# Introduction to Lie Algebras

J.I. Hall

22 December 2015

# Contents

# Preface

These are notes for my Michigan State University, Fall Semester 2015, course MTH914: *Lie Algebras*. The primary aim of the course was the introduction and discussion of the finite dimensional semisimple Lie algebras over algebraically closed fields of characteristic 0 and their representations.

Unfortunately there was not enough time to cover adequately many additional topics, including: Serre's Theorem, the proof of PBW, the construction of $\mathfrak{e}_8$, Weyl's character formula, automorphisms, and the real forms of the complex semisimple Lie algebras.

The problems and some of the proofs (particularly later in the course) are incomplete, brief, or sketched. There is also material that was covered but remains $\boxed{\text{to be included.}}$

The notation $\square$ indicates my feeling that enough proof has been provided (even when that is nothing). At the other end of the spectrum $\square\square$ indicates that the result has been stated but will not be proven. This is usually because the result is too ambitious for the course but deserves to be pointed out.

The bibliography contains a long list of references, all helpful in the preparation for the course and notes. Three of these particularly stand out:

[Eld15] A. Elduque, *Course notes: Lie algebras*, Universidad de Zaragosa, 2015, pp. 1–114.

[Maz10] V. Mazorchuk, "Lectures on $\mathfrak{sl}_2(\mathbb{C})$-modules," Imperial College Press, London, 2010.

[Ste70] I. Stewart, "Lie Algebras," Lecture Notes in Mathematics **127**, Springer-Verlag, Berlin-New York 1970.

I thank Professor V. Futorny for discussion of the topic and for pointing me toward the first two references above, and I thank Professor A. Elduque for giving me permission to use his excellent notes. The course would not have been as good or interesting without helpful suggestions from these two professors and

from Professor O. Mathieu.

Jonathan I. Hall
Department of Mathematics
Michigan State University
East Lansing, MI 48840 USA
22 December 2015

# 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 endomorphism of* $A$ *into* $\mathrm{End}_{\mathbb{K}}(A)$ *.* $\square$

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 $A$. 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 a itself 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 category ${}_{\mathbb{K}}\mathsf{Alg}$ 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

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

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 $a$ 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 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 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

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

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 be 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 $\mathrm{H}_n(\mathbb{C})$, a set which 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 $n \leq 3$, 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 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 endomorphism.

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* $\mathbb{K}$*-endomorphism* $\mathrm{ad}\colon L \longrightarrow \mathrm{End}^-_{\mathbb{K}}(L)$ *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 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 (9.2). 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.

[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.

## 1.5 Problems

**(1.7).** PROBLEM.

(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 for* $\mathbb{K}$ *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* $A$ *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

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

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

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

Of course at this point 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

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 *derived 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 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.)

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

and the algebra presented at the end of the previous section is indeed a Lie algebra, namely a copy of $\mathfrak{sl}_2(\mathbb{K})$.

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 $\mathfrak{L}=\mathfrak{L}(V,b)$. 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.** If $b$ is the usual nondegenerate orthogonal form with an orthonormal basis, then $\mathfrak{so}_n(\mathbb{K})=\mathfrak{L}$. 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 $J$ 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 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 $\mathfrak{sp}_{2l}(\mathbb{K})=\mathfrak{L}$. As matrices,

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

where $J$ 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. $\mathrm{Der}_{\mathbb{K}}(A) \leq \mathrm{End}^{-}_{\mathbb{K}}(A)$. *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 $A$ 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* $\{\, t^i \frac{d}{dt} \mid t \in \mathbb{N}\,\}$.

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

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

completing the proposition. $\square$

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 7.

(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.

### 2.4.3 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 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.

## 2.5 Other constructions

### 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 of 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 starting 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 Chapter 6 below.)

In such situations it is more usual to write 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.

### 2.5.3 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 $1 \le k \le n$ 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

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

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

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

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 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* $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[x_1, \dots x_n])$.

**(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 2.5.3. 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 2.5.3. 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 real 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 2.5.3.

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

$$x = x_1 \longrightarrow x_s = e^s x \quad y = y_1 \longrightarrow y_s = e^{s/2} y\,,$$

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 $F\colon I \longrightarrow \mathrm{Mat}_n(\mathbb{R})$ for some open interval $I$ in $\mathbb{R}$ that contains the point $a$. 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* $I = \exp(A)\exp(-A)$*.*

(c) *For all* $A \in \mathrm{Mat}_n(\mathbb{R})$ *the unique solution of the matrix ordinary differential equation*
$$f'(t) = f(t)A\,, \quad f(0) = I$$
*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 Proposition (3.2) below.)

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 of $q$ at $0$ (with $k = t^{-1}$)

$$\begin{aligned} 0 &= \lim_{t\to 0}\frac{\log g(t) - tA}{t} = \lim_{k\to\infty} k\left(\log g(k^{-1}) - k^{-1}A\right) \\ &= -A + \lim_{k\to\infty} k\log g(k^{-1}) . \end{aligned}$$

That is, $A = \lim_{k\to\infty} k\log g(k^{-1})$. 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. $\square$

## 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

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

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 ordinary differential equation

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

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

Let $G$ be a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$. 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.

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. 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 $\Lambda(G) \subseteq \mathrm{T}_I(G)$, 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 $g\left(\frac{1}{k}\right)^k \in G$. By Proposition (3.2), this 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 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 let $G$ be a closed subgroup of $\mathrm{Mat}_n(\mathbb{R})$ and set $L = \Lambda(G) = \mathrm{T}_1(G)$.

**(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 being always 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.

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

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

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

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

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

**(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 $t \in \mathbb{R}$,

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

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 $G$*.

## 3.5 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

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

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:*

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

This is the easiest theorem of the present section. Especially the candidate for the *differential* $df$ of $f$ is evident, and we give it in the matrix case. If $A \in \Lambda(G)$, then $\varphi(t) = \exp(tA)$ is a one-parameter subgroup of $G$. If we compose it with $f$, then $f\varphi(t) = f(\exp(tA))$ is a one-parameter subgroup of $H$. Therefore there is a $B \in \Lambda(H)$ with $f\varphi(t) = \exp(tB)$. We set $df(A) = B$. This is clearly unique. 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].

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 $\mathsf{B}$.

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.) 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. Thus we get a category equivalence.

**(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 the identity. 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 in that component are fundamentally the same. For example, the 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 provides a Lie group homomorphism from it to $\mathbb{S}^1$, say

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

but this map has no inverse.

**(3.15).** THEOREM. (LIE'S THIRD THEOREM) *For all finite dimensional Lie algebras $L$, there is a Lie group $G$ with $\Lambda(G)$ isomorphic to $L$.* □□

[3] This is not standard terminology.

As the Lie group $G$ is a manifold, its Lie algebra must be finite dimensional.

**(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 ${}_{\mathbb{R}}\mathsf{LieAlg}^{fd}$. $\square$

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

## 3.6 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. In particular it involves composing* log *and* exp *in the order* $\log\exp$ *as opposed to the simpler* $\exp\log$*, which we used in our proof of Proposition (3.2).*

**(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).*

Chapter 4

# Basics of Lie Algebras

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 Basic structure theory

Let $L$ be a Lie $\mathbb{K}$-algebra. A *subalgebra* of $L$ is a $\mathbb{K}$-subspace $M$ that is closed under the bracket multiplication. In this case we write $M \le L$ and $L \ge M$.

A *Lie homomorphism* is a $\mathbb{K}$-linear transformation $\varphi\colon L \longrightarrow M$ with $\varphi([x,y]) = [\varphi(x), \varphi(y)]$ for all $x, y \in L$. The *kernel* of $\varphi$ is then the kernel of $\varphi$ as a linear transformation. In view of the First Isomorphism Theorem below, the image of $\varphi$ is sometimes referred to as the *quotient algebra* of $L$ by the kernel.

The kernel $I$ is a subalgebra, indeed it is an *ideal* of $L$—a subspace of $L$ with $[x,a] \in I$ for all $x \in L$ and $a \in I$ . We do not need to distinguish right ideals from left ideals, since $[a,x] = -[x,a] \in I$; all right and left ideals are immediately 2-sided ideals.

We have the standard *Isomorphism Theorems*:

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

(a) (FIRST ISOMORPHISM THEOREM) *If $\varphi\colon L \longrightarrow M$ is a Lie homomorphism with kernel $I$, then the image algebra $\varphi(L)$ is canonically isomorphic via*

*$\varphi(a) \mapsto a + I$ to the quotient algebra $L/I$ provided with the Lie bracket $[a+I, b+I] = [a,b] + I$.*

(b) (Second Isomorphism Theorem) *Let $I$ be a subalgebra and $J$ an ideal of $L$. Then $I+J$ is a subalgebra, $I \cap J$ is an ideal of $I$ and $(I+J)/J \simeq I/(I \cap J)$.*

(c) (Third Isomorphism Theorem) *If $I$ is an ideal of $L$ contained in the ideal $K$, then $L/K$ is isomorphic to $(L/I)\big/(K/I)$. In particular, there is a bijection between the set of ideals of $L/I$ and the set of ideals of $L$ that contain $I$.* □

By definition, the subspace $I$ is an ideal precisely when it is invariant under all inner derivations $\mathrm{ad}_x$. The ideal $I$ is additionally *characteristic* in $L$ if it is invariant under all derivations of $L$, not just the inner derivations.

**(4.2).** Lemma. *Let $J$ be an ideal of $L$ and $I$ a characteristic ideal of $J$.*

(a) *$I$ is an ideal of $L$.*

(b) *If $J$ is a characteristic ideal of $L$, then $I$ is a characteristic ideal of $L$.* □

As in the Second Isomorphism Theorem, old ideals can be used to construct new ones.

**(4.3).** 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 $L$.* □

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 $b \in B$.

**(4.4).** 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 $D([a, b]) = [D(a), b] + [a, D(b)]$. □

simple
abelian algebra; largest abelian quotient $L/[L, L]$.

derived series; solvable algebras; length
Z is kernel of ad
lower central series; nilpotent algebras; class
$L^1 = L$, $L^{n+1} = [L^n, L] = [L, L^n]$
$L^{(0)} = L$, $L^{(n+1)} = [L^{(n)}, L^{(n)}]$
upper central series?
Example: $\mathfrak{b} = \mathfrak{d} \oplus \mathfrak{n}$

**(4.5).** 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 $L^{m+1}$ and the induction step coming from the Jacobi Identity:

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

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

**(4.6).** 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.* □

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

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

(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.* □

It is noteworthy that the last part of the previous lemma does not have a counterpart here; 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 Proposition (4.12) below we will introduce an additional necessary and sufficient condition for such extensions to be nilpotent.

If $L$ has a unique maximal nilpotent ideal, then it is the *nilpotent radical* of $L$. Similarly if $L$ has a unique maximal solvable ideal, then it is the *radical* or *solvable radical* of $L$. The nilpotent radical is of course 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 (4.6)(d) the quotient of $L$ by its radical is then always semisimple.

## 4.2 Basic representation theory

| |
|---|
| repn; degree; dimension<br>modules; extrinsic and intrinsic; iso and equiv |
| "universal" algebra:<br>submodule; quotient;<br>simple; irreducible; trivial; indecomposable<br>cr; semisimple module<br>composition series<br>Jordan-Hölder<br>module duals |
| adjoint module and consequence for Lie structure |

**(4.8).** THEOREM. (KRULL-SCHMIDT THEOREM) *If $V = \bigoplus_{i \in I} V_i = \bigoplus_{j \in J} V^j$ are decompositions of the $L$-module $V$ into indecomposable summands, then there is a bijection $\sigma\colon I \longrightarrow J$ with $V_i$ and $V^{\sigma(i)}$ isomorphic for all $i \in I$.* □

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

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

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

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

**(4.10).** THEOREM. *For the finite dimensional and completely reducible module $V$, let $\mathcal{I}$ be a set of representatives for the isomorphism classes of irreducible submodules of $V$ and let $V_i$ be the sum of all irreducible submodules isomorphic to $i \in I$. Then $V = \bigoplus_{i \in I} V_i$.* □

**(4.11).** THEOREM. (SCHUR'S LEMMA) *Let $V$ be a finite dimensional, irreducible $L$-module over the algebraically closed field $\mathbb{K}$. Then the scalars are the only endomorphisms of $V$ that commute with the action of $L$.* □

## 4.3 Further structure and representation

In general an extension of a nilpotent algebra by a nilpotent algebra need not be nilpotent. We do get a nilpotent algebra if we have an additional *Engel condition*, requiring the vanishing of an appropriate iterated commutator. Define $[A; B, n]$ by $[A; B, 1] = [A, B]$ and $[A; B, k+1] = [[A; B, k], B]$.

**(4.12).** 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 we consider the converse. We first claim that for all positive $n$ and $r$

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

We prove this 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. □

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.

**(4.13).** LEMMA. *Let Lie 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 $[i, j] + [m, j] + [i, n] = [i, j] + [m, j] - [n, i] \in I$.* □

**(4.14).** 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. □

We emphasize two cases.

**(4.15).** 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 second example suggests some notation. Let $\varphi\colon L \longrightarrow \mathrm{End}^-_{\mathbb{K}}(V)$ be a Lie representation of $L$. For $x \in L$, we may write $\mathrm{ad}^V_x$ 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.

**(4.16).** PROPOSITION. *Let $\delta$ be a derivation of $L$. 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.

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

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

## 4.4 Problems

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

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

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

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

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

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

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

Chapter 5

# Nilpotent representations

## 5.1 Engel's Theorem and Cartan subalgebras

**(5.1).** 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 (4.15)(ii), we calculate within the semidirect product of $V$ by $L$. By Proposition (4.16) with $y \in L$, $\delta = \mathrm{ad}_x$, $a = 0$, 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 $\mathrm{ad}_x^i(y) = 0$ for $i \geq l$ 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. □

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.

**(5.2).** Proposition. *If $\sigma$ is an nil irreducible representation of the nilpotent Lie algebra $N$, then $\sigma$ is the trivial* 1*-dimensional representation.*

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

For $k > 1$ let $x \in L^{k-1}$. 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 $y \in L$ and $w \in W$

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

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

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

**(5.4).** THEOREM. (ENGEL'S THEOREM) *If the adjoint representation of the finite dimensional Lie algebra $N$ is nil, then $N$ is nilpotent.*

PROOF. We prove this by induction on $\dim_{\mathbb{K}}(N)$ with the result clearly true in dimensions 0 and 1. Assume $N \neq 0$.

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

By Corollary (5.3) there is a 1-dimensional submodule $P/I$ for the nil action of nilpotent $I$ on $N/I$. 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 $N$ in which $I$ is an ideal of codimension 1. By maximality of $I$, $N = P$.

We now have $N = M \oplus I$ with $M = \mathbb{K}x \simeq N/I$ an abelian algebra and $I$ a nilpotent ideal. Furthermore by hypothesis $\mathrm{ad}_x^m = 0$ for some $k$, hence $[I; M, m] = [I; x, m] = 0$ . By Proposition (4.12), the algebra $N$ 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 $A = \mathrm{N}_L(A)$.

**(5.5).** 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. □

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

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

*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 $a \in \mathrm{N}_L(L_{x,0})$. Then $[x,a] \in L_{x,0}$, so $\mathrm{ad}_x^k([x,a]) = 0$ for $k = k_{x,0,[x,a]}$. But then $\mathrm{ad}_x^{k+1}(a) = \mathrm{ad}_x^k([x,a]) = 0$, hence $a \in L_{x,0}$ 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 $w \in L_{w,0}$.

**(5.7).** 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.

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

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

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

$$\gamma_\alpha(z) = z^{n-r} + \sum_{j=0}^{n-r-1} g_j(\alpha) z^j$$

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 $\alpha_1, \ldots, \alpha_{r+1}$ of $\mathbb{K}$ with $g_0(\alpha_k) \neq 0$ for $1 \le k \le r+1$. In particular $L_{w+\alpha_k h,0} \le H$ for each $k$. As $w$ is regular, this forces $L_{w+\alpha_k h,0} = H$, which is to say $\varphi_{\alpha_k}(z) = z^r$ for $1 \le k \le r+1$. But then each of the polynomials $f_i(x)$, for $1 \le i < r$, vanishes at $\alpha_1, \ldots, \alpha_{r+1}$. 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 (5.4), $H$ is nilpotent as desired. □

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 since they are all essentially equivalent. We shall address conjugacy of Cartan subalgebras of finite dimensional semisimple algebras in Corollary (8.36).

