# Introduction to Lie Algebras

J.I. Hall

8 December 2017

# Contents

# Part I

# Introduction

Chapter 1

# Introduction

## 1.1 Algebras

Let $\mathbb{K}$ be a field. A $\mathbb{K}$-*algebra* $({}_{\mathbb{K}}A, \mu)$ is a (left) $\mathbb{K}$-space $A$ equipped with a bilinear multiplication. That is, there is a $\mathbb{K}$-space homomorphism *multiplication* $\mu\colon A \otimes_{\mathbb{K}} A \longrightarrow A$. We often write $ab$ in place of $\mu(a \otimes b)$. Also we may write $A$ or $(A, \mu)$ in place of $({}_{\mathbb{K}}A, \mu)$ when the remaining pieces should be evident from the context.

If $A$ is a $\mathbb{K}$-algebra, then its *opposite algebra* $A^{\mathrm{op}}$ has the same underlying vector space but its multiplication $\mu^{\mathrm{op}}$ is given by $\mu^{\mathrm{op}}(x \otimes y) = \mu(y \otimes x)$.

**(1.1).** LEMMA. *The map $\mu\colon A \otimes_{\mathbb{K}} A \mapsto A$ is a $\mathbb{K}$-algebra multiplication if and only if the* adjoint map

$$\mathrm{ad}\colon x \mapsto \mathrm{ad}_x \quad \textit{given by} \quad \mathrm{ad}_x\, a = xa$$

*is a $\mathbb{K}$-vector space homomorphism of $A$ into* $\mathrm{End}_{\mathbb{K}}(A)$ *.* □

If $\mathcal{V} = \{\, v_i \mid i \in I \,\}$ is a $\mathbb{K}$-basis of $A$, then the algebra is completely described by the associated *multiplication coefficients* or *structure constants* $c_{ij}^k \in \mathbb{K}$ given by

$$v_i v_j = \sum_{k \in I} c_{ij}^k v_k \,,$$

for all $i, j$.

We may naturally *extend scalars* from $\mathbb{K}$ to any extension field $\mathbb{E}$. Indeed $\mathbb{E} \otimes_{\mathbb{K}} A$ has a natural $\mathbb{E}$-algebra structure with the same multiplication coefficients for the basis $\mathcal{V}$.

Going the other direction is a little more subtle. If the $\mathbb{E}$-algebra $B$ has a basis $\mathcal{V}$ for which all the $c_{ij}^k$ belong to $\mathbb{K}$, then the $\mathbb{K}$-span of the basis is a $\mathbb{K}$-algebra $A$ for which $B = \mathbb{E} \otimes_{\mathbb{K}} A$. In that case we say that $A$ is a $\mathbb{K}$-*form* of the algebra $B$. In many cases the $\mathbb{E}$-algebra $B$ has several pairwise nonisomorphic $\mathbb{K}$-forms.

Various generalizations of the above are available and often helpful. The extension field $\mathbb{E}$ of $\mathbb{K}$ is itself a special sort of $\mathbb{K}$-algebra. If $C$ is an arbitrary $\mathbb{K}$-algebra, then $C \otimes_{\mathbb{K}} A$ is a $\mathbb{K}$-algebra, with opposite algebra $A \otimes_{\mathbb{K}} C$. The relevant multiplication is $\mu = \mu_C \otimes \mu_A$:

$$\mu((c_1 \otimes a_1) \otimes (c_2 \otimes a_2)) = \mu_C(c_1 \otimes c_2) \otimes \mu_A(a_1 \otimes a_2)\,.$$

We might also wish to consider $R$-algebras for other rings $R$ with identity. For the tensor product to work reasonably, $R$ should be commutative. A middle ground would require $R$ to be an integral domain, although even in that case we must decide whether or not we wish algebras to be free as $R$-module.

Of primary interest to us is the case $R = \mathbb{Z}$. A $\mathbb{Z}$-*algebra* is a free abelian group (that is, *lattice*) $L = \bigoplus_{i \in I} \mathbb{Z}v_i$ provided with a bilinear multiplication $\mu_{\mathbb{Z}}$ which is therefore completely determined by the integral multiplication coefficients $c_{ij}^k$. From this we can construct $\mathbb{K}$-algebras $L_{\mathbb{K}} = \mathbb{K} \otimes_{\mathbb{Z}} L$ for any field $\mathbb{K}$, indeed for any $\mathbb{K}$-algebra. For instance $C \otimes_{\mathbb{Z}} \mathrm{Mat}_n(\mathbb{Z})$ is the $\mathbb{K}$-algebra $\mathrm{Mat}_n(C)$ of all $n \times n$ matrices with entries from the $\mathbb{K}$-algebra $C$.

Suppose for the basis $\mathcal{V}$ of the $\mathbb{K}$-algebra $A$ all the $c_{ij}^k$ are integers—that is, belong to the subring of $\mathbb{K}$ generated by 1. Then the $\mathbb{Z}$-algebra $L = \bigoplus_{i \in I} \mathbb{Z}v_i$ with these multiplication coefficients can be viewed as a $\mathbb{Z}$-*form* of $A$ (although we only have its quotient by $\mathrm{char}(\mathbb{K})$ as a subring of $A$). The original $\mathbb{K}$-algebra $A$ is then isomorphic to $L_{\mathbb{K}}$.

## 1.2 Types of algebras

As $\dim_{\mathbb{K}}(A \otimes_{\mathbb{K}} A) \geq \dim_{\mathbb{K}}(A)$, every $\mathbb{K}$-space admits $\mathbb{K}$-algebras. We focus on those with some sort of interesting additional structure. Examples are associative algebras, Jordan algebras, alternative algebras, composition algebras, Hopf algebras, and Lie algebras—these last being the primary focus of our study. (All the others will be discussed at least briefly.)

In most cases these algebra types naturally form subcategories of the additive ${}_{\mathbb{K}}\mathsf{Alg}$ of $\mathbb{K}$-algebras, the maps $\varphi$ of $\mathrm{Hom}_{{}_{\mathbb{K}}\mathsf{Alg}}(A, B)$ being those linear transformations $\varphi \in \mathrm{Hom}_{\mathbb{K}}(A, B)$ with $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in A$. As the category ${}_{\mathbb{K}}\mathsf{Alg}$ is additive, each morphism has a kernel and image, which are defined as usual and enjoy the usual properties.

A subcategory will often be defined initially as belonging to a particular *variety* of $\mathbb{K}$-algebras. For instance, the *associative* $\mathbb{K}$-*algebras* are precisely those $\mathbb{K}$-algebras satisfying the identical relation

$$(xy)z = x(yz)\,.$$

Alternatively, the associative $\mathbb{K}$-algebras are those whose multiplication map $\mu$ satisfies

$$\mu(\mu(x \otimes y) \otimes z) = \mu(x \otimes \mu(y \otimes z))\,.$$

As the defining identical relation is equivalent to its reverse $(zy)x = z(yx)$, the opposite of an associative algebra is also associative.

Similarly, the subcategory of *alternative* $\mathbb{K}$-*algebras* is the variety of $\mathbb{K}$-algebras given by the weak associative laws

$$x(xy) = (xx)y \quad \text{and} \quad y(xx) = (yx)x\,.$$

The opposite of an alternative algebra is also alternative.

Varietal algebras like these have nice local properties:

(i) A $\mathbb{K}$-algebra is associative if and only if all its 3-generator subalgebras are associative.

(ii) A $\mathbb{K}$-algebra is alternative if and only if all its 2-generator subalgebras are alternative.

The associative identity is linear in that each variable appears at most once in each term, while the alternative identity is not, since $x$ appears twice in each term. The linearity of an identity implies that it only need be checked on a basis of the algebra to ensure that it is valid throughout the algebra. That is, there are appropriate identities among the various $c_{ij}^k$ that are equivalent to the algebra being associative. (Exercise: find them.) This implies the (admittedly unsurprising) fact that extending the scalars of an associative algebra produces an associative algebra. It is also true that extending the scalars of an alternative algebra produces another alternative algebra, but that needs some discussion since the basic identity is not linear. (Exercise.)

The basic model for an associative algebra is $\mathrm{End}_{\mathbb{K}}(V)$ for some $\mathbb{K}$-space $V$. Indeed, most associative algebras (including all with an identity) are isomorphic to subalgebras of various $\mathrm{End}_{\mathbb{K}}(V)$. (See Proposition (1.3).) For finite dimensional $V$ we often think in matrix terms by choosing a basis for $V$ and then using that basis to define an isomorphism of $\mathrm{End}_{\mathbb{K}}(V)$ with $\mathrm{Mat}_n(\mathbb{K})$ for $n = \dim_{\mathbb{K}}(V)$.

Of course, every associative algebra is alternative, but we now construct the most famous models for alternative but nonassociative algebras. If we start with $\mathbb{K} = \mathbb{R}$, then we have the familiar construction of the complex numbers as $2 \times 2$ matrices: for $a, b \in \mathbb{K}$ we set

$$(a, b) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

with multiplication given by

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac - bd & ad + bc \\ -bc - ad & -bd + ac \end{pmatrix}$$

and conjugation given by

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^{-} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

As $\mathbb{R}$ is commutative and conjugation is trivial on $\mathbb{R}$, these can be rewritten:

*For $a, b \in \mathbb{K}$ and $a \mapsto \bar{a}$ an antiautomorphism of $\mathbb{K}$, we set*

$$(a,b)_{\mathbb{K}} = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix},$$

*with*

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} c & d \\ -\bar{d} & \bar{c} \end{pmatrix} = \begin{pmatrix} ac - \bar{d}b & da + b\bar{c} \\ -c\bar{b} - \bar{a}\bar{d} & -\bar{b}d + \bar{c}\bar{a} \end{pmatrix}$$

*and*

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}^{-} = \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix}.$$

This then gives us the complex numbers $\mathbb{C}$ as the collection of all pairs $(a,b)_{\mathbb{R}}$ of real numbers. Feeding the complex numbers back into the machine produces Hamilton's *quaternions* $\mathbb{H}$ as all pairs $(a,b)_{\mathbb{C}}$ with the multiplication and the conjugation antiautomorphism described. As $\mathbb{C}$ is commutative the quaternions are associative, but they are no longer commutative.

Finally with $\mathbb{K} = \mathbb{H}$, the resulting $\mathbb{O}$ of all pairs $(a,b)_{\mathbb{H}}$ is the *octonions* of Cayley and Graves. The octonions are indeed alternative but not associative, although this requires checking. Again conjugation is an antiautomorphism.

In each case, the $2 \times 2$ "scalar matrices" are only those with $b = 0$ and $a = \bar{a} \in \mathbb{R}$, so we have constructed $\mathbb{R}$-algebras with respective dimensions $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, $\dim_{\mathbb{R}}(\mathbb{H}) = 4$, $\dim_{\mathbb{R}}(\mathbb{O}) = 8$.

A *quadratic form* on the $\mathbb{K}$-space $V$ is a map $q \colon V \longrightarrow \mathbb{K}$ for which

$$q(ax) = a^2 q(x)$$

whenever $a \in \mathbb{K}$ and $x \in V$ and also the associated map $b \colon V \times V \longrightarrow K$ given by polarization

$$b(x,y) = q(x+y) - q(x) - q(y)$$

is a nondegenerate bilinear form. (See Appendix A for a brief discussion of quadratic and bilinear forms.)

The $\mathbb{R}$-algebras $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$ furnish examples of composition $\mathbb{R}$-algebras. A *composition algebra* is a $\mathbb{K}$-algebra $A$ with multiplicative identity, admitting a nondegenerate quadratic form $\delta \colon A \longrightarrow \mathbb{K}$ that is multiplicative:

$$\delta(x)\delta(y) = \delta(xy),$$

for all $x, y \in A$. The codimension 1 subspace $1^{\perp}$ consists of the *pure imaginary* elements of $A$, and (in characteristic not 2) the conjugation map $\overline{a1 + b} = a1 - b$, for $b \in 1^{\perp}$, is an antiautomorphism of $A$ whose fixed point subspace is $\mathbb{K}1$.

In the above $\mathbb{R}$-algebras the form $\delta$ is given by $\delta(x)1 = x\bar{x}$:

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix} = a\bar{a} + \bar{b}b \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In $\mathbb{O}$ specifically, for $a,b,c,d,e,f,g,h \in \mathbb{R}$, we find

$$\delta(((a,b)_{\mathbb{R}},(c,d)_{\mathbb{R}})_{\mathbb{C}},((e,f)_{\mathbb{R}},(g,h)_{\mathbb{R}})_{\mathbb{C}})_{\mathbb{H}} = \delta(a,b,c,d,e,f,g,h) = \\ = a^2+b^2+c^2+d^2+e^2+f^2+g^2+h^2\,.$$

Thus in $\mathbb{O}$ (and so its subalgebras $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$) all nonzero vectors have nonzero norm.

An immediate consequence of the composition law is that an invertible element of $A$ must have nonzero norm. As $\delta(x)1 = x\bar{x}$ in composition algebras, the converse is also true. Therefore if 0 is the only element of the composition algebra $A$ with norm 0, then all nonzero elements are invertible and $A$ is a *division algebra*. Prime examples are the division composition $\mathbb{R}$-algebras $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$. The following remarkable theorem of Hurwitz shows that this situation is typical

**(1.2).** Theorem. (Hurwitz' Theorem) *If $A$ is a composition algebra over $\mathbb{K}$, then $\dim_{\mathbb{K}}(A)$ is 1, 2, 4, or 8.* □□

If the composition $\mathbb{K}$-algebra $A$ is not a division algebra, then it is called *split*. It turns out that a split composition algebra over $\mathbb{K}$ is uniquely determined up to isomorphism by its dimension. In dimension 1, the algebra is $\mathbb{K}$ itself, always a division algebra. In dimension 4, a split composition $\mathbb{K}$-algebra is always $\mathrm{Mat}_2(\mathbb{K})$ with $\delta = \det$, and the diagonal matrices provide a split subalgebra of dimension 2.

Composition algebras of dimension 8 are called *octonion algebras*. The original is the real division algebra $\mathbb{O}$ presented above and due to Graves (1843, unpublished) and Cayley (1845) [SpV00, p. 23].

A split octonion algebra $\mathbb{O}^{\mathrm{sp}}(\mathbb{K})$ over any field $\mathbb{K}$ is provided by *Zorn's vector matrices* [Zor31]

$$m = \begin{pmatrix} a & \vec{b} \\ \vec{c} & d \end{pmatrix}$$

with $a,d \in \mathbb{K}$ and $\vec{b},\vec{c} \in \mathbb{K}^3$. Multiplication is given by

$$\begin{pmatrix} a & \vec{b} \\ \vec{c} & d \end{pmatrix}\begin{pmatrix} x & \vec{y} \\ \vec{z} & w \end{pmatrix} = \begin{pmatrix} ax+\vec{b}\cdot\vec{z} & a\vec{y}+w\vec{b} \\ x\vec{c}+d\vec{z} & \vec{c}\cdot\vec{y}+dw \end{pmatrix} + \begin{pmatrix} 0 & \vec{c}\times\vec{z} \\ -\vec{b}\times\vec{y} & 0 \end{pmatrix}$$

using the standard dot (inner) and cross (outer, exterior, vector) products of 3-vectors. The associated norm is

$$\delta(m) = ad - \vec{b}\cdot\vec{c}\,.$$

For any $\vec{v}$ with $\vec{v}\cdot\vec{v} = k \neq 0$ the subalgebra of all

$$m = \begin{pmatrix} a & b\vec{v} \\ ck^{-1}\vec{v} & d \end{pmatrix}$$

is a copy of the split quaternion algebra $\mathrm{Mat}_2(F)$ with norm the usual determinant.

Zorn (and others) gave a slightly different version of the vector matrices, replacing our entry $\vec{c}$ with its negative. This gives the more symmetrical norm form $\delta(m) = ad + \vec{b} \cdot \vec{c}$ but makes the connection with standard matrix multiplication and determinants less clear.

Extending coefficients in a composition algebra produces a composition algebra (although this is more than an exercise). For every composition $\mathbb{K}$-algebra $O$, there is an extension $\mathbb{E}$ of degree at most 2 over $\mathbb{K}$ with $\mathbb{E} \otimes_{\mathbb{K}} O$ a split composition $\mathbb{E}$-algebra. In particular every composition algebra over algebraically closed $\mathbb{E}$ is split and so unique up to isomorphism. The split algebra over $\mathbb{C}$ (for instance given by Zorn's vector matrices) has two isomorphism classes of $\mathbb{R}$-forms—the class of the split algebra $\mathbb{O}^{\mathrm{sp}}(\mathbb{R})$ and that of the Cayley-Graves division algebra $\mathbb{O}$.

## 1.3 Jordan algebras

As mentioned above, the basic models for associative algebras are the endomorphism algebras $\mathrm{End}_{\mathbb{K}}(V)$ for some $\mathbb{K}$-space $V$ and the related matrix algebras $\mathrm{Mat}_n(\mathbb{K})$. While Jordan and Lie algebras both have abstract varietal definitions (given below for Jordan algebras and in the next section for Lie algebras), they are first seen in canonical models coming from $\mathrm{End}_{\mathbb{K}}(V)$.

We start with the observation that any pure tensor from $V \otimes V$ is the sum of its symmetric and skew-symmetric parts:

$$v \otimes w = \frac{1}{2}(v \otimes w + w \otimes v) + \frac{1}{2}(v \otimes w - w \otimes v).$$

In 1933 P. Jordan [JvNW34] initiated the study of the $\mathbb{K}$-algebra $A^+ = (A, \mu^+) = (A, \circ)$ that is the associative $\mathbb{K}$-algebra $A$ equipped with the *Jordan product*

$$\mu^+(x \otimes y) = x \circ y = \frac{1}{2}(xy + yx)\,.$$

This requires, of course, that the characteristic of the field $\mathbb{K}$ not be 2. We could also consider the algebra without the factor of $\frac{1}{2}$, but we keep it for various reasons—in particular $x \circ x = \frac{1}{2}(xx + xx) = xx = x^2$ and $1 \circ x = \frac{1}{2}(1x + x1) = x$.

The model for all Jordan algebras is then $\mathrm{End}^+_{\mathbb{K}}(V)$, the vector space of all $\mathbb{K}$-endomorphisms of $V$ equipped with the Jordan product.

Clearly the algebra $\mathrm{End}^+_{\mathbb{K}}(V)$ is commutative. Not so obvious is the fact that we also have the identity

$$(x \circ x) \circ (y \circ x) = ((x \circ x) \circ y) \circ x\,,$$

for all $x, y \in \mathrm{End}^+_{\mathbb{K}}(V)$. (Exercise.)

We are led to the general, varietal definition: the $\mathbb{K}$-algebra $A$ is a *Jordan algebra* if it is commutative and satisfies the identical relation

$$x^2(yx) = (x^2y)x\,.$$

The canonical models are $\mathrm{End}^+_{\mathbb{K}}(V)$ and so also $\mathrm{Mat}^+_n(\mathbb{K})$ (in finite dimension).

Any subspace of $\mathrm{End}^+_{\mathbb{K}}(V)$ that is closed under the Jordan product is certainly a Jordan subalgebra. Especially if $\tau$ is an automorphism of $\mathrm{End}_{\mathbb{K}}(V)$, then its fixed-point-space is certainly closed under the Jordan product and so is a subalgebra. More subtly, if $\tau$ is an antiautomorphism of $\mathrm{End}_{\mathbb{K}}(V)$, then it induces an automorphism of $\mathrm{End}^+_{\mathbb{K}}(V)$ whose fixed points are again a Jordan subalgebra.

For instance, in the $\mathbb{K}$-algebra $\mathrm{Mat}_n(\mathbb{K})$ the transpose map is an antiautomorphism, so the symmetric matrices from $\mathrm{Mat}_n(\mathbb{K})$ form a Jordan subalgebra of $\mathrm{Mat}^+_n(\mathbb{K})$. More generally, if $A$ is a $\mathbb{K}$-algebra with an antiautomorphism $a \mapsto \bar{a}$ fixing $\mathbb{K}$, then we can try the same trick with the $\mathbb{K}$-algebra $\mathrm{Mat}_n(A)$. The transpose-conjugate map

$$\bar{\tau} \colon (a_{ij}) \mapsto (\bar{a}_{ji})$$

is then an antiautomorphism of $\mathrm{Mat}_n(A)$ (Exercise), and so the associated fixed space of Hermitian matrices

$$\mathrm{H}_n(A) = \{\, M \in \mathrm{Mat}_n(A) \mid M = M^{\bar{\tau}} \,\}$$

is closed under the Jordan product

$$M \circ N = \frac{1}{2}(MN + NM)\,.$$

If $A$ is associative then we have a Jordan algebra. Indeed this with $A = \mathbb{C}$ and $\mathbb{K} = \mathbb{R}$ was the original motivation for the physicist Jordan: in quantum mechanics the observables for the Hilbert space $\mathbb{C}^n$ are characterized by the hermitian matrices $\mathrm{H}_n(\mathbb{C})$, an $\mathbb{R}$-space that is not closed under the standard matrix product but is a real Jordan algebra under the Jordan product.

When $A$ is not associative, there is no reason to assume that this gives $\mathrm{H}_n(A)$ the structure of an (abstract) Jordan algebra. But if we choose $A$ to be an octonion algebra over $\mathbb{K}$ and let $n \le 3$, then this is in fact the case. (Recall that the alternative law is a weak version of the associative law, so this is not completely unreasonable.)

For the octonion $\mathbb{K}$-algebra $O$, the Jordan algebra $\mathrm{H}_3(O)$ is called an *Albert algebra*. Each matrix of $\mathrm{H}_3(O)$ has the shape

$$\begin{pmatrix} a & \alpha & \beta \\ \bar{\alpha} & b & \gamma \\ \bar{\beta} & \bar{\gamma} & c \end{pmatrix}$$

with $a, b, c \in \mathbb{K}$ (the fixed field of conjugation in $O$) and $\alpha, \beta, \gamma \in O$. Thus the $\mathbb{K}$-dimension of the Albert algebra $\mathrm{H}_3(O)$ is $3 + 3 \times 8 = 27$.

## 1.4 Lie algebras and linear representation

In the previous section we only discussed the symmetric part of the tensor decomposition displayed at the beginning of the section. But even at the time of Jordan, the corresponding skew part had been studied for years, starting with the Norwegian Sophus Lie and soon followed by Killing and Cartan (see [Bo01] and [Haw00]). If $A$ is an associative algebra, then we define a skew algebra $A^- = (A, \mu^-) = (A, [\cdot,\cdot])$ by furnishing $A$ with the multiplication

$$\mu^-(x \otimes y) = [x,y] = xy - yx\,.$$

(Note that the scaling factor $\frac{1}{2}$ does not appear.) The algebras $A^-$ and in particular $\mathrm{End}^-_{\mathbb{K}}(V)$ and $\mathrm{Mat}^-_n(\mathbb{K})$ are the canonical models for Lie algebras over $\mathbb{K}$.

In a given category, a representation of an object $M$ is loosely a morphism of $M$ into one of the canonical examples from the category. So a *linear representation* of a group $M$ is a homomorphism from $M$ to some $\mathrm{GL}_{\mathbb{K}}(V)$. With this in mind, we will say that a *linear representation* of an associative algebra $A$, a Jordan algebra $J$, and a Lie algebra $L$ (all over $\mathbb{K}$), respectively, is a $\mathbb{K}$-algebra homomorphism $\varphi$ belonging to, respectively, some

$$\mathrm{Hom}_{\mathbb{K}\mathsf{Alg}}(A, \mathrm{End}_{\mathbb{K}}(V))\,, \quad \mathrm{Hom}_{\mathbb{K}\mathsf{Alg}}(J, \mathrm{End}^+_{\mathbb{K}}(V))\,, \quad \mathrm{Hom}_{\mathbb{K}\mathsf{Alg}}(L, \mathrm{End}^-_{\mathbb{K}}(V))\,,$$

which in the finite dimensional case can be viewed as

$$\mathrm{Hom}_{\mathbb{K}\mathsf{Alg}}(A, \mathrm{Mat}_n(\mathbb{K}))\,, \quad \mathrm{Hom}_{\mathbb{K}\mathsf{Alg}}(J, \mathrm{Mat}^+_n(\mathbb{K}))\,, \quad \mathrm{Hom}_{\mathbb{K}\mathsf{Alg}}(L, \mathrm{Mat}^-_n(\mathbb{K}))\,.$$

The corresponding image of $\varphi$ is then a *linear associative algebra*, *linear Jordan algebra*, or *linear Lie algebra*, respectively. The representation is *faithful* if its kernel is 0. The underlying space $V$ or $\mathbb{K}^n$ is then an *$A$-module* which *carries* the representation and upon which the algebra *acts*.

It turns out that in each of these categories, many of the important examples are linear. For instance

**(1.3).** PROPOSITION. *Every associative algebra with a multiplicative identity element is isomorphic to a linear associative algebra.*

PROOF. Let $A$ be an associative algebra. For each $x \in A$, consider the map $\mathrm{ad}\colon A \longrightarrow \mathrm{End}_{\mathbb{K}}(A)$ of Lemma (1.1), given by $x \mapsto \mathrm{ad}_x$ where $\mathrm{ad}_x\, a = xa$ as before. That lemma states that ad is a vector space homomorphism.

Thus we need to check that multiplication is respected. But the associative identity

$$(xy)a = x(ya)$$

can be restated as

$$\mathrm{ad}_{xy}\, a = \mathrm{ad}_x\, \mathrm{ad}_y\, a\,,$$

for all $x, y, a \in A$. Hence $\mathrm{ad}_{xy} = \mathrm{ad}_x\, \mathrm{ad}_y$ as desired.

The kernel of ad consists of those $x$ with $xa = 0$ for all $a \in A$. In particular, the kernel is trivial if $A$ contains an identity element. □

It is clear from the proof that the multiplicative identity plays only a small role—the result should and does hold in greater generality. But for us the main message is that the adjoint map is a representation of every associative algebra. The proposition should be compared with Cayley's Theorem which proves that every group is (isomorphic to) a faithful permutation group by looking at the regular representation, which is the permutation version of adjoint action.

What about Jordan and Lie representation? Of course we still have not defined general Lie algebras, but we certainly want to include all the subalgebras of $\mathrm{End}^-_{\mathbb{K}}(V)$ and $\mathrm{Mat}^-_n(\mathbb{K})$.

As above, the multiplication map $\mu$ of an arbitrary Lie algebra $A = (A, [\cdot,\cdot])$ will be written as a bracket, in deference to the commutator product in an associative algebra:

$$\mu(x \otimes y) = [x, y]\,.$$

In the linear Lie algebras $\mathrm{End}^-_{\mathbb{K}}(V)$ and $\mathrm{Mat}^-_n(\mathbb{K})$ we always have

$$[x, x] = xx - xx = 0\,,$$

so we require that an abstract Lie algebra satisfy the *null identical relation*

$$[x, x] = 0\,.$$

This identity is not linear, but we may "linearize" it by setting $x = y + z$. We then find

$$0 = [y + z, y + z] = [y, y] + [y, z] + [z, y] + [z, z] = [y, z] + [z, y]\,,$$

giving as an immediate consequence the linear *skew identical relation*

$$[y, z] = -[z, y]\,.$$

If $\mathrm{char}\mathbb{K} \neq 2$, these two identities are equivalent. (This is typical of linearized identities: they are equivalent to the original except where neutralized by the characteristic.)

Our experience with groups and associative algebras tells us that having adjoint representations available is of great benefit, so we make an initial hopeful definition:

> *A* Lie algebra *is an algebra* $({}_{\mathbb{K}}L, [\cdot,\cdot])$ *in which all squares* $[x, x]$ *are* $0$ *and for which the* $\mathbb{K}$*-homomorphism* $\mathrm{ad}\colon L \longrightarrow \mathrm{End}^-_{\mathbb{K}}(L)$ *is a representation of* $L$.

Are $\mathrm{End}^-_{\mathbb{K}}(V)$ and $\mathrm{Mat}^-_n(\mathbb{K})$ Lie algebras in this sense? Indeed they are:

$$\begin{aligned}\mathrm{ad}_x\,\mathrm{ad}_y\,a &= \mathrm{ad}_x(ya - ay)\\ &= x(ya - ay) - (ya - ay)x\\ &= xya - xay - yax + ayx\end{aligned}$$

hence

$$\begin{aligned}[\mathrm{ad}_x, \mathrm{ad}_y]a &= (\mathrm{ad}_x \mathrm{ad}_y - \mathrm{ad}_y \mathrm{ad}_x)a \\ &= (xya - xay - yax + ayx) - (yxa - yax - xay + axy) \\ &= (xya - axy) - (yxa - ayx) \\ &= [xy, a] - [yx, a] \\ &= [xy - yx, a] \\ &= \mathrm{ad}_{[x,y]}\, a\,.\end{aligned}$$

That is, $[\mathrm{ad}_x, \mathrm{ad}_y] = \mathrm{ad}_{[x,y]}$, as desired.

Let us now unravel the consequences of the identity $\mathrm{ad}_{[x,y]} = [\mathrm{ad}_x, \mathrm{ad}_y]$ for the algebra $(L, [\cdot,\cdot])$:

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

That is,

$$[[x,y],z] + [[y,z],x] + [[z,x],y] = 0\,.$$

We arrive at the standard definition of a Lie algebra:

> *A* Lie algebra *is an algebra* $({}_{\mathbb{K}}L, [\cdot,\cdot])$ *that satisfies the two identical relations:*
>
> (i) $[x,x] = 0$*;*
>
> (ii) (Jacobi Identity) $[[x,y],z] + [[y,z],x] + [[z,x],y] = 0$.

Negating the Jacobi Identity gives us the equivalent identity

$$[z,[x,y]] + [x,[y,z]] + [y,[z,x]] = 0\,.$$

In particular, the opposite of a Lie algebra is again a Lie algebra.[1]

The Jacobi Identity and the skew law $[y,z] = -[z,y]$ are both linear, and these serve to define Lie algebras if the characteristic is not 2. This is good enough to prove that tensor product field extensions of Lie algebras are still Lie algebras as long as the characteristic is not 2

In all characteristics the null law $[x,x] = 0$ admits a weaker form of linearity. Assume that we already know $[y,y] = 0$, $[z,z] = 0$, and $[y,z] = -[z,y]$. Then for all constants $a, b$ we have

$$\begin{aligned}[ay+bz, ay+bz] &= [ay,ay] + [ay,bz] + [bz,ay] + [bz,bz] \\ &= a^2[y,y] + ab([y,z] + [z,y]) + b^2[z,z] \\ &= 0 + 0 + 0 = 0\,.\end{aligned}$$

[1] Exercise: the map $x \mapsto -x$ is an isomorphism of the Lie algebra $L$ with its opposite algebra.

This, together with the linearity of the Jacobi Identity, gives

**(1.4).** PROPOSITION. *Let $L$ be Lie $\mathbb{K}$-algebra and $\mathbb{E}$ an extension field over $\mathbb{K}$. Then $\mathbb{E} \otimes_{\mathbb{K}} L$ is a Lie $\mathbb{E}$-algebra.* □

Our discussion of representation and our ultimate definition of Lie algebras immediately give

**(1.5).** THEOREM. *For any Lie $\mathbb{K}$-algebra $L$, the map* $\mathrm{ad}\colon L \longrightarrow \mathrm{End}^{-}_{\mathbb{K}}(L)$ *is a representation of $L$. The kernel of this representation is the* center *of $L$*

$$\mathrm{Z}(L) = \{\, z \in L \mid [z, a] = 0\,,\ \textit{for all } a \in L \,\}\,. \qquad \square$$

As was the case in Proposition (1.3) the small additional requirement that the center of $A$ be trivial gives an easy proof that $A$ has a faithful representation which has finite dimension provided $A$ does. Far deeper is:

**(1.6).** THEOREM.

(a) (PBW THEOREM) *Every Lie algebra has a faithful representation as a linear Lie algebra.*

(b) (ADO-IWASAWA THEOREM) *Every finite dimensional Lie algebra has a faithful representation as a finite dimensional linear Lie algebra.* □□

Both these theorems are difficult to prove, although we will return to the easier PBW Theorem later as Theorem (9.3). Notice that the Ado-Iwasawa Theorem is not an immediate consequence of PBW. Indeed the representation produced by the PBW Theorem is almost always a representation on an infinite dimensional space.

For Jordan algebras, the efforts of this section are largely a failure. In particular the adjoint action of a Jordan algebra $A$ on itself does not give a representation in $\mathrm{End}^{+}_{\mathbb{K}}(A)$. (Exercise.)

Jordan algebras that are (isomorphic to) linear Jordan algebras are usually called *special Jordan algebras*, while those that are not linear are the *exceptional Jordan algebras*.[2] A.A. Albert [Alb34] proved that the Albert algebras—the dimension 27 Jordan $\mathbb{K}$-algebras described in Section 1.3—are exceptional rather than special. Indeed Cohn [Coh54] proved that Albert algebras are not even quotients of special algebras. Results of Birkhoff imply that the category of images of special Jordan algebras is varietal and does not contain the Albert algebras, but it is unknown what additional identical relations suffice to define this category.

## 1.5 Problems

**(1.7).** PROBLEM.

[2]So, taking a page out of the Montessori book, there are exactly two types of Jordan algebras: those that are special and those that are exceptional.

(a) *Give two linear identities that characterize alternative $\mathbb{K}$-algebras when* $\mathrm{char}\mathbb{K} \neq 2$.

(b) *Let $A$ be an alternative $\mathbb{K}$-algebra and $\mathbb{E}$ an extension field over $\mathbb{K}$. Prove that $\mathbb{E} \otimes_{\mathbb{K}} A$ is an alternative $\mathbb{E}$-algebra.*

**(1.8).** PROBLEM. *Let $A$ be an associative $\mathbb{K}$-algebra with multiplicative identity $1$, where $\mathbb{K}$ is a field of characteristic not equal to $2$.*

(a) *Prove that in general the adjoint action of a Jordan algebra does not give a representation. Consider specifically the Jordan algebra $A^+ = (A, \circ)$ and its adjoint map* $\mathrm{ad}\colon A^+ \longrightarrow \mathrm{End}^+_{\mathbb{K}}(A)$ *where you can compare* $\mathrm{ad}_{a\circ a}$ *and* $\mathrm{ad}_a \circ \mathrm{ad}_a$.

(b) *Consider the two families of maps from $A$ to itself:*

$$L_a\colon x \mapsto a \circ x = \frac{1}{2}(ax + xa)$$

*and*

$$U_a\colon x \mapsto axa\,.$$

*Prove that the $\mathbb{K}$-subspace $V$ of $A$ with $1 \in V$ is invariant under all $L_a$, for $a \in V$, if and only if it is invariant under all $U_a$, for $a \in V$.*

HINT: *The two parts of this problem are not unrelated.*

REMARK. *Observe that saying $V$ is invariant under the $L_a$ is just the statement that $V$ is a Jordan subalgebra of* $\mathrm{End}^+_{\mathbb{K}}(A)$, *the map $L_a$ being the adjoint. Therefore the problems tells us that requiring $U_a$-invariance is another way of locating Jordan subalgebras, for instance the important and motivating spaces of hermitian matrices* $\mathrm{H}_n(\mathbb{C})$ *in* $\mathrm{Mat}_n(\mathbb{C})$.

*The crucial thing about $U_a$ is that division by $2$ is not needed. Therefore the maps $U_a$ and their properties can be, and are, used to extend the study of Jordan algebras to include characteristic $2$. The appropriate structures are called* quadratic Jordan algebras, *although some care must be taken as the "multiplication" $a \star x = U_a(x)$ is not bilinear. It is linear in its second variable but quadratic in its first variable; for instance $(\alpha a) \star x = \alpha^2(a \star x)$ for $\alpha \in \mathbb{K}$.*

# Chapter 2

# Examples of Lie algebras

We give many examples of Lie algebras $({}_{\mathbb{K}}L, [\cdot,\cdot])$. These also suggest the many contexts in which Lie algebras are to be found.

## 2.1 Abelian algebras

Any $\mathbb{K}$-vector space $V$ is a Lie $\mathbb{K}$-algebra when provided with the trivial product $[v, w] = 0$ for all $v, w \in V$. These are the *abelian Lie algebras*.

## 2.2 Generators and relations

As with groups and most other algebraic systems, one effective way of producing examples is by providing a generating set and a collection of relations among the generators. For a $\mathbb{K}$-algebra that would often be through supplying a basis $\mathcal{V} = \{ v_i \mid i \in I \}$ together with appropriate equations restricting the various associated $c_{ij}^k$.

For a Lie algebra, the Jacobi Identity is linear and leads to (Exercise) the equations:

$$\sum_k c_{ij}^k c_{kl}^m + c_{jl}^k c_{kl}^m + c_{li}^k c_{kj}^m = 0\,,$$

for all $i, j, l, m \in I$.

The law $[x, x] = 0$ gives the equations

$$c_{ii}^k = 0\,.$$

Since the null law is not linear, we also must include the consequences of its linearized skew law $[x, y] = -[y, x]$; so we also require

$$c_{ij}^k = -c_{ji}^k\,.$$

An algebra whose multiplication coefficients satisfy these three sets of equations is a Lie algebra. (Exercise.)

When presenting a Lie algebra it is usual to leave the non-Jacobi equations implicit, assuming without remark that the bracket multiplication is null and skew-symmetric.

For instance, we have the $\mathbb{K}$-algebra $L = \mathbb{K}h \oplus \mathbb{K}e \oplus \mathbb{K}f$ where we state

$$[h, e] = 2e\,, \quad [h, f] = -2f\,, \quad [e, f] = h\,,$$

but in the future will not record the additional, necessary, but implied relations, which in this case are

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

Of course in order to be sure that $L$ really is a Lie algebra, we must verify the Jacobi Identity equations for all quadruples $(i, j, l, m) \in \{h, e, f\}^4$. (Exercise.)

## 2.3 Matrix algebras

### 2.3.1 Standard subalgebras of $\mathfrak{gl}_n(\mathbb{K})$

Many Lie algebras occur naturally as matrix algebras. We have already mentioned $\mathrm{Mat}_n^-(\mathbb{K})$. This is often written $\mathfrak{gl}_n(\mathbb{K})$, the *general linear algebra*, in part because it is the Lie algebra of the Lie group $\mathrm{GL}_n(\mathbb{K})$; see Theorem (3.7)(a) below. The Gothic (or Fraktur) font is also a standard for Lie algebras.

A standard matrix calculation shows that $\mathrm{tr}(MN) = \mathrm{tr}(NM)$, so the subset of matrices of trace 0 is a dimension $n^2 - 1$ subalgebra $\mathfrak{sl}_n(\mathbb{K})$ of the algebra $\mathfrak{gl}_n(\mathbb{K})$, which itself has dimension $n^2$. Indeed the *special linear algebra* $\mathfrak{sl}_n(\mathbb{K})$ is the *commutator subalgebra* $[\mathfrak{gl}_n(\mathbb{K}), \mathfrak{gl}_n(\mathbb{K})]$ spanned by all $[M, N]$ for $M, N \in \mathfrak{gl}_n(\mathbb{K})$; see Section 4.1 below.

The subalgebras $\mathfrak{n}_n^+(\mathbb{K})$ and $\mathfrak{n}_n^-(\mathbb{K})$ are, respectively, composed of all strictly upper triangular and all strictly lower triangular matrices. Both have dimension $\binom{n}{2}$. Next let $\mathfrak{d}_n(\mathbb{K})$ and $\mathfrak{h}_n(\mathbb{K})$ be the abelian subalgebras of, respectively, all diagonal matrices (dimension $n$) and all diagonal matrices of trace 0 (dimension $n - 1$). We have the *triangular decomposition*:

$$\mathfrak{gl}_n(\mathbb{K}) = \mathfrak{n}_n^+(\mathbb{K}) \oplus \mathfrak{d}_n(\mathbb{K}) \oplus \mathfrak{n}_n^-(\mathbb{K})$$

and

$$\mathfrak{sl}_n(\mathbb{K}) = \mathfrak{n}_n^+(\mathbb{K}) \oplus \mathfrak{h}_n(\mathbb{K}) \oplus \mathfrak{n}_n^-(\mathbb{K})\,.$$

This second decompositions and ones resembling it will be very important later.

Within the Lie algebra $\mathfrak{sl}_2(\mathbb{K})$, consider the three elements

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

so that $\mathfrak{h}_2(\mathbb{K}) = \mathbb{K}h$, $\mathfrak{n}_2^+(\mathbb{K}) = \mathbb{K}e$, and $\mathfrak{n}_2^-(\mathbb{K}) = \mathbb{K}f$, and

$$\mathfrak{sl}_2(\mathbb{K}) = \mathbb{K}h \oplus \mathbb{K}e \oplus \mathbb{K}f\,.$$

We then have (Exercise)

$$[h,e] = 2e\,, \quad [h,f] = -2f\,, \quad [e,f] = h\,,$$

and the algebra presented at the end of the previous section is indeed a Lie algebra, namely a copy of $\mathfrak{sl}_2(\mathbb{K})$. The basis $h, e, f$ of $\mathfrak{sl}_2(\mathbb{K})$ is called a *Chevalley basis* of this algebra.

The isomorphism of $\mathrm{Mat}_n(\mathbb{K})$ and $\mathrm{End}_{\mathbb{K}}(\mathbb{K}^n)$ lead to natural isomorphisms of the above subalgebras of $\mathfrak{gl}_n(\mathbb{K}) = \mathrm{Mat}_n^-(\mathbb{K})$ with subalgebras of $\mathrm{End}_{\mathbb{K}}^-(\mathbb{K}^n)$.

### 2.3.2 Lie algebras from forms

For the basic theory of bilinear forms, see Appendix A. For bilinear $b$, the $\mathbb{K}$-space of endomorphisms

$$\mathfrak{L}(V,b) = \{\, x \in \mathrm{End}_{\mathbb{K}}(V) \mid b(xv,w) = -b(v,xw) \text{ for all } v,w \in V \,\}$$

is then an Lie $\mathbb{K}$-subalgebra of $\mathrm{End}_{\mathbb{K}}^-(V)$. (Exercise.)

With $V = \mathbb{K}^n$ and $\mathrm{End}_{\mathbb{K}}(V) = \mathrm{Mat}_n(\mathbb{K}) = \mathfrak{gl}_n(K)$, we have some special cases of $\mathfrak{L}(V,b)$. Let $G = (b(e_i,e_j))_{i,j}$ be the *Gram matrix* of $b$ on $V$ (with respect to the usual basis). The condition above then becomes

$$\mathfrak{L}(V,b) = \{\, M \in \mathrm{Mat}_n(\mathbb{K}) \mid MG = -GM^\top \,\}\,.$$

For simplicity's sake we assume that $\mathbb{K}$ does not have characteristic 2.

(i) **Orthogonal algebras.**

(a) If $b$ is the usual nondegenerate orthogonal form with an orthonormal basis for $V$, then $\mathfrak{L}(V,b) = \mathfrak{so}_n(\mathbb{K})$. As matrices,

$$\mathfrak{so}_n(\mathbb{K}) = \{\, M \in \mathrm{Mat}_n(\mathbb{K}) \mid M = -M^\top \,\}\,.$$

If the field $\mathbb{K}$ is algebraically closed, then it is always possible to find a basis for which the Gram matrix $G$ is in split form as the $2l \times 2l$ matrix with $l$ blocks $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ down the diagonal when $n = 2l$ is even, and this same matrix with an additional single 1 on the diagonal when $n = 2l+1$ is odd.

For the split form over an arbitrary field $\mathbb{K}$, we may write $\mathfrak{so}_{2l}^+(\mathbb{K})$ in place of $\mathfrak{so}_{2l}(\mathbb{K})$.

(ii) **Symplectic algebras.** If $b$ is the usual nondegenerate (split) symplectic form on $V = \mathbb{K}^{2l}$ with symplectic basis $\mathcal{S} = \{\, v_i, w_i \mid 1 \le i \le l \,\}$ subject to

$b(v_i, v_j) = b(w_i, w_j) = 0$ and $b(v_i, w_j) = \delta_{i,j} = -b(w_j, v_i)$, then $\mathfrak{L}(V, b) = \mathfrak{sp}_{2l}(\mathbb{K})$. As matrices,

$$\mathfrak{sp}_{2l}(\mathbb{K}) = \{ M \in \mathrm{Mat}_{2l}(\mathbb{K}) \mid MG = -GM^\top \} ,$$

where $G$ is the $2l \times 2l$ matrix with $n$ blocks $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ down the diagonal.

The notation is not uniform. Especially, when $\mathbb{K} = \mathbb{R}$ the field is sometimes omitted, hence one may find

$$\mathfrak{gl}_n(\mathbb{R}) = \mathfrak{gl}(n, \mathbb{R}) = \mathfrak{gl}(n) = \mathfrak{gl}_n , \quad \mathfrak{sl}_n(\mathbb{R}) = \mathfrak{sl}(n, \mathbb{R}) = \mathfrak{sl}(n) = \mathfrak{sl}_n ;$$

and

$$\mathfrak{so}_n(\mathbb{R}) = \mathfrak{so}(n, \mathbb{R}) = \mathfrak{so}(n) = \mathfrak{so}_n .$$

More confusingly, in the case of symplectic algebras the actual definition can vary as well as the notation; see [Tu11, p. 160].

## 2.4 Derivations

A *derivation* $D$ on the $\mathbb{K}$-algebra $A$ is a linear transformation $D \in \mathrm{End}_{\mathbb{K}}(A)$ with

$$D(fg) = fD(g) + D(f)g ,$$

for all $f, g \in A$. This should be recognized as the *Leibniz product rule*. Clearly the set $\mathrm{Der}_{\mathbb{K}}(A)$ is a $\mathbb{K}$-subspace of $\mathrm{End}_{\mathbb{K}}(A)$, but in fact this provides an amazing machine for constructing Lie algebras:

**(2.1).** THEOREM. $\mathrm{Der}_{\mathbb{K}}(A) \le \mathrm{End}^-_{\mathbb{K}}(A)$. *That is, the derivation space is a Lie $\mathbb{K}$-algebra under the bracket product.*

PROOF. Let $D, E \in \mathrm{Der}_{\mathbb{K}}(A)$. Then, for all $f, g \in A$,

$$\begin{aligned} [D, E](fg) &= (DE - ED)(fg) = DE(fg) - ED(fg) \\ &= D(fEg + (Ef)g) - E(fDg + (Df)g) \\ &= D(fEg) + D((Ef)g) - E(fDg) - E((Df)g)) \\ &= fDEg + DfEg + EfDg + (DEf)g \\ &\quad - fEDg - EfDg - DfEg - (EDf)g \\ &= fDEg - fEDg + (DEf)g - (EDf)g \\ &= f([D, E]g) + ([D, E]f)g . \qquad \square \end{aligned}$$

The definition of derivations then tells us that the injection of $\mathrm{Der}_{\mathbb{K}}(A)$ into $\mathrm{End}^-_{\mathbb{K}}(A)$ gives a representation of the Lie *derivation algebra* $\mathrm{Der}_{\mathbb{K}}(A)$ on the $\mathbb{K}$-space $A$.

**(2.2).** COROLLARY. *The image of the Lie algebra $A$ under the adjoint representation is a subalgebra of* $\mathrm{Der}_{\mathbb{K}}(A)$ *and* $\mathrm{End}^-_{\mathbb{K}}(A)$.

PROOF. The image of $A$ under ad is a $\mathbb{K}$-subspace of $\mathrm{End}_{\mathbb{K}}(A)$ by our very first Lemma (1.1). It remains to check that each $\mathrm{ad}_a$ is a derivation of $A$.

We start from the Jacobi Identity:

$$[[a,y],z]+[[y,z],a]+[[z,a],y]=0\,,$$

hence

$$-[[y,z],a]=[[a,y],z]+[[z,a],y]\,.$$

That is,

$$[a,[y,z]]=[[a,y],z]+[y,[a,z]]\,,$$

or

$$\mathrm{ad}_a[y,z]=[\mathrm{ad}_a\,y,z]+[y,\mathrm{ad}_a\,z]\,. \qquad \square$$

The map $\mathrm{ad}_a$ is then an *inner derivation* of $A$, and the Lie subalgebra $\mathrm{InnDer}_{\mathbb{K}}(A)=\{\,\mathrm{ad}_a \mid a\in A\,\}$ is the *inner derivation algebra.*

We have an easy but useful observation:

**(2.3).** PROPOSITION. *Every linear transformation of* $\mathrm{End}_{\mathbb{K}}(A)$ *is a derivation of the abelian Lie algebra* $A$.

PROOF. For $D\in\mathrm{End}_{\mathbb{K}}(A)$ and $a,b\in A$

$$D[a,b]=0=0+0=[Da,b]+[a,Db]\,. \qquad \square$$

### 2.4.1 Derivations of polynomial algebras

**(2.4).** PROPOSITION.

(a) $\mathrm{Der}_{\mathbb{K}}(\mathbb{K})=0$.

(b) *If the* $\mathbb{K}$*-algebra* $A$ *has an identity element* $1$*, then for each* $D\in\mathrm{Der}_{\mathbb{K}}(A)$ *and each* $c\in\mathbb{K}1$ *we have* $D(c)=0$.

(c) $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[t])=\{\,p(t)\frac{d}{dt}\mid p(t)\in\mathbb{K}[t]\,\}$*, a Lie algebra of infinite* $\mathbb{K}$*-dimension with basis* $\{\,t^i\frac{d}{dt}\mid i\in\mathbb{N}\,\}$.

PROOF. Part (b) clearly implies (a).

(b) Let $c=c1\in\mathbb{K}1$. Then for all $x\in A$ and all $D\in\mathrm{Der}_{\mathbb{K}}(A)$ we have

$$D(cx)=cD(x)$$

as $D$ is a $\mathbb{K}$-linear transformation. But $D$ is also a derivation, so

$$D(cx)=cD(x)+D(c)x\,.$$

We conclude that $D(c)x=0$ for all $x\in A$, and so $D(c)=0$.

(c) Let $D\in\mathrm{Der}_{\mathbb{K}}(A)$. By (b) we have $D(\mathbb{K}1)=0$. As the algebra $A$ is generated by 1 and $t$, the knowledge of $D(t)$ together with the product rule should give us everything. Set $p(t)=D(t)$.

We claim that $D(t^i) = p(t)it^{i-1}$ for all $i \in \mathbb{N}$. We prove this by induction on $i$, the result being clear for $i = 0, 1$. Assume the claim for $i - 1$. Then

$$\begin{aligned} D(t^i) &= D(t^{i-1}t) = t^{i-1}D(t) + D(t^{i-1})t \\ &= t^{i-1}p(t) + p(t)(i-1)t^{i-2}t = p(t)it^{i-1}\,, \end{aligned}$$

as claimed.

