{\pm(\epsilon_1 - \epsilon_2), \pm(\epsilon_1 - \epsilon_3), \pm(\epsilon_2 - \epsilon_3), \pm\epsilon_1, \pm\epsilon_2, \pm\epsilon_3\}$, where $M_{(E_{ij},0,0)} \in L_{\epsilon_i-\epsilon_j}$ for $i \neq j$, $M_{(0,e_i,0)} \in L_{\epsilon_i}$, and $M_{(0,0,e_i)} \in L_{-\epsilon_i}$.

**8.2 Theorem.** *$L$ is simple of type $G_2$.*

*Proof.* Any proper ideal $I$ of $L$ is invariant under the adjoint action of $H$, so $I = (I\cap H) \oplus \bigl(\oplus_{\alpha\in\Phi}(I \cap L_\alpha)\bigr)$. Also, $\mathfrak{sl}_3(k)$ is isomorphic to the subalgebra $S = \{M_{(a,0,0)} : a \in \mathfrak{sl}_3(k)\}$ of $L$. If $I \cap S \neq 0$, then, since $S$ is simple, $H \subseteq S \subseteq I$, and hence $L = H + [H, L] \subseteq I$, a contradiction. On the other hand, if $I \cap S = 0$, then there is an $i = 1, 2, 3$ such that $L_\alpha \subseteq I$ with $\alpha = \pm\epsilon_i$. But $0 \neq [L_\alpha, L_{-\alpha}] \subseteq I \cap S$, a contradiction again.

Therefore $L$ is simple of rank 2 and dimension 14. Since the classical Lie algebras of rank 2 are $\mathfrak{sl}_3(k)$ of dimension 8, and $\mathfrak{so}_5(k)$ of dimension 10, the only possibility left is that $L$ must be of type $G_2$. □

**8.3 Exercise.** Compute the restriction to $H$ of the Killing form of $L$. Get a system of simple roots of $\Phi$ and check directly that $\Phi$ is the root system $G_2$.

Let us proceed now to give a construction, due to Freudenthal, of the simple Lie algebra of type $E_8$. To do so, let $V$ be a vector space of dimension 9 and $V^*$ its dual. Consider a nonzero alternating multilinear map $\det : V^9 \to k$ (the election of det to name this map is natural), which induces an isomorphism $\Lambda^9 V \cong k$, and hence another isomorphism $\Lambda^9 V^* \cong \bigl(\Lambda^9 V\bigr)^* \cong k$. Take a basis $\{e_1, \ldots, e_9\}$ of $V$ with $\det(e_1, \ldots, e_9) = 1$, and consider its dual basis $\{\varepsilon_1, \ldots, \varepsilon_9\}$ (so, under the previous isomorphisms, $\varepsilon_1 \wedge \ldots \wedge \varepsilon_9 \in \Lambda^9 V^*$ corresponds to $1 \in k$ too).

Consider now the simple Lie algebra of type $A_8$, $S = \mathfrak{sl}(V) \cong \mathfrak{sl}_9(k)$, which acts naturally on $V$. Then $V^*$ is a module too for $S$ with the action given by $x.\varphi(v) = -\varphi(x.v)$ for any $x \in S$, $v \in V$ and $\varphi \in V^*$. Consider $W = \Lambda^3 V$, which is a module too under the action given by $x.(v_1 \wedge v_2 \wedge v_3) = (x.v_1) \wedge v_2 \wedge v_3 + v_1 \wedge (x.v_2) \wedge v_3 + v_1 \wedge v_2 \wedge (x.v_3)$ for any $x \in S$ and $v_1, v_2, v_3 \in V$. The dual space (up to isomorphism) $W^* = \Lambda^3 V^*$ is likewise a module for $S$. Here $(\varphi_1 \wedge \varphi_2 \wedge \varphi_3)(v_1 \wedge v_2 \wedge v_3) = \det\bigl(\varphi_i(v_j)\bigr)$ for any $\varphi_1, \varphi_2, \varphi_3 \in V^*$ and $v_1, v_2, v_3 \in V$.

The multilinear map det induces a multilinear alternating map $T : W \times W \times W \to k$, such that

$$T\bigl(v_1 \wedge v_2 \wedge v_3, v_4 \wedge v_5 \wedge v_6, v_7 \wedge v_8 \wedge v_9) = \det(v_1, \ldots, v_9),$$

for any $v_i$'s in $V$. In the same vein we get the multilinear alternating map $T^* : W^* \times W^* \times W^* \to k$. These maps induce, in turn, bilinear maps $W \times W \to W^*$, $(w_1, w_2) \mapsto w_1 \diamond w_2 \in W^*$, with $(w_1 \diamond w_2)(w) = T(w_1, w_2, w)$, and $W^* \times W^* \to W$, $(\psi_1, \psi_2) \mapsto \psi_1 \diamond \psi_2 \in W$, with $(\psi_1 \diamond \psi_2)(\psi) = T^*(\psi_1, \psi_2, \psi)$, for any $w_1, w_2, w \in W$ and $\psi_1, \psi_2, \psi \in W^*$, and where natural identifications have been used, like $(W^*)^* \cong W$.

Take now the bilinear map $\Lambda^3 V \times \Lambda^3 V^* \to \mathfrak{sl}(V)$: $(w, \psi) \mapsto w * \psi$, given by

$$\begin{aligned}(v_1 \wedge v_2 \wedge v_3) &* (\varphi_1 \wedge \varphi_2 \wedge \varphi_3)\\ &= \frac{1}{2}\Bigl(\sum_{\sigma,\tau\in S_3} (-1)^\sigma(-1)^\tau \varphi_{\sigma(1)}(v_{\tau(1)})\varphi_{\sigma(2)}(v_{\tau(2)})v_{\tau(3)} \otimes \varphi_{\sigma(3)}\Bigr) - \frac{1}{3}\det\bigl(\varphi_i(v_j)\bigr)1_V,\end{aligned}$$

where $(-1)^\sigma$ denotes the signature of the permutation $\sigma \in S_3$, $v \otimes \varphi$ denotes the endomorphism $u \mapsto \varphi(u)v$, and $1_V$ denotes the identity map on $V$. Then for any $w \in \Lambda^3 V$, $\psi \in \Lambda^3 V^*$ and $x \in \mathfrak{sl}(V)$, the following equation holds:

$$\operatorname{trace}\bigl((w * \psi)x\bigr) = \psi(x.w).$$

(It is enough to check this for basic elements $e_J = e_{j_1} \wedge e_{j_2} \wedge e_{j_3}$, where $J = (j_1, j_2, j_3)$ and $j_1 < j_2 < j_3$, in $W$ and the elements in the dual basis of $W^*$: $\varepsilon_J = \varepsilon_{j_1} \wedge \varepsilon_{j_2} \wedge \varepsilon_{j_3}$.) Note that this equation can be used as the definition of $w * \psi$.

Now consider the vector space $L = \mathfrak{sl}(V) \oplus W \oplus W^*$ with the Lie bracket given, for any $x, y \in \mathfrak{sl}(V)$, $w, w_1, w_2 \in W$ and $\psi, \psi_1, \psi_2 \in W^*$ by:

$$\begin{aligned}
&[x, y] \text{ is the bracket in } \mathfrak{sl}(V),\\
&[x, w] = x.w \in W,\ [x, \psi] = x.\psi \in W^*,\\
&[w_1, w_2] = w_1 \diamond w_2 \in W^*,\\
&[\psi_1, \psi_2] = \psi_1 \diamond \psi_2 \in W,\\
&[w, \psi] = -w * \psi \in \mathfrak{sl}(V).
\end{aligned}$$

A lengthy computation with basic elements, shows that $L$ is indeed a Lie algebra. Its dimension is $\dim_k L = 80 + 2\binom{9}{3} = 80 + 2 \times 84 = 244$.

Let $H$ be the Cartan subalgebra of $\mathfrak{sl}(V)$ consisting of the trace zero endomorphisms with a diagonal coordinate matrix in our basis $\{e_1, \ldots, e_9\}$, and let $\delta_i : H \to k$ be the linear form such that (identifying endomorphisms with their coordinate matrices) $\delta_i\bigl(\operatorname{diag}(\alpha_1, \ldots, \alpha_9)\bigr) = \alpha_i$. Then $\delta_1 + \cdots + \delta_9 = 0$, $H$ is toral in $L$ and there is a root decomposition

$$L = H \oplus \bigl(\oplus_{\alpha \in \Phi} L_\alpha\bigr),$$

where

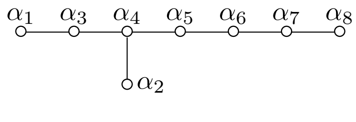
$$\Phi = \{\delta_i - \delta_j : i \neq j\} \cup \{\pm(\delta_i + \delta_j + \delta_k) : i < j < k\}.$$

Here $L_{\delta_i - \delta_j} = kE_{ij} \subseteq \mathfrak{sl}(V)$ ($E_{ij}$ denotes the endomorphism whose coordinate matrix has $(i, j)$-entry 1 and 0's elsewhere), $L_{\delta_i + \delta_j + \delta_k} = k(e_i \wedge e_j \wedge e_k) \subseteq W$ and $L_{-(\delta_i + \delta_j + \delta_k)} = k(\varepsilon_i \wedge \varepsilon_j \wedge e\varepsilon_k) \subseteq W^*$. Using that $\mathfrak{sl}(V)$ is simple, the same argument in the proof of Theorem 8.2 proves that $L$ is simple:

**8.4 Theorem.** *$L$ is simple of type $E_8$.*

*Proof.* We have shown that $L$ is simple of rank 8. The classical Lie algebras of rank 8, up to isomorphism, are $\mathfrak{sl}_9(k)$, $\mathfrak{so}_{17}(k)$, $\mathfrak{sp}_{16}(k)$ and $\mathfrak{so}_{16}(k)$, which have dimensions 80, 156, 156 and 120 respectively. Hence $L$ is not isomorphic to any of them and hence it is of type $E_8$. □

Take now the simple Lie algebra $L$ of type $E_8$ and its generators $\{h_i, x_i, y_i : i = 1, \ldots, 8\}$ as in the paragraph previous to Lemma 7.2, the indexing given by the ordering of the simple roots given in the next diagram:



Let $\kappa$ be the Killing form of $L$. Then consider the subalgebra $\hat{L}$ generated by $\{h_i, x_i, y_i : i = 1, \dots, 7\}$ and its subalgebra $\hat{H} = \oplus_{i=1}^{7} kh_i$. Since $H$ is toral in $L$, so is $\hat{H}$ in $\hat{L}$ and

$$\hat{L} = \hat{H} \oplus \Big( \oplus_{\alpha \in \Phi \cap (\mathbb{Z}\alpha_1 + \cdots \mathbb{Z}\alpha_7)} L_\alpha \Big).$$

From here it follows that $\hat{H}$ is a Cartan subalgebra of $\hat{L}$. Since the restriction of $\kappa$ to $\hat{H}$ is nondegenerate (recall that the restriction of $\kappa$ to $\sum_{i=1}^{8} \mathbb{Q}h_i$ is positive definite!), the restriction of $\kappa$ to $\hat{L}$ is nondegenerate. Thus we get a representation $\mathrm{ad} : \hat{L} \to \mathfrak{gl}(L)$ with nondegenerate trace form, and hence $\hat{L} = Z(\hat{L}) \oplus [\hat{L}, \hat{L}]$, with $[\hat{L}, \hat{L}]$ semisimple (recall Consequences 2.2). But $\hat{H} \subseteq [\hat{L}, \hat{L}]$ and $Z(\hat{L}) \subseteq C_{\hat{L}}(\hat{H}) = \hat{H}$, so $Z(\hat{L}) = 0$ and $\hat{L}$ is semisimple, with root system of type $E_7$ (which is irreducible). Theorem 6.3 shows that $\hat{L}$ is simple of type $E_7$.

The same arguments show that the Lie subalgebra $\tilde{L}$ of $L$ generated by $\{h_i, x_i, y_i : i = 1, \dots, 6\}$ is a simple Lie algebra of type $E_6$.

Finally, the existence of a simple Lie algebra of type $F_4$ will be deduced from that of $E_6$. Let now $\tilde{L}$ be the simple Lie algebra of type $E_6$ considered above, with canonical generators $\{h_i, x_i, y_i : i = 1, \dots, 6\}$. Since the multiplication in $\tilde{L}$ is determined by the Dynkin diagram, there is an automorphism $\varphi$ of $\tilde{L}$ such that

$$\begin{aligned}
&\varphi(h_1) = h_6,\ \varphi(x_1) = x_6,\ \varphi(y_1) = y_6,\\
&\varphi(h_6) = h_1,\ \varphi(x_6) = x_1,\ \varphi(y_6) = y_1,\\
&\varphi(h_3) = h_5,\ \varphi(x_3) = x_5,\ \varphi(y_3) = y_5,\\
&\varphi(h_5) = h_3,\ \varphi(x_5) = x_3,\ \varphi(y_5) = y_3,\\
&\varphi(h_2) = h_2,\ \varphi(x_2) = x_2,\ \varphi(y_2) = y_2,\\
&\varphi(h_4) = h_4,\ \varphi(x_4) = x_4,\ \varphi(y_4) = y_4.
\end{aligned}$$

In particular, $\varphi^2$ is the identity, so $\tilde{L} = \tilde{L}_{\bar{0}} \oplus \tilde{L}_{\bar{1}}$, with $\tilde{L}_{\bar{0}} = \{z \in \tilde{L} : \varphi(z) = z\}$, while $\tilde{L}_{\bar{1}} = \{z \in \tilde{L} : \varphi(z) = -z\}$, and it is clear that $\tilde{L}_{\bar{0}}$ is a subalgebra of $\tilde{L}$, $[\tilde{L}_{\bar{0}}, \tilde{L}_{\bar{1}}] \subseteq \tilde{L}_{\bar{1}}$, $[\tilde{L}_{\bar{1}}, \tilde{L}_{\bar{1}}] \subseteq \tilde{L}_{\bar{0}}$. For any $z \in \tilde{L}_{\bar{0}}$ and $z' \in \tilde{L}_{\bar{1}}$, $\kappa(z, z') = \kappa\big(\varphi(z), \varphi(z')\big) = \kappa(z, -z')$, where $\kappa$ denotes the Killing form of $\tilde{L}$. Hence $\kappa(\tilde{L}_{\bar{0}}, \tilde{L}_{\bar{1}}) = 0$ and, thus, the restriction of $\kappa$ to $\tilde{L}_{\bar{0}}$ is nondegenerate. This means that the adjoint map gives a representation $\mathrm{ad} : \tilde{L}_{\bar{0}} \to \mathfrak{gl}(\tilde{L})$ with nondegenerate trace form. As before, this gives $\tilde{L}_{\bar{0}} = Z(\tilde{L}_{\bar{0}}) \oplus [\tilde{L}_{\bar{0}}, \tilde{L}_{\bar{0}}]$, and $[\tilde{L}_{\bar{0}}, \tilde{L}_{\bar{0}}]$ is semisimple.

Consider the following elements of $\tilde{L}_{\bar{0}}$:

$$\begin{array}{llll}
\tilde{h}_1 = h_1 + h_6, & \tilde{h}_2 = h_3 + h_5, & \tilde{h}_3 = h_2, & \tilde{h}_4 = h_2,\\
\tilde{x}_1 = x_1 + x_6, & \tilde{x}_2 = x_3 + x_5, & \tilde{x}_3 = x_4, & \tilde{x}_4 = x_2,\\
\tilde{y}_1 = y_1 + y_6, & \tilde{y}_2 = y_3 + y_5, & \tilde{y}_3 = y_4, & \tilde{y}_4 = y_2.
\end{array}$$

Note that $[\tilde{x}_i, \tilde{y}_i] = \tilde{h}_i$ for any $i = 1, 2, 3, 4$. The element $\tilde{h} = 10\tilde{h}_1 + 19\tilde{h}_2 + 27\tilde{h}_3 + 14\tilde{h}_4$ satisfies

$$\begin{cases}
\alpha_1(\tilde{h}) = \alpha_6(\tilde{h}) = 20 - 19 = 1,\\
\alpha_3(\tilde{h}) = \alpha_5(\tilde{h}) = 38 - 10 - 27 = 1,\\
\alpha_4(\tilde{h}) = 54 - 38 - 14 = 2,\\
\alpha_2(\tilde{h}) = 28 - 27 = 1.
\end{cases}$$

Thus $\alpha(\tilde{h}) > 0$ for any $\alpha \in \Phi^+$, where $\Phi$ is the root system of $\tilde{L}$. In particular, $\alpha(\tilde{h}) \neq 0$ for any $\alpha \in \Phi$.

Note that $\tilde{H} = \oplus_{i=1}^6 kh_i$ is a Cartan subalgebra of $\tilde{L}$. Besides, $\varphi(\tilde{H}) = \tilde{H}$ and hence $\tilde{H} = \tilde{H}_{\bar{0}} \oplus \tilde{H}_{\bar{1}}$, with $\tilde{H}_{\bar{0}} = \tilde{H} \cap \tilde{L}_{\bar{0}} = \oplus_{i=1}^4 k\tilde{h}_i$ and $\tilde{H}_{\bar{1}} = \tilde{H} \cap \tilde{L}_{\bar{1}} = k(h_1 - h_6) \oplus k(h_3 - h_5)$. Also, for any $\alpha \in \Phi$, $x_\alpha + \varphi(x_\alpha) \in \tilde{L}_{\bar{0}}$, and this vector is a common eigenvector for $\tilde{H}_{\bar{0}}$ with eigenvalue $\alpha|_{\tilde{H}_{\bar{0}}}$, which is not zero since $\alpha(\tilde{h}) \neq 0$ for any $\alpha \in \Phi$. Hence there is a root space decomposition

$$\tilde{L}_{\bar{0}} = \tilde{H}_{\bar{0}} \oplus \big(\sum_{\alpha \in \Phi} k(x_\alpha + \varphi(x_\alpha))\big)$$

and it follows that $Z(\tilde{L}_{\bar{0}}) \subseteq C_{\tilde{L}}(\tilde{H}_{\bar{0}}) \cap L_{\bar{0}} = \tilde{H} \cap L_{\bar{0}} = \tilde{H}_{\bar{0}} \subseteq [\tilde{L}_{\bar{0}}, \tilde{L}_{\bar{0}}]$. We conclude that $Z(\tilde{L}_{\bar{0}}) = 0$, so $\tilde{L}_{\bar{0}}$ is semisimple, and $\tilde{H}_{\bar{0}}$ is a Cartan subalgebra of $\tilde{L}_{\bar{0}}$.

The root system $\tilde{\Phi}$ of $\tilde{L}_{\bar{0}}$, relative to $\tilde{H}_{\bar{0}}$, satisfies that $\tilde{\Phi} \subseteq \{\tilde{\alpha} = \alpha|_{\tilde{H}_{\bar{0}}} : \alpha \in \Phi\}$. Also $\tilde{\alpha}_i = \alpha_i|_{\tilde{H}_{\bar{0}}} \in \tilde{\Phi}$, with $\tilde{x}_i \in (\tilde{L}_{\bar{0}})_{\tilde{\alpha}_i}$ and $\tilde{y}_i \in (\tilde{L}_{\bar{0}})_{-\tilde{\alpha}_i}$ for any $i = 1, 2, 3, 4$. Moreover, $[\tilde{x}_i, \tilde{y}_i] = \tilde{h}_i$ and $\tilde{\alpha}_i(\tilde{h}_i) = 2$ for any $i$. Besides, $\tilde{\Phi} = \tilde{\Phi}^+ \cup \tilde{\Phi}^-$, with $\tilde{\Phi}^+ = \{\tilde{\alpha} \in \tilde{\Phi} : \tilde{\alpha}(\tilde{h}) > 0\} \subseteq \{\tilde{\alpha} : \alpha \in \Phi^+\}$ (and similarly with $\tilde{\Phi}^-$). We conclude that $\tilde{\Delta} = \{\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4\}$ is a system of simple roots of $\tilde{L}_{\bar{0}}$. We can compute the associated Cartan matrix. For instance,

$$[\tilde{h}_1, \tilde{x}_2] = \begin{cases} \tilde{\alpha}_2(\tilde{h}_1)\tilde{x}_2 = \langle \tilde{\alpha}_2 | \tilde{\alpha}_1 \rangle \tilde{x}_2 \\ [h_1 + h_6, x_3 + x_5] = \alpha_3(h_1 + h_6)x_3 + \alpha_5(h_1 + h_6)x_5 = -(x_3 + x_5) = -\tilde{x}_2, \end{cases}$$
$$[\tilde{h}_2, \tilde{x}_3] = \begin{cases} \tilde{\alpha}_3(\tilde{h}_2)\tilde{x}_3 = \langle \tilde{\alpha}_3 | \tilde{\alpha}_2 \rangle \tilde{x}_3 \\ [h_3 + h_5, x_4] = \alpha_4(h_3 + h_5)x_4 = -2x_4 = -2\tilde{x}_3, \end{cases}$$

which shows that $\langle \tilde{\alpha}_2 | \tilde{\alpha}_1 \rangle = -1$ and $\langle \tilde{\alpha}_3 | \tilde{\alpha}_2 \rangle = -2$. In this way we can compute the whole Cartan matrix, which turns out to be the Cartan matrix of type $F_4$:

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

thus proving that $\tilde{L}_{\bar{0}}$ is the simple Lie algebra of type $F_4$.

# Chapter 3

# Representations of semisimple Lie algebras

Unless otherwise stated, the following assumptions will be kept throughout the chapter:

- $k$ will denote an algebraically closed field of characteristic 0,
- $L$ will denote a semisimple Lie algebra over $k$,
- $H$ will be a fixed Cartan subalgebra of $L$, $\Phi$ will denote the corresponding set of roots and $L = H \oplus \left(\oplus_{\alpha\in\Phi} L_\alpha\right)$ the root space decomposition.
- $\kappa$ will denote the Killing form of $L$ and $(\,|\,) : H^* \times H^* \to k$ the induced nondegenerate bilinear form.
- For any $\alpha \in \Phi$, $t_\alpha \in H$ is defined by the relation $\alpha(h) = \kappa(t_\alpha, h)$ for any $h \in H$, and $h_\alpha = \frac{2t_\alpha}{\alpha(t_\alpha)}$.
- $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ denotes a fixed system of simple roots. Accordingly, $\Phi$ decomposes as $\Phi = \Phi^+ \cup \Phi^-$ (disjoint union), where $\Phi^+$ (respectively $\Phi^-$) is the set of positive roots (resp., negative roots). Moreover, $\Phi^- = -\Phi^+$. For $\alpha \in \Phi^+$, let $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$ with $[x_\alpha, y_\alpha] = h_\alpha$.
- $\mathcal{W}$ is the Weyl group, generated by $\{\sigma_\alpha : \alpha \in \Phi\}$.
- $L^+ = \oplus_{\alpha\in\Phi^+} L_\alpha$, $L^- = \oplus_{\alpha\in\Phi^-} L_\alpha$, so that $L = L^- \oplus H \oplus L^+$.

This chapter is devoted to the study of the finite dimensional representations of such an algebra $L$. By Weyl's theorem (Chapter 2, 2.5), any representation is completely reducible, so the attention is focused on the irreducible representations.

## § 1. Preliminaries

Let $\rho : L \to \mathfrak{gl}(V)$ be a finite dimensional representation of the Lie algebra $L$. Since the Cartan subalgebra $H$ is toral, $V$ decomposes as

$$V = \oplus_{\mu\in H^*} V_\mu,$$

where $V_\mu = \{v \in V : h.v = \mu(h)v \ \forall h \in H\}$.

**1.1 Definition.** Under these circumstances, $\mu \in H^*$ is said to be a *weight* of $V$ if $V_\mu \neq 0$. The set of weights of $V$ is denoted $P(V)$.

**1.2 Properties of $P(V)$.**

(i) For any $\alpha \in \Phi$ and $\mu \in P(V)$, $L_\alpha.V_\mu \subseteq V_{\alpha+\mu}$.

(ii) For any $\mu \in P(V)$ and $\alpha \in \Phi$, $\langle \mu | \alpha \rangle := \frac{2(\mu|\alpha)}{(\alpha|\alpha)}$ is an integer.

*Proof.* Let $S_\alpha = L_\alpha \oplus L_{-\alpha} \oplus [L_\alpha, L_{-\alpha}]$, which is isomorphic to $\mathfrak{sl}_2(k)$ and take elements $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$ such that $[x_\alpha, y_\alpha] = h_\alpha$. Then $W = \oplus_{m \in \mathbb{Z}} V_{\mu + m\alpha}$ is an $S_\alpha$-submodule of $V$. Hence the eigenvalues of the action of $h_\alpha$ on $W$ form an unbroken chain of integers:

$$(\mu + q\alpha)(h_\alpha), \dots, \mu(h_\alpha), \dots, (\mu - r\alpha)(h_\alpha), \tag{1.1}$$

with $(\mu - r\alpha)(h_\alpha) = -(\mu + q\alpha)(h_\alpha)$. But $\mu(h_\alpha) = \langle \mu | \alpha \rangle$ and $\alpha(h_\alpha) = 2$. Hence,

$$\mu(h_\alpha) = \langle \mu | \alpha \rangle = r - q \in \mathbb{Z}.$$ □

(iii) $P(V)$ is $\mathcal{W}$-invariant.

*Proof.* For any $\mu \in P(V)$ and $\alpha \in \Phi$, $\sigma_\alpha(\mu) = \mu - \langle \mu | \alpha \rangle \alpha = \mu - (r - q)\alpha \in P(V)$, since it belongs to the unbroken chain (1.1). □

(iv) Let $C = \big(\langle \alpha_i | \alpha_j \rangle\big)$ be the Cartan matrix. Then

$$P(V) \subseteq \frac{1}{\det C}\Big(\mathbb{Z}\alpha_1 + \cdots + \mathbb{Z}\alpha_n\Big) \subseteq E = \mathbb{R}\alpha_1 + \cdots + \mathbb{R}\alpha_n.$$

(Recall that $E$ is an euclidean vector space.)

*Proof.* Since $\Delta$ is a basis of $H^*$, for any $\mu \in P(V)$, there are scalars $r_1, \dots, r_n \in k$ such that $\mu = r_1\alpha_1 + \cdots + r_n\alpha_n$. Then $\langle \mu | \alpha_j \rangle = \sum_{i=1}^n \langle \alpha_i | \alpha_j \rangle r_i$, $j = 1, \dots, n$. This constitutes a system of linear equations with integer coefficients, whose matrix is $C$. Solving this system using Cramer's rule gives $r_i \in \frac{1}{\det C}\mathbb{Z}$. □

At this point it is useful to note that $\det A_n = n+1$, $\det B_n = \det C_n = 2$, $\det D_n = 4$, $\det E_6 = 3$, $\det E_7 = 2$ and $\det E_8 = \det F_4 = \det G_2 = 1$.

**1.3 Definition.**

- $\Lambda_R = \mathbb{Z}\Delta = \mathbb{Z}\Phi$ is called the *root lattice* of $L$.
- $\Lambda_W = \{\lambda \in H^* : \langle \lambda | \alpha_i \rangle \in \mathbb{Z}\ \forall i = 1, \dots, n\}$ (which is contained in $\frac{\mathbb{Z}}{\det C}\Delta$) is called the *weight lattice*.
- The elements of $\Lambda_W$ are called weights of the pair $(L, H)$.
- An element $\lambda \in \Lambda_W$ is said to be a *dominant weight* if $\langle \lambda | \alpha \rangle \geq 0$ for any $\alpha \in \Delta$. The set of dominant weights is denoted by $\Lambda_W^+$.

- For any $i = 1, \ldots, n$, let $\lambda_i \in H^*$ such that $\langle\lambda_i|\alpha_j\rangle = \delta_{ij}$ for any $j = 1, \ldots, n$. Then $\lambda_i \in \Lambda_W^+$, $\Lambda_W = \mathbb{Z}\lambda_1 + \ldots + \mathbb{Z}\lambda_n$, and $\Lambda_W^+ = \mathbb{Z}_{\geq 0}\lambda_1 + \cdots + \mathbb{Z}_{\geq 0}\lambda_n$. The weights $\lambda_1, \ldots, \lambda_n$ are called the *fundamental dominant weights*.

**1.4 Proposition.** $\Lambda_W = \{\lambda \in H^* : \langle\lambda|\alpha\rangle \in \mathbb{Z}\ \forall\alpha \in \Phi\}$. *In particular, the weight lattice does not depend on the chosen set of simple roots* $\Delta$.

*Proof.* It is trivial that $\{\lambda \in H^* : \langle\lambda|\alpha\rangle \in \mathbb{Z}\ \forall\alpha \in \Phi\} \subseteq \Lambda_W$. Conversely, let $\lambda \in \Lambda_W$ and $\alpha \in \Phi^+$. Let us check that $\langle\lambda|\alpha\rangle \in \mathbb{Z}$ by induction on $\mathrm{ht}(\alpha)$. If this height is 1, then $\alpha \in \Delta$ and this is trivial. If $\mathrm{ht}(\alpha) = n > 1$, then $\alpha = m_1\alpha_1 + \cdots + m_n\alpha_n$, with $m_1, \ldots, m_n \in \mathbb{Z}_{\geq 0}$. Since $(\alpha|\alpha) > 0$, there is at least one $i = 1, \ldots, n$ such that $(\alpha|\alpha_i) > 0$ and $\mathrm{ht}\big(\sigma_{\alpha_i}(\alpha)\big) = \mathrm{ht}\big(\alpha - \langle\alpha|\alpha_i\rangle\alpha_i\big) < \mathrm{ht}(\alpha)$. Then

$$\begin{aligned}\langle\lambda|\alpha\rangle &= \langle\sigma_{\alpha_i}(\lambda)|\sigma_{\alpha_i}(\alpha)\rangle = \big\langle\lambda - \langle\lambda|\alpha_i\rangle\alpha_i|\sigma_{\alpha_i}(\alpha)\big\rangle \\ &= \langle\lambda|\sigma_{\alpha_i}(\alpha)\rangle - \langle\lambda|\alpha_i\rangle\langle\alpha_i|\sigma_{\alpha_i}(\alpha)\rangle\end{aligned}$$

and the first summand is in $\mathbb{Z}$ by the induction hypothesis, and so is the second since $\lambda \in \Lambda_W$ and $\langle\Phi|\Phi\rangle \subseteq \mathbb{Z}$. □

**1.5 Definition.** Let $\rho : L \to \mathfrak{gl}(M)$ be a not necessarily finite dimensional representation.

(i) An element $0 \neq m \in M$ is called a *highest weight vector* if $m$ is an eigenvector for all the operators $\rho(h)$ $(h \in H)$, and $\rho(L^+)(m) = 0$.