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

**(5.8).** 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, $H$ is equal to*

$$L_{H,0} = \{\, x \in L \mid \mathrm{ad}_h^k(x) = 0 \textit{ for all } h \in H \textit{ and some } k = k_{h,0,x} \in \mathbb{N} \,\}\,,$$

*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 (5.3), there is a trivial $H$-submodule $P/H$ of dimension 1 in $L_0(H)/H$. For $x \in P \setminus H$, we have $[x, H] \leq H$. That is, $x$ is in the normalizer of $H$ but not in $H$. □

## 5.2 Weight spaces and vectors

**(5.9).** 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*

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

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 $\max_v(k_{x,\lambda,v}) \le n-1$, so

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

By Proposition (5.1), 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 $V_{x,\lambda(x)}$:

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

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 $L$,

$$V_{L,\lambda} = V_\lambda = \{\, v \in V \mid (x-\lambda(x)1)^k v = 0 \text{ for all } x \in L \text{ and some } k = k_{x,\lambda,v} \in \mathbb{N} \,\}$$

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

For every nonzero Lie algebra, the trivial representation is the *trivial weight* or *zero weight*. We have already encountered a weight space in Proposition (5.8), where the Cartan subalgebra $H$ was characterized among all nilpotent subalgebras of $L$ by being equal to its corresponding weight space $L_{H,0}$.

A nonzero weight of $L$ is a *root*.

**(5.10).** 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 weight space*

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

*is a submodule; and $V$ is the direct sum of its 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 weight spaces $V_\mu$ for some weight

[1] It should be noted that in certain places it is the eigenspaces $V^w_{L,\lambda}$ that are termed weight spaces. Perhaps the subspaces $V_{L,\lambda}$ might be called generalized weight spaces.

$\mu$ of $N$. Let the submodule $V(\mu)$ be the sum of those indecomposable summands with weight $\mu$. The previous theorem gives

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

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 $0 \neq v \in V_\lambda$. We claim:

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

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 $v_\mu \in V_\lambda \cap V(\mu) \leq V_\lambda \cap V_\mu$.

Assume $v_\mu \neq 0$. 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 $u = (x - \lambda(x)1)^{k-1} v_\mu \neq 0$, so that $(x - \lambda(x))u = 0$; that is, $xu = \lambda(x)u$. As $V(\mu)$ is a submodule, $u \in V(\mu) \leq V_\mu$; 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 $u \neq 0$ 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 $V_\lambda \leq V(\lambda) \leq V_\lambda$, hence

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

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

## 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.

**(5.11).** 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*

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

*where the weight space $L_{N,\lambda}$ for $\lambda \in (N/[N,N])^*$ is taken to be 0 when $\lambda$ is not a weight. Furthermore*

$$[L^w_{N,\alpha}, L^w_{N,\beta}] \leq L^w_{N,\alpha+\beta}\,.$$

PROOF. Let $x \in N$, $y \in L_\alpha$, and $z \in L_\beta$. Then, for $n = 2\dim_{\mathbb{K}}(N)$, by Proposition (4.16) and Theorem (5.10)

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

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

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

The most important case is that where $N = H$ is a Cartan subalgebra of $L$. Theorem (5.10) 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 weight spaces for a Cartan subalgebra $H$.

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 $H$ 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 48 and see Corollary (8.36) on semisimple algebras.

Recall that a root of $L$ is a nonzero weight. For $\alpha$ and $\beta$ roots of $L$, 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.

We have a first application of this concept.

**(5.12).** 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. Let

$$\beta - s\alpha\,, \dots, \beta\,, \dots, \beta + t\alpha$$

be the $\alpha$-string through $\beta$ and $M$ the corresponding subspace

$$M = L_{\beta - s\alpha} \oplus \cdots \oplus L_\beta \oplus \cdots \oplus L_{\beta + t\alpha}\,.$$

By Theorem (5.11) we have $[M, L_{-\alpha}] \leq M$ and $[M, L_\alpha] \leq M$.

Let $y \in L_\alpha$ and $z \in L_{-\alpha}$, and set $x = [y, z]$. As $y$ and $z$ normalize $M$, so does $x$. We have

$$\mathrm{tr}(\mathrm{ad}_x\,|_M) = \sum_{i=-s}^{t} d_i(\beta + i\alpha)(x)\,,$$

where $d_i = \dim_{\mathbb{K}}(L_{\beta+i\alpha})$. But $\mathrm{ad}_x = [\mathrm{ad}_y, \mathrm{ad}_z]$ and so it has trace 0. Therefore

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

hence

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

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

## 5.4 Problems

**(5.13).** PROBLEM. *Prove that any subalgebra of the Lie algebra $L$ that contains $L_{x,0}$ is self-normalizing.*

Chapter 6

# Killing forms and semisimple Lie algebras

## 6.1 Killing forms

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

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

where we recall our convention that $\operatorname{ad}_x^V$ is the image of $x \in L$ in $\operatorname{End}_{\mathbb{K}}(V)$. 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

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

**(6.1).** 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 $I^\perp = \{\, x \in L \mid \kappa_L^V(x, y) = 0\,, \text{ for all } y \in I \,\}$ is also an ideal of $L$.*

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)

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

(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. □

**(6.2).** 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 $I \le \mathrm{Rad}(\kappa_L^{L/I})$. The rest of the corollary follows easily. □

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 $n \ge 2$.

**(6.3).** 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

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

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 $\beta \in \Phi$. By the definition of roots, $\beta|_H \neq 0$ but $\beta|_{[H,H]} = 0$. Therefore there is an $\alpha \in \Phi_0$ with $\beta|_{[L_\alpha,L_{-\alpha}]} \neq 0$. Furthermore $\alpha$ is not the zero weight as again $\beta|_{[H,H]} = 0$. Thus by Proposition (5.12) there is a rational number $r_{\beta,\alpha}$ with

$$\beta|_{[L_\alpha,L_{-\alpha}]} = r_{\beta,\alpha}\alpha|_{[L_\alpha,L_{-\alpha}]} .$$

Choose an $x \in [L_\alpha, L_{-\alpha}]$ with $\beta(x) \neq 0$, hence $r_{\beta,\alpha} \neq 0$ and $\alpha(x) \neq 0$. Then

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

which is not equal to 0, as not all $r_{\lambda,\alpha}$ are 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. □

**(6.4).** 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 (6.2) 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 (4.6). □

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.

**(6.5).** 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 $\kappa$ be the Killing form and $R = \mathrm{Rad}(\kappa)$, an ideal by Proposition (6.1). But $\kappa|_{R\times R} = \kappa_R^R$ is identically 0, so $R$ is solvable by Cartan's Solvability Criterion (6.4). If $L$ is semisimple, then $R = 0$ and $\kappa$ 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. Therefore abelian $I$ is in $\mathrm{Rad}(\kappa_L^I)$, and also $I \leq \mathrm{Rad}(\kappa_L^{L/I})$ by Corollary (6.2). Hence by Proposition (6.1)

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

and $\kappa$ is degenerate. □

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

**(6.6).** 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 (6.1), and the restriction of $\kappa$ to it is identically 0. Therefore by Cartan's Solvablility Criterion (6.4) 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 (4.9), $L$ 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 subideal. Since no summand is trivial, each of these projections is onto the zero ideal; so $I$ itself is zero. Therefore $L$ is semisimple. □

## 6.2 Semisimple algebras I: $\mathfrak{sl}_2(\mathbb{K})$ subalgebras

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) the reduction of the classification to that of root systems;

(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 section we handle a large potion of the first part.

We first set some notation to be used throughout this section. In particular $L\,(\neq 0)$ will be a finite dimensional, semisimple Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic 0.

By Theorem (5.7) we may choose a Cartan subalgebra $H$ in $L$. By Proposition (5.8) we have $H = L_{H,0} = L_0$, the zero weight space. Let $\Phi$ be the set of all roots for $H$ on $L$, a finite set by Theorem (5.10). The set of all weights is $\Phi_0 = \{0\} \cup \Phi$.

For each $\lambda \in \Phi$, we have the weight space $L_\lambda = L_{H,\lambda}$, giving the Cartan decomposition

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

Since $L$ is nonzero and semisimple, the nilpotent Cartan subalgebra $H = L_0$ is proper in $L$, hence the root set $\Phi$ is nonempty.

The Killing form $\kappa = \kappa^L = \kappa_L = \kappa_L^L$ is nondegenerate by Cartan's Semisimplicity Criterion (6.5).

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

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

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

(c) *If* $0 \neq x \in L_\alpha$, *then* $\kappa(x, L_{-\alpha}) \neq 0$. *Especially,* $\alpha \in \Phi$ *implies* $-\alpha \in \Phi$.

PROOF. (a) Recall that for all weights $\mu, \nu$ we have $[L_\mu, L_\nu] \leq L_{\mu+\nu}$ by Theorem (5.11), and this extends to all $\lambda, \mu \in (H/[H,H])^*$ when we define $L_\lambda$ to be 0 whenever $\lambda$ is not a root.

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 $\mathrm{tr}(\mathrm{ad}_x\, \mathrm{ad}_y)$ is 0 if $\alpha+\beta$ is not equal to 0.

(b) By (a) $H^\perp = L_0^\perp \geq \bigoplus_{\lambda\in\Phi} L_\lambda$. Therefore $H^\perp \cap H \leq \mathrm{Rad}(\kappa) = 0$, and $\kappa|_{H\times H}$ is nondegenerate. Then $L = H \oplus \bigoplus_{\lambda\in\Phi} L_\lambda$ yields $H^\perp = \bigoplus_{\lambda\in\Phi} L_\lambda$.

(c) If $\kappa(x, L_{-\alpha}) = 0$, then by (a) we have $x \in \mathrm{Rad}(\kappa) = 0$. Therefore if $\alpha$ is a root, then $\kappa(L_\alpha, L_{-\alpha}) \neq 0$, hence $-\alpha$ is also a root. □

**(6.8).** 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 (6.7)(b). Therefore $H$ is abelian. □

**(6.9).** THEOREM. *We have $L = L^w$. That is, for every $\lambda \in \Phi_0$ the generalized $H$-eigenspace $L_\lambda$ is equal to the eigenspace $L^w_\lambda$.*

PROOF. By Theorem (6.6) semisimple $L$ is the direct sum of simple ideals, so we need only prove this for simple $L$. By the previous theorem $H = L^w_{H,0}$, so by Theorem (5.11) the subspace

$$H^w \oplus \bigoplus_{\alpha \in \Phi} L^w_\alpha = H \oplus \bigoplus_{\alpha \in \Phi} L^w_\alpha$$

is a nonzero subalgebra of simple $L$. That is

$$L = H^w \oplus \bigoplus_{\alpha \in \Phi} L^w_\alpha = L^w \,. \qquad \square$$

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 $t_{-\mu} = -t_\mu$.

**(6.10).** PROPOSITION.

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

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

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

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 $H = L_0$. Thus for a basis of $L$ consisting of bases for the various $L_\lambda$ (ordered appropriately using Theorem (5.10)) 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 $h \in \mathrm{Rad}(\kappa) = 0$. 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 (5.12) yields

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

and $\kappa(t_\alpha, t_\alpha) \neq 0$.

(c) By the previous theorem $\mathbb{K}x$ is a 1-dimensional $H$-submodule of $L_\alpha$. For all $h \in H$ and $y \in L_{-\alpha}$

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

We thus have $\kappa(h, [x,y] - \kappa(x,y)t_\alpha) = 0$ for all $h$. By the nondegeneracy of $\kappa$ on $H$, this gives $[x,y] = \kappa(x,y)t_\alpha \leq \mathbb{K}t_\alpha$. By Proposition (6.7)(c) we may choose $y \in L_{-\alpha}$ with $\kappa(x,y) \neq 0$, hence $[x, L_{-\alpha}] = \mathbb{K}t_\alpha$. □

Define $h_\alpha = \frac{2}{\kappa(t_\alpha,t_\alpha)} t_\alpha$, possible by Proposition (6.10)(c).

**(6.11).** THEOREM. *For each $\alpha \in \Phi$ and each $0 \neq x \in L_\alpha$, there is a $y \in L_{-\alpha}$ with*

$$\mathbb{K}x \oplus \mathbb{K}y \oplus \mathbb{K}t_\alpha = \mathbb{K}x \oplus \mathbb{K}y \oplus \mathbb{K}h_\alpha$$

*a subalgebra isomorphic to $\mathfrak{sl}_2(\mathbb{K})$.*

PROOF. For any nonzero $x \in L_\alpha$, the previous proposition allows us to choose a $y \in L_{-\alpha}$ with $\kappa(x,y) = \frac{2}{\kappa(t_\alpha,t_\alpha)}$. Then

$$\begin{aligned}[x,y] &= \kappa(x,y)t_\alpha = \frac{2}{\kappa(t_\alpha,t_\alpha)}t_\alpha = h_\alpha\,;\\ [h_\alpha,x] &= \frac{2}{\kappa(t_\alpha,t_\alpha)}[t_\alpha,x]\\ &= \frac{2}{\kappa(t_\alpha,t_\alpha)}\alpha(t_\alpha)x = \frac{2}{\kappa(t_\alpha,t_\alpha)}\kappa(t_\alpha,t_\alpha)x = 2x\,;\\ [h_\alpha,y] &= \frac{2}{\kappa(t_\alpha,t_\alpha)}[t_\alpha,y] = \frac{-2}{\kappa(t_\alpha,t_\alpha)}[t_{-\alpha},y] = \frac{-2}{\kappa(t_\alpha,t_\alpha)}(-\alpha)(t_{-\alpha})y\\ &= \frac{-2}{\kappa(t_\alpha,t_\alpha)}\alpha(t_\alpha)y = \frac{-2}{\kappa(t_\alpha,t_\alpha)}\kappa(t_\alpha,t_\alpha)y = -2y\,.\end{aligned}$$

Therefore

$$\mathbb{K}x \oplus \mathbb{K}y \oplus \mathbb{K}t_\alpha = \mathbb{K}x \oplus \mathbb{K}y \oplus \mathbb{K}h_\alpha$$

is a subalgebra, and by Section 2.3 it is a copy of $\mathfrak{sl}_2(\mathbb{K})$. □

**(6.12).** COROLLARY. *Let $L = H \oplus \bigoplus_{\alpha\in\Phi} L_\alpha$ be the Cartan decomposition for $L$ and for each $\alpha \in \Phi$ choose a basis $\{\, e_{\alpha,j} \mid 1 \leq j \leq \dim_{\mathbb{K}}(L_\alpha)\,\}$ for $L_\alpha$. For each $\alpha \in \Phi$ and $1 \leq j \leq \dim_{\mathbb{K}}(L_\alpha)$ there is a subalgebra $S_{\alpha,j}$ in $L$ that is isomorphic to $\mathfrak{sl}_2(\mathbb{K})$ and has $e_{\alpha,j} \in S_{\alpha,j}$. Furthermore $L = \sum_{\alpha,j} S_{\alpha,j}$.* □

## 6.3 Problems

**(6.13).** 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 7

# Representations of $\mathfrak{sl}_2(\mathbb{K})$

We have seen at the end of the last chapter that a finite dimensional semisimple Lie algebra $L$ over an algebraically closed field $\mathbb{K}$ of characteristic 0 is sewed together out of copies of $\mathfrak{sl}_2(\mathbb{K})$. We could have proceeded with the program outlined at the beginning of Section 6.2 toward a refined description of the Cartan decomposition, leading to the introduction of root systems. But the relevant calculations essentially come from the structure of $L$ as an $\mathfrak{sl}_2(\mathbb{K})$-module. Indeed we have already made one such calculation. Specifically, arguments involving $\alpha$-strings usually depend upon the finite dimensional representation theory of $\mathfrak{sl}_2(\mathbb{K})$. So Proposition (5.12) is actually a consequence of the fact that all finite dimensional $\mathfrak{sl}_2(\mathbb{K})$-representations can be realized over the rationals. (See Theorem (7.28)(a).)

Accordingly, in this chapter we take some time off to describe the representation theory of $\mathfrak{sl}_2(\mathbb{K})$ in a manner more detailed than actually needed for the program. (For the semisimple classification, we only need the much easier Theorem (7.22).)

In fact, $\mathfrak{sl}_2(\mathbb{K})$ is the only semisimple Lie algebra whose irreducible representations have been completely cataloged, but we do not do it in its entirety. For that, one should consult the excellent book [Maz10], which is the motivation for much in this chapter.

The irreducible, finite dimensional $\mathfrak{sl}_2(\mathbb{K})$-modules can be described quickly (and we have already seen them on page 18), but we shall also pursue certain (possibly) infinite dimensional modules. For the module properties considered we will be guided by the desire to include all finite dimensional modules, so their properties will motivate our definitions—specifically the presence of weight vectors.