As $D$ is a linear transformation, if $a(t) = \sum_{i=0}^m a_i t^i$, then

$$D(a(t)) = \sum_{i=0}^m a_i D(t^i) = \sum_{i=0}^m a_i p(t)it^{i-1} = p(t)\sum_{i=0}^m ia_i t^{i-1} = p(t)\frac{d}{dt}a(t)\,,$$

completing the proposition. □

In $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[t])$ there is the subalgebra $A = \mathbb{K}h \oplus \mathbb{K}e \oplus \mathbb{K}f$ with $e = \frac{d}{dt}$, $h = -2t\frac{d}{dt}$, $f = -t^2\frac{d}{dt}$, and relations (Exercise)

$$[e, f] = h\,, \quad [h, e] = 2e\,, \quad [h, f] = -2f\,;$$

so we have $\mathfrak{sl}_2(\mathbb{K})$ again.

We next consider $\mathbb{K}[x, y]$. A similar argument to that of the proposition proves

$$\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[x, y]) = \{\, p(x, y)\frac{\partial}{\partial x} + q(x, y)\frac{\partial}{\partial y} \mid p(x, y), q(x, y) \in \mathbb{K}[x, y]\,\}\,.$$

(See Problem (2.8).) We examine two special situations—a subalgebra and a quotient algebra.

(i) Consider the Lie subalgebra that leaves each homogeneous piece of $\mathbb{K}[x, y]$ invariant. This subalgebra has basis

$$h_x = x\frac{\partial}{\partial x}\,,\ e = x\frac{\partial}{\partial y}\,,\ f = y\frac{\partial}{\partial x}\,,\ h_y = y\frac{\partial}{\partial y}\,.$$

Set $h = h_x - h_y = x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}$. Then

$$[e, f] = h\,, \quad [h, e] = 2e\,, \quad [h, f] = -2f\,,$$

giving $\mathfrak{sl}_2(\mathbb{K})$ yet again. The 4-dimensional algebra $\mathbb{K}h_x \oplus \mathbb{K}h_y \oplus \mathbb{K}e \oplus \mathbb{K}f$ is isomorphic to $\mathfrak{gl}_2(\mathbb{K})$ with the correspondences

$$h_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad h_y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Each homogeneous piece of $\mathbb{K}[x, y]$ carries a representation of $\mathfrak{gl}_2(\mathbb{K})$ and $\mathfrak{sl}_2(\mathbb{K})$ via restriction from the action of $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[x, y])$. The degree $m$ homogeneous component $\mathbb{K}[x, y]_m$ is then a cyclic $\mathbb{K}e$- hence $\mathfrak{sl}_2(\mathbb{K})$-module $M_0(m+1)$ of dimension $m+1$ with generator $y^m$. This will be important in Chapter ??.

(ii) The algebra $\mathbb{K}[x, y]$ has as quotient the algebra $\mathbb{K}[x, x^{-1}]$ of all *Laurent polynomials* in $x$. A small extension of the arguments from Proposition (2.4)(c) (Exercise) proves that $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[x, x^{-1}])$ has $\mathbb{K}$-basis consisting of the distinct elements

$$L_m = -x^{m+1}\frac{d}{dx} \quad \text{for } m \in \mathbb{Z}\,.$$

We write the generators in this form, since they then have the nice presentation

$$[L_m, L_n] = (m - n)L_{m+n}\,.$$

All the multiplication coefficients are integers. The $\mathbb{Z}$-algebra with this presentation has infinite dimension. It is called the *Witt algebra* over $\mathbb{Z}$, just as its tensor with $\mathbb{K}$, $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[x, x^{-1}])$, is the Witt algebra over $\mathbb{K}$.

### 2.4.2 Derivations of nonassociative algebras

We may also consider derivations of the nonassociative algebras we have encountered, specifically the octonion $\mathbb{K}$-algebra $O$ and (in characteristic not 2) its related Albert algebra—the exceptional Jordan $\mathbb{K}$-algebra $\mathrm{H}_3(O)$. The derivation algebra $\mathrm{Der}_{\mathbb{K}}(O)$ has dimension 14 (when $\mathrm{char}\mathbb{K} \neq 3$) and is said to have *type* $\mathfrak{g}_2$ while the algebra of inner derivations of the Albert algebra $\mathrm{H}_3(O)$ has dimension 52 and is said to have *type* $\mathfrak{f}_4$. Especially when $\mathbb{K}$ is algebraically closed and of characteristic 0 we have the uniquely determined algebras $\mathfrak{g}_2(\mathbb{K})$ and $\mathfrak{f}_4(\mathbb{K})$, respectively.

## 2.5 New algebras from old

### 2.5.1 Extensions

As we have seen and expect, subalgebras and quotients are ways of constructing new algebras out of old algebras. We can also extend old algebras to get new ones. As with groups, central extensions are important since the information we have about a given situation may come to us, via the adjoint, in projective rather than affine form.

The *Virasoro algebra* is a central extension of the complex Witt algebra. If $W$ is the Witt $\mathbb{Z}$-algebra, then

$$\mathrm{Vir}_{\mathbb{C}} = (\mathbb{C} \otimes_{\mathbb{Z}} W) \oplus \mathbb{C}c$$

with $[w, c] = 0$ for all $w \in W$ and

$$[L_m, L_n] = (m - n)L_{m+n} + \delta_{m,-n}\frac{m(m^2 - 1)}{12}c\,.$$

The multiplication coefficients are half-integers.

The Virasoro algebra is important in applications to physics and other situations. As seen after Proposition (2.4), the Witt and Virasoro algebras both contain the subalgebra $\mathbb{C}L_{-1} \oplus \mathbb{C}L_0 \oplus \mathbb{C}L_1$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. As we shall find in Section 6.2, large parts of the finite dimensional Lie algebra theory depend upon the construction of Lie subalgebras $\mathfrak{sl}_2(\mathbb{K})$. Similarly, the infinite dimensional Lie algebras that come up in physics and elsewhere are often handled using Witt and Virasoro subalgebras, which are in a sense the infinite dimensional substitutes for the finite dimensional $\mathfrak{sl}_2(\mathbb{K})$.

Given a complex simple Lie algebra like $\mathfrak{sl}_2(\mathbb{C})$, the corresponding *affine Lie algebra* comes from a two step process. First extend scalars to the Laurent polynomials and second take an appropriate central extension. So:

$$\widehat{\mathfrak{sl}}_2(\mathbb{C}) = (\mathbb{C}[t, t^{-1}] \otimes_{\mathbb{C}} \mathfrak{sl}_2(\mathbb{C})) \oplus \mathbb{C}c$$

where the precise cocycle on the complex Lie algebra $\mathbb{C}[t, t^{-1}] \otimes_{\mathbb{C}} \mathfrak{sl}_2(\mathbb{C})$ that gives the extension is defined in terms of the Killing form on the algebra $\mathfrak{sl}_2(\mathbb{C})$. (See Section 5.4 below.)

One often writes the Lie algebra $\mathbb{C}[t, t^{-1}] \otimes_{\mathbb{C}} \mathfrak{sl}_2(\mathbb{C})$ instead as $\mathfrak{sl}_2(\mathbb{C}) \otimes_{\mathbb{C}} \mathbb{C}[t, t^{-1}]$, viewing its elements as "Laurent polynomials" with coefficients from the algebra $\mathfrak{sl}_2(\mathbb{C})$.

It is also possible to form split extensions of Lie algebras, with derivations playing the role that automorphisms play in group extensions. (See Section 4.3.) The canonical derivation $\frac{d}{dt}$ on the Laurent polynomials induces a derivation of the affine algebra which is then used to extend the affine algebra so that it has codimension 1 in the corresponding *Kac-Moody Lie algebra*.

### 2.5.2 Embeddings

We saw above that derivations of octonion and Jordan algebras give new Lie algebras. Tits, Kantor, and Koecher [Tit66] used these same nonassociative algebras to construct (the *TKK construction*) Lie algebras that are still more complicated. In particular, the space

$$\mathrm{Der}_{\mathbb{C}}(\mathbb{O}^{\mathrm{sp}}(\mathbb{C})) \;\oplus\; (\mathbb{O}^{\mathrm{sp}}(\mathbb{C})_0 \otimes_{\mathbb{C}} \mathrm{H}_3(\mathbb{O}^{\mathrm{sp}}(\mathbb{C}))_0) \;\oplus\; \mathrm{Der}_{\mathbb{C}}(\mathrm{H}_3(\mathbb{O}^{\mathrm{sp}}(\mathbb{C})))$$

of dimension $14+(8-1)\times(27-1)+52 = 248$ can be provided with a Lie algebra product (extending that of the two derivation algebra pieces) that makes it into the Lie algebra $\mathfrak{e}_8(\mathbb{C})$. Here $\mathbb{O}^{\mathrm{sp}}(\mathbb{C})_0$ is $1^{\perp}$ in $\mathbb{O}^{\mathrm{sp}}(\mathbb{C})$ and $\mathrm{H}_3(\mathbb{O}^{\mathrm{sp}}(\mathbb{C}))_0$ is a similarly defined subspace of codimension 1 in $\mathrm{H}_3(\mathbb{O}^{\mathrm{sp}}(\mathbb{C}))$. The Lie algebra $\mathfrak{e}_8(\mathbb{C})$ furthermore has the important subalgebras $\mathfrak{e}_6(\mathbb{C})$ of dimension 78 and $\mathfrak{e}_7(\mathbb{C})$ of dimension 133.

## 2.6 Other contexts

### 2.6.1 Nilpotent groups

Let $G$ be a nilpotent group with *lower central series*

$$G = \mathrm{L}^1(G) \trianglerighteq \mathrm{L}^2(G) \trianglerighteq \cdots \trianglerighteq \mathrm{L}^{n+1}(G) = 1$$

where $\mathrm{L}^{k+1}(G)$ is defined as $[G, \mathrm{L}^k(G)]$. For each $1 \leq k \leq n$ set

$$L_k = \mathrm{L}^k(G)/\mathrm{L}^{k+1}(G)\,,$$

an abelian group as is the sum

$$L = \bigoplus_{k=1}^{n} L_k\,.$$

As $G$ is nilpotent, always

$$[\mathrm{L}^i(G), \mathrm{L}^j(G)] \leq \mathrm{L}^{i+j}(G)\,.$$

This provides the relations that turn the group $L = L_G$ into a *Lie ring*—we do not require it to be free as $\mathbb{Z}$-module—within which we have

$$[L_i, L_j] \leq L_{i+j}\,.$$

Certain questions about nilpotent groups are much more amenable to study in the context of Lie rings and algebras [Hig58]. A particular important instance is the *Restricted Burnside Problem*, which states that an $m$-generated finite nilpotent group of exponent $e$ has order less than or equal to some function $f(m, e)$, dependent only on $m$ and $e$. Professor E. Zelmanov received a Fields Medal in 1994 for the positive solution of the Restricted Burnside Problem. His proof [Zel97] makes heavy use of Lie methods.

### 2.6.2 Vector fields

We shall see in the next chapter that the tangent space to a Lie group at the identity is a Lie algebra. As the group acts regularly on itself by translation, this space is isomorphic to the Lie algebra of invariant vector fields on the group.

Indeed often a *vector field* on the smooth manifold $M$ is defined to be a derivation of the algebra $\mathrm{C}^\infty(M)$ of all smooth functions; for instance, see [Hel01, p. 9]. Thus the space of all vector fields is the corresponding derivation algebra and so automatically has a Lie algebra structure.

For instance, the Lie group of rotations of the circle $\mathrm{S}^1$ is the group $\mathrm{SO}_2(\mathbb{R})$ of all matrices

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

which becomes $e^{i\theta}$ when we extend coefficients to the complex numbers. The corresponding spaces of invariant vector fields have dimension 1.

The space $\mathrm{C}^\infty(\mathrm{S}^1)$ of all smooth functions on the circle consists of those functions that can be expanded as convergent *Fourier series*

$$\sum_{m \in \mathbb{Z}} a_m \sin(m\theta) + b_m \cos(m\theta)\,,$$

which after extension to $\mathbb{C}$ becomes the simpler

$$\sum_{m\in\mathbb{Z}} c_m e^{im\theta} .$$

This space $\mathrm{C}^\infty_\mathbb{C}(\mathrm{S}^1)$ has as a dense subalgebra the space of all *Fourier polynomials*, whose canonical basis is $\{ e^{im\theta} \mid m \in \mathbb{Z} \}$.

The group of all complex orientation preserving diffeomorphisms of the circle (an "infinite dimensional Lie group") is an open subset of $\mathrm{C}^\infty_\mathbb{C}(\mathrm{S}^1)$ and has as corresponding space of smooth vector fields (not just those that are invariant) all $f\frac{d}{d\theta}$ for $f$ smooth. The dense Fourier polynomial subalgebra with basis $L_m = ie^{im\theta}\frac{d}{d\theta}$ then has

$$[L_m, L_n] = (m-n)L_{m+n} ,$$

giving the complex Witt algebra again.

## 2.7 Problems

**(2.5).** PROBLEM. *Classify up to isomorphism all Lie $\mathbb{K}$-algebras of dimension* 2. *(Of course, the abelian algebra gives the only isomorphism class in dimension* 1.*)*

**(2.6).** PROBLEM. *Prove that over an algebraically closed field $\mathbb{K}$ of characteristic not* 2*, the Lie algebra $\mathfrak{sl}_2(\mathbb{K})$ is isomorphic to $\mathfrak{so}_3(\mathbb{K})$, the orthogonal Lie algebra of $3 \times 3$ skew-symmetric matrices.*

**(2.7).** PROBLEM. *Find all subalgebras of $\mathfrak{sl}_2(\mathbb{K})$ that contain the subalgebra $H = \mathbb{K}h$.* HINT: *Small characteristic can produce anomalous results.*

**(2.8).** PROBLEM. *Calculate* $\mathrm{Der}_\mathbb{K}(\mathbb{K}[x_1, \ldots x_n])$.

**(2.9).** PROBLEM. *Consider the matrix subgroup $\mathrm{UT}_n(\mathbb{K})$ of $\mathrm{GL}_n(\mathbb{K})$, consisting of the upper unitriangular matrices—those which have* 1*'s on the diagonal, anything above the diagonal, and* 0*'s below the diagonal.*

(a) *Prove that $G = \mathrm{UT}_n(\mathbb{K})$ is a nilpotent group.*

(b) *Starting with this group $G$, construct the Lie algebra $L = L_G$ as in Section 2.6.1. Prove that $L$ is isomorphic to the Lie algebra $\mathfrak{n}_n^+(\mathbb{K})$.*

**(2.10).** PROBLEM. *Consider the subgroup $\mathrm{X}_n(\mathbb{K})$ of upper unitriangular matrices that have* 1*'s on the diagonal, anything in the nondiagonal part of the first row and last column, and* 0*'s elsewhere.*

(a) *By the previous problem $X = \mathrm{X}_n(\mathbb{K})$ is nilpotent. Prove that for $n \geq 2$ it has nilpotence class exactly* 2 *and that its center is equal to its derived group and consists only of those matrices with* 1*'s down the diagonal and the only other nonzero entries found in the upper-righthand corner.*

(b) *Starting with this group $X$, construct the Lie algebra $L = L_X$ as in Section 2.6.1. Prove that $L$ is isomorphic to the Lie algebra on the space*

$$M = \mathbb{K}z \oplus \bigoplus_{i=1}^{n-1} (\mathbb{K}x_i \oplus \mathbb{K}y_i)$$

*with relations given by*

$$[x_i, y_i] = -[y_i, x_i] = z\,,$$

*for all* $i$*, and all other brackets among generators equal to* $0$*.*

REMARK. *This Lie algebra is the* Heisenberg algebra *of dimension* $2n-1$ *over* $\mathbb{K}$*.*

# Chapter 3

# Lie groups

> (N. Jacobson [Jac79, p. 1]:) *The theory of Lie algebras is an outgrowth of the Lie theory of continuous groups.*

> (R. Carter [Car05, p. xiii]:) *Lie algebras were originally introduced by S. Lie as algebraic structures used for the study of Lie groups.*

It would be wrong for us to talk at length about Lie algebras without devoting at least some time to the way in which they arise in the theory of Lie groups. We do that in an abbreviated form in this chapter.

For us, Lie's work and the work that it motivated contain two basic observations:

(i) *If $G$ is a Lie group, then the tangent space to the identity is a Lie algebra $\Lambda(G)$.*

(ii) *The representation theory of the Lie group $G$ and of the Lie algebra $\Lambda(G)$ are essentially the same.*

The second observation displays genuine progress, since a Lie algebra is a linear object whereas the Lie group is not. This is the same advantage obtained in the passage from a nilpotent group to its associated Lie ring in Section 2.6.1.

This chapter is included in order to place Lie algebras in one of their most important contexts, historically and practically. Its material will not be used in the rest of the notes or course. Therefore for ease of presentation we assume uniformly throughout that the vector spaces, groups, and algebras we examine are defined over the real numbers. Given our later focus on algebraically closed fields of characteristic 0, it might make more sense to restrict to the complex case; but that would require more sophisticated calculus/analysis than we care to use.

## 3.1 Representation theory as spectral theory

The two observations beg the question, "What is so good about representation theory?" After all, many of our important Lie groups and algebras are already defined in terms of matrices. Why worry about more representations?

Lie and those who followed him were interested in using Lie theory to solve problems, and it is often easier to solve a problem in pieces rather than all at once. An important example is the analysis of the action of a linear transformation in terms of its eigenspaces. Such decompositions are collected together under the heading of spectral theory, and they are served by various canonical form results.

The representation theory of groups (and other algebras) can be thought of as a general form of spectral or canonical form theory. If the initial, say physical, statement of a problem has some inherent symmetry, then that symmetry should also be evident in the space of solutions. Lie noted that this action could be exploited to decompose the solution space and so perhaps find nice descriptions for the solutions. At the heart of matrix canonical form results is the feeling that matrices containing lots of zeros are the easiest to deal with.

Lie was interested in particular in solving differential equations. Dresner [Dre99, p. 16] shows how, starting from the differential equation

$$\frac{d}{dx}y = -\frac{y(y^2 - x)}{x}\,,$$

once one has noticed that the solution set is invariant under the change of variables

$$x = x_0 \longrightarrow x_s = e^s x \quad y = y_0 \longrightarrow y_s = e^{s/2} y\,,$$

for all $s \in \mathbb{R}$, it is relative easy to construct an integrating factor

$$\varphi(x, y) = \left(xy^3 - \frac{x^2 y}{2}\right)^{-1}$$

and so reach the closed form solution set

$$y = x(2x + c)^{1/2}\,.$$

The displayed symmetry group $\{\, e^s \mid s \in \mathbb{R} \,\} \simeq (\mathbb{R}, +)$ is continuous and even smooth in its variable. This type of symmetry is evident in many physical situations, and this led Lie (and others) to the study of smooth groups and their representations. We shall see in Section 3.4 that the most basic Lie group $(\mathbb{R}, +)$ is also one of the most important.

## 3.2 Lie groups and Hilbert's Fifth Problem

A *Lie group* is a smooth manifold $G$ that is also a group. These two conditions are linked by the requirements that the group multiplication $m\colon G \times G \longrightarrow G$

given by $m(x, y) = xy$ and the group inverse map $i\colon G \longrightarrow G$ given by $i(x) = x^{-1}$ are smooth maps on the manifold. Here (recalling that we are speaking of real manifolds) by *smooth* we mean $\mathrm{C}^\infty$. (For a complex manifold, smooth means holomorphic.)

Examples are provided by the *closed subgroups* of $\mathrm{GL}_n(\mathbb{R})$: those subgroups containing the limit of every sequence of group matrices for which that limit exists and is invertible. This already might be a surprise, since closure is a topological property, determined only by examining $\mathrm{C}^0$ continuity issues. The $\mathrm{C}^0$ condition is very weak when compared to the smooth $\mathrm{C}^\infty$ assumptions of the manifold definition.

If $G$ is a Lie group, then certainly

(i) $G$ is a topological group (that is, the maps $m$ and $i$ are continuous) and

(ii) $G$ is locally a finite dimensional Euclidean space.

One reading of *Hilbert's Fifth Problem* is that, in fact, the Lie groups are exactly the locally Euclidean topological groups. Once made precise, this version of the Fifth Problem was proven by Montgomery and Zippin [MoZi55] and Gleason [Gle52] in 1952. (See [Tao14] for more.)

Cartan first proved that closed subgroups of $\mathrm{GL}_n(\mathbb{R})$ are Lie groups. As such, it is reasonable to focus on such examples when initially discussing Lie groups. This is the approach take by several modern introductions to Lie groups [Eld15, Hal15, How83, vNe29, Ros02, Sti08, Tap05] and is largely what we do here. In particular, those not comfortable with manifolds need not worry—just focus on closed subgroups of $\mathrm{GL}_n(\mathbb{R})$.

Essentially everything we prove (or state) goes over to the general case, although some of the definitions and proofs would require more subtlety. In particular, in place of the concrete functions exp and log provided by convergent power series of matrices, one appeals to the uniqueness of solutions for appropriate ordinary differential equations and to the Inverse Function Theorem; see [CSM95, pp. 69-74].

## 3.3 Some matrix calculus

For the matrix $M = (m_{ij})_{ij} \in \mathrm{Mat}_{k,l}(\mathbb{R})$, set $|M| = \sqrt{\sum_{i,j} m_{ij}^2}$. This is the standard Euclidean norm on $\mathbb{R}^{kl}$; especially for $k = l = 1$ we have the usual $|(m)| = |m|$. We can then define limits of matrix functions, using this norm to determine "closeness." In turn, this gives meaning to statements that a function from one matrix space to another is continuous, for instance in our discussion above of multiplication and inversion in Lie groups.

For smoothness we need derivatives as well. The usual derivative of $f(x)$ at $x = a$ is given by

$$\lim_{t \longrightarrow 0} \frac{f(t+a) - f(a)}{t} = f'(a) = \left.\frac{df}{dx}\right|_{x=a} .$$

If we rewrite this as

$$\lim_{t \longrightarrow 0} \frac{f(t+a) - (f(a) + f'(a)t)}{t} = 0\,,$$

we are observing that near $a$ (near $t = 0$), the line $f(a) + f'(a)t$ is a good approximation to the function $f(t+a)$. This motivates the following definition of the *derivative of a matrix function*; see [Spi65, p. 16].

> *The linear transformation* $D\colon \mathrm{Mat}_{k,l}(\mathbb{R}) \longrightarrow \mathrm{Mat}_{m,n}(\mathbb{R})$ *is the derivative at* $A$ *of the matrix function* $F\colon \mathrm{Mat}_{k,l}(\mathbb{R}) \longrightarrow \mathrm{Mat}_{m,n}(\mathbb{R})$ *provided*
> $$\lim_{T \longrightarrow 0} \frac{|F(T+A) - F(A) - D(T)|}{|T|} = 0\,.$$

As derivatives are locally determined, to calculate the derivative of $F$ at $A$ we only need to know $F$ on some neighborhood of $A$ in $\mathrm{Mat}_{k,l}(\mathbb{R})$.

This definition is the appropriate one for checking properties, but our applications later in this chapter will only be concerned with the special case $k = l = 1$ and $m = n$. That is, we will consider matrix functions $F\colon (-r, r) \longrightarrow \mathrm{Mat}_n(\mathbb{R})$ for some positive $r$ with $a \in (-r, r)$. There we will use the equivalent but more familiar formulation

$$F'(a) = \lim_{t \longrightarrow 0} \frac{F(t+a) - F(a)}{t} \in \mathrm{Mat}_n(\mathbb{R})\,.$$

Once we have checked that matrix limits and derivatives behave as hoped and expected[1] (see, for instance, [Eld15],[Hal15]), we have

**(3.1).** PROPOSITION.

(a) *If the power series* $A(t) = \sum_{k=0}^{\infty} A_k t^k$ *converges for all* $|t| < r$*, then its derivative* $A'(t) = \sum_{k=0}^{\infty} k A_k t^{k-1}$ *also converges for all* $|t| < r$.

(b) $\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$ *converges*[2] *for all* $A \in \mathrm{Mat}_n(\mathbb{R})$*. For* $A, B \in \mathrm{Mat}_n(\mathbb{R})$ *with* $[A, B] = 0$ *we have* $\exp(A+B) = \exp(A)\exp(B)$*. Especially* $I = \exp(A)\exp(-A)$*, indeed* $\exp(kB) = \exp(B)^k$ *for integral* $k$.

(c) *For all* $A \in \mathrm{Mat}_n(\mathbb{R})$ *the unique solution of the matrix Initial Value Problem*

$$f'(t) = f(t)A\,, \quad f(0) = I$$

*is* $f(t) = \exp(tA)$.

(d) $\log(1+X) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{1}{k} X^k$ *converges for all* $X$ *with* $|X| < 1$*. For* $|X| < 1$*, we have* $\exp(\log(1+X)) = 1 + X$.

[1] Exercise: Check the matrix versions of Leibniz' $\frac{d}{dt}(p(t)q(t)) = p(t)q'(t) + p'(t)q(t)$ and of the chain rule.

[2] It may be of psychological and/or actual help to realize that $G(\exp A)G^{-1} = \exp GAG^{-1}$, so that Jordan Canonical Form can be used to reduce the limit parts of this calculation to the standard 1-dimensional case.

It is important that we can only guarantee $\exp(A+B) = \exp(A)\exp(B)$ when the matrices $A$ and $B$ commute. When they do, the corresponding power series multiplication goes through exactly as in the standard case. But if they do not commute, then things like $BAB$ and $AB^2$ on the lefthand side can be different, so collecting of like terms is greatly restricted.

Also note that we are defining the logarithm via its Taylor series, rather than the usual calculus definitions that use an integral or that legislate it to be the inverse function for the exponential. Thus for us it is only defined (convergent) near the identity. This will be good enough. (See the proof of the next proposition.)

The next proposition is an extension of the familiar result/definition from calculus

$$\exp(a) = \lim_{k\to\infty} \left(1 + \frac{a}{k}\right)^k ,$$

which is the special case $n = 1$ and $g(t) = 1 + at$ of the proposition.

**(3.2).** Proposition. *Let $g\colon (-r, r) \longrightarrow \mathrm{GL}_n(\mathbb{R})$ be differentiable at $0$ with $g(0) = I$ and $g'(0) = A$. Then $\lim_{k\to\infty} g\left(\frac{1}{k}\right)^k = \exp(A)$.*

Proof. Set $q(t) = \log(g(t))$ (for $t$ small enough so that $|g(t) - I| < 1$). By the chain rule, $q'(t) = g'(t)g(t)^{-1}$ (again for small $t$), so $q(0) = 0$ and $q'(0) = A$. Therefore by the definition of the matrix derivative

$$A = q'(0) = \lim_{t\to 0} \frac{\log g(t) - \log g(0)}{t} = \lim_{t\to 0} t^{-1} \log g(t) .$$

Setting $k = t^{-1}$ we gain $A = \lim_{k\to\infty} k \log g(k^{-1})$. As exponentiation is everywhere continuous,

$$\begin{aligned} \exp(A) &= \exp\left(\lim_{k\to\infty} k\log g(k^{-1})\right) = \lim_{k\to\infty} \exp\left(k\log g(k^{-1})\right) \\ &= \lim_{k\to\infty} g\left(\frac{1}{k}\right)^k , \end{aligned}$$

as desired. □

## 3.4 One-parameter subgroups

If $G$ is a Lie group, then a *one-parameter subgroup* of $G$ is a continuous homomorphism $\varphi\colon (\mathbb{R}, +) \longrightarrow G$. This links the weakest $\mathrm{C}^0$ continuity property of $G$ (and $\mathbb{R}$) with group theoretic structure. We shall see that this forces very strong continuity—not just $\mathrm{C}^\infty$ but $\mathrm{C}^\omega$ (analytic). For every $A \in \mathrm{Mat}_n(\mathbb{R})$, the analytic map $\varphi_A\colon \mathbb{R} \longrightarrow \mathrm{Mat}_n(\mathbb{R})$ given by $\varphi_A(t) = \exp(tA)$ is a one-parameter subgroup of $\mathrm{GL}_n(\mathbb{R})$ by Proposition (3.1). Surprisingly, the converse is true. This can be viewed as an important special case of Hilbert's Fifth Problem.

**(3.3).** THEOREM. *Let $\varphi\colon (\mathbb{R},+) \longrightarrow G$ be a one-parameter subgroup of the closed subgroup $G$ of $\mathrm{GL}_n(\mathbb{R})$. Then there is a unique matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ with $\varphi(t) = \exp(tA)$ for all $t \in \mathbb{R}$. In particular $\varphi$ is $\mathrm{C}^\infty$ and indeed analytic. We have $A = \varphi'(0) = \frac{d}{dt}\varphi|_{t=0}$.*

PROOF. Our proof follows [Eld15]. It has two parts. We first prove that $\varphi$ is differentiable and then prove that it is an exponential.

Set $F(t) = \int_0^t \varphi(u)\,du$. As $\varphi$ is continuous, $F$ is differentiable with $F(0) = 0$ and $F'(t) = \varphi(t)$, hence $F'(0) = I$. We use the fact that $\varphi$ is a homomorphism and make the change of variable $v = u - t$ to find

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

Next note that

$$I = F'(0) = \lim_{s\to 0} \frac{F(s)-F(0)}{s} = \lim_{s\to 0} \frac{F(s)}{s}\,,$$

hence

$$1 = \det I = \det\left(\lim_{s\to 0} \frac{F(s)}{s}\right) = \lim_{s\to 0}\left(s^{-n}\det F(s)\right)\,,$$

as det is continuous. Especially, for some small $s_0$ we must have $\det F(s_0) \neq 0$ and so $F(s_0)$ is invertible. But then the above tells us that

$$\varphi(t) = (F(t+s_0) - F(t))F(s_0)^{-1}$$

is differentiable, as desired for the first part of our argument.

We now have $\varphi$ differentiable with $\varphi(0) = I$. As $\varphi$ is a homomorphism

$$\begin{aligned}
\varphi'(t) &= \lim_{h\to 0} \frac{\varphi(t+h)-\varphi(t)}{h} = \lim_{h\to 0} \frac{\varphi(t)\varphi(h)-\varphi(t)}{h} \\
&= \lim_{h\to 0} \varphi(t)\frac{\varphi(h)-I}{h} = \varphi(t)\lim_{h\to 0} \frac{\varphi(h)-\varphi(0)}{h} \\
&= \varphi(t)\varphi'(0)
\end{aligned}$$

That is, for $\varphi'(0) = A$ the function $\varphi(t)$ solves the Initial Value Problem

$$\varphi'(t) = \varphi(t)A \quad \text{and} \quad \varphi(0) = I\,.$$

By the omnibus Proposition (3.1)(c) we have $\varphi(t) = \exp(tA)$, as claimed. □

**(3.4).** COROLLARY. $\det(\exp(tA)) = e^{t\mathrm{tr}(A)}$.

PROOF. The map $t \mapsto \det\exp(tA)$ is a one-parameter subgroup of $\mathrm{GL}_1(\mathbb{R})$. (Exercise.) Therefore there is a nonzero $a \in \mathbb{R}$ with $\det\exp(tA) = e^{ta}$ for $a = \frac{d}{dt}\det\exp(tA)|_{t=0}$.

We have $\exp(tA) = I + tA + t^2B(t)$ (for an appropriate convergent power series $B(t)$), hence with $A = (a_{ij})_{ij}$ the standard expansion of the determinant gives

$$\det\exp(tA) = 1 + t(a_{11} + \cdots + a_{nn}) + t^2c(t) = 1 + t\mathrm{tr}(A) + t^2c(t)\,.$$

Therefore $a = \frac{d}{dt}\det\exp(tA)|_{t=0} = \mathrm{tr}(A)$. □

## 3.5 The tangent space at the identity

Let $G$ be a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$. There are several ways of defining the *tangent space* at the identity element $I$ of the group $G$. We offer two—a relatively weak $\mathrm{C}^1$ (differentiable) version and a very strong $\mathrm{C}^\omega$ (analytic) condition.

A *curve* in $G$ is a differentiable map $c\colon J \longrightarrow G$, for some open interval $J$ in $\mathbb{R}$. In particular, a one-parameter subgroup is a special type of curve. Set

(i) $\mathrm{T}_I(G) = \{\, c'(0) \mid \text{curve } c\colon (-r,r) \longrightarrow G\,,\ \text{some } r \in \mathbb{R}^+,\ c(0) = I\,\}$;

(ii) $\Lambda(G) = \{\, A \mid \exp(tA) \le G\,\}$.

Clearly these tangent space candidates have the property $\Lambda(G) \subseteq \mathrm{T}_I(G)$, but we will prove in Theorem (3.6) below that we have equality. Again, this is in the spirit of Hilbert's Fifth Problem.

We first show that the tangent space is indeed a subspace.

**(3.5).** LEMMA. *$\mathrm{T}_I(G)$ is a subspace of $\mathrm{Mat}_n(\mathbb{R})$.*

PROOF. Let $A, B \in \mathrm{T}_I(G)$ and $a, b \in \mathbb{R}$. We must show that $aA + bB \in \mathrm{T}_1(G)$. Let differentiable

$$g\colon (-q,q) \longrightarrow G\,,\ g(0) = I\,,\ g'(0) = A$$

and

$$h\colon (-s,s) \longrightarrow G\,,\ h(0) = I\,,\ h'(0) = B$$

testify to $A, B \in \mathrm{T}_I(G)$.

First consider $c(t) = h(bt)$ on $(-r,r)$ with $r = |b^{-1}s|$ ($= \infty$ for $b = 0$). Then

$$c(0) = h(0) = I \quad \text{and} \quad c'(0) = bh'(0) = bB\,,$$

so $\mathrm{T}_I(G)$ is closed under scalar multiplication.

It remains to prove $A+B \in \mathrm{T}_I(G)$. For $r = \frac{1}{2}\min(q,s)$, the curve $c\colon (-r,r) \longrightarrow G$ given by

$$c(t) = \frac{1}{2}(g(2t) + h(2t))\,,$$

has

$$c(0) = \frac{1}{2}(g(0) + h(0)) = \frac{1}{2}(I + I) = I\,.$$

and

$$c'(0) = \frac{1}{2}(2g'(0) + 2h'(0)) = \frac{1}{2}(2A + 2B) = A + B\,.$$

Thus $A + B \in \mathrm{T}_1(G)$ as desired. □

**(3.6).** THEOREM. $\Lambda(G) = \mathrm{T}_I(G)$.

PROOF. We have already pointed out that $\Lambda(G) \subseteq \mathrm{T}_I(G)$. Now, for fixed but arbitrary $t \in \mathbb{R}$ and for each $B \in \mathrm{T}_I(G)$, we must prove that the matrix $\exp(tB)$ is in $G$, as then $t \mapsto \exp(tB)$ will be a one-parameter subgroup of $G$, exhibiting $B \in \Lambda(G)$ and providing the reverse containment $\Lambda(G) \supseteq \mathrm{T}_I(G)$. By the previous lemma $\mathrm{T}_I(G)$ is a $\mathbb{R}$-space, so it is enough to prove that $\exp(A) \in G$ for all $A \in \mathrm{T}_I(G)$.

For some $r \in \mathbb{R}^+$, let the curve $g\colon (-r,r) \longrightarrow G$ have $g(0) = I$ and $g'(0) = A$. Then for all integral $k$ greater than some $N$ we have $g\left(\frac{1}{k}\right) \in G$. As $G$ is a group, in turn $g\left(\frac{1}{k}\right)^k \in G$. Proposition (3.2) gives $\lim_{k\to\infty} g\left(\frac{1}{k}\right)^k = \exp(A)$, which is always invertible. As $G$ is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$, we conclude $\exp(A) \in G$ as desired. □

It is now appropriate for us to define the *tangent space* at the identity element $I$ of the group $G$, closed in $\mathrm{GL}_n(\mathbb{R})$, to be the $\mathbb{R}$-space $\Lambda(G) = \mathrm{T}_I(G)$.

Of course $\mathrm{GL}_n(\mathbb{R})$ is closed in itself. Additionally $\mathrm{SL}_n(\mathbb{R})$ is closed in $\mathrm{GL}_n(\mathbb{R})$ as it consists of all matrices $X$ with $\det(X) - 1 = 0$.

**(3.7).** THEOREM.

(a) $\Lambda(\mathrm{GL}_n(\mathbb{R})) = \mathfrak{gl}_n(\mathbb{R})$ *and* $\langle\, \exp(tA) \mid A \in \mathfrak{gl}_n(\mathbb{R})\,\rangle = \mathrm{GL}_n(\mathbb{R})^+$, *the subgroup of index* 2 *in* $\mathrm{GL}_n(\mathbb{R})$ *of all matrices with positive determinant.*

(b) $\Lambda(\mathrm{SL}_n(\mathbb{R})) = \mathfrak{sl}_n(\mathbb{R})$ *and* $\langle\, \exp(tA) \mid A \in \mathfrak{sl}_n(\mathbb{R})\,\rangle = \mathrm{SL}_n(\mathbb{R})$.

PROOF. The equality $\Lambda(\mathrm{GL}_n(\mathbb{R})) = \mathrm{Mat}_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R})$ is clear from Proposition (3.1)(b).

We next consider $A \in \Lambda(\mathrm{SL}_n(\mathbb{R}))$. By Corollary (3.4), for the one-parameter subgroup $\exp(tA)$ of $\mathrm{SL}_n(\mathbb{R})$ we have

$$1 = \det\left(\exp(tA)\right) = e^{t\mathrm{tr}(A)}\,.$$

That is, $\mathrm{tr}(A) = 0$ and $A \in \mathfrak{sl}_n(\mathbb{R})$. Conversely, for $A \in \mathfrak{sl}_n(\mathbb{R})$, by the same corollary

$$1 = e^{t\mathrm{tr}(A)} = \det\left(\exp(tA)\right)\,.$$

This is true for arbitrary $t$, so $t \mapsto \exp(tA)$ is a one-parameter subgroup of $\mathrm{SL}_n(\mathbb{R})$. Thus $A \in \Lambda(\mathrm{SL}_n(\mathbb{R}))$, hence $\Lambda(\mathrm{SL}_n(\mathbb{R})) = \mathfrak{sl}_n(\mathbb{R})$.

For each elementary matrix unit $e_{ij} \in \mathrm{Mat}_n(\mathbb{R})$ with $i \neq j$, we have $e_{ij} \in \mathfrak{sl}_n(\mathbb{R})$ and $e_{ij}^2 = 0$. Thus $\exp(te_{ij}) = I + te_{ij}$, an elementary transvection subgroup. By Gaussian elimination,

$$\begin{aligned} \langle \exp(tA) \mid A \in \mathfrak{sl}_n(\mathbb{R}) \rangle \leq \mathrm{SL}_n(\mathbb{R}) &= \langle I + te_{ij} \mid i \neq j,\, t \in \mathbb{R} \rangle \\ &\leq \langle \exp(tA) \mid A \in \mathfrak{sl}_n(\mathbb{R}) \rangle . \end{aligned}$$

Therefore $\langle \exp(tA) \mid A \in \mathfrak{sl}_n(\mathbb{R}) \rangle = \mathrm{SL}_n(\mathbb{R})$.

If $D = \mathrm{diag}(d_{11}, \dots, d_{ii}, \dots, d_{nn})$ is a diagonal matrix, then $\exp(D)$ is also diagonal with entries $e^{d_{ii}}$. Every diagonal matrix with positive entries on the diagonal can be found this way, and these together with $\mathrm{SL}_n(\mathbb{R})$ generate $\mathrm{GL}_n(\mathbb{R})^+$. By Corollary (3.4), every matrix exponential has positive determinant; so $\langle \exp(tA) \mid A \in \mathfrak{gl}_n(\mathbb{R}) \rangle = \mathrm{GL}_n(\mathbb{R})^+$. □

**(3.8).** COROLLARY. *Although* $\mathrm{GL}_n(\mathbb{R})^+$ *has index* 2 *in* $\mathrm{GL}_n(\mathbb{R})$*, the two groups have the same tangent space at the identity*

$$\Lambda(\mathrm{GL}_n(\mathbb{R})^+) = \Lambda(\mathrm{GL}_n(\mathbb{R})) = \mathfrak{gl}_n(\mathbb{R}) . \qquad \square$$

In the remaining results of this subsection, we set $L = \Lambda(G) = \mathrm{T}_1(G)$ for our closed subgroup $G$ of $\mathrm{Mat}_n(\mathbb{R})$.

**(3.9).** LEMMA. *If* $g \in G$ *and* $A \in L$*, then* $gAg^{-1} \in L$.

PROOF. As $g \in G$ and $A \in L$, the group $G$ contains $\exp(tA)$ and

$$\begin{aligned} g(\exp(tA))g^{-1} &= g\left(\sum_{k=0}^{\infty} \frac{t^k A^k}{k!}\right) g^{-1} = \sum_{k=0}^{\infty} g \left(\frac{t^k A^k}{k!}\right) g^{-1} \\ &= \sum_{k=0}^{\infty} \frac{t^k (gA^k g^{-1})}{k!} = \sum_{k=0}^{\infty} \frac{t^k (gAg^{-1})^k}{k!} \\ &= \exp(t(gAg^{-1})) . \end{aligned}$$

Therefore $gAg^{-1} \in L$. □

Thus we have the *adjoint representation* of the group $G$ on its Lie algebra $L$:

$$\mathrm{Ad}\colon G \longrightarrow \mathrm{GL}_{\mathbb{R}}(L) \quad \text{given by} \quad \mathrm{Ad}_g(A) = gAg^{-1} .$$

It should come as no surprise that in general a Lie group acts on its Lie algebra, the corresponding representation always being called adjoint.

**(3.10).** LEMMA. *For* $A, B \in L$,

$$\mathrm{Ad}_{\exp(tB)}(A) = A + t(BA - AB) + t^2 D(t)$$

*with* $D(t) = \sum_{k,l \in \mathbb{N}} d_{kl} B^k A B^l t^{k+l}$ *for* $d_{kl} \in \mathbb{R}$.

PROOF.

$$\begin{aligned}\mathrm{Ad}_{\exp(tB)}(A) &= (I + tB + t^2 B_1(t))A(I - tB + t^2 B_2(t)) \\ &= A + t(BA - AB) + t^2 D(t)\,. \qquad \square\end{aligned}$$

As written, the adjoint representation appears to involve matrix calculation of degree $\dim_{\mathbb{R}}(L)$. On the other hand already $L \leq \mathrm{Mat}_n(\mathbb{R})$; so the next result, among other things, makes the calculation more manageable.

**(3.11).** THEOREM. *For $B \in L$, $\mathrm{Ad}_{\exp(B)} = \exp(\mathrm{ad}_B)$.*

PROOF. Clearly $t \mapsto \mathrm{Ad}_{\exp(tB)}$ is a one-parameter subgroup of $\mathrm{GL}_{\mathbb{R}}(L)$, so there is an $X \in \mathrm{End}^-_{\mathbb{R}}(L)$ with $\mathrm{Ad}_{\exp(tB)} = \exp(tX)$. By the lemma

$$\mathrm{Ad}_{\exp(tB)}(A) = (I + t\,\mathrm{ad}_B + t^2 E(t))(A)\,,$$

for $E(t) = \sum_{k,l\in\mathbb{N}} d_{kl} L_{B^k} R_{B^l} t^{k+l}$. Thus

$$X = \frac{d}{dt}\,\mathrm{Ad}_{\exp(tB)}\,|_{t=0} = \frac{d}{dt}\left(I + t\,\mathrm{ad}_B + t^2 E(t)\right)|_{t=0} = \mathrm{ad}_B\ . \qquad \square$$

**(3.12).** THEOREM. *$L$ is a Lie subalgebra of $\mathrm{Mat}^-_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R})$.*

PROOF. Let $A, B \in L$. By Lemma (3.10), for all $t \in \mathbb{R}$,

$$F(t) = A + t(BA - AB) + t^2 D(t)$$

is in the $\mathbb{R}$-space $L$. Therefore, for each nonzero $t \in \mathbb{R}$,

$$t^{-1}(F(t) - A) = (BA - AB) + tD(t)$$

is also in $L$.

The Lie algebra $L$ is a subspace of $\mathrm{Mat}_n(\mathbb{R})$ and especially is closed, hence

$$\lim_{t\to 0}\,(BA - AB) + tD(t) = [B, A]$$

is in $L$. We conclude $[A, B] = -[B, A] \in L$. $\square$

Thus we have the matrix version of Lie's first observation from the beginning of the chapter:

(i) *If $G$ is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$, then the tangent space to the identity is a Lie algebra $\Lambda(G)$.*

In this case we say that $\Lambda(G)$ is the *Lie algebra of $G$*.

## 3.6 Equivalence of representation

In this section we discuss Lie's second basic observation:

(ii) *The representation theory of the Lie group $G$ and of the Lie algebra $\Lambda(G)$ are essentially the same.*

Even in the case of closed subgroups of $\mathrm{GL}_n(\mathbb{R})$, the results are more difficult than those of the previous subsections. We offer them without proof, but see [CSM95, pp. 75-81] and [Kir08, §3.8] for nice discussions of the general results and their proofs. In the closed group case, each of [Hal15, Ros02, Sti08] proves the first two theorems of this section. Serre's notes [Ser06] contain a proof of Lie's Third Theorem, which makes use of Ado's Theorem (1.6)(b).

Theorem (3.11) could be summarized by the commutative diagram

$$\begin{tikzcd} \Lambda(G) \arrow[r, "\mathrm{ad}"] \arrow[d, "\exp"'] & \Lambda(\mathrm{GL}_{\mathbb{R}}(L)) \arrow[d, "\exp"] \\ G \arrow[r, "\mathrm{Ad}"] & \mathrm{GL}_{\mathbb{R}}(L) \end{tikzcd} .$$

The next theorem provides an important extension of this.

**(3.13).** THEOREM. *If $f\colon G \longrightarrow H$ is a Lie group homomorphism, then there is a unique Lie algebra homomorphism $df\colon \Lambda(G) \longrightarrow \Lambda(H)$ with $f\exp = \exp df$. That is, we have the following commutative diagram:*

$$\begin{tikzcd} \Lambda(G) \arrow[r, dashed, "!\, df"] \arrow[d, "\exp"'] & \Lambda(H) \arrow[d, "\exp"] \\ G \arrow[r, "f"] & H \end{tikzcd} \qquad \Box\Box$$

This is the easiest theorem of the present section. As in our proof of Theorem (3.11), the candidate for the *differential* $df$ of $f$ is relatively evident. In the matrix case, if $A \in \Lambda(G)$ then $\varphi(t) = \exp(tA)$ is a one-parameter subgroup of $G$. After we compose it with $f$, the map $f\varphi(t) = f(\exp(tA))$ is a one-parameter subgroup of $H$. Therefore there is a unique $B \in \Lambda(H)$ with $f\varphi(t) = \exp(tB)$. We set $df(A) = B$. The remaining verification (in the matrix case) that this gives a Lie algebra homomorphism is achieved through calculations similar to those of the previous two sections; see [Eld15].

artan's theorem? ]remark

A functor $F$ from the category $\mathsf{A}$ to the category $\mathsf{B}$ is an *equivalence* if it is faithful, full, and dense [Jac89]:

(i) $F$ is *faithful* if the maps $F\colon \mathrm{Hom}_{\mathsf{A}}(X,Y) \longrightarrow \mathrm{Hom}_{\mathsf{B}}(F(X),F(Y))$ are always injections.

(ii) $F$ is *full* if the maps $F\colon \mathrm{Hom}_{\mathsf{A}}(X,Y) \longrightarrow \mathrm{Hom}_{\mathsf{B}}(F(X),F(Y))$ are always surjections.

(iii) $F$ is *dense*[3] if for every object $Z$ of $\mathsf{B}$ there is an object $X$ of $\mathsf{A}$ with $F(X)$ isomorphic to $Z$ in $\mathsf{B}$.

There is a natural concept of isomorphism for categories, but a more useful equivalence relation is that of equivalence. Formally, two categories are equivalent if they have isomorphic full, faithful, and dense subcategories. One should think of category equivalence as saying that the two categories are essentially the same, although the names of the isomorphism classes may have been changed. For instance, the category of all finite sets is equivalent to the category of all finite subsets of the integers.[4] In particular, equivalent categories have the same representation theory (subject to some changing of names).

Theorem (3.13) could be restated to say that $\Lambda$ with $\Lambda(f) = df$ is a faithful functor from the category of Lie groups ${}_{\mathbb{R}}\mathsf{LieGp}$ to the category of Lie algebras ${}_{\mathbb{R}}\mathsf{LieAlg}$. The next two results say that, given appropriate restrictions, $\Lambda$ is also full and dense.

**(3.14).** Theorem. (Lie's Second Theorem) *If $G$ and $H$ are Lie groups with $G$ simply connected, then for each Lie algebra homomorphism $d\colon \Lambda(G) \longrightarrow \Lambda(H)$ there is a Lie group homomorphism $f\colon G \longrightarrow H$ with $d = df$.* □□

We must restrict to *simply connected* $G$. This is a stronger requirement than *path connectivity*, which requires that, for every group element, there is a curve containing the identity and that element. Path connectivity makes sense, since our discussion of the tangent space can only reach those elements of $G$ joined to the identity by some curve. Indeed the Lie algebra of any Lie group is equal to that of the connected component of its identity element. As we saw in Corollary (3.8), the two groups $\mathrm{GL}_n(\mathbb{R})^+$ and $\mathrm{GL}_n(\mathbb{R})$ have the same Lie algebra $\mathfrak{gl}_n(\mathbb{R})$. That is because any continuous path from the identity $I$ of positive determinant 1 to a matrix of negative determinant would have to pass through a matrix of determinant 0; the path would have to leave the group $\mathrm{GL}_n(\mathbb{R})$.

A simply connected group must be path connected but also satisfy an additional requirement, which we do not give precisely. It asserts that all paths from the identity to a given element are fundamentally the same. For example, the path connected Lie groups $(\mathbb{R},+)$ and $\mathbb{S}^1 \simeq \mathrm{SO}_2(\mathbb{R})$ have the same Lie algebra, abelian of dimension 1, but they are clearly not isomorphic. The problem is that the circle $\mathbb{S}^1$ is not simply connected—going from the identity 1 to the opposite pole $-1$ via a clockwise path is fundamentally different from traveling via a counter-clockwise path. The group $(\mathbb{R},+)$ is simply connected, so Lie's Second Theorem guarantees a Lie group homomorphism from it to $\mathbb{S}^1$, for instance

$$r \mapsto \begin{pmatrix} \cos(r) & -\sin(r) \\ \sin(r) & \cos(r) \end{pmatrix},$$

[3] This is not standard terminology.