(ii) The module $M$ is said to be a *highest weight module* if it contains a highest weight vector that generates $M$ as a module for $L$.

**1.6 Proposition.** *Let* $\rho : L \to \mathfrak{gl}(V)$ *be a finite dimensional representation of* $L$. *Then*

(i) *$V$ contains highest weight vectors. If $v \in V$ is such a vector and $v \in V_\lambda$ ($\lambda \in H^*$), then $\lambda \in \Lambda_W^+$.*

(ii) *Let $0 \neq v \in V_\lambda$ be a highest weight vector. Then*

$$W = kv + \sum_{r=1}^{\infty} \sum_{1\leq i_1,\ldots,i_r\leq n} k\big(\rho(y_{\alpha_{i_1}})\cdots\rho(y_{\alpha_{i_r}})(v)\big)$$

*is the submodule of $V$ generated by $v$. Besides, $W$ is an irreducible $L$-module and $P(W) \subseteq \{\lambda - \alpha_{i_1} - \cdots - \alpha_{i_r} : r \geq 0,\ 1 \leq i_1, \ldots, i_r \leq n\}$ $\big(= \lambda - \sum_{i=1}^n \mathbb{Z}_{\geq 0}\alpha_i\big)$.*

(iii) *If $V$ is irreducible, then it contains, up to scalars, a unique highest weight vector. Its weight is called the* highest weight *of $V$.*

(iv) *(Uniqueness) For any $\lambda \in \Lambda_W^+$ there is, up to isomorphism, at most one finite dimensional irreducible $L$-module whose highest weight is $\lambda$.*

*Proof.* **(i)** Let $l \in \mathbb{Q}\Phi$ such that $(l|\alpha) > 0$ for any $\alpha \in \Delta$ (for instance, one can take $(l|\alpha) = 1$ for any $\alpha \in \Delta$), and let $\lambda \in P(V)$ such that $(l|\lambda)$ is maximum. Then for any $\alpha \in \Phi^+$, $\lambda + \alpha \notin P(V)$, so that $L_\alpha.V_\lambda = 0$. Hence $L^+.V_\lambda = 0$ and any $0 \neq v \in V_\lambda$ is a highest weight vector.

**(ii)** The subspace $W$ is invariant under the action of $L^-$ and the action of $H$ (since it is spanned by common eigenvectors for $H$). Therefore, since $L^+$ is generated by $\{x_\alpha : \alpha \in \Delta\}$, it is enough to check that $W$ is invariant under the action of $\rho(x_\alpha)$, for $\alpha \in \Delta$. But $x_\alpha.v = 0$ ($v$ is a highest weight vector) and for any $\alpha, \beta \in \Delta$ and $w \in W$, $x_\alpha.(y_\beta.w) = [x_\alpha, y_\beta].w - y_\beta.(x_\alpha.w)$, and $[x_\alpha, y_\beta]$ either is 0 or belongs to $H$. An easy induction on $r$ argument shows that $x_\alpha.\big(\rho(y_{\alpha_1})\cdots\rho(y_{\alpha_{i_r}})(v)\big) \in W$, as required.

Therefore, $W$ is an $L$-submodule and $P(W) \subseteq \{\lambda - \alpha_{i_1} - \cdots - \alpha_{i_r} : r \geq 0, 1 \leq i_1, \ldots, i_r \leq n\}$. (Note that up to now, the finite dimensionality of $V$ has played no role.)

Moreover, $V_\lambda \cap W = kv$ and if $W$ is the direct sum of two submodules $W = W' \oplus W''$, then $W_\lambda = kv = W'_\lambda \oplus W''_\lambda$. Hence either $v \in W'_\lambda$ or $v \in W''_\lambda$. Since $W$ is generated by $v$, we conclude that either $W = W'$ or $W = W''$. Now, by finite dimensionality, Weyl's Theorem (Chapter 2, 2.5) implies that $W$ is irreducible.

Besides, since for any $\alpha \in \Phi^+$ we have $\lambda + \alpha \notin P(W)$, $\langle \lambda | \alpha \rangle = r - q = r \geq 0$. This shows that $\lambda \in \Lambda_W^+$, completing thus the proof of (i).

**(iii)** If $V$ is irreducible, then $V = W$, $V_\lambda = kv_\lambda$ and for any $\mu \in P(V) \setminus \{\lambda\}$ there is an $r \geq 0$ and $1 \leq i_1, \ldots, i_r \leq n$ such that $\mu = \lambda - \alpha_{i_1} - \cdots - \alpha_{i_r}$. Hence $(l|\mu) < (l|\lambda)$. Therefore, the highest weight is the only weight with maximum value of $(l|\lambda)$.

**(iv)** If $V^1$ and $V^2$ are two irreducible highest weight modules with highest weight $\lambda$ and $v^1 \in V^1_\lambda$, $v^2 \in V^2_\lambda$ are two highest weight vectors, then $w = (v^1, v^2)$ is a highest weight vector in $V^1 \oplus V^2$, and hence $W = kw + \sum_{r=1}^{\infty} \sum_{1 \leq i_1, \ldots, i_r \leq n} k\big(\rho(y_{\alpha_1})\cdots\rho(y_{\alpha_{i_r}})(w)\big)$ is a submodule of $V^1 \oplus V^2$. Let $\pi^i : V^1 \oplus V^2 \to V^i$ denote the natural projection ($i = 1, 2$). Then $v^i \in \pi^i(W)$, so $\pi^i(W) \neq 0$ and, since both $W$ and $V^i$ are irreducible by item (ii), it follows that $\pi^i|_W : W \to V^i$ is an isomorphism ($i = 1, 2$). Hence both $V^1$ and $V^2$ are isomorphic to $W$. $\square$

There appears the natural question of existence: given a dominant weight $\lambda \in \Lambda_W^+$, does there exist a finite dimensional irreducible $L$-module $V$ whose highest weight is $\lambda$?

Note that $\lambda = m_1\lambda_1 + \cdots + m_n\lambda_n$, with $m_1, \ldots, m_n \in \mathbb{Z}_{\geq 0}$. If it can be proved that there exists and irreducible finite dimensional highest weight module $V(\lambda_i)$ of highest weight $\lambda_i$, for any $i = 1, \ldots, n$, then in the module

$$V(\lambda_1)^{\otimes m_1} \otimes \cdots \otimes V(\lambda_n)^{\otimes m_n}$$

there is a highest weight vector of weight $\lambda$ (the basic tensor obtained with the highest weight vectors of each copy of $V(\lambda_i)$), By item (ii) above this highest weight vector generates an irreducible $L$-submodule of highest weight $\lambda$. Hence it is enough to deal with the fundamental dominant weights and this can be done "ad hoc". A more abstract proof will be given here.

## § 2. Properties of weights and the Weyl group

Let us go back to the abstract situation that appeared in Chapter 2.

Let $E$ be an euclidean vector space, $\Phi$ a root system in $E$ and $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ a system of simple roots. Consider in this abstract setting the subsets we are already familiar with:

$$\Lambda_R = \mathbb{Z}\Delta = \mathbb{Z}\Phi,$$

$\Lambda_W = \{\lambda \in E : \langle\lambda|\alpha\rangle \in \mathbb{Z}\ \forall\alpha \in \Phi\} = \mathbb{Z}\lambda_1 + \cdots + \mathbb{Z}\lambda_n$ (the *weight lattice*),

$\Lambda_W^+ = \{\lambda \in \Lambda_W : \langle\lambda|\alpha\rangle \geq 0\ \forall\alpha \in \Delta\} = \mathbb{Z}_{\geq 0}\lambda_1 + \cdots + \mathbb{Z}_{\geq 0}\lambda_n$ (the set of *dominant weights*),

$\sigma_i = \sigma_{\alpha_i}$ $(i = 1, \ldots, n)$, $\mathcal{W} = \langle\sigma_\alpha : \alpha \in \Phi\rangle$ (*Weyl group*).

**2.1 Properties.**

(i) *The Weyl group is generated by $\sigma_1, \ldots, \sigma_n$.*

*Proof.* Let $\mathcal{W}_0$ be the subgroup of $\mathcal{W}$ generated by $\sigma_1, \ldots, \sigma_n$. It is enough to prove that $\sigma_\alpha \in \mathcal{W}_0$ for any $\alpha \in \Phi^+$. This will be proven by induction on $\mathrm{ht}(\alpha)$, and it is trivial if $\mathrm{ht}(\alpha) = 1$. Assume that $\mathrm{ht}(\alpha) = r$ and that $\sigma_\beta \in \mathcal{W}_0$ for any $\beta \in \Phi^+$ with $\mathrm{ht}(\beta) < r$. The arguments in the proof of Proposition 1.4 show that there is an $i = 1, \ldots, n$, such that $(\alpha|\alpha_i) > 0$, so $\beta = \sigma_{\alpha_i}(\alpha) = \alpha - \langle\alpha|\alpha_i\rangle\alpha_i$ satisfies that $\mathrm{ht}(\beta) < \mathrm{ht}(\alpha)$. Hence $\sigma_\beta \in \mathcal{W}_0$. But for any isometry $\tau$ and any $\mu \in E$:

$$\sigma_{\tau(\alpha)} \circ \tau(\mu) = \tau(\mu) - \langle\tau(\mu)|\tau(\alpha)\rangle\tau(\alpha) = \tau\Big(\mu - \langle\mu|\alpha\rangle\alpha\Big) = \tau \circ \sigma_\alpha(\mu),$$

so $\sigma_{\tau(\alpha)} = \tau \circ \sigma_\alpha \circ \tau^{-1}$. In particular, with $\tau = \sigma_i$, $\sigma_\alpha = \sigma_i\sigma_\beta\sigma_i \in \mathcal{W}_0$. □

(ii) *If $t \geq 2$, $\beta_1, \ldots, \beta_t \in \Delta$ and $\sigma_{\beta_1} \circ \ldots \circ \sigma_{\beta_{t-1}}(\beta_t) \in \Phi^-$, then there is an index $1 \leq s \leq t-1$ such that*

$$\sigma_{\beta_1} \circ \ldots \circ \sigma_{\beta_t} = \sigma_{\beta_1} \circ \ldots \circ \sigma_{\beta_{s-1}} \circ \sigma_{\beta_{s+1}} \circ \ldots \circ \sigma_{\beta_{t-1}}.$$

*(That is, there is a simpler expression as a product of generators.)*

*Proof.* Let $s$ be the largest index $(1 \leq s < t-1)$ with $\sigma_{\beta_s} \circ \ldots \circ \sigma_{\beta_{t-1}}(\beta_t) \in \Phi^-$. Thus $\sigma_{\beta_s}\big(\sigma_{\beta_{s+1}} \circ \ldots \circ \sigma_{\beta_{t-1}}(\beta_t)\big) \in \Phi^-$. But $\sigma_{\beta_s}\big(\Phi^+ \setminus \{\beta_s\}\big) = \Phi^+ \setminus \{\beta_s\}$ (Chapter 2, Proposition 6.1), so $\sigma_{\beta_{s+1}} \circ \ldots \circ \sigma_{\beta_{t-1}}(\beta_t) = \beta_s$ and, using the argument in the proof of (i),

$$\big(\sigma_{\beta_{s+1}} \circ \ldots \circ \sigma_{\beta_{t-1}}\big) \circ \sigma_{\beta_t} \circ \big(\sigma_{\beta_{t-1}} \circ \ldots \circ \sigma_{\beta_{s+1}}\big) = \sigma_{\sigma_{\beta_{s+1}} \circ \ldots \circ \sigma_{\beta_{t-1}}(\beta_t)} = \sigma_{\beta_s},$$

whence

$$\sigma_{\beta_s} \circ \sigma_{\beta_{s+1}} \circ \ldots \circ \sigma_{\beta_{t-1}} \circ \sigma_{\beta_t} = \sigma_{\beta_{s+1}} \circ \ldots \circ \sigma_{\beta_{t-1}}.$$ □

(iii) *Given any $\sigma \in \mathcal{W}$, item* (i) *implies that there are $\beta_1, \ldots, \beta_t \in \Delta$ such that $\sigma = \sigma_{\beta_1} \circ \ldots \circ \sigma_{\beta_t}$. This expression is called* reduced *if $t$ is minimum. (For $\sigma = id$, $t = 0$.) By the previous item, if the expression is reduced $\sigma(\beta_t) \in \Phi^-$. In particular, for any $id \neq \sigma \in \mathcal{W}$, $\sigma(\Delta) \neq \Delta$. Therefore, because of Chapter 2, Proposition 6.1, $\mathcal{W}$ acts simply transitively on the systems of simple roots.*

(iv) *Let $\sigma = \sigma_{i_1} \circ \cdots \circ \sigma_{i_t}$ be a reduced expression. Write $\mathrm{l}(\sigma) = t$. Also let $n(\sigma) = |\{\alpha \in \Phi^+ : \sigma(\alpha) \in \Phi^-\}|$. Then $\mathrm{l}(\sigma) = n(\sigma)$.*

*Proof.* By induction on $\mathrm{l}(\sigma) = t$. If $t = 0$ this is trivial. For $t > 0$, $\sigma = \sigma_{i_1} \circ \cdots \circ \sigma_{i_t}$ satisfies $\sigma(\alpha_{i_t}) \in \Phi^-$. But $\sigma_{i_t}\big(\Phi^+ \setminus \{\alpha_{i_t}\}\big) = \Phi^+ \setminus \{\alpha_{i_t}\}$. Hence $n\big(\sigma_{i_1} \circ \cdots \circ \sigma_{i_{t-1}}\big) = n(\sigma) - 1$ and the induction hypothesis applies. ☐

(v) *There is a unique element $\sigma_0 \in \mathcal{W}$ such that $\sigma_0(\Delta) = -\Delta$. Moreover, $\sigma_0^2 = id$ and $\mathrm{l}(\sigma_0) = |\Phi^+|$.*

*Proof.* $\mathcal{W}$ acts simply transitively on the system of simple roots, so there is a unique $\sigma_0 \in \mathcal{W}$ such that $\sigma_0(\Delta) = -\Delta$. Since $\sigma_0^2(\Delta) = \Delta$, it follows that $\sigma_0^2 = id$. Also, $\sigma_0(\Phi^+) = \Phi^-$, so $\mathrm{l}(\sigma_0) = n(\sigma_0) = |\Phi^+|$. ☐

(vi) *Define a partial order on $E$ by $\mu \leq \lambda$ if $\lambda - \mu \in \mathbb{Z}_{\geq 0}\alpha_1 + \cdots + \mathbb{Z}_{\geq 0}\alpha_n$. If $\lambda \in \Lambda_W^+$, then $|\{\mu \in \Lambda_W^+ : \mu \leq \lambda\}|$ is finite.*

*Proof.* $\lambda \in \Lambda_W$, so $\lambda = r_1\alpha_1 + \cdots + r_n\alpha_n$, with $r_1, \ldots, r_n \in \mathbb{Q}$ (see Properties 1.2). It is enough to prove that if $\lambda \in \Lambda_W^+$, then $r_i \geq 0$ for any $i$; because if $\mu \in \Lambda_W^+$ and $\mu \leq \lambda$, then $\mu = s_1\alpha_1 + \cdots + s_n\alpha_n$ with $s_i \in \mathbb{Q}$, $s_i \geq 0$ and $r_i - s_i \in \mathbb{Z}_{\geq 0}$ for any $i$, and this gives a finite number of possibilities.

Hence, it is enough to prove the following result: Let $\{v_1, \ldots, v_n\}$ be a basis of an euclidean vector space with $(v_i|v_j) \leq 0$ for any $i \neq j$, and let $v \in E$ such that $(v|v_i) \geq 0$ for any $i = 1, \ldots, n$, then $v \in \mathbb{R}_{\geq 0}v_1 + \cdots + \mathbb{R}_{\geq 0}v_n$.

To prove this, assume $v \neq 0$ and write $v = r_1v_1 + \cdots + r_nv_n$. Then $0 < (v|v) = \sum_{i=1}^n (v|r_iv_i)$, so there is an $l$ such that $r_l > 0$. Then $v' = v - r_lv_l \in V' = \mathbb{R}v_1 \oplus \cdots \hat{l} \cdots \oplus \mathbb{R}v_n$, and $(v'|v_i) = (v|v_i) - r_l(v_l|v_i) \geq 0$ for any $i \neq l$. By an inductive argument we obtain $r_i \geq 0$ for any $i = 1, \ldots, n$. ☐

(vii) *For any $\mu \in \Lambda_W$, there is a unique $\lambda \in \Lambda_W^+ \cap \mathcal{W}\mu$. That is, for any $\mu \in \Lambda_W$, its orbit under the action of $\mathcal{W}$ intersects $\Lambda_W^+$ in exactly one weight.*

*Proof.* Let $\mu = m_1\lambda_1 + \cdots + m_n\lambda_n$, with $m_i = \langle\mu|\alpha_i\rangle \in \mathbb{Z}$, $i = 1, \ldots, n$. Let us prove that there is a $\lambda \in \Lambda_W^+ \cap \mathcal{W}\mu$. If $m_i \geq 0$ for any $i$, then we can take $\lambda = \mu$. Otherwise, if $m_j < 0$ for some $j$, then $\mu_1 = \sigma_j(\mu) = \mu - m_j\alpha_j$ satisfies that $\mu_1 > \mu$ and $\mu_1 \in \mathcal{W}\mu$. If $\mu_1 \in \Lambda_W^+$ we are done, otherwise we proceed now with $\mu_1$ and obtain a chain $\mu = \mu_0 < \mu_1 < \mu_2 < \cdots$, with $\mu_i \in \mathcal{W}\mu$. Since $\mathcal{W}$ is finite, this process must stop, so there is an $r$ such that $\mu_r \in \Lambda_W^+$. (This also proves $\mu \leq \mu_r$.)

To prove the uniqueness, it is enough to prove that if $\lambda, \mu \in \Lambda_W^+$ and there exists a $\sigma \in \mathcal{W}$ with $\sigma(\lambda) = \mu$, then $\lambda = \mu$. For this, take such a $\sigma$ of minimal length and consider a reduced expression for $\sigma$: $\sigma = \sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_t}$. If $t = 0$, $\sigma = id$ and $\lambda = \mu$. Otherwise, $t > 0$ and $\sigma(\beta_t) < 0$. Hence

$$0 \leq (\lambda|\beta_t) = \big(\sigma(\lambda)|\sigma(\beta_t)\big) = \big(\mu|\sigma(\beta_t)\big) \leq 0,$$

so $(\lambda|\beta_t) = 0$, $\sigma_{\beta_t}(\lambda) = \lambda$, and $\mu = \sigma(\lambda) = \sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_{t-1}}(\lambda)$, a contradiction with the minimality of the length. ☐

(viii) *Let $\lambda \in \Lambda_W^+$ be a dominant weight. Then $\sigma(\lambda) \leq \lambda$ for any $\sigma \in \mathcal{W}$. Moreover, its stabilizer $\mathcal{W}_\lambda = \{\sigma \in \mathcal{W} : \sigma(\lambda) = \lambda\}$ is generated by $\{\sigma_i : (\lambda|\alpha_i) = 0\}$. In particular, if $\lambda$ is* strictly dominant *(that is, $\langle\lambda|\alpha_i\rangle > 0$ for any $i = 1, \ldots, n$), then $\mathcal{W}_\lambda = 1$.*

*Proof.* Let $\sigma = \sigma_{i_1} \circ \cdots \circ \sigma_{i_t}$ be a reduced expression of $id \neq \sigma \in \mathcal{W}$, and let $\lambda_s = \sigma_{i_s} \circ \cdots \circ \sigma_{i_t}(\lambda)$, $0 \leq s \leq t$. Then, for any $1 \leq s \leq t$,

$$(\lambda_s|\alpha_{i_{s-1}}) = \big(\sigma_{i_s} \circ \cdots \circ \sigma_{i_t}(\lambda)|\alpha_{i_{s-1}}\big) = \big(\lambda|\sigma_{i_t} \circ \cdots \circ \sigma_{i_s}(\alpha_{i_{s-1}})\big)$$

and this is $\geq 0$, because item (ii) shows that $\sigma_{i_t} \circ \cdots \circ \sigma_{i_s}(\alpha_{i_{s-1}}) \in \Phi^+$. Hence

$$\lambda_{s-1} = \sigma_{i_{s-1}}(\lambda_s) = \lambda_s - \langle\lambda_s|\alpha_{i_{s-1}}\rangle\alpha_{i_{s-1}} \leq \lambda_s.$$

Therefore, $\sigma(\lambda) = \lambda_1 \leq \lambda_2 \leq \cdots \lambda_t = \lambda$ and $\sigma(\lambda) = \lambda$ if and only if $\lambda_s = \lambda$ for any $s$, if and only if $(\lambda|\alpha_{i_1}) = \cdots = (\lambda|\alpha_{i_t}) = 0$. □

(ix) *A subset* $\Pi$ *of* $\Lambda_W$ *is said to be* saturated *if for any* $\mu \in \Pi$, $\alpha \in \Phi$, *and* $i \in \mathbb{Z}$ *between* $0$ *and* $\langle\mu|\alpha\rangle$, $\mu - i\alpha \in \Pi$. *In particular,* $\Pi$ *is invariant under the action of* $\mathcal{W}$. *If, in addition, there is a dominant weight* $\lambda \in \Pi$ *such that any* $\mu \in \Pi$ *satisfies* $\mu \leq \lambda$, *then* $\Pi$ *is said to be* saturated with highest weight $\lambda$.
*Let* $\lambda \in \Lambda_W^+$. *Then the subset* $\Pi$ *is saturated with highest weight* $\lambda$ *if and only if* $\Pi = \cup_{\substack{\mu\in\Lambda_W^+\\ \mu\leq\lambda}} \mathcal{W}\mu$. *In particular,* $\Pi$ *is finite in this case.*

*Proof.* For $\lambda \in \Lambda_W^+$, let $\Pi_\lambda = \cup_{\substack{\mu\in\Lambda_W^+\\ \mu\leq\lambda}} \mathcal{W}\mu$.

If $\Pi$ is saturated with highest weight $\lambda$, and $\nu \in \Pi$, there is a $\sigma \in \mathcal{W}$ such that $\sigma(\nu) \in \Lambda_W^+$. But $\Pi$ is $\mathcal{W}$-invariant, so $\mu = \sigma(\nu) \in \Lambda_W^+ \cap \Pi$, and hence $\nu \in \mathcal{W}\mu$, with $\mu \in \Lambda_W^+$ and $\mu \leq \lambda$. Therefore, $\Pi \subseteq \Pi_\lambda$. To check that $\Pi = \Pi_\lambda$ it is enough to check that any $\mu \in \Lambda_W^+$ with $\mu < \lambda$, belongs to $\Pi$. But, if $\mu' = \mu + \sum_{i=1}^n m_i\alpha_i$ is any weight in $\Pi$ with $m_i \in \mathbb{Z}_{\geq 0}$ for any $i$, and $\mu' \neq \mu$ (that is $\mu' > \mu$), then $0 < (\mu' - \mu|\mu' - \mu)$, so there is an index $j$ such that $m_j > 0$ and $(\mu' - \mu|\alpha_j) > 0$. Now, since $\mu \in \Lambda_W^+$, $\langle\mu|\alpha_j\rangle \geq 0$, so $\langle\mu'|\alpha_j\rangle > 0$, and since $\Pi$ is saturated, $\mu'' = \mu' - \alpha_j \in \Pi$. Starting with $\mu' = \lambda$ and proceeding in this way, after a finite number of steps we obtain that $\mu \in \Pi$.

Conversely, we have to prove that for any $\lambda \in \Lambda_W^+$, $\Pi_\lambda$ is saturated. By its very definition, $\Pi_\lambda$ is $\mathcal{W}$-invariant. Let $\mu \in \Pi_\lambda$, $\alpha \in \Phi$ and $i \in \mathbb{Z}$ between $0$ and $\langle\mu|\alpha\rangle$. It has to be proven that $\mu - i\alpha \in \Pi_\lambda$. Take $\sigma \in \mathcal{W}$ such that $\sigma(\mu) \in \Lambda_W^+$. Since $\langle\sigma(\mu)|\sigma(\alpha)\rangle = \langle\mu|\alpha\rangle$, we may assume that $\mu \in \Lambda_W^+$. Also, changing if necessary $\alpha$ by $-\alpha$, we may assume that $\alpha \in \Phi^+$. Besides, with $m = \langle\mu|\alpha\rangle$, $\sigma_\alpha(\mu - i\alpha) = \mu - (m - i)\alpha$, so it is enough to assume that $0 < i \leq \lfloor\frac{m}{2}\rfloor$. Then $\langle\mu - i\alpha|\alpha\rangle = m - 2i \geq 0$.

If $\langle\mu - i\alpha|\alpha\rangle > 0$ and $\sigma \in \mathcal{W}$ satisfies $\sigma(\mu - i\alpha) \in \Lambda_W^+$, then $0 < \langle\sigma(\mu - i\alpha)|\sigma(\alpha)\rangle$, so $\sigma(\alpha) \in \Phi^+$ and $\sigma(\mu - i\alpha) = \sigma(\mu) - i\sigma(\alpha) < \sigma(\mu) \leq \mu$, since $\mu$ is dominant. Hence $\sigma(\mu - i\alpha) \in \Pi_\lambda$ and so does $\mu - i\alpha$. On the other hand, if $m$ is even and $i = \frac{m}{2}$, then $\langle\mu - i\alpha|\alpha\rangle = 0$. Take again a $\sigma \in \mathcal{W}$ such that $\sigma(\mu - i\alpha) \in \Lambda_W^+$. If $\sigma(\alpha) \in \Phi^+$, the same argument applies and $\sigma(\mu - i\alpha) < \sigma(\mu) \leq \mu$. But if $\sigma(\alpha) \in \Phi^-$, take $\tau = \sigma \circ \sigma_\alpha$, then $\tau(\mu - i\alpha) = \sigma(\mu - i\alpha) \in \Lambda_W^+$ and $\tau(\alpha) = \sigma(-\alpha) \in \Phi^+$, so again the same argument applies. □

(x) *Let* $\rho = \frac{1}{2}\sum_{\alpha\in\Phi^+}\alpha$ *be the* Weyl vector *(see Chapter 2, 6.1), and let* $\lambda \in \Lambda_W^+$ *and* $\mu \in \mathcal{W}\lambda$. *Then* $(\mu + \rho|\mu + \rho) \leq (\lambda + \rho|\lambda + \rho)$, *and they are equal if and only if* $\mu = \lambda$. *The same happens for any* $\mu \in \Lambda_W^+$ *with* $\mu \leq \lambda$. *Hence, in particular,* $(\mu + \rho|\mu + \rho) < (\lambda + \rho|\lambda + \rho)$ *for any* $\mu \in \Pi_\lambda \setminus \{\lambda\}$.

*Proof.* Since $\sigma_i(\rho) = \rho - \alpha_i$ for any $i = 1, \dots, n$, it follows that $\langle \rho | \alpha_i \rangle = 1$ for any $i$, so $\rho = \lambda_1 + \cdots + \lambda_n \in \Lambda_W^+$. Let $\mu \in \mathcal{W}\lambda \setminus \{\lambda\}$ and let $\sigma \in \mathcal{W}$ such that $\mu = \sigma(\lambda)$. Then,

$$(\lambda+\rho|\lambda+\rho) - (\mu+\rho|\mu+\rho) = (\lambda+\rho|\lambda+\rho) - (\sigma(\lambda)+\rho|\sigma(\lambda)+\rho) = 2(\lambda-\sigma(\lambda)|\rho).$$

But $\sigma(\lambda) < \lambda$ (item (viii)) and $\rho$ is strictly dominant, so $(\rho|\lambda - \sigma(\lambda)) > 0$, and the first assertion follows.

Now, if $\mu \in \Lambda_W^+$ with $\mu \leq \lambda$, then

$$(\lambda+\rho|\lambda+\rho) - (\mu+\rho|\mu+\rho) = (\lambda+\mu|\lambda-\mu) + 2(\lambda-\mu|\rho) \geq 0$$

since $\lambda + \mu \in \Lambda_W^+$, $\lambda - \mu \geq 0$ and $\rho$ is strictly dominant. Besides, this is 0 if and only if $\lambda - \mu = 0$. □

Later on, it will be proven that if $V$ is any irreducible finite dimensional module over $L$, then its set of weights $P(V)$ is a saturated set of weights.

## § 3. Universal enveloping algebra

In this section infinite dimensional vector spaces will be allowed.

Given a vector space $V$, recall that its *tensor algebra* is the direct sum

$$T(V) = k \oplus V \oplus (V \otimes_k V) \oplus \cdots \oplus V^{\otimes n} \oplus \cdots,$$

with the associative multiplication determined by

$$(v_1 \otimes \cdots \otimes v_n)(w_1 \otimes \cdots \otimes w_m) = v_1 \otimes \cdots \otimes v_n \otimes w_1 \otimes \cdots \otimes w_m.$$

Then $T(V)$ is a unital associative algebra over $k$.

Given the Lie algebra $L$, let $I$ be the ideal generated by the elements

$$x \otimes y - y \otimes x - [x,y] \in L \oplus (L \otimes L) \subseteq T(L),$$

where $x, y \in L$. The quotient algebra

$$U(L) = T(L)/I$$

is called the *universal enveloping algebra* of $L$. For $x_1, \dots, x_n \in L$, write $x_1 \cdots x_n = x_1 \otimes \cdots \otimes x_l + I$ and denote by juxtaposition the multiplication in $U(L)$. Let us denote by $\iota : L \to U(L)$ the natural map: $x \mapsto \iota(x) = x + I$. The universal property of the tensor algebra immediately gives:

**3.1 Universal property.** *Given a unital associative algebra $A$ over $k$, let $A^-$ be the Lie algebra defined on $A$ by means of $[x,y] = xy - yx$, for any $x, y \in A$. Then for any Lie algebra homomorphism $\varphi : L \to A^-$, there is a unique homomorphism of unital algebras $\phi : U(L) \to A$ such that the following diagram is commutative*

$$\begin{tikzcd}
L \arrow[r, "\iota"] \arrow[dr, "\varphi"'] \arrow[r, phantom, shift right=4ex, "\circlearrowleft"] & U(L) \arrow[d, dashed, "\phi"] \\
& A
\end{tikzcd}$$

**Remark.** The universal enveloping algebra makes sense for any Lie algebra, not just for the semisimple Lie algebras over algebraically closed fields of characteristic 0 considered in this chapter.