Throughout this chapter, we let $\mathbb{K}$ be an algebraically closed field of characteristic 0.

## 7.1 Weight modules

Within the Lie algebra $\mathfrak{sl}_2(\mathbb{K})$, we have focused on three elements

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

which provide the presentation

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

for the 3-dimensional algebra $\mathfrak{sl}_2(\mathbb{K}) = \mathbb{K}h \oplus \mathbb{K}e \oplus \mathbb{K}f$. This is the Cartan decomposition of $L = \mathfrak{sl}_2(\mathbb{K})$ for the Cartan subalgebra $H = \mathbb{K}h = L_0$ with weight spaces $\mathbb{K}e = L_2$ and $\mathbb{K}f = L_{-2}$.

The following lemma tells us these 1-spaces can be characterized extrinsically (as Cartan subalgebra and its weight spaces), so this presentation is in a sense canonical.

**(7.1).** LEMMA. *The algebra $\mathfrak{sl}_2(\mathbb{K})$ is simple. Every Cartan subalgebra of $\mathfrak{sl}_2(\mathbb{K})$ is equivalent under $\mathrm{Aut}(\mathfrak{sl}_2(\mathbb{K}))$ to $\mathbb{K}h$, and the only subalgebras containing $\mathbb{K}h$ are*

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

PROOF.[1] With respect to the basis $\{h, e, f\}$ the Gram matrix of the Killing form $\kappa$ is

$$\begin{pmatrix} 8 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{pmatrix}$$

so $\kappa$ is nondegenerate as $\mathrm{char}\mathbb{K} \neq 2$. Therefore by Cartan's Semisimplicity Criterion (6.5) the algebra $\mathfrak{sl}_2(\mathbb{K})$ of dimension 3 is semisimple. Were it not simple, it would have an abelian ideal of dimension 1, which is not the case.

A Cartan subalgebra is thus abelian (Theorem (6.8)) and so is contained in the normalizer of the subalgebra generated by each of its elements. By Jordan Canonical Form, every nonzero element of $\mathfrak{sl}_2(\mathbb{K})$ is conjugate under $\mathrm{GL}_2(\mathbb{K}) \leq \mathrm{Aut}(\mathfrak{sl}_2(\mathbb{K}))$ to one of

$$\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} b & 0 \\ 0 & -b \end{pmatrix}$$

for nonzero $a, b \in \mathbb{K}$. The normalizer of the subalgebra with the first shape contains elements of the second type and is not nilpotent. The abelian subalgebra $\mathbb{K}h$ of all matrices of the second type is self-normalizing and so is a Cartan subalgebra by Proposition (5.8).

The rest of the Lemma follows by Problem (2.7) or by calculation with respect to the Cartan decomposition given by $\mathbb{K}h$. □

The Cartan subalgebra $\mathbb{K}h$ of $\mathfrak{sl}_2(\mathbb{K})$ has dimension 1, so the elements $\lambda$ of its dual space are described entirely by the element $\lambda(h) \in \mathbb{K}$ via $\lambda(kh) = k\lambda(h)$ for

[1] Exercise: This proof is overkill. Do the calculations needed to make it elementary—that is, free of reference to results like Proposition (5.8) and Theorems (6.5) and (6.8).

all $k \in \mathbb{K}$. We therefore abuse notation and terminology somewhat by setting $\lambda(h) = \lambda \in \mathbb{K}$ and saying that the element $\lambda$ of $\mathbb{K}$ is a weight when, more properly, it is the associated linear functional $\lambda\colon \mathbb{K}h \longrightarrow \mathbb{K}$ given by $kh \mapsto \lambda k$ that is the weight.

Let $V$ be an $\mathfrak{sl}_2(\mathbb{K})$-module. We let the images of $h$, $e$, and $f$ in $\mathrm{End}_{\mathbb{K}}^{-}(V)$ be, respectively, $\mathsf{H}^V$, $\mathsf{E}_+^V$, $\mathsf{E}_-^V$, usually abbreviated to $\mathsf{H}$, $\mathsf{E}_+$, $\mathsf{E}_-$. (In the notation introduced on page 42, these are $\mathrm{ad}_h^V$, $\mathrm{ad}_e^V$, and $\mathrm{ad}_f^V$.)

Recall that we write $V_\lambda^w$ for those $x \in V_\lambda$ with $[h,x] = \lambda(h)x$, such a nonzero $x$ being a weight vector for the weight $\lambda$. The space of weight vectors $V_\lambda^w$ is the actual $\mathsf{H}$-eigenspace, a subspace of the generalized $\mathsf{H}$-eigenspace $V_\lambda$. If finite dimensional $V_\lambda \neq 0$ then $V_\lambda^w \neq 0$ by Jordan Canonical Form.

The $\mathfrak{sl}_2(\mathbb{K})$-module $V$ is *torsion-free* if, for all nonzero $f(x) \in \mathbb{K}[x]$ and all nonzero $v \in V$, we have $f(\mathsf{H})v \neq 0$.

Torsion-free modules and weight vectors provide us with a basic dichotomy.

**(7.2).** PROPOSITION. *If $V$ is an irreducible $\mathfrak{sl}_2(\mathbb{K})$-module, then either $V$ is torsion-free or $V$ contains a weight vector.*

PROOF. Assume $V$ has torsion, so that there are $0 \neq v \in V$ and $0 \neq f(x) \in \mathbb{K}[x]$ with $f(\mathsf{H})v = 0$. Choose $f(x)$ to be monic and of minimal degree subject to this. Thus $f(x) = \prod_{i=1}^n (x - \alpha_i)$ for distinct $\alpha_i \in \mathbb{K}$. With $w = \prod_{i=1}^{n-1}(\mathsf{H} - \alpha_i)v$, we then have $w$ a weight vector for $\lambda = \alpha_n$. $\square$

Our goal in this chapter is to classify all irreducible modules containing a weight vector. As $v, \mathsf{H}v, \mathsf{H}^2v, \dots, \mathsf{H}^iv, \dots$ must be linearly dependent in finite dimensional $V$, this classification will include the classification of all irreducible finite dimensional $\mathfrak{sl}_2(\mathbb{K})$-modules.

**(7.3).** LEMMA. *Let $V$ be an $\mathfrak{sl}_2(\mathbb{K})$-module.*

(a) $\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon} - \mathsf{E}_{-\epsilon}\mathsf{E}_\epsilon = \epsilon\mathsf{H}$; $\mathsf{H}\mathsf{E}_\epsilon - \mathsf{E}_\epsilon\mathsf{H} = 2\epsilon\mathsf{E}_\epsilon$.

(b) $\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon} = \epsilon\mathsf{H} + \mathsf{E}_{-\epsilon}\mathsf{E}_\epsilon$; $\mathsf{H}\mathsf{E}_\epsilon = \mathsf{E}_\epsilon(\mathsf{H} + 2\epsilon)$; $\mathsf{E}_\epsilon\mathsf{H} = (\mathsf{H} - 2\epsilon)\mathsf{E}_\epsilon$.

(c) *If $f(x) \in \mathbb{K}[x]$ is a polynomial and $n \in \mathbb{N}$, then*

$$f(\mathsf{H})\mathsf{E}_\epsilon^n = \mathsf{E}_\epsilon^n f(\mathsf{H} + 2\epsilon n) \;\; \textit{and} \;\; \mathsf{E}_\epsilon^n f(\mathsf{H}) = f(\mathsf{H} - 2\epsilon n)\mathsf{E}_\epsilon^n\,.$$

(d) *If $v \in V_\lambda$ then $\mathsf{E}_\epsilon v \in V_{\lambda+2\epsilon}$.*

(e) *If $v \in V_\lambda^w$ then $\mathsf{E}_\epsilon v \in V_{\lambda+2\epsilon}^w$ and $\mathsf{E}_{-\epsilon}\mathsf{E}_\epsilon v \in V_\lambda^w$.*

PROOF. The first part consists of the equations demonstrating that $V$ is an $\mathfrak{sl}_2(\mathbb{K})$-module. The second part is then a rewritten version of the first, and the third part follows by induction. (Exercise.)

Let $v \in V_\lambda$. Then by the previous part

$$(\mathsf{H} - (\lambda + 2\epsilon))^k(\mathsf{E}_\epsilon v) = ((\mathsf{H} - (\lambda + 2\epsilon))^k\mathsf{E}_\epsilon)v = (\mathsf{E}_\epsilon(\mathsf{H} - \lambda)^k)v = \mathsf{E}_\epsilon((\mathsf{H} - \lambda)^k v)\,,$$

which is 0 for large enough $k$, hence $\mathsf{E}_\epsilon v \in V_{\lambda+2\epsilon}$. Further let $v \in V_\lambda^w$. Then the equalities holds and are 0 for $k = 1$, proving that $\mathsf{E}_\epsilon v \in V_{\lambda+2\epsilon}^w$, as claimed at the beginning of the last part. The end follows directly. □

Parts of the lemma also follow directly from Theorem (5.11), when we consider the weight spaces for the nilpotent subalgebra $\mathbb{K}h\,(= N)$ of the semidirect product algebra $\mathfrak{sl}_2(\mathbb{K}) \oplus V\,(= L)$, as discussed under Example (4.15)(b).

A *weight module* for $\mathfrak{sl}_2(\mathbb{K})$ is a module that is generated by weight vectors, which for an irreducible module is equivalent to containing a weight vector. By the previous result, the sum of all the spaces $V_\lambda^w$ of weight vectors is a submodule of $V$, so $V$ is a weight module if and only if $V_\lambda^w = V_\lambda$ for all $\lambda$. This in turn is equivalent to the statement that $\mathsf{H}^V = \mathsf{H}$ is "diagonal" or "semisimple" in its action on $V$. These remarks include the usual definitions of a weight module; see [Maz10, p. 59].

If $v$ is a $\lambda$-weight vector, then always

$$\lambda v = \mathsf{H}v = (\mathsf{E}_+\mathsf{E}_- - \mathsf{E}_-\mathsf{E}_+)v = \mathsf{E}_+\mathsf{E}_-v - \mathsf{E}_-\mathsf{E}_+v$$

where by the lemma $\mathsf{E}_+\mathsf{E}_-v$ and $\mathsf{E}_-\mathsf{E}_+v$ both belong to $V_\lambda^w$. We say that the weight vector $v$ is *coherent*[2] if there is are constants $\lambda_+$ and $\lambda_-$ in $\mathbb{K}$ with

$$\mathsf{E}_+\mathsf{E}_-v = \lambda_+ v \quad \text{and} \quad \mathsf{E}_-\mathsf{E}_+v = \lambda_- v\,,$$

where necessarily $\lambda_+ - \lambda_- = \lambda$.

**(7.4).** LEMMA. *If* $\dim_{\mathbb{K}}(V_\lambda^w) = 1$*, then each weight vector* $v \in V_\lambda^w$ *is coherent.* □

We shall see (in Proposition (7.24)) that weight vectors in irreducible $\mathfrak{sl}_2(\mathbb{K})$-modules are always coherent. Irreducible modules are always cyclic, so our classification results will come from careful study of cyclic modules generated by coherent weight vectors.

Two special types of coherent weight vectors are of particular note. In the $\mathfrak{sl}_2(\mathbb{K})$-module $V$, we will call the nonzero vector $v$ a *highest weight vector* for the weight $\lambda \in \mathbb{K}$ provided $v$ a weight vector in $V_\lambda^w$ for $\mathsf{H}$ and additionally $\mathsf{E}_+(v) = 0$, so that $\mathsf{E}_-\mathsf{E}_+(v) = 0$, $\lambda_- = 0$, and $\lambda_+ = \lambda$. Equivalently, $\mathbb{K}v \le V_\lambda$ is not just a $\mathbb{K}h$-submodule but is also a submodule for the Borel subalgebra $B^+ = \mathbb{K}h \oplus \mathbb{K}e$. Similarly nonzero $v$ is a *lowest weight vector* for the weight $\lambda \in K$ provided $\mathbb{K}v \le V_\lambda$ is a $B^-$-submodule; that is, $v$ is weight vector of $V_\lambda^w$ for $\mathsf{H}$ and also $\mathsf{E}_-(v) = 0$, hence $\mathsf{E}_+\mathsf{E}_-(v) = 0$, $\lambda_+ = 0$, and $\lambda_- = \lambda$.

**(7.5).** THEOREM. *Let $V$ be a finite dimensional module for $\mathfrak{sl}_2(\mathbb{K})$. Then $V$ contains highest weight vectors and lowest weight vectors.*

PROOF. Suppose $V_\lambda \neq 0$, and choose $0 \neq v \in V_\lambda^w$ (possible, as mentioned above). By Lemma (7.3)(d) and finite dimensionality, there are integers $t$ with

[2] Unlike the other terminology in this chapter, this is not standard. But it is somewhat related to a concept of coherence within the representation theory.

$0 \neq \mathsf{E}_+^t v \in V^w_{\lambda+2t}$ but $0 = \mathsf{E}_+^{t+1} v$, and $s$ with $0 \neq \mathsf{E}_-^s v \in V^w_{\lambda-2s}$ but $0 = \mathsf{E}_-^{s+1} v$. Thus $v_+ = \mathsf{E}_+^t v$ is a highest weight vector and $v_- = \mathsf{E}_-^s v$ is a lowest weight vector. □

An elementary arithmetic lemma will be of help.

**(7.6).** LEMMA. *Let $\lambda, \lambda_+, \lambda_- \in \mathbb{K}$ with $\lambda_+ - \lambda_- = \lambda$. Set $\epsilon \in \{\pm\} = \{\pm 1\}$, and consider the two sequences $a_\epsilon(i)$, for $i \in \mathbb{Z}$, where $a_\epsilon(i) = 1$ for $-\epsilon i \in \mathbb{N}$. The following are equivalent:*

(1) $a_\epsilon(i) = (i-\epsilon)(\lambda - i) + \lambda_\epsilon = i(\lambda - i + \epsilon) + \lambda_{-\epsilon}$ *for all* $\epsilon i \in \mathbb{Z}^+$.

(2) $a_\epsilon(\epsilon) = \lambda_\epsilon$ *and* $a_\epsilon(i) - a_\epsilon(i-\epsilon) = \epsilon(\lambda - 2(i-\epsilon))$ *for all* $\epsilon i \in \mathbb{Z}^+$.

(3) $a_\epsilon(\epsilon) = \lambda_\epsilon$ *and* $a_-(i)a_+(i+1) - a_+(i)a_-(i-1) = \lambda - 2i$ *for all* $i \in \mathbb{Z}$.

PROOF. (Exercise.) □

**(7.7).** COROLLARY. *Let $\lambda = \lambda_+ - \lambda_-$ and $\epsilon \in \{\pm\} = \{\pm 1\}$. From the two half-infinite sequences*

$$a_\epsilon(i) = (i-\epsilon)(\lambda - i) + \lambda_\epsilon = i(\lambda - i + \epsilon) + \lambda_{-\epsilon} \text{ for all } \epsilon i \in \mathbb{Z}^+$$

*create the new doubly infinite sequence*

$$b(j) = a_-(j) \text{ for } j \in \mathbb{Z}^- \text{ and } b(j) = a_+(j+1) \text{ for } j \in \mathbb{N}\,.$$

*Then $b(j) = j(\lambda - j - 1) + \lambda_+$ for all $j \in \mathbb{Z}$.* □

We now investigate the structure of a cyclic submodule.