[4] S. Awodey nicely describes category equivalence as "isomorphism up to isomorphism."

but this map has no inverse.

**(3.15).** Theorem. (Lie's Third Theorem) *For each finite dimensional Lie algebra $L$, there is a Lie group $G$ with $\Lambda(G)$ isomorphic to $L$.* □□

The new hypothesis that the Lie algebra be finite dimensional is necessary since the Lie group $G$ is a manifold.

**(3.16).** Theorem. *The functor $\Lambda$ gives a category equivalence of the category of simply connected Lie groups ${}_{\mathbb{R}}\mathsf{LieGp}^{sc}$ and the category of finite dimensional Lie algebras ${}_{\mathbb{R}}\mathsf{LieAlg}^{fd}$.*

Proof. We have already observed that Theorem (3.13) says that $\Lambda$ is faithful. By Lie's Second Theorem (3.14) it is full on ${}_{\mathbb{R}}\mathsf{LieGp}^{sc}$, and by Lie's Third Theorem (3.14) it is dense to ${}_{\mathbb{R}}\mathsf{LieAlg}^{fd}$. □

In particular, we now know that the (appropriately restricted) Lie group $G$ and Lie algebra $\Lambda(G)$ have essentially the same representation theory.

## 3.7 Problems

**(3.17).** Problem.

(a) *In $\mathrm{GL}_n(\mathbb{R})$ prove that $\exp(tA)\exp(tB) = \exp(t(A+B))$, for all $t \in \mathbb{R}$, if and only if $[A,B] = 0$.*

Hint: *The function $\exp(t(A+B)) - \exp(tA)\exp(tB)$ is smooth on $\mathbb{R}$.*

(b) *Let $A = e_{12}$ and $B = e_{23}$ be matrix units in $\mathrm{Mat}_3(\mathbb{R})$. Do the calculations in $\mathrm{SL}_3(\mathbb{R})$ and $\mathfrak{sl}_3(\mathbb{R})$ that exhibit $A + B \in \mathfrak{sl}_3(\mathbb{R})$ but $\exp(A)\exp(B) \neq \exp(A+B)$.*

Remark. *For small enough values of $t$, the smooth curve $\exp(tA)\exp(tB)$ has norm less than 1, so $\log(\exp(tA)\exp(tB))$ exists. Its precise calculation in terms of $A$ and $B$ is the content of the* Campbell-Baker-Hausdorff Theorem, *which begins*

$$\log(\exp(tA)\exp(tB)) = t(A+B) + \frac{1}{2}t^2[A,B] + t^3(\cdots)\,.$$

*As such, it also provides a proof of* (a). *Even at this level it is more sophisticated than what we have done up to now.*

**(3.18).** Problem. *Consider the group $X = \mathrm{X}_n(\mathbb{R})$ of Problem (2.10). Prove that its Lie algebra is a Heisenberg algebra isomorphic to $L_X$.*

**(3.19).** Problem. *Let $G$ be a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$. Prove that if $c : I \longrightarrow \Lambda(G)$ is a curve, differentiable on the open interval $I$, then $c'(t) \in \Lambda(G)$ for all $t \in I$.*
Hint: *Examine the proof of Theorem (3.12).*

# Part II

# Classification

# Chapter 4

# Lie Algebra Basics

The previous chapters were, in a sense, introduction and justification. The actual work starts here. We repeat our basic definition: a *Lie algebra* is a $\mathbb{K}$-algebra $({}_{\mathbb{K}}A, [\cdot,\cdot])$ that satisfies the two identical relations:

(i) $[x, x] = 0$;

(ii) (Jacobi Identity) $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$.

Our overall goals are to classify and understand Lie algebras and their representations under suitable additional hypotheses. We will focus on finite dimensional Lie algebras over algebraically closed fields of characteristic 0, but various parts of what we say are valid in a more general context. In particular, in this chapter we make no restriction on dimension or field, except where expressly noted.

## 4.1 Solvable and nilpotent Lie algebras

By definition, the subspace $A$ is an ideal of the Lie algebra $L$ precisely when it is invariant under all inner derivations $\mathrm{ad}_x$ for $x \in L$. The ideal $A$ is additionally *characteristic* in $L$ if it is invariant under all derivations of $L$, not just the inner derivations. If $J$ is a characteristic ideal of the ideal $A$ in the Lie algebra $L$, then $J$ is an ideal of $L$. (Exercise.)

For subspaces $A$ and $B$ of $L$ we let the *commutator* $[A, B]$ be the subspace of $L$ spanned by $[a, b]$ for all $a \in A$ and $b \in B$. As $[a, b] = -[b, a]$ always $[A, B] = [B, A]$.

**(4.1).** Lemma. *Let $L$ be a Lie $\mathbb{K}$-algebra.*

(a) *If $A$ and $B$ are ideals of $L$, then $[A, B]$ is an ideal of $L$.*

(b) *If $A$ and $B$ are characteristic ideals of $L$, then $[A, B]$ is a characteristic ideal of $L$.*

PROOF. For each derivation $D$, we have $D([a,b]) = [D(a),b] + [a,D(b)]$. □

Especially $[L,L]$ is a characteristic ideal, and $L/[L,L]$ is the largest abelian quotient of $L$:

**(4.2).** LEMMA. *For the Lie algebra $L$, the quotient $L/[L,L]$ is abelian. If $\varphi\colon L \longrightarrow A$ is a homomorphism from $L$ to abelian $A$, then $[L,L] \le \ker\varphi$ so that $\varphi$ factors through $L/[L,L]$ via*

$$x \mapsto x + [L,L] \mapsto \varphi(x)\,. \qquad \square$$

The *derived series* $L^{(k)}$, $k \in \mathbb{N}$, of the Lie algebra $L$ is then given by

$$L^{(0)} = L\,, \quad L^{(n+1)} = [L^{(n)}, L^{(n)}]\,.$$

By Lemma (4.1)(b) each $L^{(k)}$ is characteristic in $L$, and each subquotient $L^{(k)}/L^{(k+1)}$ is abelian. The algebra $L$ is *solvable* provided there is a $k \in \mathbb{N}$ with $L^{(k)} = 0$, in which case the smallest such $k$ is the *derived length* of $L$.

Similarly the *lower central series* $L^k$, $kin\mathbb{Z}^+$, of $L$ is given by

$$L^1 = L\,, \quad L^{n+1} = [L^n, L^n]\,.$$

Again each $L^k$ is characteristic in $L$, and each subquotient $L^k/L^{k+1}$ is abelian. The algebra $L$ is *nilpotent* provided there is a $k \in \mathbb{N}$ with $L^{(k+1)} = 0$, in which case the smallest such $k$ is the *nilpotence class* of $L$.

As examples, consider the *Borel algebra* in $\mathrm{Mat}_n(\mathbb{K})$

$$\mathfrak{b} = \mathfrak{d} \oplus \mathfrak{n}\,,$$

which is the direct sum of its abelian diagonal subalgebra $\mathfrak{d}$ and the ideal $\mathfrak{n}$ of strictly upper-triangular matrices.

$$[\mathfrak{b},\mathfrak{b}] = [\mathfrak{d},\mathfrak{d}] + [\mathfrak{d},\mathfrak{n}] + [\mathfrak{n},\mathfrak{d}] + [\mathfrak{n},\mathfrak{n}] \le \mathfrak{n}\,.$$

As seen earlier (in Problem ) the ideal $\mathfrak{n}$ is nilpotent hence solvable (by the next proposition). Therefore the Borel algebra is solvable.

**(4.3).** PROPOSITION. *Let $L$ be a Lie $\mathbb{K}$-algebra.*

(a) $[L^m, L^n] \le L^{m+n}$.

(b) $L^{(m)} \le L^{2^m}$.

(c) *If $L$ is nilpotent, then $L$ is solvable.*

PROOF. Part (a) follows from induction on $n$, with $n = 1$ given by the definition of $L^{m+1}$. The induction step comes from combining the Jacobi Identity with the definition (twice) and induction (twice):

$$\begin{aligned}[L^m, L^{n+1}] &= [L^m, [L^n, L]] \\ &\leq [L^n, [L, L^m]] + [L, [L^m, L^n]] \\ &\leq [L^n, L^{m+1}] + [L, L^{m+n}] \\ &\leq L^{n+m+1} + L^{1+m+n} \\ &\leq L^{m+n+1}\,.\end{aligned}$$

Now (b) follows from (a) and (c) follows from (b). □

**(4.4).** LEMMA. *Let $L$ be a Lie $\mathbb{K}$-algebra.*

(a) *Subalgebras and quotient algebras of solvable $L$ are solvable.*

(b) *The sum of solvable ideals in $L$ is a solvable ideal of $L$*

(c) *If $\dim_{\mathbb{K}}(L)$ is finite, then $L$ has a unique maximal solvable ideal.*

(d) *If the ideal $I$ and the quotient $L/I$ are solvable, then $L$ is solvable.* □

**(4.5).** LEMMA. *Let $L$ be a Lie $\mathbb{K}$-algebra.*

(a) *Subalgebras and quotient algebras of nilpotent $L$ are nilpotent.*

(b) *The sum of nilpotent ideals in $L$ is a nilpotent ideal of $L$.*

(c) *If $\dim_{\mathbb{K}}(L)$ is finite, then $L$ has a unique maximal nilpotent ideal.* □

As is always true for finite dimensional $L$, when $L$ has a unique maximal solvable ideal, then it is the *radical* or *solvable radical* of $L$. Similarly when $L$ has a unique maximal nilpotent ideal, it is the *nilpotent radical* of $L$.

By the proposition, the nilpotent radical is contained in the solvable radical. On the other hand, the last term in the derived series of a solvable ideal is an abelian ideal and so is nilpotent. Therefore the solvable radical is 0 if and only if the nilpotent radical is 0.

A Lie algebra is *semisimple* if its (solvable) radical is 0. By Lemma (4.4)(d) the quotient of $L$ by its radical is then always semisimple.

The last part of the lemma on solvable extensions does not have a nilpotent counterpart—the extension of a nilpotent Lie algebra by a nilpotent Lie algebra need not be nilpotent. (Otherwise, all solvable Lie algebras would also be nilpotent.) In the following proposition we do get a nilpotent extension algebra under an additional *Engel condition*, requiring the vanishing of an appropriate iterated commutator. Define $[A; B, n]$ by $[A; B, 1] = [A, B]$ and $[A; B, k+1] = [[A; B, k], B]$.

**(4.6).** PROPOSITION. *Let the Lie algebra $L$ contain an ideal $I$ such that $I$ and $L/I$ are nilpotent. Further assume $L$ has a subalgebra $M$ such that $L = I + M$. Then $L$ is nilpotent if and only if there is a positive $m$ with $[I; M, m] = 0$.*

PROOF. See [Ste70, Lemma 2.1].

If $L$ is nilpotent, then letting $m$ be the class of $L$ gives the required condition.

Now consider the converse. We first claim that for all positive $n$ and $r$

$$[I^n; L, r] \leq I^{n+1} + [I^n; M, r].$$

As $I^n$ is characteristic in the ideal $I$, it is an ideal of $L$; so at least we have $[I^n; L, r] \leq I^n$.

We prove the claim by induction on $r$, the result being clear for $r = 1$ as $L = I + M$. Assume the result for $r$. Then

$$\begin{aligned}
[I^n; L, r+1] &= [[I^n; L, r], L] \\
&\leq [I^{n+1} + [I^n; M, r], I + M] \\
&\leq [I^{n+1}, I + M] + [[I^n; M, r], I] + [[I^n; M, r], M].
\end{aligned}$$

The first two summands are in $I^{n+1}$ (as $I^{n+1}$ and $I^n$ are ideals of $L$) and the last is equal to $[I^n; M, r+1]$. This gives the claim.

Let $k$ be the maximum of $m$ and the nilpotence class of $L/I$. We prove $L^{kn} \leq I^n$ by induction on $n$, with the case $n = 1$ valid by the definition of $k$. By definition, induction, the claim, and hypothesis

$$L^{kn+k} = [L^{kn}; L, k] \leq [I^n; L, k] \leq I^{n+1} + [I^n; M, k] \leq I^{n+1} + [I; M, k] = I^{n+1},$$

as desired.

For large enough $n$, nilpotent $I$ has $I^n = 0$. Thus $L^{kn} = 0$, and $L$ is nilpotent. □

## 4.2 Representation and modules

that a $\mathbb{K}$-*representation* of the Lie $\mathbb{K}$-algebra $L$ is a Lie homomorphism $\varphi \colon L \longrightarrow A^-$ for some associative $\mathbb{K}$-algebra $A$. Often $\varphi \colon L \longrightarrow \mathrm{End}^-_{\mathbb{K}}(V)$ for a $\mathbb{K}$-space $V$. The space $V$ is then an $M$-*module* and the action $L \times V \longrightarrow V$ of $L$ on that module is characterized by

$$[a, b]v = \varphi([a, b])(v) = \varphi(a)(\varphi(b)(v)) - \varphi(b)(\varphi(a)(v)) = a(bv) - b(av)$$

for $a, b \in L$ and $v \in V$.

The modules for the Lie algebra $L$ form a category ${}_L\mathsf{Mod}$. There is a natural bridge between this context and that of module categories over associative algebras.

Let $V$ be a $\mathbb{K}$-space. For $f \in \mathbb{N}$, let $V^{\otimes f}$ be the $f^{\mathrm{th}}$ tensor power of the module $V$ (with $V^{\otimes 0} = \mathbb{K}$ and $V^{\otimes 1} = V$). The *tensor algebra* $\mathrm{T}(V)$ is the associative $\mathbb{K}$-algebra

$$\mathrm{T}(V) = \bigoplus_{n \in \mathbb{N}} V^{\otimes n}$$

with multiplication determined by the linear extension of

$$\mu(v_1 \otimes \cdots \otimes v_k, w_1 \otimes \cdots \otimes w_m) = v_1 \otimes \cdots \otimes v_k \otimes w_1 \otimes \cdots \otimes w_m .$$

If $V$ happens to be the Lie algebra $L$ over $\mathbb{K}$, then its *universal enveloping algebra* $\mathrm{U}(L)$ is the quotient $\mathrm{T}(L)/I$ where $I$ is the ideal in $\mathrm{T}(L)$ generated by all the elements $x \otimes y - y \otimes x - [x, y]$ for $x, y \in L$. The construction gives us a natural representation $\iota_L \colon L \longrightarrow \mathrm{U}(L)^-$. Especially $\mathrm{U}(L)$-modules are $L$-modules. This correspondence is readily seen to be universal in at least two senses. (See Chapter 9 for further discussion.)

**(4.7).** THEOREM.

(a) *If for associative $A$ the map $\varphi \colon L \longrightarrow A^-$ is a representation of $L$, then there is a unique associative algebra morphism $\varphi_A \colon \mathrm{U}(L) \longrightarrow A$ with $\varphi = \varphi_A \iota_L$.*

(b) *The two module categories ${}_L\mathsf{Mod}$ and ${}_{\mathrm{U}(L)}\mathsf{Mod}$ are isomorphic.* □

That is, $L$-modules and $\mathrm{U}(L)$-modules are the same thing.

One advantage is immediate. For associative algebras $A$, every cyclic $A$-module is a quotient of ${}_AA$. As irreducible modules are always cyclic, every cyclic and irreducible $L$- and $\mathrm{U}(L)$-module is a quotient of $\mathrm{U}(L)$. This is an improvement. For instance in Section 2.4.1 we found that 3-dimensional $\mathfrak{sl}_2(\mathbb{K})$ has irreducible modules of arbitrarily large finite dimension. (Among other things, this implies that the universal enveloping algebra for $\mathfrak{sl}_2(\mathbb{K})$ is infinite dimensional.) This also demonstrates a downside—the algebra $\mathfrak{sl}_2(\mathbb{K})$ is very small and manageable, while we have just seen that its universal algebra has numerous distinct quotients.

Large portions of standard terminology and results module theory for associative algebras go over to Lie algebras effortlessly via the universal enveloping algebra. For instance, two $L$-modules are isomorphic precisely when they are isomorphic as $\mathrm{U}(L)$-modules.

The $L$-submodules and quotient modules of an $L$-module $V$ over $\mathbb{K}$ will be exactly those spaces that are $\mathrm{U}(L)$-submodules and quotient modules. A module will then be *irreducible* or *simple* if it has no proper and nontrivial $L$-submodule, which is to say $\mathrm{U}(L)$-submodule. The dual of the $L$-module $V$ is its dual $V^*$ as $\mathrm{U}(L)$-module.

Finite dimensional $L$-modules have *composition series*, and these satisfy the Jordan-Hölder Theorem, in that the multiset of composition factors is the same for any two composition series of the module.

Indecomposable modules for $L$ are those for $\mathrm{U}(L)$. The Krull-Schmidt Theorem remains valid, saying that any two decompositions into indecomposable summands are essentially the same.

The concept of *complete reducibility* for modules remains relevant:

**(4.8).** THEOREM. *Let $V$ be a finite dimensional module for the Lie algebra $L$. Then the following are equivalent:*

(1) *for every submodule $W$ of $V$, there is a complementary submodule $W'$ with $V = W \oplus W'$;*

(2) *$V$ is a sum of irreducible submodules;*

(3) *$V$ is a direct sum of irreducible submodules.* $\square$

Particularly important is Schur's Lemma:

**(4.9).** Theorem. (Schur's Lemma) *Let $L$ be a Lie algebra and $V$ a finite dimensional, irreducible $L$-module over the algebraically closed field $\mathbb{K}$. Then the scalars of $\mathbb{K}$ are the only endomorphisms of $V$ that commute with the action of $L$ on $V$.* $\square$

## 4.3 Semidirect products

The next lemma describes the elementary *internal semidirect product* for Lie algebras. The corresponding *external semidirect product* or *split extension* of Lie algebras is then the construction of the proposition that follows.

**(4.10).** Lemma. *Let Lie $\mathbb{K}$-algebra $L = M \oplus I$ where $M$ is a subalgebra and $I$ is an ideal. Then for $m, n \in M$ and $i, j \in I$ we have*

$$[m + i, n + j] = [m, n] + [i, j] + [m, j] + [i, n]\,,$$

*where $[m, n] \in M$ and $[i, j] + [m, j] + [i, n] = [i, j] + [m, j] - [n, i] \in I$.* $\square$

**(4.11).** Proposition. *Let $M$ and $I$ be Lie $\mathbb{K}$-algebras, and let $\delta\colon M \longrightarrow \mathrm{Der}_{\mathbb{K}}(I)$ be a Lie homomorphism of $M$ into the derivation algebra of $I$ given by $m \mapsto \delta_m$. Then $M \oplus I$ with bracket multiplication*

$$[(m, i), (n, j)] = ([m, n], [i, j] + \delta_m(j) - \delta_n(i))$$

*is a Lie $\mathbb{K}$-algebra in which $0 \oplus I$ is an ideal isomorphic to $I$ and $M \oplus 0$ is a subalgebra isomorphic to $M$. Furthermore, for each $m \in M$, $\mathrm{ad}_{(m,0)}$ induces $(0, \delta_m)$ on $0 \oplus I$.*

Proof. Exercise: the only difficulty is the verification of the Jacobi Identity. In doing that, the corresponding calculation from the lemma can be used as a guide. $\square$

We emphasize two cases.

**(4.12).** Example.

(a) *If $\delta$ is a derivation of the Lie algebra $A$, then with $M = \mathbb{K}\delta$ and $I = A$ we make $L = \mathbb{K}\delta \oplus A$ into a Lie algebra as in the proposition. Here $A$ is an ideal of codimension $1$ upon which the derivation $\delta$ is now induced by the inner derivation $\mathrm{ad}_\delta$ of the new algebra $L$.*

(b) *Let $V$ be a module for the Lie algebra $M$. As in the proposition $L = M \oplus V$ becomes a Lie algebra after we declare $V(= I)$ to be an abelian Lie algebra: $[V, V] = 0$. (Any endomorphism of an abelian Lie algebra is a derivation by Proposition (2.3).)*

The original definition of modules is extrinsic—the module arrives as an accessory to a representation. But modules often arise intrinsically—the module itself is the focal point and gives rise to a representation. The most basic of such situations is the adjoint action of $M$ on itself. This generalizes to the action of $M$ on the subquotient $A/B$ where $A$ and $B$ are ideals of $M$ with $A \geq B$. The theory of solvable and nilpotent algebras in the previous section comes from the study of abelian subquotients.

These remarks and the second example above suggest some useful notation. Let $\varphi\colon M \longrightarrow \mathrm{End}^-_{\mathbb{K}}(V)$ be a Lie representation of $M$. As the action of $\varphi$ on $V$ is induced in $L = M \oplus V$ by the restriction of $L$-adjoint action to the ideal $0 \oplus V$, we sometimes write $\mathrm{ad}^V$ for $\varphi$ and $\mathrm{ad}^V_x$ for $\varphi(x)$. In particular $\mathrm{ad}^L_x$ is the usual adjoint action $\mathrm{ad}_x$ of $x$ on $L$ in the adjoint representation.

## 4.4 Problems

**(4.13).** PROBLEM. *?? lower central series;* Z *is kernel of* ad

**(4.14).** PROBLEM. *For $L = \mathfrak{n}$ calculate $L^k$ and $L^{(k)}$.*

**(4.15).** PROBLEM. *Field indep* $\exp(\delta)$ *auto; see [Ros02, p. 51].*

**(4.16).** PROBLEM. *nilpotent derivations and automorphisms*

**(4.17).** PROBLEM. *Jordan-Chevalley decomposition*

**(4.18).** PROBLEM. *Lie algebra central extension.*

**(4.19).** PROBLEM. *Action of $L$ on $V \otimes_{\mathbb{K}} W$.*

**(4.20).** PROBLEM. *Action of $L$ on $V^*$, given action on $V$.*

Chapter 5

# The Cartan decomposition

## 5.1 Engel's Theorem and Cartan subalgebras

We begin with a helpful calculation.

**(5.1).** Proposition. *Let $\delta$ be a derivation of the Lie algebra $L$. For $x, y \in L$ and $a, b \in \mathbb{K}$:*

$$(\delta - a1 - b1)^n[x, y] = \sum_{i=0}^{n} \binom{n}{i} [(\delta - a1)^{n-i}(x), (\delta - b1)^i(y)] \,.$$

Proof. We prove this by induction on $n$ with the case $n = 0$ being trivial and the case $n = 1$ following from the definition of a derivation.

$$\begin{aligned}
(\delta - a1 - b1)^n[x, y] &= (\delta - a1 - b1)((\delta - a1 - b1)^{n-1}[x, y]) \\
&= (\delta - a1 - b1) \sum_{i=0}^{n-1} \binom{n-1}{i} [(\delta - a1)^{n-1-i}(x), (\delta - b1)^i(y)] \\
&= \sum_{i=0}^{n-1} \binom{n-1}{i} \delta[(\delta - a1)^{n-1-i}(x), (\delta - b1)^i(y)] \\
&\quad + (-a1 - b1) \sum_{i=0}^{n-1} \binom{n-1}{i} [(\delta - a1)^{n-1-i}(x), (\delta - b1)^i(y)] \\
&= \sum_{i=0}^{n-1} \binom{n-1}{i} [\delta(\delta - a1)^{n-1-i}(x), (\delta - b1)^i(y)] \\
&\quad + \sum_{i=0}^{n-1} \binom{n-1}{i} [(\delta - a1)^{n-1-i}(x), \delta(\delta - b1)^i(y)]
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=0}^{n-1}\binom{n-1}{i}[-a(\delta-a1)^{n-1-i}(x),(\delta-b1)^i(y)]\\
&+\sum_{i=0}^{n-1}\binom{n-1}{i}[(\delta-a1)^{n-1-i}(x),-b(\delta-b1)^i(y)]\\
=&\sum_{i=0}^{n-1}\binom{n-1}{i}[(\delta-a1)^{n-i}(x),(\delta-b1)^i(y)]\\
&+\sum_{i=0}^{n-1}\binom{n-1}{i}[(\delta-a1)^{n-1-i}(x),(\delta-b1)^{i+1}(y)]\\
=&\sum_{i=0}^{n-1}\binom{n-1}{i}[(\delta-a1)^{n-i}(x),(\delta-b1)^i(y)]\\
&+\sum_{j=1}^{n}\binom{n-1}{j-1}[(\delta-a1)^{n-j}(x),(\delta-b1)^j(y)]\\
=&\sum_{k=0}^{n}\binom{n}{k}[(\delta-a1)^{n-k}(x),(\delta-b1)^k(y)]\,. \qquad \square
\end{aligned}
$$

**(5.2).** PROPOSITION. *Let $N$ be a nilpotent Lie algebra and $V$ a $\mathbb{K}$-module. For each element $x$ of $N$ and each $\lambda \in \mathbb{K}$, the generalized eigenspace*

$$V_x^\lambda = \{\, v \in V \mid (x-\lambda 1)^k v = 0\,, \textit{some } k = k_{x,\lambda,v} \in \mathbb{N}\,\}$$

*for $x$ on $V$ is an $N$-submodule of $V$.*

PROOF. Let $N$ have nilpotence class $l$. For $v \in V_x^\lambda$ set $n = l + k_{x,\lambda,v}$. As in Example (4.12)(b), we calculate within the semidirect product $L = N \oplus V$ with $N$ acting on $V$. By Proposition (5.1) with $y \in N$, $\delta = \mathrm{ad}_x$, $a = 0$, and $b = \lambda$,

$$
\begin{aligned}
(x-\lambda 1)^n(yv) &= (\mathrm{ad}_x - \lambda 1)^n[y,v]\\
&= \sum_{i=0}^{n}\binom{n}{i}[\mathrm{ad}_x^{n-i}(y),(\mathrm{ad}_x-\lambda 1)^i(v)] = 0\,,
\end{aligned}
$$

since $\mathrm{ad}_x^j(y) = 0$ for $j \geq l$ and $(\mathrm{ad}_x - \lambda 1)^i(v) = 0$ for $i \geq k_{x,\lambda,v}$.

This shows that $yV_x^\lambda \leq V_x^\lambda$, hence the subspace $V_x^\lambda$ of $V$ is in fact a submodule. $\square$

An endomorphism is *nil* if some power of it is 0, and a *nil representation* of the Lie algebra $N$ is one in which each element of $N$ acts as a nil endomorphism.

**(5.3).** PROPOSITION. *If $\sigma\colon N \longrightarrow \mathrm{End}_{\mathbb{K}}^-(V)$ is a nil irreducible representation of the nilpotent Lie algebra $N$, then $\sigma$ is the trivial 1-dimensional representation.*

PROOF. Certainly $N^n V = 0$, where $n$ is the nilpotence class of $N$. Suppose $N^k V = 0$. If $k = 1$, then $NV = 0$ and irreducible $V$ has dimension 1, as desired.

For $k > 1$ let $x \in N^{k-1}$. As the representation is nil, for nonzero $u \in V$ there is a positive $n_u$ with $x^{n_u} u = 0$. For minimal such $n_u$, the element $w = x^{n_u - 1} u$ is nonzero with $xw = 0$. Thus $W = \{ v \in V \mid xv = 0 \}$ is nonzero. For all $y \in N$ and $w \in W$

$$x(yw) = y(xw) - [x, y]w = 0 - 0 = 0$$

as $w \in W$ and $[x, y] \in N^k$. Therefore $yw \in W$, which is thus a nonzero submodule. By irreducibility $W = V$, hence $xV = 0$. But this implies $N^{k-1}V = 0$, and we are done. □

**(5.4).** Corollary. *If $\sigma$ is a nil representation of the nilpotent Lie algebra $N$ on the finite dimensional module $V$, then $V$ has an $N$-composition series with all factors of dimension 1 and trivial.* □

**(5.5).** Theorem. (Engel's Theorem) *If the adjoint representation of the finite dimensional Lie algebra $L$ is nil, then $L$ is nilpotent.*

Proof. We prove this by induction on $\dim_{\mathbb{K}}(L)$ with the result clearly true in dimensions 0 and 1. Assume $\dim_{\mathbb{K}}(L) \geq 2$.

Let $I$ be a maximal proper subalgebra of $L$. As $\mathrm{ad}_x^I = \mathrm{ad}_x^L |_I$ for $x \in I$, the adjoint representation of $I$ is nil. Therefore by induction $I$ is nilpotent.

By Corollary (5.4) there is a 1-dimensional submodule $P/I$ for the nil action of nilpotent $I$ on $L/I$. Let $x \in P \setminus I$ and $M = \mathbb{K}x$. Then

$$\begin{aligned} [P, P] &= [M + I, M + I] \\ &= [M, M] + [M, I] + [I, M] + [I, I] \\ &= [M, I] + [I, M] + [I, I] \\ &\leq I\,, \end{aligned}$$

so $P$ is a subalgebra of $L$ in which $I$ is an ideal of codimension 1. By maximality of $I$, $L = P$.

We now have $L = M \oplus I$ with $M = \mathbb{K}x \simeq L/I$ an abelian algebra and $I$ a nilpotent ideal. Furthermore by hypothesis $\mathrm{ad}_x^m = 0$ for some $m$, hence $[I; M, m] = [I; x, m] = 0$. By Proposition (4.6), the algebra $L$ is nilpotent. □

If $A$ is a subspace of the Lie algebra $L$, then the *normalizer* of $A$ in $L$, $\mathrm{N}_L(A)$, is $\{ x \in L \mid [x, A] \leq A \}$. The subspace $A$ is then *self-normalizing* if $A = \mathrm{N}_L(A)$.

**(5.6).** Lemma.

(a) *If $A$ is a subspace of the Lie algebra $L$, then $\mathrm{N}_L(A)$ is a subalgebra.*

(b) *If $A$ is a self-normalizing subspace of the Lie algebra $L$, then $A$ is a subalgebra.*

Proof. For $x, y \in \mathrm{N}_L(A)$ and $a \in A$, the Jacobi Identity gives

$$[[x, y], a] = -[[y, a], x] - [[a, x], y] \in A\,,$$

so the vector space $\mathrm{N}_L(A)$ is a subalgebra. The second part then follows from the first. □

**(5.7).** LEMMA. *Let $L$ be a Lie algebra, $x \in L$, and*

$$L_x^0 = \{\, y \mid \mathrm{ad}_x^k(y) = 0 \textit{ for some } k = k_{x,0,y} \in \mathbb{N} \,\}$$

*be the generalized eigenspace for $x$ acting on $L$ in the adjoint representation with eigenvalue $0$. Then $L_x^0$ is a self-normalizing subalgebra of $L$.*

PROOF. Let $a \in \mathrm{N}_L(L_x^0)$. As $[x,x]=0$, we have $x \in L_x^0$, hence $[x,a] \in L_x^0$. Therefore $\mathrm{ad}_x^k([x,a]) = 0$ for $k = k_{x,0,[x,a]}$. But then $\mathrm{ad}_x^{k+1}(a) = \mathrm{ad}_x^k([x,a]) = 0$, whence $a \in L_x^0$m, and the subspace $L_x^0$ is self-normalizing. By the previous lemma it is then a subalgebra. □

The element $w$ of the finite dimensional Lie algebra is said to be *regular* in $L$ if the dimension of the subalgebra $L_w^0$ is minimal. This dimension is then the *rank* of $L$. As long as $L \neq 0$ this is positive since always $w \in L_w^0$.

**(5.8).** THEOREM. *Assume $\mathbb{K}$ has characteristic $0$. Let $w$ be a regular element of the finite dimensional Lie algebra $L$ and set $H = L_w^0$. Then $H$ is a nilpotent and self-normalizing subalgebra of $L$.*

PROOF. We follow [Eld15].

By the previous lemma, $H$ is a self-normalizing subalgebra. We must prove it to be nilpotent. Let $n$ be the dimension of $L$ and $r = \dim_{\mathbb{K}}(H)$, its rank.

For fixed but arbitrary $h \in H$ and all $\alpha \in \mathbb{K}$, the element $w + \alpha h$ belongs to $H$. Consider the linear transformation $\mathrm{ad}_{w+\alpha h} (= \mathrm{ad}_w + \alpha\, \mathrm{ad}_h)$ of $L$, which leaves the subspace $H$ invariant and so also acts on the quotient space $L/H$. Its characteristic polynomial $\chi^L_{w+\alpha h}(z) \in \mathbb{K}[z]$ is then $\varphi_\alpha(z)\gamma_\alpha(z)$ where

$$\varphi_\alpha(z) = \chi^H_{w+\alpha h}(z) = z^r + \sum_{i=0}^{r-1} f_i(\alpha) z^i$$

is the characteristic polynomial of $\mathrm{ad}_{w+\alpha h}$ on $H$ and

$$\gamma_\alpha(z) = \chi^{L/H}_{w+\alpha h}(z) = z^{n-r} + \sum_{j=0}^{n-r-1} g_j(\alpha) z^j$$

is the characteristic polynomial of $\mathrm{ad}_{w+\alpha h}$ on $L/H$. The standard calculation of the characteristic polynomial as a determinant reveals the polynomials $f_i(x)$ of $\mathbb{K}[x]$ to have degree at most $r$ while the $g_j(x)$ have degree at most $n-r$ .

As $H = L_w^0$ we have $\gamma_0(0) \neq 0$ hence $g_0(0) \neq 0$. Especially the polynomial $g_0(x)$ of degree at most $n-r$ is not identically 0. As $\mathbb{K}$ has characteristic 0 we have $|\mathbb{K}| > n$, so there are distinct elements $\alpha_1, \dots, \alpha_{r+1}$ of $\mathbb{K}$ with $g_0(\alpha_k) \neq 0$ for $1 \leq k \leq r+1$. In particular $L^0_{w+\alpha_k h} \leq H$ for each $k$. As $w$ is regular, this forces $L^0_{w+\alpha_k h} = H$, which is to say $\varphi_{\alpha_k}(z) = z^r$ for $1 \leq k \leq r+1$. But then each of the polynomials $f_i(x)$, for $0 \leq i < r$, vanishes at $\alpha_1, \dots, \alpha_{r+1}$. As these polynomials have degree at most $r$, they must be identically 0.

Therefore $\varphi_\alpha(z) = z^r$ for all values of $\alpha \in \mathbb{K}$, and every $w + \alpha h$ is nil on $H$. As $h$ was fixed but arbitrary, we find that every element of $H$ is nil on $H$. By Engel's Theorem (5.5), $H$ is nilpotent as desired. $\square$

A *Cartan subalgebra* of the Lie algebra $L$ is a nilpotent, self-normalizing subalgebra. The theorem tells us that Cartan subalgebras always exist in finite dimension and characteristic 0. More is true: for finite dimensional Lie algebras over algebraically closed fields of characteristic 0, the automorphism group of $L$ is transitive on the Cartan subalgebras (so all arise as in the theorem); see [Jac79, p. 273]. At times we may abuse notation or terminology by not mentioning the specific Cartan subalgebra being used. Indeed often they are all essentially equivalent; we shall address conjugacy of Cartan subalgebras of finite dimensional semisimple algebras in Corollary (8.24).

There are many characterizations of Cartan subalgebras. The following is important here.

**(5.9).** PROPOSITION. *Suppose $H$ is a nilpotent subalgebra of the finite dimensional Lie algebra $L$. Then $H$ is a Cartan subalgebra if and only if in the action of $H$ on $L$ via the adjoint we have*

$$\begin{aligned} H = L_H^0 &= \bigcap_{h \in H} L_h^0 \\ &= \{\, x \in L \mid \mathrm{ad}_h^k(x) = 0 \textit{ for all } h \in H \textit{ and some } k = k_{h,0,x} \in \mathbb{N} \,\}, \end{aligned}$$

*the largest subspace of $L$ upon which $H$ is nil.*

PROOF. The nilpotent algebra $H$ is certainly contained in $L_H^0$. We show that $H$ is proper in $L_H^0$ if and only if $H$ is not self-normalizing. As the Cartan subalgebras are by definition the self-normalizing nilpotent subalgebras, this will give the result.

Let $x \in \mathrm{N}_L(H) \setminus H$. Then, for each $h \in H$ we have $[h, x] \in H$. As $H$ is nilpotent, $\mathrm{ad}_h^k[h, x]$ is 0 for sufficiently large $k = k_h$. But then $\mathrm{ad}_h^{k+1}(x) = 0$ and $x$ is in $L_H^0$ but not in $H$.

Suppose $L_H^0 > H$. By Corollary (5.4), there is a trivial $H$-submodule $P/H$ of dimension 1 in $L_H^0/H$. For $x \in P \setminus H$, we have $[x, H] \le H$. That is, $x$ is in the normalizer of $H$ but not in $H$. $\square$

## 5.2 Weight spaces and vectors

**(5.10).** THEOREM. *Assume $\mathbb{K}$ is algebraically closed of characteristic 0. Let $V$ be an indecomposable $\mathbb{K}N$-module for the nilpotent Lie algebra $N$ with $0 < n = \dim_{\mathbb{K}}(V)$. Then there is a 1-dimensional Lie homomorphism $\lambda\colon N \longrightarrow \mathbb{K}$ with*

$$V = \{\, v \in V \mid (x - \lambda(x)1)^n v = 0 \textit{ for all } x \in N \,\}.$$

PROOF. We may replace $N$ with its image in $\mathrm{End}_{\mathbb{K}}^-(V) \simeq \mathrm{Mat}_n^-(\mathbb{K})$. As $\mathbb{K}$ is algebraically closed, all $x \in N$ have eigenvalues in their action on $V$.

By standard linear algebra (say, Jordan Canonical Form), for each $x \in N$ the module $V$ is the direct sum of its generalized eigenspaces

$$V_x^\lambda = \{\, v \in V \mid (x - \lambda 1)^k v = 0\,, \text{some } k = k_{x,\lambda,v} \in \mathbb{N}\,\}\,.$$

Indeed $\max_v(k_{x,\lambda,v}) \le n$, so

$$V_x^\lambda = \{\, v \in V \mid (x - \lambda 1)^n v = 0\,\}\,.$$

By Proposition (5.2), indecomposability, and the above remarks, each $x \in N$ has a unique eigenvalue $\lambda(x)$ on $V$, and for every $x$ the whole space $V$ is equal to the generalized $x$-eigenspace $V_x^{\lambda(x)}$:

$$V = V_x^{\lambda(x)} = \{\, v \in V \mid (x - \lambda(x)1)^n v = 0\,\}\,.$$

In particular $\mathrm{tr}(x) = n\lambda(x)$. As $\mathbb{K}$ has characteristic 0, we find that $\lambda(x) = n^{-1}\mathrm{tr}(x)$ is a linear map $\lambda\colon N \longrightarrow \mathbb{K}$. Furthermore

$$\lambda([x,y]) = n^{-1}\mathrm{tr}(xy - yx) = 0\,,$$

for all $x, y \in N$; that is, $\lambda|_{[N,N]} = 0$. Therefore the linear transformation $\lambda\colon N \longrightarrow \mathbb{K}$ is a 1-dimensional representation of the abelian Lie algebra $N/[N,N]$ and so of $N$ itself. □

A 1-dimensional representation of a Lie algebra $L$ is called a *weight* of the algebra. All weights of $L$ belong to the dual of the $\mathbb{K}$-space $L/[L,L]$. For an $L$-module $V$ and weight $\lambda$ of $L$,

$$\begin{aligned} V^\lambda = V_L^\lambda &= \bigcap_{x \in L} V_x^{\lambda(x)} \\ &= \{\, v \in V \mid (x - \lambda(x)1)^k v = 0 \text{ for all } x \in L \text{ and some } k = k_{x,\lambda,v} \in \mathbb{N}\,\} \end{aligned}$$

is the corresponding *generalized weight space* in $V$. These are the generalized eigenspaces for the action of $L$. A nonzero vector $v \in V^\lambda$ is a *weight vector* if it is an actual eigenvector for all $L$ ($k_{x,\lambda,v} \le 1$ for all $x \in L$). The corresponding eigenspace of weight vectors is then the *weight space*[1]

$$V_\lambda = V_{L,\lambda} = \bigcap_{x \in L} V_{x,\lambda} = \{\, v \in V \mid (x - \lambda(x)1)v = 0 \text{ for all } x \in L\,\}\,.$$

Notice that $V_\lambda$ is nonzero if and only if $V^\lambda$ is nonzero.

For every nonzero Lie algebra, the trivial representation is the *trivial weight* or *zero weight*. We have already encountered a generalized weight space in Proposition (5.9), where the Cartan subalgebra $H$ was characterized among all nilpotent subalgebras of $L$ by being equal to its corresponding generalized weight space $L_H^0$.

[1] In the older literature it is often the generalized eigenspaces $V_L^\lambda$ that are termed weight spaces.

**(5.11).** THEOREM. *Assume $\mathbb{K}$ is algebraically closed of characteristic* 0. *For the nilpotent Lie algebra $N$ and the $N$-module $V$ of finite dimension $n$, $N$ has only finitely many weights on $V$; each generalized weight space*

$$V_N^\lambda = V^\lambda = \{\, v \in V \mid (x - \lambda(x)1)^n v = 0 \textit{ for all } x \in N \,\}$$

*is a submodule; and $V$ is the direct sum of its generalized weight spaces.*

PROOF. As $V$ is finite dimensional, we can write $V$ as a direct sum of finitely many nonzero indecomposable submodules. By the previous theorem, each of these summands is contained in one of the the generalized weight spaces $V^\mu$ for some weight $\mu$ of $N$. Let the submodule $V(\mu)$ be the sum of those indecomposable summands with weight $\mu$. The previous theorem gives

$$V(\mu) \le \{\, v \in V \mid (x - \mu(x)1)^n v = 0 \text{ for all } x \in N \,\} \le V^\mu\,.$$

Now we have

$$V = \bigoplus_{\mu \in J} V(\mu)\,,$$

where $J$ is a finite set of weights for $N$ on $V$. In particular, every $v \in V$ can be uniquely written $v = \sum_{\mu \in M} v_\mu$ with $v_\mu \in V(\mu)$.

Let $\lambda$ be an arbitrary weight of $N$ on $V$, and consider $0 \ne v \in V^\lambda$. We claim:

$$v_\mu \ne 0 \implies \mu = \lambda\,.$$

As the various nonzero $v_\mu$ are linearly independent and each $V(\mu)$ is a submodule, $(x - \lambda(x)1)^m v = 0$ implies $(x - \lambda(x)1)^m v_\mu = 0$ and so $v_\mu \in V^\lambda \cap V(\mu) \le V^\lambda \cap V^\mu$.

Assume $v_\mu \ne 0$. For fixed but arbitrary $x \in N$, choose $k\,(= k_{x,\lambda(x),v}) \in \mathbb{N}$ minimal with $(x - \lambda(x)1)^k v_\mu = 0$. Set $u = (x - \lambda(x)1)^{k-1} v_\mu \ne 0$, so that $(x - \lambda(x))u = 0$; that is, $xu = \lambda(x)u$. As $V(\mu)$ is a submodule, $u \in V(\mu) \le V^\mu$; so there is an $m \in \mathbb{Z}^+$ with $(x - \mu(x)1)^m u = 0$. But

$$(x - \mu(x)1)u = xu - \mu(x)u = \lambda(x)u - \mu(x)u = (\lambda(x) - \mu(x))u\,,$$

hence

$$0 = (x - \mu(x)1)^m u = (\lambda(x) - \mu(x))^m u\,.$$

Now $u \ne 0$ forces $\lambda(x) - \mu(x) = 0$. That is, for all $x \in N$ we have $\lambda(x) = \mu(x)$, hence $\lambda = \mu$ as claimed.

For every weight $\lambda$, each nonzero $v \in V^\lambda$ must project nontrivially onto at least one of the summands $V(\mu)$ for $\mu \in M$. By the claim, there is only one such summand, namely $V(\lambda)$, and $v \in V(\lambda)$. Thus $\lambda \in J$ and there are only finitely many weights for $N$ on $V$. Also $V^\lambda \le V(\lambda) \le V^\lambda$, hence

$$V(\lambda) = \{\, v \in V \mid (x - \lambda(x)1)^n v = 0 \text{ for all } x \in N \,\} = V^\lambda\,.$$

Finally $V$ is the direct sum of the submodules $V(\mu)$, so it is equally well the direct sum of the generalized weight spaces $V^\lambda$, each a submodule. $\square$

## 5.3 The Cartan decomposition

We can use the results of the previous sections to consider a Lie algebra as a module for any of its nilpotent subalgebras.

**(5.12).** THEOREM. *Let $L$ be a finite dimensional Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic $0$. Let $\alpha$ and $\beta$ be weights of $L$ for the nilpotent subalgebra $N$. Then the generalized weight spaces satisfy*

$$[L_N^\alpha, L_N^\beta] \leq L_N^{\alpha+\beta}\,,$$

*where $L_N^\lambda$ for $\lambda \in (N/[N,N])^*$ is taken to be $0$ when $\lambda$ is not a weight. For the corresponding weight spaces we have*

$$[L_{N,\alpha}, L_{N,\beta}] \leq L_{N,\alpha+\beta}\,.$$

PROOF. Let $x \in N$, $y \in L_N^\alpha$, and $z \in L_N^\beta$. Then, for $n = 2\dim_{\mathbb{K}}(N)$, by Proposition (5.1) and Theorem (5.11)

$$(\mathrm{ad}_x - \alpha 1 - \beta 1)^n [y,z] = \sum_{i=0}^{n} \binom{n}{i} [(\mathrm{ad}_x - \alpha 1)^{n-i}(y), (\mathrm{ad}_x - \beta 1)^i(z)] = 0\,.$$

Therefore, $[y,z] \in L_N^{\alpha+\beta}$.

If additionally $y \in L_{N,\alpha}$, and $z \in L_{N,\beta}$, then the identity holds with $n = 1$, hence $[L_{N,\alpha}, L_{N,\beta}] \leq L_{N,\alpha+\beta}$. □

The most important case is that where $N = H$ is a Cartan subalgebra of $L$. Theorem (5.11) tells us that

$$L = \bigoplus_\lambda L_H^\lambda = \bigoplus_\lambda L^\lambda\,,$$

where $\lambda$ runs over the finite set of weights of $H (= L_H^0)$ on $L$. This is a *Cartan decomposition* of the Lie algebra $L$—the decomposition of $L$ as the direct sum of its generalized weight spaces for a Cartan subalgebra $H$.

A nonzero weight for the action of $H$ on $L$ is a *root*. Therefore the Cartan decomposition is often written

$$L = H \oplus \bigoplus_{\lambda \in \Phi} L^\lambda\,,$$

where $\lambda$ runs over the finite set $\Phi$ of roots of $H$ on $L$.

Here and above we see the common abuse of notation and terminology that refers to the weights and weight spaces of $L$ without specifying the Cartan subalgebra $H$ being used, say, writing $L^\lambda$ in place of $L_H^\lambda$. Usually $H\,(= L_H^0 = L^0)$ will be clear from the context, and in the cases of most interest to us all Cartan algebras are equivalent; see the remarks on page 55 and see Corollary (8.24) on semisimple algebras.

By the theorem $[L^\alpha, L^{-\alpha}] \leq L^0 = H$, clearly a special situation. Right now we do not even know that $-\alpha$ is a root for every root $\alpha$. It will turn out that this is the case and that subalgebras generated by elements $x \in L^\alpha$ and $y \in L^{-\alpha}$ and the related action on submodules of $L$ are extremely important. We begin with a technical lemma of that nature, which we then use for the first of several times.

**(5.13).** LEMMA. *Let $L$ be a finite dimensional Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic $0$. Assume that $\alpha$ and $-\alpha$ are roots.*

*Let $M = \bigoplus_{\lambda\in\Lambda} M^\lambda$, for $\Lambda \subseteq \Phi$, be a $\mathbb{K}$-subspace of $L$ that is invariant under $\mathrm{ad}_x$ and $\mathrm{ad}_y$ for $x \in L^\alpha$ and $y \in L^{-\alpha}$ and has $M^\lambda = M \cap L^\lambda$ for each $\lambda \in \Lambda$. Then $h = [x, y]$ is in $H$ and*

$$0 = \sum_{\lambda\in\Lambda} \dim_{\mathbb{K}}(M^\lambda)\lambda(h)\,.$$

PROOF. As $x$ and $y$ are in the subalgebra $\mathrm{N}_L(M)$, so is $h = [x, y]$. Here $h$ is an element of $H$ by the previous theorem; indeed by the theorem $h$ normalizes each $M^\lambda = M \cap L^\lambda$.

The linear transformation $\mathrm{ad}_h = [\mathrm{ad}_x, \mathrm{ad}_y]$ is a commutator and so has trace $0$ on the subspace $M$ normalized by $x$ and $y$. That is,

$$0 = \mathrm{tr}(\mathrm{ad}_h|_M) = \sum_{i=a}^{b} \mathrm{tr}(\mathrm{ad}_h|_{M^\lambda}) = \sum_{i=a}^{b} \dim_{\mathbb{K}}(M^\lambda)\lambda(h)\,,$$

as desired. □

**(5.14).** PROPOSITION. *Let $L$ be a finite dimensional Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic $0$, and let $\alpha$ and $\beta$ be roots of $L$. Then $\beta$ is a rational multiple of $\alpha$ when restricted to the subspace $[L^\alpha, L^{-\alpha}]$.*

PROOF. This is [Ste70, Lemma 3.2].

The result is trivial if $-\alpha$ is not a root, so we may assume it is.

Choose nonnegative integers $s$ and $t$ so that neither $\beta-(s+1)\alpha$ nor $\beta+(t+1)\alpha$ is a root. As there are only finitely many roots, such $s$ and $t$ certainly exist. Define $M$ to be the corresponding subspace

$$M = L^{\beta-s\alpha} \oplus \cdots \oplus L^\beta \oplus \cdots \oplus L^{\beta+t\alpha} = \bigoplus_{i=-s}^{t} L^{\beta+i\alpha}\,.$$

Set $d_i = \dim_{\mathbb{K}}(L^{\beta+i\alpha})$, so that $d_i \geq 0$ with $d_0 > 0$ as $\beta$ is a root.

Our choice of $s$ and $t$ and Theorem (5.12) give $[M, L^{-\alpha}] \leq M$ and $[M, L^\alpha] \leq M$. That is, $L^{-\alpha}$ and $L^\alpha$ normalize $M$. Let $x \in L^\alpha$ and $y \in L^{-\alpha}$, and set $h = [x, y]$. By the technical lemma above

$$0 = \sum_{i=-s}^{t} d_i(\beta + i\alpha)(h)\,,$$

hence

$$\beta(h) = \frac{d}{e}\alpha(h)$$

for $d = -\sum_{i=-s}^{t} id_i$ and $e = \sum_{i=-s}^{t} d_i \neq 0$. By linearity, this holds for all $h$ in $[L^\alpha, L^{-\alpha}]$. □

## 5.4 Killing forms

Let $L$ be a Lie $\mathbb{K}$-algebra and $V$ a finite dimensional $L$-module. The *Killing form* of $L$ on $V$, $\kappa_L^V : L \times L \longrightarrow \mathbb{K}$ is is a bilinear form given by

$$\kappa_L^V(x, y) = \mathrm{tr}(\mathrm{ad}_x^V \, \mathrm{ad}_y^V)\,,$$

where we recall our convention that $\mathrm{ad}_x^V$ is the image of $x \in L$ in $\mathrm{End}_{\mathbb{K}}(V)$. For the basic theory of bilinear forms, refer to Appendix A.