**3.2 Poincaré-Birkhoff-Witt Theorem.** *Let $L$ be an arbitrary Lie algebra over a field $k$ and let $\{x_i : i \in I\}$ be a basis of $L$ with ordered set $I$ (that is, $I$ is endowed with a total order). Then the 'monomials'*

$$x_{i_1} \cdots x_{i_n}, \quad n \geq 0,\ i_1, \dots, i_n \in I,\ i_1 < \cdots < i_n,\ e_1, \dots, e_n \in \mathbb{N},$$

*with the understanding that the empty product equals* 1*, is a basis of* $U(L)$*.*

*Proof.* It is clear that these monomials span $U(L)$, so we must show that they are linearly independent.

Given a monomial $x_{i_1} \otimes \cdots \otimes x_{i_n}$ in $T(L)$ define its *index* as the number of pairs $(j,k)$, with $1 \leq j < k \leq n$, such that $i_j > i_k$. Therefore, we must prove that the image in $U(L)$ of the monomials of index 0 are linearly independent.

Since the monomials form a basis of $T(L)$, a linear map $T(L) \to T(L)$ is determined by the images of the monomials. Also, $T(L)$ is the direct sum of the subspaces $T(L)_n$ spanned by the monomials of degree $n$ ($n \geq 0$). Define a linear map $\varphi : T(L) \to T(L)$ as follows:

$$\begin{aligned}
&\varphi(1) = 1, \quad \varphi(x_i) = x_i \ \forall i \in I, \\
&\varphi(x_{i_1} \otimes \cdots \otimes x_{i_n}) = x_{i_1} \otimes \cdots \otimes x_{i_n} \quad \text{if } n \geq 2 \text{ and the index is } 0,
\end{aligned}$$

and, with $n, s \geq 2$, assuming $\varphi$ has been defined for monomials of degree $< n$ (hence in $\bigoplus_{r=0}^{n-1} T(L)_r$), and for monomials of degree $n$ and index $< s$, define

$$\begin{aligned}
\varphi(x_{i_1} \otimes \cdots \otimes x_{i_n}) &= \varphi(x_{i_1} \otimes \cdots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \cdots \otimes x_{i_n}) \\
&\qquad + \varphi(x_{i_1} \otimes \cdots \otimes [x_{i_j}, x_{i_{j+1}}] \otimes \cdots \otimes x_{i_n}),
\end{aligned} \tag{3.2}$$

if the index of $x_{i_1} \otimes \cdots \otimes x_{i_n}$ is $s$ and $j$ is the lowest index such that $i_j > i_{j+1}$. (Note that the index of $x_{i_1} \otimes \cdots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \cdots \otimes x_{i_n}$ is $s-1$ and $x_{i_1} \otimes \cdots \otimes [x_{i_j}, x_{i_{j+1}}] \otimes \cdots \otimes x_{i_n} \in T(L)_{n-1}$, so the right hand side of (3.2) is well defined.)

Let us prove that $\varphi$ satisfies the condition in (3.2) for any $n \geq 2$, any monomial $x_{i_1} \otimes \cdots \otimes x_{i_n}$, and any index $1 \leq j \leq n-1$ with $i_j > i_{j+1}$. (If this is true then, by anticommutativity, (3.2) is satisfied for any monomial of degree $n \geq 2$ and any index $1 \leq j \leq n-1$.)

This is trivial if the index of $x_{i_1} \otimes \cdots \otimes x_{i_n}$ is 1. In particular for $n = 2$. Assume this is true for degree $< n$ and for degree $n$ and index $< s$, with $n \geq 3$ and $s \geq 2$. If the index of $x_{i_1} \otimes \cdots \otimes x_{i_n}$ is $s$, let $j$ be the lowest index with $i_j > i_{j+1}$ and let $j'$ be another index with $i_{j'} > i_{j'+1}$.

If $j+1<j'$, then $n\geq 4$ and the indices $j, j+1, j', j'+1$ are different. Then,

$$\begin{aligned}
&\varphi(x_{i_1}\otimes\cdots\otimes x_{i_n})\\
&\quad=\varphi(x_{i_1}\otimes\cdots\otimes x_{i_{j+1}}\otimes x_{i_j}\otimes\cdots\otimes x_{i_n})+\varphi(x_{i_1}\otimes\cdots\otimes[x_{i_j},x_{i_{j+1}}]\otimes\cdots\otimes x_{i_n}),\\
&\quad=\varphi(x_{i_1}\otimes\cdots\otimes x_{i_{j+1}}\otimes x_{i_j}\otimes\cdots\otimes x_{i_{j'+1}}\otimes x_{i_{j'}}\otimes\cdots\otimes x_{i_n})\\
&\qquad+\varphi(x_{i_1}\otimes\cdots\otimes x_{i_{j+1}}\otimes x_{i_j}\otimes\cdots\otimes[x_{i_{j'}},x_{i_{j'+1}}]\otimes\cdots\otimes x_{i_n})\\
&\qquad+\varphi(x_{i_1}\otimes\cdots\otimes[x_{i_j},x_{i_{j+1}}]\otimes\cdots\otimes x_{i_{j'+1}}\otimes x_{i_{j'}}\otimes\cdots\otimes x_{i_n})\\
&\qquad+\varphi(x_{i_1}\otimes\cdots\otimes[x_{i_j},x_{i_{j+1}}]\otimes\cdots\otimes[x_{i_{j'}},x_{i_{j'+1}}]\otimes\cdots\otimes x_{i_n})\\
&\quad=\varphi(x_{i_1}\otimes\cdots\otimes x_{i_{j'+1}}\otimes x_{i_{j'}}\otimes\cdots\otimes x_{i_n})+\varphi(x_{i_1}\otimes\cdots\otimes[x_{i_{j'}},x_{i_{j'+1}}]\otimes\cdots\otimes x_{i_n})
\end{aligned}$$

The first equality works by definition of $\varphi$ and the second and third because the result is assumed to be valid for degree $n$ and index $<s$ and for degree $<n$.

Finally, if $j+1=j'$ nothing is lost in the argument if we assume $j=1$, $j'=2$, and $n=3$. Write $x_{i_1}=x$, $x_{i_2}=y$, and $x_{i_3}=z$. Hence we have:

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

by definition of $\varphi$ and because (3.2) is valid for lower index. But since $\varphi$ satisfies (3.2) in degree $<n$,

$$\begin{aligned}
&\varphi([y,z]\otimes x)+\varphi(y\otimes[x,z])+\varphi([x,y]\otimes z)\\
&\qquad=\varphi([[y,z],x])+\varphi([y,[x,z]])+\varphi([[x,y],z])\\
&\qquad\qquad+\varphi(x\otimes[y,z])+\varphi([x,z]\otimes y)+\varphi(z\otimes[x,y])\\
&\qquad=\varphi(x\otimes[y,z])+\varphi([x,z]\otimes y)+\varphi(z\otimes[x,y]),
\end{aligned}$$

because $[[y,z],x]+[y,[x,z]]+[[x,y],z]=[[y,z],x]+[[z,x],y]+[[x,y],z]=0$. Thus,

$$\begin{aligned}
&\varphi(x\otimes y\otimes z)\\
&\quad=\varphi(z\otimes y\otimes x)+\varphi(z\otimes[x,y])+\varphi([x,z]\otimes y)+\varphi(x\otimes[y,z])\\
&\quad=\varphi(z\otimes x\otimes y)+\varphi([x,z]\otimes y)+\varphi(x\otimes[y,z])\\
&\quad=\varphi(x\otimes z\otimes y)+\varphi(x\otimes[y,z]),
\end{aligned}$$

because (3.2) is valid for index $<s$.

Therefore (3.2) is satisfied for any $n\geq 2$, any monomial $x_{i_1}\otimes\cdots\otimes x_{i_n}$, and any $1\leq j\leq n-1$. Since the ideal $I$ is spanned by the elements

$$x_{i_1}\otimes\cdots\otimes x_{i_n}-x_{i_1}\otimes\cdots\otimes x_{i_{j+1}}\otimes x_{i_j}\otimes\cdots\otimes x_{i_n}-x_{i_1}\otimes\cdots\otimes[x_{i_j},x_{i_{j+1}}]\otimes\cdots\otimes x_{i_n},$$

it follows that $\varphi(I)=0$. On the other hand, $\varphi$ is the identity on the span of the monomials of index 0, so the linear span of these monomials intersects $I$ trivially, and hence it maps bijectively on $U(L)$, as required. □

**3.3 Corollary.** *The natural map $\iota: L\to U(L)$ is one-to-one.*

If $S$ is a subalgebra of a Lie algebra $L$, then the inclusion map $S \hookrightarrow L \xrightarrow{\iota} U(L)$ extends to a homomorphism $U(S) \to U(L)$, which is one-to-one by Theorem 3.2 (as any ordered basis of $S$ can be extended to an ordered basis of $L$). In this way, $U(S)$ will be identified to a subalgebra of $U(L)$.

Moreover, if $L = S \oplus T$ for subalgebras $S$ and $T$, then the union of an ordered basis of $S$ and an ordered basis of $T$ becomes an ordered basis of $L$ by imposing that the elements of $S$ are lower than the elements in $T$. Then Theorem 3.2 implies that the linear map $U(S) \otimes_k U(T) \to U(L)$, $x \otimes y \mapsto xy$, is an isomorphism of vector spaces.

## § 4. Irreducible representations

By the universal property of of $U(L)$, any representation $\phi : L \to \mathfrak{gl}(V)$ induces a representation of $U(L)$: $\tilde{\phi} : U(L) \to \mathrm{End}_k(V)$. Therefore a module for $L$ is the same thing as a left module for the associative algebra $U(L)$.

Now, given a linear form $\lambda \in H^*$, consider:

- $J(\lambda) = \sum_{\alpha\in\Phi^+} U(L)x_\alpha + \sum_{i=1}^n U(L)(h_i - \lambda(h_i)1)$, which is a left ideal of $U(L)$, where $h_i = h_{\alpha_i}$ for any $i = 1, \dots, n$.

  Theorem 3.2 implies that $J(\lambda) \neq U(L)$. Actually, $L = L^- \oplus B$, where $B = H \oplus L^+$ ($B$ is called a Borel subalgebra), so $U(L)$ is linearly isomorphic to $U(L^-)\otimes_k U(B)$. Then with $\tilde{J}(\lambda) = \sum_{\alpha\in\Phi^+} U(B)x_\alpha + \sum_{i=1}^n U(B)(h_i-\lambda(h_i)1)$ (a left ideal of $U(B)$), we get $J(\lambda) = U(L^-)\tilde{J}(\lambda)$. Now the Lie algebra homomorphism $\rho : B \to k$ such that $\rho(x_\alpha) = 0$, for any $\alpha \in \Phi^+$, and $\rho(h_i) = \lambda(h_i)$, for any $i = 1, \dots, n$, extends to a homomorphism of unital algebras $\tilde{\rho} : U(B) \to k$, and $\tilde{J}(\lambda) \subseteq \ker\tilde{\rho}$. Hence $\tilde{J}(\lambda) \neq U(B)$ and, therefore, $J(\lambda) \neq U(L)$.

- $M(\lambda) = U(L)/J(\lambda)$, which is called the associated *Verma module*. (It is a left module for $U(L)$, hence a module for $L$.)

- $\theta : U(L) \to M(\lambda)$, $u \mapsto u + J(\lambda)$, the canonical homomorphism of modules.

- $m_\lambda = \theta(1) = 1 + J(\lambda)$, the canonical generator: $M(\lambda) = U(L)m_\lambda$.

Then $x_\alpha \in J(\lambda)$ for any $\alpha \in \Phi^+$, so $x_\alpha m_\lambda = 0$. Therefore, $L^+.m_\lambda = 0$. Also, $h_i - \lambda(h_i)1 \in J(\lambda)$, so $h_i m_\lambda = \lambda(h_i)m_\lambda$ for any $i$, and hence $hm_\lambda = \lambda(h)m_\lambda$ for any $h \in H$.

Therefore, as in the proof of Proposition 1.6

$$M(\lambda) = km_\lambda + \sum_{r=1}^{\infty}\sum k\big(y_{\alpha_{i_1}}.(y_{\alpha_{i_2}}\cdots(y_{\alpha_{i_r}}.m_\lambda))\big).$$

(Note that $y_{\alpha_{i_1}}.(y_{\alpha_{i_2}}\cdots(y_{\alpha_{i_r}}.m_\lambda)) \in M(\lambda)_{\lambda-\alpha_{i_1}-\cdots-\alpha_{i_r}}$.)

Let $K(\lambda)$ be the sum of all the proper submodules of $M(\lambda)$. Then $m_\lambda \notin K(\lambda)$, so $K(\lambda) \neq M(\lambda)$ and $V(\lambda) = M(\lambda)/K(\lambda)$ is an irreducible $L$-module (although, in general, of infinite dimension). However,

$$\dim M(\lambda)_{\lambda-\alpha_{i_1}-\cdots-\alpha_{i_r}} \leq |\{(\beta_{j_1}, \dots, \beta_{j_r}) \in \Delta^r : \beta_{j_1} + \cdots + \beta_{j_r} = \alpha_{i_1} + \cdots + \alpha_{i_r}\}| \leq r!,$$

so for any $\mu \in H^*$, the dimension of the weight space $V(\lambda)_\mu$ is finite.

**4.1 Theorem.** *For any $\lambda \in H^*$, $\dim V(\lambda)$ is finite if and only if $\lambda \in \Lambda_W^+$.*

*Proof.* The vector $v_\lambda = m_\lambda + K(\lambda)$ is a highest weight vector of $V(\lambda)$ of weight $\lambda$, and hence by Proposition 1.6, if $\dim V(\lambda)$ is finite, then $\lambda \in \Lambda_W^+$.

Conversely, assume that $\lambda \in \Lambda_W^+$. Let $x_i = x_{\alpha_i}$, $y_i = y_{\alpha_i}$ and $h_i = h_{\alpha_i}$, $i = 1, \dots, n$, be the standard generators of $L$. Denote by $\phi : L \to \mathfrak{gl}\big(V(\lambda)\big)$ the associated representation. For any $i = 1, \dots, n$, $m_i = \langle \lambda | \alpha_i \rangle \in \mathbb{Z}_{\geq 0}$, because $\lambda$ is dominant. Several steps will be followed now:

1. $\phi(y_i)^{m_i+1}(v_\lambda) = 0$ for any $i = 1, \dots, n$.

   *Proof.* Let $u_i = \phi(y_i)^{m_i+1}(v_\lambda) = y_i^{m_i+1} v_\lambda$ (as usual we denote by juxtaposition the action of an associative algebra, in this case $U(L)$, on a left module, here $V(\lambda)$). For any $j \neq i$, $[x_j, y_i] = 0$, so $x_j u_i = y_i^{m_i+1}(x_j v_\lambda) = 0$. By induction, it is checked that in $U(L)$, $x_i y_i^{m+1} = y_i^{m+1} x_i + (m+1) y_i^m h_i - m(m+1) y_i^m$ for any $m \in \mathbb{Z}_{\geq 0}$. Hence

   $$\begin{aligned} x_i u_i = x_i y_i^{m_i+1} v_\lambda &= y_i^{m_i+1} x_i v_\lambda + (m_i+1) y_i^{m_i} h_i v_\lambda - m_i(m_i+1) y_i^{m_i} v_\lambda \\ &= 0 + y_i^{m_i}\Big((m_i+1)\lambda(h_i) v_\lambda - m_i(m_i+1) v_\lambda\Big) = 0 \end{aligned}$$

   Thus $L^+ u_i = 0$ and hence $u_i$ is a highest weight vector of weight $\mu_i = \lambda - (m_i+1)\alpha_i$. Then $W_i = k u_i + \sum_{i=1}^r \sum k\big(y_{\alpha_{i_1}} \cdots y_{\alpha_{i_r}} u_i\big)$ is a proper submodule of $V(\lambda)$, and hence it is 0. In particular $u_i = 0$, as required. □

2. Let $S_i = L_{\alpha_i} \oplus L_{-\alpha_i} \oplus [L_{\alpha_i}, L_{-\alpha_i}] = k x_i + k y_i + k h_i$, which is a subalgebra of $L$ isomorphic to $\mathfrak{sl}_2(k)$. Then $V(\lambda)$ is a sum of finite dimensional $S_i$-submodules.

   *Proof.* The linear span of $v_\lambda, y_i v_\lambda, \dots, y_i^{m_i} v_\lambda$ is an $S_i$-submodule. Hence the sum $V'$ of the finite dimensional $S_i$-submodules of $V(\lambda)$ is not 0. But if $W$ is a finite dimensional $S_i$-submodule, then consider $\bar{W} = LW = \sum zW$, where $z$ runs over a fixed basis of $L$. Hence $\dim \bar{W} < \infty$. But for any $w \in W$, $x_i(zw) = [x_i, z]w + z(x_i w) \in LW = \bar{W}$ and, also, $y_i(zw) \in LW = \bar{W}$. Therefore, $\bar{W} = LW$ is a finite dimensional $S_i$-submodule, and hence contained in $V'$. Thus, $V'$ is a nonzero $L$-submodule of the irreducible module $V(\lambda)$, so $V' = V(\lambda)$, as required. □

3. The set of weights $P\big(V(\lambda)\big)$ is invariant under $\mathcal{W}$.

   *Proof.* It is enough to see that $P\big(V(\lambda)\big)$ is invariant under $\sigma_i$, $i = 1, \dots, n$. Let $\mu \in P\big(V(\lambda)\big)$ and $0 \neq v \in V(\lambda)_\mu$. Then $v \in V(\lambda) = V'$, so by complete reducibility (Weyl's Theorem, Chapter 2, 2.5) there are finite dimensional $S_i$-submodules $W_1, \dots, W_m$ such that $v \in W_1 \oplus \cdots \oplus W_m$. Thus, $v = w_1 + \cdots + w_m$, with $w_j \in W_j$ for any $j$, and we may assume that $w_1 \neq 0$. Since $h_i v = \mu(h_i) v$, it follows that $h_i w_j = \mu(h_i) w_j$ for any $j$. Hence $\mu(h_i)$ is an eigenvalue of $h_i$ in $W_1$, and the representation theory of $\mathfrak{sl}_2(k)$ shows that $-\mu(h_i)$ is another eigenvalue. Besides, if $\mu(h_i) \geq 0$, then $0 \neq y_i^{\mu(h_i)} w_1 \in (W_1)_{-\mu(h_i)}$, so $0 \neq y_i^{\mu(h_i)} v \in V(\lambda)_{\mu - \mu(h_i)\alpha_i} = V(\lambda)_{\sigma_i(\mu)}$; while if $\mu(h_i) < 0$, then $0 \neq x_i^{-\mu(h_i)} w_1 \in (W_1)_{-\mu(h_i)}$, so $0 \neq x_i^{-\mu(h_i)} v \in V(\lambda)_{\mu - \mu(h_i)\alpha_i} = V(\lambda)_{\sigma_i(\mu)}$. In any case $\sigma_i(\mu) \in P\big(V(\lambda)\big)$. □

4. For any $\mu \in P\big(V(\lambda)\big) \subseteq \{\lambda - \alpha_{i_1} - \cdots - \alpha_{i_r} : r \geq 0,\ 1 \leq i_1, \ldots, i_r \leq n\} \subseteq \Lambda_W$, there is a $\sigma \in \mathcal{W}$ such that $\sigma(\mu) \in \Lambda_W^+$. Hence, by the previous item, $\sigma(\mu) \in P\big(V(\lambda)\big)$, so $\sigma(\mu) \leq \lambda$. Therefore, $P\big(V(\lambda)\big) \subseteq \cup_{\substack{\mu \in \Lambda_W^+ \\ \mu \leq \lambda}} \mathcal{W}\mu$. Hence $P\big(V(\lambda)\big)$ is finite, and since all the weight spaces of $V(\lambda)$ are finite dimensional, we conclude that $V(\lambda)$ is finite dimensional. □

**4.2 Corollary.** *The map*

$$\begin{aligned} \Lambda_W^+ &\to \{\textit{isomorphism classes of finite dimensional irreducible } L\textit{-modules}\} \\ \lambda &\mapsto \quad \textit{the class of } V(\lambda), \end{aligned}$$

*is a bijection.*

**4.3 Proposition.** *For any $\lambda \in \Lambda_W^+$, $P\big(V(\lambda)\big) = \Pi_\lambda$ (the saturated set of weights with highest weight $\lambda$, recall Properties 2.1).*

*Moreover, for any $\mu \in P\big(V(\lambda)\big)$, $\dim V(\lambda)_\mu = \dim V(\lambda)_{\sigma(\mu)}$ for any $\sigma \in \mathcal{W}$, and $(\mu + \rho|\mu + \rho) \leq (\lambda + \rho|\lambda + \rho)$, with equality only if $\mu = \lambda$.*

*Proof.* For any $\mu \in P\big(V(\lambda)\big)$ and any $\alpha \in \Phi$, $\oplus_{m \in \mathbb{Z}} V(\lambda)_{\mu + m\alpha}$ is a module for $S_\alpha = L_\alpha \oplus L_{-\alpha} \oplus [L_\alpha, L_{-\alpha}] \cong \mathfrak{sl}_2(k)$. Hence its weights form a chain: the *$\alpha$-string of $\mu$*: $\mu + q\alpha, \ldots, \mu, \ldots, \mu - r\alpha$ with $\langle \mu|\alpha \rangle = r - q$. Therefore, $P\big(V(\lambda)\big)$ is a saturated set of weights with highest weight $\lambda$, and thus $P\big(V(\lambda)\big) = \Pi_\lambda$ by Properties 2.1.

The last part also follows from Properties 2.1.

Now, if $\phi : L \to \mathfrak{gl}\big(V(\lambda)\big)$ is the associated representation, for any $\alpha \in \Phi^+$, $\operatorname{ad}\phi(x_\alpha) \in \operatorname{End}_k\big(\phi(L)\big)$ and $\phi(x_\alpha) \in \operatorname{End}_k\big(V(\lambda)\big)$ are nilpotent endomorphisms. Moreover, $\operatorname{ad}\phi(x_\alpha)$ is a derivation of the Lie algebra $\phi(L)$. Hence $\exp\big(\operatorname{ad}\phi(x_\alpha)\big)$ is an automorphism of $\phi(L)$, while $\exp\phi(x_\alpha) \in GL\big(V(\lambda)\big)$. The same applies to $\phi(y_\alpha)$. Consider the maps:

$$\begin{aligned} \tau_\alpha &= \exp\big(\operatorname{ad}\phi(x_\alpha)\big)\exp\big(-\operatorname{ad}\phi(y_\alpha)\big)\exp\big(\operatorname{ad}\phi(x_\alpha)\big) \in \operatorname{Aut}\phi(L), \\ \eta_\alpha &= \exp\phi(x_\alpha)\exp(-\phi(y_\alpha))\exp\phi(x_\alpha) \in GL\big(V(\lambda)\big). \end{aligned}$$

In $\operatorname{End}_k(V(\lambda))$, $\exp\big(\operatorname{ad}\phi(x_\alpha)\big) = \exp\big(L_{\phi(x_\alpha)} - R_{\phi(x_\alpha)}\big) = \exp L_{\phi(x_\alpha)} \exp\big(-R_{\phi(x_\alpha)}\big)$, where $L_a$ and $R_a$ denote the left and right multiplication by the element $a \in \operatorname{End}_k\big(V(\lambda)\big)$, which are commuting endomorphisms. Hence, for any $z \in L$,

$$\exp\big(\operatorname{ad}\phi(x_\alpha)\big)(\phi(z)) = \big(\exp\phi(x_\alpha)\big)\phi(z)\big(\exp(-\phi(x_\alpha))\big)$$

and $\tau_\alpha(\phi(z)) = \eta_\alpha \phi(z) \eta_\alpha^{-1}$.

For any $h \in H$, $\exp\big(\operatorname{ad}\phi(x_\alpha)\big)(\phi(h)) = \phi(h) + [\phi(x_\alpha), \phi(h)] = \phi(h - \alpha(h)x_\alpha)$. Hence

$$\begin{aligned} &\exp\big(-\operatorname{ad}\phi(y_\alpha)\big)\exp\big(\operatorname{ad}\phi(x_\alpha)\big)(\phi(h)) \\ &\qquad = \phi(h - \alpha(h)x_\alpha) + \phi\big([h - \alpha(h)x_\alpha, y_\alpha]\big) + \tfrac{1}{2}\phi\big([[h - \alpha(h)x_\alpha, y_\alpha], y_\alpha]\big) \\ &\qquad = \phi\Big(h - \alpha(h)x_\alpha - \alpha(h)y_\alpha - \alpha(h)h_\alpha + \tfrac{1}{2}2\alpha(h)y_\alpha\Big) \\ &\qquad = \phi\big(h - \alpha(h)h_\alpha - \alpha(h)x_\alpha\big) \end{aligned}$$

and, finally,

$$\begin{aligned}\tau_\alpha\bigl(\phi(h)\bigr) &= \exp\bigl(\operatorname{ad}\phi(x_\alpha)\bigr)\bigl(h-\alpha(h)h_\alpha-\alpha(h)x_\alpha\bigr)\\ &= \phi\Bigl(h-\alpha(h)h_\alpha-\alpha(h)x_\alpha+[x_\alpha,h-\alpha(h)h_\alpha-\alpha(h)x_\alpha]\Bigr)\\ &= \phi\Bigl(h-\alpha(h)h_\alpha-\alpha(h)x_\alpha-\alpha(h)x_\alpha+2\alpha(h)x_\alpha\Bigr)\\ &= \phi\bigl(h-\alpha(h)h_\alpha\bigr),\end{aligned}$$

so for any $0\neq v\in V(\lambda)_\mu$ and $h\in H$,

$$\sigma_\alpha(\mu)(h)v=\bigl(\mu-\langle\mu|\alpha\rangle\alpha\bigr)(h)v=\mu\bigl(h-\alpha(h)h_\alpha\bigr)v=\tau_\alpha\bigl(\phi(h)\bigr)(v)=\eta_\alpha\phi(h)\bigl(\eta_\alpha^{-1}(v)\bigr).$$

That is, $\phi(h)\bigl(\eta_\alpha^{-1}(v)\bigr)=\sigma_\alpha(\mu)(h)\eta_\alpha^{-1}(v)$ for any $h\in H$, so $\eta_\alpha^{-1}(v)\in V(\lambda)_{\sigma_\alpha(\mu)}$ and $\eta_\alpha^{-1}\bigl(V(\lambda)_\mu\bigr)\subseteq V(\lambda)_{\sigma_\alpha(\mu)}$. But also, $\eta_\alpha^{-1}\bigl(V(\lambda)_{\sigma_\alpha(\mu)}\bigr)\subseteq V(\lambda)_{\sigma_\alpha^2(\mu)}=V(\lambda)_\mu$. Therefore, $\eta_\alpha\bigl(V(\lambda)_\mu\bigr)=V(\lambda)_{\sigma_\alpha(\mu)}$ and both weight spaces have the same dimension. □

## § 5. Freudenthal's multiplicity formula

Given a dominant weight $\lambda\in\Lambda_W^+$ and a weight $\mu\in\Lambda_W$, the dimension of the associated weight space, $m_\mu=\dim V(\lambda)_\mu$, is called the *multiplicity* of $\mu$ in $V(\lambda)$. Of course, $m_\mu=0$ unless $\mu\in P\bigl(V(\lambda)\bigr)$.

The multiplicity formula due to Freudenthal gives a recursive method to compute these multiplicities:

**5.1 Theorem.** ***(Freudenthal's multiplicity formula, 1954)*** *For any $\lambda\in\Lambda_W^+$ and $\mu\in\Lambda_W$:*

$$\Bigl((\lambda+\rho|\lambda+\rho)-(\mu+\rho|\mu+\rho)\Bigr)m_\mu=2\sum_{\alpha\in\Phi^+}\sum_{j=1}^{\infty}(\mu+j\alpha|\alpha)m_{\mu+j\alpha}.$$

(Note that the sum above is finite since there are only finitely many weights in $P\bigl(V(\lambda)\bigr)$. Also, starting with $m_\lambda=1$, and using Proposition 4.3, this formula allows the recursive computation of all the multiplicities.)

*Proof.* Let $\phi:L\to\mathfrak{gl}\bigl(V(\lambda)\bigr)$ be the associated representation and denote also by $\phi$ the representation of $U(L)$, $\phi:U(L)\to\operatorname{End}_k\bigl(V(\lambda)\bigr)$. Let $\{a_1,\dots,a_m\}$ and $\{b_1,\dots,b_m\}$ be dual bases of $L$ relative to the Killing form (that is, $\kappa(a_i,b_j)=\delta_{ij}$ for any $i,j$). Then for any $z\in L$, $[a_i,z]=\sum_{j=1}^m\alpha_i^ja_j$ for any $i$ and $[b_j,z]=\sum_{i=1}^m\beta_j^ib_i$ for any $j$. Hence, inside $U(L)$,

$$\sum_{i=1}^m[a_ib_i,z]=\sum_{i=1}^m\bigl([a_i,z]b_i+a_i[b_i,z]\bigr)=\sum_{i,j=1}^n\bigl(\alpha_i^j+\beta_j^i\bigr)a_jb_i,$$

but

$$0=\kappa([a_i,z],b_j)+\kappa(a_i,[b_j,z])=\alpha_i^j+\beta_j^i,$$

so $[\sum_{i=1}^m a_ib_i,L]=0$. Therefore, the element $c=\sum_{i=1}^m a_ib_i$ is a central element in $U(L)$, which is called the *universal Casimir element* (recall that a Casimir element was used in the proof of Weyl's Theorem, Chapter 2, 2.5). By the well-known Schur's Lemma, $\phi(c)$ is a scalar.