**(7.8).** THEOREM. *Let $v$ be a weight vector for the weight $\lambda$ in the $\mathfrak{sl}_2(\mathbb{K})$-module $V$. Set $v_0 = v$, and let $\epsilon \in \{\pm\} = \{\pm 1\}$.*

*For all $i \in \mathbb{Z}^+$ define*

$$v_{\epsilon i} = \mathsf{E}^i_{-\epsilon}(v) = \mathsf{E}^i_{-\epsilon}(v_0) = \mathsf{E}_{-\epsilon}(v_{\epsilon(i-1)})\,,$$

*which we may rewrite as:*

$$\mathsf{E}_\epsilon(v_i) = v_{i-\epsilon} \quad \text{for } \epsilon i \in -\mathbb{N}$$

*Then*

(a) $\mathsf{H}(v_i) = (\lambda - 2i)v_i$ *for all* $i \in \mathbb{Z}$;

(b) *Assume $v$ is coherent. Set $\mathsf{E}_\epsilon \mathsf{E}_{-\epsilon} v = \lambda_\epsilon v$, hence $\lambda_+ - \lambda_- = \lambda$. Then*

$$\mathsf{E}_\epsilon(v_i) = a_\epsilon(i) v_{i-\epsilon} \quad \text{for } \epsilon i \in \mathbb{Z}^+$$

*where $a_\epsilon(i) = (i-\epsilon)(\lambda - i) + \lambda_\epsilon$, as in Lemma (7.6).*

Proof. (a) We prove this by induction on $|i|$, the case $|i| = 0$ holding by definition. For an $i$ with $|i| > 0$, define $\epsilon \in \{\pm\} = \{\pm 1\}$ by $i = \epsilon|i|$. We use Lemma (7.3) and induction to calculate

$$\begin{aligned}
\mathsf{H}v_i &= \mathsf{H}\mathsf{E}_{-\epsilon}v_{i-\epsilon} \\
&= \mathsf{E}_{-\epsilon}(\mathsf{H} - 2\epsilon)v_{i-\epsilon} \\
&= \mathsf{E}_{-\epsilon}((\lambda - 2(i-\epsilon)) - 2\epsilon)v_{i-\epsilon} \\
&= (\lambda - 2i)\mathsf{E}_{-\epsilon}v_{i-\epsilon} \\
&= (\lambda - 2i)v_i\,.
\end{aligned}$$

(b) We proceed by induction on the positive integer $\epsilon i$. The case $\epsilon i = 1$ serves to define the two constants $a_\epsilon(\epsilon) = \lambda_\epsilon$, where $\lambda_\epsilon - \lambda_{-\epsilon} = \epsilon\lambda$ by coherence of $v$. Assume $\epsilon i \geq 2$. Then

$$\begin{aligned}
\mathsf{E}_\epsilon(v_i) &= \mathsf{E}_\epsilon\mathsf{E}_{-\epsilon}v_{i-\epsilon} \\
&= (\epsilon\mathsf{H} + \mathsf{E}_{-\epsilon}\mathsf{E}_\epsilon)v_{i-\epsilon} && \text{by Lemma (7.3)} \\
&= \epsilon\mathsf{H}v_{i-\epsilon} + \mathsf{E}_{-\epsilon}\mathsf{E}_\epsilon v_{i-\epsilon} \\
&= \epsilon(\lambda - 2(i-\epsilon))v_{i-\epsilon} + \mathsf{E}_{-\epsilon}a_\epsilon(i-\epsilon)v_{i-2\epsilon} && \text{by (a) and induction} \\
&= \epsilon(\lambda - 2(i-\epsilon))v_{i-\epsilon} + a_\epsilon(i-\epsilon)v_{i-\epsilon} && \text{as } -\epsilon(i-2\epsilon) \in -\mathbb{N} \\
&= a_\epsilon(i)v_{i-\epsilon} && \text{by Lemma (7.6).}
\end{aligned}$$

This completes the induction and proof. □

## 7.2 Verma modules

Theorem (7.8) motivates the following construction:

**(7.9).** Definition. *Let $\lambda, \lambda_+, \lambda_- \in \mathbb{K}$ with $\lambda_+ - \lambda_- = \lambda$. Set $M(\lambda, \lambda_+, \lambda_-) = \bigoplus_{i\in\mathbb{Z}} \mathbb{K}v_i$. Define the linear transformations $\mathsf{H}$, $\mathsf{E}_+$, and $\mathsf{E}_-$ on $M(\lambda, \lambda_+, \lambda_-)$ by*

$$\begin{aligned}
\mathsf{H}(v_i) &= (\lambda - 2i)v_i\,, \\
\mathsf{E}_\epsilon(v_i) &= a_\epsilon(i)v_{i-\epsilon}\,,
\end{aligned}$$

*for $i \in \mathbb{Z}$ and $\epsilon \in \{\pm\} = \{\pm 1\}$ with*

$$\begin{aligned}
a_\epsilon(i) &= (i-\epsilon)(\lambda - i) + \lambda_\epsilon && \text{for } \epsilon i \in \mathbb{Z}^+ \text{ and} \\
a_\epsilon(i) &= 1 && \text{for } \epsilon i \in -\mathbb{N}.
\end{aligned}$$

In this action and with respect to the basis $\{\, v_i \mid i \in \mathbb{Z}\,\}$, the linear transformation $\mathsf{H}$ is "diagonal," in the sense that it takes each 1-space $\mathbb{K}v_i$ to itself. Similarly $\mathsf{E}_+$ is "lower diagonal," always taking $\mathbb{K}v_i$ to $\mathbb{K}v_{i-1}$, and $\mathsf{E}_-$ is "upper diagonal," taking $\mathbb{K}v_i$ to $\mathbb{K}v_{i+1}$. The corresponding nonzero coefficients, the *transformation coefficients*, are

$$\mathsf{H}\colon v_i \xrightarrow{\lambda-2i} v_i \qquad \mathsf{E}_+\colon v_i \xrightarrow{a_+(i)} v_{i-1} \qquad \mathsf{E}_-\colon v_i \xrightarrow{a_-(i)} v_{i+1}$$

We exhibit the actions on $M(\lambda, \lambda_+, \lambda_-)$ pictorially as below. In the picture every basis vector appears twice—once on the line displaying the action of $\mathsf{E}_-$ and once on the line displaying the action of $\mathsf{E}_+$. Passage between the two lines gives the action of $\mathsf{H}$.

$$
\begin{tikzcd}
\mathsf{E}_- & {} \arrow[r, dashed] & v_{-2} \arrow[r, dashed, "a_-(-2)"] \arrow[d, "\lambda+4"] & v_{-1} \arrow[r, dashed, "\lambda_-"] \arrow[d, "\lambda+2"] & v_0 \arrow[r, "1"] \arrow[d, "\lambda"] & v_1 \arrow[r, "1"] \arrow[d, "\lambda-2"] & v_2 \arrow[r] \arrow[d, "\lambda-4"] & {} \\
\mathsf{E}_+ & {} & v_{-2} \arrow[l] & v_{-1} \arrow[l, "1"'] & v_0 \arrow[l, "1"'] & v_1 \arrow[l, dashed, "\lambda_+"'] & v_2 \arrow[l, dashed, "a_+(2)"'] & {} \arrow[l, dashed]
\end{tikzcd}
$$

**(7.10).** Theorem. *The maps*

$$h \mapsto \mathsf{H}\,, \quad e \mapsto \mathsf{E}_+\,, \quad f \mapsto \mathsf{E}_-$$

*give $M(\lambda, \lambda_+, \lambda_-)$, as defined in (7.9) above, the structure of a cyclic $\mathfrak{sl}_2(\mathbb{K})$-module generated by the coherent weight vector $v_0$ for the weight $\lambda$ and having $\mathsf{E}_\epsilon \mathsf{E}_{-\epsilon} v_0 = \lambda_\epsilon v_0$.*

Proof. Within the definition (7.9) we find $\mathsf{E}_\epsilon \mathsf{E}_{-\epsilon} v_0 = a_\epsilon(\epsilon) v_0 = \lambda_\epsilon v_0$.

To verify $[\mathsf{H}, \mathsf{E}_\epsilon] = \mathsf{H}\mathsf{E}_\epsilon - \mathsf{E}_\epsilon \mathsf{H}$ we check equality on the basis vectors $v_i$:

$$
\begin{aligned}
[\mathsf{H}, \mathsf{E}_\epsilon] v_i &= (\mathsf{H}\mathsf{E}_\epsilon - \mathsf{E}_\epsilon \mathsf{H}) v_i \\
&= \mathsf{H}\mathsf{E}_\epsilon v_i - \mathsf{E}_\epsilon \mathsf{H} v_i \\
&= \mathsf{H} a_\epsilon(i) v_{i-\epsilon} - \mathsf{E}_\epsilon (\lambda - 2i) v_i \\
&= a_\epsilon(i) \mathsf{H} v_{i-\epsilon} - (\lambda - 2i) \mathsf{E}_\epsilon v_i \\
&= a_\epsilon(i)(\lambda - 2(i-\epsilon)) v_{i-\epsilon} - (\lambda - 2i) a_\epsilon(i) v_{i-\epsilon} \\
&= \left((\lambda - 2(i-\epsilon)) - (\lambda - 2i)\right) a_\epsilon(i) v_{i-\epsilon} \\
&= 2\epsilon \mathsf{E}_\epsilon v_i\,;
\end{aligned}
$$

so $[\mathsf{H}, \mathsf{E}_\epsilon] = 2\epsilon \mathsf{E}_\epsilon$, as desired.

We must also verify $[\mathsf{E}_+, \mathsf{E}_-] = \mathsf{H}$:

$$
\begin{aligned}
[\mathsf{E}_+, \mathsf{E}_-] v_i &= \mathsf{E}_+ \mathsf{E}_- v_i - \mathsf{E}_- \mathsf{E}_+ v_i \\
&= \mathsf{E}_+ a_-(i) v_{i+1} - \mathsf{E}_- a_+(i) v_{i-1} \\
&= a_-(i) a_+(i+1) v_i - a_+(i) a_-(i-1) v_i \\
&= \left(a_-(i) a_+(i+1) - a_+(i) a_-(i-1)\right) v_i \\
&= (\lambda - 2i) v_i = \mathsf{H} v_i\,,
\end{aligned}
$$

where between the last two lines we have used Lemma (7.6). Accordingly $[\mathsf{E}_+, \mathsf{E}_-] = \mathsf{H}$, which together with the preceding paragraph proves that we have a representation and module. □

The module $M_+(\lambda, \lambda_+, \lambda_-)$ is a *generalized Verma module* for $\mathfrak{sl}_2(\mathbb{K})$ with weight $\lambda$. As an immediate consequence of the previous two results we have a universal property for generalized Verma modules:

**(7.11).** COROLLARY. *Let $v$ be a coherent weight vector for the weight $\lambda$ in the $\mathfrak{sl}_2(\mathbb{K})$-module $V$ with $\mathsf{E}^V_\epsilon \mathsf{E}^V_{-\epsilon} v = \lambda_\epsilon v$. Then the map $v_0 \mapsto v$ extends to a Lie module homomorphism taking the generalized Verma module $M(\lambda, \lambda_+, \lambda_-)$ onto the $\mathfrak{sl}_2(\mathbb{K})$-submodule of $V$ generated by $v$.* □

**(7.12).** COROLLARY.

(a) $V = M(\lambda, \lambda_+, \lambda_-) = \bigoplus_{\mu \in \lambda + 2\mathbb{Z}} V_\mu$ *with* $\dim_{\mathbb{K}}(V_\mu) = 1$ *for all* $\mu \in \lambda + 2\mathbb{Z}$.

(b) *Every weight vector in $V$ is coherent.*

(c) *Any $\mathsf{H}$-submodule, and so any $\mathfrak{sl}_2(\mathbb{K})$-submodule, of $V$ is a sum $\bigoplus_{\mu \in I} V_\mu$, for some subset $I \subseteq \lambda + 2\mathbb{Z}$.* □

As already mentioned, each of the operators $\mathsf{H}$ and $\mathsf{E}_\epsilon$ is "nearly diagonal" on the the basis $\{ v_i \mid i \in \mathbb{Z} \}$. If we replace various of the $v_i$ with nonzero scalar multiples, this will not change the near-diagonal structures, but it will change the values of the certain of the transition coefficients. A particular case is of interest.

**(7.13).** PROPOSITION. *For constants $0 \neq b, d \in \mathbb{K}$ with $bd = 1$, and an integer $j$, consider the new basis $\{ v'_i \mid i \in \mathbb{Z} \}$ for $M_+(\lambda, \lambda_+, \lambda_-)$ given by*

$$\begin{aligned} v'_i &= v_i && \text{if } i > j \\ &= bv_i && \text{if } i \leq j\,. \end{aligned}$$

*Then the transition parameters for this basis are equal to the transition parameters with respect to the original basis with only two exceptions:*

$$a'_-(j) = ba_-(j) \text{ and } a'_+(j+1) = da_+(j+1)\,.$$

PROOF. If $i > j+1$ or $i = j+1$ and $\epsilon = -$

$$\mathsf{E}_\epsilon(v'_i) = \mathsf{E}_\epsilon(v_i) = a_\epsilon(i) v_{i-\epsilon} = a_\epsilon(i) v'_{i-\epsilon}\,.$$

Similarly if $i < j$ or $i = j$ and $\epsilon = +$

$$\mathsf{E}_\epsilon(v'_i) = \mathsf{E}_\epsilon(bv_i) = a_\epsilon(i) b v_{i-\epsilon} = a_\epsilon(i) v'_{i-\epsilon}\,.$$

On the other hand,

$$\mathsf{E}_+(v'_{j+1}) = \mathsf{E}_+(v_{j+1}) = a_+(j+1) v_j = bda_+(j+1) v_j = da_+(j+1) v'_j$$

and

$$\mathsf{E}_-(v'_j) = \mathsf{E}_-(bv_j) = b\mathsf{E}_-(v_j) = ba_-(j) v_{j+1} = ba_-(j) v'_{j+1}\,. \qquad \square$$

**(7.14).** THEOREM. *For a given $\delta \in \{\pm 1\}$, if $a_\delta(\delta) = \lambda_\delta$ is not equal to $0$, then the map $\lambda_\delta^{-1} v_\delta \mapsto v''_0$ extends to an isomorphism of $M(\lambda, \lambda_+, \lambda_-)$ with $M(\lambda'', \lambda''_+, \lambda''_-)$ where $\lambda'' = \lambda + 2\delta$ and $\lambda''_{-\delta} = \lambda_\delta$.*

PROOF. For $M(\lambda, \lambda_+, \lambda_-)$ we start with

$$\begin{tikzcd}
\mathsf{E}_- & {} \arrow[r, dashed] & v_{-2} \arrow[r, dashed, "a_-(-2)"] \arrow[d, "\lambda+4"] & v_{-1} \arrow[r, dashed, "\lambda_-"] \arrow[d, "\lambda+2"] & v_0 \arrow[r, "1"] \arrow[d, "\lambda"] & v_1 \arrow[r, "1"] \arrow[d, "\lambda-2"] & v_2 \arrow[r] \arrow[d, "\lambda-4"] & {} \\
\mathsf{E}_+ & {} & v_{-2} \arrow[l] & v_{-1} \arrow[l, "1"'] & v_0 \arrow[l, "1"'] & v_1 \arrow[l, dashed, "\lambda_+"'] & v_2 \arrow[l, dashed, "a_+(2)"'] & {} \arrow[l, dashed]
\end{tikzcd}$$

Assume $\lambda_- \neq 0$, and in the proposition set $j = -1$, $b = \lambda_-^{-1}$, and $d = \lambda_-$. We then find

$$\begin{tikzcd}
\mathsf{E}_- & {} \arrow[r, dashed] & v'_{-2} \arrow[r, dashed, "a_-(-2)"] \arrow[d, "\lambda+4"] & v'_{-1} \arrow[r, "1"] \arrow[d, "\lambda+2"] & v'_0 \arrow[r, "1"] \arrow[d, "\lambda"] & v'_1 \arrow[r, "1"] \arrow[d, "\lambda-2"] & v'_2 \arrow[r] \arrow[d, "\lambda-4"] & {} \\
\mathsf{E}_+ & {} & v'_{-2} \arrow[l] & v'_{-1} \arrow[l, "1"'] & v'_0 \arrow[l, dashed, "\lambda_-"'] & v'_1 \arrow[l, dashed, "\lambda_+"'] & v'_2 \arrow[l, dashed, "a_+(2)"'] & {} \arrow[l, dashed]
\end{tikzcd}$$

Therefore the map $v'_i \mapsto v''_{i+1}$ gives an isomorphism of $M(\lambda, \lambda_+, \lambda_-)$ with the generalized Verma module $M(\lambda + 2, \lambda_-, a_-(-2))$.

On the other hand, if $\lambda_+ \neq 0$ then in the proposition we set $j = 0$, $b = \lambda_+$, and $d = \lambda_+^{-1}$ to reveal an isomorphism of $M(\lambda, \lambda_+, \lambda_-)$ with the generalized Verma module $M(\lambda - 2, a_+(2), \lambda_+)$. $\square$

**(7.15).** COROLLARY.

(a) *If $a_\epsilon(i)$ is nonzero for all pairs $(\epsilon, i)$, then $M(\lambda, \lambda_+, \lambda_-)$ is irreducible.*

(b) *If $M(\lambda, \lambda_+, \lambda_-)$ is reducible, then there is a $\mu \in \lambda + 2\mathbb{Z}$ with $M(\lambda, \lambda_+, \lambda_-)$ isomorphic to $M(\mu, \mu, 0)$ or to $M(\mu, 0, -\mu)$.*

PROOF. (a) By the theorem, under these circumstances $M(\lambda, \lambda_+, \lambda_-)$ is cyclically generated by every $\mathbb{K}v_i$; so by Corollary (7.12) there are no nonzero, proper submodules.