If the relevant Lie algebra $L$ should be evident from the context, then we write $\kappa^V$. Finally, if $V = L$, the representation being the adjoint, we may drop reference to $V$ as well, since we then have the usual definition of the Killing form

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

**(5.15).** PROPOSITION.

(a) *The Killing form $\kappa_L^V$ is a symmetric, bilinear form on $L$.*

(b) *If $W$ is an $L$-submodule of $V$, then*

$$\kappa_L^V = \kappa_L^W + \kappa_L^{V/W}\,.$$

(c) *The Killing form is an* invariant form *(or* associative form*): for all $x, y, z \in L$*

$$\kappa_L^V([x, y], z) = \kappa_L^V(x, [y, z])\,.$$

(d) *If $I$ is an ideal of $L$, then $I^\perp = \{\, x \in L \mid \kappa_L^V(x, y) = 0\,,$ for all $y \in I \,\}$ is also an ideal of $L$. Especially $L^\perp = \mathrm{Rad}(L, \kappa)$ is an ideal.*

PROOF.

(a) The trace is linear in its argument with target $\mathbb{K}$, and multiplication in $\mathrm{End}_{\mathbb{K}}(V)$ is bilinear; so $\kappa_L^V$ is a bilinear form on $L$. It is symmetric since $\mathrm{tr}(ab) = \mathrm{tr}(ba)$ in $\mathrm{End}_{\mathbb{K}}(V)$.

(b) This is evident if we write the module action in matrix form, using a basis that extends a basis of $W$ to one for all $V$.

(c)

$$\begin{aligned}\kappa_L^V([x,y],z) &= \mathrm{tr}(\mathrm{ad}^V_{[x,y]}\,\mathrm{ad}^V_z)\\ &= \mathrm{tr}((\mathrm{ad}^V_x\,\mathrm{ad}^V_y - \mathrm{ad}^V_y\,\mathrm{ad}^V_x)\,\mathrm{ad}^V_z)\\ &= \mathrm{tr}(\mathrm{ad}^V_x\,\mathrm{ad}^V_y\,\mathrm{ad}^V_z - \mathrm{ad}^V_y\,\mathrm{ad}^V_x\,\mathrm{ad}^V_z)\\ &= \mathrm{tr}(\mathrm{ad}^V_x\,\mathrm{ad}^V_y\,\mathrm{ad}^V_z) - \mathrm{tr}(\mathrm{ad}^V_y\,\mathrm{ad}^V_x\,\mathrm{ad}^V_z)\\ &= \mathrm{tr}(\mathrm{ad}^V_x\,\mathrm{ad}^V_y\,\mathrm{ad}^V_z) - \mathrm{tr}(\mathrm{ad}^V_x\,\mathrm{ad}^V_z\,\mathrm{ad}^V_y)\\ &= \mathrm{tr}(\mathrm{ad}^V_x(\mathrm{ad}^V_y\,\mathrm{ad}^V_z - \mathrm{ad}^V_z\,\mathrm{ad}^V_y))\\ &= \mathrm{tr}(\mathrm{ad}^V_x\,\mathrm{ad}^V_{[y,z]})\\ &= \kappa_L^V(x,[y,z])\,.\end{aligned}$$

(d) For all $a \in I$, $y \in L$, and $b \in I^\perp$ we have by (c)

$$0 = \kappa_L^V([a,y],b) = \kappa_L^V(a,[y,b])\,.$$

That is, $[y,b] \in I^\perp$ for all $y \in L$ and $b \in I^\perp$; so $I^\perp$ is an ideal. □

**(5.16).** Corollary. *Let $I$ be an ideal of the finite dimensional Lie algebra $L$. Then $I \le \mathrm{Rad}(\kappa_L^{L/I})$ and $\kappa_I^I = \kappa_L^I|_{I\times I} = \kappa_L^L|_{I\times I}$*

Proof. From the second part of the proposition

$$\kappa_L^L = \kappa_L^I + \kappa_L^{L/I}\,.$$

As $I$ acts as 0 on $L/I$, we certainly have $I \le \mathrm{Rad}(\kappa_L^{L/I})$. Thus on $I \times I$ we have $\kappa_L^L = \kappa_L^I + 0 = \kappa_L^I$. As $I$ is an ideal of $L$ also $\kappa_I^I = \kappa_L^I|_{I\times I}$. □

Some care must be taken in the use of this result. In $\mathfrak{sl}_n(\mathbb{K})$ the Borel algebra $\mathfrak{b}_n(\mathbb{K}) = \mathfrak{n}_n^+(\mathbb{K}) \oplus \mathfrak{h}_n(\mathbb{K})$ is the split extension of its derived subalgebra $\mathfrak{n}_n^+(\mathbb{K}) = [\mathfrak{b}_n(\mathbb{K}), \mathfrak{b}_n(\mathbb{K})]$ by the Cartan subalgebra $\mathfrak{h}_n(\mathbb{K})$. Let $L = \mathfrak{b}_n(\mathbb{K})$ and $I = \mathfrak{n}_n^+(\mathbb{K})$. Then nilpotent $I$ consists of strictly upper triangular matrices, so $\kappa_I^I$ is identically 0; $L/I \simeq \mathfrak{h}_n(\mathbb{K})$ is abelian and so $\kappa_{L/I}^{L/I}$ is identically 0. Nevertheless $\kappa_L^L = \kappa_L^I + \kappa_L^{L/I}$ is not identically 0 on solvable $\mathfrak{b}_n(\mathbb{K})$ provided $n \ge 2$.

**(5.17).** Theorem. *Let $L\,(\neq 0)$ be a finite dimensional Lie algebra over the field $\mathbb{K}$ of characteristic 0. If $L = [L,L]$ then the Killing form $\kappa$ is not identically 0.*

Proof. For any extension field $\mathbb{E}$ of $\mathbb{K}$, if $\kappa^L$ is identically 0, then so is $\kappa^{\mathbb{E}\otimes_{\mathbb{K}} L}$. Therefore in proving the theorem we may assume that $\mathbb{K}$ is algebraically closed.

Let $L = \bigoplus_{\lambda\in\Phi_0} L^\lambda$ be the Cartan decomposition for $L$ relative to the Cartan subalgebra $H = L^0$ and finite set of weights $\Phi_0$. Thus

$$L = [L,L] = \left[\bigoplus_{\lambda\in\Phi_0} L^\lambda, \bigoplus_{\mu\in\Phi_0} L^\mu\right] = \bigoplus_{\lambda,\mu\in\Phi_0} [L^\lambda, L^\mu]$$

In particular

$$H = \bigoplus_{\lambda \in \Phi_0} [L^\lambda, L^{-\lambda}]\,.$$

As nonzero nilpotent $H > [H,H]$ and $L = [L,L]$, we have $H < L$; so the set of roots $\Phi = \Phi_0 \setminus \{0\}$ is nonempty.

Let $\lambda \in \Phi$. By the definition of roots, $\lambda|_H \neq 0$ but $\lambda|_{[H,H]} = 0$. Therefore there is an $\alpha \in \Phi_0$ with $\lambda|_{[L^\alpha, L^{-\alpha}]} \neq 0$. Furthermore $\alpha$ is not the zero weight as again $\lambda|_{[H,H]} = \lambda|_{[L^0, L^{-0}]} = 0$. Thus by Proposition (5.14) there is a rational number $r_{\lambda,\alpha}$ with

$$\lambda|_{[L^\alpha, L^{-\alpha}]} = r_{\lambda,\alpha}\alpha|_{[L^\alpha, L^{-\alpha}]}\,.$$

For some fixed $\beta \in \Phi$, choose an $x \in [L^\alpha, L^{-\alpha}] \leq H$ with $\beta(x) \neq 0$, hence $r_{\beta,\alpha} \neq 0$ and $\alpha(x) \neq 0$. Then

$$\begin{aligned}
\kappa(x,x) &= \operatorname{tr}(\operatorname{ad}_x \operatorname{ad}_x)\\
&= \sum_{\lambda \in \Phi_0} \lambda(x)^2 \dim_{\mathbb{K}}(L^\lambda)\\
&= 0 + \sum_{\lambda \in \Phi} \lambda(x)^2 \dim_{\mathbb{K}}(L^\lambda)\\
&= \alpha(x)^2 \sum_{\lambda \in \Phi} r_{\lambda,\alpha}^2 \dim_{\mathbb{K}}(L^\lambda)\,,
\end{aligned}$$

which is not equal to 0—as above $\alpha(x) \neq 0$; $r_{\beta,\alpha}$ is not zero; and all $\dim_{\mathbb{K}}(L^\lambda)$ are positive integers. Since $\kappa(x,x) \neq 0$, the form $\kappa$ is not identically 0 on $L$, as desired. □

**(5.18).** COROLLARY. (CARTAN'S SOLVABILITY CRITERION) *Let $L$ be a finite dimensional Lie algebra over the field $\mathbb{K}$ of characteristic 0. If the Killing form is identically 0, then $L$ is solvable.*

PROOF. Assume the Killing form $\kappa$ is identically 0. The proof is by induction on $\dim_{\mathbb{K}}(L)$, with the dimension 0 and 1 cases clear. By the Theorem $L \neq [L,L]$. By Corollary (5.16) the Killing form for $[L,L]$ comes from restriction of the Killing form for $L$ and so is also identically 0. Therefore by induction $[L,L]$ is solvable, and then $L$ is as well by Lemma (4.4). □

A slightly more complicated condition on $\kappa$ is both necessary and sufficient for solvability; see [Eld15]: $L$ is solvable if and only if $\kappa|_{L \times [L,L]}$ is identically 0. A case in point is that of the Borel algebras $\mathfrak{b}_n(\mathbb{K})$, mentioned above, which, are solvable with a nonzero Killing form whose restriction to the derived subalgebra $[\mathfrak{b}_n(\mathbb{K}), \mathfrak{b}_n(\mathbb{K})]$ is identically 0.

We then have the natural result that lives at the opposite end of the solvability and degeneracy spectrum.

**(5.19).** THEOREM. (CARTAN'S SEMISIMPLICITY CRITERION) *Let $L$ be a finite dimensional Lie algebra over the field $\mathbb{K}$ of characteristic 0. Then $L$ is semisimple if and and only if its Killing form is nondegenerate.*

PROOF. Let $\kappa = \kappa_L^L$ be the Killing form of $L$ and $R = \mathrm{Rad}(L, \kappa)$, an ideal by Proposition (5.15). But $\kappa|_{R\times R} = \kappa_R^R$ is identically 0, so $R$ is solvable by Cartan's Solvability Criterion (5.18). Assuming $L$ to be semisimple, we find $R = 0$ so that $\kappa$ is nondegenerate.

Now let $S$ be a nonzero solvable ideal of $L$, and take $I$ to be the last nonzero term in its derived series. As $I$ is characteristic in $S$, it is an abelian ideal of $L$. Therefore $\mathrm{ad}_I^I = 0$, and $I \leq \mathrm{Rad}(\kappa_L^I)$. Also $I \leq \mathrm{Rad}(\kappa_L^{L/I})$ by Corollary (5.16). Hence by Proposition (5.15)

$$0 \neq I \leq \mathrm{Rad}(\kappa_L^I) \cap \mathrm{Rad}(\kappa_L^{L/I}) = \mathrm{Rad}(\kappa_L^I + \kappa_L^{L/I}) = \mathrm{Rad}(\kappa)$$

and $\kappa$ is degenerate. □

We also have a result that resolves a possible confusion involving terminology.

**(5.20).** THEOREM. *Let $L$ be a finite dimensional Lie algebra over the field $\mathbb{K}$ of characteristic 0. Then $L$ is semisimple if and only if, as $L$-module, it is completely reducible with no trivial 1-dimensional ideals.*

*In this case all minimal ideals (irreducible submodules) are nontrivial simple subalgebras, and they are pairwise perpendicular with respect to the Killing form.*

PROOF. Let $\kappa$ be the Killing form on $L$, and let $I$ be an ideal in semisimple $L$. Then $I \cap I^\perp$ is an ideal by Proposition (5.15), and the restriction of $\kappa$ to it is identically 0. Therefore by Cartan's Solvability Criterion (5.18) the ideal $I \cap I^\perp$ is solvable and hence 0 in semisimple $L$. Therefore finite dimensional $L = I \oplus I^\perp$, and every ideal $I$ is complemented in $L$.

By Theorem (4.8), $L$ is completely reducible as $L$-module. In particular, minimal ideals and irreducible submodules are the same and are simple. If any of these were trivial simple ideals, they would be solvable ideals, which is not the case. Finally for the minimal ideal $I$, the complement $I^\perp$ must be the sum of all other minimal ideals, so these simple summands are pairwise perpendicular.

Conversely, assume that $L$ is completely reducible with the decomposition $L = \bigoplus_{i=0}^m S_i$ into simple ideals with no summand trivial. Any solvable ideal $I$ projects onto each summand $S_i$ as a solvable ideal. Since no summand is trivial, each of these projections is onto the zero ideal; $I$ itself must be zero. Therefore $L$ is semisimple. □

## 5.5 Problems

**(5.21).** PROBLEM. *Prove that any subalgebra of the Lie algebra $L$ that contains $L_x^0$ is self-normalizing.*

dd exercise: ]associative/invariant forms and properties

**(5.22).** PROBLEM. *Let $L$ be a finite dimensional Lie algebra in characteristic 0. Prove that $L$ is solvable if $\kappa|_{L\times[L,L]}$ is identically 0.*

# Chapter 6

# Semisimple Lie Algebras: Basic Structure

We take the view that the classification of finite dimensional, semisimple Lie algebras over algebraically closed fields of characteristic 0 has four basic parts:

(i) for each algebra, the construction of a root system that functions as a skeleton;

(ii) the classification of root systems;

(iii) the uniqueness of Lie algebras corresponding to the various root systems;

(iv) the existence of Lie algebras corresponding to the various root systems.

In this chapter we handle the first part.

We first set some notation to be used throughout. Especially

*$L$ is a nonzero, finite dimensional, semisimple Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic 0.*

By Theorem (5.8) we may choose and fix a Cartan subalgebra $H$ in $L$. By Proposition (5.9) we have $H = L^0_H = L^0$, the zero generalized weight space. Let $\Phi$ be the set of all roots for $H$ on $L$, a finite set by Theorem (5.11). The set of all weights is $\Phi_0 = \{0\} \cup \Phi$.

For each $\lambda \in \Phi$, we have the generalized weight space $L^\lambda = L^\lambda_H$, giving the Cartan decomposition

$$L = \bigoplus_{\lambda \in \Phi_0} L^\lambda = H \oplus \bigoplus_{\lambda \in \Phi} L^\lambda .$$

Since $L$ is nonzero and semisimple, the nilpotent Cartan subalgebra $H = L^0$ is proper in $L$, hence the root set $\Phi$ is nonempty.

For arbitrary $\lambda \in (H/[H,H])^*$, we define $L^\lambda$ to be 0 whenever $\lambda$ is not a weight.

The Killing form $\kappa = \kappa_L^L$ is nondegenerate by Cartan's Semisimplicity Criterion (5.19).

## 6.1 Toral subalgebras

**(6.1).** PROPOSITION. *Let $\alpha$ and $\beta$ be weights.*

(a) $\kappa(L^\alpha, L^\beta) = 0$ *if* $\alpha + \beta \neq 0$.

(b) *For every* $0 \neq x \in L^\alpha$ *we have* $\kappa(x, L^{-\alpha}) \neq 0$. *Especially* $-\alpha \in \Phi$.

(c) $\kappa|_{H\times H}$ *is nondegenerate, and* $H^\perp = \bigoplus_{\lambda\in\Phi} L^\lambda$.

PROOF. (a) Recall that for all weights $\mu, \nu$ we have $[L^\mu, L^\nu] \leq L^{\mu+\nu}$ by Theorem (5.12), and this extends to all $\lambda, \mu \in (H/[H,H])^*$ as we have defined $L^\lambda$ to be 0 whenever $\lambda$ is not a weight.

For $x \in L^\alpha$, $y \in L^\beta$, and $\gamma \in \Phi_0$,

$$\mathrm{ad}_x \, \mathrm{ad}_y \, L^\gamma = [x, [y, L^\gamma]] \leq [L^\alpha, [L^\beta, L^\gamma]] \leq [L^\alpha, L^{\beta+\gamma}] \leq L^{\alpha+\beta+\gamma} .$$

Therefore if $\alpha + \beta$ is not equal to 0, then all diagonal entries of $\mathrm{ad}_x \, \mathrm{ad}_y$ are 0 and $\mathrm{tr}(\mathrm{ad}_x \, \mathrm{ad}_y) = \kappa(x, y)$ is 0.

(b) If $x \in L^\alpha$ with $\kappa(x, L^{-\alpha}) = 0$, then by (a)

$$x \in \Big(\bigoplus_{\lambda\in\Phi_0} L^\lambda\Big)^\perp = L^\perp = \mathrm{Rad}(L, \kappa) = 0 .$$

As $\alpha$ is a root, a nonzero $x$ in in $L^\alpha$ exists. Hence $\kappa(x, L^{-\alpha}) \neq 0$; this requires $L^{-\alpha} \neq 0$, so $-\alpha \in \Phi$.

(c) By (b) with $\alpha = 0$ we have $H \cap H^\perp = 0$, so $H$ is nondegenerate for $\kappa$. By (a) $H^\perp$ contains $\bigoplus_{\lambda\in\Phi} L^\lambda$, and now we have equality. □

**(6.2).** THEOREM. *The Cartan subalgebra $H$ is abelian.*

PROOF. Let $x, y \in H$. Then

$$\kappa(x, y) = \mathrm{tr}(\mathrm{ad}_x \, \mathrm{ad}_y) = \sum_{\lambda\in\Phi} \lambda(x)\lambda(y) \dim L^\lambda .$$

If $w \in [H, H]$, then $\lambda(w) = 0$ for all $\lambda \in \Phi$; so $\kappa(w, y) = 0$ for all $w \in [H, H]$ and $y \in H$. That is, $[H, H] \leq H \cap H^\perp = 0$ by Proposition (6.1)(b). Therefore $H$ is abelian. □

The theorem can be restated as $L^0 = L_0$ $(= H)$; the generalized weight space for the zero weight of $H$ is equal to its weight space. We shall see below that this is true for all weight spaces.

As abelian $H\,(= H/[H,H])$ is finite dimensional and nondegenerate under $\kappa$, for every linear functional $\mu \in H^*$ there is a unique $t_\mu \in H$ with

$$\kappa(t_\mu, h) = \mu(h)$$

for all $h \in H$. Especially $t_0 = 0$ and indeed $t_{a\mu} = at_\mu$ for all $a \in \mathbb{K}$.

**(6.3).** PROPOSITION.

(a) $H = \sum_{\lambda\in\Phi} \mathbb{K}t_\lambda$.

(b) *For each* $\alpha \in \Phi$ *we have* $\alpha(t_\alpha) = \kappa(t_\alpha, t_\alpha) \neq 0$.

PROOF. (a) Let $J = \sum_{\lambda\in\Phi} \mathbb{K}t_\lambda \leq H$ and choose $h \in J^\perp \cap H$. Then $\lambda(h) = \kappa(t_\lambda, h) = 0$ for all $\lambda \in \Phi$ and indeed for all $\lambda \in \Phi_0$ since $t_0 = 0$ with $H = L^0$. Thus for a basis of $L$ consisting of bases for the various $L^\lambda$ (ordered appropriately using Theorem (5.11)) every $\mathrm{ad}_x$, for $x \in H$, is represented by a matrix that is upper triangular and $\mathrm{ad}_h$ itself is strictly upper triangular. But then $\mathrm{ad}_h \mathrm{ad}_x$ is always strictly upper triangular, hence $h \in \mathrm{Rad}(H, \kappa) = 0$. Therefore $J^\perp \cap H = 0$ with $J \leq H$, so $J = H$ because $L$ has finite dimension.

(b) By nondegeneracy of $\kappa$ on $H$ and (a), there is a root $\beta$ with $0 \neq \kappa(t_\beta, t_\alpha) = \beta(t_\alpha)$. Then Proposition (5.14) yields

$$0 \neq \beta(t_\alpha) = r_{\beta,\alpha}\alpha(t_\alpha) = r_{\beta,\alpha}\kappa(t_\alpha, t_\alpha)\,,$$

so $\kappa(t_\alpha, t_\alpha) \neq 0$. □

**(6.4).** THEOREM.

(a) $L^\alpha = L_\alpha$ *has dimension* 1 *for each* $\alpha \in \Phi$.

(b) *For* $\alpha \in \Phi$ *we have* $\mathbb{K}\alpha \cap \Phi = \{\pm\alpha\}$.

(c) *For* $\alpha \in \Phi$, $x \in L_\alpha$, *and* $y \in L_{-\alpha}$ *we have* $[x, y] = \kappa(x, y)t_\alpha$.

PROOF. We first prove (c) in the form:

*For* $\alpha \in \Phi$, $x \in L^\alpha$, *and* $y \in L_{-\alpha}$ *we have* $[x, y] = \kappa(x, y)t_\alpha$.

By Theorem (5.12) in any event $[x, y] \in H$. We have, for all $h \in H$,

$$\begin{aligned}
\kappa([x, y], h) &= \kappa(x, [y, h]) && \kappa \text{ is associative}\\
&= \kappa(x, \alpha(h)y) && y \in L_{-\alpha}\\
&= \alpha(h)\kappa(x, y) &&\\
&= \kappa(t_\alpha, h)\kappa(x, y) && \text{definition of } t_\alpha\\
&= \kappa(\kappa(x, y)t_\alpha, h)\,. &&
\end{aligned}$$

Therefore for all $h \in H$

$$\kappa([x, y] - \kappa(x, y)t_\alpha, h) = 0\,.$$

By Proposition (6.1)(c), the form $\kappa$ is nondegenerate on $H$; thus $[x, y] = \kappa(x, y)t_\alpha$, as claimed.

Next we make a start on (b) with the

*Claim:* $\mathbb{K}\alpha \cap \Phi = \mathbb{Q}\alpha \cap \Phi$.

By Proposition (6.1)(b) for $y \in L_{-\alpha}$ there is an $x \in L^\alpha$ with $\kappa(x, y) \neq 0$. Thus by the above we have $0 \neq t_\alpha \in [L^\alpha, L^{-\alpha}]$. Next from Proposition (6.3)(b) we have $\alpha(t_\alpha) \neq 0$. We conclude that $\alpha|_{[L^\alpha, L^{-\alpha}]}$ is nonzero.

Let $\beta = r\alpha$ be a root for $0 \neq r \in \mathbb{K}$, so that

$$\beta|_{[L^\alpha, L^{-\alpha}]} = r\alpha|_{[L^\alpha, L^{-\alpha}]}$$

is nonzero. By Proposition (5.14) the scalar $r$ in this equation is rational. We conclude that $\mathbb{K}\alpha \cap \Phi = \mathbb{Q}\alpha \cap \Phi$, as claimed.

We are now in a position to prove (a) and finish (b) simultaneously.

Consider $\mathbb{K}\alpha \cap \Phi = \mathbb{Q}\alpha \cap \Phi$. Replacing $\alpha$ by any of the members of this finite set does not change the set itself. Therefore we may assume that $\alpha$ is minimal within $\mathbb{Q}^+\alpha$, which is to say

$$[0, 1]\alpha \cap \Phi = \{\alpha\}.$$

As before, choose $x \in L^\alpha$ and $y \in L_{-\alpha}$ with $\kappa(x, y) \neq 0$. Define the $\mathbb{K}$-subspace of $L$

$$M = \mathbb{K}y \oplus H \oplus \bigoplus_{i \in \mathbb{Q}^{\geq 1}} L^{i\alpha}\,,$$

where the last piece has only finitely many nonzero summands, one of which is $L^\alpha$. We claim that $M$ is invariant under both $\mathrm{ad}_x$ and $\mathrm{ad}_y$. By Theorem (5.12) this is immediately clear for $\mathrm{ad}_x$ since $\mathrm{ad}_x(L^{j\alpha}) \leq L^{(j+1)\alpha}$ with $\mathbb{K}y \leq L^{-\alpha}$ and $H = L^0$.

Similarly $\mathrm{ad}_y(L^{j\alpha}) \leq L^{(j-1)\alpha}$. Here $\mathrm{ad}_y(\mathbb{K}y) = 0$ and $\mathrm{ad}_y(H) \leq \mathbb{K}y \leq M$ as $y \in L_{-\alpha}$. Finally, since $\alpha$ was chosen to be minimal

$$\mathrm{ad}_y\Big(\bigoplus_{i \in \mathbb{Q}^{\geq 1}} L^{i\alpha}\Big) \leq \bigoplus_{i \in \mathbb{Q}^{\geq 0}} L^{i\alpha} = H \oplus \bigoplus_{i \in \mathbb{Q}^{\geq 1}} L^{i\alpha} \leq M\,.$$

This gives our claim that $M$ is invariant under $\mathrm{ad}_x$ and $\mathrm{ad}_y$. Set $h = [x, y]$, and invoke the Technical Lemma (5.13):

$$\begin{aligned}
0 &= \dim(\mathbb{K}y)(-\alpha(h)) + (\dim H)0 + \sum_{i \in \mathbb{Q}^{\geq 1}} \dim(L^{i\alpha})\, i\alpha(h) \\
&= \alpha(h)\Big(-1 + 0 + \sum_{i \in \mathbb{Q}^{\geq 1}} i \dim(L^{i\alpha})\Big) \\
&= \alpha(h)\Big((-1 + \dim L^\alpha) + \sum_{i \in \mathbb{Q}^{>1}} i \dim(L^{i\alpha})\Big)
\end{aligned}$$

The space $L^\alpha$ has dimension at least 1 and

$$\alpha(h) = \alpha(\kappa(x, y)t_\alpha) = \kappa(x, y)\alpha(t_\alpha) \neq 0\,,$$

as observed above. We conclude that $\dim L^\alpha = 1$ and so additionally $L_\alpha = L^\alpha$, as in (a).

Furthermore $\dim L^{i\alpha} = 0$ for $i > 1$. Therefore $\mathbb{Q}^{\geq 0}\alpha \cap \Phi = \{\alpha\}$. After swapping $-\alpha$ for $\alpha$ we also find $\mathbb{Q}^{\geq 0}(-\alpha) \cap \Phi = \{-\alpha\}$, which is to say $\mathbb{Q}^{\leq 0}\alpha \cap \Phi = \{-\alpha\}$. These combine to $\mathbb{Q}\alpha \cap \Phi = \{\pm\alpha\}$. We add to this the claim $\mathbb{K}\alpha \cap \Phi = \mathbb{Q}\alpha \cap \Phi$, verified above, and so conclude $\mathbb{K}\alpha \cap \Phi = \{\pm\alpha\}$. This is the desired (b). □

We have an important corollary:

**(6.5).** Corollary. *A nonzero, finite dimensional, semisimple Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic $0$ with rank $l$ and root system $\Phi$ has dimension $l + |\Phi|$.*

Proof. $L = H \oplus \bigoplus_{\lambda\in\Phi} L_\lambda$ where $H$ has dimension $l$ and each $L_\lambda$ has dimension 1. □

## 6.2 $\mathfrak{sl}_2(\mathbb{K})$ subalgebras

**(6.6).** Theorem. *For each $\alpha \in \Phi$ the subalgebra generated by $L_\alpha$ and $L_{-\alpha}$ is isomorphic to $\mathfrak{sl}_2(\mathbb{K})$.*

*More specifically, for each $0 \neq x \in L_\alpha$ (respectively, $0 \neq y \in L_{-\alpha}$) there is a $0 \neq y \in L_{-\alpha}$ (respectively, $0 \neq x \in L_\alpha$) such that $(x, y, h_\alpha)$ is a Chevalley basis, where $h_\alpha = \frac{2}{\kappa(t_\alpha,t_\alpha)} t_\alpha$.*

Proof. $\kappa(t_\alpha, t_\alpha) \neq 0$ by Proposition (6.3)(b), so we may define $h_\alpha = \frac{2}{\kappa(t_\alpha,t_\alpha)} t_\alpha$. First

$$\begin{aligned}
[h_\alpha, x] &= \frac{2}{\kappa(t_\alpha,t_\alpha)}[t_\alpha, x] \\
&= \frac{2}{\kappa(t_\alpha,t_\alpha)}\alpha(t_\alpha)x = \frac{2}{\kappa(t_\alpha,t_\alpha)}\kappa(t_\alpha,t_\alpha)x = 2x\,; \\
[h_\alpha, y] &= \frac{2}{\kappa(t_\alpha,t_\alpha)}[t_\alpha, y] = \frac{-2}{\kappa(t_\alpha,t_\alpha)}[t_{-\alpha}, y] = \frac{-2}{\kappa(t_\alpha,t_\alpha)}(-\alpha(t_{-\alpha}))y \\
&= \frac{-2}{\kappa(t_\alpha,t_\alpha)}\alpha(t_\alpha)y = \frac{-2}{\kappa(t_\alpha,t_\alpha)}\kappa(t_\alpha,t_\alpha)y = -2y\,.
\end{aligned}$$

Also $\kappa(x, y) \neq 0$ by the previous theorem, so

$$0 \neq [x, y] = \kappa(x, y)t_\alpha = \kappa(x, y)\left(\frac{\kappa(t_\alpha,t_\alpha)}{2}\right)h_\alpha\,.$$

Therefore, given $x$ or $y$ the other may be scaled so that $\kappa(x, y) = \frac{2}{\kappa(t_\alpha,t_\alpha)}$, creating the Chevalley basis $(h_\alpha, x, y)$ for

$$\mathbb{K}[x, y] \oplus \mathbb{K}x \oplus \mathbb{K}y = \mathbb{K}t_\alpha \oplus \mathbb{K}x \oplus \mathbb{K}y = \mathbb{K}h_\alpha \oplus \mathbb{K}x \oplus \mathbb{K}y\,.$$

which is isomorphic to $\mathfrak{sl}_2(\mathbb{K})$, as desired. □

We now have a fundamental property of the Cartan decomposition.

**(6.7).** COROLLARY. *Let $L = H \oplus \bigoplus_{\alpha\in\Phi} L_\alpha$ be the Cartan decomposition for $L$. For each $\alpha \in \Phi$ the subalgebra $S_\alpha = S_{-\alpha}$ generated by $L_\alpha$ and $L_{-\alpha}$ is a copy of $\mathfrak{sl}_2(\mathbb{K})$, and $L = \sum_{\alpha\in\Phi} S_\alpha$.* □

PROOF. Proposition (6.3)(a) tells us that $H = \sum_{\alpha\in\Phi} \mathbb{K}t_\alpha$, and the theorem provides the rest. □

Theorem (6.4) also has a profound effect on the representation theory of semisimple Lie algebras.

**(6.8).** THEOREM. *Let $L$ be a semisimple Lie algebra over the algebraically closed field or characteristic $0$. If $V = \bigoplus_{\lambda\in\Lambda} V^\lambda$ is a nonzero finite dimensional $L$-module, written as the direct sum of its generalized weight modules $V^\lambda$ for the Cartan subalgebra $H$ of $L$, then $\bigoplus_{\lambda\in\Lambda} V_\lambda$ is a nonzero $L$-submodule of $V$. Especially, if $V$ is irreducible then $V = \bigoplus_{\lambda\in\Lambda} V_\lambda$.*

PROOF. By Theorem (6.4) the algebra $L$ is the direct sum of its weight spaces $\bigoplus_{\alpha\in\Phi} L_\alpha$. If we consider the action of nilpotent $H$ inside the semidirect product $M = L \oplus V$, then by Theorem (5.12) $M = \bigoplus_{\pi\in\Phi\cup\Lambda} M^\pi$, and furthermore for all $\alpha \in \Phi$ and all $\lambda \in \Lambda$

$$L_\alpha V_\lambda = [L_\alpha, V_\lambda] \le [M_\alpha, M_\lambda] \cap V \le M_{\alpha+\lambda} \cap V = V_{\alpha+\lambda}\,.$$

Thus

$$L\Big(\bigoplus_{\lambda\in\Lambda} V_\lambda\Big) = \bigoplus_{\alpha\in\Phi,\lambda\in\Lambda} L_\alpha V_\lambda \le \bigoplus_{\lambda\in\Lambda} V_\lambda\,.$$

As $V_\lambda \neq 0$ if and only if $V^\lambda \neq 0$ in finite dimensional $V$, we are done. □

If the $L$-module $V$ is equal to $\bigoplus_{\lambda\in\Lambda} V_\lambda$, then $V$ is said to be a *weight module* for $L$.

## 6.3 The root system of a semisimple Lie algebra

As $h \mapsto t_h$ is a $\mathbb{K}$-isomorphism of $H$ and $H^*$, we may define on $H^*$ the symmetric bilinear form

$$(x, y) = \kappa^*(x, y) = \kappa(t_x, t_y)\,.$$

Set $E_\mathbb{Q} = \sum_{\alpha\in\Phi} \mathbb{Q}\alpha \le H^*$. The $\mathbb{Q}$-space $E_\mathbb{Q}$, equipped with the restriction of the form $\kappa^*$ and the special spanning set $\Phi$, will provide us with the basic example of a root system.

Here we develop some of its properties, and then in the next chapter we classify all spaces that enjoy these properties.

Following Theorem (6.6), for each root $\alpha \in \Phi$ we define $h_\alpha = \frac{2}{\kappa(t_\alpha,t_\alpha)} t_\alpha$ (possible since $\kappa(t_\alpha, t_\alpha) = (\alpha, \alpha) \neq 0$ by Proposition (6.3)(b)). Similarly for each $\alpha \in \Phi$, we let $\alpha^\vee = \frac{2}{(\alpha,\alpha)}\alpha$, the *coroot* corresponding to the root $\alpha$. With this definition, for all $\mu \in H^*$

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

Especially, for all roots $\alpha$ we have $\alpha(h_\alpha) = (\alpha, \alpha^\vee) = 2$.

We first see that $(E_{\mathbb{Q}}, \kappa^*|_{E_{\mathbb{Q}}})$ is a rational form of the orthogonal space $(H^*, \kappa^*)$. (From now on, we will use the notation $(x, y)$ exclusively in place of $\kappa^*(x, y)$.)

For $\alpha, \beta \in \Phi$ the *$\alpha$-string through $\beta$* is the longest string of roots

$$\beta - s\alpha\,, \dots\,, \beta - i\alpha\,, \dots\,, \beta\,, \dots\,, \beta + j\alpha\,, \dots\,, \beta + t\alpha\,.$$

That is, all the maps in the string are roots, but $\beta - (s+1)\alpha$ and $\beta + (t+1)\alpha$ are not roots. As $\Phi$ is finite, such a string always exists (although it may consist solely of $\beta$).

**(6.9).** THEOREM. *Let $\alpha, \beta \in \Phi$ with $\beta \neq \pm\alpha$, and let $\beta - s\alpha, \dots, \beta, \dots, \beta + t\alpha$ be the $\alpha$-string of roots through $\beta$.*

(a) $\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha = \beta - (\beta, \alpha^\vee)\alpha \in \Phi$.

(b) $\beta(h_\alpha) = \frac{2(\beta,\alpha)}{(\alpha,\alpha)} = (\beta, \alpha^\vee) = s - t \in \mathbb{Z}$.

(c) $[L_\alpha, L_\beta] = L_{\alpha+\beta}$.

(d) *For $j$ an integer, $\beta + j\alpha$ is a root if and only if $j$ is in the interval* $[-s, t]$.

PROOF. (a) Assume (b); then

$$\beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha = \beta - (\beta, \alpha^\vee)\alpha = \beta - (s - t)\alpha$$

is a root in the $\alpha$-string through $\beta$.

For (b) the definitions and calculations above give $\beta(h_\alpha) = \frac{2(\beta,\alpha)}{(\alpha,\alpha)} = (\beta, \alpha^\vee)$, so we must prove that this is equal to the integer $s - t$.

The proofs of (b), (c), and (d) use specializations of one calculation.

By Theorem (6.4)(a) every $L_\gamma$ has dimension 0 or 1 with only finitely many being nonzero. For a finite subset $J$ of the integers, we consider the subspace $M_J = \bigoplus_{j \in J} L_{\beta + j\alpha}$ within which each $L_{\beta+j\alpha}$ is required to have dimension 1.

For $0 \neq x \in L_\alpha$ and $0 \neq y \in L_{-\alpha}$, the element $h_\alpha$ is a multiple of $[x, y]$ and has $\alpha(h_\alpha) = 2$. Suppose that $M_J$ is invariant under $\mathrm{ad}_x$ and $\mathrm{ad}_y$. The Technical

Lemma (5.13) then gives

$$\begin{aligned} 0 &= \sum_{j\in J} (\dim L_{\beta+j\alpha})(\beta+j\alpha)(h_\alpha) = \sum_{j\in J}(\beta+j\alpha)(h_\alpha) \\ &= \sum_{j\in J}\beta(h_\alpha) + \sum_{j\in J} j\alpha(h_\alpha) = |J|\beta(h_\alpha) + 2\sum_{j\in J} j\,. \end{aligned}$$

That is,

$$|J|\beta(h_\alpha) = -2\sum_{j\in J} j\,.$$

For (b) we set $J = \mathbb{Z}\cap[-s,t]$ of size $s+t+1$. By Theorem (5.12) always $\mathrm{ad}_x(L_{\beta+i\alpha}) \le L_{\beta+(i+1)\alpha}$ and $\mathrm{ad}_y(L_{\beta+i\alpha}) \le L_{\beta+(i-1)\alpha}$. Because $L_{\beta+(t+1)\alpha} = 0 = L_{\beta-(s+1)\alpha}$ the space $M_J$ is invariant under $\mathrm{ad}_x$ and $\mathrm{ad}_y$.

Therefore

$$(s+t+1)\beta(h_\alpha) = -2\left(\sum_{j=1}^{s}(-j) + 0 + \sum_{i=1}^{t} i\right),$$

which yields $\beta(h_\alpha) = s-t$, as claimed.

(c) By Theorem (5.12) always $[L_\alpha, L_\beta] \le L_{\alpha+\beta}$. If $[L_\alpha, L_\beta] \ne 0$ then $\alpha+\beta$ is a root. But each root space has dimension 1 (again by by Theorem (6.4)(a)), so $[L_\alpha, L_\beta] = L_{\alpha+\beta}$.

Assume $[L_\alpha, L_\beta] = 0$ so that $\mathrm{ad}_x(L_\beta) = 0$. Then for $J = \mathbb{Z}\cap[-s,0]$ the subspace $M_J = \bigoplus_{i=-s}^{0} L_{\beta+i\alpha}$ is invariant under both $\mathrm{ad}_x$ and $\mathrm{ad}_y$ (as in the proof of (b)). Therefore

$$(s+1)\beta(h_\alpha) = -2\sum_{j=1}^{s}(-j) + 0\,,$$

so that $\beta(h_\alpha) = s$. From (b) we already know $\beta(h_\alpha) = s-t$, so $t = 0$. That is, $\beta+\alpha$ is not in the $\alpha$-string through $\beta$. Therefore $\alpha+\beta$ is not a root, and $[L_\alpha, L_\beta] = 0 = L_{\alpha+\beta}$ in this case as well.

(d) The set of $j$ for which $\beta+j\alpha$ is a root is the disjoint union of finitely many segments of the integers, one of which is $[-s,t]$, corresponding to the $\alpha$-string through $\beta$. To prove (c) we must show that this is the only interval.

Assume, for a contradiction, that there are nonempty integer intervals $[a,b]$ and $[c,d]$ with $b+1<c$ and $\beta+j\alpha$ a root for all members of the two intervals but $\beta+(a-1)\alpha$, $\beta+(b+1)\alpha$, $\beta+(c-1)\alpha$, and $\beta+(d+1)\alpha$ not roots.

The set of all roots $\beta+j\alpha$, for $j$ an integer, is the same as the set of all roots $\beta+(k-b)\alpha$, for $k$ an integer. That is, without loss of generality we may replace $\beta$ by $\beta+b\alpha$ and thereby assume $b=0$, hence also $1<c\le d$.

Our assumptions imply that $M_{[a,b]} = M_{[a,0]}$ and $M_{[c,d]}$ are both invariant under $\mathrm{ad}_x$ and $\mathrm{ad}_y$. Therefore

$$|[a,0]|\beta(h_\alpha) = -2\sum_{j=a}^{0} j$$

is nonnegative, while

$$|[c,d]|\beta(h_\alpha) = -2\sum_{j=c}^{d} j$$

is negative. This is the desired contradiction. $\square$

**(6.10).** Theorem.

(a) *The form $(\cdot,\cdot)$ is positive definite from $E_{\mathbb{Q}}$ to $\mathbb{Q}$.*

(b) *Any $\mathbb{Q}$-basis of $E_{\mathbb{Q}}$ is a $\mathbb{K}$-basis of $H^*$.*

Proof. (a) For every $\gamma \in E_{\mathbb{Q}}$ we have

$$(\gamma,\gamma) = \kappa(t_\gamma, t_\gamma) = \sum_{\beta\in\Phi} \beta(t_\gamma)^2$$

since always $\dim L_\beta$ is 1 by Theorem (6.4). By Proposition (6.3)(b), for each $\alpha \in \Phi$

$$0 \neq \kappa(t_\alpha, t_\alpha) = (\alpha,\alpha) = \sum_{\beta\in\Phi} \beta(t_\alpha)^2 = \sum_{\beta\in\Phi} (\beta,\alpha)^2\,,$$

hence

$$\frac{4}{(\alpha,\alpha)} = \sum_{\beta\in\Phi} \left(\frac{2(\beta,\alpha)}{(\alpha,\alpha)}\right)^2,$$

which is a positive integer by the previous theorem. Thus $(\alpha,\alpha) \in \mathbb{Q}^+$ and furthermore

$$(\beta,\alpha) = \frac{(\alpha,\alpha)}{2}\cdot\frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Q}\,.$$

Therefore on $E_{\mathbb{Q}} = \sum_{\alpha\in\Phi} \mathbb{Q}\alpha$ the form $(\cdot,\cdot)$ has all its values in $\mathbb{Q}$.

Let $\gamma = \sum_{\alpha\in\Phi} \gamma_\alpha \alpha \in E_{\mathbb{Q}}$ with $\gamma_\alpha \in \mathbb{Q}$. Then, as above,

$$(\gamma,\gamma) = \sum_{\beta\in\Phi} \beta(t_\gamma)^2 = \sum_{\beta\in\Phi}\left(\sum_{\alpha\in\Phi}\gamma_\alpha\beta(t_\alpha)\right)^2 = \sum_{\beta\in\Phi}\left(\sum_{\alpha\in\Phi}\gamma_\alpha(\beta,\alpha)\right)^2 \geq 0\,,$$

as it is a sum of rational squares by the previous paragraph. Furthermore if $(\gamma,\gamma) = 0$ then the rational $\beta(t_\gamma)$ is 0 for all $\beta \in \Phi$. That is, $\kappa(t_\beta, t_\gamma) = 0$ for all $\beta \in \Phi$. Since the $t_\beta$ span $H$ by Proposition (6.3), this in turn gives $t_\gamma \in H \cap H^\perp = 0$ (by Proposition (6.1)), hence $\gamma = 0$. The form is positive definite.

(b) Let $\{\, b_i \mid i \in I \,\}$ be a $\mathbb{Q}$-basis for $E_{\mathbb{Q}}$. As $\Phi \subset E_{\mathbb{Q}}$, we have $H^* = \sum_{i\in I} \mathbb{K}b_i$. Thus there is a finite subset $J \subseteq I$ with $\{\, b_j \mid j \in J \,\}$ a $\mathbb{K}$-basis of $H^*$.

Suppose $h \in (\bigoplus_{j\in J} \mathbb{Q}b_j)^\perp \cap E_{\mathbb{Q}}$. Then $H^* = \bigoplus_{j\in J} \mathbb{K}b_j \leq h^\perp$. By Proposition (6.1) and the definition of our form, it is nondegenerate on $H^*$; so we must have $h = 0$. But now in nondegenerate $E_{\mathbb{Q}} = \bigoplus_{i\in I} \mathbb{Q}b_i$ we have

$\left(\bigoplus_{j\in J}\mathbb{Q}b_j\right)^{\perp} = 0$ for the finite dimensional subspace $\bigoplus_{j\in J}\mathbb{Q}b_j$. We conclude that $J = I$ and $\bigoplus_{j\in J}\mathbb{Q}b_j = E_{\mathbb{Q}}$. Thus its $\mathbb{Q}$-basis $\{\, b_i \mid i \in I \,\} = \{\, b_j \mid j \in J \,\}$ is a $\mathbb{K}$-basis of $H^*$. $\square$

Chapter 7

# Classification of root systems

## 7.1 Abstract root systems

Let $E$ be a finite dimensional Euclidean space, and let $0 \neq v \in E$. The linear transformation

$$r_v\colon x \mapsto x - \frac{2(x,v)}{(v,v)}v$$

is the *reflection* with center $v$.

**(7.1).** LEMMA. *Let $0 \neq v \in E$.*

(a) $r_v^2 = 1$.

(b) *If $\alpha \in E^\times$ then $r_{\alpha v} = r_v$.*

(c) *$r_v \in \mathrm{O}(E)$, the orthogonal group of isometries of $E$.*

(d) *If $g \in \mathrm{O}(E)$ then $r_v^g = r_{g(v)}$.*

(e) *For the subspace $W \leq E$ we have $W^{r_x} \leq W$ if and only if $x \in W$ or $(x,W) = 0$.* □

**(7.2).** DEFINITION. *Let $E$ be a finite dimensional real space equipped with a Euclidean positive$G_2$ definite form $(\cdot,\cdot)$. Let $\Phi$ be a subset of $E$ with the following properties:*

(i) *$0 \notin \Phi$ and finite $\Phi$ spans $E$;*

(ii) *for any $\alpha \in \Phi$ we have $\mathbb{R}\alpha \cap \Phi = \{\pm\alpha\}$;*

(iii) *for each $\alpha \in \Phi$ the reflection $r_\alpha\colon x \mapsto x - \frac{2(x,\alpha)}{(\alpha,\alpha)}\alpha$ takes $\Phi$ to itself;*

(iv) (CRYSTALLOGRAPHIC CONDITION) *for each $\alpha, \beta \in \Phi$ we have $\frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$.*

*Then* $(E, \Phi)$ *is an* abstract root system *with the elements of* $\Phi$ *the* roots. *Its* rank *is* $\dim_{\mathbb{R}}(E)$. $G_2$ *The subgroup* $\mathrm{W}(\Phi)$ *equal to* $\langle\, r_\alpha \mid \alpha \in \Phi \,\rangle$ *is the* Weyl group *of the system. More generally, for any* $\Sigma \subseteq \Phi$, *we set* $\mathrm{W}(\Sigma) = \langle\, r_\alpha \mid \alpha \in \Sigma \,\rangle$.

As before, for each $\alpha \in \Phi$ the element $\alpha^\vee = \frac{2}{(\alpha,\alpha)}\alpha$ is the corresponding *coroot*. Then $\Phi^\vee$ is itself an abstract root system with $\mathrm{W}(\Phi^\vee) = \mathrm{W}(\Phi)$. (Exercise.) The Crystallographic Condition takes the form: $\frac{2(\beta,\alpha)}{(\alpha,\alpha)} = (\beta, \alpha^\vee) \in \mathbb{Z}$.

The perpendicular direct sum of abstract root systems is still an abstract root system. We say that $(E, \Phi)$ is *irreducible* if it is not possible to write $E$ as the direct sum of systems of smaller dimension. That is, we cannot have $E = E_1 \perp E_2$, with each $E_i$ spanned by nonempty $\Phi_i = \Phi \cap E_i$.

We say that two abstract root systems $(E, \Phi)$ and $(E', \Phi')$ are *equivalent root systems* if there is an invertible linear transformation $\varphi$ from $E$ to $E'$ taking $\Phi$ to $\Phi'$ and such that, for each $\alpha, \beta \in \Phi$ we have $(\alpha, \beta^\vee) = (\varphi(\alpha), \varphi(\beta)^\vee)$. Equivalence does not change the Weyl group. Equivalence is slightly weaker than isomorphism, where $\varphi$ is an isometry of $E$ and $E'$. Equivalence respects irreducibility. Indeed every equivalence becomes an isomorphism after we rescale each irreducible component of $\Phi'$ by an appropriate constant. (Exercise.)

The motivation for the current section is:

**(7.3).** THEOREM. *Let* $L$ *be a finite dimensional semisimple Lie algebra over the algebraically closed field* $\mathbb{K}$ *of characteristic* $0$. *For* $\Phi = \Phi^L$ *the set of roots with respect to the Cartan subalgebra* $H$, *set* $E_{\mathbb{Q}} = \sum_{\alpha\in\Phi} \mathbb{Q}\alpha$ *and* $E^L = \mathbb{R} \otimes_{\mathbb{Q}} E_{\mathbb{Q}}$. *Then* $(E^L, \Phi^L)$ *is an abstract root system.*

PROOF. By Theorem (6.10) the space $E^L$ with its form is positive definite and spanned by $\Phi$. Theorem (6.4)(b) tells us the only roots that are scalar multiples of the root $\alpha$ are $\pm\alpha$. Finally the Crystallographic Condition is Theorem (6.9)(a). □

We shall often abuse the terminology by talking of a *root system* rather than an abstract root system. The more precise terminology is designed to distinguish between an intrinsic root system $(E^L, \Phi^L)$, as in the theorem, and an extrinsic root system—an abstract root system.

We may also abuse notation by saying that $\Phi$ is a root system, leaving the enveloping Euclidean space $E$ implicit.

Let $v_1, \dots, v_n$ be a basis of $E$. We give the elements of $E$ (and so $\Phi$) the *lexicographic ordering*:

(i) *for* $0 \neq x = \sum_{i=1}^n x_i v_i$, *we set* $0 < x$ *if and only if* $0 < x_j$ *and* $x_i = 0$ *when* $i < j$, *for some* $1 \leq j \leq n$;

(ii) *for* $x \neq y$, *we set* $y < x$ *if and only if* $0 < x - y$;

(iii) *for* $x \neq y$, *we set* $x > y$ *if and only if* $y < x$.