Take a basis $\{g_1, \dots, g_n\}$ of $H$ with $\kappa(g_i, g_j) = \delta_{ij}$ for any $i, j$. For any $\mu \in H^*$, let $t_\mu \in H$, such that $\kappa(t_\mu, .\,) = \mu$. Then $t_\mu = r_1 g_1 + \cdots + r_n g_n$, with $r_i = \kappa(t_\mu, g_i) = \mu(g_i)$ for any $i$. Hence,

$$(\mu|\mu) = \mu(t_\mu) = \sum_{i=1}^{n} r_i \mu(g_i) = \sum_{i=1}^{n} \mu(g_i)^2.$$

For any $\alpha \in \Phi^+$, take $x_\alpha \in L_\alpha$ and $x_{-\alpha} \in L_{-\alpha}$ such that $\kappa(x_\alpha, x_{-\alpha}) = 1$ (so that $[x_\alpha, x_{-\alpha}] = t_\alpha$). Then the element

$$c = \sum_{i=1}^{n} g_i^2 + \sum_{\alpha\in\Phi^+} (x_\alpha x_{-\alpha} + x_{-\alpha} x_\alpha) = \sum_{i=1}^{n} g_i^2 + \sum_{\alpha\in\Phi^+} t_\alpha + 2 \sum_{\alpha\in\Phi^+} x_{-\alpha} x_\alpha$$

is a universal Casimir element.

Let $0 \neq v_\lambda \in V(\lambda)_\lambda$ be a highest weight vector, then since $x_\alpha v_\lambda = 0$ for any $\alpha \in \Phi^+$,

$$\phi(c) v_\lambda = \Bigl(\sum_{i=1}^{n} \lambda(g_i)^2 + \sum_{\alpha\in\Phi^+} \lambda(t_\alpha)\Bigr) v_\lambda = \Bigl((\lambda|\lambda) + 2(\lambda|\rho)\Bigr) v_\lambda = (\lambda|\lambda + 2\rho) v_\lambda,$$

because $2\rho = \sum_{\alpha\in\Phi^+} \alpha$. Therefore, since $\phi(c)$ is a scalar, $\phi(c) = (\lambda|\lambda + 2\rho) id$.

For simplicity, write $V = V(\lambda)$. Then $\operatorname{trace}_{V_\mu} \phi(c) = (\lambda|\lambda + 2\rho) m_\mu$.

Also, for any $v \in V_\mu$,

$$\begin{aligned}
\phi(c) v &= \Bigl(\sum_{i=1}^{n} \mu(g_i)^2\Bigr) v + \Bigl(\sum_{\alpha\in\Phi^+} \mu(t_\alpha)\Bigr) v + 2 \Bigl(\sum_{\alpha\in\Phi^+} \phi(x_{-\alpha}) \phi(x_\alpha)\Bigr) v \\
&= (\mu|\mu + 2\rho) v + 2 \sum_{\alpha\in\Phi^+} \phi(x_{-\alpha}) \phi(x_\alpha) v.
\end{aligned}$$

Recall that if $f : U_1 \to U_2$ and $g : U_2 \to U_1$ are linear maps between finite dimensional vector spaces, then $\operatorname{trace}_{U_1} gf = \operatorname{trace}_{U_2} fg$. In particular,

$$\begin{aligned}
\operatorname{trace}_{V_\mu} \phi(x_{-\alpha}) \phi(x_\alpha) &= \operatorname{trace}_{V_{\mu+\alpha}} \phi(x_\alpha) \phi(x_{-\alpha}) \\
&= \operatorname{trace}_{V_{\mu+\alpha}} \Bigl(\phi(t_\alpha) + \phi(x_{-\alpha}) \phi(x_\alpha)\Bigr) \\
&= (\mu + \alpha|\alpha) m_{\mu+\alpha} + \operatorname{trace}_{V_{\mu+\alpha}} \phi(x_{-\alpha}) \phi(x_\alpha) \\
&= \sum_{j=1}^{\infty} (\mu + j\alpha|\alpha) m_{\mu+j\alpha}.
\end{aligned}$$

(The argument is repeated until $V_{\mu+j\alpha} = 0$ for large enough $j$.) Therefore,

$$(\lambda|\lambda + 2\rho) m_\mu = (\mu|\mu + 2\rho) m_\mu + 2 \sum_{\alpha\in\Phi^+} \sum_{j=1}^{\infty} (\mu + j\alpha|\alpha) m_{\mu+j\alpha},$$

and this is equivalent to Freudenthal's multiplicity formula. □

**5.2 Remark.** Freudenthal's multiplicity formula remains valid if the inner product is scaled by a nonzero factor.

**5.3 Example.** Let $L$ be the simple Lie algebra of type $G_2$ and write $\Delta = \{\alpha, \beta\}$. The Cartan matrix is $\left(\begin{smallmatrix} 2 & -1 \\ -3 & 2 \end{smallmatrix}\right)$, so we may scale the inner product so that $(\alpha|\alpha) = 2$, $(\beta|\beta) = 6$ and $(\alpha|\beta) = -3$. The set of positive roots is (check it!):

$$\Phi^+ = \{\alpha, \beta, \alpha + \beta, 2\alpha + \beta, 3\alpha + \beta, 3\alpha + 2\beta\}.$$

Let $\lambda_1, \lambda_2$ be the fundamental dominant weights, so:

$$\begin{aligned} &\langle \lambda_1|\alpha\rangle = 1,\ \langle \lambda_1|\beta\rangle = 0, \quad \text{so} \quad \lambda_1 = 2\alpha + \beta, \\ &\langle \lambda_2|\alpha\rangle = 0,\ \langle \lambda_2|\beta\rangle = 1, \quad \text{so} \quad \lambda_2 = 3\alpha + 2\beta. \end{aligned}$$

Consider the dominant weight $\lambda = \lambda_1 + \lambda_2 = 5\alpha + 3\beta$ (see figure 5.1).

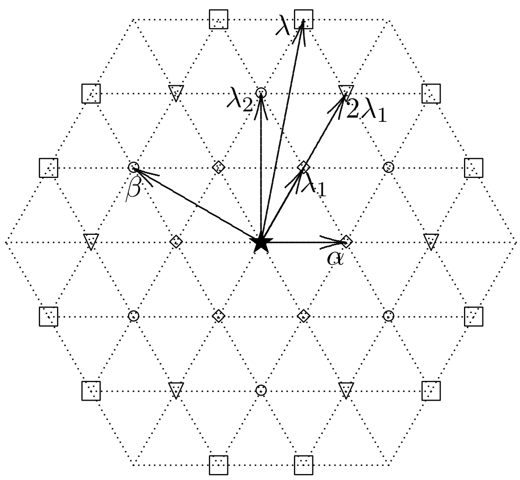



Figure 5.1: $G_2$, roots and weights

Then

$$\{\mu \in \Lambda_W^+ : \mu \leq \lambda\} = \{\lambda, 2\lambda_1, \lambda_2, \lambda_1, 0\},$$

and in order to compute the weight multiplicities of $V(\lambda)$ it is enough to compute $m_\lambda = 1$, $m_{2\lambda_1}$, $m_{\lambda_2}$, $m_{\lambda_1}$ and $m_0$.

The Weyl group $\mathcal{W}$ is generated by $\sigma_\alpha$ and $\sigma_\beta$, which are the reflections along the lines trough the origin and perpendicular to $\alpha$ and $\beta$ respectively. The composition $\sigma_\alpha\sigma_\beta$ is the counterclockwise rotation of angle $\frac{\pi}{3}$. Thus $\mathcal{W}$ is easily seen to be the dihedral group of order 12. Therefore, $P\big(V(\lambda)\big)$ consists of the orbits of the dominant weights $\leq \lambda$ (4.3), which are the weights marked in Figure 5.1.

A simple computation gives that $\lambda = \rho = \lambda_1 + \lambda_2$, $(\lambda_1|\lambda_1) = 2$, $(\lambda_2|\lambda_2) = 6$, $(\lambda_1|\lambda_2) = 3$ and

$$\begin{aligned} &(\lambda + \rho|\lambda + \rho) = 4(\lambda|\lambda) = 56, \\ &(2\lambda_1 + \rho|2\lambda_1 + \rho) = (3\lambda_1 + \lambda_2|3\lambda_1 + \lambda_2) = 42, \\ &(\lambda_2 + \rho|\lambda_2 + \rho) = (\lambda_1 + 2\lambda_2|\lambda_1 + 2\lambda_2) = 38, \\ &(\lambda_1 + \rho|\lambda_1 + \rho) = (2\lambda_1 + \lambda_2|2\lambda_1 + \lambda_2) = 26, \\ &(\rho|\rho) = 14. \end{aligned}$$

We start with $m_\lambda = 1$, then Freudenthal's multiplicity formula gives:

$$\begin{aligned}(56-42)m_{2\lambda_1} &= 2\sum_{\gamma\in\Phi^+}(2\lambda_1+\gamma|\gamma)m_{2\lambda_1+\gamma}\\ &= 2\Big((2\lambda_1+\alpha|\alpha)m_{2\lambda_1+\alpha}+(2\lambda_1+\beta|\beta)m_{2\lambda_1+\beta}\\ &\qquad+(2\lambda_1+\alpha+\beta|\alpha+\beta)m_{2\lambda_1+\alpha+\beta}\Big)\\ &= 2\Big((5\alpha+2\beta|\alpha)+(4\alpha+3\beta|\beta)+(5\alpha+3\beta|\alpha+\beta)\Big)\\ &= 2\Big((10-6)+(-12+18)+(10-24+18)\Big)=28,\end{aligned}$$

and we conclude that $m_{2\lambda_1}=\frac{28}{14}=2$. Thus the multiplicity of the weight spaces, whose weight is conjugated to $2\lambda_1$ is 2. (These are the weights marked with a $\triangledown$ in Figure 5.1.)

In the same vein,

$$\begin{aligned}(56-38)m_{\lambda_2} &= 2\sum_{\gamma\in\Phi^+}\sum_{j=1}^{\infty}(\lambda_2+j\gamma|\gamma)m_{\lambda_2+j\gamma}\\ &= 2\Big((\lambda_2+\alpha|\alpha)m_{\lambda_2+\alpha}+(\lambda_2+2\alpha|\alpha)m_{\lambda_2+2\alpha}\\ &\quad+(\lambda_2+\alpha+\beta|\alpha+\beta)m_{\lambda_2+\alpha+\beta}+(\lambda_2+2\alpha+\beta|2\alpha+\beta)m_{\lambda_2+2\alpha+\beta}\Big)\\ &= 2\Big((4\alpha+2\beta|\alpha)2+(5\alpha+2\beta|\alpha)+(4\alpha+3\beta|\alpha+\beta)+(5\alpha+3\beta|2\alpha+\beta)\Big)\\ &= 2\Big((8-6)2+(10-6)+(8-12-9+18)+(20-15-18+18)\Big)=36,\end{aligned}$$

and we conclude that $m_{\lambda_2}=2$. Now,

$$\begin{aligned}(56-26)m_{\lambda_1} &= 2\sum_{\gamma\in\Phi^+}\sum_{j=1}^{\infty}(\lambda_1+j\gamma|\gamma)m_{\lambda_1+j\gamma}\\ &= 2\Big((\lambda_1+\alpha|\alpha)m_{\lambda_1+\alpha}+(\lambda_1+2\alpha|\alpha)m_{\lambda_1+2\alpha}+(\lambda_1+\beta|\beta)m_{\lambda_1+\beta}\\ &\quad+(\lambda_1+\alpha+\beta|\alpha+\beta)m_{\lambda_1+\alpha+\beta}+(\lambda_1+2(\alpha+\beta)|\alpha+\beta)m_{\lambda_1+2(\alpha+\beta)}\\ &\quad+(\lambda_1+2\alpha+\beta|2\alpha+\beta)m_{\lambda_1+2\alpha+\beta}+(\lambda_1+3\alpha+\beta|3\alpha+\beta)m_{\lambda_1+3\alpha+\beta}\\ &\quad+(\lambda_1+3\alpha+2\beta|3\alpha+2\beta)m_{\lambda_1+3\alpha+2\beta}\Big)\\ &= 2\Big((3\alpha+\beta|\alpha)2+(4\alpha+\beta|\alpha)1+(2\alpha+2\beta|\beta)2+(3\alpha+2\beta|\alpha+\beta)2\\ &\quad+(4\alpha+3\beta|\alpha+\beta)1+(4\alpha+2\beta|2\alpha+\beta)2+(5\alpha+2\beta|3\alpha+\beta)1\\ &\quad+(5\alpha+3\beta|3\alpha+2\beta)1\Big)=120,\end{aligned}$$

so $m_{\lambda_1}=4$. Finally,

$$\begin{aligned}(56-14)m_0 &= 2\sum_{\gamma\in\Phi^+}\sum_{j=1}^{\infty}(j\gamma|\gamma)m_{j\gamma}=2\sum_{\gamma\in\Phi^+}\sum_{j=1}^{\infty}(\gamma|\gamma)jm_{j\gamma}\\ &= 2\Big(2(4+2\cdot 2)+6\cdot 2+2(4+2\cdot 2)+2(4+2\cdot 2)+6\cdot 2+6\cdot 2\Big)=168,\end{aligned}$$

so $m_0=4$.

Taking into account the sizes of the orbits, we get also that

$$\dim_k V(\lambda) = 12 \cdot 1 + 6 \cdot 2 + 6 \cdot 2 + 6 \cdot 4 + 4 = 64.$$

In the computations above, we made use of the symmetry given by the Weyl group. This can be improved.

**5.4 Lemma.** *Let $\lambda \in \Lambda_W^+$, $\mu \in P\bigl(V(\lambda)\bigr)$ and $\alpha \in \Phi$. Then*

$$\sum_{j\in\mathbb{Z}} (\mu + j\alpha|\alpha) m_{\mu+j\alpha} = 0.$$

*Proof.* $\oplus_{j\in\mathbb{Z}} V(\lambda)_{\mu+j\alpha}$ is a module for $S_\alpha$ (notation as in the proof of Proposition 4.3). But $S_\alpha = [S_\alpha, S_\alpha]$, since it is simple, hence the trace of the action of any of its elements is 0. In particular,

$$0 = \operatorname{trace}_{\oplus_{j\in\mathbb{Z}} V(\lambda)_{\mu+j\alpha}} \phi(t_\alpha) = \sum_{j\in\mathbb{Z}} (\mu + j\alpha|\alpha) m_{\mu+j\alpha}. \qquad \square$$

Now, Freudenthal's formula can be changed slightly using the previous Lemma and the fact that $2\rho = \sum_{\alpha\in\Phi^+} \alpha$.

**5.5 Corollary.** *For any $\lambda \in \Lambda_W^+$ and $\mu \in \Lambda_W$:*

$$(\lambda|\lambda + 2\rho) m_\mu = \sum_{\alpha\in\Phi} \sum_{j=1}^{\infty} (\mu + j\alpha|\alpha) m_{\mu+j\alpha} + (\mu|\mu) m_\mu.$$

We may change $j = 1$ for $j = 0$ in the sum above, since $(\mu|\alpha)m_\mu + (\mu| - \alpha)m_\mu = 0$ for any $\alpha \in \Phi$.

Now, if $\lambda \in \Lambda_W^+$, $\mu \in P\bigl(V(\lambda)\bigr) \cap \Lambda_W^+$ and $\sigma \in \mathcal{W}_\mu$ (the stabilizer of $\mu$, which is generated by the $\sigma_i$'s with $(\mu|\alpha_i) = 0$ by 2.1), then for any $\alpha \in \Phi$ and $j \in \mathbb{Z}$, $m_{\mu+j\alpha} = m_{\mu+j\sigma(\alpha)}$.

Let $I$ be any subset of $\{1, \dots, n\}$ and consider

$$\begin{aligned} &\Phi_I = \Phi \cap \bigl(\oplus_{i\in I} \mathbb{Z}\alpha_i\bigr) \quad \text{(a root system in } \oplus_{i\in I}\mathbb{R}\alpha_i\,!)\\ &\mathcal{W}_I, \text{ the subgroup of } \mathcal{W} \text{ generated by } \sigma_i,\ i \in I,\\ &\mathcal{W}_I^-, \text{ the group generated by } \mathcal{W}_I \text{ and } -id. \end{aligned}$$

For any $\alpha \in \Phi$, let $\mathcal{O}_{I,\alpha} = \mathcal{W}_I^- \alpha$. Then, if $\alpha \in \Phi_I$, $-\alpha = \sigma_\alpha(\alpha) \in \mathcal{W}_I\alpha$, so $\mathcal{W}_I^-\alpha = \mathcal{W}_I\alpha$. However, if $\alpha \notin \Phi_I$, then $\alpha = \sum_{i=1}^n r_i\alpha_i$ and there is an index $j \notin I$ with $r_j \neq 0$. For any $\sigma \in \mathcal{W}_I$, $\sigma(\alpha) = r_j\alpha_j + \sum_{i\neq j} r_i'\alpha_i$ (the coefficient of $\alpha_j$ does not change). In particular, if $\alpha \in \Phi^\pm$, then $\mathcal{W}_I\alpha \subseteq \Phi^\pm$. Therefore, $\mathcal{W}_I^-\alpha$ is the disjoint union of $\mathcal{W}_I\alpha$ and $-\mathcal{W}_I\alpha$.

**5.6 Proposition.** ***(Moody-Patera)*** *Let $\lambda \in \Lambda_W^+$ and $\mu \in P\bigl(V(\lambda)\bigr) \cap \Lambda_W^+$. Consider $I = \bigl\{i \in \{1, \dots, n\} : (\mu|\alpha_i) = 0\bigr\}$ and the orbits $\mathcal{O}_1, \dots \mathcal{O}_r$ of the action of $\mathcal{W}_I^-$ on $\Phi$. Take representatives $\gamma_i \in \mathcal{O}_i \cap \Phi^+$ for any $i = 1, \dots, r$. Then,*

$$\Bigl((\lambda + \rho|\lambda + \rho) - (\mu + \rho|\mu + \rho)\Bigr) m_\mu = \sum_{i=1}^{r} |\mathcal{O}_i| \sum_{j=1}^{\infty} (\mu + j\gamma_i|\gamma_i) m_{\mu+j\gamma_i}.$$

*Proof.* Arrange the orbits so that $\Phi_I = \mathcal{O}_1 \cup \cdots \cup \mathcal{O}_s$ and $\Phi \setminus \Phi_I = \mathcal{O}_{s+1} \cup \cdots \cup \mathcal{O}_r$. Hence $\mathcal{O}_i = \mathcal{W}_I\gamma_i$ for $i = 1, \dots, s$, while $\mathcal{O}_i = \mathcal{W}_i\gamma_i \cup -\mathcal{W}_I\gamma_i$ for $i = s+1, \dots, r$. Then, using the previous Lemma,

$$\begin{aligned}
&\sum_{\alpha\in\Phi}\sum_{j=1}^{\infty}(\mu + j\alpha|\alpha)m_{\mu+j\alpha} \\
&\qquad = \sum_{i=1}^{s}|\mathcal{W}_I\gamma_i|\sum_{j=1}^{\infty}(\mu + j\gamma_i|\gamma_i)m_{\mu+j\gamma_i} \\
&\qquad\quad + \sum_{i=s+1}^{r}|\mathcal{W}_I\gamma_i|\sum_{j=1}^{\infty}\Bigl((\mu + j\gamma_i|\gamma_i)m_{\mu+j\gamma_i} + (\mu - j\gamma_i| - \gamma_i)m_{\mu-j\gamma_i}\Bigr) \\
&\qquad = \sum_{i=1}^{s}|\mathcal{W}_I\gamma_i|\sum_{j=1}^{\infty}(\mu + j\gamma_i|\gamma_i)m_{\mu+j\gamma_i} \\
&\qquad\quad + \sum_{i=s+1}^{r}|\mathcal{W}_I\gamma_i|\Bigl(2\sum_{j=1}^{\infty}(\mu + j\gamma_i|\gamma_i)m_{\mu+j\gamma_i} + (\mu|\gamma_i)m_\mu\Bigr).
\end{aligned}$$

But, for any $i = s+1, \dots, r$, $2|\mathcal{W}_I\gamma_i| = |\mathcal{O}_i|$ and

$$\sum_{i=s+1}^{r}|\mathcal{W}_I\gamma_i|(\mu|\gamma_i) = \sum_{\alpha\in\Phi^+\setminus\Phi_I^+}(\mu|\alpha) = \sum_{\alpha\in\Phi^+}(\mu|\alpha) = 2(\mu|\rho).$$

Now, substitute this in the formula in Corollary 5.5 to get the result. □

**5.7 Example.** In the previous Example, for $\mu = 0$, $\mathcal{W}_I = \mathcal{W}$ and there are two orbits: the orbit of $\lambda_1$ (the short roots) and the orbit of $\lambda_2$ (the long roots), both of size 6. Hence,

$$\begin{aligned}
\Bigl((\lambda + \rho|\lambda + \rho) - (\rho|\rho)\Bigr)m_0 &= (56 - 14)m_0 \\
&= 6\Bigl((\lambda_1|\lambda_1)m_{\lambda_1} + (2\lambda_1|\lambda_1)m_{2\lambda_1} + (\lambda_2|\lambda_2)m_{\lambda_2}\Bigr) \\
&= 6\bigl(2\cdot 4 + 4\cdot 2 + 6\cdot 2\bigr) = 168,
\end{aligned}$$

so again we get $m_0 = 4$.

## § 6. Characters. Weyl's formulae

Consider the group algebra $\mathbb{R}\Lambda_W$. To avoid confusion between the binary operation (the addition) in $\Lambda_W$ and the addition in $\mathbb{R}\Lambda_W$, multiplicative notation will be used for $\Lambda_W$. Thus any $\lambda \in \Lambda_W$, when considered as an element of $\mathbb{R}\Lambda_W$, will be denoted by the formal symbol $e^\lambda$, and the binary operation (the addition) in $\Lambda_W$ becomes the product $e^\lambda e^\mu = e^{\lambda+\mu}$ in $\mathbb{R}\Lambda_W$. Hence,

$$\mathbb{R}\Lambda_W = \Bigl\{\sum_{\mu\in\Lambda_W} r_\mu e^\mu : r_\mu \in \mathbb{R},\ r_\mu = 0 \text{ for all but finitely many } \mu\text{'s}\Bigr\}.$$

Since $\Lambda_W$ is freely generated, as an abelian group, by the fundamental dominant weights: $\Lambda_W = \mathbb{Z}\lambda_1 \oplus \cdots \oplus \mathbb{Z}\lambda_n$, $\mathbb{R}\Lambda_W$ is isomorphic to the ring of Laurent polynomials in $n$ variables by means of:

$$\begin{aligned} \mathbb{R}[X_1^{\pm 1}, \ldots, X_n^{\pm 1}] &\to \mathbb{R}\Lambda_W \\ p(X_1, \ldots, X_n) &\mapsto p\bigl(e^{\lambda_1}, \ldots, e^{\lambda_n}\bigr). \end{aligned}$$

In particular, $\mathbb{R}\Lambda_W$ is an integral domain.

There appears a natural action of the Weyl group $\mathcal{W}$ on $\mathbb{R}\Lambda_W$ by automorphisms:

$$\begin{aligned} \mathcal{W} &\hookrightarrow \operatorname{Aut}\bigl(\mathbb{R}\Lambda_W\bigr) \\ \sigma &\mapsto \bigl(e^{\mu} \mapsto \sigma \cdot e^{\mu} = e^{\sigma(\mu)}\bigr). \end{aligned}$$

An element $p \in \mathbb{R}\Lambda_W$ is said to be *symmetric* if $\sigma \cdot p = p$ for any $\sigma \in \mathcal{W}$, and it is said *alternating* if $\sigma \cdot p = (-1)^{\sigma} p$ for any $\sigma \in \mathcal{W}$, where $(-1)^{\sigma} = \det \sigma \; (= \pm 1)$.

Consider the alternating map

$$\begin{aligned} \mathcal{A} : \mathbb{R}\Lambda_W &\to \mathbb{R}\Lambda_W \\ p &\mapsto \sum_{\sigma \in \mathcal{W}} (-1)^{\sigma} \sigma \cdot p. \end{aligned}$$

Then,

(i) For any $p \in \mathbb{R}\Lambda_W$, $\mathcal{A}(p)$ is alternating.

(ii) If $p \in \mathbb{R}\Lambda_W$ is alternating, $\mathcal{A}(p) = |\mathcal{W}| p$.

(iii) The alternating elements are precisely the linear combinations of the elements $\mathcal{A}(e^{\mu})$, for *strictly dominant* $\mu$ (that is, $\langle \mu | \alpha \rangle > 0$ for any $\alpha \in \Phi^+$). These form a basis of the subspace of alternating elements.

*Proof.* For any $\mu \in \Lambda_W$, there is a $\sigma \in \mathcal{W}$ such that $\sigma(\mu) \in \Lambda_W^+$ (Properties 2.1), and $\mathcal{A}(e^{\mu}) = (-1)^{\sigma} \mathcal{A}(e^{\sigma(\mu)})$. But if there is a simple root $\alpha_i$ such that $\langle \mu | \alpha_i \rangle = 0$, then $\mu = \sigma_i(\mu)$, so $\mathcal{A}(e^{\mu}) = (-1)^{\sigma_i} \mathcal{A}(e^{\sigma_i(\mu)}) = -\mathcal{A}(e^{\mu}) = 0$. Now, item (ii) finishes the proof. (The linear independence is clear.) $\square$

**6.1 Lemma.** *Let* $\rho = \frac{1}{2} \sum_{\alpha \in \Phi^+} \alpha$ *be the Weyl vector, and consider the element*

$$q = e^{-\rho} \prod_{\alpha \in \Phi^+} \bigl(e^{\alpha} - 1\bigr) = e^{\rho} \prod_{\alpha \in \Phi^+} \bigl(1 - e^{-\alpha}\bigr)$$

*in* $\mathbb{R}\Lambda_W$*. Then* $q = \mathcal{A}(e^{\rho})$.

*Proof.* For any simple root $\gamma \in \Delta$, $\sigma_\gamma\bigl(\Phi^+ \setminus \{\gamma\}\bigr) = \Phi^+ \setminus \{\gamma\}$ (Proposition 6.1). Hence $\sigma_\gamma(\rho) = \rho - \gamma$ and

$$\begin{aligned} \sigma_\gamma(q) &= e^{\rho - \gamma}(1 - e^{\gamma}) \prod_{\alpha \in \Phi^+ \setminus \{\gamma\}} \bigl(1 - e^{-\alpha}\bigr) \\ &= e^{\rho}\bigl(e^{-\gamma} - 1\bigr) \prod_{\alpha \in \Phi^+ \setminus \{\gamma\}} \bigl(1 - e^{-\alpha}\bigr) = -q. \end{aligned}$$

Thus, $q$ is alternating.

But, by its own definition, $q$ is a real linear combination of elements $e^\mu$, with $\mu = \rho - \sum_{\alpha\in\Phi^+}\epsilon_\alpha\alpha \leq \rho$ (where $\epsilon_\alpha$ is either 0 or 1). Hence

$$q = \frac{1}{|\mathcal{W}|}\mathcal{A}(q) = \sum_{\substack{\mu\in\Lambda_W^+ \\ \mu \text{ strictly dominant}}} c_\mu \mathcal{A}(e^\mu),$$

for some real scalars $c_\mu$ such that $c_\mu \neq 0$ only if $\mu$ is strictly dominant, $\mu \leq \rho$ and $\mu = \rho - \sum_{\alpha\in\Phi^+}\epsilon_\alpha\alpha$ as above. But then, for any such $\mu$ and $i = 1, \ldots, n$,

$$\langle \rho - \mu | \alpha_i\rangle = 1 - \langle \mu|\alpha_i\rangle \leq 0,$$

because $\langle\mu|\alpha_i\rangle \geq 1$, as $\mu$ is strictly dominant. Hence $(\rho-\mu|\alpha) \leq 0$ for any $\alpha\in\Phi^+$ and

$$0 \leq (\rho-\mu|\rho-\mu) = (\rho - \mu | \sum_{\alpha\in\Phi^+}\epsilon_\alpha\alpha) \leq 0,$$

so $\mu = \rho$. We conclude that $q = c\mathcal{A}(e^\rho)$ for some scalar $c$, but the definition of $q$ shows that

$$q = e^\rho + \text{a linear combination of terms } e^\nu, \text{ with } \nu < \rho,$$

so $c = 1$ and $q = \mathcal{A}(e^\rho)$. □

Consider the euclidean vector space $E = \mathbb{R}\otimes_{\mathbb{Q}}\mathbb{Q}\Phi$, and the $\mathbb{R}\Lambda_W$-module $\mathbb{R}\Lambda_W\otimes_{\mathbb{R}} E$. Extend the inner product $(\,.\,|\,.\,)$ on $E$ to a $\mathbb{R}\Lambda_W$-bilinear map

$$\big(\mathbb{R}\Lambda_W\otimes_{\mathbb{R}} E\big)\times\big(\mathbb{R}\Lambda_W\otimes_{\mathbb{R}} E\big) \to \mathbb{R}\Lambda_W,$$

and consider the $\mathbb{R}$-linear maps defined by:

$$\begin{aligned}
&\text{(GRADIENT)} \quad && \operatorname{grad}: \mathbb{R}\Lambda_W \to \mathbb{R}\Lambda_W\otimes_{\mathbb{R}} E \\
& && \qquad\quad e^\mu \mapsto e^\mu\otimes\mu, \\
&\text{(LAPLACIAN)} \quad && \qquad \Delta: \mathbb{R}\Lambda_W \to \mathbb{R}\Lambda_W \\
& && \qquad\quad e^\mu \mapsto (\mu|\mu)e^\mu,
\end{aligned}$$

which satisfy, for any $f, g\in\mathbb{R}\Lambda_W$:

$$\begin{cases} \operatorname{grad}(fg) = f\operatorname{grad}(g) + g(\operatorname{grad}(f), \\ \Delta(fg) = f\Delta(g) + g\Delta(f) + 2\big(\operatorname{grad}(f)|\operatorname{grad}(g)\big). \end{cases}$$

**6.2 Definition.** Let $V$ be a finite dimensional module for $L$, the element

$$\chi_V = \sum_{\mu\in\Lambda_W}(\dim_k V_\mu)e^\mu$$

of $\mathbb{R}\Lambda_W$ is called the *character* of $V$.

For simplicity, we will write $\chi_\lambda$ instead of $\chi_{V(\lambda)}$, for any $\lambda\in\Lambda_W^+$.