(b) By (a) if $M(\lambda, \lambda_+, \lambda_-)$ is reducible, then there is at least one pair $(\epsilon, i)$ with $a_\epsilon(i) = 0$. Choose the smallest $|i| = \epsilon i$ for which this is true. Then $a_\epsilon(\epsilon j) \neq 0$ for $1 \leq j < |i|$, and by the theorem $M(\lambda, \lambda_+, \lambda_-)$ is isomorphic to $M(\mu, \mu_+, \mu_-)$ with $\mu = \lambda - 2i$ and $\mu_\epsilon = a_\epsilon(i) = 0$. $\square$

We therefore must analyse the submodule structure of the modules $M(\mu, \mu, 0)$ and $M(\mu, 0, -\mu)$. For this we have two important definitions.

**(7.16).** DEFINITION. *Let $\lambda \in \mathbb{K}$ and set $M_+(\lambda) = \bigoplus_{i \in \mathbb{N}} \mathbb{K}v_i$ with $v_{-1} = 0 \in M_+(\lambda)$. Define the linear transformations $\mathsf{H}$, $\mathsf{E}_+$, and $\mathsf{E}_-$ on $M_+(\lambda)$ by*

$$\mathsf{H}(v_i) = (\lambda - 2i)v_i\,,$$
$$\mathsf{E}_\epsilon(v_i) = a_\epsilon(i)v_{i-\epsilon}\,,$$

*for $i \in \mathbb{N}$, $\epsilon \in \{\pm\} = \{\pm 1\}$, and $a_\epsilon(i)$ by*

$$a_+(i) = i(\lambda - i + 1) \qquad \textit{for } i \in \mathbb{Z}^+ \textit{ and}$$
$$a_-(i) = 1 \qquad \textit{for } i \in \mathbb{N}\ :$$

$$\begin{tikzcd}
\mathsf{E}_- & v_0 \arrow[r, "1"] \arrow[d, "\lambda"] & v_1 \arrow[r, "1"] \arrow[d, "\lambda-2"] & v_2 \arrow[r] \arrow[d, "\lambda-4"] & {} \\
\mathsf{E}_+ & v_0 & v_1 \arrow[l, dashed, "\lambda"'] & v_2 \arrow[l, dashed, "a_+(2)"'] & {} \arrow[l, dashed]
\end{tikzcd}$$

**(7.17).** Definition. *Let $\lambda \in \mathbb{K}$ and set $M_-(\lambda) = \bigoplus_{i \in \mathbb{N}} \mathbb{K}v_{-i}$ with $v_1 = 0 \in M_-(\lambda)$. Define the linear transformations $\mathsf{H}$, $\mathsf{E}_+$, and $\mathsf{E}_-$ on $M_-(\lambda)$ by*

$$\begin{aligned}
\mathsf{H}(v_i) &= (\lambda - 2i)v_i\,, \\
\mathsf{E}_\epsilon(v_i) &= a_\epsilon(i)v_{i-\epsilon}\,,
\end{aligned}$$

*for $i \in -\mathbb{N}$, $\epsilon \in \{\pm\} = \{\pm 1\}$, and $a_\epsilon(i)$ by*

$$\begin{aligned}
a_-(i) &= i(\lambda - i - 1) & &\textit{for } i \in -\mathbb{Z}^+ \textit{ and} \\
a_+(i) &= 1 & &\textit{for } i \in -\mathbb{N}\ :
\end{aligned}$$

$$\begin{tikzcd}
\mathsf{E}_- & {} \arrow[r, dashed] & v_{-2} \arrow[r, dashed, "a_-(-2)"] \arrow[d, "\lambda+4"] & v_{-1} \arrow[r, dashed, "-\lambda"] \arrow[d, "\lambda+2"] & v_0 \arrow[d, "\lambda"] \\
\mathsf{E}_+ & {} & v_{-2} \arrow[l] & v_{-1} \arrow[l, "1"] & v_0 \arrow[l, "1"]
\end{tikzcd}$$

The space $M_+(\lambda)$, defined above and with the described action, is the *Verma module* for $\mathfrak{sl}_2(\mathbb{K})$ with highest weight $\lambda$. Similarly $M_-(\lambda)$ with the described action is the *Verma module* for $\mathfrak{sl}_2(\mathbb{K})$ with lowest weight $\lambda$. At this point, these names are presumptive, since we have not proven that the maps $h \mapsto \mathsf{H}$, $e \mapsto \mathsf{E}_+$, and $f \mapsto \mathsf{E}_-$ give $M_+(\lambda)$ or $M_-(\lambda)$ the structure of a cyclic $\mathfrak{sl}_2(\mathbb{K})$-module generated by the weight vector $v_0$ with highest or lowest weight $\lambda$. This will be a consequence of the next theorem, where we will see each of these as quotient modules and submodules of appropriate generalized Verma modules.

**(7.18).** Theorem. *For each $\lambda \in \mathbb{K}$ we have following nonsplit exact sequences of $\mathfrak{sl}_2(\mathbb{K})$-modules:*

(a)
$$0 \longrightarrow M_-(\lambda+2) \longrightarrow M(\lambda,\lambda,0) \longrightarrow M_+(\lambda) \longrightarrow 0\,;$$

(b)
$$0 \longrightarrow M_+(\lambda-2) \longrightarrow M(\lambda,0,-\lambda) \longrightarrow M_-(\lambda) \longrightarrow 0\,.$$

Proof. For $M(\lambda,\lambda,0)$ we have

$$\begin{tikzcd}
\mathsf{E}_- & {} \arrow[r, dashed] & v_{-2} \arrow[r, dashed, "a_-(-2)"] \arrow[d, "\lambda+4"] & v_{-1} \arrow[d, "\lambda+2"] & v_0 \arrow[r, "1"] \arrow[d, "\lambda"] & v_1 \arrow[r, "1"] \arrow[d, "\lambda-2"] & v_2 \arrow[r] \arrow[d, "\lambda-4"] & {} \\
\mathsf{E}_+ & {} & v_{-2} \arrow[l] & v_{-1} \arrow[l, "1"] & v_0 \arrow[l, "1"] & v_1 \arrow[l, dashed, "\lambda"'] & v_2 \arrow[l, dashed, "a_+(2)"'] & {} \arrow[l, dashed]
\end{tikzcd}$$

Thus $K_- = \sum_{i\in\mathbb{Z}^-} \mathbb{K}v_i$ is a proper submodule and is, in fact, isomorphic to $M_-(\lambda+2)$. The quotient $M(\lambda,\lambda,0)/K_-$ is next seen to be a copy of $M_+(\lambda)$. The extension is nonsplit since, by Corollary (7.12)(c), the only possible submodule complement to $K_-$ would be $\sum_{i\in\mathbb{N}} \mathbb{K}v_i$ whereas $\mathsf{E}_+v_0 \in K_-$.

For $M(\lambda,0,-\lambda)$ instead

$$
\begin{tikzcd}
\mathsf{E}_- & {} \arrow[r, dashed] & v_{-2} \arrow[r, dashed, "a_-(-2)"] \arrow[d, "\lambda+4"] & v_{-1} \arrow[r, dashed, "\lambda"] \arrow[d, "\lambda+2"] & v_0 \arrow[r, "1"] \arrow[d, "\lambda"] & v_1 \arrow[r, "1"] \arrow[d, "\lambda-2"] & v_2 \arrow[r] \arrow[d, "\lambda-4"] & {} \\
\mathsf{E}_+ & {} & v_{-2} \arrow[l] & v_{-1} \arrow[l, "1"'] & v_0 \arrow[l, "1"'] & v_1 & v_2 \arrow[l, dashed, "a_+(2)"] & {} \arrow[l, dashed]
\end{tikzcd}
$$

Here $K_+ = \sum_{i\in\mathbb{Z}^+} \mathbb{K}v_i$ is a proper submodule and is isomorphic to $M_+(\lambda-2)$. The quotient $M(\lambda,0,\lambda)/K_+$ is a copy of $M_-(\lambda)$. Again the extension is nonsplit as the only possible complement to $K_+$ would be $\sum_{i\in\mathbb{N}} \mathbb{K}v_{-i}$ but $\mathsf{E}_-v_0 \in K_+$.
□

Especially, the spaces $M_\epsilon(\lambda)$ are indeed $\mathfrak{sl}_2(\mathbb{K})$-modules, as presumed above.

**(7.19).** Corollary.

(a) *Let $v$ be a highest weight vector for the weight $\lambda$ in the $\mathfrak{sl}_2(\mathbb{K})$-module $V$. Then the map $v_0 \mapsto v$ extends to a Lie module homomorphism taking the Verma module $M_+(\lambda)$ with highest weight $\lambda$ onto the $\mathfrak{sl}_2(\mathbb{K})$-submodule of $V$ generated by $v$.*

(b) *Let $v$ be a lowest weight vector for the weight $\lambda$ in the $\mathfrak{sl}_2(\mathbb{K})$-module $V$. Then the map $v_0 \mapsto v$ extends to a Lie module homomorphism taking the Verma module $M_-(\lambda)$ with lowest weight $\lambda$ onto the $\mathfrak{sl}_2(\mathbb{K})$-submodule of $V$ generated by $v$.* □

**(7.20).** Theorem.

(a) *The module $M_\epsilon(\lambda)$ is irreducible if and only if $\epsilon\lambda \notin \mathbb{N}$.*

(b) *If $\lambda = n \in \epsilon\mathbb{N}$, then $M_\epsilon(n)$ is indecomposable with two composition factors:*

$$0 \longrightarrow M_\epsilon(-n-2\epsilon) \longrightarrow M_\epsilon(n) \longrightarrow M_0(\epsilon n+1) \longrightarrow 0\,,$$

*with $M_\epsilon(-n-2\epsilon)$ irreducible of infinite dimension and $M_0(\epsilon n+1)$ irreducible of finite dimension $\epsilon n+1$.*

Proof. By Corollary (7.12) every submodule is $\sum_{i\in I} \mathbb{K}v_i$ for some $I \subseteq \epsilon\mathbb{N}$. As it is $\mathsf{E}_{-\epsilon}$-invariant, a nonzero submodule must be $\sum_{k\leq i\in\mathbb{N}} \mathbb{K}v_{\epsilon i}$ for some $k \in \mathbb{N}$. When $a_\epsilon(i)$ is nonzero for all $i \in \epsilon\mathbb{N}$, the module $M_\epsilon(\lambda)$ itself is the only such $\mathsf{E}_\epsilon$-invariant subspace. Therefore $M_\epsilon(\lambda)$ is irreducible unless

$$a_\epsilon(i) = i(\lambda - i + \epsilon) = i((\lambda+\epsilon) - i) = 0$$

for some $i \in \epsilon\mathbb{Z}^+$. For a fixed $\lambda$ and $\epsilon$ this can only happen for

$$i = \lambda + \epsilon \in \epsilon\mathbb{Z}^+\,,$$

which is to say

$$\lambda = i - \epsilon \in \epsilon\mathbb{N}\,.$$

Especially if $\lambda \notin \epsilon\mathbb{N}$, then $M_\epsilon(\lambda)$ is irreducible.

Now suppose $\lambda = n = \epsilon m \in \epsilon\mathbb{N}$ for $m \in \mathbb{N}$, so that $\epsilon(m+1) = \lambda + \epsilon$ with $a_\epsilon(\lambda+\epsilon) = 0$. Then $K_\epsilon = \bigoplus_{\epsilon j \geq m+1} \mathbb{K}v_j$ is a submodule of $M_\epsilon(\lambda)$, complemented by the $(m+1)$-subspace $\bigoplus_{\epsilon j=0}^{m} \mathbb{K}v_j$. In particular $M_\epsilon(\lambda)$ is reducible, completing (a).

For the submodule $K_\epsilon = \bigoplus_{\epsilon j \geq m+1} \mathbb{K}v_j$ we calculate

$$\mathsf{H}(v_{\epsilon(m+1)}) = \mathsf{H}(v_{\lambda+\epsilon}) = \lambda - 2(\lambda+\epsilon) = -\lambda - 2\epsilon = -n - 2\epsilon\,.$$

Thus $K_\epsilon$ is isomorphic to $M_\epsilon(-n-2\epsilon)$, an infinite dimensional irreducible module as $\epsilon(-n-2\epsilon) = -\epsilon n - 2 \in \mathbb{Z}^-$. The extension is nonsplit since, by Corollary (7.12)(c) and the Third Isomorphism Theorem, the only possible complementary submodule would be $\bigoplus_{\epsilon j=0}^{m} \mathbb{K}v_j$, whereas $\mathsf{E}_{-\epsilon}v_{\epsilon m} \in \mathbb{K}_\epsilon$.

For $m \in \mathbb{N}$, the quotients $M_\epsilon(\epsilon m)/K_\epsilon$ have dimension $m+1$ and are irreducible, since there are no further solutions to $i((\lambda+\epsilon) - i) = 0$. By Theorem (7.5) every finite dimensional irreducible $\mathfrak{sl}_2(\mathbb{K})$-module has both high weight vectors and low weight vectors. In particular each $M_\epsilon(\epsilon m)/K_\epsilon$ must also be $M_{-\epsilon}(-\epsilon k)/K_{-\epsilon}$ for some $k$, dimension considerations forcing $m = k$. That is, the two finite dimensional modules $M_\epsilon(\epsilon m)/K_\epsilon$ are isomorphic. We conclude that, up to isomorphism, there is a unique irreducible $\mathfrak{sl}_2(\mathbb{K})$-module of each positive dimension $m+1$. This we have denoted $M_0(m+1)$. □

For example, with $\epsilon = +$ and $\lambda = m = 3$ we have

$$\begin{tikzcd}
\mathsf{E}_- & v_0 \arrow[r, "1"] \arrow[d, "3"] & v_1 \arrow[r, "1"] \arrow[d, "1"] & v_2 \arrow[r, "1"] \arrow[d, "-1"] & v_3 \arrow[r, "1"] \arrow[d, "-3"] & v_4 \arrow[r, "1"] \arrow[d, "-5"] & v_5 \\
\mathsf{E}_+ & v_0 & v_1 \arrow[l, dashed, "3"] & v_2 \arrow[l, dashed, "4"] & v_3 \arrow[l, dashed, "3"] & v_4 & v_5 \arrow[l, dashed, "-5"]
\end{tikzcd}$$

and the submodule $K_+ = \bigoplus_{j \geq 4} \mathbb{K}v_j$ of $M_+(3)$ is revealed as a copy of $M_+(-5)$, while the quotient module is $M_0(4)$ with dimension 4 and weights $\{3, 1, -1, -3\}$.

We have already seen on page 18 a version of the module $M_0(m+1)$. Let

$$e \mapsto x\frac{\partial}{\partial y} \quad \text{and} \quad f \mapsto y\frac{\partial}{\partial x} \quad \text{and} \quad h \mapsto x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}$$

in $\mathrm{Der}_{\mathbb{K}}(\mathbb{K}[x,y])$ acting on $\mathbb{K}[x,y]_m$, the space of homogeneous polynomials of total degree $m$ in $\mathbb{K}[x,y]$. The polynomial $x^m$ is a weight vector with highest weight $m$, while $y^m$ is a lowest weight vector with weight $-m$; we have a module isomorphism with $M_+(m)/K_+$ given by $v_i + K_+ \mapsto x^{m-i}y^i$ , for $0 \leq i \leq n$.

The next result is a corollary of the previous one and is of fundamental importance. Versions of it are at the heart of the representation theory for all semisimple Lie algebras over algebraically closed fields of characteristic 0. In particular, see Theorem (9.10) below.

**(7.21).** THEOREM.
(CLASSIFICATION OF IRREDUCIBLE HIGHEST WEIGHT $\mathfrak{sl}_2(\mathbb{K})$-MODULES) *For every $\lambda \in \mathbb{K}$, up to isomorphism there is a unique irreducible $\mathfrak{sl}_2(\mathbb{K})$-module $L_+(\lambda)$ with highest weight $\lambda$. Indeed*

(a) *if $\lambda \notin \mathbb{N}$, then $L_+(\lambda) = M_+(\lambda)$ of infinite dimension;*

(b) *if $\lambda = n \in \mathbb{N}$, then $L_+(\lambda) = M_0(n+1)$ of finite dimension $n+1$.*

*These are pairwise nonisomorphic.* □

There is, of course, a corresponding result for irreducible lowest weight modules.

Every irreducible finite dimensional $\mathfrak{sl}_2(\mathbb{K})$-module is a highest weight module, so we also have the following result which was mentioned in the introduction to this chapter and will aid us to complete the classification of semisimple Lie algebras over $\mathbb{K}$.