This gives us a partition of $\Phi$ into the *positive roots* $\Phi^+ = \{\, \alpha \in \Phi \mid 0 < \alpha \,\}$ and the *negative roots* $\Phi^- = \{\, \alpha \in \Phi \mid 0 > \alpha \,\} = -\Phi^+$. The positive root $\delta$ is then a *simple root* or *fundamental root* if it is not possible to write $\delta$ as $\alpha+\beta$ with $\alpha,\beta \in \Phi^+$. We let $\Delta = \{\delta_1,\dots,\delta_l\}$ be the set of simple roots in $\Phi^+$.

**(7.4).** THEOREM. *Let $(E,\Phi)$ be a root system with $\Delta = \{\delta_1,\dots,\delta_l\}$ the set of simple roots in $\Phi^+$.*

(a) $\Phi^+ = \Phi \cap \sum_{i=1}^{l} \mathbb{N}\delta_i$.

(b) *For distinct $\alpha,\beta \in \Delta$ we have $(\alpha,\beta) \le 0$.*

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

PROOF. (a) The lexicographic ordering gives us a total order on $\Phi^+$, say $\alpha_1 < \cdots < \alpha_k < \cdots < \alpha_N$ where $N = |\Phi^+|$. We prove $\alpha_k \in \Phi \cap \sum_{i=1}^{l} \mathbb{N}\delta_i$ by induction on $k$. If $\alpha$, $\beta$, and $\alpha+\beta$ are all in $\Phi^+$, then $\alpha < \alpha+\beta > \beta$. Especially $\alpha_1$ is simple. Now consider $\alpha_k$. If it is simple, we are done. Otherwise $\alpha_k = \alpha_i+\alpha_j$ with $i,j < k$. By induction $\alpha_i$ and $\alpha_j$ are both in $\Phi \cap \sum_{i=1}^{l} \mathbb{N}\delta_i$, so $\alpha_k$ is as well.

(b) Consider

$$(\alpha,\beta^\vee)(\beta,\alpha^\vee) = \frac{4(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)} = 4\cos(\theta_{\alpha,\beta})^2 \in \mathbb{Z}\,,$$

where $\theta_{\alpha,\beta}$ is the angle between the vectors $\alpha$ and $\beta$.

As $0 \le \cos(\theta_{\alpha,\beta})^2 \le 1$, this must be one of $0,1,2,3,4$ with 4 occurring only when $\alpha = -\beta$. We only need consider $1,2,3$, so at least one of the integers $(\alpha,\beta^\vee)$ and $(\beta,\alpha^\vee)$ is $\pm1$. Without loss, we may assume $(\alpha,\beta^\vee)$ is $\pm1$.

Suppose $(\alpha,\beta^\vee) = 1$, so that $\alpha^{r_\beta} = \alpha - (\alpha,\beta^\vee)\beta = \alpha-\beta$ is a root. If $\alpha-\beta$ is positive, then $\alpha = \beta + (\alpha-\beta)$ contradicts $\alpha \in \Delta$. If $\alpha-\beta$ is negative, then $\beta-\alpha$ is a positive root and $\beta = \alpha + (\beta-\alpha)$ contradicts $\beta \in \Delta$. We conclude that

$$-1 = (\alpha,\beta^\vee) = \frac{2(\alpha,\beta)}{(\beta,\beta)}$$

and so $(\alpha,\beta) < 0$.

(c) By (a) the set $\Delta$ spans $\Phi^+$ hence $\Phi$ and so all $E$ (by the definition (7.2)(i)). We must show it to be linearly independent.

Suppose $\sum_{k=1}^{l} d_k\delta_k = 0$ with $d_k \in \mathbb{R}$. We rewrite this as

$$x = \sum_{i\in I} d_i\delta_i = \sum_{j\in J} d'_j\delta_j$$

where all $d_i$ and $d'_j = -d_j$ are nonnegative and $\{1,\dots,l\}$ is the disjoint union of $I$ and $J$.

First

$$(x,x) = \left(\sum_{i\in I} d_i\delta_i, \sum_{j\in J} d'_j\delta_j\right) = \sum_{i\in I, j\in J} d_i d'_j(\delta_i,\delta_j) \le 0$$

by (b), so we must have $x = 0$. On the other hand, the definition of our ordering tells us that if any of the nonnegative $d_i$ for $i \in I$ or $d'_j$ for $j \in J$ are nonzero, then $x = \sum_{i \in I} d_i \delta_i = \sum_{j \in J} d'_j \delta_j > 0$. Therefore $d_i = 0$ for all $i \in I$ and $d'_j = d_j = 0$ for all $j \in J$. That is, $\Delta$ is linearly independent. □

We may describe $\Delta$ as an *obtuse basis* since $(\alpha, \beta) \leq 0$ for distinct $\alpha, \beta \in \Delta$. If the root $\alpha$ has its unique expression $\alpha = \sum_{i=1}^{l} d_i \delta_i$ for integers $d_i$ then the *height* of the root $\alpha$ (relative to $\Delta$) is the integer $\mathrm{ht}(\alpha) = \sum_{i=1}^{l} d_i$, positive for positive roots and negative for negative roots.

The number $l$ that appears above as the dimension of $E$ and the cardinality of the simple basis $\Delta$ is the *rank* of the root system $(E, \Phi)$ and the Weyl group $\mathrm{W}(\Phi)$ which we shall see below is equal to $\mathrm{W}(\Delta)$.

**(7.5).** COROLLARY.

(a) *$\Phi$ is the disjoint union of $\Phi^+$ and $\Phi^- = -\Phi^+$ where, for each $\epsilon = \pm$, each sum of roots from $\Phi^\epsilon$ is either not a root or is a root in $\Phi^\epsilon$.*

(b) *For the set $\Delta$ of simple roots in $\Phi^+$, every root $\alpha$ has a unique representation $\sum_{i=1}^{l} d_i \delta_i$ where all the $d_i$ are nonnegative integers when $\alpha$ is a positive root and all the $d_i$ are nonpositive integers when $\alpha$ is negative.* □

The original choice of positive system $\Phi^+$ (and so its associated simple basis $\Delta$) seemed relatively arbitrary, coming from an ordering determined by a fixed but arbitrary choice of basis for $E$. The next theorem puts these into a more geometric context, characterizing $\Phi^+$ and $\Delta$ in terms of the properties presented in the corollary.

A basis of $E$ composed of roots is a *simple basis* if whenever we write a root as a linear combination of its elements, all coefficients are nonnegative integers or all coefficients are nonpositive integers. The motivating example is the set $\Delta$ of simple roots.

In the root system $(E, \Phi)$ if $\Phi$ is the disjoint union of $\Gamma^+$ and $\Gamma^- = -\Gamma^+$ where, for each $\epsilon = \pm$, each sum of roots from $\Gamma^\epsilon$ is either not a root or is a root in $\Gamma^\epsilon$, then we say that $\Gamma^+$ is a *positive system* in $(E, \Phi)$. The basic example is $\Phi^+$, or more generally the set of "positive roots" with respect to any simple basis, within which the members of that basis are appropriately minimal.

**(7.6).** THEOREM.

(a) *For $\delta \in \Delta$, we have $(\Phi^+ \setminus \delta)^{r_\delta} = \Phi^+ \setminus \delta$.*

(b) *For every positive system $\Gamma^+$ in $(E, \Phi)$ there is a $w \in \mathrm{W}(\Phi)$ with $\Gamma^+ = (\Phi^+)^w$.*

(c) *For every simple basis $\Pi$ in $(E, \Phi)$ there is a $w \in \mathrm{W}(\Phi)$ with $\Pi = \Delta^w$.*

PROOF. (a) Let $\alpha \in \Phi^+ \setminus \delta$ be given as $\alpha = \sum_{\gamma \in \Delta} a_\gamma \gamma$. Thus

$$\alpha^{r_\delta} = \alpha - (\alpha, \delta^\vee)\delta = (a_\delta - (\alpha, \delta^\vee))\delta + \sum_{\delta \neq \gamma \in \Delta} a_\gamma \gamma \,.$$

Since $\alpha \in \Phi^+ \setminus \delta$ there is some simple root $\gamma$ other that $\delta$ with $a_\gamma > 0$, hence $\alpha^{r_\delta}$ remains in $\Phi^+ \setminus \delta$ by Theorem (7.4)(a). Thus $(\Phi^+ \setminus \delta)^{r_\delta} \subseteq \Phi^+ \setminus \delta$; indeed $(\Phi^+ \setminus \delta)^{r_\delta} = \Phi^+ \setminus \delta$ because $\Phi$ is finite.

(b) The finite root set $\Phi$ is the disjoint union both of $\Phi^+$ and $\Phi^- = -\Phi^+$ and of $\Gamma^+$ and $\Gamma^- = -\Gamma^+$. The proof is by induction on $|\Phi^+ \cap \Gamma^-|$. If this is 0 then $\Phi^+ = \Gamma^+$, and we set $w = 1 \in \mathrm{W}(\Phi)$.

Let $0 < k = |\Phi^+ \cap \Gamma^-|$. Then $\Delta$ is not contained in $\Gamma^+$, so there is a $\delta \in \Delta \cap \Gamma^-$. By (a)

$$k - 1 = |(\Phi^+)^{r_\delta} \cap \Gamma^-| = |\Phi^+ \cap (\Gamma^-)^{r_\delta}|$$

As $(\Gamma^-)^{r_\delta} = -(\Gamma^+)^{r_\delta}$, by induction there is a $v \in \mathrm{W}(\Phi)$ with $(\Gamma^+)^{r_\delta} = (\Phi^+)^v$. That is, $\Gamma^+ = (\Phi^+)^{v r_\delta}$, as desired.

(c) $\Pi$ is the set of roots in the positive system $\Gamma^+ = \Phi \cap \left(\sum_{\pi \in \Pi} \mathbb{N}\pi\right)$ that cannot be written as a sum of two other roots in $\Gamma^+$, so the $w$ of (b) works here as well. □

Thus positive systems and simple bases are uniquely determined up to the action of the Weyl group. Conversely, each simple basis determines the Weyl group.

**(7.7).** THEOREM. *Let $(E, \Phi)$ be a root system and $\Pi = \{\pi_1, \ldots, \pi_l\}$ a simple basis. Then $\mathrm{W}(\Phi) = \mathrm{W}(\Pi)$ is a finite group with every element of $\{\, r_\alpha \mid \alpha \in \Phi \,\}$ conjugate to some element of $\{\, r_\pi \mid \pi \in \Pi \,\}$.*

PROOF. The Weyl group $\mathrm{W}(\Phi)$ permutes the finite set $\Phi$ and so induces a finite group of permutations. This permutation group is a faithful representation of $\mathrm{W}(\Phi)$ since $\Phi$ spans $E$.

By the transitivity result of the previous proposition, we may assume that $\Pi = \Delta$ and $\pi_k = \delta_k$.

As $\alpha^{r_\alpha} = -\alpha$ and $r_\alpha^g = r_{\alpha^g}$ (as in Lemma (7.1)(b)), it is enough to show that for each $\alpha \in \Phi^+$ there is an element $w$ of $\mathrm{W}(\Delta)$ with $\alpha^w \in \Delta$. We do this by induction on the height $\mathrm{ht}(\alpha)$. If $\mathrm{ht}(\alpha) = 1$, then $\alpha \in \Delta$ and there is nothing to prove.

Assume $\mathrm{ht}(\alpha) > 1$. Let $\alpha = \sum_{i=1}^{l} d_i \delta_i$ with $d_i \in \mathbb{N}$ by Theorem (7.4)(a). As

$$0 < (\alpha, \alpha) = \left(\alpha, \sum_{i=1}^{l} d_i \delta_i\right) = \sum_{i=1}^{l} d_i (\alpha, \delta_i)$$

there is an $j$ with $d_j > 0$ and $(\alpha, \delta_j) > 0$ hence $(\alpha, \delta_j^\vee) > 0$. Without loss we may take $j = 1$.

Since $\mathrm{ht}(\alpha) > 1$ there must be a second index $k \neq 1$ with $d_k > 0$. Then

$$\begin{aligned} \alpha^{r_{\delta_1}} &= \alpha - (\alpha, \delta_1^\vee)\delta_1 \\ &= (d_1 - (\alpha, \delta_1^\vee))\delta_1 + \sum_{i=2}^{l} d_i \delta_i \,. \end{aligned}$$

Because $d_k > 0$ the root $\alpha^{r_{\delta_1}}$ remains positive, but since $(\alpha, \delta_1^\vee) > 0$ its height is less than that of $\alpha$. Therefore, by induction there is a $u \in \mathrm{W}(\Delta)$ with $(\alpha^{r_{\delta_1}})^u \in \Delta$, hence $\alpha^w \in \Delta$ for $w = r_{\gamma_1} u \in \mathrm{W}(\Delta)$. $\square$

## 7.2 Graphs and diagrams

**(7.8).** LEMMA. *Let $\alpha$ and $\beta$ be independent vectors in the Euclidean space $E$. Then $\langle r_\alpha, r_\beta \rangle$ is a dihedral group in which the rotation $r_\alpha r_\beta$ generates a normal subgroup of index 2 and order $m_{\alpha,\beta}$ (possibly infinite) and the nonrotation elements are all reflections of order 2. In particular, the group $\langle r_\alpha, r_\beta \rangle$ is finite, of order $2m_{\alpha,\beta}$, if and only if the 1-spaces spanned by $\alpha$ and $\beta$ meet at the acute angle $\frac{\pi}{m_{\alpha,\beta}}$.* $\square$

The *Coxeter graph* of the set of simple roots $\Delta$ has $\Delta$ as vertex set, with $\alpha$ and $\beta$ connected by a bond of strength $m_{\alpha,\beta} - 2$ where $\langle r_\alpha, r_\beta \rangle$ is dihedral of order $2m_{\alpha,\beta}$. In particular, distinct $\alpha$ and $\beta$ are not connected if and only if they commute. The Coxeter graph is *irreducible* if it is connected.

**(7.9).** LEMMA. *If $\Sigma$ is an irreducible component of the Coxeter graph of $\Delta$, then $E = \sum_{\sigma \in \Sigma} \mathbb{R}\sigma \perp \sum_{\gamma \in \Delta \setminus \Sigma} \mathbb{R}\gamma$ and*

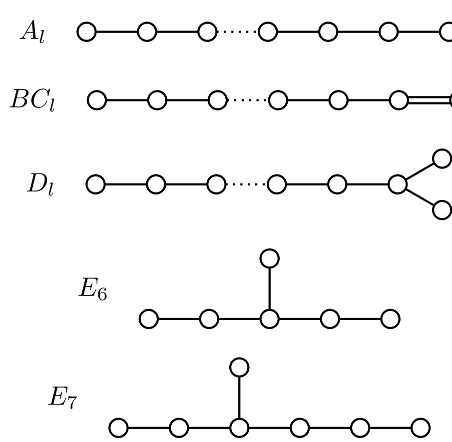
$$\mathrm{W}(\Phi) = \mathrm{W}(\Delta) = \mathrm{W}(\Sigma) \oplus \mathrm{W}(\Delta \setminus \Sigma) = \mathrm{W}(\Phi_\Sigma) \oplus \mathrm{W}(\Phi_{\Delta \setminus \Sigma})$$

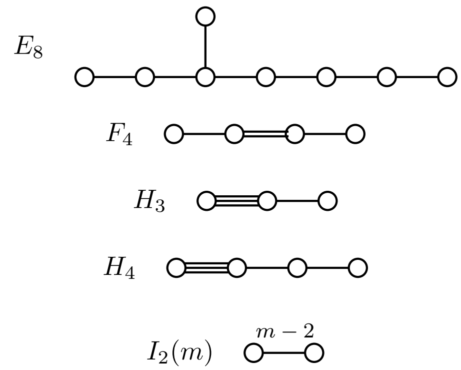
*where $\Phi_\Sigma = \Sigma^{\mathrm{W}(\Phi)} = \Sigma^{\mathrm{W}(\Sigma)}$ and $\Phi_{\Delta\setminus\Sigma} = (\Delta \setminus \Sigma)^{\mathrm{W}(\Phi)} = (\Delta \setminus \Sigma)^{\mathrm{W}(\Delta\setminus\Sigma)}$. Here $\Phi_\Sigma$ and $\Phi_{\Delta\setminus\Sigma}$ are perpendicular and have union $\Phi$.*

PROOF. This is an immediate consequence of Lemmas (7.1) and (7.8) and of Theorem (7.7). $\square$

We repeat Theorem B-(2.3) from Appendix B.

**(7.10).** THEOREM. *The Coxeter graph for an irreducible finite group generated by the $l$ distinct Euclidean reflections for an obtuse basis is one of the following:*






$\square$

It is not at all clear which Coxeter graphs actually correspond to root systems. Two properties of root systems, especially the Crystallographic Condition, play no role in the proof of the previous theorem. We next see that only a few of the graphs $I_2(m)$ can actually occur if the Coxeter graph comes from a root system.

**(7.11).** PROPOSITION. *Let $\alpha, \beta \in \Phi$ with $\alpha \neq \pm\beta$. Then, up to order of $\alpha, \beta$ and admissible rescaling, we have one of*

| $(\alpha,\beta^\vee)(\beta,\alpha^\vee)$ | $\cos(\pi/m_{\alpha,\beta})$ | $m_{\alpha,\beta}$ | $(\alpha,\beta^\vee)$ | $(\beta,\alpha^\vee)$ | $(\alpha,\alpha)$ | $(\beta,\beta)$ | $(\alpha,\beta)$ |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 2 | 0 | 0 | $*$ | 1 | 0 |
| 1 | $\frac{1}{2}$ | 3 | $-1$ | $-1$ | 1 | 1 | $-\frac{1}{2}$ |
| 2 | $\frac{\sqrt{2}}{2}$ | 4 | $-2$ | $-1$ | 2 | 1 | $-1$ |
| 3 | $\frac{\sqrt{3}}{2}$ | 6 | $-3$ | $-1$ | 3 | 1 | $-\frac{3}{2}$ |

PROOF. For all $\alpha, \beta \in \Phi$ we have

$$(\alpha,\beta^\vee)(\beta,\alpha^\vee) = \frac{4(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)} = 4\left(\cos\left(\frac{\pi}{m_{\alpha,\beta}}\right)\right)^2 \in \mathbb{Z}\,.$$

This must be an integer factorization $(\alpha,\beta^\vee)(\beta,\alpha^\vee)$ in the range 0 to 4. Indeed 4 could only happen for $\alpha = \pm\beta$, which has been excluded. Therefore we have the four possibilities of the first column.

In the second column, we then have $\cos(\pi/m_{\alpha,\beta}) = \frac{1}{2}\sqrt{(\alpha,\beta^\vee)(\beta,\alpha^\vee)}$, where we are in the first quadrant since $m_{\alpha,\beta}$, the order of $r_\alpha r_\beta$, is at least 2. We then have $m_{\alpha,\beta} = \frac{\pi}{\arccos(c)}$, where $c$ is the cosine value from the preceding column.

We have not yet chosen order or scaling for $\alpha$ and $\beta$, and we do that in the next two columns while choosing the factorization of $(\alpha,\beta^\vee)(\beta,\alpha^\vee)$. If necessary, we replace $\beta$ by $-\beta$ so that both $(\alpha,\beta^\vee)$ and $(\beta,\alpha^\vee)$ are nonpositive.

Next we rescale the pair $\alpha, \beta$ so that $(\beta,\beta) = 1$ always and note that

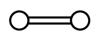
$$\frac{(\alpha,\beta^\vee)}{(\beta,\alpha^\vee)} = \frac{(\alpha,\alpha)}{(\beta,\beta)}\,.$$

This gives us the next two columns of the table, although in the first line we have no information about the squared length of $\alpha$.

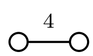
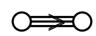
Finally as $(\beta, \beta) = 1$, we have

$$(\alpha, \beta) = \frac{1}{2}\left(\frac{2(\alpha, \beta)}{(\beta, \beta)}\right) = \frac{1}{2}(\alpha, \beta^\vee)\,. \qquad \square$$

The *Dynkin diagram* of $\Delta$ is essentially a directed version of its Coxeter graph. In accordance with the previous proposition, each two node subgraph of the Coxeter graph is replaced with a new, possibly directed, edge in the Dynkin diagram. All $A_1 \times A_1$ edges (that is, nonedges) and $A_2$ edges (single bonds) are left undisturbed. On the other hand

$BC_2$  becomes $B_2 = C_2$ 

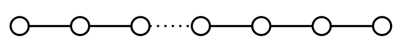
Similarly

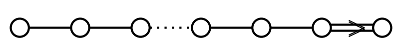
$I_2(6)$  becomes $G_2$ 

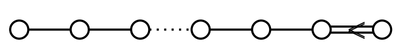
The arrow (or "greater than") sign on the edge is there to indicate that the root at the tail (or "big") end is longer than the root at the tip ("small") end. Also notice that $G_2$ has three bonds rather than four. In all cases the number of bonds between the notes $\alpha$ and $\beta$ is the integer $(\alpha, \beta^\vee)(\beta, \alpha^\vee)$. This change in notation reminds us that ion $G_2$ the *long root* has squared length 3 times that of the *short root*, as in the table of the proposition. Similarly in $B_2 = C_2$, the long root has squared length equal to twice that of the short root. (The roots at the two ends of $A_2$ have equal length.)

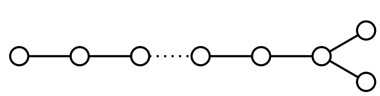
By the proposition, in classifying Dynkin diagrams we need only consider Coxeter graphs for which all $m_{\alpha,\beta}$ come from $2, 3, 4, 6$. In particular $H_3$ and $H_4$ do not lead to root systems nor do the $I_2(m)$, except for $A_2 = I_2(2)$, $B_2 = C_2 = I_2(4)$, and $G_2 = I_2(6)$. The need for both names $B_2 = C_2$ becomes clearer when we combine the previous two results to find:

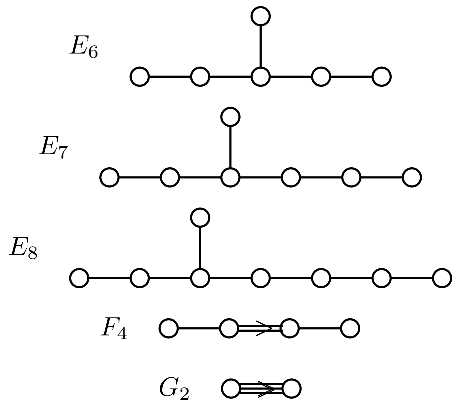
**(7.12).** THEOREM. *The Dynkin diagram for an irreducible abstract root system of rank $l$ is one of the following:*

$A_l$ 

$B_l$ 

$C_l$ 

$D_l$ 




□

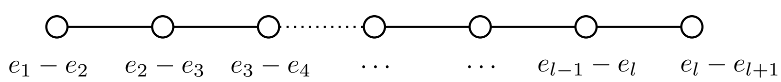
## 7.3 Existence of root systems

### 7.3.1 The classical root systems—$A_l$, $B_l$, $C_l$, $D_l$

A simply-laced Coxeter graph is its own Dynkin diagram. For the graphs/diagrams $A_l$ and $D_l$ we found (in Appendix Section B.2) the groups generated by reflections to be the symmetric group $\mathrm{W}(A_l) = \mathrm{Sym}(l+1)$ (acting on the $l$-space $\mathbb{R}^l = \mathbb{R}^{l+1} \cap \mathbf{1}^\perp$) and $\mathrm{W}(D_l) = 2^{l-1} : \mathrm{Sym}(l)$ acting monomially on $\mathbb{R}^l$. The corresponding root systems are:

$$\Phi_{A_l} = \{\, \pm(e_i - e_j) \mid 1 \le i < j \le l+1 \,\}$$

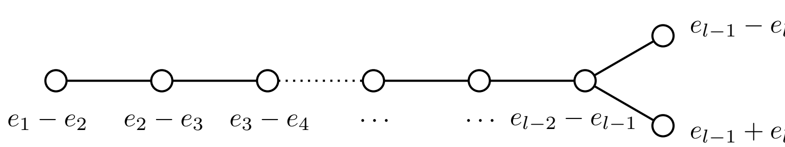
with simple basis




and

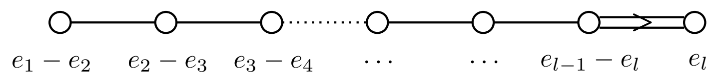
$$\Phi_{D_l} = \{\, \pm e_i \pm e_j \mid \ 1 \le i < j \le l \,\}$$

with simple basis




In contrast to this, the Coxeter graph $BC_l$ admits two Dynkin diagrams, depending upon whether or not the tail consists of long roots ($B_l$) or short roots ($C_l$). In both cases the Weyl group consists of all $\pm 1$-monomial matrices $\mathrm{W}(B_l) = \mathrm{W}(C_n) = 2^l : \mathrm{Sym}(l)$. (Again, see Appendix Section B.2.) But the root systems and simple bases are subtly different.

For $B_l$ the simple basis is given by




So $\Phi_{B_l}$ is the union $\Phi_{B_l}^{\text{long}} \cup \Phi_{B_l}^{\text{short}}$ of long root set

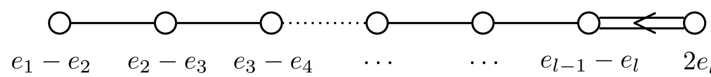
$$\Phi_{B_l}^{\text{long}} = \{\, \pm(e_i - e_j) \mid 1 \leq i < j \leq l \,\}$$

with squared length 2 and short root set

$$\Phi_{B_l}^{\text{short}} = \{\, \pm e_i \mid 1 \leq i \leq l \,\}$$

with squared length 1.

In contrast, for $C_n$ we have




The root system $\Phi_{B_l}$ is then the union $\Phi_{C_l}^{\text{short}} \cup \Phi_{C_l}^{\text{long}}$ of short root set

$$\Phi_{C_l}^{\text{short}} = \{\, \pm(e_i - e_j) \mid 1 \leq i < j \leq l \,\}$$

with squared length 2 and long root set

$$\Phi_{C_l}^{\text{long}} = \{\, \pm 2e_i \mid 1 \leq i \leq l \,\}$$

with squared length 4.

### 7.3.2 $E_8$

Our (candidate) root system $\Phi$ is the union $\Phi_{D_8} \cup \Phi_{\text{even}}$ of

$$\Phi_{D_8} = \{\, \pm e_i \pm e_j \mid 1 \leq i < j \leq 8 \,\}$$

(a root system of type $D_8$) and

$$\Phi_{\text{even}} = \{\, \frac{1}{2}(\pm e_1 \pm e_2 \pm e_3 \pm e_4 \pm e_5 \pm e_6 \pm e_7 \pm e_8) \mid \text{even number of minus signs} \,\}.$$

This set has cardinality

$$|\Phi_{D_8}| + |\Phi_{\text{even}}| = 4\binom{8}{2} + 2^7 = 112 + 128 = 240\,.$$

All the roots have squared length 2 (especially, for each root $\alpha = \alpha^\vee$). To verify that this is an abstract root system we must check the four conditions of Definition (7.2). The first two are trivial: (i) $\Phi$ spans $\mathbb{R}^8$ (indeed its subsystem $\Phi_{D_8}$ does already) and (ii) the only scalar multiples of the root $\alpha$ that are roots are $\pm\alpha$ (since all roots have squared length 2).

The reflection in the root $\alpha \in \Phi$ is given by

$$r_\alpha \colon v \longrightarrow v - (v, \alpha^\vee)\alpha = v - (v, \alpha)\alpha$$

(as $\alpha = \alpha^\vee$). We then see easily that the even monomial group in $\mathrm{GL}_8(\mathbb{R})$

$$\mathrm{W}(\Phi_{D_8}) = \mathrm{W}(D_8) = 2^7 : \mathrm{Sym}(8)\,,$$

leaves invariant both $\Phi_{D_8}$ (clearly) and $\Phi_{\text{even}}$. Indeed it is transitive on $\Phi_{\text{even}}$ (as it contains all even sign changes).

By transitivity, to check (7.2)(iii) invariance under reflections and (7.2)(iv) the Crystallographic Condition we need only test a single root from $\Phi_{\text{even}}$, $\alpha = \frac{1}{2}(e_1+e_2+e_3+e_4+e_5+e_6+e_7+e_8)$ being a good choice. Clearly this has integral inner product in $\{-1, 0, +1\}$ with each root of $\mathrm{W}(\Phi_{D_8})$. Also, since each root $\beta$ of $\Phi_{\text{even}}$ contains an even number of minus signs, the inner product of $\alpha$ and $\beta$ is also an integer, again from $\{-1, 0, +1\}$. This verifies the Crystallographic Condition.

To check (7.2)(iii) we must show that for each for root $\beta \in \Phi$, the vector $r_\alpha(\beta)$ is also a root of $\Phi$. There are five cases:

(a) $\beta \in \Phi_{\text{even}}$ contains 0 or 8 minus signs;

(b) $\beta \in \Phi_{\text{even}}$ contains 2 or 6 minus signs;

(c) $\beta \in \Phi_{\text{even}}$ contains 4 minus signs;

(d) $\beta = \pm(e_i - e_j) \in \Phi_{D_8}$;

(e) $\beta = \pm(e_i + e_j) \in \Phi_{D_8}$.

In the first case, $\beta = \pm\alpha$ and there is nothing to prove. In cases (c) and (d) we have $(\beta, \alpha) = 0$ so that $r_\alpha(\beta) = \beta \in \Phi$, as desired.

In case (b) we combine the cases of two and six minus signs together as these are negatives of each other. If a root with six minus signs is taken to a root, then its negative (with two minus signs) is also taken to a root—the negative of the previous result. Similarly under (e) we only need to consider the various $e_i + e_j$. We calculate

$$r_\alpha(e_1 + e_2) = \frac{1}{2}\Big(e_1 + e_2 - e_3 - e_4 - e_5 - e_6 - e_7 - e_8\Big) \in \Phi\,,$$

hence

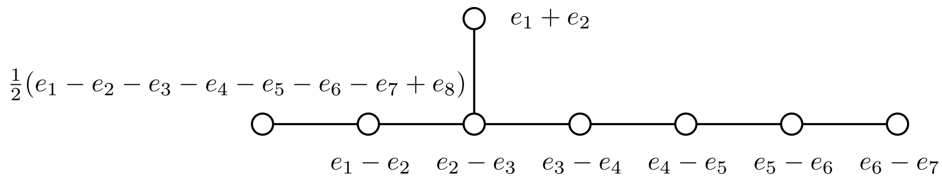
$$r_\alpha\Big(\frac{1}{2}\Big(e_1 + e_2 - e_3 - e_4 - e_5 - e_6 - e_7 - e_8\Big)\Big) = e_1 + e_2 \in \Phi\,.$$

All other cases follow by symmetry, as the group $\mathrm{W}(D_8)$ contains $\mathrm{Sym}(8)$.

We have now proven $\Phi$ to be an abstract root system. In $\mathbb{R}^8$ we order the orthonormal basis:

$$e_1 > e_2 > e_3 > e_4 > e_5 > e_6 > e_7 > e_8\,.$$

This provides us with a unique partition into positive and negative roots $\Phi^+ \cup \Phi^-$ and a unique simple basis $\Delta$ within $\Phi^+$. This basis is given by:



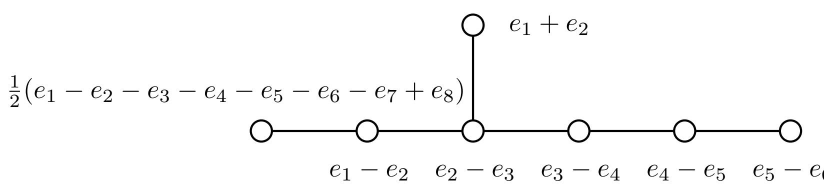

We see, in particular, that $\Delta$ and $\Phi$ belong to the Dynkin diagram $E_8$. We now can set $\Phi = \Phi_{E_8}$, a root system of type $E_8$.

### 7.3.3 $E_7$

The root system $\Phi_{E_7}$ of type $E_7$ is the intersection $\Phi_{E_8} \cap \mathbb{R}^7$ of the root system $\Phi_{E_8}$ with the appropriate hyperplane $\mathbb{R}^7 = \mathbb{R}^8 \cap (e_7 + e_8)^\perp$ of $\mathbb{R}^8$. The number of its roots is $2 + 4\binom{6}{2} + 2^6 = 126$.

We need not check the conditions for an abstract root system since they are inherited from $\Phi_{E_8}$. (Certainly the Crystallographic Condition still holds. And any of the reflections centered in $\Phi_{E_7}$ fix $\Phi_{E_8}$ but also fix $(e_7 + e_8)^\perp$.)

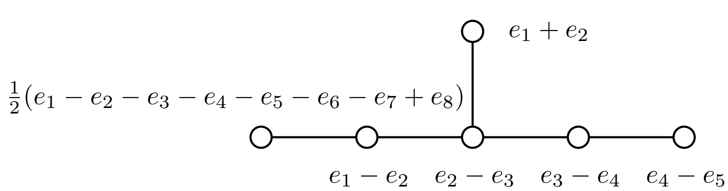
The appropriate simple base is now:




### 7.3.4 $E_6$

The root system $\Phi_{E_6}$ of type $E_6$ is the intersection $\Phi_{E_8} \cap \mathbb{R}^6$ of the root systems $\Phi_{E_8}$ and $\Phi_{E_7}$ with the appropriate hyperplane $\mathbb{R}^6 = \mathbb{R}^8 \cap (e_7 + e_8)^\perp \cap (e_6 - e_7)^\perp$ of the space $\mathbb{R}^7 = \mathbb{R}^8 \cap (e_7 + e_8)^\perp$. The number of its roots is $4\binom{5}{2} + 2^5 = 72$.

Again the conditions for an abstract root system are inherited from $\Phi_{E_8}$ and $\Phi_{E_7}$. The appropriate simple base is:

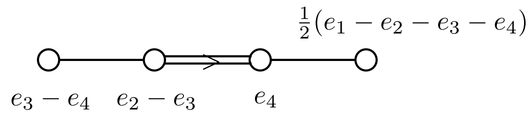


### 7.3.5 $F_4$

The root system $\Phi_{F_4}$ consists of the roots of squared length 1 and 2 in $\mathbb{Z}^4 \cup \mathbb{Z}(\frac{1}{2}(e_1+e_2+e_3+e_4))$. Thus we have the $8+4\binom{4}{2}+2^4 = 48$ roots of

$$\Phi_{F_4} = \{\, \pm e_i\,,\ \pm e_i \pm e_j\,,\ \frac{1}{2}(\pm e_1 \pm e_2 \pm e_3 \pm e_4) \mid 1 \le i < j \le 4 \,\}$$

with Dynkin diagram

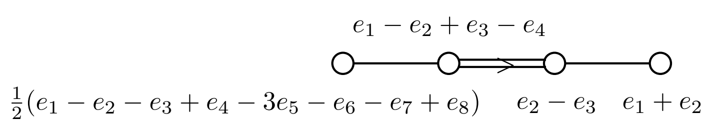



The first two defining properties of a root system are immediate, and the remaining two can be checked easily in a similar manner to our treatment of $\Phi_{E_8}$ above. The second pair of defining properties can also be derived from those for $\Phi_{E_8}$, actually $\Phi_{E_6}$, if we think of $F_4$ as the fixed points of the graph automorphism for $E_6$ given by

$$\frac{1}{2}(e_1 - e_2 - e_3 - e_4 - e_5 - e_6 - e_7 + e_8) \quad \longleftrightarrow \quad e_4 - e_5$$

$$e_1 - e_2 \quad \longleftrightarrow \quad e_3 - e_4$$

with $e_1 + e_2$ and $e_2 - e_3$ fixed. Summing the orbits of length two, we find the $F_4$ Dynkin diagram and a suitable basis of simple roots:



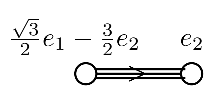

within

$$\mathbb{R}^4 = \mathbb{R}^8 \cap (e_7+e_8)^\perp \cap (e_6-e_7)^\perp \cap (e_1-e_2+e_3+3e_4)^\perp \cap (e_5-e_6-e_7+e_8)^\perp\,.$$

This system is not isometric to the previous one but is similar, as all roots here have squared length equal to twice what they have in the previous rendition.

### 7.3.6 $G_2$

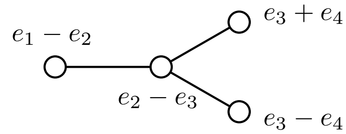
It is easy to draw the root systems of rank 2—those with dihedral Weyl group. Here the group is dihedral of order 12, and the root system contains exactly 12 roots. An appropriate simple basis is:



As was the case for $F_4$, there is a less direct but very helpful second way of constructing the $G_2$ root system. Consider the root system

$$\Phi_{D_4} = \{\, \pm e_i \pm e_j \mid 1 \leq i < j \leq 4 \,\}$$

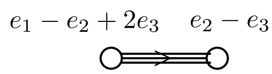
with Dynkin diagram and simple basis




This diagram has a *triality graph automorphism* of order three which rotates the diagram clockwise. If, as before, we add the roots from the various orbits, then the root $e_2 - e_3$ is fixed, while the orbit of length three gives us the new root

$$(e_1 - e_2) + (e_1 + e_4) + (e_3 - e_4) = e_1 - e_2 + 2e_3$$

and we have the Dynkin diagram of the new root system of fixed points:




Again this has type $G_2$ but is now living in the Euclidean space

$$\mathbb{R}^2 = \mathbb{R}^3 \cap (e_1 - e_2 - e_3)^\perp = \mathbb{R}^4 \cap (e_1 - e_2 - e_3)^\perp \cap e_4^\perp \,.$$

These two renditions of the $G_2$ root system are not isometric, but they are equivalent: in the first, the two square lengths are 3 and 1 while in the second they are 6 and 2. As must be the case, they both have 3 as the ratio of the length of a long root to that of a short root.

## 7.4 The Cartan matrix and uniqueness I

The integers $(\alpha, \beta^\vee)$ with $\alpha, \beta$ in the root system $\Phi$ are the *Cartan integers*. For the simple basis $\Delta$ the *Cartan matrix* $\mathrm{Cart}(\Delta)$ of $\Delta$ is the $l \times l$ integer matrix with $(i, j)$ entry the Cartan integer $c_{i,j} = (\delta_i, \delta_j^\vee)$. Especially all diagonal entries are $(\delta, \delta^\vee) = 2$.

We record the fundamental properties of the Cartan matrix:

**(7.13).** Lemma.

(a) $c_{i,j} \in \mathbb{Z}$;

(b) $c_{i,i} = 2$ *for all* $i$;

(c) $c_{i,j} \leq 0$ *for all* $i \neq j$;

(d) $c_{i,j} \neq 0$ *if and only if* $c_{j,i} \neq 0$. □

**(7.14).** THEOREM.

(a) *The Dynkin diagram and the Cartan matrix contain exactly the same information about the root system* $(E, \Phi)$.

(b) *The Dynkin diagram and Cartan matrix determine the root system* $(E, \Phi)$ *uniquely up to equivalence.*

PROOF. (a) By Proposition (7.11), in the pair of positions $(\alpha, \beta)$ and $(\beta, \alpha)$ of the Cartan matrix the entries are

$$\{(\alpha, \beta^\vee), (\beta, \alpha^\vee)\} \in \{\{0,0\}, \{-1,-1\}, \{-2,-1\}, \{-3,-1\}\}\,.$$

The number of bonds between the two simple roots in the Dynkin diagram is the product of the two elements of the pair. An arrow on an edge between $\alpha$ and $\beta$ is pointed away from $\alpha$ precisely when $|(\alpha, \beta^\vee)| > 1$.

(b) We may assume the Dynkin diagram to be irreducible. As in (a), once we choose a squared length for one of the simple roots all the squared lengths are known (invoking equivalence). The Gram matrix for the simple basis is thus determined (as the Cartan matrix times the diagonal matrix with entries $(\beta, \beta)/2$). That is, the simple basis is determined up to an isometry. By Theorem (7.7), the simple roots determine the reflections of a generating set for the Weyl group and then the root system itself consists of the images of the simple roots under that Weyl group. □

# Chapter 8

# Semisimple Lie Algebras: Classification

We continue to examine finite dimensional semisimple Lie algebras over algebraically closed fields of characteristic 0. The previous two chapters effectively complete the first two steps of the classification, as described at the beginning of Chapter 6:

(i) for each algebra, the construction of a root system that functions as a skeleton;

(ii) the classification of root systems.

More precisely, we have shown that each algebra is an amalgamation of subalgebras isomorphic to $\mathfrak{sl}_2(\mathbb{K})$ with the amalgamation encoded by one of a known collection of root systems.

It remains to prove that for each root system (Dynkin diagram/Cartan matrix) and each $\mathbb{K}$ a corresponding simple Lie algebra does exist and is unique. This is (largely) accomplished in this chapter.

**(8.1).** Theorem. (Classification of semisimple Lie algebras) *Let $L$ be a finite dimensional semisimple Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic $0$. Then $L$ can be expressed uniquely as a direct sum of simple subalgebras. Each simple subalgebra is isomorphic to exactly one of the following, where in each case the rank is $l$:*

(a) $\mathfrak{a}_l(\mathbb{K}) \simeq \mathfrak{sl}_{l+1}(\mathbb{K})$, *for rank* $l \geq 1$, *of dimension* $l^2 + 2l$;

(b) $\mathfrak{b}_l(\mathbb{K}) \simeq \mathfrak{so}_{2l+1}(\mathbb{K})$, *for rank* $l \geq 3$, *of dimension* $2l^2 + l$;

(c) $\mathfrak{c}_l(\mathbb{K}) \simeq \mathfrak{sp}_{2l}(\mathbb{K})$, *for rank* $l \geq 2$, *of dimension* $2l^2 + l$;

(d) $\mathfrak{d}_l(\mathbb{K}) \simeq \mathfrak{so}_{2l}(\mathbb{K})$, *for rank* $l \geq 4$, *of dimension* $2l^2 - l$;

(e) $\mathfrak{e}_6(\mathbb{K})$ *of rank* $l = 6$ *and dimension* 78*;*

(f) $\mathfrak{e}_7(\mathbb{K})$ *of rank* $l = 7$ *and dimension* 133*;*

(g) $\mathfrak{e}_8(\mathbb{K})$ *of rank* $l = 8$ *and dimension* 248*;*

(h) $\mathfrak{f}_4(\mathbb{K})$ *of rank* $l = 4$ *and dimension* 52*;*

(i) $\mathfrak{g}_2(\mathbb{K})$ *of rank* $l = 2$ *and dimension* 14.

*None of these simple algebras is isomorphic to one from another case or to any other algebra from the same case. All exist.*

## 8.1 Reduction to the irreducible, simple case

We have encountered various concepts of *irreducibility.* A reflection group is irreducible if it acts irreducibly on its underlying space. A Coxeter graph or Dynkin diagram is irreducible if it is connected. A root system is irreducible if it is not the perpendicular direct sum of two proper subsystems. A Cartan matrix (see below) is irreducible if it cannot be written as a direct sum of two smaller Cartan matrices.

In the context of interest to us, semisimple Lie algebras, all of these concepts are equivalent.[1] The philosophy is always that in a classification one should easily reduce to the irreducible case. This remains true with our semisimple Lie algebras.

**(8.2).** THEOREM. *A finite dimensional semisimple Lie algebra over the algebraically closed field* $\mathbb{K}$ *is the perpendicular direct sum of its minimal ideals, all simple Lie algebras.*

*A semisimple algebra is simple if and only if its Dynkin diagram is irreducible, and the simple summands of the previous paragraph are in bijection with with irreducible components of the Dynkin diagram of the algebra.*

PROOF. The first paragraph is essentially a restatement of Theorem (5.20).

Let $I$ be an ideal of the semisimple Lie algebra $L$. As the Cartan subalgebra $H$ is diagonal in its adjoint action on $L$ (by Theorem (6.4)(a)), the ideal $I$ is the direct sum of its intersection $H \cap I$ and the $L_\lambda$ (of dimension 1 by the same theorem) for $\lambda$ in some subset $\Lambda_I$ of $\Phi$. Furthermore, as $L$ is generated as an algebra by the $L_\delta$ for $\delta \in \Delta$, we must have $\Delta_I = \Lambda_I \cap \Delta$ nonempty.

By Theorem (5.20) there is an ideal $J$ with $L = I \oplus J$. If $\delta \in \Delta_I$ and $\gamma \in \Delta_J$, then

$$[L_\delta, L_\gamma] \leq L_{\delta+\gamma} \leq I \cap J = 0\,.$$

Therefore $[L_\delta, L_\gamma] = 0$, so $\delta + \gamma \notin \Phi$ by Theorem (6.9)(c). Thus $\delta$ and $\gamma$ are not connected in the Dynkin diagram of $\Delta$ by Proposition (7.11). That is, $\Delta_I$ is a union of irreducible components of $\Delta$.

[1] Luckily.

Conversely, suppose that $\Sigma$ is an irreducible component of $\Delta$ and hence of the corresponding Coxeter graph. Then by Lemma (7.9) the root system $\Phi$ is the union of the perpendicular subsystems $\Phi_\Sigma = \Phi \cap \bigoplus_{\sigma\in\Sigma} \mathbb{Z}\sigma$ and $\Phi_{\Delta\setminus\Sigma} = \Phi \cap \bigoplus_{\delta\in\Delta\setminus\Sigma} \mathbb{Z}\delta$. Therefore $\Phi_\Sigma$ is the root system for the subalgebra $L_\Sigma$ generated by the $L_\sigma$ for $\sigma \in \pm\Sigma$, an ideal of $L$.

We have now shown that ideals come from disjoint unions of irreducible components of $\Delta$ and that irreducible subdiagrams correspond to ("generate") ideals perpendicular to all others. In particular, the simple ideals are in bijection with the irreducible components of the Dynkin diagram. □

## 8.2 The Cartan matrix and uniqueness II

We assume the notation of Theorem (8.1). Additionally, in the root system $(E^L, \Phi^L) = (E, \Phi)$ we choose a partition $\Phi = \Phi^+ \cup \Phi^-$ associated with the simple basis $\Delta = \{\delta_1, \dots, \delta_l\}$. The integers $(\alpha, \beta^\vee)$ with $\alpha, \beta \in \Phi$ are the *Cartan integers.*

**(8.3).** PROPOSITION. *Let $\alpha \in \Phi^+$. Then with $k$ the height of $\alpha$ there are $\alpha_a \in \Delta$ for $1 \le a \le k$ with*

$$\sum_{a=1}^{b} \alpha_a \in \Phi^+ \text{ for each } 1 \le b \le k \quad \text{and} \quad \alpha = \sum_{a=1}^{k} \alpha_a \,.$$

PROOF. The proof is by induction on $k = \mathrm{ht}(\alpha)$. If $k = 1$, then $\alpha = \alpha_1 \in \Delta$, and we are done. Assume $k > 1$. Let $\alpha = \sum_{i=1}^{l} d_i\delta_i$.

We have

$$0 < (\alpha, \alpha) = \sum_{i=1}^{l} d_i(\alpha, \delta_i)\,,$$

so some $(\alpha, \delta_j)$ is positive as is the integer $(\alpha, \delta_j^\vee)$ . Without loss we may assume $j = 1$.

The root

$$\alpha^{r_{\delta_1}} = \alpha - (\alpha, \delta_1^\vee)\delta_1 = (d_1 - (\alpha, \delta_1^\vee))\delta_1 + \sum_{i=2}^{l} d_i\delta_i$$

belongs to the $\delta_1$-string through $\alpha$, as does $\alpha$ itself. By Theorem (6.9)

$$\beta = \alpha - \delta_1 = (d_1 - 1)\delta_1 + \sum_{i=2}^{l} d_i\delta_i$$

is also a root in that string and has height $k - 1 > 0$. Especially it is positive. Therefore by induction there are $\beta_a \in \Delta$ for $1 \le a \le k - 1$ with

$$\sum_{a=1}^{b} \beta_a \in \Phi^+ \text{ for each } 1 \le b \le k-1 \quad \text{and} \quad \beta = \sum_{a=1}^{k-1} \beta_a \,.$$

As $\alpha = \beta + \delta_1$, with $\alpha_a = \beta_a$ for $1 \le a \le k-1$ and $\alpha_k = \delta_1$, we have the result. □

**(8.4).** COROLLARY. *Let $\gamma \in \Phi^-$. Then with $k$ the height of $\gamma$ there are $\gamma_a \in -\Delta$ for $1 \le a \le -k$ with*

$$\sum_{a=1}^{b} \gamma_a \in \Phi^- \text{ for each } 1 \le b \le -k \quad \text{and} \quad \gamma = \sum_{a=1}^{-k} \gamma_a \,.$$

PROOF. Set $\alpha = -\gamma$ and then $\gamma_a = -\alpha_a$. □

Recall that the integers $(\alpha, \beta^\vee)$ with $\alpha, \beta \in \Phi$ are the *Cartan integers*. For the simple basis $\Delta$ the *Cartan matrix* $\mathrm{Cart}(\Delta)$ of $\Delta$ is the $l \times l$ integer matrix with $(i,j)$ entry the Cartan integer $c_{i,j} = (\delta_i, \delta_j^\vee)$. so that diagonal entries are $(\delta, \delta^\vee) = 2$. The Cartan matrix of $\Delta$ is often called the Cartan matrix of $L$, although this terminology is currently loose for us since we have not shown that all Cartan subalgebras are equivalent (but see Corollary (8.24)).

Choose $e_i \in L_{\delta_i}$ and $e_{-i} \in L_{-\delta_i}$ and set $h_i = [e_i, e_{-i}]$. Do this in accordance with Theorem (6.6) so that $S_i = \mathbb{K}h_i \oplus \mathbb{K}e_i \oplus \mathbb{K}e_{-i}$ is isomorphic to $\mathfrak{sl}_2(\mathbb{K})$ with the standard relations, which we record along with others in the next proposition.

For $\delta_i, \delta_j \in \Delta$ let $c_{i,j} = (\delta_i, \delta_j^\vee)$ be the associated Cartan integer.

**(8.5).** PROPOSITION. *The Lie algebra $L$ is generated by the elements $e_i, e_{-i}$ for $1 \le i \le l$. We have the following relations in $L$:*

(i) $[h_i, h_j] = 0$ *for all* $1 \le i, j \le l$;

(ii) $[h_i, e_j] = c_{j,i} e_j$ *and* $[h_i, e_{-j}] = -c_{j,i} e_{-j}$ *for all* $1 \le i, j \le l$;

(iii) $[e_i, e_{-i}] = h_i$ *for all* $1 \le i \le l$;

(iv) $[e_i, e_{-j}] = 0$ *for all* $i \ne j$;

(v) $\mathrm{ad}_{e_i}^{1-c_{j,i}}(e_j) = 0$ *and* $\mathrm{ad}_{e_{-i}}^{1-c_{j,i}}(e_{-j}) = 0$ *for* $1 \le i, j \le l$ *with* $i \ne j$.