**6.3 Theorem.** ***(Weyl's character formula)*** *For any $\lambda \in \Lambda_W^+$,*

$$\chi_\lambda \mathcal{A}(e^\rho) = \mathcal{A}(e^{\lambda+\rho}).$$

In theory, this allows the computation of $\chi_\lambda$ as a closed quotient in $\mathbb{R}\Lambda_W$. In practice, Freudenthal's multiplicity formula is more efficient.

*Proof.* Note (Corollary 5.5) that Freudenthal's multiplicity formula is equivalent to

$$(\lambda|\lambda+2\rho)m_\mu = \sum_{\alpha\in\Phi}\sum_{j=0}^{\infty}(\mu+j\alpha|\alpha)m_{\mu+j\alpha} + (\mu|\mu)m_\mu \tag{6.3}$$

for any $\mu \in \Lambda_W$. Multiply by $e^\mu$ and sum on $\mu$ to get

$$(\lambda|\lambda+2\rho)\chi_\lambda = \sum_{\mu\in\Lambda_W}\sum_{\alpha\in\Phi}\sum_{j=0}^{\infty}(\mu+j\alpha|\alpha)m_{\mu+j\alpha}e^\mu + \Delta(\chi_\lambda).$$

Now,

$$\prod_{\alpha\in\Phi}(e^\alpha - 1) = \prod_{\alpha\in\Phi^+}(e^\alpha-1)(e^{-\alpha}-1) = \epsilon q^2$$

with $\epsilon = \pm 1$. Multiply by $\epsilon q^2$ to obtain

$$\begin{aligned}
&\epsilon(\lambda|\lambda+2\rho)\chi_\lambda q^2 - \epsilon\Delta(\chi_\lambda)q^2 \\
&\quad = \sum_{\alpha\in\Phi}\sum_{\mu\in\Lambda_W}\sum_{j=0}^{\infty}(\mu+j\alpha|\alpha)m_{\mu+j\alpha}(e^{\mu+\alpha}-e^\mu)\prod_{\substack{\beta\in\Phi\\ \beta\neq\alpha}}(e^\beta-1) \\
&\quad = \sum_{\alpha\in\Phi}e^\alpha\prod_{\substack{\beta\in\Phi\\ \beta\neq\alpha}}(e^\beta-1)\sum_{\mu\in\Lambda_W}\Big(\sum_{j=0}^{\infty}(\mu+j\alpha|\alpha)m_{\mu+j\alpha} - \sum_{j=0}^{\infty}(\mu+\alpha+j\alpha|\alpha)m_{\mu+\alpha+j\alpha}\Big)e^\mu \\
&\quad = \sum_{\alpha\in\Phi}e^\alpha\prod_{\substack{\beta\in\Phi\\ \beta\neq\alpha}}(e^\beta-1)\sum_{\mu\in\Lambda_W}m_\mu(\mu|\alpha)e^\mu \\
&\quad = \Big(\sum_{\alpha\in\Phi}(e^\alpha\prod_{\substack{\beta\in\Phi\\ \beta\neq\alpha}}(e^\beta-1)\otimes\alpha\Big) \mid \sum_{\mu\in\Lambda_W}m_\mu e^\mu\otimes\mu\Big) \\
&\quad = \big(\mathrm{grad}(\epsilon q^2)|\,\mathrm{grad}(\chi_\lambda)\big) = 2\epsilon q\big(\mathrm{grad}(q)|\,\mathrm{grad}(\chi_\lambda)\big) \\
&\quad = \epsilon q\big(\Delta(\chi_\lambda q) - \chi_\lambda\Delta(q) - q\Delta(\chi_\lambda)\big).
\end{aligned}$$

That is,

$$(\lambda|\lambda+2\rho)\chi_\lambda q = \Delta(\chi_\lambda q) - \chi_\lambda\Delta(q).$$

But $q = \sum_{\sigma\in\mathcal{W}}(-1)^\sigma e^{\sigma(\rho)}$ by the previous Lemma, so

$$\Delta(q) = \sum_{\sigma\in\mathcal{W}}\big(\sigma(\rho)|\sigma(\rho)\big)(-1)^\sigma e^{\sigma(\rho)} = (\rho|\rho)q,$$

so

$$(\lambda+\rho|\lambda+\rho)\chi_\lambda q = \Delta(\chi_\lambda q). \tag{6.4}$$

Now, $\chi_\lambda q$ is a linear combination of some $e^{\mu+\sigma(\rho)}$'s, with $\mu \in P\big(V(\lambda)\big)$ and $\sigma \in \mathcal{W}$, and

$$\begin{aligned}\Delta\big(e^{\mu+\sigma(\rho)}\big) &= \big(\mu+\sigma(\rho)|\mu+\sigma(\rho)\big)e^{\mu+\sigma(\rho)}\\ &= \big(\sigma^{-1}(\mu)+\rho|\sigma^{-1}(\mu)+\rho\big)e^{\mu+\sigma(\rho)}.\end{aligned}$$

Therefore, $e^{\mu+\sigma(\rho)}$ is an eigenvector of $\Delta$ with eigenvalue $\big(\sigma^{-1}(\mu)+\rho|\sigma^{-1}(\mu)+\rho\big)$, which equals $(\lambda+\rho|\lambda+\rho)$ because of (6.4). This implies (Properties 2.1) that $\sigma^{-1}(\mu)=\lambda$, or $\mu=\sigma(\lambda)$ and, hence, $\chi_\lambda q$ is a linear combination of $\{e^{\sigma(\lambda+\rho)} : \sigma \in \mathcal{W}\}$.

Since $\chi_\lambda$ is symmetric, and $q$ is alternating, $\chi_\lambda q$ is alternating. Also, $\sigma(\lambda+\rho)$ is strictly dominant if and only if $\sigma = id$. Hence $\chi_\lambda q$ is a scalar multiple of $\mathcal{A}(e^{\lambda+\rho})$, and its coefficient of $e^{\lambda+\rho}$ is 1. Hence, $\chi_\lambda q = \mathcal{A}(e^{\lambda+\rho})$, as required. □

If $p \in \mathbb{R}\Lambda_W$ is symmetric, then $p\mathcal{A}(e^\rho)$ is alternating. Then Weyl's character formula shows that $\{\chi_\lambda : \lambda \in \Lambda_W^+\}$ is a basis of the subspace of symmetric elements.

Weyl's character formula was derived by Weyl in 1926 in a very different guise.

**6.4 Corollary.** ***(Weyl's dimension formula)*** *For any $\lambda \in \Lambda_W^+$,*

$$\dim_k V(\lambda) = \prod_{\alpha\in\Phi^+} \frac{(\alpha|\lambda+\rho)}{(\alpha|\rho)} = \prod_{\alpha\in\Phi^+} \frac{\langle\lambda+\rho|\alpha\rangle}{\langle\rho|\alpha\rangle}.$$

*Proof.* Let $\mathbb{R}[[t]]$ be the ring of formal power series on the variable $t$, and for any $\nu \in \Lambda_W$ consider the homomorphism of real algebras given by:

$$\begin{aligned}\zeta_\nu : \mathbb{R}\Lambda_W &\longrightarrow \mathbb{R}[[t]]\\ e^\mu &\mapsto \exp\big((\mu|\nu)t\big) = \sum_{s=0}^{\infty} \frac{1}{s!}\big((\mu|\nu)t\big)^s.\end{aligned}$$

For any $\mu,\nu \in \Lambda_W$,

$$\begin{aligned}\zeta_\nu\big(\mathcal{A}(e^\mu)\big) &= \sum_{\sigma\in\mathcal{W}} (-1)^\sigma \exp\big((\sigma(\mu)|\nu)t\big)\\ &= \sum_{\sigma\in\mathcal{W}} (-1)^\sigma \exp\big((\mu|\sigma^{-1}(\nu))t\big)\\ &= \zeta_\mu\big(\mathcal{A}(e^\nu)\big).\end{aligned}$$

The homomorphism $\zeta_\rho$ will be applied now to Weyl's character formula. First,

$$\begin{aligned}\zeta_\rho\big(\mathcal{A}(e^\mu)\big) &= \zeta_\mu\big(\mathcal{A}(e^\rho)\big) = \zeta_\mu(q)\\ &= \zeta_\mu\big(e^{-\rho}\big) \prod_{\alpha\in\Phi^+} \big(\zeta_\mu(e^\alpha - 1)\big)\\ &= \exp\big((-\rho|\mu)t\big) \prod_{\alpha\in\Phi^+} \big(\exp\big((\alpha|\mu)t\big) - 1\big)\\ &= \prod_{\alpha\in\Phi^+} \Big(\exp\big(\tfrac{1}{2}(\alpha|\mu)t\big) - \exp\big(-\tfrac{1}{2}(\alpha|\mu)t\Big).\end{aligned}$$

Hence,

$$\zeta_\rho(\chi_\lambda q) = \zeta_\rho(\chi_\lambda) \prod_{\alpha\in\Phi^+} \Big(\exp\big(\tfrac{1}{2}(\alpha|\rho)t\big) - \exp\big(-\tfrac{1}{2}(\alpha|\rho)t\Big),$$

while
$$\zeta_\rho\bigl(\mathcal{A}(e^{\lambda+\rho})\bigr) = \prod_{\alpha\in\Phi^+}\Bigl(\exp\bigl(\tfrac{1}{2}(\alpha|\lambda+\rho)t\bigr) - \exp\bigl(-\tfrac{1}{2}(\alpha|\lambda+\rho)t\bigr).$$

With $N = |\Phi^+|$,
$$\prod_{\alpha\in\Phi^+}\Bigl(\exp\bigl(\tfrac{1}{2}(\alpha|\mu)t\bigr) - \exp\bigl(-\tfrac{1}{2}(\alpha|\mu)t\bigr) = \Bigl(\prod_{\alpha\in\Phi^+}(\alpha|\mu)\Bigr)t^N + \text{higher degree terms},$$

so if we look at the coefficients of $t^N$ in $\zeta_\rho\bigl(\chi_\lambda q\bigr) = \zeta_\rho\bigl(\mathcal{A}(e^{\lambda+\rho})\bigr)$ we obtain, since the coefficient of $t^0$ in $\zeta_\rho(\chi_\lambda)$ is $\dim_k V(\lambda)$, that
$$\dim_k V(\lambda)\prod_{\alpha\in\Phi^+}(\alpha|\rho) = \prod_{\alpha\in\Phi^+}(\alpha|\lambda+\rho). \qquad \square$$

**6.5 Example.** If $L$ is the simple Lie algebra of type $G_2$, $\Phi^+ = \{\alpha, \beta, \alpha+\beta, 2\alpha+\beta, 3\alpha+\beta, 3\alpha+2\beta\}$ (see Example 5.3) . Take $\lambda = n\lambda_1 + m\lambda_2$. Then Weyl's dimension formula gives
$$\begin{aligned}\dim_k V(\lambda) &= \tfrac{(n+1)3(m+1)\bigl(n+1+3(m+1)\bigr)\bigl(2(n+1)+3(m+1)\bigr)\bigl(3(n+1)+3(m+1)\bigr)\bigl(3(n+1)+6(m+1)\bigr)}{1\cdot 3\cdot 4\cdot 5\cdot 6\cdot 9}\\ &= \frac{1}{120}(n+1)(m+1)(n+m+2)(n+2m+3)(n+3m+4)(2n+3m+5).\end{aligned}$$

In particular, $\dim_k V(\lambda_1) = 7$ and $\dim_k V(\lambda_2) = 14$.

**6.6 Remark.** Weyl's dimension formula is extremely easy if $\lambda$ is a multiple of $\rho$. Actually, if $\lambda = m\rho$, then
$$\dim_k V(\lambda) = \prod_{\alpha\in\Phi^+}\frac{\bigl((m+1)\rho|\alpha)\bigr)}{(\rho|\alpha)} = (m+1)^{|\Phi^+|}.$$

For instance, with $\lambda = \lambda_1 + \lambda_2$ for $G_2$, $\dim_k V(\lambda) = 2^6 = 64$ (compare with Example 5.3).

Two more formulae to compute multiplicities will be given. First, for any $\mu \in \Lambda_W$ consider the integer:
$$\mathrm{p}(\mu) = \Bigl|\bigl\{(r_\alpha)_{\alpha\in\Phi^+} \in \mathbb{Z}_{\geq 0}^{|\Phi^+|} : \mu = \sum_{\alpha\in\Phi^+} r_\alpha\alpha\bigr\}\Bigr|.$$

Thus $\mathrm{p}(0) = 1 = \mathrm{p}(\alpha)$ for any $\alpha \in \Delta$. Also, if $\alpha, \beta \in \Delta$, with $\alpha \neq \beta$ and $(\alpha|\beta) \neq 0$, then $\mathrm{p}(\alpha+\beta) = 2$, as $\alpha+\beta$ can be written in two ways as a $\mathbb{Z}_{\geq 0}$-linear combination of positive roots: $1\cdot\alpha + 1\cdot\beta + 0\cdot(\alpha+\beta)$ and $0\cdot\alpha + 0\cdot\beta + 1\cdot(\alpha+\beta)$. Note, finally, that $\mathrm{p}(\mu) = 0$ if $\mu \notin \mathbb{Z}_{\geq 0}\Delta$.

**6.7 Theorem.** ***(Kostant's formula, 1959)*** *For any $\lambda \in \Lambda_W^+$ and $\mu \in \Lambda_W$,*
$$\dim_k V(\lambda)_\mu = \sum_{\sigma\in\mathcal{W}}(-1)^\sigma\,\mathrm{p}\bigl(\sigma(\lambda+\rho) - (\mu+\rho)\bigr).$$

*Proof.* Take the formal series

$$\sum_{\mu\in\Lambda_W} \mathrm{p}(\mu)e^{\mu} = \prod_{\alpha\in\Phi^+}\bigl(1+e^{\alpha}+e^{2\alpha}+\cdots\bigr) = \prod_{\alpha\in\Phi^+}\bigl(1-e^{\alpha}\bigr)^{-1},$$

in the natural completion of $\mathbb{R}\Lambda_W$ (which is naturally isomorphic to the ring of formal Laurent series $\mathbb{R}[[X_1^{\pm 1},\dots,X_n^{\pm 1}]]$). Thus,

$$\Bigl(\sum_{\mu\in\Lambda_W}\mathrm{p}(\mu)e^{\mu}\Bigr)\Bigl(\prod_{\alpha\in\Phi^+}(1-e^{\alpha})\Bigr) = 1.$$

Let $\theta : \mathbb{R}\Lambda_W \to \mathbb{R}\Lambda_W$ be the automorphism given by $\theta(e^{\mu}) = e^{-\mu}$ for any $\mu\in\Lambda_W$. If this is applied to Weyl's character formula (recall that $q = \mathcal{A}(e^{\rho}) = e^{\rho}\prod_{\alpha\in\Phi^+}\bigl(1-e^{-\alpha}\bigr) = e^{-\rho}\prod_{\alpha\in\Phi^+}\bigl(1-e^{\alpha}\bigr)$), we obtain

$$\Bigl(\sum_{\mu\in\Lambda_W} m_\mu e^{-\mu}\Bigr)e^{-\rho}\prod_{\alpha\in\Phi^+}(1-e^{\alpha}) = \sum_{\sigma\in\mathcal{W}}(-1)^{\sigma}e^{-\sigma(\lambda+\rho)}.$$

Multiply this by $e^{\rho}\bigl(\sum_{\nu\in\Lambda_W}\mathrm{p}(\nu)e^{\nu}\bigr)$ to get

$$\begin{aligned}\sum_{\mu\in\Lambda_W} m_\mu e^{-\mu} &= \Bigl(\sum_{\sigma\in\mathcal{W}}(-1)^{\sigma}e^{\rho-\sigma(\lambda+\rho)}\Bigr)\Bigl(\sum_{\nu\in\Lambda_W}\mathrm{p}(\nu)e^{\nu}\Bigr)\\ &= \sum_{\sigma\in\mathcal{W}}\sum_{\nu\in\Lambda_W}(-1)^{\sigma}\,\mathrm{p}(\nu)e^{\rho+\nu-\sigma(\lambda+\rho)},\end{aligned}$$

which implies that

$$\begin{aligned} m_\mu &= \sum_{\sigma\in\mathcal{W}}(-1)^{\sigma}\,\mathrm{p}(\nu_\sigma), \quad \text{with } \nu_\sigma \text{ such that } \rho+\nu_\sigma-\sigma(\lambda+\rho) = -\mu\\ &= \sum_{\sigma\in\mathcal{W}}(-1)^{\sigma}\,\mathrm{p}\bigl(\sigma(\lambda+\rho)-(\mu+\rho)\bigr). \end{aligned}$$

□

**6.8 Corollary.** *For any* $0\neq\mu\in\Lambda_W$*,*

$$\mathrm{p}(\mu) = -\sum_{1\neq\sigma\in\mathcal{W}}(-1)^{\sigma}\,\mathrm{p}\bigl(\mu-(\rho-\sigma(\rho))\bigr).$$

*Proof.* Take $\lambda = 0$ in Kostant's formula. Then $V(0) = k$ and

$$0 = \dim_k V(0)_{-\mu} = \sum_{\sigma\in\mathcal{W}}(-1)^{\sigma}\,\mathrm{p}\bigl(\sigma(\rho)-(-\mu+\rho)\bigr).$$

□

**6.9 Theorem.** ***(Racah's formula, 1962)*** *For any* $\lambda\in\Lambda_W^+$ *and* $\mu\in P\bigl(V(\lambda)\bigr)$*, with* $\mu\neq\lambda$*,*

$$m_\mu = -\sum_{1\neq\sigma\in\mathcal{W}}(-1)^{\sigma}m_{\mu+\rho-\sigma(\rho)}.$$

*(*$m_\nu = \dim_k V(\lambda)_\nu$ *for any* $\nu\in\Lambda_W$*).)*

Note that, since $\rho$ is strictly dominant, $\sigma(\rho) < \rho$ for any $1\neq\sigma\in\mathcal{W}$, hence $\mu+\rho-\sigma(\rho) > \mu$ and thus Racah's formula gives a recursive method starting with $m_\lambda = 1$.

*Proof.* If $\sigma \in \mathcal{W}$ satisfies $\sigma(\lambda+\rho) = \mu+\rho$, then $(\mu+\rho|\mu+\rho) = (\lambda+\rho|\lambda+\rho)$ and $\mu = \lambda$ (Properties 2.1). Now, by Kostant's formula and the previous Corollary,

$$\begin{aligned} m_\mu &= \sum_{\sigma\in\mathcal{W}} (-1)^\sigma \,\mathrm{p}\big(\sigma(\lambda+\rho) - (\mu+\rho)\big) \\ &= -\sum_{\sigma\in\mathcal{W}} \sum_{1\neq\tau\in\mathcal{W}} (-1)^\sigma (-1)^\tau \,\mathrm{p}\big(\sigma(\lambda+\rho) - (\mu+\rho) - (\rho-\tau(\rho))\big) \\ &= -\sum_{1\neq\tau\in\mathcal{W}} (-1)^\tau \sum_{\sigma\in\mathcal{W}} (-1)^\sigma \,\mathrm{p}\Big(\sigma(\lambda+\rho) - \big((\mu+\rho-\tau(\rho)) + \rho\big)\Big) \\ &= -\sum_{1\neq\tau\in\mathcal{W}} (-1)^\tau m_{\mu+\rho-\tau(\rho)}. \end{aligned}$$ □

**6.10 Example.** Consider again the simple Lie algebra of type $G_2$, and $\lambda = \rho = \lambda_1 + \lambda_2$. For the rotations $1 \neq \sigma \in \mathcal{W}$ one checks that (see Figure 5.1):

$$\rho - \sigma(\rho) = \alpha + 2\beta, 6\alpha + 2\beta, 10\alpha + 2\beta, 9\alpha + 4\beta, 4\alpha + \beta,$$

while for the symmetries in $\mathcal{W}$,

$$\rho - \sigma(\rho) = \alpha, \beta, 4\alpha + \beta, 9\alpha + 6\beta, 10\alpha + 5\beta, 6\alpha + 2\beta.$$

Starting with $m_\lambda = 1$ we obtain,

$$m_{2\lambda_1} = m_{2\lambda_1+\alpha} + m_{2\lambda_1+\beta} = m_\lambda + m_\lambda = 2,$$

since both $2\lambda_1 + \alpha$ and $2\lambda_1 + \beta$ are conjugated, under the action of $\mathcal{W}$, to $\lambda$. In the same spirit, one can compute:

$$\begin{aligned} m_{\lambda_2} &= m_{\lambda_2+\alpha} = m_{2\lambda_1} = 2, \\ m_{\lambda_1} &= m_{\lambda_1+\alpha} + m_{\lambda_1+\beta} = m_{\lambda_2} + m_{2\lambda_1} = 4, \\ m_0 &= m_\alpha + m_\beta - m_{\alpha+2\beta} - m_{4\alpha+\beta} = m_{\lambda_1} + m_{\lambda_2} - m_\lambda - m_\lambda = 4. \end{aligned}$$

## § 7. Tensor products decompositions

Given two dominant weights $\lambda', \lambda'' \in \Lambda_W^+$, Weyl's Theorem on complete reducibility shows that the tensor product $V(\lambda') \otimes_k V(\lambda'')$ is a direct sum of irreducible modules:

$$V(\lambda') \otimes_k V(\lambda'') \cong \oplus_{\lambda\in\Lambda_W^+} n_\lambda V(\lambda).$$

Moreover, for any $\mu \in \Lambda_W$,

$$\Big(V(\lambda') \otimes_k V(\lambda'')\Big)_\mu = \oplus_{\nu\in\Lambda_W} V(\lambda')_\nu \otimes_k V(\lambda'')_{\mu-\nu},$$

which shows that $\chi_{V(\lambda')\otimes_k V(\lambda'')} = \chi_{\lambda'}\chi_{\lambda''}$. Hence

$$\chi_{\lambda'}\chi_{\lambda''} = \sum_{\lambda\in\Lambda_W^+} n_\lambda \chi_\lambda.$$

The purpose of this section is to provide methods to compute the multiplicities $n_\lambda$.

**7.1 Theorem.** ***(Steinberg, 1961)*** *For any $\lambda', \lambda'' \in \Lambda_W^+$,*

$$n_\lambda = \sum_{\sigma,\tau\in\mathcal{W}} (-1)^{\sigma\tau}\, \mathrm{p}\big(\sigma(\lambda'+\rho)+\tau(\lambda''+\rho)-(\lambda+2\rho)\big).$$

*Proof.* From $\chi_{\lambda'}\chi_{\lambda''} = \sum_{\lambda\in\Lambda_W^+} n_\lambda \chi_\lambda$, we get

$$\chi_{\lambda'}\Big(\chi_{\lambda''}\mathcal{A}(e^\rho)\Big) = \sum_{\lambda\in\Lambda_W^+} n_\lambda\big(\chi_\lambda\mathcal{A}(e^\rho)\big),$$

which, by Weyl's character formula, becomes

$$\Big(\sum_{\mu\in\Lambda_W} m'_\mu e^\mu\Big)\Big(\sum_{\tau\in\mathcal{W}}(-1)^\tau e^{\tau(\lambda''+\rho)}\Big) = \sum_{\lambda\in\Lambda_W^+} n_\lambda\Big(\sum_{\tau\in\Lambda_W}(-1)^\tau e^{\tau(\lambda+\rho)}\Big). \tag{7.5}$$

The coefficient of $e^{\lambda+\rho}$ on the right hand side of (7.5) is $n_\lambda$, since in each orbit $\mathcal{W}(\lambda+\rho)$ there is a unique dominant weight, namely $\lambda+\rho$.

On the other hand, by Kostant's formula, the left hand side of (7.5) becomes:

$$\begin{aligned}\Big(\sum_{\mu\in\Lambda_W}\sum_{\sigma\in\mathcal{W}}(-1)^\sigma\, \mathrm{p}\big(\sigma(\lambda'+\rho)-\mu-\rho\big)e^\mu\Big)&\Big(\sum_{\tau\in\mathcal{W}}(-1)^\tau e^{\tau(\lambda''+\rho)}\Big)\\ &= \sum_{\mu\in\Lambda_W}\sum_{\sigma,\tau\in\mathcal{W}}(-1)^{\sigma\tau}\, \mathrm{p}\big(\sigma(\lambda'+\rho)-\mu-\rho\big)e^{\mu+\tau(\lambda''+\rho)}.\end{aligned}$$

Note that $\mu+\tau(\lambda''+\rho)=\lambda+\rho$ if and only if $-\mu-\rho=\tau(\lambda''+\rho)-(\lambda+2\rho)$, so the coefficient of $e^{\lambda+\rho}$ on the left hand side of (7.5) is

$$\sum_{\sigma,\tau\in\mathcal{W}} (-1)^{\sigma\tau}\, \mathrm{p}\big(\sigma(\lambda'+\rho)+\tau(\lambda''+\rho)-(\lambda+2\rho)\big),$$

as required. □

**7.2 Corollary.** ***(Racah, 1962)*** *For any $\lambda, \lambda', \lambda'' \in \Lambda_W^+$ the multiplicity of $V(\lambda)$ in $V(\lambda')\otimes_k V(\lambda'')$ is*

$$n_\lambda = \sum_{\sigma\in\mathcal{W}}(-1)^\sigma m'_{\lambda+\rho-\sigma(\lambda''+\rho)}.$$

*(For any weight $\mu$, $m'_\mu$ denotes the multiplicity of $\mu$ in $V(\lambda')$.)*

*Proof.*

$$\begin{aligned} n_\lambda &= \sum_{\sigma,\tau\in\mathcal{W}} (-1)^{\sigma\tau}\, \mathrm{p}\big(\sigma(\lambda'+\rho)+\tau(\lambda''+\rho)-(\lambda+2\rho)\big) \quad \text{(Steinberg)}\\ &= \sum_{\tau\in\mathcal{W}}(-1)^\tau\Big(\sum_{\sigma\in\mathcal{W}}(-1)^\sigma\, \mathrm{p}\big(\sigma(\lambda'+\rho)-(\lambda+\rho-\tau(\lambda''+\rho)+\rho)\big)\Big)\\ &= \sum_{\tau\in\mathcal{W}}(-1)^\tau m'_{\lambda+\rho-\tau(\lambda''+\rho)} \quad \text{(Kostant)}. \end{aligned}$$

□

To give a last formula to compute $\chi_{\lambda'}\chi_{\lambda''}$ some more notation is needed. First, by Weyl's character formula 6.3, for any $\lambda \in \Lambda_W^+$, $\chi_\lambda = \frac{\mathcal{A}(e^{\lambda+\rho})}{\mathcal{A}(e^\rho)}$. Let us extend this, by *defining* $\chi_\lambda$ for any $\lambda \in \Lambda_W$ by means of this formula. For any weight $\mu \in \Lambda_W$, recall that $\mathcal{W}_\mu$ denotes the stabilizer of $\mu$ in $\mathcal{W}$: $\mathcal{W}_\mu = \{\sigma \in \mathcal{W} : \sigma(\mu) = \mu\}$. If this stabilizer is trivial, then there is a unique $\sigma \in \mathcal{W}$ such that $\sigma(\mu) \in \Lambda_W^+$ (Properties 2.1). Consider then

$$\mathrm{s}(\mu) = \begin{cases} 0 & \text{if } \mathcal{W}_\mu \neq 1 \\ (-1)^\sigma & \text{if } \mathcal{W}_\mu = 1, \text{ and } \sigma(\mu) \in \Lambda_W^+. \end{cases}$$

Denote also by $\{\mu\}$ the unique dominant weight which is conjugate to $\mu$. Let $\sigma \in \mathcal{W}$ such that $\sigma(\mu) = \{\mu\}$. If $\{\mu\}$ is strictly dominant, then $\mathcal{A}(e^\mu) = (-1)^\sigma \mathcal{A}(e^{\{\mu\}}) = \mathrm{s}(\mu)\mathcal{A}(e^{\{\mu\}})$, otherwise there is an $i = 1, \ldots, n$ such that $\sigma_i\sigma(\mu) = \sigma(\mu)$ (Properties 2.1) so $\sigma^{-1}\sigma_i\sigma \in \mathcal{W}_\mu$ and $\mathrm{s}(\mu) = 0$; also $\mathcal{A}(e^{\{\mu\}}) = \mathcal{A}(\sigma_i \cdot e^{\{\mu\}}) = -\mathcal{A}(e^{\{\mu\}}) = 0$ and $\mathcal{A}(e^\mu) = 0$ too. Hence $\mathcal{A}(e^\mu) = \mathrm{s}(\mu)\mathcal{A}(e^{\{\mu\}})$ for any $\mu \in \Lambda_W$. Therefore, for any $\lambda \in \Lambda_W$, $\mathcal{A}(e^{\lambda+\rho}) = \mathrm{s}(\lambda+\rho)\mathcal{A}(e^{\{\lambda+\rho\}})$, and

$$\chi_\lambda = \mathrm{s}(\lambda+\rho)\chi_{\{\lambda+\rho\}-\rho}\,.$$

**7.3 Theorem.** ***(Klymik, 1968)*** *For any $\lambda', \lambda'' \in \Lambda_W^+$,*

$$\chi_{\lambda'}\chi_{\lambda''} = \sum_{\mu \in P\left(V(\lambda')\right)} m'_\mu \chi_{\mu+\lambda''}\,.$$

Note that this can be written as

$$\chi_{\lambda'}\chi_{\lambda''} = \sum_{\mu \in P\left(V(\lambda')\right)} m'_\mu\, \mathrm{s}(\mu+\lambda''+\rho)\chi_{\{\mu+\lambda''+\rho\}-\rho}. \tag{7.6}$$

By Properties 2.1, if $\nu \in \Lambda_W$ and $\mathrm{s}(\nu) \neq 0$, then $\{\nu\}$ is strictly dominant, and hence $\{\nu\} - \rho \in \Lambda_W^+$, so all the weights $\{\mu+\lambda''+\rho\} - \rho$ that appear with nonzero coefficient on the right hand side of the last formula are dominant.