**(7.22).** THEOREM. *Up to isomorphism the finite dimensional irreducible $\mathfrak{sl}_2(\mathbb{K})$-modules are the modules $M_0(m+1)$ for $m \in \mathbb{N}$. These modules are all self-dual. In $M_0(m+1)$ the weights are $-m, -m+2, \ldots m-2, m$, and each weight space has dimension 1.* □

## 7.3 The Casimir operator

**(7.23).** PROPOSITION. *Let $V$ be an $\mathfrak{sl}_2(\mathbb{K})$-module, and in $\mathrm{End}_{\mathbb{K}}(V)$ define the element*

$$\mathsf{C} = \mathsf{C}^V = (\mathsf{H}^V)^2 + 1 + 2(\mathsf{E}_+^V\mathsf{E}_-^V + \mathsf{E}_-^V\mathsf{E}_+^V)\,.$$

(a) $\mathsf{C} = (\mathsf{H} - \epsilon)^2 + 4\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon}$.

(b) $\mathsf{CH} = \mathsf{HC}$ *and* $\mathsf{CE}_\epsilon = \mathsf{E}_\epsilon\mathsf{C}$.

PROOF. We make frequent use of Lemma (7.3):

$$\begin{aligned}\mathsf{C} &= \mathsf{H}^2 + 1 + 2(\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon} + \mathsf{E}_{-\epsilon}\mathsf{E}_\epsilon)\\ &= \mathsf{H}^2 + 1 + 2(\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon} - \epsilon\mathsf{H} + \mathsf{E}_\epsilon\mathsf{E}_{-\epsilon})\\ &= (\mathsf{H}^2 - 2\epsilon\mathsf{H} + 1) + 4\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon}\\ &= (\mathsf{H} - \epsilon)^2 + 4\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon}\,.\end{aligned}$$

$$\begin{aligned}\mathsf{HC} &= \mathsf{H}(\mathsf{H} - \epsilon)^2 + 4(\mathsf{HE}_\epsilon)\mathsf{E}_{-\epsilon}\\ &= (\mathsf{H} - \epsilon)^2\mathsf{H} + 4\mathsf{E}_\epsilon((\mathsf{H} + 2\epsilon)\mathsf{E}_{-\epsilon})\\ &= (\mathsf{H} - \epsilon)^2\mathsf{H} + 4\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon}(\mathsf{H} + 2\epsilon - 2\epsilon)\\ &= \mathsf{CH}\,.\end{aligned}$$

$$\begin{aligned}
\mathsf{C}\mathsf{E}_\epsilon &= (\mathsf{H}-\epsilon)^2\mathsf{E}_\epsilon + 4\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon}\mathsf{E}_\epsilon \\
&= \mathsf{E}_\epsilon(\mathsf{H}-\epsilon+2\epsilon)^2 + 4\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon}\mathsf{E}_\epsilon \\
&= \mathsf{E}_\epsilon((\mathsf{H}+\epsilon)^2 + 4\mathsf{E}_{-\epsilon}\mathsf{E}_\epsilon) \\
&= \mathsf{E}_\epsilon\mathsf{C}\,. \qquad \square
\end{aligned}$$

The somewhat mysterious $\mathsf{C} = \mathsf{C}^V$ is the *Casimir operator*[3] on $V$. It has its uses.

**(7.24).** PROPOSITION. *In an irreducible weight module for $\mathfrak{sl}_2(\mathbb{K})$ every weight vector $v$ for $\lambda$ is coherent.*

PROOF. If the $\mathfrak{sl}_2(\mathbb{K})$-module $V$ is irreducible, then by Schur's Lemma (4.11) the Casimir operator $\mathsf{C}$ acts as a scalar: there is a $c \in \mathbb{K}$ with $\mathsf{C}v = cv$ for all $v \in V$. Let $v$ be a weight vector in $V$, say, for the weight $\lambda$. As $\mathsf{C} = (\mathsf{H}-\epsilon)^2+4\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon}$, we have $\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon} = \frac{1}{4}(\mathsf{C}-(\mathsf{H}-\epsilon)^2)$. Thus

$$\begin{aligned}
\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon}v &= \frac{1}{4}(\mathsf{C}-(\mathsf{H}-\epsilon)^2)v \\
&= \frac{1}{4}(\mathsf{C}v-(\mathsf{H}-\epsilon)^2v) \\
&= \frac{c-(\lambda-\epsilon)^2}{4}v\,.
\end{aligned}$$

That is, $v$ is a coherent weight vector for $\lambda$ with $\lambda_\epsilon = \frac{c-(\lambda-\epsilon)^2}{4}$. $\square$

**(7.25).** COROLLARY. *Let $V$ be an irreducible weight module for $\mathfrak{sl}_2(\mathbb{K})$ containing the weight vector $v$ for $\lambda$ with $\mathsf{E}_\epsilon\mathsf{E}_{-\epsilon}v = \lambda_\epsilon v$. Then the Casimir operator acts as scalar multiplication by $4\lambda_\epsilon + (\lambda-\epsilon)^2$.* $\square$

The *spectrum* $\mathrm{Spec}(V)$ of an $\mathfrak{sl}_2(\mathbb{K})$-module $V$ is the set of weights associated with weight vectors in the module.

**(7.26).** THEOREM. *An irreducible weight module for $\mathfrak{sl}_2(\mathbb{K})$ is isomorphic to one of the following:*

(1) *$M_0(m+1)$ for $m \in \mathbb{N}$ with spectrum $[-m,m]\cap(m+2\mathbb{Z})$;*

(2) *$M_+(\lambda)$ for $\lambda \notin \mathbb{N}$ with spectrum $\lambda-2\mathbb{N}$;*

(3) *$M_-(\lambda)$ for $-\lambda \notin \mathbb{N}$ with spectrum $\lambda+2\mathbb{N}$;*

(4) *$M(\lambda,\lambda_+,\lambda_-)$ with $\lambda_- = \lambda_+ - \lambda$ and spectrum $\lambda+2\mathbb{Z}$, such that the polynomial $x^2-(\lambda-1)x-\lambda_+ \in \mathbb{K}[x]$ has no integer roots.*

[3] In other places one my find our Casimir operator $C$ replaced by $aC+b$ for constants $a,b \in \mathbb{K}$. This has no affect on its uses. We follow the convention of [Maz10], which gives the nice renditions of Proposition (7.23)(a).

*All these modules are irreducible and every weight space has dimension* 1.

*No module from one case is isomorphic to a module from one of the other cases. Within each of the first three cases, two modules are isomorphic if and only if they have the same parameter. In the last case,* $M(\lambda, \lambda_+, \lambda_-)$ *is isomorphic to* $M(\mu, \mu_+, \mu_-)$ *if and only if* $\mu - \lambda \in 2\mathbb{Z}$ *and* $\mu_+ - \lambda_+ = \frac{1}{4}(\lambda - \mu)(\lambda + \mu - 2)$.

Proof. By the previous proposition, an irreducible weight module $V$ is generated by a coherent weight vector. If the module is not isomorphic to some $M(\lambda, \lambda_+, \lambda_-)$ (as in the last case), then by Corollary (7.15) it is a quotient of a Verma module $M_\epsilon(\lambda)$ . By Theorem (7.20) the module $V$ is then isomorphic to one of the examples in the first three conclusions, all irreducible. In any event, all weight spaces have dimension 1.

The various spectra are also clear, and show that no module from one case is isomorphic to one from another case, nor can different parameters in any one of the first three cases produce isomorphic modules.

It remains to decide under what circumstances $M(\lambda, \lambda_+, \lambda_-)$ is irreducible and when two such modules can be isomorphic.

By Corollaries (7.7) and (7.15)(a) the module $M(\lambda, \lambda_+, \lambda_-)$ is irreducible if and only if $b(j) = j(\lambda - j - 1) + \lambda_+$ is nonzero for all $j \in \mathbb{Z}$. This is the case precisely when

$$x(\lambda - x - 1) + \lambda_+ = -x^2 + (\lambda - 1)x + \lambda_+ \in \mathbb{K}[x]$$

has no integral roots.

Suppose that $M(\lambda, \lambda_+, \lambda_-)$ and $M(\mu, \mu_+, \mu_-)$ are isomorphic. By spectral considerations, we must have $\lambda + 2\mathbb{Z} = \mu + 2\mathbb{Z}$. By symmetry we may assume that $\mu = \lambda - 2i$ for some $i \in \mathbb{N}$. In that case, isomorphism holds if and only if, in terms of the transformation coefficients for $M(\lambda, \lambda_+, \lambda_-)$, we have

$$\mu_+ = a_+(i+1) = ((i+1) - 1)(\lambda - (i+1)) + \lambda_+ = i(\lambda - i - 1) + \lambda_+ \,.$$

As $\mu = \lambda - 2i$ we have $i = \frac{\lambda - \mu}{2}$, so this becomes $\mu_+ - \lambda_+ = \frac{1}{4}(\lambda - \mu)(\lambda + \mu - 2)$, as claimed. □

If $V$ is an irreducible module containing the (coherent) weight vector $v$ for $\lambda$ with $\mathsf{E}_\epsilon \mathsf{E}_{-\epsilon} v = \lambda_\epsilon$, then the Casimir scalar is $c = 4\lambda_\epsilon + (\lambda - \epsilon)^2$ by Corollary (7.25). For instance, for $V = M_0(m+1)$ we have $\lambda = m$ and $\lambda_- = 0$ so that $c = 4(0) + (m+1)^2 = (m+1)^2$.

This also allows effective relabelling of the irreducible modules $M(\lambda, \lambda_+, \lambda_-)$. Indeed this module can now be characterized by the two parameters $\lambda + 2\mathbb{Z} \in (\mathbb{K}, +)/2\mathbb{Z}$ and $c\,(= 4\lambda_\epsilon + (\lambda - \epsilon)^2)$, two such irreducible modules being isomorphic if and only if they have the same parameter pair. (Exercise.) This is the approach taken in [Maz10, Theorem 3.32]. The corresponding requirement for irreducibility is that $c \neq (\mu + 1)^2$ for all $\mu \in \lambda + 2\mathbb{Z}$. (Exercise.)

## 7.4 Finite dimensional $\mathfrak{sl}_2(\mathbb{K})$-modules

We have a second hidden use of the Casimir operator.

**(7.27).** THEOREM. (WEYL'S THEOREM) *Every finite dimensional $\mathfrak{sl}_2(\mathbb{K})$-module is completely reducible.*

PROOF. Equivalently, an extension $V$ of a finite dimensional irreducible $\mathfrak{sl}_2(\mathbb{K})$-module $W \simeq M_0(m+1)$ by a second finite dimensional irreducible module is always split. Consider

$$0 \longrightarrow M_0(m+1) \longrightarrow V \longrightarrow M_0(n+1) \longrightarrow 0\,.$$

By passing to the dual of $V$, if necessary, we may assume $m \leq n$. If $m < n$, then the weight space $V_n$ has dimension 1. But then (say, by Corollary (7.19) and Theorem (7.20)), $V_n$ generates a submodule $U$ of $V$ that is isomorphic to $M_0(n+1)$ and splits the extension as $V = W \oplus U$. Therefore we may assume $m = n$.

Each of the weight spaces $V_k$ for $k \in [-m, m] \cap (m + 2\mathbb{Z})$ has dimension 2 with $V_m = \ker \mathsf{E}_+$ and $V_{-m} = \ker \mathsf{E}_-$. By Lemma (7.3)(d) we have

$$V_m = \mathsf{E}_+^m V = \mathsf{E}_+^m V_{-m}\,,$$

with $\ker \mathsf{E}_+^m = \bigoplus_{i=-m+2}^m V_i$ and $\mathsf{A} = \mathsf{E}_+^m$ an invertible linear transformation from $V_{-m}$ to $V_m$.

For $\mathsf{C} = \mathsf{C}^V$, the Casimir operator, we calculate the map $\mathsf{CA} = \mathsf{AC}$ from $V_{-m}$ to $V_m$ in two ways. For $v \in V_{-m}$ we first have

$$\begin{aligned}(\mathsf{CA})v &= \mathsf{C}(\mathsf{A}v)\\ &= ((\mathsf{H}+1)^2 + 4\mathsf{E}_-\mathsf{E}_+)(\mathsf{A}v)\\ &= (\mathsf{H}+1)^2(\mathsf{A}v)\,,\end{aligned}$$

as $\mathsf{A}v \in V_m = \ker \mathsf{E}_+$. Similarly as $v \in V_{-m} = \ker \mathsf{E}_-$,

$$\begin{aligned}(\mathsf{AC})v &= \mathsf{A}(\mathsf{C}v)\\ &= \mathsf{A}((\mathsf{H}-1)^2 + 4\mathsf{E}_+\mathsf{E}_-)v\\ &= \mathsf{A}(\mathsf{H}-1)^2 v\\ &= \mathsf{E}_+^m(\mathsf{H}-1)^2 v\\ &= (\mathsf{H}-1-2m)^2\mathsf{E}_+^m v && \text{by Lemma (7.3)(c)}\\ &= (\mathsf{H}-1-2m)^2(\mathsf{A}v)\end{aligned}$$

Therefore, from $V_{-m}$ to $V_m$, the map $\mathsf{AC} = \mathsf{CA}$ is equal to

$$(\mathsf{H}+1)^2\mathsf{A} = (\mathsf{H}-1-2m)^2\mathsf{A}\,.$$

Since $\mathsf{A}$ is a bijection from $V_{-m}$ to $V_m$, this says that, as a map from $V_m$ to $V_m$,

$$0 = (\mathsf{H}+1)^2 - (\mathsf{H}-1-2m)^2 = (4m+4)(\mathsf{H}-m)\,.$$

We already knew that $(\mathsf{H}-m)^2 V_m = 0$, because the weight space $V_m$ has dimension 2. Now we have learned that $(\mathsf{H}-m)V_m = 0$, so $V_m = V_m^w$ consists entirely of weight vectors (and 0).

Let $u \in V_m \setminus W$. Then again by Corollary (7.19) and Theorem (7.20)), the weight vector $u$ generates a submodule $U$ of $V$ that is isomorphic to $M_0(m+1)$ and splits the extension as $V = W \oplus U$. □

Weyl's Theorem is valid for all finite dimensional semisimple algebras over $\mathbb{K}$, not just $\mathfrak{sl}_2(\mathbb{K})$. The proof uses the appropriate generalization of the current Casimir operator.

From Theorem (7.26)(a) and Weyl's Theorem (7.27) we immediately have:

**(7.28).** THEOREM.

(a) *The spectrum of a finite dimensional $\mathfrak{sl}_2(\mathbb{K})$-module $V$ has one of the following types:*

  (i) $[-m, m] \cap (m + 2\mathbb{Z})$ *for some* $m \in \mathbb{N}$;

  (ii) $([-m, m] \cap (m + 2\mathbb{Z})) \cup ([-n, n] \cap (n + 2\mathbb{Z}))$ *for some even* $m$ *and some odd* $n$ *from* $\mathbb{N}$.

  *In particular, $V$ contains a $\mathfrak{sl}_2(\mathbb{K})$-invariant $\mathbb{Q}$-submodule $V_{\mathbb{Q}}$ with $V = \mathbb{K} \otimes_{\mathbb{Q}} V_{\mathbb{Q}}$.*

(b) *The number of composition factors, indeed irreducible summands, in a finite dimensional $\mathfrak{sl}_2(\mathbb{K})$-module $V$ is* $\dim_{\mathbb{K}}(V_0) + \dim_{\mathbb{K}}(V_1)$. □

## 7.5 Problems

**(7.29).** PROBLEM. *Let $\lambda, \lambda_+, \lambda_- \in \mathbb{K}$ with $\lambda_+ - \lambda_- = \lambda$. Set $M_\star(\lambda, \lambda_+, \lambda_-) = \bigoplus_{i \in \mathbb{Z}} \mathbb{K} v_i$. Define the linear transformations $\mathsf{H}$, $\mathsf{E}_+$, and $\mathsf{E}_-$ on $M_\star(\lambda, \lambda_+, \lambda_-)$ by*

$$\begin{aligned} \mathsf{H}(v_i) &= (\lambda - 2i)v_i\,; \\ \mathsf{E}_+(v_i) &= v_{i-1}\,; \\ \mathsf{E}_-(v_i) &= b(i)v_{i+1}\,, \end{aligned}$$

*for all $i \in \mathbb{Z}$ with $b(i) = i(\lambda - i - 1) + \lambda_+$ (as in Corollary (7.7)).*

(a) *Prove that the maps*

$$h \mapsto \mathsf{H}\,, \quad e \mapsto \mathsf{E}_+\,, \quad f \mapsto \mathsf{E}_-$$

*give $M_\star(\lambda, \lambda_+, \lambda_-)$ the structure of an $\mathfrak{sl}_2(\mathbb{K})$-module in which each weight space has dimension 1.*

(b) *Prove that if the generalized Verma module $M(\lambda, \lambda_+, \lambda_-)$ is irreducible, then it is isomorphic to $M_\star(\lambda, \lambda_+, \lambda_-)$.*

(c) *Prove that $M_0(m+1)$ is never a quotient of $M_\star(\lambda, \lambda_+, \lambda_-)$. In particular if the generalized Verma module $M(\lambda, \lambda_+, \lambda_-)$ is reducible, then it need not be isomorphic to $M_\star(\lambda, \lambda_+, \lambda_-)$.*

REMARK. *The modules $M_\star(\lambda, \lambda_+, \lambda_-)$ are those used in Mazorchuk's excellent book [Maz10, Chapter 3].*

**(7.30).** PROBLEM. *CG decompose $M_0(m+1) \otimes_{\mathbb{K}} M_0(n+1)$.*

# Chapter 8

# Semisimple Lie algebras

We return to the classification of finite dimensional, semisimple Lie algebras over algebraically closed fields of characteristic 0, begun in Section 6.2.