PROOF. We have the Cartan decomposition

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

By Proposition (6.3) and Theorem (7.4)(c), the $L_\alpha$ have dimension 1 and the $h_i$ generate $H$. By Theorem (6.4)(a) always $[L_\alpha, L_\beta] = L_{\alpha+\beta}$ for $\alpha, \beta \in \Phi$. Therefore by induction on the height of $\gamma \in \Phi$ and using Proposition (8.3) and its corollary, we find that $L_\gamma$ is in the subalgebra generated by the various $e_i, e_{-i}$. That subalgebra is therefore $L$ itself.

Parts (i) and (iii) are part of the definitions for the generating set. Part (d) holds as $\delta_i - \delta_j$ is never a root for $\delta_i, \delta_j \in \Delta$.

For part (ii) with $\epsilon = \pm$

$$[h_i, e_{\epsilon j}] = \delta_{\epsilon j}(h_i)e_{\epsilon j} = \epsilon(\delta_j, \delta_i^\vee)e_{\epsilon j} = \epsilon c_{j,i}e_{\epsilon j}\,.$$

Finally in (v), for $\delta_i, \delta_j \in \Delta$ the $\delta_i$-string through $\delta_j$ is

$$\delta_j\,,\ \delta_j + \delta_i\,,\ \ldots\,,\ \delta_j - (\delta_j, \delta_i^\vee)\delta_i$$

by Theorem (6.9)(a). Noting that $c_{j,i} = c_{-j,-i}$, we have

$$\mathrm{ad}_{e_i}^{1-c_{j,i}}(e_j) \in L_{(1-(\delta_j,\delta_i^\vee))+\delta_j} = L_{(\delta_j-(\delta_j,\delta_i^\vee))+1} = 0\,. \qquad \square$$

The following remarkable result gives uniqueness and existence at the same time for Lie algebras over $\mathbb{K}$ and every abstract root system $(E, \Phi)$. We do not[2] prove this difficult theorem, but we do use its relations (from the proposition) as the entry to our uniqueness proof for $L$.

**(8.6).** Theorem. (Serre's Theorem) *Let $\mathbb{K}$ be an algebraically closed field of characteristic $0$, and let $C = (c_{i,j})_{i,j}$ be the Cartan matrix of the abstract root system $(E, \Phi)$. Then the generators and relations of Proposition (8.5) give a presentation of a semisimple Lie algebra $L$ over $\mathbb{K}$ with Cartan matrix $C$ and root system equivalent to $(E, \Phi)$.* $\square\square$

Serre's Theorem thus reduces the existence and uniqueness question to the easier existence and uniqueness problem for root systems, which we solved in the previous chapter.

## 8.3 Uniqueness

**Some of the arguments here are incomplete. This section will be dropped or changed dramatically at some point.**

**(8.7).** Theorem. *Let $L_1$ and $L_2$ be finite dimensional semisimple Lie algebras over the algebraically closed field $\mathbb{K}$ of characteristic $0$. Then the following are equivalent.*

(1) *$L_1$ and $L_2$ are isomorphic;*

(2) *the associated root systems $(E_1, \Phi_1)$ and $(E_2, \Phi_2)$ are equivalent;*

(3) *the associated simple bases $\Delta_1$ and $\Delta_2$ have isomorphic Dynkin diagrams;*

(4) *the associated simple bases $\Delta_1$ and $\Delta_2$ have equivalent Cartan matrices; that is, there is a permutation matrix $P$ with $\mathrm{Cart}(\Delta_2) = P\,\mathrm{Cart}(\Delta_1)P^\top$.*

[2] at least for now

Parts (3) and (4) are equivalent and both imply (2); see Theorem (7.14). On the other hand (2) implies (3) and (4) by Theorem (7.6)(b).

That (1) implies (2) requires the result (already mentioned) that two Cartan subalgebras are equivalent under an automorphism of $L$. We will prove this later in Corollary (8.24), in an *ad hoc* and after-the-fact manner. See page 109 for our ultimate proof of the theorem.

At present we will deal with the crucial (2) $\Longrightarrow$ (1) part of the theorem above:

**(8.8).** THEOREM. *Let $L$ and $L'$ be finite dimensional semisimple Lie algebras over the algebraically closed field $\mathbb{K}$ of characteristic $0$. Let the associated root systems $(E, \Phi)$ and $(E', \Phi')$ be equivalent. Then $L$ and $L'$ are isomorphic.*

*Indeed the isomorphism of $(E, \Phi)$ and $(E', \Phi')$ extends to an isomorphism of $L$ and $L'$ that takes the Cartan subalgebra $H$ associated with $(E, \Phi)$ to the Cartan subalgebra $H'$ associated with $(E', \Phi')$.*

Of course, this is a direct consequence of Serre's Theorem (10.15), but we have not proved that yet.

Before proving this, we point out an interesting and helpful corollary.

**(8.9).** COROLLARY. *Any nontrivial automorphism of the Dynkin diagram of semisimple $L$ extends to a nontrivial automorphism of $L$.* □

Such automorphisms are usually referred to as *graph automorphisms*.

Our uniqueness proof is motivated by that of [Eld15]. The basic observation is that, with respect to the Cartan basis $\{\, h_i, e_\alpha \mid 1 \leq i \leq l\,,\ \alpha \in \Phi \,\}$, most of the adjoint actions are nearly monomial. We then show (starting as in Proposition (8.3)) that, for an appropriate choice of the basis vectors, the actual multiplication coefficients are rational and depend somewhat canonically upon the root system $\Phi$.

An example is the following working lemma.

**(8.10).** LEMMA. *Let $\delta \in \Delta \cup -\Delta$ and $\beta \in \Phi$ with $\beta \neq \pm\delta$, and let $\beta - s\delta, \ldots, \beta, \ldots, \beta + t\delta$ be the $\delta$-string of roots through $\beta$. Let $S_\delta = \mathbb{K}h_\delta \oplus \mathbb{K}e_\delta \oplus \mathbb{K}e_{-\delta}$ be isomorphic to $\mathfrak{sl}_2(\mathbb{K})$ with the standard relations from Proposition (8.5). Then for $x \in L_\beta$ we have $[e_\delta, [e_{-\delta}, x]] = t(s+1)x$.*

PROOF.

□

We could rephrase this to say: there is a nonzero rational constant $\chi(\delta, \beta)$ depending only on $\delta$ and $\beta$ with

$$\mathrm{ad}_{e_{-\delta}} \mathrm{ad}_{e_\delta} e_\beta = \chi(\delta, \beta) e_\beta\,.$$

This is the model for our uniqueness results below, in particular Theorem (8.12).

For each $\delta \in \Delta \cup -\Delta$, set $a_\delta = \mathrm{ad}_{e_\delta}$. Consider words $w = w_k \dots w_1$ in the alphabet

$$\mathcal{A} = \mathcal{A}^+ \cup \mathcal{A}^- \text{ for } \mathcal{A}^+ = \{\, a_\delta \mid \delta \in \Delta \,\}\,,\ \mathcal{A}^- = \{\, a_\delta \mid \delta \in -\Delta \,\}.$$

If $w = a_{\delta_{i_k}} \cdots a_{\delta_{i_1}}$, then we define $\|w\| = \sum_{j=1}^k \delta_{i_j}$.

For each such word $w$ we set

$$e(w) = w_k \cdots w_2\, e(w_1)\,,$$

where we initialize with $e(a_\delta) = e_\delta$. Note that $e(w) \in L_{\|w\|}$.

By Proposition (8.3) and its corollary, for every $\alpha \in \Phi^\epsilon$ there is a word $w$ in the alphabet $\mathcal{A}^\epsilon$ with $\mathbb{K}e(w) = L_\alpha$. Indeed it is possible to do this with $k = |\operatorname{ht}(\alpha)|$. For each $\alpha$, choose and fix one such word $w_\alpha$ and set $e_\alpha = e(w_\alpha)$. If below we say that something "depends on $\alpha$" we may actually mean that it depends upon $\alpha$ and the fixed choice of representative word $w_\alpha$.

**(8.11).** LEMMA. *For each word $w$ from the alphabet $\mathcal{A}^\epsilon$ there is a constant $\chi_w \in \mathbb{Q}$ with $e(w) = \chi_w e_{\|w\|}$.*

PROOF. SKETCH: Let $w = w_k w_{k-1} \cdots w_1$ and set $w_k = a_\delta$. Use Lemma (8.10) and induction on $k$, with $k = 1$ being immediate. For $\delta, \gamma \in \epsilon\Delta$ always $-\delta + \gamma \notin \Phi$. Thus as endomorphisms $a_{-\delta} a_\gamma = a_\gamma a_{-\delta}$ unless $\gamma = \delta$. □

**(8.12).** THEOREM. *We have $L = \bigoplus_{i=1}^l \mathbb{K}h_i \oplus \bigoplus_{\alpha \in \Phi} \mathbb{K}e_\alpha$ with*

(i) $[h_i, h_j] = 0$;

(ii) $[h_i, e_\alpha] = (\alpha, \delta_i^\vee) e_\alpha$;

(iii) $[e_\alpha, e_\beta] = \chi_{\alpha,\beta} e_{\alpha+\beta}$, $\chi_{\alpha,\beta} \in \mathbb{Q}$ *if* $\alpha \neq -\beta$;

(iv) $[e_\alpha, e_{-\alpha}] = \sum_{j=1}^l \chi_{j,\alpha} h_j$, $\chi_{j,\alpha} \in \mathbb{Q}$.

*Here the constants $\chi_\star$ only depend upon the appropriate configuration (that is, $w_\alpha$, $w_\beta$, $j$) from the root system $(E^L, \Phi^L)$.*

PROOF. The first two are immediate. Now we consider the various $[e_\alpha, e_\beta]$, which we verify by induction on $\min(|\operatorname{ht}(\alpha)|, |\operatorname{ht}(\beta)|)$. As $[e_\alpha, e_\beta] = -[e_\beta, e_\alpha]$ we may assume $|\operatorname{ht}(\alpha)| \le |\operatorname{ht}(\beta)|$).

First suppose $1 = |\operatorname{ht}(\alpha)|$; that is, $\alpha \in \epsilon\Delta$ ($\epsilon \in \pm$). If $\beta = \alpha$ then $[e_\alpha, e_\beta] = 0 e_\alpha$, and if $\beta = -\alpha \in -\epsilon\Delta$ then $[e_\alpha, e_\beta] = h_\alpha = \sum_{j=1}^l \chi_{j,\alpha} h_j$ (with all but one of the constants equal to 0). For $\beta \neq \pm\alpha$, we have $[e_\alpha, e_\beta] = e(w)$ for $w = a_\alpha w_\beta$; so $[e_\alpha, e_\beta] = \chi_w e_{\alpha+\beta} = \chi_{\alpha,\beta} e_{\alpha+\beta}$ by the lemma.

Now assume $1 < k = |\operatorname{ht}(\alpha)| \le |\operatorname{ht}(\beta)|$. Let $w_\alpha = w_k w_{k-1} \cdots w_1$, and set $w_k = a_\delta$ and $w = w_{k-1} \cdots w_1 (\neq \emptyset)$. Furthermore let $\gamma = \|w\|$. Note that $1 \le |\operatorname{ht}(\gamma)| < |\operatorname{ht}(\alpha)| \le |\operatorname{ht}(\beta)|$, and especially $\gamma \neq -\beta \neq \delta$.

We calculate (using induction and the lemma)

$$\begin{aligned}
[e_\alpha, e_\beta] &= [e(w_\alpha), e_\beta] \\
&= [[e_\delta, e(w)], e_\beta] \\
&= [e_\delta, [e(w), e_\beta]] - [e(w), [e_\delta, e_\beta]] \\
&= \chi_w([e_\delta, [e_\gamma, e_\beta]] - [e_\gamma, [e_\delta, e_\beta]]) \\
&= \chi_w(\chi_{\gamma,\beta}[e_\delta, e_{\gamma+\beta}] - \chi_{\delta,\beta}[e_\gamma, e_{\delta+\beta}])\,.
\end{aligned}$$

At this point, there are two cases to consider, depending upon whether or not

$$\alpha + \beta = \delta + \gamma + \beta = \gamma + \delta + \beta$$

is equal to 0.

If $\alpha + \beta \neq 0$ then by induction

$$\begin{aligned}
[e_\alpha, e_\beta] &= \chi_w(\chi_{\gamma,\beta}[e_\delta, e_{\gamma+\beta}] - \chi_{\delta,\beta}[e_\gamma, e_{\delta+\beta}]) \\
&= \chi_w(\chi_{\gamma,\beta}\chi_{\delta,\gamma+\beta}e_{\delta+\gamma+\beta} - \chi_{\delta,\beta}\chi_{\gamma,\delta+\beta}e_{\gamma+\delta+\beta}) \\
&= \chi_w(\chi_{\gamma,\beta}\chi_{\delta,\gamma+\beta} - \chi_{\delta,\beta}\chi_{\gamma,\delta+\beta})e_{\gamma+\delta+\beta} \\
&= \chi_{\alpha,\beta}\, e_{\alpha+\beta}
\end{aligned}$$

where the rational constant

$$\chi_{\alpha,\beta} = \chi_w(\chi_{\gamma,\beta}\chi_{\delta,\gamma+\beta} - \chi_{\delta,\beta}\chi_{\gamma,\delta+\beta})$$

depends only on $\alpha$ and $\beta$ (and the associated $w_\alpha = a_\delta w$ with $\gamma = \|w\|$).

If $\alpha + \beta = 0$ then $-\delta = \gamma + \beta$ and $-\gamma = \delta + \beta$. By induction again

$$\begin{aligned}
[e_\alpha, e_{-\alpha}] &= [e_\alpha, e_\beta] \\
&= \chi_w(\chi_{\gamma,\beta}[e_\delta, e_{\gamma+\beta}] - \chi_{\delta,\beta}[e_\gamma, e_{\delta+\beta}]) \\
&= \chi_w(\chi_{\gamma,-\alpha}[e_\delta, e_{-\delta}] - \chi_{\delta,-\alpha}[e_\gamma, e_{-\gamma}]) \\
&= \chi_w\left(\chi_{\gamma,-\alpha}\left(\sum_{j=1}^{l}\chi_{j,\delta}h_j\right) - \chi_{\delta,-\alpha}\left(\sum_{j=1}^{l}\chi_{j,\gamma}h_j\right)\right) \\
&= \chi_w\sum_{j=1}^{l}(\chi_{\gamma,-\alpha}\chi_{j,\delta} - \chi_{\delta,-\alpha}\chi_{j,\gamma})h_j \\
&= \sum_{j=1}^{l}\chi_{j,\alpha}h_j\,,
\end{aligned}$$

where the rational constants

$$\chi_{j,\alpha} = \chi_w(\chi_{\gamma,-\alpha}\chi_{j,\delta} - \chi_{\delta,-\alpha}\chi_{j,\gamma})$$

are entirely determined by $j$, $\alpha$, and the associated $w_\alpha = a_\delta w$ with $\gamma = \|w\|$. □

PROOF OF THEOREM (8.8).

The isomorphism between the root systems $(E, \Phi)$ and $(E', \Phi')$ gives rise (by Theorem (7.6)(b)) to a map $h_i \mapsto h'_i$ $(1 \le i \le l)$ and $e_\alpha \mapsto e'_{\alpha'}$ $(\alpha \in \Phi)$ that by the theorem extends to an isomorphism of the Lie algebras $L$ and $L'$. □

## 8.4 Existence

### 8.4.1 $\mathfrak{sl}_{l+1}(\mathbb{K})$

**(8.13).** EXAMPLE. *Let $L = \mathfrak{sl}_{l+1}(\mathbb{K})$, the Lie algebra of trace $0$ matrices in $\mathrm{Mat}_{l+1}(\mathbb{K})$ for $l \in \mathbb{Z}^+$.*

(a) *$L$ is simple of type $\mathfrak{a}_l(\mathbb{K})$ and dimension $l^2 + 2l$.*

(b) *All Cartan subalgebras have rank $l$ and are conjugate under $\mathrm{SL}_{l+1}(\mathbb{K}) \le \mathrm{Aut}(L)$ to $H$, the abelian and dimension $l$ subalgebra of all diagonal matrices with trace $0$.*

(c) *The $H$-root spaces are the various $\mathbb{K}e_{i,j}$ for $1 \le i \ne j \le l+1$ with corresponding root $\varepsilon_i - \varepsilon_j$ within $H^*$ and its rational form $\mathbb{Q}^l$ $(= \mathbb{Q}^{l+1} \cap \mathbf{1}^\perp)$.*

(d) *The simple roots of $\Delta$ are $\delta_i = \varepsilon_i - \varepsilon_{i+1} = \delta_i^\vee$ for $1 \le i \le l$, and so the Dynkin diagram is $A_l$.*

(e) *The Weyl reflection $r_{\varepsilon_i - \varepsilon_j}$ induces on $\mathbb{R}^l = \mathbb{R}^{l+1} \cap \mathbf{1}^\perp \le \mathbb{R}^{l+1}$ the permutation $(i, j)$ of the Weyl group $\mathrm{W}(A_l) \simeq \mathrm{Sym}(l+1)$.*

PROOF. (a) The dimension is $(l+1)^2 - 1$, as the only restriction is on the trace. Indeed, at least as vector space $L$ is $H \oplus \bigoplus_{i \ne j} \mathbb{K}e_{i,j}$. The rest of this part then follows from (d) and Theorem (8.2).

(b) $L$ is irreducible on the natural module $V = \mathbb{K}^{l+1}$ (for instance, because the range of $e_{i,j}$ is the basis subspace $\mathbb{K}e_i$). Therefore by Theorem (6.8) the module $V$ is a weight module, which is to say that every Cartan subalgebra $C$ of $L$ can be diagonalized. Thus there is a $g \in \mathrm{GL}_{l+1}(\mathbb{K})$ and indeed in $\mathrm{SL}_{l+1}(\mathbb{K})$ (as $l \ge 1$) with the Cartan subalgebra $C^g$ in $H$. But a self-normalizing subalgebra of $L$ within abelian $H$ must be $H$ itself, so $H$ is a Cartan subalgebra and $C^g = H$.

(c) If $h = \mathrm{diag}(h_1, \ldots, h_{l+1}) \in H$, then $[h, e_{i,j}] = (h_i - h_j)e_{i,j}$. Therefore $\mathbb{K}e_{i,j}$ is a root space $L_\alpha$. When we let the canonical basis of $\mathbb{Q}^{l+1} \le V^* = \mathbb{K}^{l+1}$ be $\varepsilon_i, \ldots, \varepsilon_{l+1}$, we find $\alpha(h) = (\varepsilon_i - \varepsilon_j)(h)$; that is, $\alpha$ becomes the root $\varepsilon_i - \varepsilon_j$ in the Euclidean $l$-space $\mathbb{R}^{l+1} \cap \mathbf{1}^\perp$.

(d) The lexicographic order induced by $\varepsilon_1 > \varepsilon_2 > \cdots > \varepsilon_{l+1}$ yields the simple base $\Delta$ described. Note that all roots $\alpha$ have $\alpha^\vee = \alpha$. If $i < j$ then

$(\delta_i, \delta_j^\vee)$ is 0 unless $j = i+1$ where it is $-1$. Thus the Dynkin diagram of $\Delta$ and $L$ is $A_l$.

(e) For $1 \le k \le l+1$

$$\begin{aligned} r_{\varepsilon_i - \varepsilon_j}(\varepsilon_k) &= \varepsilon_k - (\varepsilon_k, (\varepsilon_i - \varepsilon_j)^\vee)(\varepsilon_i - \varepsilon_j) \\ &= \varepsilon_k - (\varepsilon_k, \varepsilon_i - \varepsilon_j)(\varepsilon_i - \varepsilon_j)\,. \end{aligned}$$

Thus $r_{\varepsilon_i - \varepsilon_j}(\varepsilon_k) = \varepsilon_k$ if $k \notin \{i, j\}$ while $r_{\varepsilon_i - \varepsilon_j}(\varepsilon_i) = \varepsilon_j$ and $r_{\varepsilon_i - \varepsilon_j}(\varepsilon_j) = \varepsilon_i$. That is, $r_{\varepsilon_i - \varepsilon_j}$ induces the 2-cycle $(\varepsilon_i, \varepsilon_j)$ on the set $\{\varepsilon_1, \ldots, \varepsilon_{l+1}\}$. These generate the symmetric group. □

### 8.4.2 Spaces with forms

**(8.14).** THEOREM. *Let $b$ be the nondegenerate $\eta$-symmetric form on finite dimensional space $V$ over algebraically closed $\mathbb{K}$ of characteristic 0. The Lie algebra $L$ consists of those $x \in \mathrm{End}_{\mathbb{K}}(V) \simeq \mathrm{Mat}_n(\mathbb{K})$ with*

$$b(xv, w) = -b(v, xw)$$

*for all $v, w \in V$. Assume additionally that $L$ is semisimple and irreducible on $V$.*

*Let $C$ be a Cartan subalgebra of $L$. Then we are in one of three cases:*

(a) $L = \mathfrak{so}_{2l}(\mathbb{K})$ *with* $(n, \eta) = (2l, +1)$;

(b) $L = \mathfrak{sp}_{2l}(\mathbb{K})$ *with* $(n, \eta) = (2l, -1)$;

(c) $\mathfrak{so}_{2l+1}(\mathbb{K})$ *with* $(n, \eta) = (2l+1, +1)$.

*In all cases, there is a basis consisting of $C$-weight vectors possessing a split Gram matrix*

$$\bigoplus_{k=1}^{l} \begin{pmatrix} 0 & 1 \\ \eta & 0 \end{pmatrix}$$

*in the two even dimensional cases and*

$$(1) \oplus \bigoplus_{k=1}^{l} \begin{pmatrix} 0 & 1 \\ \eta & 0 \end{pmatrix}$$

*in the odd dimensional case.*

**(8.15).** THEOREM. *Let $L$ be one of the Lie algebras $\mathfrak{so}_{2l}(\mathbb{K})$ with $(n, \eta) = (2l, +1)$ or $\mathfrak{sp}_{2l}(\mathbb{K})$ with $(n, \eta) = (2l, -1)$ or $\mathfrak{so}_{2l+1}(\mathbb{K})$ with $(n, \eta) = (2l+1, +1)$. Set $V = \mathbb{K}^n$ to be the natural module for $L$. Let $C$ be a Cartan subalgebra for $L$. Then, in the action of $L$ on $V$, there is a basis of $C$-weight vectors with Gram matrix in split form as the $2l \times 2l$ matrix with $l$ blocks $\begin{pmatrix} 0 & 1 \\ \eta & 0 \end{pmatrix}$ down the diagonal when $n = 2l$ is even, and this same matrix with an additional single 1 on the diagonal when $n = 2l+1$ is odd.*

PROOF. In all cases $L$ is irreducible on $V$, so by Theorem (6.8) the module $V$ is a weight module for all choices of Cartan subalgebra $C$.

Let $b$ be the nondegenerate $\eta$-symmetric form on $V$ for $\eta = \pm 1$ with $L$ equal to those $x \in \mathrm{End}_{\mathbb{K}}(V) \simeq \mathrm{Mat}_n(\mathbb{K})$ with

$$b(xv, w) = -b(v, xw)$$

for all $v, w \in V$. Let $v \in V_{C,\lambda}$, and $w \in V_{C,\mu}$. Then for all $h \in C$

$$\lambda(h)b(v, w) = b(hv, w) = -b(v, hw) = -\mu(h)b(v, w)\,.$$

That is, $(\lambda+\mu)(h)\, b(v, w)$ is identically 0 for $h \in C$. In particular, if $\lambda \neq -\mu$ then $b(v, w) = 0$ and $V_{C,\lambda}$ and $V_{C,\mu}$ are perpendicular. The space $V$ is nondegenerate, so for all weights $\lambda$ of $C$ on $V$ we must have $(V_{C,\lambda}, V_{C,-\lambda}) \neq 0$.

Let $\lambda \neq 0$, and choose $0 \neq v \in V_{C,\lambda}$. As $v \notin \mathrm{Rad}(V, b)$ there is a $w \in V_{C,-\lambda}$ with $b(v, w) \neq 0$. We have $b(v, v) = 0 = b(w, w)$ (as $\lambda \neq -\lambda$). Therefore we may rescale one of the pair $\{v, w\}$ so that the Gram matrix of the nondegenerate 2-space $W = \mathbb{K}v \oplus \mathbb{K}w$ has the stated form $\begin{pmatrix} 0 & 1 \\ \eta & 0 \end{pmatrix}$. As $C$ leaves $W = W_1$ invariant, it also acts on $V_1 = W^{\perp}$. Continuing in this fashion we leave $V$ written as a perpendicular direct sum $W_1 \oplus W_2 \oplus \cdots W_m \oplus V_0$ where the basis $\{v_i, w_i\}$ of $W_i$ consists of $\lambda_i$- and $-\lambda_i$-weight vectors for $\lambda_i \neq 0$ and $V_0$ is the 0-weight space, nondegenerate if nonzero. If $V_0$ has dimension 0, then $m = l$, $n = 2l$, and we are done. If $V_0 = \mathbb{K}v$ has dimension 1, then $m = l$, and $n = 2l + 1$. As $b$ is nondegenerate and $\mathbb{K}$ is algebraically closed, we may rescale to $b(v, v) = 1$, and again we are done.

If $\dim_{\mathbb{K}}(V_0) \geq 2$, then for any nondegenerate 2-space $W_0$ of $V_0$, by Lemma A-(1.7) (of Appendix A) there is again a basis $\{v_0, w_0\}$ of weight vectors in $W_0$ with the same Gram matrix $\begin{pmatrix} 0 & 1 \\ \eta & 0 \end{pmatrix}$. We continue in this fashion within $W_0^{\perp}$ until we exhaust $V_0$ ($n = 2l$) or reach a subspace of dimension 1 ($n = 2l + 1$), and we are done. □

**(8.16).** EXAMPLES. *For $\eta \in \{\pm\} = \{\pm 1\}$, let the $\mathbb{K}$-space $V = V_\eta = \mathbb{K}^{2l}$ have basis $\{\, e_i, e_{-i} \mid 1 \leq i \leq l \,\}$ and be is equipped with the split $\eta$-symmetric form $b = b_\eta$ given by*

$$b(e_i, e_{-i}) = 1\,, b(e_{-i}, e_i) = \eta\,,\ \textit{otherwise}\ b(e_a, e_b) = 0\,.$$

*The Lie algebra $L = L_\eta$ is then composed of all $x \in \mathrm{End}_{\mathbb{K}}(V) \simeq \mathrm{Mat}_{2l}(\mathbb{K})$ with*

$$b_\eta(xv, w) = -b_\eta(v, xw)$$

*for all $v, w \in V$. Thus $L_+$ is the orthogonal Lie algebra $\mathfrak{so}_{2l}(\mathbb{K})$, and $L_-$ is the symplectic Lie algebra $\mathfrak{sp}_{2l}(\mathbb{K})$*

(i) $\mathfrak{so}_{2l}(\mathbb{K})$*: orthogonal case $\eta = +1$.*

(a) *The algebra $L_+ = \mathfrak{so}_{2l}(\mathbb{K})$ is simple of type $\mathfrak{d}_l(\mathbb{K})$ and dimension $2l^2 - l$.*

(b) *All Cartan subalgebras have rank $l$ and are conjugate under* $\mathrm{Aut}(\mathfrak{so}_{2l}(\mathbb{K}))$ *to $H$, the abelian and dimension $l$ subalgebra of all diagonal matrices with basis $e_{i,i} - e_{-i,-i}$ for $1 \le i \le l$.*

(c) *For $h = \sum_{k=1}^{l} h_k(e_{k,k} - e_{-k,-k}) \in H$ we let $\varepsilon_k \colon h \mapsto h_k$ give the chosen basis for $H^*$ and $\mathbb{Q}^l$. The $H$-root spaces are spanned by the following weight vectors and have the corresponding roots:*

| *Vector* | *Root* |
|---|---|
| $e_{i,j} - e_{-j,-i}$ | $\varepsilon_i - \varepsilon_j$ |
| $e_{-i,-j} - e_{j,i}$ | $-(\varepsilon_i - \varepsilon_j)$ |
| $e_{i,-j} - e_{j,-i}$ | $\varepsilon_i + \varepsilon_j$ |
| $e_{-i,j} - e_{-j,i}$ | $-(\varepsilon_i + \varepsilon_j)$ |

(d) *The simple roots of $\Delta$ are $\delta_i = \varepsilon_i - \varepsilon_{i+1} = \delta_i^\vee$ for $1 \le i \le l-1$ and $\delta_l = \varepsilon_{l-1} + \varepsilon_l = \delta_l^\vee$, and so the Dynkin diagram is $D_l$.*

(e) *The Weyl reflection $r_{\varepsilon_i - \varepsilon_{i+1}}$ induces on $\mathbb{R}^l$ the permutation $(i, i+1)$ while $r_{\varepsilon_{l-1}+\varepsilon_l}$ fixes $\varepsilon_k$ for $k < l-1$ but has $r_{\varepsilon_{l-1}+\varepsilon_l}(\varepsilon_{l-1}) = -\varepsilon_l$ and $r_{\varepsilon_{l-1}+\varepsilon_l}(\varepsilon_{l-1}) = -\varepsilon_l$. So the Weyl group* $\mathrm{W}(D_l)$ *is* $2^{l-1}\colon \mathrm{Sym}(l)$.

(ii) $\mathfrak{sp}_{2l}(\mathbb{K})$*: symplectic case $\eta = -1$.*

(a) *The algebra $L_- = \mathfrak{sp}_{2l}(\mathbb{K})$ is simple of type $\mathfrak{c}_l(\mathbb{K})$ and dimension $2l^2 + l$.*

(b) *All Cartan subalgebras have rank $l$ and are conjugate under* $\mathrm{Aut}(\mathfrak{sp}_{2l}(\mathbb{K}))$ *to $H$, the abelian and dimension $l$ subalgebra of all diagonal matrices with basis $e_{i,i} - e_{-i,-i}$ for $1 \le i \le l$.*

(c) *For $h = \sum_{k=1}^{l} h_k(e_{k,k} - e_{-k,-k}) \in H$ we let $\varepsilon_k \colon h \mapsto h_k$ give the chosen basis for $H^*$ and $\mathbb{Q}^l$. The $H$-root spaces are spanned by the following weight vectors and have the corresponding roots:*

| *Vector* | *Root* |
|---|---|
| $e_{i,j} - e_{-j,-i}$ | $\varepsilon_i - \varepsilon_j$ |
| $e_{-i,-j} - e_{j,i}$ | $-(\varepsilon_i - \varepsilon_j)$ |
| $e_{i,-j} + e_{j,-i}$ | $\varepsilon_i + \varepsilon_j$ |
| $e_{-i,j} + e_{-j,i}$ | $-(\varepsilon_i + \varepsilon_j)$ |
| $e_{i,-i}$ | $2\varepsilon_i$ |
| $e_{-i,i}$ | $-2\varepsilon_i$ |

(d) *The simple roots of $\Delta$ are $\delta_i = \varepsilon_i - \varepsilon_{i+1} = \delta_i^\vee$ for $1 \le i \le l-1$ and $\delta_l = 2\varepsilon_l$ (with $\delta_l^\vee = \varepsilon_l$), and so the Dynkin diagram is $C_l$.*

(e) *The Weyl reflection $r_{\varepsilon_i - \varepsilon_{i+1}}$ induces on $\mathbb{R}^l$ the permutation $(i, i+1)$ while $r_{2\varepsilon_l}$ is the diagonal reflection taking $\varepsilon_l$ to $-\varepsilon_l$. So the Weyl group* $\mathrm{W}(C_l)$ *is* $2^l\colon \mathrm{Sym}(l)$.

Proof. (a) It is helpful to consider the $2l \times 2l$ matrices of $\mathrm{Mat}_{2l}(\mathbb{K})$ as $l \times l$ matrices whose entries are the various $2 \times 2$ submatrices $\begin{pmatrix} a_{i,j} & b_{i,-j} \\ b_{-i,j} & a_{-i,-j} \end{pmatrix}$.

The requirements for such a matrix to be in $L_\eta$ are then

$$\begin{pmatrix} a_{i,j} & b_{i,-j} \\ b_{-i,j} & a_{-i,-j} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ \eta & 0 \end{pmatrix} = - \begin{pmatrix} 0 & 1 \\ \eta & 0 \end{pmatrix} \begin{pmatrix} a_{j,i} & b_{-j,i} \\ b_{j,-i} & a_{-j,-i} \end{pmatrix},$$

which is to say then

$$\begin{pmatrix} \eta b_{i,-j} & a_{i,j} \\ \eta a_{-i,-j} & b_{-i,j} \end{pmatrix} = \begin{pmatrix} -b_{j,-i} & -a_{-j,-i} \\ -\eta a_{j,i} & -\eta b_{-j,i} \end{pmatrix}.$$

Thus

$$\begin{aligned} \eta b_{i,-j} &= -b_{j,-i} \\ a_{i,j} &= -a_{-j,-i} \\ \eta a_{-i,-j} &= -\eta a_{j,i} \\ b_{-i,j} &= -\eta b_{-j,i}\,. \end{aligned}$$

We rewrite and view these as four separate equations subject to the restriction $1 \leq i \leq j \leq l$:

$$\begin{aligned} a_{i,j} &= -a_{-j,-i} \\ a_{-i,-j} &= -a_{j,i} \\ b_{i,-j} &= -\eta b_{j,-i} \\ b_{-i,j} &= -\eta b_{-j,i}\,. \end{aligned}$$

Thus the matrices of $L_\eta$ can have anything above the diagonal $2 \times 2$ blocks (where $i < j$), these entries determining those below the diagonal blocks. This contributes $4(l(l-1)/2) = 2l^2 - 2l$ to the dimension, the relevant basis elements being, for $1 \leq i < j \leq l$,

$$e_{i,j} - e_{-j,-i}\,, \quad e_{-i,-j} - e_{j,i}\,, \quad e_{i,-j} - \eta e_{j,-i}\,, \quad e_{-i,j} - \eta e_{-j,i}\,.$$

In the diagonal blocks $i = j$ we must have

$$\begin{aligned} a_{i,i} &= -a_{-i,-i} \\ a_{-i,-i} &= -a_{i,i} \\ b_{i,-i} &= -\eta b_{i,-i} \\ b_{-i,i} &= -\eta b_{-i,i}\,. \end{aligned}$$

The first two equations are equivalent and contribute $l$ to the overall dimension, the corresponding basis elements being $e_{i,i} - e_{-i,-i}$ for $1 \leq i \leq l$. In the second two equations, if $\eta = +1$ there are no nonzero solutions (as $\mathbb{K}$ has characteristic 0), while if $\eta = -1$ the equations are trivially valid and so contribute a full $2l$ to the dimension, the additional basis elements being $e_{i,-i}$ and $e_{-i,i}$ for $1 \leq i \leq l$.

Therefore

$$\dim_{\mathbb{K}}(L_+) = (2l^2 - 2l) + l = 2l^2 - l$$

and

$$\dim_{\mathbb{K}}(L_-) = (2l^2 - 2l) + l + 2l = 2l^2 + l\,.$$

The rest of (a) then will follow from (d) and Theorem (8.2).

(b) Theorem (8.15) proves that any Cartan subalgebra is conjugate under $\mathrm{Aut}(L_\eta)$ into the diagonal subalgebra of the algebra. But this diagonal subalgebra is abelian, so the self-normalizing Cartan subalgebra within it must be the whole diagonal subalgebra. As we saw under (a) it has basis $e_{i,i} - e_{-i,i}$ for $1 \le i \le l$.

(c) The basis we described under (a) turns out (unsurprisingly) to be a basis of weight vectors. For instance:

$$\begin{aligned}
\left[\sum_{k=1}^{l} h_k(e_{k,k} - e_{-k,-k})\,,\, e_{i,-j} - \eta e_{j,-i}\right] &= \sum_{k=1}^{l} h_k[e_{k,k} - e_{-k,-k}\,,\, e_{i,-j} - \eta e_{j,-i}] \\
&= \sum_{k=1}^{l} h_k((e_{k,k} - e_{-k,-k})(e_{i,-j} - \eta e_{j,-i}) - (e_{i,-j} - \eta e_{j,-i})(e_{k,k} - e_{-k,-k})) \\
&= \sum_{k=1}^{l} h_k((e_{k,k}e_{i,-j} - e_{k,k}\eta e_{j,-i}) - (-e_{i,-j}e_{-k,-k} + \eta e_{j,-i}e_{-k,-k})) \\
&= (h_i e_{i,-j} - h_j \eta e_{j,-i}) - (-h_j e_{i,-j} + \eta h_i e_{j,-i}) \\
&= (h_i + h_j)e_{i,-j} - (h_i + h_j)\eta e_{j,-i} \\
&= (h_i + h_j)(e_{i,-j} - \eta e_{j,-i})\,.
\end{aligned}$$

Therefore $e_{i,-j} - \eta e_{j,-i}$ is a weight vector for the root $\varepsilon_i + \varepsilon_j$.

The other entries in the tables follow by similar calculations. For instance:

$$\begin{aligned}
\left[\sum_{k=1}^{l} h_k(e_{k,k} - e_{-k,-k})\,,\, e_{-i,i}\right] &= \sum_{k=1}^{l} h_k[e_{k,k} - e_{-k,-k}\,,\, e_{-i,i}] \\
&= \sum_{k=1}^{l} h_k(e_{k,k}e_{-i,i} - e_{-k,-k}e_{-i,i} - e_{-i,i}e_{k,k} + e_{-i,i}e_{-k,-k}) \\
&= -h_i e_{-i,-i}e_{-i,i} - h_i e_{-i,i}e_{i,i} \\
&= -2h_i e_{-i,i}\,.
\end{aligned}$$

Thus in the symplectic ($\eta = -1$) case $e_{-i,i}$ is a weight vector for the root $-2\varepsilon_i$.

(d) Lexicographic order is induced by $\varepsilon_1 > \varepsilon_2 > \cdots > \varepsilon_l$. The simple roots are then evident. Note that in the symplectic case $\varepsilon_{l-1} + \varepsilon_l$ remains a positive root, but it is no longer simple as $\varepsilon_{l-1} + \varepsilon_l = (\varepsilon_{l-1} - \varepsilon_l) + 2\varepsilon_l$.

(e) The reflections in $\varepsilon_i - \varepsilon_j$ were calculated under Example (8.13)(e), and the reflection in $2\varepsilon_l$ is clear. All that needs checking is

$$r_{\varepsilon_{l-1}+\varepsilon_l}(\varepsilon_{l-1}) = \varepsilon_{l-1} - (\varepsilon_{l-1}, \varepsilon_{l-1} + \varepsilon_l)(\varepsilon_{l-1} + \varepsilon_l) = \varepsilon_{l-1} - (\varepsilon_{l-1} + \varepsilon_l) = -\varepsilon_l\,. \quad \square$$

**(8.17).** EXAMPLE. *Let the $\mathbb{K}$-space $V = V_\eta = \mathbb{K}^{2l+1}$ (for $\eta = +1$) have the basis $\{ e_0, e_i, e_{-i} \mid 1 \le i \le l \}$ and be equipped with the split orthogonal form $b$ given by*

$$b(e_0, e_0) = 1\,,\ b(e_i, e_{-i}) = b(e_{-i}, e_i) = 1\,,\ \textit{otherwise } b(e_a, e_b) = 0\,.$$

*The Lie algebra $L$ is the orthogonal Lie algebra $\mathfrak{so}_{2l+1}(\mathbb{K})$, composed of all $x \in \mathrm{End}_{\mathbb{K}}(V) \simeq \mathrm{Mat}_{2l+1}(\mathbb{K})$ with*

$$b(xv, w) = -b(v, xw)$$

*for all $v, w \in V$.*

(a) *The algebra $L = \mathfrak{so}_{2l+1}(\mathbb{K})$ is simple of type $\mathfrak{b}_l(\mathbb{K})$ and dimension $2l^2 + l$.*

(b) *All Cartan subalgebras have rank $l$ and are conjugate under $\mathrm{Aut}(\mathfrak{so}_{2l+1}(\mathbb{K}))$ to $H$, the abelian and dimension $l$ subalgebra of all diagonal matrices with basis $e_{i,i} - e_{-i,-i}$ for $1 \le i \le l$.*

(c) *For $h = \sum_{k=1}^{l} h_k(e_{k,k} - e_{-k,-k}) \in H$ we let $\varepsilon_k \colon h \mapsto h_k$ give the chosen basis for $H^*$. The $H$-root spaces are spanned by the following weight vectors and have the corresponding roots:*

| *Vector* | *Root* |
|---|---|
| $e_{i,j} - e_{-j,-i}$ | $\varepsilon_i - \varepsilon_j$ |
| $e_{-i,-j} - e_{j,i}$ | $-(\varepsilon_i - \varepsilon_j)$ |
| $e_{i,-j} + e_{j,-i}$ | $\varepsilon_i + \varepsilon_j$ |
| $e_{-i,j} + e_{-j,i}$ | $-(\varepsilon_i + \varepsilon_j)$ |
| $e_{i,0} - e_{0,-i}$ | $\varepsilon_i$ |
| $e_{-i,0} - e_{0,i}$ | $-\varepsilon_i$ |

(d) *The simple roots of $\Delta$ are $\delta_i = \varepsilon_i - \varepsilon_{i+1} = \delta_i^\vee$ for $1 \le i \le l-1$ and $\delta_l = \varepsilon_l$ (with $\delta_l^\vee = 2\varepsilon_l$) and so the Dynkin diagram is $B_l$.*

(e) *The Weyl reflection $r_{\varepsilon_i - \varepsilon_{i+1}}$ induces on $\mathbb{R}^l$ the permutation $(i, i+1)$ while $r_{\varepsilon_l}$ is the diagonal reflection taking $\varepsilon_l$ to $-\varepsilon_l$. So the Weyl group $\mathrm{W}(B_l)$ is $2^l \colon \mathrm{Sym}(l)$.*

PROOF. As the Gram matrices indicate, the algebra $\mathfrak{so}_{2l+1}(\mathbb{K})$ can be thought of as an extension of $\mathfrak{so}_{2l}(\mathbb{K})$. As such, most of the arguments from the previous example (case $\eta = +1$) are valid here. Furthermore the ultimate similarity of the root systems means that the symplectic case $\eta = -1$ of the previous example is also relevant here. (Perhaps all three algebras should be handled at once.)

(a) We think of the Gram matrix $G_{2l+1}$ as the Gram matrix $G_{2l}$ for $\mathfrak{so}_{2l}(\mathbb{K})$ with a new row and column indexed 0, corresponding to the basis element $e_0$ of $V = \mathbb{K}^{2l+1}$, the diagonal entry being $b(e_0, e_0) = 1$ and all other entries in the new row and column being 0. Then $MG_{2l+1} = -G_{2l+1}M^\top$ if and only if

(1) $M_{0,0} = 0$;

(2) the rest of row $M_{0,*}$ contains any vector $v \in \mathbb{K}^{2l}$;

(3) the rest of column $M_{*,0}$ contains $-vG_{2l}$;

(4) deleting row 0 and column 0 from $M$ leaves a matrix of $\mathfrak{so}_{2l}(\mathbb{K})$, as described in Example (8.16)(i).

Thus a basis for $L$ is that for $\mathfrak{so}_{2l}(\mathbb{K})$ from Example (8.16)(i), supplemented with the $2l$ elements $e_{i,0} - e_{0,-i}$ and $e_{-i,0} - e_{0,i}$. The dimension is then $2l^2 - l + 2l = 2l^2 + l$. The rest of (a) will then follow from (d) and Theorem (8.2) as before.

(b) Again by Theorem (8.15) a Cartan subalgebra is conjugate under $\mathrm{Aut}(L)$ into and then to the abelian diagonal subalgebra of the algebra, which remains the rank $l$ space with basis $e_{i,i} - e_{-i,-i}$ for $1 \le i \le l$.

(c) The weight vectors and roots for the subalgebra $\mathfrak{so}_{2l}(\mathbb{K})$ are unchanged. We must additionally calculate:

$$\begin{aligned}
\left[\sum_{k=1}^{l} h_k(e_{k,k} - e_{-k,-k})\,,\, e_{i,0} - e_{0,-i}\right] &= \sum_{k=1}^{l} h_k[e_{k,k} - e_{-k,-k}\,,\, e_{i,0} - e_{0,-i}] \\
&= \sum_{k=1}^{l} h_k(e_{k,k}e_{i,0} - e_{0,-i}e_{-k,-k}) \\
&= h_i e_{i,0} - h_i e_{0,-i} \\
&= h_i(e_{i,0} - e_{0,-i})\,.
\end{aligned}$$

Parts (d) and (e) follow, as in Example (8.16)(ii). □

### 8.4.3 $\mathfrak{e}_8$, $\mathfrak{e}_7$, $\mathfrak{e}_6$

**(8.18).** PROPOSITION.

(a) *If* $\mathfrak{e}_8(\mathbb{K})$ *exists, then it has dimension* 248 *and a proper subalgebra* $\mathfrak{e}_7(\mathbb{K})$.

(b) *If* $\mathfrak{e}_7(\mathbb{K})$ *exists, then it has dimension* 133 *and a proper subalgebra* $\mathfrak{e}_6(\mathbb{K})$.

(c) *If* $\mathfrak{e}_6(\mathbb{K})$ *exists, then it has dimension* 78.

PROOF. If $\Gamma$ is a subsystem of the root system $\Phi$ of $L = H \oplus \bigoplus_{\alpha\in\Phi} L_\alpha$, then $M = H \oplus \bigoplus_{\alpha\in\Gamma} L_\alpha$ is a subalgebra by Theorem (5.12). Indeed by Proposition (8.5) the subspace $\bigoplus_{\alpha\in\Gamma} L_\alpha$ generates a semisimple subalgebra $M_0 = (H \cap M_0) \oplus \bigoplus_{\alpha\in\Gamma} L_\alpha$ with root system $\Gamma$. Thus the containments are clear from our construction of the root systems $E_8 \supset E_7 \supset E_6$ in Section 7.3.

We know by Corollary (6.5) that the rank $l$ semisimple Lie algebra with root system $(E^L, \Phi^L)$ has dimension $l + |\Phi|$. From the same section, we then find:

(a) $\dim_{\mathbb{K}}(\mathfrak{e}_8(\mathbb{K})) = 8 + 240 = 248$;

(b) $\dim_{\mathbb{K}}(\mathfrak{e}_7(\mathbb{K})) = 7 + 126 = 133$;

(c) $\dim_{\mathbb{K}}(\mathfrak{e}_6(\mathbb{K})) = 6 + 72 = 78$; □

We leave unproven:

**(8.19).** THEOREM. *The Lie algebra $\mathfrak{e}_8(\mathbb{K})$ exists.*

As we have shown that the root system $E_8$ exists, this theorem is in fact a consequence of Serre's Theorem (10.15).

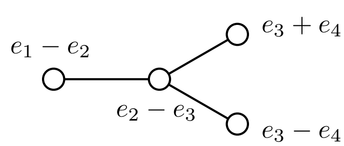
### 8.4.4 $\mathfrak{g}_2$ and $\mathfrak{f}_4$

In the previous chapter we constructed certain root systems by examining the fixed points of automorphisms of related Dynkin diagrams. Serre's Theorem (10.15) or the uniqueness Theorem (8.7) tell us that this extends to the construction of appropriate Lie algebras:

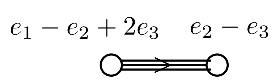
**(8.20).** THEOREM. *Any automorphism of the Dynkin diagram of semisimple $L$ extends to a nontrivial automorphism of $L$ acting in the same manner on the Cartan subalgebra and associated root spaces.* □

**(8.21).** THEOREM. *The Lie algebra $\mathfrak{d}_4(\mathbb{K})$ has a graph automorphism of order 3. Its fixed points contain a Lie algebra of type $\mathfrak{g}_2(\mathbb{K})$. Especially $\mathfrak{g}_2(\mathbb{K})$ of dimension 14 exists.*

PROOF. In Section 7.3.6, beginning with the root system of type $D_4$ with Dynkin diagram and simple basis




we used the triality graph automorphism to construct in $\mathbb{R}^2 = \mathbb{R}^3 \cap (e_1 - e_2 - e_3)^\perp$ a $G_2$ root system of fixed points:




By Theorem (8.20) this leads to a triality automorphism of the Lie algebra $\mathfrak{d}_4$ of dimension $28 = 4+24$. The fixed points correspond to the fixed subalgebra $H_0$ of the Cartan subalgebra (as dictated by the fixed root subsystem), the root spaces for the six fixed roots, and the fixed spaces on the diagonal of the six nontrivial orbits on root spaces. The fixed space thus has dimension $2 + 6 + 6 = 14$. In view of its root system $G_2$ and inherited Cartan decomposition, this algebra must of type $\mathfrak{g}_2$. □

In Section 7.3.5 we saw that the root system of type $E_6$ has a graph automorphism that fixes a root system of type $F_4$, which contains 48 roots. Theorem (8.20) and arguments similar to those above for $\mathfrak{g}_2$ then yield

**(8.22).** THEOREM. *If $\mathfrak{e}_6(\mathbb{K})$ exists then it has a proper subalgebra $\mathfrak{f}_4(\mathbb{K})$ of dimension* $52 = 4 + 48$. $\square$

## 8.5 Semisimple algebras V: Classification

We now can prove almost all of:

**(8.23).** THEOREM. (CLASSIFICATION OF SEMISIMPLE LIE ALGEBRAS) *Let $L$ be a finite dimensional semisimple Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic $0$. Then $L$ can be expressed uniquely as a direct sum of simple subalgebras. Each simple subalgebra is isomorphic to exactly one of the following, where in each case the rank is $l$:*

(a) $\mathfrak{a}_l(\mathbb{K}) \simeq \mathfrak{sl}_{l+1}(\mathbb{K})$, *for rank $l \geq 1$, of dimension $l^2 + 2l$;*

(b) $\mathfrak{b}_l(\mathbb{K}) \simeq \mathfrak{so}_{2l+1}(\mathbb{K})$, *for rank $l \geq 3$, of dimension $2l^2 + l$;*

(c) $\mathfrak{c}_l(\mathbb{K}) \simeq \mathfrak{sp}_{2l}(\mathbb{K})$, *for rank $l \geq 2$, of dimension $2l^2 + l$;*

(d) $\mathfrak{d}_l(\mathbb{K}) \simeq \mathfrak{so}_{2l}(\mathbb{K})$, *for rank $l \geq 4$, of dimension $2l^2 - l$;*

(e) $\mathfrak{e}_6(\mathbb{K})$ *of rank $l = 6$ and dimension $78$;*

(f) $\mathfrak{e}_7(\mathbb{K})$ *of rank $l = 7$ and dimension $133$;*

(g) $\mathfrak{e}_8(\mathbb{K})$ *of rank $l = 8$ and dimension $248$;*

(h) $\mathfrak{f}_4(\mathbb{K})$ *of rank $l = 4$ and dimension $52$;*

(i) $\mathfrak{g}_2(\mathbb{K})$ *of rank $l = 2$ and dimension $14$.*

*None of these simple algebras is isomorphic to one from another case or to any other algebra from the same case. All exist.*

PROOF. A simple algebra must be of one of these eight types by Theorems (7.12) and (8.2). (The rank restrictions in the first four classic cases are made to avoid diagram duplication such as $B_2 = C_2$ and $A_3 = D_3$.) In each case there is, up to isomorphism, at most one example by Theorem (8.8).