*Proof.* As in the proof of Steinberg's Theorem, with $P' = P\big(V(\lambda')\big)$,

$$\begin{aligned}
\chi_{\lambda'}\big(\chi_{\lambda''}\mathcal{A}(e^\rho)\big) &= \chi_{\lambda'}\mathcal{A}(e^{\lambda''+\rho}) \\
&= \Big(\sum_{\mu\in P'} m'_\mu e^\mu\Big)\Big(\sum_{\sigma\in\mathcal{W}} (-1)^\sigma e^{\sigma(\lambda''+\rho)}\Big) \\
&= \sum_{\sigma\in\mathcal{W}} (-1)^\sigma \Big(\sum_{\mu\in P'} m'_\mu e^{\sigma(\mu)}\Big) e^{\sigma(\lambda''+\rho)} \quad (\text{as } m'_\mu = m'_{\sigma(\mu)}\ \forall \sigma \in \mathcal{W}) \\
&= \sum_{\mu\in P'} m'_\mu \sum_{\sigma\in\mathcal{W}} (-1)^\sigma e^{\sigma(\mu+\lambda''+\rho)} \\
&= \sum_{\mu\in P'} m'_\mu \mathcal{A}\big(e^{\mu+\lambda''+\rho}\big).
\end{aligned}$$

□

**7.4 Corollary.** *Let $\lambda, \lambda', \lambda'' \in \Lambda_W^+$. If $V(\lambda)$ is (isomorphic to) a submodule of $V(\lambda') \otimes_k V(\lambda'')$, then there exists $\mu \in P\big(V(\lambda')\big)$ such that $\lambda = \mu + \lambda''$.*

*Proof.* Because of (7.6), if $V(\lambda)$ is isomorphic to a submodule of $V(\lambda') \otimes_k V(\lambda'')$ there is a $\mu \in P\bigl(V(\lambda')\bigr)$ such that $\{\mu + \lambda'' + \rho\} = \lambda + \rho$. Take $\mu \in P\bigl(V(\lambda')\bigr)$ and $\sigma \in \mathcal{W}$ such that $\sigma(\mu+\lambda''+\rho) = \lambda+\rho$ and $\sigma$ has minimal length. It is enough to prove that $\mathrm{l}(\sigma) = 0$. If $\mathrm{l}(\sigma) = t \geq 1$, let $\sigma = \sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_t}$ be a reduced expression. Then $\sigma(\beta_t) \in \Phi^-$ by Properties 2.1 and

$$\begin{aligned} 0 &\geq \bigl(\lambda + \rho | \sigma(\beta_t)\bigr) = \bigl(\sigma^{-1}(\lambda + \rho) | \beta_t\bigr) \\ &= \bigl(\mu + \lambda'' + \rho | \beta_t\bigr) = \bigl(\mu | \beta_t\bigr) + \bigl(\lambda'' + \rho | \beta_t\bigr) \geq \bigl(\mu | \beta_t\bigr), \end{aligned}$$

since $\lambda + \rho$ and $\lambda'' + \rho$ are dominant. Hence $0 \geq \langle \mu + \lambda'' + \rho | \beta_t \rangle \geq \langle \mu | \beta_t \rangle$ and $\hat{\mu} = \mu - \langle \mu + \lambda'' + \rho | \beta_t \rangle \beta_t \in P\bigl(V(\lambda')\bigr)$. Therefore,

$$\begin{aligned} \lambda + \rho = \sigma(\mu + \lambda'' + \rho) &= (\sigma \circ \sigma_{\beta_t})\bigl(\sigma_{\beta_t}(\mu + \lambda'' + \rho)\bigr) \\ &= (\sigma \circ \sigma_{\beta_t})\bigl(\mu + \lambda'' + \rho - \langle \mu + \lambda'' + \rho | \beta_t \rangle \beta_t\bigr) \\ &= \sigma \circ \sigma_{\beta_t}(\hat{\mu} + \lambda'' + \rho), \end{aligned}$$

a contradiction with the minimality of $\mathrm{l}(\sigma)$, as $\sigma \circ \sigma_{\beta_t} = \sigma_{\beta_1} \circ \cdots \circ \sigma_{\beta_{t-1}}$. □

**7.5 Example.** As usual, let $L$ be the simple Lie algebra of type $G_2$. Let us decompose $V(\lambda_1) \otimes_k V(\lambda_2)$ using Klymik's formula.

Recall that $\lambda_1 = 2\alpha + \beta$, $\lambda_2 = 3\alpha + 2\beta$, so $\alpha = 2\lambda_1 - \lambda_2$ and $\beta = -3\lambda_1 + 2\lambda_2$. Scaling so that $(\alpha|\alpha) = 2$, one gets $(\lambda_1|\alpha) = 1$, $(\lambda_2|\beta) = 3$, and $(\lambda_1|\beta) = 0 = (\lambda_2|\alpha)$.

Also, $P\bigl(V(\lambda_1)\bigr) = \mathcal{W}\lambda_1 \cup \mathcal{W}0 = \{0, \pm\alpha, \pm(\alpha + \beta), \pm(2\alpha + \beta)\}$ (the short roots and 0). The multiplicity of any short root equals the multiplicity of $\lambda_1$, which is 1.

Freudenthal's formula gives

$$\bigl((\lambda_1 + \rho | \lambda_1 + \rho) - (\rho|\rho)\bigr) m_0 = 2 \sum_{\gamma \in \Phi^+} \sum_{j=1}^{\infty} (j\gamma|\gamma) m_{j\gamma} = 2 \sum_{\substack{\gamma \in \Phi^+ \\ \gamma \text{ short}}} (\gamma|\gamma) = 12,$$

since $m_{\lambda_1} = 1$, so $m_\gamma = 1$ for any short $\gamma$, as all of them are conjugate. But $(\lambda_1 + \rho | \lambda_1 + \rho) - (\rho|\rho) = (\lambda_1 | \lambda_1 + 2\rho) = (\lambda_1 | 3\lambda_1 + 2\lambda_2) = (3\lambda_1 + 2\lambda_2 | 2\alpha + \beta) = 12$. Thus, $m_0 = 1$.

Hence all the weights of $V(\lambda_1)$ have multiplicity 1, and Klymik's formula gives then

$$\chi_{\lambda_1} \chi_{\lambda_2} = \sum_{\mu \in P(V(\lambda_1))} \mathrm{s}(\mu + \lambda_2 + \rho) \chi_{\{\mu + \lambda_2 + \rho\} - \rho}.$$

Let us compute the contribution to this sum of each $\mu \in P\bigl(V(\lambda_1)\bigr)$:

- $0 + \lambda_2 + \rho$ is strictly dominant, so $\mathrm{s}(0 + \lambda_2 + \rho) = 1$, and we obtain the summand $1 \cdot \chi_{\lambda_2}$,
- $\alpha + \lambda_2 + \rho = 2\lambda_1 + \rho$ is strictly dominant, so $\mathrm{s}(\alpha + \lambda_2 + \rho) = 1$ and we get $1 \cdot \chi_{2\lambda_1}$,
- $-\alpha + \lambda_2 + \rho = -\lambda_1 + 3\lambda_2$ is not dominant, and $\sigma_\alpha(-\lambda_1 + 3\lambda_2) = -\lambda_1 + 3\lambda_2 + \alpha = \lambda_1 + 2\lambda_2 = \lambda_2 + \rho$ is strictly dominant, so $\mathrm{s}(-\alpha + \lambda_2 + \rho) = -1$ and get $(-1) \cdot \chi_{\lambda_2}$,
- $\alpha + \beta + \lambda_2 + \rho = 3\lambda_2$ is stabilized by $\sigma_\alpha$, so $\mathrm{s}(\alpha + \beta + \lambda_2 + \rho) = 0$,
- $-(\alpha + \beta) + \lambda_2 + \rho = \lambda_1 + \rho$ is strictly dominant, so we get $1 \cdot \chi_{\lambda_1}$,

- $2\alpha + \beta + \lambda_2 + \rho = \lambda_1 + \lambda_2 + \rho$ is strictly dominant, so we get $1 \cdot \chi_{\lambda_1+\lambda_2}$,
- $-(2\alpha + \beta) + \lambda_2 + \rho = \lambda_2$ is stabilized by $\sigma_\alpha$.

Therefore, Klymik's formula gives:

$$V(\lambda_1) \otimes_k V(\lambda_2) \cong V(\lambda_1 + \lambda_2) \oplus V(2\lambda_1) \oplus V(\lambda_1). \tag{7.7}$$

With some insight, we could have proceeded in a different way. First, the multiplicity of the highest possible weight $\lambda' + \lambda''$ in $V(\lambda') \otimes_k V(\lambda'')$ is always 1, so $V(\lambda' + \lambda'')$ always appears in $V(\lambda') \otimes_k V(\lambda'')$ with multiplicity 1.

In the example above, if $\mu \in P\big(V(\lambda_1)\big)$ and $\mu + \lambda_2 \in \Lambda_W^+$, then $\mu + \lambda_2 \in \{\lambda_1 + \lambda_2, 2\lambda_1, \lambda_1, \lambda_2\}$. Hence,

$$V(\lambda_1) \otimes_k V(\lambda_2) \cong V(\lambda_1 + \lambda_2) \oplus pV(2\lambda_1) \oplus qV(\lambda_1) \oplus rV(\lambda_2),$$

and $\dim_k V(\lambda_1) \otimes_k V(\lambda_2) = 7 \times 14 = 98$, $\dim_k V(\lambda_1 + \lambda_2) = \dim_k V(\rho) = 2^6 = 64$. Weyl's dimension formula gives

$$\dim_k V(2\lambda_1) = \prod_{\gamma \in \Phi^+} \frac{(2\lambda_1 + \rho|\gamma)}{(\rho|\gamma)} = \frac{3 \cdot 3 \cdot 6 \cdot 9 \cdot 12 \cdot 15}{1 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 9} = 27.$$

The only possibility of $98 = 64 + p \cdot 27 + q \cdot 7 + r \cdot 14$ is $p = q = 1$, $r = 0$, thus recovering (7.7).

**7.6 Exercise.** Let $\lambda', \lambda'' \in \Lambda_W^+$. Prove that $P\big(V(\lambda') \otimes V(\lambda'')\big)$ equals $P\big(V(\lambda' + \lambda'')\big)$.

# Appendix A

# Simple real Lie algebras

Let $L$ be a simple real Lie algebra. By Schur's Lemma, the centralizer algebra $\mathrm{End}_L(L)$ is a real division algebra, but for any $\alpha, \beta \in \mathrm{End}_L(L)$ and $x, y \in L$,

$$\alpha\beta\big([x,y]\big) = \alpha\big[x, \beta y] = [\alpha x, \beta y] = \beta\big([\alpha x, y]\big) = \beta\alpha\big([x,y]\big),$$

and, since $L = [L, L]$, it follows that $\mathrm{End}_L(L)$ is commutative. Hence $\mathrm{End}_L(L)$ is (isomorphic to) either $\mathbb{R}$ or $\mathbb{C}$.

In the latter case, $L$ is then just a complex simple Lie algebra, but considered as a real Lie algebra.

In the first case, $\mathrm{End}_L(L) = \mathbb{R}$, so $\mathrm{End}_{L^{\mathbb{C}}}(L^{\mathbb{C}}) = \mathbb{C}$, where $L^{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} L = L \oplus iL$. Besides, $L^{\mathbb{C}}$ is semisimple because its Killing form is the extension of the Killing form of $L$, and hence it is nondegenerate. Moreover, if $L^{\mathbb{C}}$ is the direct sum of two proper ideals $L^{\mathbb{C}} = L_1 \oplus L_2$, then $\mathbb{C} = \mathrm{End}_{L^{\mathbb{C}}}(L^{\mathbb{C}}) \supseteq \mathrm{End}_{L^{\mathbb{C}}}(L_1) \oplus \mathrm{End}_{L^{\mathbb{C}}}(L_2)$, which has dimension at least 2 over $\mathbb{C}$, a contradiction. Hence $L^{\mathbb{C}}$ is simple. In this case, $L$ is said to be *central simple* and a *real form* of $L^{\mathbb{C}}$. (More generally, a simple Lie algebra over a field $k$ is said to be central simple, if its scalar extension $\bar{k} \otimes_k L$ is a simple Lie algebra over $\bar{k}$, an algebraic closure of $k$.)

Consider the natural antilinear automorphism $\sigma$ of $L^{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} L = L \oplus iL$ given by $\sigma = - \otimes id$ ($\alpha \mapsto \bar{\alpha}$ is the standard conjugation in $\mathbb{C}$). That is,

$$\begin{aligned} \sigma : L^{\mathbb{C}} &\to L^{\mathbb{C}} \\ x + iy &\mapsto x - iy. \end{aligned}$$

Then $L$ is the fixed subalgebra by $\sigma$, which is called the *conjugation associated to $L$*.

Therefore, in order to get the real simple Lie algebras, it is enough to obtain the real forms of the complex simple Lie algebras.

## § 1. Real forms

**1.1 Definition.** Let $L$ be a real semisimple Lie algebra.

- $L$ is said to be *split* if it contains a Cartan subalgebra $H$ such that $\mathrm{ad}_h$ is diagonalizable (over $\mathbb{R}$) for any $h \in H$.
- $L$ is said to be *compact* if its Killing form is definite.

- $L$ is said to be a *real form* of a complex Lie algebra $S$ if $L^{\mathbb{C}}$ is isomorphic to $S$ (as complex Lie algebras).

**1.2 Proposition.** *The Killing form of any compact Lie algebra is negative definite.*

*Proof.* Let $\kappa$ be the Killing form of the compact Lie algebra $L$ with $\dim_{\mathbb{R}} L = n$. For any $0 \neq x \in L$, let $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ be the eigenvalues of $\mathrm{ad}_x$ (possibly repeated). If for some $j = 1, \dots, n$, $\lambda_j \in \mathbb{R} \setminus \{0\}$, then there exists a $0 \neq y \in L$ such that $[x, y] = \lambda_j y$. Then the subalgebra $T = \mathbb{R}x + \mathbb{R}y$ is solvable and $y \in [T, T]$. By Lie's Theorem (Chapter 2, 1.9) $\mathrm{ad}_y$ is nilpotent, so $\kappa(y, y) = 0$, a contradiction with $\kappa$ being definite. Thus, $\lambda_j \notin \mathbb{R} \setminus \{0\}$ for any $j = 1, \dots, n$. Now, if $\lambda_j = \alpha + i\beta$ with $\beta \neq 0$, then $\bar{\lambda}_j = \alpha - i\beta$ is an eigenvalue of $\mathrm{ad}_x$ too, and hence there are elements $y, z \in L$, not both 0, such that $[x, y] = \alpha y + \beta z$ and $[x, z] = -\beta y + \alpha z$. Then

$$[x, [y, z]] = [[x, y], z] + [y, [x, z]] = 2\alpha[y, z].$$

The previous argument shows that either $\alpha = 0$ or $[y, z] = 0$. In the latter case $T = \mathbb{R}x + \mathbb{R}y + \mathbb{R}z$ is a solvable Lie algebra with $0 \neq \alpha y + \beta z \in [T, T]$. But this gives again a contradiction.

Therefore, $\lambda_1, \dots, \lambda_n \in \mathbb{R}i$ and $\kappa(x, x) = \sum_{j=1}^{n} \lambda_j^2 \leq 0$. □

**1.3 Theorem.** *Any complex semisimple Lie algebra contains both a split and a compact real forms.*

*Proof.* Let $S$ be a complex semisimple Lie algebra and let $\{h_j, x_j, y_j : j = 1, \dots, n\}$ be a set of canonical generators of $S$ relative to a Cartan subalgebra and an election of a simple system $\Delta$ of roots, as in Chapter 2, § 7. For any $\alpha \in \Phi^+$ choose $I_\alpha = (j_1, \dots, j_m)$ ($\mathrm{ht}(\alpha) = m$) such that $0 \neq \mathrm{ad}_{x_{j_m}} \cdots \mathrm{ad}_{x_{j_2}}(x_{j_1}) \in S_\alpha$ and take $x_\alpha = \mathrm{ad}_{x_{j_m}} \cdots \mathrm{ad}_{x_{j_2}}(x_{j_1})$ and $y_\alpha = \mathrm{ad}_{y_{j_m}} \cdots \mathrm{ad}_{y_{j_2}}(y_{j_1})$. Then $\{h_1, \dots, h_n, x_\alpha, y_\alpha : \alpha \in \Phi^+\}$ is a basis of $S$ and its structure constants are rational numbers that depend on the Dynkin diagram. Therefore,

$$L = \sum_{j=1}^{n} \mathbb{R}h_j + \sum_{\alpha \in \Phi^+} \big(\mathbb{R}x_\alpha + \mathbb{R}y_\alpha\big)$$

is a split real form of $S = L \oplus iL$. Its associated conjugation $\sigma : S \to S$ is determined by $\sigma(x_j) = x_j$ and $\sigma(y_j) = y_j$ for any $j = 1, \dots, n$.

But there is a unique automorphism $\omega \in \mathrm{Aut}_{\mathbb{C}} S$ such that $\omega(x_j) = -y_j$ and $\omega(y_j) = -x_j$ for any $j = 1, \dots, n$, because $-\Delta$ is another simple system of roots. Note that $\omega(h_j) = -h_j$ for any $j$ and $\omega^2 = id$. Then

$$\sigma\omega\sigma(x_j) = \omega(x_j) \quad \text{and} \quad \sigma\omega\sigma(y_j) = \omega(y_j)$$

for any $j$, so $\sigma\omega\sigma = \omega$, or $\sigma\omega = \omega\sigma$. Consider the antilinear involutive automorphism $\tau = \sigma\omega = \omega\sigma$ of $S$. Let us check that $\tau$ is the conjugation associated to a compact real form of $S$. Denote by $\kappa$ the Killing form of $S$.

First, by induction on $\mathrm{ht}(\alpha)$, let us prove that $\kappa(x_\alpha, \omega(x_\alpha))$ is a negative rational number:

- If $\mathrm{ht}(\alpha) = 1$, then $\alpha = \alpha_j$ for some $j$, $x_\alpha = x_j$ and $\kappa(x_j, \omega(x_j)) = -\kappa(x_j, y_j) = -\frac{2}{(\alpha_j|\alpha_j)} < 0$, since $h_\alpha = \frac{2}{(\alpha|\alpha)} t_\alpha = [x_\alpha, y_\alpha] = \kappa(x_\alpha, y_\alpha) t_\alpha$ (see Chapter 2, § 5.) for any $\alpha \in \Phi$, and the bilinear form $( . \mid . )$ is positive definite on $\mathbb{R}\Phi$.

- If $\mathrm{ht}(\alpha) = m + 1$, then $x_\alpha = q[x_j, x_\beta]$ for some $j = 1, \dots, n$ and $q \in \mathbb{Q}$, with $\mathrm{ht}(\beta) = m$, then

$$\begin{aligned}
\kappa\bigl(x_\alpha, \omega(x_\alpha)\bigr) &= q^2 \kappa\bigl([x_j, x_\beta], [\omega(x_j), \omega(x_\beta)]\bigr) \\
&= q^2 \kappa\bigl(x_\beta, [x_j, [y_j, \omega(x_\beta)]]\bigr) \\
&\in \mathbb{Q}_{>0} \kappa\bigl(x_\beta, \omega(x_\beta)\bigr) \quad \text{(by Chapter 2, Lemma 7.1)} \\
&\in \mathbb{Q}_{<0} \quad \text{(by the induction hypothesis).}
\end{aligned}$$

Now take $K$ the fixed subalgebra $S^\tau$ of $\tau$. Hence,

$$K = \sum_{j=1}^{n} \mathbb{R}(ih_j) + \sum_{\alpha \in \Phi^+} \Bigl( \mathbb{R}\bigl(x_\alpha + \omega(x_\alpha)\bigr) + \mathbb{R}i\bigl(x_\alpha - \omega(x_\alpha)\bigr) \Bigr),$$

which is a real form of $S = K \oplus iK$. Note that

- $\kappa(ih_r, ih_s) = -\kappa(h_r, h_s)$, and the restriction of $\kappa$ to $\sum_{j=1}^{n} \mathbb{R}h_j$ is positive definite,
- $\kappa\bigl(x_\alpha + \omega(x_\alpha), x_\alpha + \omega(x_\alpha)\bigr) = 2\kappa\bigl(x_\alpha, \omega(x_\alpha)\bigr) < 0$, by the previous argument,
- $\kappa\bigl(i(x_\alpha - \omega(x_\alpha)), i(x_\alpha - \omega(x_\alpha))\bigr) = 2\kappa\bigl(x_\alpha, \omega(x_\alpha)\bigr) < 0$, and
- $\kappa\bigl(x_\alpha + \omega(x_\alpha), i(x_\alpha - \omega(x_\alpha))\bigr) = i\kappa\bigl(x_\alpha + \omega(x_\alpha), x_\alpha - \omega(x_\alpha)\bigr) = 0.$

Hence the Killing form of $K$, which is obtained by restriction of $\kappa$, is negative definite, and hence $K$ is compact. □

**1.4 Remark.** The signature of the Killing form of the split form $L$ above is $\operatorname{rank} L$, while for the compact form $K$ is $-\dim K$.

**1.5 Definition.** Let $S$ be a complex semisimple Lie algebra and let $\sigma_1, \sigma_2$ be the conjugations associated to two real forms. Then:

- $\sigma_1$ and $\sigma_2$ are said to be *equivalent* if the corresponding real forms $S^{\sigma_1}$ and $S^{\sigma_2}$ are isomorphic.

- $\sigma_1$ and $\sigma_2$ are said to be *compatible* if they commute: $\sigma_1\sigma_2 = \sigma_2\sigma_1$.

Given a complex semisimple Lie algebra and a conjugation $\sigma$, this is said to be *split* or *compact* if so is its associated real form $S^\sigma$.

Note that the split $\sigma$ and compact $\tau$ conjugations considered in the proof of Theorem 1.3 are compatible.

**1.6 Proposition.** *Let $S$ be a complex semisimple Lie algebra and let $\sigma_1, \sigma_2$ be the conjugations associated to two real forms. Then:*

(i) *$\sigma_1$ and $\sigma_2$ are equivalent if and only if there is an automorphism $\varphi \in \mathrm{Aut}_{\mathbb{C}} S$ such that $\sigma_2 = \varphi\sigma_1\varphi^{-1}$.*

(ii) *$\sigma_1$ and $\sigma_2$ are compatible if and only if $\theta = \sigma_1\sigma_2$ (which is an automorphism of $S$) is involutive ($\theta^2 = id$). In this case $\theta$ leaves invariant both real forms ($\theta|_{S^{\sigma_i}} \in \mathrm{Aut}_{\mathbb{R}}(S^{\sigma_i})$, $i = 1, 2$).*

(iii) *If $\sigma_1$ and $\sigma_2$ are compatible and compact, then $\sigma_1 = \sigma_2$.*

*Proof.* For (i), if $\psi : S^{\sigma_1} \to S^{\sigma_2}$ is an isomorphism, then $\psi$ induces an automorphism $\varphi : S = S^{\sigma_1} \oplus iS^{\sigma_1} \to S = S^{\sigma_2} \oplus iS^{\sigma_2}$ ($\varphi(x + iy) = \psi(x) + i\psi(y)$ for any $x, y \in S^{\sigma_1}$). Moreover, it is clear that $\varphi\sigma_1 = \sigma_2\varphi$ as this holds trivially for the elements in $S^{\sigma_1}$. Conversely, if $\sigma_2 = \varphi\sigma_1\varphi^{-1}$ for some $\varphi \in \mathrm{Aut}_{\mathbb{C}} S$, then $\varphi\big(S^{\sigma_1}\big) \subseteq S^{\sigma_2}$ and the restriction $\varphi|_{S^{\sigma_1}}$ gives an isomorphism $S^{\sigma_1} \to S^{\sigma_2}$.

For (ii), it is clear that if $\sigma_1$ and $\sigma_2$ are compatible, then $\theta = \sigma_1\sigma_2$ is $\mathbb{C}$-linear (as a composition of two antilinear maps) and involutive ($\theta^2 = \sigma_1\sigma_2\sigma_1\sigma_2 = \sigma_1^2\sigma_2^2 = id$). Conversely, if $\theta^2 = id$, then $\sigma_1\sigma_2\sigma_1\sigma_2 = id = \sigma_1^2\sigma_2^2$, so $\sigma_1\sigma_2 = \sigma_2\sigma_1$ (as $\sigma_1$ and $\sigma_2$ are invertible).

Finally, assume that $\sigma_1$ and $\sigma_2$ are compatible and compact, and let $\theta = \sigma_1\sigma_2$, which is an involutive automorphism which commutes with both $\sigma_1$ and $\sigma_2$. Then $S^{\sigma_1} = S^{\sigma_1}_+ \oplus S^{\sigma_1}_-$, where $S^{\sigma_1}_\pm = \{x \in S^{\sigma_1} : \theta(x) = \pm x\}$. Let $\kappa$ be the Killing form of $S$, which restricts to the Killing forms of $S^{\sigma_i}$ ($i = 1, 2$). For any $x \in S^{\sigma_1}_-$, $0 \geq \kappa(x, x) = -\kappa(x, \theta(x)) = -\kappa(x, \sigma_2(x))$, as $\theta(x) = \sigma_1\sigma_2(x) = \sigma_2\sigma_1(x) = \sigma_2(x)$. But the map

$$\begin{aligned} h_{\sigma_2} : S \times S &\longrightarrow \mathbb{C} \\ (u, v) &\mapsto -\kappa(u, \sigma_2(v)) \end{aligned}$$

is hermitian, since $\kappa\big(\sigma_2(u), \sigma_2(v)\big) = \overline{\kappa(u, v)}$ for any $u, v$, because $\sigma_2$ is an antilinear automorphism, and it is also positive definite since the restriction of $h_{\sigma_2}$ to $S^{\sigma_2} \times S^{\sigma_2}$ equals $-\kappa|_{S^{\sigma_2} \times S^{\sigma_2}}$, which is positive definite, since $S^{\sigma_2}$ is compact. Therefore, for any $x \in S^{\sigma_1}_-$, $0 \geq \kappa(x, x) = h_{\sigma_2}(x, x) \geq 0$, so $\kappa(x, x) = 0$, and $x = 0$, since $S^{\sigma_1}$ is compact. Hence $S^{\sigma_1}_- = 0$ and $id = \theta|_{S^{\sigma_1}}$, so $\theta = id$ as $S = S^{\sigma_1} \oplus iS^{\sigma_1}$ and $\sigma_1 = \sigma_2$. □

**1.7 Theorem.** *Let $S$ be a complex semisimple Lie algebra, and let $\sigma$ and $\tau$ be two conjugations, with $\tau$ being compact. Then there is an automorphism $\varphi \in \mathrm{Aut}_{\mathbb{C}} S$ such that $\sigma$ and $\varphi\tau\varphi^{-1}$ (which is compact too) are compatible. Moreover, $\varphi$ can be taken of the form* $\exp(i\,\mathrm{ad}_u)$ *with $u \in K = S^\tau$.*

*Proof.* Consider the positive definite hermitian form

$$\begin{aligned} h_\tau : S \times S &\longrightarrow \mathbb{C} \\ (x, y) &\mapsto -\kappa\big(x, \tau(y)\big) \end{aligned}$$

and the automorphism $\theta = \sigma\tau \in \mathrm{Aut}_{\mathbb{C}} S$. For any $x, y \in S$,

$$h_\tau\big(\theta(x), y\big) = -\kappa\big(\sigma\tau(x), \tau(y)\big) = -\kappa\big(x, \theta^{-1}\tau(y)\big) = -\kappa\big(x, \tau\sigma\tau(y)\big) = h_\tau\big(x, \theta(y)\big).$$

Thus, $\theta$ is selfadjoint relative to $h_\tau$ and, hence, there is an orthonormal basis $\{x_1, \ldots, x_N\}$ of $S$ over $\mathbb{C}$, relative to $h_\tau$, formed by eigenvectors for $\theta$. The corresponding eigenvalues are all real and nonzero. Identify endomorphisms with matrices through this basis to get the diagonal matrices:

$$\theta = \mathrm{diag}(\lambda_1, \ldots, \lambda_N), \quad \theta^2 = \mathrm{diag}(|\lambda_1|^2, \ldots, |\lambda_N|^2) = \exp\Big(\mathrm{diag}\big(2\log|\lambda_1|, \ldots, 2\log|\lambda_N|\big)\Big).$$

For any $r, s = 1, \dots, N$, $[x_r, x_s] = \sum_{j=1}^{N} c_{rs}^j x_j$ for suitable structure constants. With $\mu_j = |\lambda_j|^2 = \lambda_j^2$ for any $j = 1, \dots, N$, and since $\theta^2$ is an automorphism, we get $\mu_r \mu_s c_{rs}^j = \mu_j c_{rs}^j$ for any $r, s, j = 1, \dots, N$, and hence (either $c_{rs}^j = 0$ or $\mu_r \mu_s = \mu_j$) for any $t \in \mathbb{R}$, $\mu_r^t \mu_s^t c_{rs}^j = \mu_j^t c_{rs}^j$, which shows that, for any $t \in \mathbb{R}$,

$$\varphi_t = \operatorname{diag}(\mu_1^t, \dots, \mu_N^t) = \exp\Bigl(\operatorname{diag}\bigl(2t \log|\lambda_1|, \dots, 2t \log|\lambda_N|\bigr)\Bigr)$$

is an automorphism of $S$.

On the other hand, $\tau\theta = \tau\sigma\tau = \theta^{-1}\tau$, so $\tau\varphi_1 = \tau\theta^2 = \theta^{-2}\tau = \varphi_{-1}\tau$. This shows that $\tau \operatorname{diag}(\mu_1, \dots, \mu_N) = \operatorname{diag}(\mu_1^{-1}, \dots, \mu_N^{-1})\tau$ and, as before, this shows that $\tau\varphi_t = \varphi_{-t}\tau$ for any $t \in \mathbb{R}$. Let $\tau' = \varphi_t \tau \varphi_{-t}$. We will look for a value of $t$ that makes $\sigma\tau' = \tau'\sigma$. But,

$$\begin{aligned}
\sigma\tau' &= \sigma\varphi_t\tau\varphi_{-t} = \sigma\tau\varphi_{-2t} = \theta\varphi_{-2t}, \\
\tau'\sigma &= \varphi_t\tau\varphi_{-t}\sigma = \varphi_{2t}\tau\sigma = \varphi_{2t}\theta^{-1} = \theta^{-1}\varphi_{2t}.
\end{aligned}$$

($\theta^{-1}$ and $\varphi_{2t}$ commute as they both are diagonal.)