We recall some notation to be used throughout this chapter (except Section 8.2). Again $L\,(\neq 0)$ will be a finite dimensional, semisimple Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic 0.

By Theorem (5.7) there is a Cartan subalgebra $H$ in $L$, and $H$ is abelian by Theorem (6.8). By Proposition (5.8) we have $H = L_{H,0} = L_0$, the zero weight space. Let $\Phi$ be the set of all roots for $H$ on $L$, a finite set by Theorem (5.10).

For each $\lambda \in \Phi$, we have the weight space $L_\lambda = L_{H,\lambda} = L_\lambda^w$ (by Theorem (6.9)), giving the Cartan decomposition

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

Since $L$ is nonzero and semisimple, the abelian Cartan subalgebra $H = L_0$ is proper in $L$, hence the root set $\Phi$ is nonempty.

The Killing form $\kappa = \kappa^L = \kappa_L = \kappa_L^L$ is nondegenerate by Cartan's Semisimplicity Criterion (6.5). Furthermore the restriction of $\kappa$ to abelian $H$ is nondegenerate by Proposition (6.7)(b). Thus for every linear functional $\mu \in H^*$, and especially for every root in $\Phi$, there is a unique $t_\mu \in H$ with $\kappa(t_\mu, h) = \mu(h)$ for all $h \in H$.

Let $E_{\mathbb{Q}} = \sum_{\alpha \in \Phi} \mathbb{Q}\alpha \leq H^*$. Define on $H^*\,(\geq E_{\mathbb{Q}})$ the symmetric bilinear form

$$(x, y) = \kappa(t_x, t_y)\,.$$

For the root $\alpha \in \Phi$, define $h_\alpha = \frac{2}{\kappa(t_\alpha, t_\alpha)} t_\alpha$, possible by Proposition (6.10)(c). Similarly for each $\alpha \in \Phi$, we let $\alpha^\vee = \frac{2}{\kappa(t_\alpha, t_\alpha)}\alpha = \frac{2}{(\alpha,\alpha)}\alpha$, the *coroot* corresponding to the root $\alpha$. Then

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

## 8.1 Semisimple algebras II: Root systems

Our notation is that of the introduction to this chapter.

On page 51 we introduced $\alpha$-strings. We return to the idea, now in a broader context.

**(8.1).** PROPOSITION. *Let $V$ ($\neq 0$) be a finite dimensional $L$-module with weight set $\Phi_V$ (with respect to the Cartan subalgebra $H$). Let $\alpha \in \Phi$ and $\mu \in \Phi_V$.*

(a) *$L_\alpha V_\mu \le V_{\mu+\alpha}$ and $L^w_\alpha V^w_\mu \le V^w_{\mu+\alpha}$ with $L^w_\alpha V^w_\mu \neq 0$ if $\mu+\alpha \in \Phi(V)$.*

(b) *If $V_\mu \neq 0$ then $V^w_\mu \neq 0$, hence $0 \neq V^w \le V$. In particular, if $V$ is irreducible then $V = V^w$.*

(c) *There are $s,t \in \mathbb{N}$ with $V_{\mu+k\alpha} \neq 0$ if and only if $k \in [-s,t]$.*

(d) *$\mu(h_\alpha) = (\mu,\alpha^\vee) = s-t \in \mathbb{Z}$.*

(e) *If $(\mu,\alpha^\vee) < 0$, then $\mu+\alpha \in \Phi_V$.*

(f) *$\mu - (\mu,\alpha^\vee)\alpha \in \Phi_V$.*

PROOF. Part (a) is primarily an application of Theorem (5.11) in the semidirect product $T = L \oplus V$ (as described in Example (4.15)(b)). We have $L_\alpha \le T_\alpha$ and $V_\mu \le T_\mu$, so

$$L_\alpha V_\mu = [L_\alpha, V_\mu] \le [T_\alpha, T_\mu] \le T_{\mu+\alpha}\,.$$

As $V$ is an ideal of $T$, we further have $L_\alpha V_\mu = [L_\alpha, V_\mu] \le T_{\mu+\alpha} \cap V = V_{\mu+\alpha}$. Theorem (5.11) gives $L^w_\alpha V^w_\mu \le V^w_{\mu+\alpha}$ in the same way. All that remains to be shown from (a) is that $L^w_\alpha V^w_\mu \neq 0$ provided $\mu+\alpha \in \Phi(V)$, and that will follow from arguments below.

If $V_\mu \neq 0$ then, by standard linear algebra, the pairwise commuting endomorphisms $h_1,\dots,h_l$ have a common nonzero eigenvector.[1] Therefore if $V_\mu \neq 0$ then $V^w_\mu \neq 0$. By this and (a), $V^w$ is a nonzero submodule of $V$. Thus when $V$ is irreducible it is equal to $V^w$. This gives (b).

By (b) from now on we may assume that $V_{\mu+k\alpha} = V^w_{\mu+k\alpha}$.

Following Theorem (6.11) we choose $x \in L_\alpha$ and $y \in L_{-\alpha}$ with $h_\alpha = [x,y]$ and $S = \mathbb{K}h_\alpha \oplus \mathbb{K}x \oplus \mathbb{K}y$ a subalgebra of $L$ isomorphic to $\mathfrak{sl}_2(\mathbb{K})$. The subspace

$$M = \sum_{k\in\mathbb{Z}} V_{\mu+k\alpha}$$

[1] Exercise: induction on $l$ with $l = 1$ coming from Jordan Canonical Form.

of $V$ is then an $S$-module. For each $k$ we have

$$(\mu + k\alpha)(h_\alpha) = \mu(h_\alpha) + k\alpha(h_\alpha) = (\mu, \alpha^\vee) + 2k\,,$$

so by Theorem (7.28)(a) we have $(\mu, \alpha^\vee) \in \mathbb{Z}$ and

$$\mathrm{Spec}(M) = [-m, m] \cap (m + 2\mathbb{Z}) = [-m, m] \cap ((\mu, \alpha^\vee) + 2\mathbb{Z})$$

where $m \in \mathbb{N}$ satisfies

$$-m = (\mu - s\alpha)(h_\alpha) \quad \text{and} \quad m = (\mu + t\alpha)(h_\alpha)$$

for appropriate $s, t \in \mathbb{N}$. In particular $V_{\mu+k\alpha} = M_{\mu+k\alpha} \neq 0$ if and only if $k \in [-s, t]$. Especially when $t > 0$ we have $0 \neq xM_\mu \leq M_{\mu+\alpha} = V_{\mu+\alpha}$ (completing (a)).

We solve

$$-(\mu - s\alpha)(h_\alpha) = (\mu + t\alpha)(h_\alpha)$$

to find $(\mu, \alpha^\vee) = \mu(h_\alpha) = s - t$. If $(\mu, \alpha^\vee)$ is negative, then $t > 0$ and $\mu + \alpha \in \Phi_V$.

Finally

$$\mu - (\mu, \alpha^\vee) = \mu - (s - t)\alpha = \mu + (-s + t)\alpha$$

with $k = (-s + t) \in [-s, t]$, so $V_{\mu-(\mu,\alpha^\vee)} \neq 0$ and $\mu - (\mu, \alpha^\vee)\alpha \in \Phi_M \subseteq \Phi_V$. □

We now view $L$ as a finite dimensional module for itself and its various $\mathfrak{sl}_2(\mathbb{K})$ subalgebras (found in Theorem (6.11)).

**(8.2).** Theorem.

(a) *Let $\alpha, \beta \in \Phi$ with $\beta \neq \pm\alpha$, and let $\beta - s\alpha, \ldots, \beta, \ldots, \beta + t\alpha$ be the $\alpha$-string of roots through $\beta$. Then $\beta(h_\alpha) = (\beta, \alpha^\vee) = s - t \in \mathbb{Z}$.*

(b) *Let $\alpha, \beta, \alpha + \beta \in \Phi$. Then $[L_\alpha, L_\beta] = L_{\alpha+\beta}$.*

(c) *For $\alpha \in \Phi$ we have $\dim_{\mathbb{K}}(L_\alpha) = 1$ and $\Phi \cap \mathbb{K}\alpha = \{\pm\alpha\}$ in $H^*$.*

Proof. The first part follows immediately from the previous proposition when we consider the adjoint action of $L$ on itself. For the second part, Theorem (5.11) and the proposition give $0 \neq [L_\alpha, L_\beta] \leq L_{\alpha+\beta}$. Therefore once we have proven in (c) that all $L_\alpha$ have dimension 1, parts (a) and (b) will be complete.

The final part effectively comes from taking $\alpha = \mu$ in the proposition. As before, Theorem (6.11) provides us with $x \in L_\alpha$ and $y \in L_\alpha$ with $h_\alpha = [x, y]$ and $S = \mathbb{K}h_\alpha \oplus \mathbb{K}x \oplus \mathbb{K}y$ isomorphic to $\mathfrak{sl}_2(\mathbb{K})$.

Consider the subspace

$$N = \mathbb{K}y \oplus \mathbb{K}h_\alpha \oplus \bigoplus_{i \in \mathbb{Z}^+} L_{i\alpha}\,.$$

We claim that $N$ is an $S$-module. By Proposition (6.10)(c) we have $[x, \mathbb{K}y] = \mathbb{K}h_\alpha = [y, L_\alpha]$. To see that $N$ is $x$-invariant we check

$$[x, \mathbb{K}y] = \mathbb{K}h_\alpha\,,\ [x, \mathbb{K}h_\alpha] \leq L_\alpha,\ \text{and}\ [x, L_{(i-1)\alpha}] \leq L_{i\alpha} \ \text{ for } i > 1\,.$$

Similarly $N$ is a $\mathbb{K}y$-submodule as

$$[y, \mathbb{K}y] = 0\,,\ [y, \mathbb{K}h_\alpha] \le \mathbb{K}y\,,\ [y, L_\alpha] = \mathbb{K}h_\alpha\,,\ \text{and } L_{(i-1)\alpha} \ge [y, L_{i\alpha}] \ \text{ for } i > 1\,.$$

Therefore $N$ is an $S$-module. Its weights are

$$-\alpha(h_\alpha) = -2\,,\ 0(h_\alpha) = 0\,,\ \alpha(h_\alpha) = 2\,,$$

and

$$(i\alpha)(h_\alpha) = 2i \text{ when } i > 1 \text{ and } N_{i\alpha} = L_{i\alpha} \neq 0\,.$$

By Theorem (7.28)(a) we must have $N_{i\alpha} = L_{i\alpha} = 0$ for all $i > 1$. Furthermore $N_0 = \mathbb{K}h_\alpha$ has dimension 1, so by Theorem (7.28)(b) the module $N$ is irreducible and a copy of $M_0(3)$. (Indeed $N = \mathbb{K}y \oplus \mathbb{K}h_\alpha \oplus \mathbb{K}x = S$, the adjoint module with $S$ acting on itself.) Especially the dimension of $L_\alpha = N_2$ is 1 for every $\alpha \in \Phi$. □

**(8.3).** Theorem.

(a) *For $\alpha, \beta \in \Phi$ we have $(\beta, \alpha) \in \mathbb{Q}, (\beta, \alpha^\vee) = \frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$, and $\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha = \beta - (\beta, \alpha^\vee)\alpha \in \Phi$.*

(b) *The form $(\cdot, \cdot)$ is positive definite on $E_\mathbb{Q}$.*

(c) *Any $\mathbb{Q}$-basis of $E_\mathbb{Q}$ is a $\mathbb{K}$-basis of $H^*$.*

Proof. (a) We apply Proposition (8.1)(c,e) with $V = L$ and $\mu = \beta$. This gives (a) directly except for the claim that $(\beta, \alpha) \in \mathbb{Q}$.

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 the previous theorem). Especially for $\alpha \in \Phi$

$$0 < (\alpha, \alpha) = \sum_{\beta \in \Phi} \beta(t_\alpha)^2 = \sum_{\beta \in \Phi} (\beta, \alpha)^2$$

hence

$$\frac{4}{(\alpha, \alpha)} = \sum_{\beta \in \Phi} \left( \frac{2(\beta, \alpha)}{(\alpha, \alpha)} \right)^2 \in \mathbb{Z}^+\,.$$

Thus $(\alpha, \alpha) \in \mathbb{Q}$ and indeed

$$(\beta, \alpha) = \frac{(\alpha, \alpha)}{2} \cdot \frac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \mathbb{Q}\,.$$

(b) Let $\gamma = \sum_{\alpha \in \Phi} \gamma_\alpha \alpha \in E_\mathbb{Q}$ with $\gamma_\alpha \in \mathbb{Q}$. Then, as above,

$$(\gamma, \gamma) = \sum_{\beta \in \Phi} \beta(t_\gamma)^2 = \sum_{\beta \in \Phi} \left( \sum_{\alpha \in \Phi} \gamma_\alpha \beta(t_\alpha) \right)^2 = \sum_{\beta \in \Phi} \left( \sum_{\alpha \in \Phi} \gamma_\alpha (\beta, \alpha) \right)^2 \ge 0\,,$$

as it is a sum of rational squares by (a). Furthermore if $(\gamma,\gamma)=0$ then $\beta(t_\gamma)$ 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 (6.10), this in turn gives $t_\gamma\in H\cap H^\perp=0$ (by Proposition (6.7)), hence $\gamma=0$.

(c) 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 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 $H^*=\bigoplus_{j\in J}\mathbb{K}b_j\le h^\perp$. By Proposition (6.7) and the definition of our form, it is nondegenerate on $H^*$; so we must have $h=0$. But now in nondegenerate (indeed, positive definite) $E_\mathbb{Q}=\bigoplus_{i\in I}\mathbb{Q}b_i$ 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 $H^*$. □

## 8.2 Classification of root systems

This section and its notation are independent of the rest of the chapter.

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 $v$.

**(8.4).** LEMMA. *Let $0\neq v\in E$.*

(a) *$r_v\in \mathrm{O}(E)$, the orthogonal group of isometries of $E$.*

(b) *If $g\in\mathrm{O}(E)$ then $r_v^g=r_{g(v)}$.*

(c) *If $\mathbb{R}v^{r_x}=\mathbb{R}v$ if and only if $v\in\mathbb{R}x$ or $(v,x)=0$.* □

**(8.5).** DEFINITION. *Let $E$ be a finite dimensional real space equipped with a Euclidean positive 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 each $\alpha\in\Phi$ the reflection $r_\alpha\colon x\mapsto x-\frac{2(x,\alpha)}{(\alpha,\alpha)}\alpha$ takes $\Phi$ to itself;*

(iii) *for any $\alpha\in\Phi$ we have $\mathbb{R}\alpha\cap\Phi=\{\pm\alpha\}$;*

(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* $\dim_\mathbb{R}(E)$.

*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 let $\mathrm{W}(\Sigma)=\langle\, r_\alpha\mid\alpha\in\Sigma\,\rangle$.*

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:

**(8.6).** 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. This follows by Theorem (8.2)(a) and Theorem (8.3). □

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 $v_1, \ldots, v_n$ 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 $\Delta = \{\delta_1, \ldots, \delta_l\}$ be the set of simple roots in $\Phi^+$.

**(8.7).** THEOREM. *Let $(E, \Phi)$ be a root system with $\Delta = \{\delta_1, \ldots, \delta_l\}$ 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* $(\alpha, \beta) \leq 0$.