In the four classical cases, each exists by Examples (8.13), (8.16), and (8.17) with the given rank and dimension. These results also show that no algebra from any one classical case is isomorphic to any other from its own case or from any other case. Indeed all Cartan subalgebras have rank $l$ and are conjugate under the corresponding automorphism groups, so the rank and dimension reveal the only possible collisions to be $\mathfrak{b}_l(\mathbb{K})$ and $\mathfrak{c}_l(\mathbb{K})$ (for $l \geq 3$) and also possibly $\mathfrak{e}_6$ (when $l = 6$). But the root systems $B_l$ and $C_l$ are distinct for $l \geq 3$, and both possess roots of two different lengths, while $E_6$ has only one root length.

The rank 2 algebra $\mathfrak{g}_2(\mathbb{K})$ exists and has dimension 14 by Theorem (8.21).

Leaving aside existence for the moment, by Proposition (8.18) the exceptional algebras $\mathfrak{e}_8(\mathbb{K})$, $\mathfrak{e}_7(\mathbb{K})$, $\mathfrak{e}_6(\mathbb{K})$, and $\mathfrak{f}_4(\mathbb{K})$ all (if they exist) have different dimensions and so cannot in any case be isomorphic to each other. Furthermore, none is isomorphic to a classical algebra (as mentioned above) or to $\mathfrak{g}_2(\mathbb{K})$ of dimension 14.

Existence for the algebra $\mathfrak{e}_8(\mathbb{K})$ is stated in Theorem (8.19) (our main unproven result). Given this, existence for $\mathfrak{e}_7(\mathbb{K})$, $\mathfrak{e}_6(\mathbb{K})$, and $\mathfrak{f}_4(\mathbb{K})$ come from Proposition (8.18) and Theorem (8.22). □

As mentioned in the proof, the only part of the theorem that we have left unproven is the existence of $\mathfrak{e}_8(\mathbb{K})$ as in Theorem (8.19). (Although the existence of $\mathfrak{e}_8$ is used to confirm existence of the other exceptional algebras.) For the following corollary the theorem is not necessary as Proposition (8.18) suffices.

**(8.24).** Corollary. *Let $L$ be a finite dimensional semisimple Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic 0. Then all Cartan subalgebras of $L$ are conjugate under the action of $\mathrm{Aut}(L)$.*

Proof. By Theorem (8.8) or Serre's Theorem (10.15), if two Cartan subalgebras of semisimple $L$ give rise to isomorphic root systems, then the subalgebras are conjugate under $\mathrm{Aut}(L)$. Therefore if $L$ contains two nonconjugate Cartan subalgebras, this must arise from one of the simple algebras in the theorem being isomorphic to another simple algebra with a different root system and hence in a different case. But, as the theorem states, this does not happen. □

Proof of Theorem (8.7).

Directly after the statement of the theorem we observed that its parts (2), (3), and (4) are equivalent. Serre's Theorem (10.15) or Theorem (8.8) then prove that (2) implies (1). Now that we know that all Cartan subalgebras are conjugate via an automorphism, we cannot have two isomorphic algebras with nonisomorphic root systems; that is, (1) implies (2). □

## 8.6 Problems

**(8.25).** Problem. *Prove: $\Phi^\vee$ is a root system with simple basis $\Delta^\vee$ and $\mathrm{W}(\Phi^\vee) = \mathrm{W}(\Phi)$.*

**(8.26).** Problem. *We may consider $\alpha$-strings in the more general context of abstract root systems $(E, \Phi)$. Let $\alpha$ and $\beta$ ($\neq \pm\alpha$) be roots in $\Phi$. Prove that the integers $k$ for which $\beta + k\alpha$ is a root are those from an interval $[-s, t]$ with $s, t \in \mathbb{N}$ and that $(\beta, \alpha^\vee) = s - t$.*

Remark. *Compare this with Theorem (6.9).*

**(8.27).** Problem. *Totally positive word or totally negative word is the same as minimal word.*

**(8.28).** Problem. *Highest root. $\Phi^+ \longrightarrow \Phi^-$.*

# Part III

# Important results and constructions

Chapter 9

# PBW and Free Lie Algebras

For the set $X$, the $\mathbb{K}$-space $V_X = \bigoplus_{x\in X} \mathbb{K}x$ has the property

> *Given any $\mathbb{K}$-space $V$ and set map $f\colon X \longrightarrow V$, there is a unique $\mathbb{K}$-linear transformation $f_V\colon V_X \longrightarrow V$ such that $f$ factors through $V_X$. That is, $f = f_V \iota_X$, where $\iota_X$ is the natural map taking $X$ to $V_X$.*

The appropriate commutative diagram is:

$$\begin{tikzcd} & V_X \arrow[d, dashed, "!\, f_V"] \\ X \arrow[ur, "\iota_X"] \arrow[r, "f"'] & V \end{tikzcd}$$

Thus in $V_X$ we easily find both the constructive and conceptual versions of the notion "free on $X$" in the category of $\mathbb{K}$-spaces. More impressively, but not a lot harder to prove, is that the tensor algebra $\mathrm{T}(V_X) = \bigoplus_{n\in\mathbb{N}} (V_X)^{\otimes n}$ is free on $X$ in the category of associative $\mathbb{K}$-algebras (always with identity).

The tensor algebra is also the constructive solution to the conceptual problem raised by a second universal property:

> *Given any associative $\mathbb{K}$-algebra $A$ and any $\mathbb{K}$-linear transformation $f\colon V \longrightarrow A$, there is a unique $\mathbb{K}$-algebra map $f_A\colon \mathrm{T}(V) \longrightarrow A$ such that $f$ factors through $\mathrm{T}(V)$. That is, $f = f_A \iota_V$, where $\iota_V$ is the natural linear transformation taking $V$ to $\mathrm{T}(V)$:*

$$\begin{tikzcd} & \mathrm{T}(V) \arrow[d, dashed, "!\, f_A"] \\ V \arrow[ur, "\iota_V"] \arrow[r, "f"'] & A \end{tikzcd}$$

Here the initial map $f$ is not just a set map but has "decoration"—it is a linear transformation. So we might phrase this by saying "$\mathrm{T}(V)$ is the free associative algebra on the $\mathbb{K}$-space $V$." (But we will avoid such terminology.)

In this chapter we deal with two related problems regarding Lie algebras. The first is the natural issue of free Lie algebras. Given the set $X$, the pair $(L_X, \iota_X)$, consisting of a Lie $\mathbb{K}$-algebra $L_X$ and a set map $\iota_X : X \longrightarrow L_X$, is *free on* $X$ provided it has the following universal property:

> *Given any Lie $\mathbb{K}$-algebra $L$ and set map $f\colon X \longrightarrow L$, there is a unique Lie algebra homomorphism $f_L\colon L_X \longrightarrow L$ such that $f$ factors through $L_X$. That is, $f = f_L \iota_X$:*

$$\begin{tikzcd} & L_X \arrow[d, dashed, "! f_L"] \\ X \arrow[ur, "\iota_X"] \arrow[r, "f"'] & L \end{tikzcd}$$

This conceptual definition is elegant enough to prove two of the standard results (Exercise) on universal objects:

**(9.1).** Lemma.

(a) *If $(L_X^{(1)}, \iota^{(1)})$ and $(L_X^{(2)}, \iota^{(2)})$ are two free Lie algebras on $X$ then they are isomorphic via a map $\varphi^{(1,2)}$ with $\iota^{(2)} = \varphi^{(1,2)}\iota^{(1)}$.*

(b) *If $(L_X, \iota_X)$ is a free Lie algebra on $X$ then $\iota_X(X)$ generates $L_X$ as Lie algebra.* □

As there are arbitrarily large Lie $\mathbb{K}$-algebras,

**(9.2).** Lemma. *If $(L_X, \iota_X)$ is a free Lie algebra on $X$ then $\iota_X$ is an injection.* □

Crucially, these say nothing regarding the actual existence of such a free algebra. The constructive aspect is missing and will be provided below in Section 9.3.

Our second universal object is the *universal enveloping algebra* for the Lie algebra $L$ over $\mathbb{K}$. This is an associative $\mathbb{K}$-algebra $\mathrm{U}(L)$ equipped with a Lie homomorphism $\iota_L\colon L \longrightarrow \mathrm{U}(L)^-$. (That is, $\iota_L$ is a homomorphism from $L$ to the linear Lie algebra $\mathrm{U}(L)^-$.) The universal enveloping algebra is conceptualized by:

> *Given any associative $\mathbb{K}$-algebra $A$ and any Lie algebra homomorphism $f\colon L \longrightarrow A^-$, there is a unique associative algebra homomorphism $f_A\colon \mathrm{U}(L) \longrightarrow A$ such that $f$ factors through $\mathrm{U}(A)$. That is, $f = f_A \iota_L$:*

$$\begin{tikzcd} & \mathrm{U}(L) \arrow[d, dashed, "! f_A"] \\ L \arrow[ur, "\iota_L"] \arrow[r, "f"'] & A \end{tikzcd}$$

The universal property again gives uniqueness up to isomorphism and generation by the image of $\iota_L$, but here does not immediately address injectivity of $\iota_L$.

The universal enveloping algebra $\mathrm{U}(L)$ was introduced in Section 4.2, where we constructed it as the quotient of the tensor algebra $\mathrm{T}(L)$ by the ideal $I_L$ in $\mathrm{T}(M)$ generated by all the elements $x \otimes y - y \otimes x - [x, y]$ for $x, y \in L$. This construction is not a great deal of immediate help, since we know little of the structure of $\mathrm{U}(L)$. We will describe it more carefully below in the Poincaré-Birkhoff-Witt (PBW) Theorem (9.3) and Section 9.2, including injectivity for $\iota_L$.

This does at least suggest that tensor algebras—the solutions to universal problems in the associative context—can also be of use in solving universal problems in the Lie context. We might make the wild guess that the Lie algebra $\mathrm{T}(V_X)^-$ is free on $X$. And we would be (essentially) right! But it is not at all clear that this Lie algebra solves the universal problem. The proof of this in Section 9.3 will invoke the PBW Theorem in a nontrivial way.

## 9.1 The Poincaré-Birkhoff-Witt Theorem

**(9.3).** THEOREM. (POINCARÉ-BIRKHOFF-WITT THEOREM) *Let the Lie algebra $L$ have the $\mathbb{K}$-basis $\{ v_i \mid i \in I \}$ for some totally ordered set $(I, \leq)$. For each $v_i$, let $\bar{v}_i$ be its image in $\mathrm{U}(L)$.*

(a) (WEAK PBW) *The universal enveloping algebra $\mathrm{U}(L)$ has as $\mathbb{K}$-spanning set the collection of all monomials $\bar{v}_{i_1} \cdots \bar{v}_{i_n}$ for $n \in \mathbb{N}$ and $i_i \leq \cdots \leq i_n$ (where $n = 0$ corresponds to the monomial $1$).*

(b) (STRONG PBW) *The universal enveloping algebra $\mathrm{U}(L)$ has as $\mathbb{K}$-basis the collection of all monomials $\bar{v}_{i_1} \cdots \bar{v}_{i_n}$ for $n \in \mathbb{N}$ and $i_i \leq \cdots \leq i_n$ (where $n = 0$ corresponds to the monomial $1$).*

As universal $\mathrm{U}(L)$ is generated by the image of $L$ as an associative algebra, it is clear that the set of all monomials $\bar{v}_{i_1} \cdots \bar{v}_{i_n}$, for all $n \in \mathbb{N}$, spans $\mathrm{U}(L)$. Therefore the Weak PBW Theorem (9.3)(a) follows directly from the following proposition:

**(9.4).** PROPOSITION. *Let $\sigma \in \mathrm{Sym}(n)$. Then in $\mathrm{U}(L)$ the element*

$$\bar{v}_{i_1} \cdots \bar{v}_{i_n} - \bar{v}_{i_{\sigma(1)}} \cdots \bar{v}_{i_{\sigma(n)}}$$

*is a linear combination of various monomials $\bar{v}_{j_1} \cdots \bar{v}_{j_k}$, each having $k$ less than $n$.*

PROOF. Let $[v_a, v_b] = \sum_{l \in I} c^l_{a,b} v_l$ in $L$. Then

$$\begin{aligned}
\bar{v}_{k_1} \cdots \bar{v}_{i_{j-1}}\, \bar{v}_a \bar{v}_b\, \bar{v}_{i_{j+2}} \cdots \bar{v}_{k_m} &- \bar{v}_{k_1} \cdots \bar{v}_{i_{j-1}} \bar{v}_b \bar{v}_a\, \bar{v}_{i_{j+2}} \cdots \bar{v}_{k_m} \\
&= \bar{v}_{k_1} \cdots \bar{v}_{i_{j-1}} \Big( \sum_{l \in I} c^l_{a,b} \bar{v}_l \Big) \bar{v}_{i_{j+2}} \cdots \bar{v}_{k_m} \\
&= \sum_{l \in I} c^l_{a,b} \bar{v}_{k_1} \cdots \bar{v}_{i_{j-1}} \bar{v}_l \bar{v}_{i_{j+2}} \cdots \bar{v}_{k_m}\,.
\end{aligned}$$

Thus the result is true for $\sigma = (j, j+1)$. The full symmetric group is generated by such 2-cycles. $\square$

Our proof of the PBW Theorem, in its strong version Theorem (9.3)(b) is motivated by that of Serre [Ser06, §3.4]. The basic idea is to construct a module for $L$, with nondecreasing monomials as basis, that is ultimately revealed as the adjoint module for associative $\mathrm{U}(L)$ acting on itself.

We let $\mathrm{M}(I)$ be the set of all finite multisets chosen from the totally ordered set $I$.[1] The connection with the discussion above is that each $n$-multiset $M$ has a unique nondecreasing representation $i_i \leq \cdots \leq i_n$. In this situation we will write $\bar{v}_M$ for the nondecreasing monomial $\prod_{i=1}^n \bar{v}_{i_j}$ (with $\bar{v}_\emptyset = 1$). The PBW Theorem (9.3)(b) states that these form a $\mathbb{K}$-basis for $\mathrm{U}(L)$.

For each $M \in \mathrm{M}(I)$ we let $x_M$ be a basis vector in the $\mathbb{K}$-space $X = \bigoplus_{M \in \mathrm{M}(I)} \mathbb{K} x_M$. If $M$ is the multiset represented by $i_i \leq \cdots \leq i_n$, then $n$ is the *degree* of $M$ and of $x_M$. For each degree $d \in \mathbb{N}$, we define the subspaces $X^d = \bigoplus_{\deg(M)=d} \mathbb{K} x_M$ and $X_d = \bigoplus_{i=1}^d X^d$. Here the vectors of $X_d \setminus X_{d-1}$ are said to have degree $d$. Especially, the only vectors of degree 0 are those of $\mathbb{K} x_\emptyset$.

We use the term degree because as $\mathbb{K}$-space $X$ is isomorphic to the $\mathbb{K}$-space $\mathbb{K}[x_i,\ i \in I]$ of polynomials in the indeterminates $x_i$. The isomorphism is given by $x_M \longleftrightarrow \prod_{i \in M} x_i$ and preserves degree. We largely avoid polynomials, since for us their algebra structure plays only a notational role. For each $j \in I$ and $M \in \mathrm{M}(I)$ we will write $x_{jM}$ in place of $x_{\{j\} \cup M} \longleftrightarrow x_j \prod_{i \in M} x_i$.

We begin our proof of the PBW Theorem by defining, for each $i \in I$, a linear transformation $a_i \in \mathrm{End}_{\mathbb{K}}(X)$. (The notation is to suggest the adjoint action of $\bar{v}_i$ on $\mathrm{U}(L)$.) These linear transformations are given by their action on the canonical basis of $X$.

For each $q \in I$ the action of $a_q$ on $x_Q$ is defined recursively by degree—assuming that, for all $t \in I$ and all $T \in \mathrm{M}(I)$ of degree less than $\deg(Q)$, the vector $a_t x_T$ has already been defined:

**(9.5).** Definition. *For $q \in I$ and $Q \in \mathrm{M}(I)$ we set*

$$a_q\, x_Q = x_{qQ} + y_Q^q$$

*where*

(a) $y_Q^q = 0$ *if $q \leq \min(Q)$;*

(b) $y_Q^q = a_r y_R^q + a_{[q,r]} x_R$ *if $q > \min(R) = r$ with $Q = rR$.*

Here we define $a_{[q,r]}$ to be $\sum_{l \in I} c_{q,r}^l a_l$ when $[v_q, v_r] = \sum_{l \in I} c_{q,r}^l v_l$ in $L$.

For $Q = \emptyset$ we are always in case (a) with $a_q x_Q = x_q$ as $q < \min(\emptyset)$ for all $q \in I$ (by convention). Furthermore, we always have $\deg(y_T^t) \leq \deg(T)$: this is clear for (a), while under (b) we have

$$\deg(y_Q^q) \leq 1 + \max(\deg(y_R^q), \deg(x_R)) \leq \deg(Q)$$

[1]Thus $\mathrm{M}(I)$ equals $\bigoplus_{i \in I} \mathbb{Z} i$, the set of all functions $F$ from $I$ to $\mathbb{Z}$ having only finitely many nonzero coordinate values $F_i$, the multiplicity of $i$ in the multiset $F$.

by induction on degree. Especially the recursion can proceed as given.

It is also helpful to realize that (b) holds for $q \geq \min(R)$. (Exercise.)

On its own, this definition lacks motivation. It says in part that $a_q x_Q$ and $x_{qQ}$ are equal, modulo terms of smaller degree. This fits with Proposition (9.4), since $a_q$ is designed to mimic the adjoint action of $\mathrm{U}(L)$ on itself. Part (a) is consistent, while the next lemma rephrases and motivates (b).

**(9.6).** LEMMA. *Let $q, r \in I$ and $R \in \mathrm{M}(I)$ with $q \geq r \leq \min(R)$. Then*

$$(a_q a_r - a_r a_q) x_R = a_{[q,r]} x_R\,.$$

PROOF.

$$\begin{aligned} a_q a_r x_R &= a_q x_{rR} \\ &= x_{qrR} + a_r y^q_R + a_{[q,r]} x_R \\ &= a_r x_{qR} + a_r y^q_R + a_{[q,r]} x_R \\ &= a_r (x_{qR} + y^q_R) + a_{[q,r]} x_R \\ &= a_r (a_q x_R) + a_{[q,r]} x_R\,. \end{aligned}$$ □

Having defined the various $a_i$, we now extend them to a linear transformation $a \colon L \longrightarrow \mathrm{End}_{\mathbb{K}}(X)$ via

$$a(v) = \sum_{i \in I} \alpha_i a_i \quad \text{when} \quad v = \sum_{i \in I} \alpha_i v_i \quad \text{with} \quad \alpha_i \in \mathbb{K}\,.$$

Especially $a(v_i) = a_i$ and $a([v_i, v_j]) = a_{[i,j]}$, as defined above. We shall also use the notation $a([[v_i, v_j], v_k]) = a_{[[i,j],k]}$.

**(9.7).** THEOREM. *The linear transformation $a \colon L \longrightarrow \mathrm{End}_{\mathbb{K}}(X)$ is a Lie homomorphism $a \colon L \longrightarrow \mathrm{End}^-_{\mathbb{K}}(X)$. That is,*

$$[a(v), a(w)] = a([v, w])$$

*for all $v, w \in L$.*

PROOF. For this we need only show

$$(a_i a_j - a_j a_i) x_M = a_{[i,j]} x_M$$

for all $i, j \in I$ and all $M \in \mathrm{M}(I)$.

We break this verification into two cases:

(i) $\min(M) \geq \min(i, j)$;

(ii) $\min(M) < \min(i, j)$.

Case (i) comes immediately from the motivational Lemma (9.6) when we set $R = M$, $r = \min(i, j)$, and $q = \max(i, j)$.

To complete the proof we must handle Case (ii). This we do by induction on $\deg(M)$, assuming that

$$[a(v), a(w)] x_T = a([v, w]) x_T$$

for all $v, w \in L$ and for all $T$ of degree smaller than that of $M$. Especially $[a_{[d,e]}, a_f]x_T = a_{[[d,e],f]}x_T$ for all $d, e, f \in I$ and such $T$. (Exercise.)

As $\min(M) < i$, Case (ii) is vacuous when $M$ is empty. This gets the induction started and also allows us to set $M = kN$ for $k = \min(M)$ and $\deg(N) = \deg(M) - 1$. We thus must verify

$$a_i a_j a_k \, x_N - a_j a_i a_k \, x_N = a_{[i,j]} a_k x_N$$

for all $i, j, k \in I$ and $N \in \mathrm{M}(I)$, subject to $k \le \min(i, j, N)$ and $\deg(N) < \deg(M)$.

With these restrictions, consider the related

$$\begin{aligned} a_j a_k a_i \, x_N - a_k a_j a_i \, x_N &= [a_j, a_k] a_i x_N \\ &= [a_j, a_k] x_{iN} + [a_j, a_k] y^i_N \\ &= a_{[j,k]} x_{iN} + a_{[j,k]} y^i_N && \text{by Case (i) and induction} \\ &= a_{[j,k]} a_i \, x_N \\ &= a_i a_{[j,k]} \, x_N + [a_{[j,k]}, a_i] x_N \\ &= a_i a_j a_k \, x_N - a_i a_k a_j \, x_N + a_{[[j,k],i]} x_N && \text{by induction.} \end{aligned}$$

That is,

$$a_j a_k a_i \, x_N - a_k a_j a_i \, x_N = a_i a_j a_k \, x_N - a_i a_k a_j \, x_N + a_{[[j,k],i]} x_N \, .$$

Symmetry in $i$ and $j$ gives

$$a_i a_k a_j \, x_N - a_k a_i a_j \, x_N = a_j a_i a_k \, x_N - a_j a_k a_i \, x_N + a_{[[i,k],j]} x_N \, .$$

After subtracting the second of these from the first and dropping common terms, we are left with

$$a_k a_i a_j \, x_N - a_k a_j a_i \, x_N = a_i a_j a_k \, x_N - a_j a_i a_k \, x_N + a_{[[j,k],i]} x_N - a_{[[i,k],j]} x_N \, .$$

That is,

$$a_k [a_i, a_j] \, x_N - a_{[[j,k],i]} x_N + a_{[[i,k],j]} x_N = a_i a_j a_k \, x_N - a_j a_i a_k \, x_N \, .$$

By induction and the Jacobi identity in $L$ on the lefthand side

$$a_k a_{[i,j]} \, x_N + a_{[[i,j],k]} x_N = a_i a_j a_k \, x_N - a_j a_i a_k \, x_N \, ,$$

which is the desired

$$a_{[i,j]} a_k \, x_N = a_i a_j a_k \, x_N - a_j a_i a_k \, x_N \, .$$

This completes Case (ii) and so our proof of the theorem. □

Theorem (9.7) provides the Lie homomorphism $a$ from $L$ to $\mathrm{End}^-_{\mathbb{K}}(X)$. Universality then provides a unique associative algebra homomorphism $\bar{a}\colon \mathrm{U}(L) \longrightarrow \mathrm{End}_{\mathbb{K}}(X)$ with $a = \bar{a}\iota_L$, where $\iota_L$ is the canonical map from $L$ to $\mathrm{U}(L)$, which we have rendered $v \mapsto \iota_L(v) = \bar{v}$. Recall that for the multiset $M$ represented by $i_i \le \cdots \le i_n$, the element $\bar{v}_M$ of $\mathrm{U}(L)$ is the monomial $\prod_{i=1}^n \bar{v}_{i_j}$ with $\bar{v}_\emptyset = 1$.

**(9.8).** PROPOSITION. *As* $\mathrm{U}(L)$*-module,* $X$ *is isomorphic to the adjoint module* ${}_{\mathrm{U}(L)}\,\mathrm{U}(L)$ *via an isomorphism taking each* $x_M$ *to* $\bar{v}_M$ *for* $M \in \mathrm{M}(I)$.

PROOF. We have $a(v_i) = \bar{a}(\bar{v}_i) = a_i$, for all $i \in I$, and $\bar{a}\big(\prod_{j=1}^n \bar{v}_{i_j}\big) = \prod_{j=1}^n a_{i_j}$. Especially $\bar{a}(\bar{v}_i)x_\emptyset = x_i$. By induction on degree, for the multiset $M$ with representation $i_1 \le \cdots \le i_n$,

$$\bar{a}(\bar{v}_M)x_\emptyset = \bar{a}\Big(\prod_{j=1}^n \bar{v}_{i_j}\Big)x_\emptyset = x_M\,;$$

indeed the action of each $\bar{a}(\bar{v}_{i_k}) = a_{i_k}$ on $\bar{a}(\prod_{j=k+1}^n \bar{v}_{i_j})x_\emptyset$ is calculated under (a) of Definition (9.5) since the representation is nondecreasing. We conclude that $X$ is a cyclic $\mathrm{U}(L)$-module generated by $x_\emptyset$. That is, there is a $\mathrm{U}(L)$-module homomorphism $A\colon \mathrm{U}(L) \longrightarrow X$ given by $1 \mapsto x_\emptyset$ hence $\bar{v}_i \mapsto x_i$, for all $i$, and $\bar{v}_M \mapsto x_M$ as displayed.

Suppose $\sum_{M\in\mathrm{M}(I)} c_M\bar{v}_M$ is in the kernel of $A$. Then in $X$

$$0 = A\Big(\sum_{M\in\mathrm{M}(I)} c_M\bar{v}_M\Big) = \sum_{M\in\mathrm{M}(I)} c_M A(\bar{v}_M) = \sum_{M\in\mathrm{M}(I)} c_M x_M\,.$$

By definition the $x_M$ are linearly independent in $X$, so $c_M = 0$ for all $M$. Thus the kernel is trivial, and $A$ is a module isomorphism whose inverse is described in the statement of the proposition. □

PROOF OF PBW THEOREM (9.3).

By universality $\mathrm{U}(L)$ is generated by the $\bar{v}_i$, for $i \in I$, so it is spanned by the set of all monomials $\prod_{j=1}^n \bar{v}_{i_j}$ for all $n \in \mathbb{N}$ (where we include 1 of degree $n = 0$). Therefore by Proposition (9.4) the nondecreasing monomials $\bar{v}_M$ span $\mathrm{U}(L)$. By Proposition (9.8) these are in turn linearly independent since their images $x_M$ in the isomorphic $\mathrm{U}(L)$-module $X$ are. Therefore they form a basis. This completes the proof of the PBW Theorem (9.3). □

## 9.2 Consequences for universal enveloping algebras

**(9.9).** THEOREM. *For the Lie algebra* $L$*, the map* $\iota_L\colon L \longrightarrow \mathrm{U}(L)$ *is an injection. Especially* $\iota_L$ *is a faithful representation of* $L$ *in* $\mathrm{U}(L)^-$.

PROOF. By design and definition, the map $\iota_L\colon L \longrightarrow \mathrm{U}(L)^-$ is a Lie homomorphism. For the basis $\{\, v_i \mid i \in I\,\}$ of $L$, the images $\iota_L(v_i) = \bar{v}_i$ are linearly independent in $\mathrm{U}(L)$ by the PBW Theorem (9.3). Thus the kernel of the homomorphism $\iota_L$ is trivial, and it is an injection. □

**(9.10).** THEOREM.

(a) *If* $A$ *is a subalgebra of the Lie algebra* $L$ *then the injection of* $A$ *into* $L$ *induces an injection of* $\mathrm{U}(A)$ *into* $\mathrm{U}(L)$.

(b) *If the Lie algebra $L$ is the vector space direct sum $A \oplus B$ of two subalgebras $A$ and $B$, then* $\mathrm{U}(L)$ *is isomorphic as vector space to the tensor product* $\mathrm{U}(A) \otimes_{\mathbb{K}} \mathrm{U}(B)$ *of the two subalgebras given by* (a).

PROOF. (a) Choose a totally ordered basis $\{ v_i \mid i \in I \}$ for $L$ that includes (say, as its initial segment) the totally ordered basis $\{ u_j \mid j \in J \}$ for $A$ with $J \subseteq I$. Then the universal property of $\mathrm{U}(A)$ gives a homomorphism $\iota_U$ from $\mathrm{U}(A)$ to $\mathrm{U}(L)$ taking each nondecreasing monomial $\bar{u}_M$ to the corresponding element $\bar{v}_M$. By the PBW Theorem (9.3) this basis of $\mathrm{U}(A)$ is mapped to a linearly independent set in $\mathrm{U}(L)$. The kernel of $\iota_U$ is therefore trivial, and $\iota_U$ is an injection.

(b) Choose bases $\{ u_j \mid j \in J \}$ for $A$ and $\{ w_k \mid k \in K \}$ for $B$ with the totally ordered set $I$ the disjoint union of $J$ and $K$ and with $j < k$ for all $j \in J$ and $k \in K$. Then by the PBW Theorem (9.3) the algebra $\mathrm{U}(A)$ has the basis $\{ \bar{u}_M \mid M \in \mathrm{M}(J) \}$; $\mathrm{U}(B)$ has the basis $\{ \bar{w}_N \mid N \in \mathrm{M}(K) \}$; and $\mathrm{U}(L)$ has the basis $\{ \bar{u}_M \bar{w}_N \mid M \in \mathrm{M}(J), N \in \mathrm{M}(K) \}$. This reveals $\mathrm{U}(L)$ as a copy of the $\mathbb{K}$-space $\mathrm{U}(A) \otimes_{\mathbb{K}} \mathrm{U}(B)$. □

## 9.3 Free Lie algebras

**(9.11).** THEOREM. *Let $X$ be a set, and define $V_X = \bigoplus_{x \in X} \mathbb{K}x$, the $\mathbb{K}$-space with $X$ as basis. Let $L_X$ be the Lie subalgebra of $\mathrm{T}(V_X)^-$ generated by $X$. If $\iota_X$ is the inclusion of $X$ in $L_X$, then $(L_X, \iota_X)$ is a free Lie $\mathbb{K}$-algebra on $X$.*

PROOF. We must prove that for any Lie $\mathbb{K}$-algebra $L$ and set map $f \colon X \longrightarrow L$ there is a unique Lie homomorphism $f_L$ with $f = f_L \iota_X$ :

$$\begin{tikzcd} & L_X \arrow[d, dashed, "!\, f_L"] \\ X \arrow[ur, "\iota_X"] \arrow[r, "f"] & L \end{tikzcd}$$

By definition $L_X$ is generated by $\iota_X(X)$, so if such a map $f_L$ exists then it has to be unique.

Let $j_x$ be the injection of $L_X$ into $\mathrm{T}(V_X)$ and $g$ the canonical map from $L$ to $\mathrm{U}(L)$. As $T(V_X)$ is free associative on $X$, there is a unique map $g_U$ from $\mathrm{T}(V_X)$ to $\mathrm{U}(L)$ with $g_U j_X i_X = gf$:

$$\begin{tikzcd} & L_X \arrow[r, "j_X"] & \mathrm{T}(V_X) \arrow[d, "!\, g_U"] \\ X \arrow[ur, "\iota_X"] \arrow[r, "f"] & L \arrow[r, "g"] & \mathrm{U}(L) \end{tikzcd}$$

The image of $gf$ is clearly in $g(L)$. As the diagram commutes, this is also true of $g_U j_X i_X$ and $g_U j_X$. Therefore we may truncate the diagram to:

$$\begin{tikzcd} & L_X \arrow[dr, "g_U j_X"] & \\ X \arrow[ur, "\iota_X"] \arrow[r, "f"] & L \arrow[r, "g"] & g(L) \end{tikzcd}$$

By the PBW Theorem (9.3) and especially Theorem (9.9), the map $g$ is an injection of $L$ into $\mathrm{U}(L)$ that is a Lie isomorphism of $L$ and $g(L)$ (as subalgebra of $\mathrm{U}(L)^-$). Thus $g$ has a left inverse $h$ that is a Lie isomorphism when restricted to $g(L)$. But then $hg_U j_X i_X = hgf = f$, and $f_L = hg_U j_X$ is the desired Lie homomorphism:

$$\begin{tikzcd} & L_X \arrow[d, dotted, "f_L"] \arrow[dr, "g_U j_X"] & \\ X \arrow[ur, "\iota_X"] \arrow[r, "f"] & L & g(L) \arrow[l, "h"'] \end{tikzcd}$$

□

# Chapter 10

# Kac-Moody Lie Algebras and Serre's Theorem

Let $L$ be a finite dimensional semisimple Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic 0 with Cartan subalgebra $H$ having basis $\{\, h_j \in H \mid 1 \leq j \leq l \,\}$. Further let $\{\, \alpha_i \in H^* \mid 1 \leq i \leq l \,\}$ be a basis of simple roots of the root system $\Phi$ contained in $H^*$ (a largely cosmetic change in terminology and notation). The $l \times l$ matrix $C(\Phi) = (c_{i,j})_{i,j}$ with entries from $\mathbb{K}$ (indeed from $\mathbb{Z}$) is the associated Cartan matrix where $c_{i,j} = \alpha_i(h_j)$ $(= (\alpha_i, \alpha_j^\vee)$; see Section 6.3).

From Proposition (8.5) we have:

**(10.1).** Proposition. *The Lie algebra $L$ is generated by elements $h_i, e_i, e_{-i}$ for $1 \leq i \leq l$ that satisfy the relations:*

(a) $[h_i, h_j] = 0$ *for all* $1 \leq i, j \leq l$*;*

(b) $[e_i, e_{-j}] = \delta_{i,j} h_j$ *for all* $1 \leq i, j \leq l$*;*

(c) $[h_i, e_j] = c_{j,i} e_j$ *for all* $1 \leq i, j \leq l$*;*

(d) $[h_i, e_{-j}] = -c_{j,i} e_{-j}$ *for all* $1 \leq i, j \leq l$*;*

(e) $\mathrm{ad}_{e_i}^{1-c_{j,i}}(e_j) = 0$ *and* $\mathrm{ad}_{e_{-i}}^{1-c_{j,i}}(e_{-j}) = 0$ *for* $1 \leq i, j \leq l$ *with* $i \neq j$*.* □

As discussed in Section 8.2, these relations are those that were used by Serre to prove simultaneously the uniqueness and existence of a semisimple Lie algebras for each Cartan matrix arising from a root system. But the importance of the first four of these relations had been noted earlier, going back to Weyl even, and had been used by Chevalley, Harish-Chandra, and Jacobson—to study uniqueness and existence issues regarding these algebras and their representation—before Serre's observation that the final relation pair is a critical addition.

## 10.1 Kac-Moody Lie algebras

An interesting fact, noted by both Kac and Moody, is that the final Serre relation is the only one dependent upon $C(\Phi)$ being integral with nonpositive entries off the diagonal. They investigated the following algebras, generalizing those studied earlier:

**(10.2).** PRESENTATION. *Given $C \in \mathrm{Mat}_l(\mathbb{K})$, the Lie $\mathbb{K}$-algebra $\mathrm{Ch}'(C)$ is generated by the elements $h_i, e_i, e_{-i}$ for $1 \leq i \leq l$ subject to the relations:*

(a) $[h_i, h_j] = 0$ *for all* $1 \leq i, j \leq l$;

(b) $[e_i, e_{-j}] = \delta_{i,j} h_j$ *for all* $1 \leq i, j \leq l$;

(c) $[h_i, e_j] = c_{j,i} e_j$ *for all* $1 \leq i, j \leq l$;

(d) $[h_i, e_{-j}] = -c_{j,i} e_{-j}$ *for all* $1 \leq i, j \leq l$.

We call these algebras (and the related $\mathrm{Ch}(C)$ below) *Chevalley algebras* because Chevalley (at least according to Harish-Chandra) was the first to notice their value is situations more general than that of $C$ arising from a root system.

In $C(\Phi)$ the matrix entries $c_{i,j}$ are the entries $\alpha_i(h_j)$ of a Gram matrix for a pairing of a $\mathbb{K}$-space $H$ and its dual. But for every matrix $C \in \mathrm{Mat}_l(\mathbb{K})$ there is a basis $\{ \alpha_i \mid 1 \leq i \leq l \}$ for $(\mathbb{K}^l)^*$ and a subset of vectors $\{ h_j \mid 1 \leq j \leq l \}$ in $\mathbb{K}^l$ with $C = (\alpha_i(h_j))_{i,j}$. (For instance, let $\alpha_i$ be the the standard basis of $(\mathbb{K}^l)^*$ and let $h_j$ be the $j^{\text{th}}$ column of $C$.) The problem is that when $C$ is singular, the set $\{ h_j \mid 1 \leq j \leq l \}$ cannot be linearly independent; it is not a basis. The cure is to pass to a bigger space $\mathbb{K}^n$ and its dual in order to find bases whose Gram matrix has $C$ as a submatrix. For this to happen, we must have $n \geq 2l - r$, where $r$ is the rank of $C$, and for $n = 2l - r$ there is an essentially unique solution. (Exercise.)

A second more general presentation is now given by:

**(10.3).** PRESENTATION. *Let $H$ be a $\mathbb{K}$-space containing the linearly independent set $\{ h_j \in H \mid 1 \leq j \leq l \}$. Further let $\{ \alpha_i \in H^* \mid 1 \leq i \leq l \}$ be linearly independent in $H^*$. Set $C = (\alpha_i(h_j))_{i,j} \in \mathrm{Mat}_l(\mathbb{K})$ (and assume that $\dim_{\mathbb{K}}(H) = 2l - \mathrm{rank}(C)$).*

*The Lie algebra $\mathrm{Ch}(C)$ is generated by the $\mathbb{K}$-space $H$ and the elements $e_i, e_{-i}$ for $1 \leq i \leq l$ subject to the relations:*

(a) $[g, h] = 0$ *for all* $g, h \in H$;

(b) $[e_i, e_{-j}] = \delta_{i,j} h_j$ *for all* $1 \leq i, j \leq l$;

(c) $[h, e_j] = \alpha_j(h) e_j$ *for all* $h \in H$ *and* $1 \leq j \leq l$;

(d) $[h, e_{-j}] = -\alpha_j(h) e_{-j}$ *for all* $h \in H$ *and* $1 \leq j \leq l$.

In dealing with these algebras, we will follow the usual abuse of notation, using the same notation for the elements of the generating set as for their images in the presented algebra. In the previous chapter we saw that for the appropriate free algebras the mappings $\iota_{\mathrm{Ch}'}$ and $\iota_{\mathrm{Ch}}$ were injective on the generating sets, but that does imply the same for those free algebras modulo their relation ideals $I_{\mathrm{Ch}'}$ and $I_{\mathrm{Ch}}$. For instance $h \in H$ might become 0 as $\iota_{\mathrm{Ch}}(h) + I_{\mathrm{Ch}} \in \mathrm{Ch}(C)$. Luckily in Theorem (10.5) below we shall see that the induced maps are injections on generators in all cases.

Further notice that in Presentation (10.3) the statement "generated by the $\mathbb{K}$-space $H$" indicates that we are assuming (without statement) all the relations required for the image of $H$ in $\mathrm{Ch}(C)$ to retain its natural structure as $\mathbb{K}$-space:

$$\alpha(\iota_{\mathrm{Ch}}(g)) + \beta(\iota_{\mathrm{Ch}}(h)) = \iota_{\mathrm{Ch}}(\alpha g + \beta h)\,,$$

for all $\alpha, \beta \in \mathbb{K}$ and $g, h \in H$.

**(10.4).** PROPOSITION.

(a) *The map given by $e_i \mapsto -e_{-i}$ and $h \mapsto -h$, for all $h \in H$, induces an order 2 automorphism of* $\mathrm{Ch}(C)$ *and of* $\mathrm{Ch}'(C)$, *the* Cartan involution $\omega$.

(b) *Always* $\mathrm{Ch}'(C) \leq [\mathrm{Ch}(C), \mathrm{Ch}(C)]$. *Indeed we have equality provided $C$ has no row consisting only of 0's. Furthermore* $\mathrm{Ch}'(C) = \mathrm{Ch}(C)$ *if $C$ is nonsingular.*

(c) $\mathrm{Ch}(C)$ *and* $\mathrm{Ch}'(C)$ *are determined up to isomorphism by $C$.*

(d) *The map $\iota_C \colon e_i \mapsto e_i$ $\iota_C \colon e_{-i} \mapsto e_{-i}$ and $\iota_C \colon h_j \mapsto h_j$ extends to a Lie algebra short exact sequence*

$$1 \longrightarrow \mathrm{Ch}'(C) \xrightarrow{\iota_C} \mathrm{Ch}(C) \longrightarrow K \longrightarrow 1$$

*whose restriction to $H$ is*

$$1 \longrightarrow D \xrightarrow{\iota_C} H \longrightarrow K \longrightarrow 1$$

*with $D = H \cap \mathrm{Ch}'(C) = \bigoplus_{j=1}^{l} \mathbb{K}h_j$ and $H = D \oplus K$.*

PROOF. (a) This is clear from the defining relations.

(b) This is clear from the defining relations.

(c) This is clear for $\mathrm{Ch}'(C)$. For $\mathrm{Ch}(C)$ it is a consequence of the "essential uniqueness" mentioned before the presentation.

(d) One direction is clear. The map $e_i \mapsto 0$, $h = d + k \mapsto k$ extends to a Lie homomorphism of $\mathrm{Ch}(C)$ onto $K$. The kernel is clearly an image of $\mathrm{Ch}'(C)$.

More work shows that the kernel is isomorphic to $\mathrm{Ch}'(C)$. This is done by showing that there is a split extension $\mathrm{Ch}'(C) \rtimes K$ that is a homomorphic image of and so isomorphic to $\mathrm{Ch}(C)$. $\square$

While the Chevalley algebras $\mathrm{Ch}(C)$ and $\mathrm{Ch}'(C)$ need not be the same, it is clear from the proposition that there is very little difference between them.

It turns out that $\mathrm{Ch}(C)$ is slightly easier to deal with, so we focus on it. We now (and for the remainder of this chapter) adapt the notation of Presentation (10.3). While many of the results are valid more generally, we make the uniform assumption that $\mathbb{K}$ is algebraically closed of characteristic 0.

We also let the *root lattice* $\Lambda_C = \sum_{i=1}^{l} \mathbb{Z}\alpha_i \, (\leq H^*)$, and set $\Lambda_C^+ = \sum_{i=1}^{l} \mathbb{Z}^+\alpha_i$ and $\Lambda_C^- = \sum_{i=1}^{l} \mathbb{Z}^-\alpha_i$. For a subspace $W$ and $\alpha \in \Lambda_C$

$$W_\alpha = \{\, w \in W \mid h(w) = \alpha(h)w,\ h \in H \,\}\,.$$

Let $N^+$ be the subalgebra of $\mathrm{Ch}(C)$ (and $\mathrm{Ch}'(C)$) generated by the $e_i$ for $1 \leq i \leq l$, and let $N^-$ be the corresponding subalgebra generated by the $e_{-i}$ for $1 \leq i \leq l$.

The main result of the current section is:

**(10.5).** THEOREM.

(a) $\mathrm{Ch}(C) = N^- \oplus H \oplus N^+$

(b) *$N^-$ is free on $\{\, e_{-i} \mid 1 \leq i \leq l \,\}$, and $N^+$ is free on $\{\, e_i \mid 1 \leq i \leq l \,\}$.*

(c) *$N^+ = \sum_{\alpha \in \Lambda_C^+} N_\alpha^+$ and $N^- = \sum_{\alpha \in \Lambda_C^-} N_\alpha^-$. Always $N_\alpha^\epsilon$ has finite dimension. Each $N_{\alpha_i}^+$ and $N_{\alpha_{-i}}^-$ has dimension 1, while if $k$ is an integer with $|k| \geq 2$ always $N_{k\alpha_i}^+ = 0 = N_{k\alpha_{-i}}^-$*

(d) *$\mathrm{Ch}(C)$ contains a unique ideal $R$ that is such maximal subject to $R \cap H = 0$. This ideal is left invariant by $\omega$.*

Before discussing the proof of this, we make a fundamental definition. The algebra $\mathrm{KM}(C) = \mathrm{Ch}(C)/R$ is the *Kac-Moody Lie algebra* with Cartan matrix $C$.

By the previous proposition and the last part of the theorem, the ideal $R$ is contained in the subalgebra $\mathrm{Ch}'(C)$. In the literature $\mathrm{KM}'(C) = \mathrm{Ch}'(C)/R$ is often called a *Kac-Moody Lie algebra*. Again the distinction between the two is not large.

The proof of the structure theorem for Chevalley algebras is based upon the following representation theoretic construction:

**(10.6).** THEOREM. *Let $V = \bigoplus_{i=1}^{l} \mathbb{K}v_i$ be a $\mathbb{K}$-space of dimension $l$. Choose $\lambda \in H^*$. Then $\mathrm{Ch}(C)$ has a representation $\chi_\lambda$ on $\mathrm{T}(V)$ such that:*

(a) *For $1 \leq i \leq l$ and for each $k \in \mathbb{N}$ we have $\chi_\lambda(e_{-i})\colon \mathrm{T}(V)^k \longrightarrow \mathrm{T}(V)^{k+1}$, given by $e_{-i}(a) = v_i a$ for all $a \in \mathrm{T}(V)$.*

(b) *For each $k \in \mathbb{N}$ we have $\chi_\lambda(h)\colon \mathrm{T}(V)^k \longrightarrow \mathrm{T}(V)^k$ acting diagonally; for instance $h(1) = \lambda(h)1$ and $h(v_i) = (\lambda - \alpha_i)(h)v_i$.*

(c) *For each $k \in \mathbb{N}$ we have $\chi_\lambda(e_i)\colon \mathrm{T}(V)^k \longrightarrow \mathrm{T}(V)^{k-1}$, especially $e_i(1) = 0$.*

PROOF. ...

PROOF OF THEOREM (10.5).

(a) Under (c) we will see that $\mathrm{Ch}(C) = N^- + H + N^+$, so here we must prove this sum to be direct.

Suppose $u = n^- + h + n^+$ is 0 in $\mathrm{Ch}(C)$. Then for all $\lambda$ we get

$$0 = \chi_\lambda(u)(1) = n^-(1) + \lambda(h)\,.$$

The constant (degree $\leq 0$) part of the righthand side is $\lambda(h)$, so this is 0 for all $\lambda$. This implies that $h = 0$. Especially, the algebra $H$ is embedded isomorphically in $\mathrm{Ch}(C)$. Furthermore $n^-(1) = 0$. Under (b) we will see that $N^-$ acts freely on $\mathrm{T}(V)$ in all the representations $\chi_\lambda$. Therefore $n^-(1) = 0$ gives $n^- = 0$, hence $n^+ = 0$ as well.

(b) Via $e_{-i} \leftrightarrow v_i$ we can view $V$ as the free vector space $V_X$ on the set $X = \{\, e_{-i} \mid 1 \leq i \leq l \,\}$, so $\mathrm{T}(V) = \mathrm{T}(V_X)$ becomes the free associative algebra on $X$ with $\chi_\lambda(e_{-i}).a = v_i a$ describing the left action of $\mathrm{T}(V_X)$ on itself. As in the previous chapter, $N_-$ then is identified with the corresponding free Lie algebra on $X$ within $T(V_X)^-$.

As $N^-$ is the free Lie algebra on $\{\, e_{-i} \mid 1 \leq i \leq l \,\}$, the action of the Cartan involution $\omega$ tells us that $N^+$ is the free Lie algebra on $\{\, e_i \mid 1 \leq i \leq l \,\}$.

(c) $\left(\sum_{\alpha\in\Lambda_C^-} N_\alpha^-\right) + H + \left(\sum_{\alpha\in\Lambda_C^+} N_\alpha^+\right)$ is left invariant under all $\mathrm{ad}(e_{-i})$, $\mathrm{ad}(e_i)$, and $\mathrm{ad}(h)$, and it contains all these generators. Therefore the sum is equal to $\mathrm{Ch}(C)$, whence $N^-$ and $N^+$ are as described (except for the remarks on dimension).

(d) Standard linear algebra tells us that each ideal $I$ in $\mathrm{Ch}(C) = \bigoplus_{\alpha\in\Lambda} \mathrm{Ch}(C)_\alpha$ is equal to $\bigoplus_{\alpha\in\Lambda} I_\alpha$. Especially, those that meet $H$ trivially are all in $\bigoplus_{0\neq\alpha\in\Lambda} \mathrm{Ch}(C)_\alpha$ and generate a unique maximal such ideal $R$. $\Box$

**(10.7).** LEMMA. $\mathrm{C}_{N^\epsilon}(N^{-\epsilon}) = 0$.

PROOF. ...

## 10.2 Generalized Cartan matrices

To say more about the Kac-Moody algebra $\mathrm{KM}(C)$, we make additional assumptions about the Cartan matrix $C = (c_{i,j})_{i,j} = (\alpha_i(h_j))_{i,j}$, all of which hold when $C = C(\Phi)$ for a root system $\Phi$.

**(10.8).** DEFINITION. *A matrix* $C = (c_{i,j})_{i,j} \in \mathrm{Mat}_l(\mathbb{K})$ *is a* generalized Cartan matrix *provided:*

(i) $c_{i,i} = 2$ *for all* $1 \leq i \leq l$*;*

(ii) $c_{i,j}$ *is a nonpositive integer for all* $1 \leq i, j \leq l$ *with* $i \neq j$*;*

(iii) $c_{i,j} \neq 0$ *if and only if* $c_{j,i} \neq 0$*.*

We continue with the notation of the notation and terminology of the previous section.such Additionally from now on we assume that $C$ is a generalized Cartan matrix. For the following result we only need Condition (10.8)(i).