Hence $\sigma\tau' = \tau'\sigma$ if and only if $\theta^2 = \varphi_{4t}$, if and only if $t = \frac{1}{4}$. Thus we take

$$\varphi = \varphi_{\frac{1}{4}} = \operatorname{diag}\bigl(\mu_1^{\frac{1}{4}}, \dots, \mu_N^{\frac{1}{4}}\bigr) = \exp\Bigl(\operatorname{diag}\bigl(\tfrac{1}{2} \log|\lambda_1|, \dots, \tfrac{1}{2} \log|\lambda_N|\bigr)\Bigr) = \exp d,$$

with $d = \operatorname{diag}\bigl(\frac{1}{2} \log|\lambda_1|, \dots, \frac{1}{2} \log|\lambda_N|\bigr)$. But $\varphi_t \in \operatorname{Aut}_{\mathbb{C}} S$ for any $t \in \mathbb{R}$, so $\exp td \in \operatorname{Aut}_{\mathbb{C}} S$ for any $t \in \mathbb{R}$ and, by differentiating at $t = 0$, this shows that $d$ is a derivation of $S$ so, by Chapter 2, Consequences 2.2, there is a $z \in S$ such that $d = \operatorname{ad}_z$. Note that $d = \operatorname{ad}_z$ is selfadjoint ($(\operatorname{ad}_z)^* = \operatorname{ad}_z$) relative to the hermitian form $h_\tau$. But $S = K \oplus iK$ and for any $u \in K$ and $x, y \in S$

$$\begin{aligned}
h_\tau\bigl([u, x], y\bigr) &= -\kappa\bigl([u, x], \tau(y)\bigr) \\
&= \kappa\bigl(x, [u, \tau(y)]\bigr) \\
&= \kappa\bigl(x, \tau([u, y])\bigr) \quad \text{(since } \tau(u) = u\text{)} \\
&= -h_\tau\bigl(x, [u, y]\bigr),
\end{aligned}$$

so $\operatorname{ad}_u$ is skew relative to $h_\tau$. Therefore, $\operatorname{ad}_z$ is selfadjoint if and only if $z \in iK$. □

**1.8 Remark.** Under the conditions of the proof above, for any $\psi \in \operatorname{Aut}_{\mathbb{R}} S$ such that $\psi\sigma = \sigma\psi$ and $\psi\tau = \tau\psi$, one has $\psi\theta = \theta\psi$ and hence (working with the real basis $\{x_1, ix_1, \dots, x_N, ix_N\}$) one checks that $\psi\varphi_t = \varphi_t\psi$ for any $t \in \mathbb{R}$ so, in particular, $\psi\varphi = \varphi\psi$. That is, the automorphism $\varphi$ commutes with any real automorphism commuting with $\sigma$ and $\tau$.

**1.9 Corollary.** *Let $S$ be a complex semisimple Lie algebra and let $\sigma, \tau$ be two compact conjugations. Then $\sigma$ and $\tau$ are equivalent. That is, up to isomorphism, $S$ has a unique compact form.*

*Proof.* By Theorem 1.7, there is an automorphism $\varphi$ such that $\sigma$ and $\varphi\tau\varphi^{-1}$ are compatible (and compact!). By Proposition 1.6, $\sigma = \varphi\tau\varphi^{-1}$. □

**1.10 Theorem.** *Let $S$ be a complex semisimple Lie algebra, $\theta$ and involutive automorphism of $S$ and $\tau$ a compact conjugation. Then there is an automorphism $\varphi \in \operatorname{Aut}_{\mathbb{C}} S$ such that $\theta$ commutes with $\varphi\tau\varphi^{-1}$. Moreover, $\varphi$ can be taken of the form* $\exp(i\,\mathrm{ad}_u)$ *with $u \in K = S^\tau$. In particular, there is a compact form, namely $\varphi(K)$, which is invariant under $\theta$.*

*Proof.* First note that $(\theta\tau)^2$ is an automorphism of $S$ and for any $x, y \in S$,

$$\begin{aligned} h_\tau\big((\theta\tau)^2(x), y\big) &= -\kappa\big((\theta\tau)^2(x), \tau(y)\big) \\ &= -\kappa\big(x, (\theta\tau)^{-2}\tau(y)\big) \\ &= -\kappa\big(x, \tau(\theta\tau)^2(y)\big) \quad ((\theta\tau)^{-1} = \tau\theta) \\ &= h_\tau(x, (\theta\tau)^2(y)\big), \end{aligned}$$

so $(\theta\tau)^2$ is selfadjoint. Besides,

$$\begin{aligned} h_\tau\big((\theta\tau)^2(x), x\big) &= -\kappa\big((\theta\tau)^2(x), \tau(x)\big) \\ &= -\kappa\big(\theta(\tau\theta\tau)(x), \tau(x)\big) \\ &= -\kappa\big(\tau\theta\tau(x), \theta\tau(x)\big) \quad (\theta \in \operatorname{Aut}_{\mathbb{C}} S \text{ and } \theta^2 = id) \\ &= h_\tau\big(\theta\tau(x), \theta\tau(x)\big) \geq 0, \end{aligned}$$

so $(\theta\tau)^2$ is selfadjoint and positive definite. Hence there is an orthonormal basis of $S$ in which the matrix of $(\theta\tau)^2$ is $\mathrm{diag}(\mu_1, \ldots, \mu_N)$ with $\mu_j > 0$ for any $j = 1, \ldots, N$. Identifying again endomorphisms with their coordinate matrices in this basis, consider the automorphism $\varphi_t = \mathrm{diag}(\mu_1^t, \ldots, \mu_N^t)$ for any $t \in \mathbb{R}$.

Since $\tau(\theta\tau)^2 = (\theta\tau)^{-2}\tau$, it follows that $\tau\varphi_t = \varphi_{-t}\tau$ and, as in the proof of Theorem 1.3, take $\tau' = \varphi_t\tau\varphi_{-t}$. Then,

$$\begin{aligned} \theta\tau' &= \theta\varphi_t\tau\varphi_{-t} = \theta\tau\varphi_{-2t}, \\ \tau'\theta &= \varphi_t\tau\varphi_{-t}\theta = \varphi_{2t}(\theta\tau)^{-1} = (\theta\tau)^{-1}\varphi_{2t}, \end{aligned}$$

where it has been used that, since $(\theta\tau)^2$ commutes with $\theta\tau$, so does $\varphi_t$ for any $t$. Hence $\theta\tau' = \tau'\theta$ if and only if $t = \frac{1}{4}$.

The rest follows as in the proof of Theorem 1.7. □

Now, a map can be defined for any complex semisimple Lie algebra $S$:

$$\Psi : \left\{ \begin{array}{c} \text{Isomorphism classes of} \\ \text{real forms of } S \end{array} \right\} \longrightarrow \left\{ \begin{array}{c} \text{Conjugation classes in } \operatorname{Aut}_{\mathbb{C}} S \\ \text{of involutive automorphisms} \end{array} \right\}$$

$$[\sigma] \qquad \mapsto \qquad [\sigma\tau]$$

where $[\,.\,]$ denotes the corresponding conjugation class and $\tau$ is a compact conjugation that commutes with $\sigma$ (see 1.7). Note that we are identifying any real form with the conjugation class in $\operatorname{Aut}_{\mathbb{C}} S$ of the corresponding conjugation (Proposition 1.6).

**1.11 Theorem.** *The map $\Psi$ above is well defined and bijective.*

*Proof.* If $\sigma$ is a conjugation and $\tau_1$ and $\tau_2$ are compact conjugations commuting with $\sigma$, then there is a $\varphi \in \operatorname{Aut}_{\mathbb{C}} S$ such that $\tau_2 = \varphi\tau_1\varphi^{-1}$ (Corollary 1.9) and $\varphi$ commutes with any real automorphism commuting with $\tau_1$ and $\tau_2$ (Remark 1.8). Hence $\sigma\tau_2 = \sigma\varphi\tau_1\varphi^{-1} = \varphi(\sigma\tau_1)\varphi^{-1}$. Hence $[\sigma\tau_2] = [\sigma\tau_1]$ and, therefore, the image of $[\sigma]$ does not depend on the compact conjugation chosen.

Now, if $\sigma_1, \sigma_2$ are equivalent conjugations and $\varphi \in \operatorname{Aut}_{\mathbb{C}} S$ satisfies $\sigma_2 = \varphi\sigma_1\varphi^{-1}$, if $\tau_1$ is a compact conjugation commuting with $\sigma_1$, then $\tau_2 = \varphi\tau_1\varphi^{-1}$ is a compact conjugation commuting with $\sigma_2$, and $\sigma_2\tau_2 = \varphi\sigma_1\varphi^{-1}\varphi\tau_1\varphi^{-1} = \varphi\sigma_1\tau_1\varphi^{-1}$. Hence, $\Psi$ is well defined.

Let $\theta \in \operatorname{Aut}_{\mathbb{C}} S$ be an involutive automorphism, and let $\tau$ be a compact conjugation commuting with $\theta$ (Theorem 1.10). Then $\sigma = \theta\tau$ is a conjugation commuting with $\tau$ and $\Psi\bigl([\sigma]\bigr) = [\sigma\tau] = [\theta\tau^2] = [\theta]$.

Finally, to check that $\Psi$ is one-to-one, let $\sigma_1, \sigma_2$ be two conjugations and let $\tau_1, \tau_2$ be two compact conjugations with $\tau_i\sigma_i = \sigma_i\tau_i$, $i = 1, 2$. Write $\theta_i = \sigma_i\tau_i$. Assume that there is a $\varphi \in \operatorname{Aut}_{\mathbb{C}} S$ such that $\theta_2 = \varphi\theta_1\varphi^{-1}$. Is $[\sigma_1] = [\sigma_2]$?

By Corollary 1.9, there exists $\psi \in \operatorname{Aut}_{\mathbb{C}} S$ such that $\tau_2 = \psi\tau_1\psi^{-1}$. Thus, $[\sigma_1] = [\psi\sigma_1\psi^{-1}]$ and $\Psi\bigl([\sigma_1]\bigr) = [\psi\sigma_1\psi^{-1}\psi\tau_1\psi^{-1}] = [\psi\sigma_1\psi^{-1}\tau_2]$. Hence we may assume that $\tau_1 = \tau_2 = \tau$, so $\theta_i = \sigma_i\tau$, $i = 1, 2$.

Now, by Theorem 1.7 and Remark 1.8, there is an automorphism $\gamma \in \operatorname{Aut}_{\mathbb{C}} S$ such that $\gamma\tau\gamma^{-1}$ and $\varphi^{-1}\tau\varphi$ are compatible and $\gamma$ commutes with $\theta_1$, since $\theta_1$ commutes with $\tau$, and also $\theta_1 = \varphi^{-1}\theta_2\varphi$ commutes with $\varphi^{-1}\tau\varphi$. But two compatible compact conjugations coincide (Proposition 1.6), so $\gamma\tau\gamma^{-1} = \varphi^{-1}\tau\varphi$. Then,

$$\begin{aligned}
\sigma_2 = \theta_2\tau &= \varphi\theta_1\varphi^{-1}\tau \\
&= \varphi\theta_1(\varphi^{-1}\tau\varphi)\varphi^{-1} = \varphi\theta_1\gamma\tau\gamma^{-1}\varphi^{-1} \\
&= \varphi\gamma\theta_1\tau(\varphi\gamma)^{-1} \\
&= (\varphi\gamma)\sigma_1(\varphi\gamma)^{-1}.
\end{aligned}$$

Hence $[\sigma_1] = [\sigma_2]$. $\square$

**1.12 Remark.**

(i) The proof of Theorem 1.3 shows that $\Psi([\text{‘split form’}]) = [\omega]$ ($\omega(x_j) = -y_j$, $\omega(y_j) = -x_j$ for any $j$). Trivially, $\Psi([\text{‘compact form’}]) = [id]$.

(ii) Let $\theta \in \operatorname{Aut}_{\mathbb{C}} S$ be an involutive automorphism, and let $\tau$ be a compact conjugation commuting with $\theta$. Take $\sigma = \theta\tau$. Then $S^\tau = K = K_{\bar 0} \oplus K_{\bar 1}$, where $K_{\bar 0} = \{x \in K : \theta(x) = x\}$ and $K_{\bar 1} = \{x \in K : \theta(x) = -x\}$. Then the real form corresponding to $\sigma$ is $S^\sigma = L = K_{\bar 0} \oplus iK_{\bar 1}$, and its Killing form is

$$\kappa_L = \kappa|_L = \kappa|_{K_{\bar 0}} \perp \kappa|_{iK_{\bar 1}} \cong \kappa|_{K_{\bar 0}} \perp \Bigl(-\kappa|_{K_{\bar 1}}\Bigr).$$

Since $\kappa|_{K_{\bar 0}}$ and $\kappa|_{K_{\bar 1}}$ are negative definite, the signature of $\kappa_L$ is $\dim_{\mathbb{R}} K_{\bar 1} - \dim_{\mathbb{R}} K_{\bar 0} = \dim_{\mathbb{C}} S_{\bar 1} - \dim_{\mathbb{C}} S_{\bar 0}$, where $S_{\bar 0} = \{x \in S : \theta(x) = x\}$ and $S_{\bar 1} = \{x \in S : \theta(x) = -x\}$.

The decomposition $L = K_{\bar 0} \oplus iK_{\bar 1}$ is called a *Cartan decomposition* of $L$.

(iii) To determine the real simple Lie algebras it is enough then to classify the involutive automorphisms of the simple complex Lie algebras, up to conjugation. This will be done over arbitrary algebraically closed fields of characteristic 0 by a process based on the paper by A.W. Knapp: "A quick proof of the classification of simple real Lie algebras", Proc. Amer. Math. Soc. **124** (1996), no. 10, 3257–3259.

## § 2. Involutive automorphisms

Let $k$ be an algebraically closed field of characteristic 0, and let $L$ be a semisimple Lie algebra over $k$, $H$ a fixed Cartan subalgebra of $L$, $\Phi$ the corresponding root system and $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ a system of simple roots. Let $x_1, \ldots, x_n, y_1, \ldots, y_n$ be the canonical generators that are being used throughout.

(i) For any subset $J \subseteq \{1, \ldots, n\}$, there is a unique involutive automorphism $\theta_J$ of $L$ such that

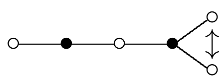
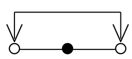
$$\begin{cases} \theta_J(x_i) = x_i, & \theta_J(y_i) = y_i, & \text{if } i \notin J, \\ \theta_J(x_i) = -x_i, & \theta_J(y_i) = -y_i, & \text{if } i \in J. \end{cases}$$

We will say that $\theta_J$ corresponds to the Dynkin diagram of $(\Phi, \Delta)$, where the nodes corresponding to the roots $\alpha_i$, $i \in J$, are shaded.

(ii) Also, if $\omega$ is an 'involutive automorphism' of the Dynkin diagram of $(\Phi, \Delta)$, that is, a bijection among the nodes of the diagram that respects the Cartan integers, and if $J$ is a subset of $\{1, \ldots, n\}$ consisting of fixed nodes by $\omega$, then there is a unique involutive automorphism $\theta_{\omega,J}$ of $L$ given by,

$$\begin{cases} \theta_{\omega,J}(x_i) = x_{\omega(i)}, & \theta_{\omega,J}(y_i) = y_{\omega(i)}, & \text{if } i \notin J, \\ \theta_{\omega,J}(x_i) = -x_i, & \theta_{\omega,J}(y_i) = -y_i, & \text{if } i \in J. \end{cases}$$

We will say that $\theta_{\omega,J}$ corresponds to the Dynkin diagram of $(\Phi, \Delta)$ with the nodes in $J$ shaded and where $\omega$ is indicated by arrows, like the following examples:

These diagrams, where some nodes are shaded and a possible involutive diagram automorphism is specified by arrows, are called *Vogan diagrams* (see A.W. Knapp: *Lie groups beyond an Introduction*, Birkhäuser, Boston 1996).

**2.1 Theorem.** *Let $k$ be an algebraically closed field of characteristic* 0*. Then, up to conjugation, the involutive automorphisms of the simple Lie algebras are the automorphisms that correspond to the Vogan diagrams that appear in Tables A.1, A.2.*

In these tables, one has to note that for the orthogonal Lie algebras of small dimension over an algebraically closed field of characteristic 0, one has the isomorphisms $\mathfrak{so}_3 \cong A_1$, $\mathfrak{so}_4 \cong A_1 \times A_1$ and $\mathfrak{so}_6 \cong A_3$. Also, $Z$ denotes a one-dimensional Lie algebra, and $\mathfrak{so}_{r,s}(\mathbb{R})$ denotes the orthogonal Lie algebra of a nondegenerate quadratic form of dimension $r+s$ and signature $r-s$. Besides, $\mathfrak{so}^*_{2n}(\mathbb{R})$ denotes the Lie algebra of the skew matrices relative to a nondegenerate antihermitian form on a vector space over the quaternions: $\mathfrak{so}^*_{2n}(\mathbb{R}) = \{x \in \mathrm{Mat}_n(\mathbb{H}) : x^t h + h\bar{x} = 0\}$, where $h = \mathrm{diag}(i,\dots,i)$. In the same vein, $\mathfrak{sp}_n(\mathbb{H}) = \{x \in \mathrm{Mat}_n(\mathbb{H}) : x^t + \bar{x} = 0\}$, while $\mathfrak{sp}_{r,s}(\mathbb{H}) = \{x \in \mathrm{Mat}_{r+s}(\mathbb{H}) : x^t h + h\bar{x} = 0\}$, where $h = \mathrm{diag}(1,\dots,1,-1,\dots,-1)$ ($r$ 1's and $s$ $-1$'s). Finally, an expression like $E_{8,-24}$ denotes a real form of $E_8$ such that the signature of its Killing form is $-24$.

*Proof.* Let $L$ be a simple Lie algebra over $k$ and let $\varphi \in \mathrm{Aut}\, L$ be an involutive automorphism. Then $L = S \oplus T$, with $S = \{x \in L : \varphi(x) = x\}$ and $T = \{x \in L : \varphi(x) = -x\}$. The subspaces $S$ and $T$ are orthogonal relative to the Killing form (since the Killing form $\kappa$ is invariant under $\varphi$).

**(i)** There exists a Cartan subalgebra $H$ of $L$ which contains a Cartan subalgebra of $S$ and is invariant under $\varphi$:

In fact, the adjoint representation $\mathrm{ad} : S \to \mathfrak{gl}(L)$ has a nondegenerate trace form, so $S = Z(S) \oplus [S,S]$ and $[S,S]$ is semisimple (Chapter 2, 2.2). Besides, for any $x \in Z(S)$, $x = x_s + x_n$ with $x_s, x_n \in N_L(T) \cap C_L(S) = S \cap C_L(S) = Z(S)$ (as the normalizer $N_L(T)$ is invariant under $\varphi$ and $N_L(T) \cap T$ is an ideal of $L$ and hence trivial). Besides, $\kappa(x_n, S) = 0$, so $x_n = 0$. Hence $Z(S)$ is a toral subalgebra, and there is a Cartan subalgebra $H_S$ of $S$ with $H_S = Z(S) \oplus \bigl(H_S \cap [S,S]\bigr)$. Then $H_S$ is toral on $L$. Let $H = C_L(H_S) = H_S \oplus H_T$, where $H_T = C_L(H_S) \cap T$. Then $[H,H] = [H_T,H_T] \subseteq S$. Hence $[[H,H],H] = 0$, so $H$ is a nilpotent subalgebra. Thus, $[H,H]$ acts both nilpotently and semisimply on $L$. Therefore, $[H,H] = 0$ and $H$ is a Cartan subalgebra of $L$, since for any $x \in H_T$, $x_n \in H$, $\kappa(x_n, H) = 0$ and, as $H$ is the zero weight space relative to $H_S$, the restriction of $\kappa$ to $H$ is nondegenerate, hence $x_n = 0$ and $H$ is toral.

**(ii)** Fix one such Cartan subalgebra $H$ and let $\Phi$ be the associated set of roots. Then $\varphi$ induces a linear map $\varphi^* : H^* \to H^*$, $\alpha \mapsto \bar{\alpha} = \alpha \circ \varphi|_H$. Since $\varphi$ is an automorphism, $\varphi(L_\alpha) = L_{\bar{\alpha}}$ for any $\alpha \in \Phi$, so $\bar{\Phi} = \Phi$. Besides, for any $\alpha \in \Phi$ and any $h \in H$, $\bar{\alpha}(h) = \alpha(\varphi(h)) = \kappa(t_\alpha, \varphi(h)) = \kappa(\varphi(t_\alpha), h)$, so $\varphi(t_\alpha) = t_{\bar{\alpha}}$ for any $\alpha \in \Phi$. This shows that $\sum_{\alpha\in\Phi} \mathbb{Q} t_\alpha$ is invariant under $\varphi$.

**(iii)** Consider the subsets $\Phi_S = \{\alpha \in \Phi : L_\alpha \subseteq S\}$ and $\Phi_T = \{\alpha \in \Phi : L_\alpha \subseteq T\}$. Then $\Phi_S \cup \Phi_T = \{\alpha \in \Phi : \bar{\alpha} = \alpha\}$:

Actually, $[H_T, S] \subseteq T$ and $[H_T, T] \subseteq S$, so for any $\alpha \in \Phi_S \cup \Phi_T$, $\alpha(H_T) = 0$ and $\alpha = \bar{\alpha}$. Conversely, if $\alpha(H_T) = 0$, then $L_\alpha = (L_\alpha \cap S) \oplus (L_\alpha \cap T)$ and, since $\dim L_\alpha = 1$, either $L_\alpha \subseteq S$ or $L_\alpha \subseteq T$.

**(iv)** The rational vector space $\check{E} = \sum_{\alpha\in\Phi} \mathbb{Q} t_\alpha$ is invariant under $\varphi$ and $\kappa|_{\check{E}}$ is positive definite (taking values on $\mathbb{Q}$). Hence $\check{E} = \check{E}_S \perp \check{E}_T$, where $\check{E}_S = \check{E} \cap S$ and $\check{E}_T = \check{E} \cap T$. Also, $\Phi \subseteq E = \sum_{\alpha\in\Phi} \mathbb{Q}\alpha = E_S \oplus E_T$, where $E_S = \{\alpha \in E : \alpha(H_T) = 0\}$ and $E_T = \{\alpha \in E : \alpha(H_S) = 0\}$, with $E_S$ and $E_T$ orthogonal relative to the positive definite symmetric bilinear form $(\,.\,|\,.\,)$ induced by $\kappa$. Moreover, $\Phi \cap E_T = \emptyset$:

In fact, if $\alpha \in \Phi$ and $\alpha(H_S) = 0$, then for any $x = x_S + x_T \in L_\alpha$ ($x_S \in S$, $x_T \in T$), and any $h \in H_S$, $[h, x_S + x_T] = \alpha(h)(x_S + x_T) = 0$. Hence $x_S \in C_S(H_S) = H_S$ and $[H, x_S] = 0$. Now, for any $h \in H_T$, $\alpha(h)(x_S + x_T) = [h, x_S + x_T] = [h, x_T] \in S$. Hence $x_T = 0 = x_S$, a contradiction.

**(v)** There is a system of simple roots $\Delta$ such that $\bar{\Delta} = \Delta$:

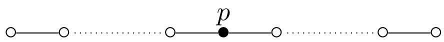
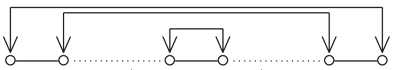
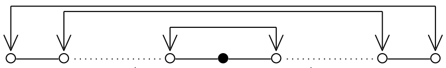
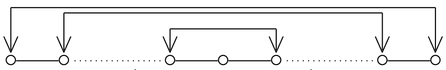
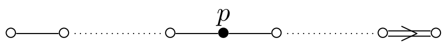
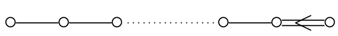
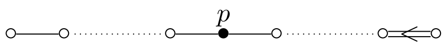
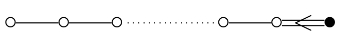
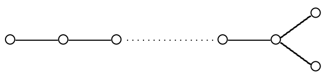
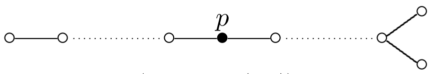
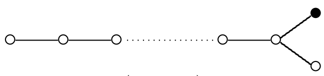
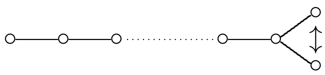
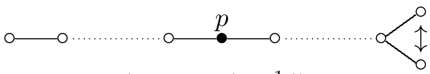
| **Type** | **Vogan diagram** | **Fixed subalgebra** | **Real form** ($k=\mathbb{C}$) |
|---|---|---|---|
| $A_n$ |  | $A_n \quad (\varphi = id)$ | $\mathfrak{su}_{n+1}(\mathbb{R})$ |
| |  $(1 \le p \le \left[\frac{n}{2}\right])$ | $A_{p-1} \times A_{n-p} \times Z$ $(A_0 = 0)$ | $\mathfrak{su}_{p,n+1-p}(\mathbb{R})$ |
| |  $(n = 2r > 1)$ | $\mathfrak{so}_{2r+1}$ | $\mathfrak{sl}_{n+1}(\mathbb{R})$ |
| |  $(n = 2r-1 > 1)$ | $\mathfrak{so}_{2r}$ | $\mathfrak{sl}_{n+1}(\mathbb{R})$ |
| |  $(n = 2r-1 > 1)$ | $\mathfrak{sp}_{2r}$ | $\mathfrak{sl}_{r}(\mathbb{H})$ |
| $B_n$ |  | $B_n \quad (\varphi = id)$ | $\mathfrak{so}_{2n+1}(\mathbb{R})$ |
| |  $(1 \le p \le n)$ | $\mathfrak{so}_{2n+1-p} \times \mathfrak{so}_p$ $\big(\mathfrak{so}_1 = 0,\ \mathfrak{so}_2 = Z\big)$ | $\mathfrak{so}_{2n+1-p,p}(\mathbb{R})$ |
| $C_n$ |  | $C_n \quad (\varphi = id)$ | $\mathfrak{sp}_n(\mathbb{H})$ |
| |  $(1 \le p \le \lfloor\frac{n}{2}\rfloor)$ | $\mathfrak{sp}_{2p} \times \mathfrak{sp}_{2(n-p)}$ | $\mathfrak{sp}_{n-p,p}(\mathbb{H})$ |
| |  | $A_{n-1} \times Z$ | $\mathfrak{sp}_{2n}(\mathbb{R})$ |
| $D_n$ |  | $D_n \quad (\varphi = id)$ | $\mathfrak{so}_{2n}(\mathbb{R})$ |
| |  $(1 \le p \le \lfloor\frac{n}{2}\rfloor)$ | $\mathfrak{so}_{2(n-p)} \times \mathfrak{so}_{2p}$ | $\mathfrak{so}_{2(n-p),2p}(\mathbb{R})$ |
| |  $(n > 4)$ | $A_{n-1} \times Z$ | $\mathfrak{so}^*_{2n}(\mathbb{R})$ |
| |  | $B_{n-1}$ | $\mathfrak{so}_{2n-1,1}(\mathbb{R})$ |
| |  $(1 \le p \le \lfloor\frac{n-1}{2}\rfloor)$ | $\mathfrak{so}_{2n-2p-1} \times \mathfrak{so}_{2p+1}$ | $\mathfrak{so}_{2n-2p-1,2p+1}(\mathbb{R})$ |

Table A.1: Involutive automorphisms: classical cases

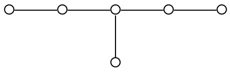
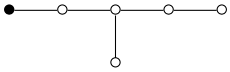
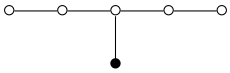
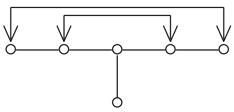
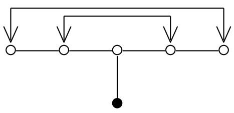
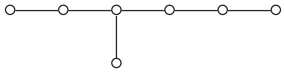
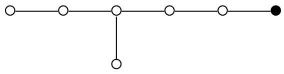
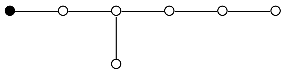
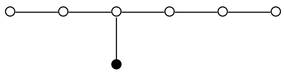
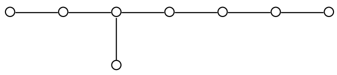
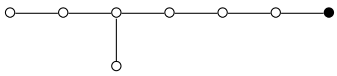
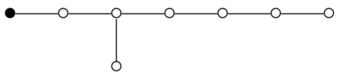
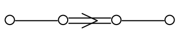
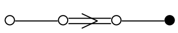
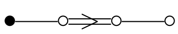
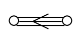
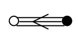
| **Type** | **Vogan diagram** | **Fixed subalgebra** | **Real form** ($k = \mathbb{C}$) |
|---|---|---|---|
| $E_6$ |  | $E_6$ | $E_{6,-78}$ |
| |  | $D_5 \times Z$ | $E_{6,-14}$ |
| |  | $A_5 \times A_1$ | $E_{6,2}$ |
| |  | $F_4$ | $E_{6,-26}$ |
| |  | $C_4$ | $E_{6,6}$ |
| $E_7$ |  | $E_7$ | $E_{7,-133}$ |
| |  | $E_6 \times Z$ | $E_{7,-25}$ |
| |  | $D_6 \times A_1$ | $E_{7,-5}$ |
| |  | $A_7$ | $E_{7,7}$ |
| $E_8$ |  | $E_8$ | $E_{8,-248}$ |
| |  | $E_7 \times A_1$ | $E_{8,-24}$ |
| |  | $D_8$ | $E_{8,8}$ |
| $F_4$ |  | $F_4$ | $F_{4,-52}$ |
| |  | $B_4$ | $F_{4,-20}$ |
| |  | $C_3 \times A_1$ | $F_{4,4}$ |
| $G_2$ |  | $G_2$ | $G_{2,-14}$ |
| |  | $A_1 \times A_1$ | $G_{2,2}$ |

Table A.2: Involutive automorphisms: exceptional cases

For any $\alpha \in \Phi$, $\alpha = \alpha_S + \alpha_T$, with $\alpha_S \in E_S$, $\alpha_T \in E_T$ and $\alpha_S \neq 0$ because of **(iv)**. Choose $\beta \in E_S$ such that $(\beta|\alpha) = (\beta|\alpha_S) \neq 0$ for any $\alpha \in \Phi$. Then $(\alpha|\beta) = (\bar\alpha|\bar\beta) = (\bar\alpha|\beta)$ for any $\alpha \in \Phi$, so that in the total order on $\Phi$ given by $\beta$, $\Phi^+ = \overline{\Phi^+}$ and $\alpha \in \Phi^+$ is simple if and only if so is $\bar\alpha$.

**(vi)** Let $\Delta$ be a system of simple roots invariant under $\varphi$, hence

$$\Delta = \{\alpha_1, \dots, \alpha_s, \alpha_{s+1}, \dots, \alpha_{s+2r}\},$$

with $\alpha_i = \bar\alpha_i$, for $i = 1, \dots, s$, and $\bar\alpha_{s+2i-1} = \alpha_{s+2i}$ for $i = 1, \dots, r$. Let $\alpha = m_1\alpha_1 + \cdots + m_{s+2r}\alpha_{s+2r}$ be a root with $\bar\alpha = \alpha$ and assume that $s \geq 1$. Then $\alpha \in \Phi_S$ (respectively $\alpha \in \Phi_T$) if and only if $\sum_{\alpha_i \in \Phi_T} m_i$ is even (respectively odd):

To prove this, it can be assumed that $\alpha \in \Phi^+$. We will proceed by induction on $\mathrm{ht}(\alpha)$. If $\mathrm{ht}(\alpha) = 1$, then $\alpha = \bar\alpha$, so there is an index $i = 1, \dots, s$ such that $\alpha = \alpha_i$ and the result is trivial. Hence assume that $\mathrm{ht}(\alpha) = n > 1$ and that $\alpha = \bar\alpha$. If there is an $i = 1, \dots, s$ such that $(\alpha|\alpha_i) > 0$, then $\alpha = \beta + \alpha_i$, for some $\beta \in \Phi^+$ with $\beta = \bar\beta$. Besides $L_\alpha = [L_\beta, L_{\alpha_i}]$ and the induction hypothesis applies. Otherwise, there is an index $j > s$ such that $(\alpha|\alpha_j) > 0$, so $(\alpha|\bar\alpha_j) = (\bar\alpha|\bar\alpha_j) = (\alpha|\alpha_j) > 0$ and $(\alpha - \alpha_j|\bar\alpha_j) = (\alpha|\bar\alpha_j) - (\alpha_j|\bar\alpha_j) > 0$, since $(\alpha_j|\bar\alpha_j) \leq 0$. Note that, since $s \geq 1$, $\alpha_j$ and $\bar\alpha_j$ are not connected in the Dynkin diagram, since $\varphi^*$ induces an automorphism of the diagram (the only possibility for $\alpha_j$ and $\bar\alpha_j$ to be connected would be in a diagram $A_{2r}$, but with $s = 0$), hence $\alpha_j + \bar\alpha_j \notin \Phi$, so if $L_{\alpha_j} = k(x_S + x_T)$, then $L_{\bar\alpha_j} = k(x_S - x_T)$ and $0 = [x_S + x_T, x_S - x_T] = -2[x_S, x_T]$. Therefore, $\alpha = \beta + \alpha_j + \bar\alpha_j$, with $\beta, \beta + \alpha_j \in \Phi$, $L_\alpha = [L_{\bar\alpha_j}, [L_{\alpha_j}, L_\beta]]$ and $\beta = \bar\beta$. Hence $L_\alpha = \mathrm{ad}_{x_S - x_T}\mathrm{ad}_{x_S + x_T}(L_\beta) = \big(\mathrm{ad}^2_{x_S} - \mathrm{ad}^2_{x_T}\big)(L_\beta)$. Thus, $L_\alpha$ is contained in $S$ (respectively $T$) if and only if so is $L_\beta$.

Once we have such a system of simple roots, it is clear that canonical generators of $L$ can be chosen so that $\varphi$ becomes the automorphism associated to a Vogan diagram (if $\varphi^*(\alpha_i) = \alpha_j$ with $i \neq j$, then it is enough to take $x_j = \varphi(x_i)$ and $y_j = \varphi(y_i)$). Let us check that it is possible to choose such a system $\Delta$ so that the corresponding Vogan diagram is one of the diagrams that appear in Tables A.1, A.2, where there is at most a node shaded and this node has some restrictions.

**(vii)** Let $\Lambda = \{\mu \in E_S : (\alpha|\mu) \in \mathbb{Z}\ \forall \alpha \in \Phi \text{ and } (\alpha|\mu) \in 2\mathbb{Z} + 1\ \forall \alpha \in \Phi_T\}$. Then, if $s \geq 1$, $\Lambda \neq \emptyset$:

Note that with $\Delta$ as above, $E_S = \sum_{i=1}^s \mathbb{Q}\alpha_i + \sum_{j=1}^r \mathbb{Q}\big(\alpha_{s+2j-1} + \alpha_{s+2j}\big)$, while $E_T = \sum_{j=1}^r \mathbb{Q}\big(\alpha_{s+2j-1} - \alpha_{s+2j}\big)$. Let $\{\mu_i\}_{i=1}^{s+2r}$ be the dual basis of $\Delta$. Then $\mu_1, \dots, \mu_s$ are orthogonal to $\alpha_{s+2j-1}$ and $\alpha_{s+2j}$ for any $j = 1, \dots, r$, so $\mu_1, \dots, \mu_s \in E_T^\perp = E_S$. Also, the invariance of $(\,.\,|\,.\,)$ under the automorphism induced by $\varphi$ shows that $\mu_{s+2j} = \bar\mu_{s+2j-1}$ for any $j = 1, \dots, r$. Let $\mu = \sum_{\{i : \alpha_i \in \Phi_T\}} \mu_i$, which satisfies that $(\alpha_i|\mu) = 1$ for any $i$ with $\alpha_i \in \Phi_T$ and $(\alpha_j|\mu) = 0$ otherwise. Hence by **(vi)** $(\alpha|\mu) \in 2\mathbb{Z} + 1$ for any $\alpha \in \Phi_T$.

**(viii)** Note that $\Lambda \subseteq \{\mu \in E : (\mu|\alpha) \in \mathbb{Z}\ \forall \alpha \in \Delta\} = \mathbb{Z}\mu_1 + \cdots + \mathbb{Z}\mu_{s+2r}$, which is a discrete subset of $E$. Let $0 \neq \mu \in \Lambda$ of minimal norm. Then there exists a system of simple roots $\Delta'$ such that $\Delta' = \bar\Delta'$ and with $(\mu|\alpha) \geq 0$ for any $\alpha \in \Delta'$:

Let $\beta \in E_S$ as in **(v)**, take a positive and large enough $r \in \mathbb{Q}$ such that, for any $\alpha \in \Phi$, $(\alpha|\beta + r\mu)$ is $> 0$ if and only if, either $(\alpha|\mu) > 0$ or $(\alpha|\mu) = 0$ and $(\alpha|\beta) > 0$. Then consider the total order in $\Phi$ given by $\beta' = \beta + r\mu$ ($\beta' = \bar\beta'$). The associated system of simple roots $\Delta'$ satisfies the required conditions.

**(ix)** Take $\mu$ and the system of simple roots $\Delta'$ in **(viii)**. Then

$$\Delta' = \{\alpha'_1, \dots, \alpha'_{s'}, \alpha'_{s'+1}, \dots, \alpha'_{s'+2r'}\},$$

with $\bar{\alpha}'_i = \alpha'_i$, $i = 1, \dots, s'$ and $\bar{\alpha}'_{s'+2j-1} = \alpha'_{s'+2j}$, $j = 1, \dots, r'$. Let $\{\mu'_i\}_{i=1}^{s'+2r'}$ be the dual basis to $\Delta'$. Since $(\mu|\alpha) \geq 0$ for any $\alpha \in \Phi^+$ ($\mu$ is said *dominant* then), and $\bar{\mu} = \mu$,

$$\mu = \sum_{i=1}^{s'} m_i \mu'_i + \sum_{j=1}^{r'} m_{s'+j}\big(\mu'_{s'+2j-1} + \mu'_{s'+2j}\big),$$

with $m_1, \dots, m_{s'+r'} \in \mathbb{Z}_{\geq 0}$.

Note that if $0 \neq h_1, h_2 \in \sum_{\alpha \in \Phi^+} \mathbb{Q}t_\alpha$ with $\alpha(h_i) \geq 0$ for any $\alpha \in \Phi^+$ and $i = 1, 2$, then $\kappa(h_1, h_2) = \operatorname{trace}\big(\operatorname{ad}_{h_1} \operatorname{ad}_{h_2}\big) = 2\sum_{\alpha \in \Phi^+} \alpha(h_1)\alpha(h_2) > 0$ (use Exercise 6.15 in Chapter 2). As a consequence, the inner product of any two nonzero dominant elements of $E$ is $> 0$.

Hence if some $m_i > 0$, $i = 1, \dots, s'$, then $\mu - \mu'_i$ is dominant, so $(\mu - \mu'_i|\mu'_i) \geq 0$ and this is 0 if and only if $\mu = \mu'_i$. Now, $\mu - 2\mu'_i \in \Lambda$ and $(\mu - 2\mu'_i|\mu - 2\mu'_i) = (\mu|\mu) - 4(\mu - \mu'_i|\mu'_i) \leq (\mu|\mu)$. By the minimality of $\mu$, we conclude that $(\mu - \mu'_i|\mu'_i) = 0$ and $\mu = \mu'_i$. Therefore $\Delta' \cap \Phi_T \subseteq \{\alpha'_i\}$.

On the other hand, if $m_i = 0$, for any $i = 1, \dots, s'$, then $(\mu|\alpha'_i) = 0$ (even!), so $\Delta' \cap \Phi_T = \emptyset$.

Therefore there is at most one shaded node in the associated Vogan diagram. More precisely, either $\Delta' \cap \Phi_T = \emptyset$, or $\Delta' \cap \Phi_T = \{\alpha'_i\}$ and $\mu = \mu'_i$ for some $i = 1, \dots, s'$. In this latter case, for any $i \neq j = 1, \dots, s'$, $(\mu - \mu'_j|\mu'_j) \leq 0$ (otherwise $\mu - 2\mu'_j \in \Lambda$ with $\big(\mu - 2\mu'_j|\mu - 2\mu'_j\big) < (\mu|\mu)$). Also, if for some $j = 1, \dots, r'$,

$$\Big(\mu - \frac{1}{2}(\mu'_{s'+2j-1} + \mu'_{s'+2j}) \mid \mu'_{s'+2j-1} + \mu'_{s'+2j}\Big) > 0,$$

we would have

$$\Big(\mu - (\mu'_{s'+2j-1} + \mu'_{s'+2j}) \mid \mu - (\mu'_{s'+2j-1} + \mu'_{s'+2j})\Big) < (\mu|\mu),$$

a contradiction with the minimality of $\mu$, since $\mu - (\mu'_{s'+2j-1} + \mu'_{s'+2j}) \in \Lambda$, because for any $\alpha \in \Phi_T$, $\big(\mu'_{s'+2j-1}|\alpha\big) = \big(\bar{\mu}'_{s'+2j-1}|\bar{\alpha}\big) = \big(\mu'_{s'+2j}|\alpha\big)$.

Therefore, if $\Delta' \cap \Phi_T = \{\alpha'_i\}$ for some $i = 1, \dots, r$, then

$$\begin{aligned}
&\mu = \mu'_i,\\
(2.1)\qquad &(\mu - \mu'_j|\mu'_j) \leq 0 \text{ for any } i \neq j = 1, \dots, s,\\
&\big(\mu - \tfrac{1}{2}(\mu'_{s'+2j-1} + \mu'_{s'+2j})|\mu'_{s'+2j-1} + \mu'_{s'+2j}\big) \leq 0, \text{ for any } j = 1, \dots, r.
\end{aligned}$$

(The last condition in (2.1) does not appear in Knapp's article.)

**(x)** Looking at Tables A.1, A.2, what remains to be proved is to check that for Vogan diagrams associated to the Lie algebras of type $C_n$, $D_n$, $E_6$, $E_7$, $E_8$, $F_4$ or $G_2$, in case there is a shaded node, this node satisfies the requirements in the Tables A.1, A.2. This can be deduced easily case by case from (2.1):

- For $C_n$, order the roots as follows,

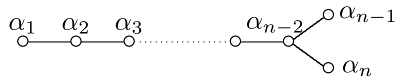

  Here $\alpha_i = \epsilon_i - \epsilon_{i+1}$, $i = 1, \dots, n-1$ and $\alpha_n = 2\epsilon_n$ where, up to a nonzero scalar, $(\epsilon_i|\epsilon_j) = \delta_{ij}$ for any $i, j$. Hence $\mu'_i = \epsilon_1 + \cdots + \epsilon_i$ for $i = 1, \dots, n-1$ and $\mu'_n = \frac{1}{2}(\epsilon_1 + \cdots + \epsilon_n)$.

For any $i = 1, \ldots, n-1$,

$$\big(\mu_i' - \mu_n' | \mu_n'\big) = \frac{1}{2}\big(i - (n-i)\big) = \frac{1}{2}(2i - n),$$

so (2.1) is satisfied if and only if $i \leq \left\lfloor \frac{n}{2} \right\rfloor$.

- For $D_n$,

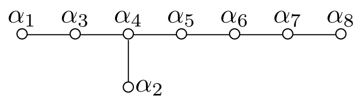



Here either $\varphi^*$ is the identity, or $\bar{\alpha}_{n-1} = \alpha_n$ ($\bar{\alpha}_i = \alpha_i$ for $i \leq n-1$). Also, $\alpha_i = \epsilon_i - \epsilon_{i+1}$ for $i = 1, \ldots, n-1$ and $\alpha_n = \epsilon_{n-1} + \epsilon_n$ where, up to a scalar, $(\epsilon_i | \epsilon_j) = \delta_{ij}$. Hence $\mu_i' = \epsilon_1 + \cdots + \epsilon_i$, for $i = 1, \ldots, n-2$, $\mu_{n-1}' = \frac{1}{2}(\epsilon_1 + \cdots + \epsilon_{n-1} - \epsilon_n)$ and $\mu_n' = \frac{1}{2}(\epsilon_1 + \cdots + \epsilon_{n-1} + \epsilon_n)$.
For any $i = 1, \ldots, n-2$,

$$\big(\mu_i' - \mu_n' | \mu_n'\big) = \frac{1}{4}(2i - n), \quad \Big(\mu_i' - \frac{1}{2}(\mu_{n-1}' + \mu_n') \mid \mu_{n-1}' + \mu_n'\Big) = \frac{1}{4}\big(2i - (n-1)\big),$$

so if $\varphi^* = id$, then (2.1) is satisfied if $i \leq \left\lfloor \frac{n}{2} \right\rfloor$, while if $\varphi^* \neq id$, (2.1) is satisfied if $i \leq \left\lfloor \frac{n-1}{2} \right\rfloor$.

- For $E_8$, take the simple roots as follows:




Here $\Delta_8 = \{\alpha_1, \ldots, \alpha_8\}$ with

$$\begin{aligned}
\alpha_1 &= \frac{1}{2}(\epsilon_1 - \epsilon_2 - \cdots - \epsilon_7 + \epsilon_8), \\
\alpha_2 &= \epsilon_1 + \epsilon_2, \\
\alpha_i &= \epsilon_{i-1} - \epsilon_{i-2}, \ i = 3, \ldots, 8,
\end{aligned}$$

for an orthonormal basis (up to a scaling of the inner product) $\{\epsilon_i : i = 1, \ldots, 8\}$. Hence

$$\begin{aligned}
\mu_1' &= 2\epsilon_8, \\
\mu_2' &= \frac{1}{2}(\epsilon_1 + \cdots + \epsilon_7 + 5\epsilon_8), \\
\mu_3' &= \frac{1}{2}(-\epsilon_1 + \epsilon_2 + \cdots + \epsilon_7 + 7\epsilon_8), \\
\mu_i' &= \epsilon_{i-1} + \cdots + \epsilon_7 + (9-i)\epsilon_8, \ i = 4, \ldots, 8.
\end{aligned}$$

For any $i = 2, \ldots, 6$,

$$\big(\mu_i' - \mu_1' | \mu_1'\big) > 0, \quad \big(\mu_i' - \mu_8' | \mu_8'\big) > 0,$$

so if (2.1) is satisfied, then $i = 1$ or $i = 8$.

- For $E_7$, $\Delta_7 = \Delta_8 \setminus \{\alpha_8\}$. It follows that

$$
\begin{aligned}
\mu'_1 &= \epsilon_8 - \epsilon_7,\\
\mu'_2 &= \frac{1}{2}\big(\epsilon_1 + \cdots + \epsilon_6 + 2(\epsilon_8 - \epsilon_7)\big),\\
\mu'_3 &= \frac{1}{2}\big(-\epsilon_1 + \cdots + \epsilon_6 + 3(\epsilon_8 - \epsilon_7)\big),\\
\mu'_4 &= \epsilon_3 + \cdots + \epsilon_6 + 2(\epsilon_8 - \epsilon_7),\\
\mu'_5 &= \epsilon_4 + \epsilon_5 + \epsilon_6 + \frac{3}{2}(\epsilon_8 - \epsilon_7),\\
\mu'_6 &= \epsilon_5 + \epsilon_6 + (\epsilon_8 - \epsilon_7),\\
\mu'_7 &= \epsilon_6 + \frac{1}{2}(\epsilon_8 - \epsilon_7).
\end{aligned}
$$

  Hence $\big(\mu'_i - \mu'_7|\mu'_7\big) > 0$ for $i = 3, 4, 5, 6$, so (2.1) imply that $i = 1, 2$ or 7.

- For $E_6$, take $\Delta_6 = \Delta_7 \setminus \{\alpha_7\}$. Here either $\varphi^* = id$ or it interchanges $\alpha_1$ and $\alpha_6$, and $\alpha_3$ and $\alpha_5$. Besides,

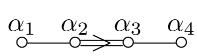
$$
\begin{aligned}
\mu'_1 &= \frac{2}{3}(\epsilon_8 - \epsilon_7 - \epsilon_6),\\
\mu'_2 &= \frac{1}{2}\big(\epsilon_1 + \cdots + \epsilon_5 + (\epsilon_8 - \epsilon_7 - \epsilon_6)\big),\\
\mu'_3 &= \frac{1}{2}\big(-\epsilon_1 + \cdots + \epsilon_5\big) + \frac{5}{6}\big(\epsilon_8 - \epsilon_7 - \epsilon_6\big),\\
\mu'_4 &= \epsilon_3 + \epsilon_4 + \epsilon_5 + (\epsilon_8 - \epsilon_7 - \epsilon_6),\\
\mu'_5 &= \epsilon_4 + \epsilon_5 + \frac{2}{3}(\epsilon_8 - \epsilon_7 - \epsilon_6),\\
\mu'_6 &= \epsilon_5 + \frac{1}{3}(\epsilon_8 - \epsilon_7 - \epsilon_6).
\end{aligned}
$$

  Moreover,

$$
\big(\mu'_3 - \mu'_1|\mu'_1\big) > 0, \quad \big(\mu'_5 - \mu'_6|\mu'_6\big) > 0, \quad \big(\mu'_4 - \mu'_2|\mu'_2\big) > 0,
$$

  so if $\varphi^* = id$, then (2.1) implies that $i = 1, 2$ or 6, so the symmetry of the diagram shows that after reordering, it is enough to consider the cases of $i = 1$ or $i = 2$. On the other hand, if $\varphi^* \neq id$, then $i = 2$ is the only possibility.

- For $F_4$, consider the ordering of the roots given by




  Here

$$
\begin{aligned}
\alpha_1 &= \epsilon_2 - \epsilon_3,\\
\alpha_2 &= \epsilon_3 - \epsilon_4,\\
\alpha_3 &= \epsilon_4,\\
\alpha_4 &= \frac{1}{2}(\epsilon_1 - \epsilon_2 - \epsilon_3 - \epsilon_4),
\end{aligned}
$$

for a suitable orthonormal basis. Hence,

$$\begin{aligned} \mu_1' &= \epsilon_1 + \epsilon_2, \\ \mu_2' &= 2\epsilon_1 + \epsilon_2 + \epsilon_3, \\ \mu_3' &= 3\epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4, \\ \mu_4' &= 2\epsilon_1, \end{aligned}$$

so

$$\big(\mu_2' - \mu_1' | \mu_1'\big) > 0, \ \big(\mu_3' - \mu_4' | \mu_4'\big) > 0,$$

and (2.1) imply that $i = 1$ or $i = 4$.

- For $G_2$, order the roots as follows:


Then $\alpha_1 = \epsilon_2 - \epsilon_1$, $\alpha_2 = \frac{1}{2}(\epsilon_1 - 2\epsilon_2 + \epsilon_3)$, where $\{\epsilon_1, \epsilon_2, \epsilon_3\}$ is an orthonormal basis of a three-dimensional inner vector space and $\Delta' = \{\alpha_1, \alpha_2\}$ generates a two-dimensional vector subspace. Then $\mu_1' = \epsilon_3 - \epsilon_1$ and $\mu_2' = \frac{1}{3}(-\epsilon_1 - \epsilon_2 + 2\epsilon_3)$, so $\big(\mu_1' - \mu_2' | \mu_2'\big) > 0$, and hence (2.1) forces $i = 2$.

**(xi)** The assertions on the third column in Tables A.1, A.2 follows by straightforward computations, similar to the ones used for the description of the exceptional simple Lie algebra of type $F_4$ in Chapter 2, Section § 8. (Some more information will be given in the next section.) The involutive automorphisms that appear in these tables for each type are all nonconjugate, since their fixed subalgebras are not isomorphic. □

## § 3. Simple real Lie algebras

What is left is to check that the information on the fourth column in Tables A.1, A.2 is correct.

First, because of item (ii) in Remark 1.12, the signature of the Killing form of the real form of a simple complex Lie algebra $S$ associated to an involutive automorphism $\theta \in \operatorname{Aut}_{\mathbb{C}} S$ is $\dim_{\mathbb{C}} S_{\bar{1}} - \dim_{\mathbb{C}} S_{\bar{0}}$, and this shows that the third column in Table A.2 determines completely the fourth. Thus, it is enough to deal with the classical cases. Here, only the type $A_n$ will be dealt with, leaving the other types as an exercise.

Let $S = \mathfrak{sl}_{n+1}(\mathbb{C})$ be the simple complex Lie algebra of type $A_n$. The special unitary Lie algebra

$$\mathfrak{su}_{n+1}(\mathbb{R}) = \{x \in \mathfrak{sl}_{n+1}(\mathbb{C}) : \bar{x}^t = -x\}$$

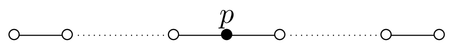
is a compact real subalgebra of $S$ (here the 'bar' denotes complex conjugation), as for any $x \in \mathfrak{su}_{n+1}(\mathbb{R})$,

$$\kappa(x, x) = 2(n+1)\operatorname{trace}(x^2) = -2(n+1)\operatorname{trace}(x\bar{x}^t) < 0$$

(see Equation (6.5) in Chapter 2). Let $\tau$ be the associated compact conjugation: $\tau(x) = -\bar{x}^t$. For any Vogan diagram, we must find an involutive automorphism $\varphi \in \operatorname{Aut}_{\mathbb{C}} \mathfrak{sl}_{n+1}(\mathbb{C})$ associated to it and that commutes with $\tau$. Then $\sigma = \varphi\tau$ is the conjugation associated to the corresponding real form.

(a) For $\varphi = id$, $\sigma = \tau$ and the real form is $S^\sigma = \mathfrak{su}_{n+1}(\mathbb{R})$.

(b) Let $a_p = \operatorname{diag}(1, \dots, 1, -1, \dots, -1)$ be the diagonal matrix with $p$ 1's and $(n + 1 - p)$ $-1$'s. Then $a_p^2 = I_{n+1}$ (the identity matrix). The involutive automorphism $\varphi_p : x \mapsto a_p x a_p = a_p x a_p^{-1}$ of $\mathfrak{sl}_{n+1}(\mathbb{C})$ commutes with $\tau$, its fixed subalgebra is formed by the block diagonal matrices with two blocks of size $p$ and $n+1-p$, so $S^{\varphi_p} \cong \mathfrak{sl}_p(\mathbb{C}) \oplus \mathfrak{sl}_{n+1-p}(\mathbb{C}) \oplus Z$, where $Z$ is a one-dimensional center. Moreover, the usual Cartan subalgebra $H$ of the diagonal matrices in $S$ (see Equation (6.4) in Chapter 2) contains a Cartan subalgebra of the fixed part and is invariant under $\varphi_p$. Here $x_i = E_{i,i+1}$ (the matrix with a 1 on the $(i, i+1)$ position and 0's elsewhere). Then $\varphi_p(x_p) = -x_p$, while $\varphi_p(x_j) = x_j$ for $j \neq p$, so the associated Vogan diagram is




Now, with $\sigma_p = \varphi_p \tau$, the associated real form is

$$\{x \in \mathfrak{sl}_{n+1}(\mathbb{C}) : x^t a_p + a_p \bar{x} = 0\} = \mathfrak{su}_{p,n+1-p}(\mathbb{R}).$$

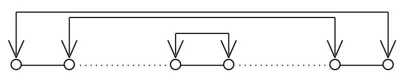
(c) With $n = 2r$, consider the symmetric matrix of order $n + 1 = 2r + 1$

$$b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_r \\ 0 & I_r & 0 \end{pmatrix},$$

which satisfies $b^2 = I_n$, and the involutive automorphism $\varphi_b : x \mapsto -bx^t b$, which commutes with $\tau$. The fixed subalgebra by $\varphi_b$ is precisely $\mathfrak{so}_{2r+1}(\mathbb{C})$. Again $\varphi_b$ preserves the by now usual Cartan subalgebra $H$. With the description of the root system in Chapter 2, (6.4), it follows that $\varphi_b^*(\epsilon_1) = \epsilon_1 \circ \varphi_b = -\epsilon_1$, while $\varphi^*(\epsilon_i) = -\epsilon_{r+i}$, for $i = 2, \dots, r+1$. Take the system of simple roots

$$\Delta' = \{\epsilon_2 - \epsilon_3, \epsilon_3 - \epsilon_4, \dots, \epsilon_r - \epsilon_{r+1}, \epsilon_{r+1} - \epsilon_1, \epsilon_1 - \epsilon_{2r+1}, \epsilon_{2r+1} - \epsilon_{2r}, \dots, \epsilon_{r+3} - \epsilon_{r+2}\}$$

which is invariant under $\varphi^*$ and shows that the associated Vogan diagram is



Finally, consider the regular matrix

$$a = \begin{pmatrix} 1 & 0 & 0 \\ 0 & I_r & I_r \\ 0 & iI_r & -iI_r \end{pmatrix},$$

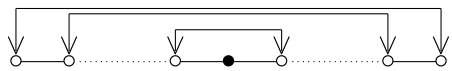
which satisfies that $b = a^{-1}\bar{a} = \bar{a}^{-1}a$, and the associated conjugation $\sigma_b = \varphi_b \tau$. Its real form is

$$\begin{aligned} S^{\sigma_b} &= \{x \in \mathfrak{sl}_{n+1}(\mathbb{C}) : b\bar{x}b = x\} \\ &= \{x \in \mathfrak{sl}_{n+1}(\mathbb{C}) : \overline{axa^{-1}} = axa^{-1}\} \\ &= a^{-1}\mathfrak{sl}_{n+1}(\mathbb{R})a \cong \mathfrak{sl}_{n+1}(\mathbb{R}). \end{aligned}$$

(d) In the same vein, with $n = 2r - 1$, consider the symmetric matrix $d = \left(\begin{smallmatrix} 0 & I_r \\ I_r & 0 \end{smallmatrix}\right)$ and the involutive automorphism $\varphi_d : x \mapsto -dx^t d = -dx^t d^{-1}$. Here the fixed subalgebra is $\mathfrak{so}_{2r}(\mathbb{C})$, and $\varphi_d^*(\epsilon_i) = -\epsilon_{r+i}$ for $i = 1, \ldots, r$. A suitable system of simple roots is

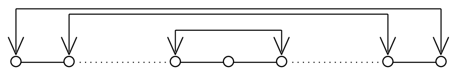
$$\Delta' = \{\epsilon_r - \epsilon_{r-1}, \epsilon_{r-1} - \epsilon_{r-2}, \ldots, \epsilon_2 - \epsilon_1, \epsilon_1 - \epsilon_{r+1}, \epsilon_{r+1} - \epsilon_{r+2}, \ldots, \epsilon_{2r-1} - \epsilon_{2r}\}.$$

The only root in $\Delta'$ fixed by $\varphi_d$ is $\epsilon_1 - \epsilon_{r+1}$ and $\varphi_d(E_{1,r+1}) = -E_{1,r+1}$, which shows that the associated Vogan diagram is



As before, with $\sigma_d = \varphi_d \tau$, one gets the real form $S^{\sigma_d} \cong \mathfrak{sl}_{n+1}(\mathbb{R})$.

(e) Finally, with $n = 2r - 1$, consider the skew-symmetric matrix $c = \left(\begin{smallmatrix} 0 & I_r \\ -I_r & 0 \end{smallmatrix}\right)$ and the involutive automorphism $\varphi_c : x \mapsto cx^t c = -cx^t c^{-1}$. Here the fixed subalgebra is $\mathfrak{sp}_{2r}(\mathbb{C})$, and $\varphi_c^*(\epsilon_i) = -\epsilon_{r+i}$ for $i = 1, \ldots, r$. The same $\Delta'$ of the previous item works here but $\varphi_c(E_{1,r+1}) = E_{1,r+1}$, which shows that the associated Vogan diagram is



With $\sigma_c = \varphi_c \tau$, one gets the real form

$$\begin{aligned} S^{\sigma_c} &= \{x \in \mathfrak{sl}_{n+1}(\mathbb{C}) : -c\bar{x}c = x\} \\ &= \left\{ \left(\begin{smallmatrix} p & q \\ -\bar{q} & \bar{p} \end{smallmatrix}\right) : p, q \in \mathfrak{gl}_r(\mathbb{C}) \text{ and } \operatorname{Re}\big(\operatorname{trace}(p)\big) = 0 \right\} \\ &\cong \{p + jq \in \mathfrak{gl}_r(\mathbb{H}) : \operatorname{Re}\big(\operatorname{trace}(p)\big) = 0\} = \mathfrak{sl}_r(\mathbb{H}), \end{aligned}$$

where $j \in \mathbb{H}$ satisfies $j^2 = -1$ and $ij = -ji$ and Re denotes the real part.