**(10.9).** THEOREM. *In* $\mathrm{Ch}(C)$ *and* $\mathrm{KM}(C)$*, for each $i$ the subspace* $\mathbb{K}h_i \oplus \mathbb{K}e_i \oplus \mathbb{K}e_{-i}$ *of* $\mathrm{Ch}(C)$ *and* $\mathrm{KM}(C)$ *is a subalgebra isomorphic to* $\mathfrak{sl}_2(\mathbb{K})$.

PROOF. We have the appropriate relations for $\mathfrak{sl}_2(\mathbb{K})$:

$$[e_i, e_{-i}] = h_i\,, \quad [h_i, e_i] = c_{i,i}e_i = 2e_i\,, \quad [h_i, e_{-i}] = -c_{i,i}e_{-i} = -2e_{-i}\,,$$

therefore these three elements generate an image of $\mathfrak{sl}_2(\mathbb{K})$ which is nontrivial by Theorem (10.5). As $\mathfrak{sl}_2(\mathbb{K})$ is simple, we have the result. □

The next result requires Condition (10.8)(ii) in order to make sense. Its proof also makes use of (10.8)(iii).

**(10.10).** THEOREM. *In the Kac-Moody algebra* $\mathrm{KM}(C)$ *we have the Serre relations*

$$\mathrm{ad}_{e_i}^{1-c_{j,i}}(e_j) = 0 \quad \textit{and} \quad \mathrm{ad}_{e_{-i}}^{1-c_{j,i}}(e_{-j}) = 0$$

*for all* $1 \le i, j \le l$ *with* $i \ne j$.

PROOF. By the action of the Cartan involution and Lemma (10.7) it is enough to prove

$$[e_{-k}, \mathrm{ad}_{e_i}^{1-c_{j,i}}(e_j)] = 0$$

for all $k, i, j$. If $k \notin \{i, j\}$, then this is clear by relation (10.3)(b).

If $j = k \ne i$ then $e_{-j}$ commutes with $e_i$, so

$$\begin{aligned}
[e_{-j}, \mathrm{ad}_{e_i}^{1-c_{j,i}}(e_j)] &= \mathrm{ad}_{e_i}^{1-c_{j,i}}[e_{-j}, e_j] \\
&= \mathrm{ad}_{e_i}^{-c_{j,i}}[e_i, -h_j] \\
&= \mathrm{ad}_{e_i}^{-c_{j,i}}\, c_{i,j}e_i\,.
\end{aligned}$$

If $c_{j,i}$ is not equal to 0, then this is 0 as $\mathrm{ad}(e_i)e_i = 0$. If $c_{j,i} = 0$ then by Condition (iii) above also $c_{i,j} = 0$, and again we find 0.

It remains to consider the case $i = k \ne j$. This is done by looking at the action of the $\mathfrak{sl}_2(\mathbb{K})$ subalgebra $\mathbb{K}h_i + \mathbb{K}e_i + \mathbb{K}e_{-i}$ provided by Theorem (10.9). □

Define the *Serre algebra* $\mathrm{Se}(C)$ to be the quotient of $\mathrm{Ch}(C)$ by the ideal $I_{\mathrm{Se}}$ generated by all $\mathrm{ad}_{e_i}^{1-c_{j,i}}(e_j)$ and $\mathrm{ad}_{e_{-i}}^{1-c_{j,i}}(e_{-j})$ for $1 \le i, j \le l$ with $i \ne j$. We defined the Kac-Moody algebra $\mathrm{KM}(C)$ to be the quotient of $\mathrm{Ch}(C)$ by its ideal $I_{\mathrm{KM}}$, unique maximal subject to intersecting $H$ trivially. Thus $I_{\mathrm{KM}} \ge I_{\mathrm{Se}}$ and $\mathrm{Se}(C)$ has $\mathrm{KM}(C)$ as the quotient by its unique maximal ideal $I_{\mathrm{KM}}/I_{\mathrm{Se}}$ meeting $H$ trivially. The algebra $\mathrm{Se}(C)$ admits the Cartan automorphism $\omega$. In most situations of interest $\mathrm{Se}(C) = \mathrm{KM}(C)$.

## 10.3 The Weyl group of a generalized Cartan matrix

Continuing with the notation and terminology of this chapter, we extend the definition of the Weyl group for a root system (see Section 7.1) to arbitrary generalized Cartan matrices. On the $\mathbb{K}$-space $H^*$ define, for each $1 \le i \le l$, the linear transformation

$$r_i \colon \beta \mapsto \beta - \beta(h_i)\alpha_i \,.$$

This acts trivially on the annihilator of $h_i$ in $H^*$ and it takes $\alpha_i$ to $-\alpha_i$ as $\alpha_i(h_i) = c_{i,i} = 2$; that is, $r_i$ is a reflection on $H^*$ with center $\alpha_i$. We define the *Weyl group* of $C$, $\mathrm{W}(C)$, to be the subgroup of $\mathrm{GL}(H^*)$ generated by reflections $r_i$ for $1 \le i \le l$. It is in general not finite, and it can only be irreducible if $C$ is nonsingular. On the positive side, for each $1 \le j \le l$, we have

$$r_i(\alpha_j) = \alpha_j - \alpha_j(h_i)\alpha_i = \alpha_i - c_{j,i}\alpha_i \in \alpha_i + \mathbb{Z}\alpha_i \,,$$

so that $\mathrm{W}(C)$ leaves invariant the root lattice $\Lambda_C = \sum_{i=1}^{l} \mathbb{Z}\alpha_i$.

We next use the two theorems of the previous section to embed the Weyl group $\mathrm{W}(C)$ naturally in the Kac-Moody Lie algebra $\mathrm{KM}(C)$.

An endomorphism $c$ of $V$ is *locally nilpotent* if for every $v \in V$ there is an $n \in \mathbb{N}$ with $c^n(v) = 0$. As we are in characteristic 0, this allows to define $\exp(c) = \sum_{k=0}^{\infty} \frac{1}{k!} c^k$. This exponentiation process can be thought of as "integrating" a derivation to produce an automorphism, as in Chapter 3.

**(10.11).** LEMMA. *If the derivation $d$ of the Lie algebra $L$ over $\mathbb{K}$ is locally nilpotent, then $\exp(\mathrm{ad}(d))$ is an automorphism of $L$.*

PROOF. This is a relatively easy consequence of the Leibniz Proposition (5.1). □

An immediate consequence of Theorem (10.10) is that the Serre and Kac-Moody Lie algebras coming from a generalized Cartan matrix is *integrable*: the endomorphisms $\mathrm{ad}_{e_i}$ and $\mathrm{ad}_{e_{-i}}$ are locally nilpotent and so can be integrated to give automorphisms:

**(10.12).** COROLLARY. *The invertible linear transformations $\exp(\mathrm{ad}(e_i))$ and $\exp(\mathrm{ad}(e_{-i}))$ are automorphisms of $\mathrm{Se}(C)$ and $\mathrm{KM}(C)$.* □

PROOF. By Presentation (10.3)(b) the endomorphism $\mathrm{ad}_{e_i}$ is nilpotent on $e_{-j}$ as long as $j \neq i$. In that case $\mathrm{ad}_{e_i}(e_{-i}) \in H$; for each $h \in H$ we have $\mathrm{ad}_{e_i}(h) \in \mathbb{K}e_i$. Therefore $\mathrm{ad}^2_{e_i}(h) = 0$ and $\mathrm{ad}^3_{e_i}(e_{-j}) = 0$. Theorem (10.10) then completes the proof that $\mathrm{ad}_{e_i}$ is locally nilpotent. The Cartan involution reveals $\mathrm{ad}(e_{-i})$ to be locally nilpotent as well, and the lemma completes the argument. □

Let us pause to do some calculations with $2 \times 2$ matrices. The standard rendering of $\mathfrak{sl}_2(\mathbb{K}) = \mathbb{K}h_i + \mathbb{K}e_i + \mathbb{K}e_{-i}$ is as

$$h_i = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e_i = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_{-i} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

so that

$$\exp(e_i) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \exp(e_{-i}) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

An instance of Whitehead's Lemma then gives

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

or

$$\exp(e_i)\exp(-e_{-i})\exp(e_i) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Therefore at this level the Weyl reflection $r_i$ (see Section 8.4.1) is induced by the linear transformation $\exp(e_i)\exp(-e_{-i})\exp(e_i)$. This remains true if we move to the 3-dimensional adjoint representation of $\mathfrak{sl}_2(\mathbb{K})$ and the corresponding map $\exp(\mathrm{ad}_{e_i})\exp(\mathrm{ad}(-e_{-i}))\exp(\mathrm{ad}_{e_i})$. (Here, as in Theorem (3.11), $\exp(-a)\, b\, \exp(a) = \exp(\mathrm{ad}_a)(b)$.)

We now return to the Lie algebras $\mathrm{Se}(C)$ and $\mathrm{KM}(C)$. For each $1 \le i \le l$ set

$$\rho_i = \exp(\mathrm{ad}_{e_i})\exp(\mathrm{ad}(-e_{-i}))\exp(\mathrm{ad}_{e_i}),$$

an automorphism by Corollary (10.12), which the previous paragraph suggests might emulate the Weyl reflection $r_i$. This is indeed the case. Set $\mathrm{W}_\star(C) = \langle\, \rho_i \mid 1 \le i \le l\, \rangle \le \mathrm{Aut}(\mathrm{Se}(C))$.

**(10.13).** PROPOSITION.

(a) $\rho_i(h) = h$ *for each* $h \in H$.

(b) $\rho_i(\mathrm{Se}(C)_\alpha) = \mathrm{Se}(C)_{r_i(\alpha)}$.

(c) $\rho_i(\mathrm{KM}(C)_\alpha) = \mathrm{KM}(C)_{r_i(\alpha)}$.

(d) *The map* $\rho_i \mapsto r_i$ *extends to a homomorphism from* $\mathrm{W}_\star(C)$ *onto* $\mathrm{W}(C)$.

PROOF. Parts (a) and (b) come from calculation and then imply (c). □

The proposition and Theorem (10.5)(c) immediately give

**(10.14).** COROLLARY. *Let* $w \in \mathrm{W}(C)$.

(a) $\dim \mathrm{Se}(C)_\alpha = \dim \mathrm{Se}(C)_{w(\alpha)}$. *Especially,* $\dim \mathrm{Se}(C)_{w(\alpha_i)} = 1$.

(b) $\dim \mathrm{KM}(C)_\alpha = \dim \mathrm{KM}(C)_{w(\alpha)}$. *Especially,* $\dim \mathrm{KM}(C)_{w(\alpha_i)} = 1$. □

## 10.4 Serre's Theorem

**(10.15).** THEOREM. (SERRE'S THEOREM) *Let* $\mathbb{K}$ *be an algebraically closed field of characteristic* 0*, and let* $C = (c_{i,j})_{i,j}$ *be the Cartan matrix of the abstract root system* $(E, \Phi)$*. Then the generators and relations of Proposition (10.1) give a presentation of a semisimple Lie algebra* $L$ *over* $\mathbb{K}$ *with Cartan matrix* $C$ *and root system equivalent to* $(E, \Phi)$.

Indeed, in this case $L$ is isomorphic to $\mathrm{Se}(C) = \mathrm{KM}(C)$.

As $C$ is invertible we have $\mathrm{Ch}(C) = \mathrm{Ch}'(C)$ and $H = D$ in Proposition (10.4). Corollary (10.14) and (ii) below tell us that $H \oplus \bigoplus_{\alpha\in\Phi} \mathrm{KM}(C)_\alpha$ is a subspace of $\mathrm{KM}(C)$ of dimension $\dim H + |\Phi|$, which we know is the target dimension for $L$. Thus we want to prove that this is all of $\mathrm{Se}(C)$ and $\mathrm{KM}(C)$.

The main thing that needs to be proven is that $\mathrm{Se}(C)_\beta = 0$ for all $\beta \in \Lambda \setminus \Phi$. This is done by showing that for any $\beta$ in the root lattice but not in $\Phi$, there is a $w \in \mathrm{W}(C) = \mathrm{W}(\Phi)$ such that $w(\beta)$, when written as an integral linear combination of the various $\alpha_i$, has some positive coefficients and some negative coefficients. By Theorem (10.5) and Corollary (10.14) we get $0 = \dim \mathrm{Se}(C)_{w(\beta)} = \dim \mathrm{Se}(C)_\beta$, as desired.

The steps in the proof:

(i) Using the $\mathfrak{sl}_2(\mathbb{K})$ of Theorem (10.9) we prove that $\mathrm{Se}(C)$ and $\mathrm{KM}(C)$ are semisimple and $H$ is a Cartan subgroup.

(ii) Each $\mathrm{Se}(C)_{\alpha_i}$ is in $\mathrm{Se}(C) \setminus I_{\mathrm{KM}}$ (again via $\mathfrak{sl}_2(\mathbb{K})$).

(iii) By Theorem (10.5) we have $\mathrm{Ch}(C)_{k\alpha_I} \neq 0$ for $k \in \mathbb{Z}$ if and only if $k \in \{-1, 0, 1\}$

(iv) Let $\beta \in \Lambda \setminus \left(\bigcup_{i=1}^{l} \mathbb{Z}\alpha_i\right)$. Then is a $w \in \mathrm{W}(C) = \mathrm{W}(\Phi)$ such that $w(\beta)$, when written as an integral linear combination of the various $\alpha_i$, has some positive coefficients and some negative coefficients. Thus $\mathrm{Ch}(C)_\beta = 0$ by Theorem (10.5)(c).

(v) $\mathrm{Se}(C) = \mathrm{KM}(C)$ of dimension $\dim H + |\Phi|$ and $C = C(\Phi)$.

# Part IV

# Appendices

Appendix A

# Forms

## A.1 Basics

Let $\sigma$ be an automorphism of $\mathbb{K}$ with fixed field $\mathbb{F}$. For the $\mathbb{K}$-space $V$, the map $b\colon V \times V \longrightarrow K$ is a *$\sigma$-sesquilinear form* provided it is biadditive and

$$b(pv, qw) = pb(v, w)q^{\sigma}$$

for all $v, w \in V$ and $p, q \in \mathbb{K}$. The case $\sigma = 1$ is that of *bilinear forms.*

The form is *reflexive* if

$$b(v, w) = 0 \iff b(w, v) = 0 \ .$$

Important examples are the *$(\sigma, \eta)$-hermitian forms*: those $\sigma$-sesquilinear forms with always

$$b(v, w) = \eta b(w, v)^{\sigma}$$

for some fixed nonzero $\eta$. Observe that

$$b(v, w) = \eta b(w, v)^{\sigma} = \eta(\eta b(v, w)^{\sigma})^{\sigma} = \eta\eta^{\sigma} b(v, w)^{\sigma^2} \ .$$

Assuming that $b$ is not identically 0, there are $v, w$ with $b(v, w) = 1$; so $\eta\eta^{\sigma} = 1$. But then for all $a \in \mathbb{K}$

$$a = b(av, w) = b(av, w)^{\sigma^2} = a^{\sigma^2} \ ,$$

and $\sigma^2 = 1$.

For a $(\sigma, \eta)$-hermitian form that is bilinear we have $\sigma = 1$, and so $1 = \eta\eta^{\sigma} = \eta^2$, giving $\eta = \pm 1$. The case $(\sigma, \eta) = (1, 1)$ is that of *symmetric bilinear forms* or *orthogonal forms*, while $(\sigma, \eta) = (1, -1)$ gives *alternating forms* or *symplectic forms.*

For $S \subseteq V$ write $S^\perp$ for the subspace $\{\, v \in V \mid b(v,s) = 0\,,\ \text{for all } s \in S\,\}$ and say that $V$ and $b$ are *nondegenerate* provided its *radical*

$$\mathrm{Rad}(V,b) = \mathrm{Rad}(V) = \mathrm{Rad}(b) = V^\perp$$

is equal to $\{0\}$. If $\mathbb{E} \leq \mathbb{R}$ and $b$ is an orthogonal form, we say that $b$ is *positive definite* if it has the property

$$b(x,x) \geq 0 \text{ always and } b(x,x) = 0 \iff x = 0\,.$$

This is stronger than nondegeneracy.

The form $b$ restricts to a form on each subspace $U$ of $V$, and $U$ is a *nondegenerate subspace* provided its radical under this restriction is 0; that is, $U \cap U^\perp = 0$.

**(1.1).** LEMMA. *For the* $(\mathrm{Id},\eta)$*-hermitian form* $b\colon V \times V \longrightarrow \mathbb{E}$ *the map* $\rho^b\colon w \mapsto b(\cdot,w)$ *is a* $\mathbb{E}$*-homomorphism of* $V$ *into* $V^*$ *and the map* $\lambda^b\colon v \mapsto b(v,\cdot)$ *is a* $\mathbb{E}$*-homomorphism of* $V$ *into* $V^*$*. Here* $\ker \rho^b = V^\perp = \ker \lambda^b$. □

**(1.2).** LEMMA. *For the nondegenerate* $(\mathrm{Id},\eta)$*-hermitian form* $b\colon V \times V \longrightarrow \mathbb{E}$ *let* $U$ *be a finite dimensional subspace of* $V$*.*

(a) *The codimension of* $U^\perp$ *in* $V$ *is equal to the dimension of* $U$*, and* $U^{\perp\perp} = U$*.*

(b) *The restriction of* $h$ *to* $U$ *is nondegenerate if and only if* $V = U \oplus U^\perp$*.* □

Write the vector $v = \sum_{i \in I} v_i x_i$ for the basis $\mathcal{X} = \{\, x_i \mid i \in I\,\}$ as the column $I$-tuple $v = (\ldots, v_i, \ldots)$. The *Gram matrix* $G = G_{\mathcal{X}}$ of the form $b$ is the $I \times I$ matrix $(b(x_i,x_j))_{i,j}$, and we have a matrix representation of the form $b$:

$$b(v,w) = v^\top G\, w\,.$$

If $\mathcal{Y}$ is a second basis and $A$ is the $I \times I$ base change matrix that takes vectors written in the basis $\mathcal{Y}$ to their corresponding representation in the basis $\mathcal{X}$, then $G_{\mathcal{Y}} = A^\top G_{\mathcal{X}} A$.

**(1.3).** COROLLARY. *The nondegenerate* $(\mathrm{Id},\eta)$*-hermitian form* $b\colon V \times V \longrightarrow \mathbb{E}$ *on the finite dimensional space* $V$ *is nondegenerate if and only if its Gram matrix is invertible.* □

This point of view makes it clear that if $b\colon V \times V \longrightarrow \mathbb{E}$ is nondegenerate and $\mathbb{F}$ is and extension field of $\mathbb{E}$, then we have an induced nondegenerate form $b^{\mathbb{F}}\colon (\mathbb{F} \otimes_{\mathbb{E}} V) \times (\mathbb{F} \otimes_{\mathbb{E}} V) \longrightarrow \mathbb{F}$.

## A.2 Orthogonal geometry

Throughout this chapter $F$ will be a commutative field and $V$ will be a finite dimensional vector space over $F$. For any subset $W$ of $V$, we let $\langle W\rangle \leq V$ be the $F$-subspace of $V$ spanned by $W$.

Let $q\colon V \longrightarrow F$ be a *quadratic form* on the $F$-space $V$. That is,

$$q(\alpha x) = \alpha^2 q(x)\,,$$

for all $\alpha \in F$ and $x \in V$, and the associated form $h = h_q\colon V \times V \longrightarrow F$, given by

$$h(x,y) = q(x+y) - q(x) - q(y)\,,$$

is bilinear (and symmetric). For any subspace $W$ of $V$, the restriction of $q$ to $W$ is a quadratic form on $W$. We call $(V,q)$ an *orthogonal space* or a *quadratic space*. The associated bilinear form $h_q$ will typically be abbreviated to $h$.

Always $h(x,x) = 2q(x)$. So in characteristic other than 2, the bilinear form $h$ determines $q$. That is not the case in characteristic 2 where $h(x,x)$ is always 0: $h$ is a *symplectic form*.

If $K$ is an extension of $F$, then $q$ extends naturally to a quadratic form $q|^K$ on the tensor product $K \otimes_F V = V|^K$. Indeed for any totally ordered set $(I,<)$ and basis $\mathcal{I} = \{\, x_i \mid i \in I \,\}$ of the $E$-space $W$, any map $q_I\colon \mathcal{I} \longrightarrow E$ and Gram matrix $\{\, h(x_i,x_j) \in E \mid i < j \,\}$ extends by "linearity" to a unique quadratic form $q^W$ on $W$.

For $W \subseteq V$, we let $W^\perp = \{x \in V \mid h(x,w) = 0,\, w \in W\,\}$, an $F$-subspace of $V$. The form $q$ is *nondegenerate* if $V^\perp = 0$.

**(1.4).** Lemma. *Let $q$ be a quadratic form on the finite dimensional $F$-space $V$ with associated bilinear form $h$.*

(a) *For each $x \in V$, let $\lambda_x\colon V \longrightarrow F$ be given by $y^{\lambda_x} = h(x,y)$. Then $\lambda\colon V \longrightarrow V^*$ given by $x \mapsto \lambda_x$ is an homomorphism of $F$-vector spaces. It is an isomorphism if and only if $(V,q)$ is nondegenerate.*

(b) *If $(V,q)$ is nondegenerate then $\dim_F U + \dim_F U^\perp = \dim_F V$ for each subspace $U$.*

(c) *If $(V,q)$ is nondegenerate and $U \cap U^\perp = 0$, then $V = U \oplus U^\perp$ (which we may write as $U \perp U^\perp$).*

(d) *$(V,q)$ is nondegenerate if and only if $(K \otimes_F V, q|^K)$ is nondegenerate.*

Proof. The first part is routine, given the definitions. The rest then follows directly. □

A subset $S$ of $V$ is *singular* (or sometimes even *totally singular*) if the restriction of $q$ to $S$ is identically 0. If $U$ is a singular subspace, then $q$ induces a quadratic form on the quotient space $U^\perp/U$, nondegenerate if $(V,q)$ is nondegenerate.

A vector that is not singular is *nonsingular*, and a space $(V,q)$ in which all nonzero vectors are nonsingular is an *asingular space*.

Let $(V,q_V)$ and $(W,q_W)$ be quadratic spaces over $F$. An *isometry* from $(V,q_V)$ to $(W,q_W)$ is an invertible $g \in \mathrm{Hom}_F(V,W)$ with

$$q_W(v^g) = q_V(v)\,, \text{ for all } v \in V\,.$$

Thus two quadratic $F$-spaces are essentially the same precisely when they are *isometric.*

One dimensional quadratic spaces $Fx$ are easy to describe: for all $y = \alpha x \in Fx$ we have $q(y) = d\alpha^2$ for the constant $d = q(x)$. (Characteristic 2 quadratic 1-spaces are always degenerate.) The structure of 2-dimensional spaces is crucial.

**(1.5).** PROPOSITION. *Let $(V, q)$ be a quadratic $F$-space of dimension* 2*.*

(a) *If $0 \neq x \in V$ is singular with $x^\perp = V$, then $(V, q)$ is degenerate and, for $y \in V \setminus Fx$, we have $q(\beta x + \gamma y) = e\gamma^2$ where $e = q(y)$ is a constant.*

(b) *If $0 \neq x \in V$ is singular with $x^\perp \neq V$, then $(V, q)$ is nondegenerate and there are exactly two* 1*-spaces in $V$ consisting of singular vectors. In this case, we have a basis of singular vectors $x$ and $y$ with $h(x, y) = 1$, hence $q(\beta x + \gamma y) = \beta\gamma$. Especially, for each $\alpha \in F$ there are $z \in V$ with $q(z) = \alpha$.*

(c) *If all nonzero vectors of $V$ are nonsingular, then there is a quadratic extension $K$ of $F$ for which the extension $q|^K$ of $q$ to $K \otimes_F V = V|^K$ has nonzero singular vectors and so falls under* (a) *or* (b)*.*

*In this case $(V, q)$ is isometric to $K$ (as $F$-space) provided with the quadratic form $q_K(\kappa) = d\kappa\bar{\kappa}$, where the bar denotes Galois conjugation in $K$ over $F$ and $d \in F$ is fixed and nonzero. If $K$ is separable over $F$ then $(V, q)$ is nondegenerate; if $K$ is inseparable over $F$ (which forces* $\mathrm{char}F = 2$*) then $V = V^\perp$.*

PROOF. (a) This is immediate from the remarks about spaces of dimension 1.

(b) As $q(x) = 0$, $h(x, x) = 0$; so for $w \notin \langle x \rangle = x^\perp$ we have $h(x, w) \neq 0$. If necessary, replace $w$ by a scalar multiple so that $h(x, w) = 1$. Consider $y = \beta x + w$. Then

$$h(x, y) = h(x, w) = 1\,, \text{ and } q(y) = q(\beta x) + q(w) + h(\beta x, w) = q(w) + \beta\,.$$

Therefore $\beta = -q(w)$ gives a second 1-space $\langle y \rangle$ of singular vectors and all other nonzero vectors are nonsingular. Finally

$$q(\beta x + \gamma y) = q(\beta x) + q(\gamma y) + h(\beta x, \gamma y) = 0 + 0 + \beta\gamma = \beta\gamma\,.$$

In particular $q(\alpha x + y) = \alpha$.

(c) Choose a basis $\{u, v\}$ of $V$ with $q(u) = d$, $q(v) = f$, and $h(u, v) = e$. Then $q(\beta u + \gamma v) = d\beta^2 + e\beta\gamma + f\gamma^2$. As there are no singular vectors in $V$, the polynomial $dz^2 + ez + f$ is irreducible of degree 2 in $F[z]$ but has a root $\alpha$ in the quadratic extension $K = F(\alpha)$ of $F$.

When we identify $V$ with the $F$-space $K$ via the linear isomorphism given by $u \mapsto 1$ and $v \mapsto -\alpha$, so that $\beta u + \gamma v \mapsto \beta - \alpha\gamma = \kappa$, we find

$$q_K(\kappa) = q(\beta u + \gamma v) = \\ d\beta^2 + e\beta\gamma + f\gamma^2 = d(\beta - \alpha\gamma)(\beta - \bar{\alpha}\gamma) = d(\beta - \alpha\gamma)\overline{(\beta - \alpha\gamma)} = d\kappa\bar{\kappa}\,.$$

The space $(V|^K, q|^K)$ contains the singular 1-space spanned by $\alpha u+v$ and so comes under (a) or (b). We have $V|^K = K(\alpha u+v) \oplus Kv$ with $h([,|^{)}K]\alpha u + v\alpha u + v = 0$. We calculate

$$\begin{aligned} h([,|^{)}K]\alpha u + v{-}v &= q|^K(\alpha u + v - v) - q|^K(\alpha u + v) - q|^K(-v) \\ &= \alpha^2 q(u) - 0 - q(-v) \\ &= d\alpha^2 - f \\ &= d\alpha^2 - f - (d\alpha^2 + e\alpha + f) \\ &= -e\alpha - 2f\,. \end{aligned}$$

As $dz^2+ez+f$ is irreducible of degree 2 in $F[z]$, necessarily $d \neq 0 \neq f \in F$. But $\alpha \notin F$, so the quantity $-e\alpha - 2f$ is zero if and only if $e = 0$ and $\mathrm{char}(F) = 2$. This is in turn the case if and only if the polynomial and $K$ are both inseparable over $F$.

Thus if $K$ is separable over $F$ then $(V|^K, q|^K)$ is nondegenerate as in (b), and $(V, q)$ is also nondegenerate by Lemma (1.4)(d). If $K$ is inseparable over $F$, then $V|^K = K(\alpha u + v) \perp Kv$ with $h([,|^{)}K]vv = h(v, v) = 2q(v) = 0$. Thus $h|^K$ and $h$ as well are identically 0, and $V = V^\perp$. □

## A.3 Hyperbolic orthogonal spaces

The orthogonal space $(V, q)$ admits the *hyperbolic basis* $\mathcal{H} = \{\dots, f_i, g_i, \dots\}$ $(1 \leq i \leq m)$ provided for all $i, j, l$:

$$q(f_i) = q(g_j) = h(f_i, f_l) = h(g_j, g_l) = 0\,,\ h(f_i, g_j) = \delta_{i,j}\,.$$

Especially the dimension $2m$ of $V$ is even and $q$ is nondegenerate. The integer $m$ is the *index* of the form.

A hyperbolic 2-space of course provides an example, but so does the 4-dimensional $F$-space $\mathrm{Mat}_2(F)$ of $2 \times 2$ matrices over $F$ with $q$ the determinant function There the four matrix units form a hyperbolic basis (up to sign).

If $(V, q)$ has a hyperbolic basis, then we say that $q$ and $V$ are *split* or *hyperbolic.*

**(1.6).** PROPOSITION. *If $q$ is a nondegenerate quadratic form on the $F$-space $V$ of finite dimension, then the following are equivalent:*

(1) *$V$ has a hyperbolic basis.*

(2) *$V$ is a perpendicular direct sum of hyperbolic 2-spaces.*

(3) *Every maximal singular subspace has dimension $\dim_F(V)/2$.*

(4) *There are maximal singular subspaces $M$ and $N$ with $V = M \oplus N$.*

(5) *There is a singular subspace of dimension at least $\dim_F(V)/2$.*

(6) *For any basis $\chi$ of the totally singular subspace $X$, $V$ has a hyperbolic basis containing $\chi$.*

PROOF. (1) and (2) are clearly equivalent, and both are consequences of (6). (5) is a consequence of all the others. If the hyperbolic basis of (1) is the one given above, then the spaces $M = \langle \dots, f_i \dots \rangle$ and $N = \langle \dots, g_i, \dots \rangle$ are maximal singular with $V = M \oplus N$, as in (4).

Also (6) implies (3) as every singular subspace spanned by a subset of a hyperbolic basis is contained in such a maximal singular subspace of dimension $\dim_F(V)/2$.

It remains to prove that (5) implies (6), which we do by induction on $\dim(V)$ with Proposition (1.5) providing the initial step. (The case of dimension 1 being trivial since nondegenerate 1-spaces contain no nonzero singular vectors.) If $M$ is a singular subspace of dimension at least $\dim(V)/2$ and $z$ is singular, then $z^\perp \cap M$ contains a hyperplane of $M$ and singular $\langle z, z^\perp \cap M\rangle$ has dimension at least that of $M$. Thus, if necessary replacing $M$ or enlarging $\chi$, we may assume that $M \cap \chi$ is nonempty. Let $x \in M \cap \chi$. Then, for any $y$ in $(\chi \setminus \{x\})^\perp$ but not its hyperplane $\chi^\perp$, the 2-space $\langle x, y\rangle$ is hyperbolic by Proposition (1.5). Nondegenerate $\langle x, y\rangle^\perp$ contains $M \cap y^\perp$ and $\chi \setminus \{x\}$. By induction $\chi \setminus \{x\}$ embeds in a hyperbolic basis of $\langle x, y\rangle^\perp$, and therefore $\chi$ is in a hyperbolic basis of $V$. □

## A.4 Canonical forms

One natural example of an orthogonal form on $V$ is one that has an *orthonormal basis*; that is, the Gram matrix is the identity matrix.

In many situations, particularly over algebraically closed fields, other bases are of interest. We next define the *split forms* of orthogonal and symplectic type:

> For $\eta \in \{\pm\} = \{\pm 1\}$, the $\mathbb{K}$-space $V = V_\eta = \mathbb{K}^{2l}$ has basis $\{ e_i, e_{-i} \mid 1 \le i \le l \}$ and is equipped with the split $(\mathrm{Id}, \eta)$-form $b = b_\eta$ given by
> $$b(e_i, e_{-i}) = 1\,, b(e_{-i}, e_i) = \eta\,,\ \text{otherwise } b(e_a, e_b) = 0\,.$$
> The form is split orthogonal when $\eta = +1$ and split symplectic when $\eta = -1$.
>
> The $\mathbb{K}$-space $V = V_\eta = \mathbb{K}^{2l+1}$ has basis $\{ e_0, e_i, e_{-i} \mid 1 \le i \le l \}$ and is equipped with the split orthogonal form $b$ given by
> $$b(e_0, e_0) = 1\,,\ b(e_i, e_{-i}) = b(e_{-i}, e_i) = 1\,,\ \text{otherwise } b(e_a, e_b) = 0\,.$$

**(1.7).** LEMMA. *Consider the* $(\mathrm{Id}, \eta)$*-hermitian form* $b\colon V \times V \longrightarrow \mathbb{E}$ *on the* $\mathbb{E}$*-space* $V$ *of dimension* 2 *with* $\mathrm{char}\mathbb{E} \neq 2$. *Suppose* $b(x, x) = 0$ *but* $x \notin \mathrm{Rad}(V, b)$. *Then* $V$ *is nondegenerate, and there is a second vector* $y$ *with* $b(y, y) = 0$, $b(x, y) = 1$, *and* $V = \mathbb{E}x \oplus \mathbb{E}y$. *That is, the Gram matrix for* $b$ *in the basis* $\{x, y\}$ *of* $V$ *is* $\begin{pmatrix} 0 & 1 \\ \eta & 0 \end{pmatrix}$. □

**(1.8).** THEOREM. *Consider the nondegenerate symplectic form* $b\colon V \times V \longrightarrow \mathbb{E}$ *on the finite dimensional* $\mathbb{E}$*-space* $V$. *The form is split.*

PROOF. For a symplectic for $b(x, x) = 0$ always. Use the lemma and induction. $\square$

**(1.9).** THEOREM. *Consider the nondegenerate orthogonal form* $b\colon V \times V \longrightarrow \mathbb{E}$ *on the finite dimensional* $\mathbb{E}$*-space* $V$ *over the algebraically closed field* $\mathbb{E}$ *of characteristic not* 2.

(a) *If* $\dim_{\mathbb{E}}(V) \geq 2$*, then* $V$ *contains nonzero vectors* $x$ *with* $b(x, x) = 0$*.*

(b) *The form is split.*

PROOF. The first part allows the second part to be proved by induction using the lemma. $\square$

Appendix B

# Finite Groups Generated by Reflections

Let $E$ be a finite dimensional Euclidean space, and let $0 \neq v \in E$. The linear transformation

$$r_v \colon x \mapsto x - \frac{2(x,v)}{(v,v)} v$$

is the *reflection* with *center* $\mathbb{R}v$.

**(2.1).** LEMMA. *Let $0 \neq v \in E$.*

(a) $r_v^2 = 1$.

(b) *If $\alpha \in E^\times$ then $r_{\alpha v} = r_v$.*

(c) *$r_v \in \mathrm{O}(E)$, the orthogonal group of isometries of $E$.*

(d) *If $g \in \mathrm{O}(E)$ then $r_v^g = r_{g(v)}$.*

(e) *For the subspace $W \leq E$ we have $W^{r_x} \leq W$ if and only if $x \in W$ or $(x,W) = 0$.* □

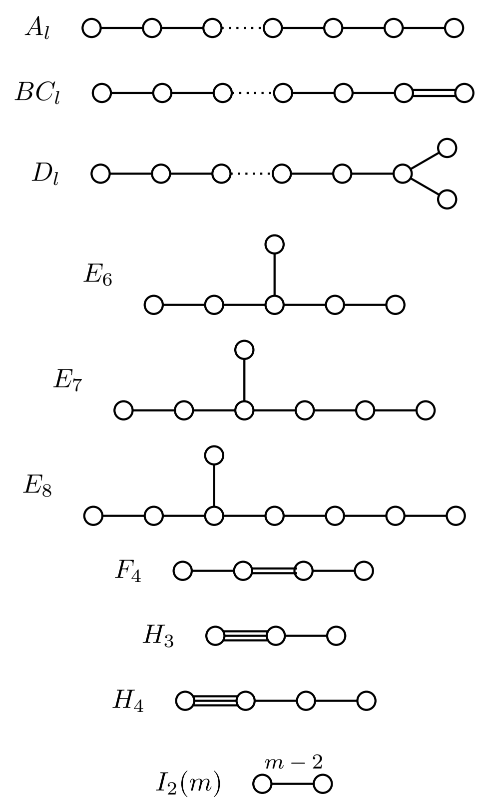
**(2.2).** LEMMA. *Let $\alpha$ and $\beta$ be independent vectors in the Euclidean space $E$. Then $\langle r_\alpha, r_\beta \rangle$ is a dihedral group in which the rotation $r_\alpha r_\beta$ generates a normal subgroup of index 2 and order $m_{\alpha,\beta}$ (an integer at least two or infinite) and the nonrotation elements are all reflections of order 2. In particular, the group $\langle r_\alpha, r_\beta \rangle$ is finite, of order $2m_{\alpha,\beta}$, if and only if the 1-spaces spanned by $\alpha$ and $\beta$ meet at the acute angle $\frac{\pi}{m_{\alpha,\beta}}$.* □

## B.1 Coxeter graphs

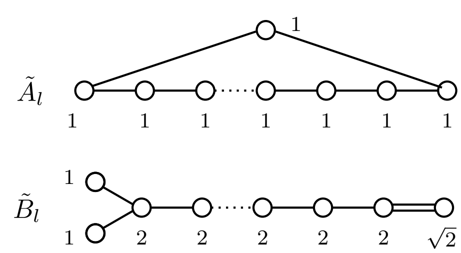
We are concerned in this appendix with finite subgroups of $\mathrm{O}(E)$ generated by a set $\{ r_v \mid v \in \Delta \}$ of reflections (necessarily finite itself).

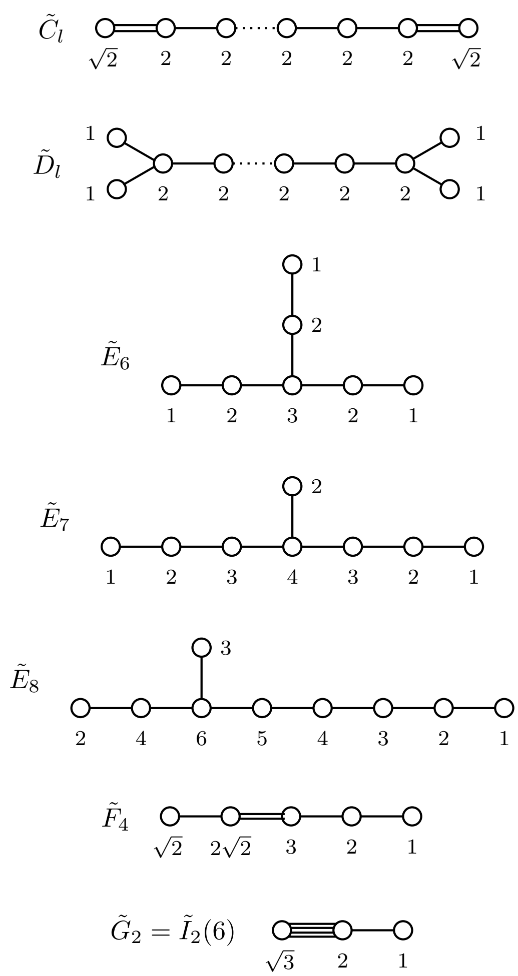
The *Coxeter graph* of this reflection set has $\Delta$ as vertex set, with $\alpha$ and $\beta$ connected by a bond of strength $m_{\alpha,\beta} - 2$ where $\langle r_\alpha, r_\beta \rangle$ is dihedral of order $2m_{\alpha,\beta}$, for the positive integer $m_{\alpha,\beta} \geq 2$. In particular, distinct $\alpha$ and $\beta$ are not connected if and only if $m_{\alpha,\beta} = 2$ if and only if they commute.

**(2.3).** THEOREM. *The Coxeter graph for an irreducible finite group generated by the $l$ distinct Euclidean reflections for an obtuse basis is one of the following:*



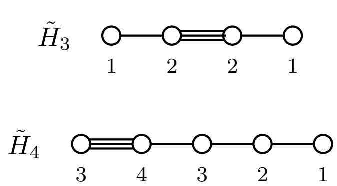

**(2.4).** PROPOSITION. *Let $G$ be the Gram matrix associated with one of the graphs below. Then the positive vector $x$ whose coordinate entries are under or adjacent to the corresponding node of the graph has $Gx = 0$. Especially, the associated form is not positive definite.*






□

**(2.5).** Proposition. *Let $G$ be the Gram matrix associated with one of the two graphs below. Then the positive vector $x$ whose coordinate entries are under the corresponding node of the graph has $x^\top Gx < 0$.*




*Note that* $\cos(\frac{\pi}{5}) = \frac{1+\sqrt{5}}{4}$. □

## B.2 Some finite groups generated by reflections

We describe the three "classical" families of finite groups generated by reflections.

The center $\mathbb{R}v$ of the reflection $r_v : x \mapsto x - \frac{2(x,v)}{(v,v)} v$ is characterized as the range $[r_v, E]$ of the linear transformation $r_v - 1$ in its action on $E$. Indeed, the only linear isometry $r$ of $E$ with one dimensional range $[r, E] = \mathbb{R}v$ is the reflection $r_v$, since $E = \mathbb{R}v \perp v^\perp$ with $r$ trivial on $v^\perp$ ($r$-invariant but not containing the range) and taking $v$ to $-v$ ($r$-invariant with $r$ inducing a nontrivial isometry).

### B.2.1 The symmetric group and $A_l$

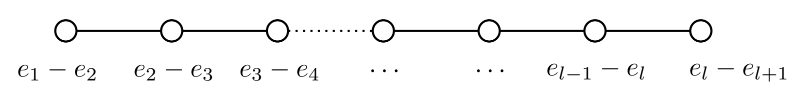
The 2-cycle permutation $(1,2)$ acting on the permutation module $\bigoplus_{i=1}^{l+1} \mathbb{R}e_i$ has corresponding matrix

$$r = \begin{pmatrix} 0 & 1 & \vdots & 0 \\ 1 & 0 & \vdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \vdots & I_{l-1} \end{pmatrix}$$

so that

$$r - I_{l+1} = \begin{pmatrix} -1 & 1 & \vdots & 0 \\ 1 & -1 & \vdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \vdots & 0 \end{pmatrix}$$

with range $\mathbb{R}(e_1 - e_2)$. Similarly the 2-cycle $(i, i+1)$ is a reflection with center $\mathbb{R}(e_i - e_{i+1})$. As the symmetric group $\mathrm{Sym}(l+1)$ is generated by the $l$ distinct 2-cycles $(1,2)$ through $(l, l+1)$, it is a finite group generated by $l$ reflections having Coxeter graph $A_l$ :



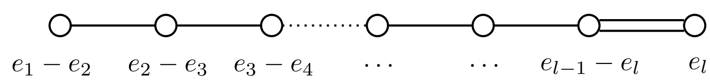

This representation is reducible, since in $\mathbb{R}^{l+1}$ the symmetric group leaves invariant the 1-space $\mathbb{R}\mathbf{1}$ of constant vectors and its complement $\mathbb{R}^{l+1} \cap \mathbf{1}^\perp$ (with respect to the standard dot product) consisting of vectors with sum 0.

### B.2.2 The $\pm 1$-monomial groups and $BC_l$

All reflections on $\mathbb{R}^l$ are conjugate to

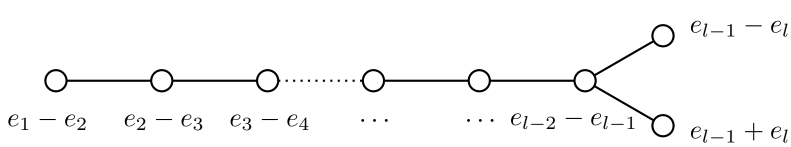
$$r = \begin{pmatrix} -1 & 0 & \vdots & 0 \\ 0 & 1 & \vdots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \vdots & I_{l-2} \end{pmatrix}$$

with center spanned by $e_1$. The symmetric group $\mathrm{Sym}(l)$ (as described above) acts on this space via permutation matrices and normalizes the diagonal subgroup and especially its subgroup $D$ having all diagonal entries $\pm 1$. The subgroup $D \simeq 2^l$ has as $\mathbb{F}_2$-basis the reflections with centers $e_i$, these being permuted by $\mathrm{Sym}(l)$. Therefore the $\pm 1$-monomial group $D : \mathrm{Sym}(l) = 2^l : \mathrm{Sym}(l)$ is a finite group generated by $l$ reflections having the Coxeter graph $BC_l$ :




### B.2.3 The even monomial groups and $D_l$

Similarly to the action of $\mathrm{Sym}(l+1)$ on $\mathbb{R}^{l+1}$ discussed above, the subgroup $\mathrm{Sym}(l)$ of the group of type $BC_l$ is reducible on its $\mathbb{F}_2$-module $D$—it leaves invariant the subgroup $\pm I$ and its "perp" $E$ consisting of all $\pm 1$-diagonal matrices containing an even number of $-1$'s. The subgroup $E : \mathrm{Sym}(l) = 2^{l-1} : \mathrm{Sym}(l)$ is again a finite group generated by $l$ reflections. Its Coxeter graph is $D_l$ with generating centers given by:



# Bibliography


[Alb34] A.A. Albert, *On a certain algebra of quantum mechanics,* Ann. of Math. (2) **35** (1934), 65–73.

[Bo01] A. Borel, "Essays in the History of Lie Groups and Algebraic Groups," History of Mathematics **21**, American Mathematical Society, Providence, RI and London Mathematical Society, Cambridge, 2001.

[Car05] R. Carter, "Lie Algebras of Finite and Affine Type," Cambridge Studies in Advanced Mathematics **96**, Cambridge University Press, Cambridge, 2005.

[CSM95] R. Carter, G. Segal, and I. Macdonald, "Lectures on Lie Groups and Lie Algebras," London Mathematical Society Student Texts **32**, Cambridge University Press, Cambridge, 1995.

[Coh54] P.M. Cohn, *On homomorphic images of special Jordan algebras*, Canadian J. Math. **6** (1954), 253–264.

[Dre99] L. Dresner, "Applications of Lie's Theory of Ordinary and Partial Differential Equations," Institute of Physics Publishing, Bristol, 1999.

[Eld15] A. Elduque, *Course notes: Lie algebras*, Univ. de Zaragosa, 2015, pp. 1–114.

[Gle52] A.M. Gleason, "Groups without small subgroups," Ann. of Math. (2) **56** (1952), 193–212.

[Hal15] B. Hall, "Lie groups, Lie algebras, and Representations. An Elementary Introduction," Second edition, Graduate Texts in Mathematics **222**, Springer 2015.

[Haw00] T. Hawkins, "Emergence of the Theory of Lie Groups, An Essay in the History of Mathematics 1869–1926," Sources and Studies in the History of Mathematics and Physical Sciences, Springer-Verlag, New York, 2000.

[Hel01] S. Helgason, "Differential Geometry, Lie Groups, and Symmetric Spaces," Graduate Studies in Mathematics **34**, American Mathematical Society, Providence, RI, 2001.

[Hig58] G. Higman, *Lie ring methods in the theory of finite nilpotent groups* in: "Proc. Internat. Congress Math. 1958," Cambridge Univ. Press, New York, 1960, 307–312.

[How83] R. Howe, *Very basic Lie theory*, Amer. Math. Monthly **90** (1983), 600–623.

[Hum78] J.E. Humphreys, "Introduction to Lie Algebras and Representation Theory," Graduate Texts in Mathematics **9**, Springer-Verlag, New York-Berlin, 1978.

[Jac79] N. Jacobson, "Lie Algebras," Dover Publications, Inc., New York, 1979.

[Jac89] N. Jacobson, "Basic Algebra II," W.H. Freeman and Company, New York, 1989.

[JvNW34] P. Jordan, J. von Neumann, and E. Wigner, *On an algebraic generalization of the quantum mechanical formalism*, Ann. of Math. **35** (1934), 29–64.

[KhW09] B. Khesin and R. Wendt. "The geometry of infinite-dimensional groups," Ergebnisse der Mathematik und ihrer Grenzgebiete, **3** Folge **51**, Springer-Verlag, Berlin, 2009.

[Kir08] A. Kirillov, Jr., "An Introduction to Lie Groups and Lie Algebras," Cambridge Studies in Advanced Mathematics **113**, Cambridge University Press, Cambridge, 2008.

[Lee13] J.M. Lee, "Introduction to Smooth Manifolds," Graduate Texts in Mathematics **218**, Springer, New York, 2013.

[Maz10] V. Mazorchuk, "Lectures on $\mathfrak{sl}_2(\mathbb{C})$-modules," Imperial College Press, London, 2010.

[MoZi55] D. Montgomery and L. Zippin, "Topological Transformation Groups," Interscience Publishers, New York-London, 1955.

[vNe29] J. v. Neumann, *Über die analytischen Eigenschaften von Gruppen linearer Transformationen und ihrer Darstellungen*, Math. Z. **30** (1929), 3–42.

[Ros02] W. Rossmann, "Lie Groups. An Introduction Through Linear Groups," Oxford Graduate Texts in Mathematics **5**, Oxford University Press, Oxford, 2002.

[Ser06] J.-P. Serre, "Lie Algebras and Lie Groups," Lecture Notes in Mathematics **1500**, Springer-Verlag, Berlin, 2006.

[Spi65] M. Spivak, "Calculus on Manifolds," Mathematics Monograph Series, W.A. Benjamin, 1965.

[SpV00] T.A. Springer and F.D. Veldkamp, "Octonions, Jordan Algebras and Exceptional Groups," Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000.

[Ste70] I. Stewart, "Lie Algebras," Lecture Notes in Mathematics **127**, Springer-Verlag, Berlin-New York 1970.

[Sti08] J. Stillwell, "Naive Lie Theory," Undergraduate Texts in Mathematics, Springer, New York, 2008.

[Tao14] T. Tao, "Hilbert's Fifth Problem and Related Topics," Graduate Studies in Mathematics **153**, American Mathematical Society, Providence, RI, 2014.

[Tap05] K. Tapp, "Matrix Groups for Undergraduates," Student Mathematical Library **29**, American Mathematical Society, Providence, RI, 2005

[Tit66] J. Tits, *Algèbres alternatives, algèbres de Jordan et algèbres de Lie exceptionnelles. I. Construction,* Nederl. Akad. Wetensch. Proc. Ser. A **69** (= Indag. Math. **28**) (1966), 223–237.

[Tu11] L.W. Tu, "An Introduction to Manifolds," Universitext, Springer, New York, 2011.

[Vin03] E.B. Vinberg, "A Course in Algebra," Graduate Studies in Mathematics **56**, American Mathematical Society, Providence, RI, 2003.

[Zel97] E. Zelmanov, *On the restricted Burnside problem*, "Fields Medallists' Lectures," , World Sci. Ser. 20th Century Math. **5**, World Sci. Publ., River Edge, NJ, 1997, 623–632.

[Zor31] M. Zorn, *Theorie der alternativen Ringe*, Abh. Math. Sem. Hamburg Univ. **8** (1931), 123–147