(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 induct on the index $k$. 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 $\beta$.

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 $\pm 1$. Without loss, we may assume $(\alpha, \beta^\vee)$ is $\pm 1$.

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 (8.5)(iii)). 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 $\{1, \ldots, l\}$ is the disjoint union of $I$ and $J$.

First

$$(x, x) = \left(\sum_{i \in I} d_i\delta_i, \sum_{j \in J} d'_j\delta_j\right) = \sum_{i \in I, j \in J} d_i d'_j(\delta_i, \delta_j) \leq 0$$

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. □

**(8.8).** 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 set $\Delta$ is the *simple basis* in $\Phi^+$ for $\Phi$ and $E$, uniquely determined by $\Phi^+$, and $l = n = \dim E$ is the *rank* of the system. We also 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 $\alpha$ is the integer $\mathrm{ht}(\alpha) = \sum_{i=1}^{l} d_i$, positive for positive roots and negative for negative roots.

In the root system $(E, \Phi)$ if $\Phi$ is the disjoint union of $F^+$ and $F^- = -F^+$ where, for each $\epsilon = \pm$, each sum of roots from $F^\epsilon$ is either not a root or is a root in $F^\epsilon$, then we say that $F^+$ is a *positive system* in $(E, \Phi)$. From the corollary, the basic example is $\Phi^+$. We next see that, up to the action of $\mathrm{W}(\Phi)$, this is the only example.

**(8.9).** Proposition.

(a) *For* $\delta \in \Delta$*, we have* $(\Phi^+ \setminus \delta)^{r_\delta} = \Phi^+ \setminus \delta$.

(b) *Let* $F^+$ *be a positive system and set* $\Delta_0 = F^+ \cap \Delta$*. If* $\delta \in \Delta \setminus \Delta_0$ *then* $(F^+)^{r_\delta}$ *is a positive system with* $(F^+)^{r_\delta} \cap \Delta = \{\delta\} \cup \Delta_0$.

(c) *For every positive system* $F^+$ *in* $(E, \Phi)$ *there is a* $w \in \mathrm{W}(\Phi)$ *with* $(F^+)^w = \Phi^+$.

(d) $(\Phi^+)^w = \Phi^-$ *if* $w = \prod_{i=1}^{l} r_{\delta_{\sigma(i)}}$ *for any permutation* $\sigma \in \mathrm{Sym}(l)$.

Proof. Sketch: (a) comes from Theorem (8.7)(b). This then gives (b) which in turn gives (c) (as $\Delta \subseteq F^+$ implies $\Phi^+ = F^+$) and (d). □

Thus the set $\Phi^+$ of positive roots in $\Phi$ is determined uniquely up to the action of the Weyl group. This in turn means that simple bases for $\Phi$ are all equivalent up to the action of the Weyl group. Conversely, each simple basis determines the Weyl group.

**(8.10).** Theorem. *Let* $(E, \Phi)$ *be a root system and* $\Delta = \{\delta_1, \ldots, \delta_l\}$ *a simple basis in* $\Phi^+$*. Then* $\mathrm{W}(\Phi) = \mathrm{W}(\Delta)$ *is a finite group with every element of* $\{\, r_\alpha \mid \alpha \in \Phi \,\}$ *conjugate to some element of* $\{\, r_\delta \mid \delta \in \Delta \,\}$.

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 $E$.

As $\alpha^{r_\alpha} = -\alpha$ and $r_\alpha^g = r_{\alpha^g}$ (as in Lemma (8.4)(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 (8.7)(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 $\mathrm{ht}(\alpha) > 1$, by Definition (8.5)(iii) 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 $w = r_{\delta_1} u \in \mathrm{W}(\Delta)$. □

**(8.11).** 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 $\frac{\pi}{m_{\alpha,\beta}}$.* □

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.

**(8.12).** 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*

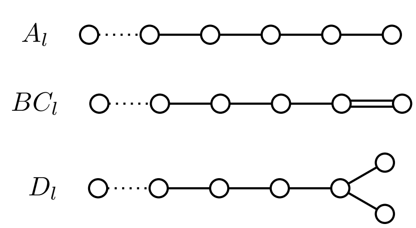
$$\mathrm{W}(\Phi) = \mathrm{W}(\Delta) = \mathrm{W}(\Sigma) \oplus \mathrm{W}(\Delta \setminus \Sigma) = \mathrm{W}(\Phi_\Sigma) \oplus \mathrm{W}(\Phi_{\Delta \setminus \Sigma})$$

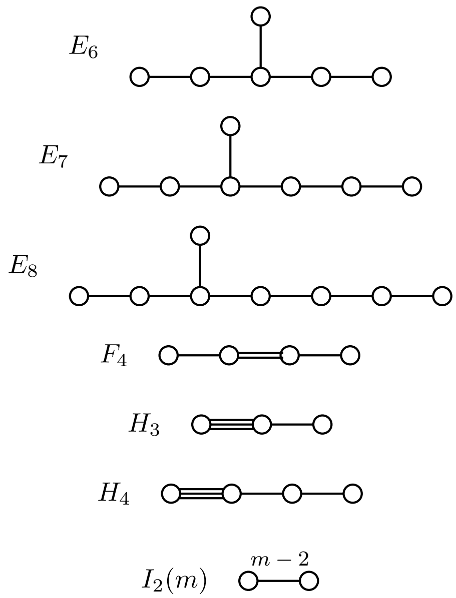
*where $\Phi_\Sigma = \Sigma^{\mathrm{W}(\Phi)} = \Sigma^{\mathrm{W}(\Sigma)}$ and $\Phi_{\Delta \setminus \Sigma} = (\Delta \setminus \Sigma)^{\mathrm{W}(\Phi)} = (\Delta \setminus \Sigma)^{\mathrm{W}(\Delta \setminus \Sigma)}$. Here $\Phi_\Sigma$ and $\Phi_{\Delta \setminus \Sigma}$ are perpendicular and have union $\Phi$.*

PROOF. This is an immediate consequence of Lemmas (8.4) and (8.11) and of Theorem (8.10). □

We repeat Theorem B-(2.3) from Appendix B.

**(8.13).** 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:*






□

It is not at all clear which Coxeter graphs actually correspond to root systems. The last two properties of root systems 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.

**(8.14).** 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

$$(\alpha,\beta^\vee)(\beta,\alpha^\vee) = \frac{4(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)} = 4\cos\left(\frac{\pi}{m_{\alpha,\beta}}\right)^2 \in \mathbb{Z}\,.$$

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

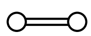
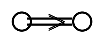
$$\frac{(\alpha, \beta^\vee)}{(\beta, \alpha^\vee)} = \frac{(\alpha, \alpha)}{(\beta, \beta)}\,.$$

This gives us the next two columns of the table, although in the first line we have no information about the squared length of $\alpha$.

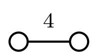
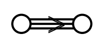
Finally as $(\beta, \beta) = 1$, we have

$$(\alpha, \beta) = \frac{1}{2}\left(\frac{2(\alpha, \beta)}{(\beta, \beta)}\right) = \frac{1}{2}(\alpha, \beta^\vee)\,. \qquad \square$$

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 the Dynkin diagram. All $A_1 \times A_1$ edges (that is, a nonedge) and $A_2$ edges (single bond) are left undisturbed. On the other hand

$BC_2$  becomes $B_2 = C_2$ 

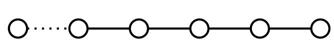
Similarly

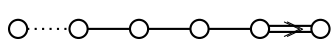
$I_2(6)$  becomes $G_2$ 

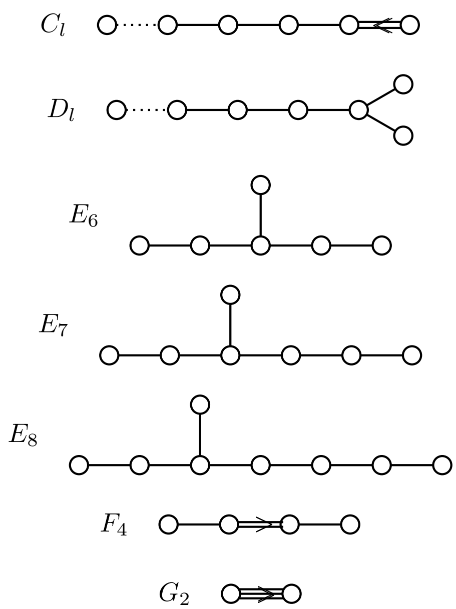
The arrow (or "greater than") sign on the edge is there to indicate that the root at the tail (or "big") end is longer than the root at the tip ("small") end. Also notice that $G_2$ has three bonds rather than 4. This change in notation indicates that the *long root* has squared length 3 times that of 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:

**(8.15).** THEOREM. *The Dynkin diagram for an irreducible abstract root system of rank $l$ is one of the following:*

$A_l$ 

$B_l$ 




□

It turns out that every Dynkin diagram in the theorem does come from a (unique up to equivalence) root system. We will deal with this existence issue in our discussion of existence of semisimple Lie algebras.

## 8.3 Semisimple algebras III: Uniqueness

We resume the notation of the introduction to the chapter. Additionally, in the root system $(E^L, \Phi^L) = (E, \Phi)$ we choose (as in the previous section) a partition $\Phi = \Phi^+ \cup \Phi^-$ associated with the simple basis $\Delta = \{\delta_1, \ldots, \delta_l\}$. The integers $(\alpha, \beta^\vee)$ with $\alpha, \beta \in \Phi$ are the *Cartan integers*. Then 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 $c_{i,j} = (\delta_i, \delta_j^\vee)$. All 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 (8.36)).

**(8.16).** 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* $\mathrm{Cart}(\Delta_2) = P\,\mathrm{Cart}(\Delta_1)P^\top$.

It is reasonably clear that (3) and (4) are equivalent and both imply (2). On the other hand (2) implies (3) and (4) by Proposition (8.9)(c).

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 (8.36), in an ad hoc and after-the-fact manner. See page 105 for the ultimate proof of the theorem.

At present we will deal with the crucial (2) $\implies$ (1) part of the theorem above:

**(8.17).** 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 isomorphic. 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 $(E', \Phi')$.*

Before proving this, we point out an interesting and helpful corollary.

**(8.18).** 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*.

**(8.19).** 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 (8.2)

$$\beta = \alpha - \delta_1 = (d_1 - 1)\delta_1 + \sum_{i=2}^{l} d_i\delta_i$$

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. □

**(8.20).** 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*

$$\sum_{a=1}^{b} \gamma_a \in \Phi^- \textit{ for each } 1 \le b \le -k \quad \textit{and} \quad \gamma = \sum_{a=1}^{-k} \gamma_a\,.$$

PROOF. Set $\alpha = -\gamma$ and then $\gamma_a = -\alpha_a$. □

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 (6.11) 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.

**(8.21).** PROPOSITION. *The Lie algebra $L$ is generated by the elements $h_i, e_i, e_{-i}$ for $1 \le i \le l$. We have the following relations in $L$:*

(a) $[h_i, h_j] = 0$ *for all* $1 \le i, j \le l$;

(b) $[h_i, e_j] = c_{j,i} e_j$ *and* $[h_i, e_{-j}] = -c_{j,i} e_{-j}$ *for all* $1 \le i, j \le l$;

(c) $[e_i, e_{-i}] = h_i$ *for all* $1 \le i \le l$;

(d) $[e_i, e_{-j}] = 0$ *for all* $i \ne j$;

(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* $1 \le i, j \le l$ *with* $i \ne j$.

PROOF. We have the Cartan decomposition

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

By Proposition (6.10) and Theorem (8.7)(c), the $L_\alpha$ have dimension 1 and the $h_i$ generate $H$. By Theorem (5.11) always $[L_\alpha, L_\beta] \le L_{\alpha+\beta}$ for $\alpha, \beta \in \Phi$. As all the $L_\alpha$ have dimension 1, this is true with equality by Theorem (8.2). Therefore by induction on the height of $\gamma \in \Phi$ and using the previous proposition and its corollary, we find that $L_\gamma$ is in the subalgebra generated by the various $h_i, e_i, e_{-i}$. That subalgebra is therefore $L$ itself.

Parts (a) and (c) 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 (b) 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 (e), 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 (8.2). 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 prove this difficult theorem, but we do use its relations (from the proposition) as the entry to our uniqueness proof for $L$.

**(8.22).** 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 (8.21) give a presentation of a semisimple Lie algebra $L$ over $\mathbb{K}$ with Cartan matrix $C$ and root system equivalent to $(E, \Phi)$.* $\square\square$

Our uniqueness proof is motivated by that of [Eld15]. The basic observation is that, with respect to the Cartan basis $\{\, h_i, e_\alpha \mid 1 \le i \le l\,,\ \alpha \in \Phi \,\}$, most of the adjoint actions are nearly monomial. We then show (starting as in the proposition) 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.

**(8.23).** 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 (8.21). Then for $x \in L_\beta$ we have $[e_\delta, [e_{-\delta}, x]] = t(s+1)x$.*

Proof. In the notation of Chapter 7 (for instance Definition (7.16)), we may take $x = v_i$ with $i = t$ and $\lambda = m = s + t$ so that the coefficient $i(\lambda - i + 1) = t(s + t - t + 1) = t(s + 1)$. $\square$

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 (8.25).

For each $\delta \in \Delta \cup \Delta$, set $a_\delta = \mathrm{ad}_{e_\delta}$. Consider words $w = w_k \ldots w_1$ 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 (8.19) 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 $w_\alpha$.

**(8.24).** 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 (8.23) 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 $\gamma = \delta$. □

**(8.25).** 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 $[e_\alpha, e_\beta] = 0e_\alpha$, 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 (8.17).

The isomorphism between the root systems $(E, \Phi)$ and $(E', \Phi')$ gives rise (by Proposition (8.9)(c)) 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 Semisimple algebras IV: Existence

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 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.[2] The philosophy is always that in a classification one should easily reduce to the irreducible case. This remains true with our semisimple Lie algebras.

**(8.26).** 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 (6.6).

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 (8.2)), the ideal $I$ is the direct sum of its intersection $H \cap I$ and the $L_\lambda$ 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 (6.6) 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 (8.2). Thus $\delta$ and $\gamma$ are not connected in the Dynkin diagram of $\Delta$ by Proposition (8.14). That is, $\Delta_I$ is a union of irreducible components of $\Delta$.

Conversely, suppose that $\Sigma$ is an irreducible component of $\Delta$ and hence of the corresponding Coxeter graph. Then by Lemma (8.12) 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 ("span") ideals perpendicular to all others. In particular, the simple ideals are in bijection with the irreducible components of the Dynkin diagram. □

**(8.27).** 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 $\mathrm{SL}_{l+1}(\mathbb{K}) \leq \mathrm{Aut}(L)$ to $H$, the abelian and dimension $l$ subalgebra of all diagonal matrices with trace $0$.*

[2] Luckily.

(c) *The $H$-root spaces are the various $\mathbb{K}e_{i,j}$ for $1 \leq i \neq j \leq l+1$ with corresponding root $\varepsilon_i - \varepsilon_j$ in the Euclidean space $\mathbb{R}^{l+1} \cap \mathbf{1}^\perp \simeq \mathbb{R}^l$.*

(d) *The simple roots of $\Delta$ are $\delta_i = \varepsilon_i - \varepsilon_{i+1} = \delta_i^\vee$ for $1 \leq i \leq l$, and so the Dynkin diagram is $A_l$.*

(e) *The Weyl reflection $r_{\varepsilon_i - \varepsilon_j}$ induces on $\mathbb{R}^l \leq \mathbb{R}^{l+1}$ 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 $H \oplus \bigoplus_{i \neq j} \mathbb{K}e_{i,j}$. The rest of this part then follows from (d) and Theorem (8.26).

(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 (8.2)(c) we have $V = V^w$, 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 $l \geq 1$) 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 $V^* = \mathbb{R}^{l+1}$ be $\varepsilon_i, \ldots, \varepsilon_{l+1}$, we find $\alpha(h) = (\varepsilon_i - \varepsilon_j)(h)$; that is, $\alpha = \varepsilon_i - \varepsilon_j$ from 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 $1 \leq k \leq l+1$

$$\begin{aligned} r_{\varepsilon_i - \varepsilon_j}(\varepsilon_k) &= \varepsilon_k - (\varepsilon_k, (\varepsilon_i - \varepsilon_j)^\vee)(\varepsilon_i - \varepsilon_j) \\ &= \varepsilon_k - (\varepsilon_k, \varepsilon_i - \varepsilon_j)(\varepsilon_i - \varepsilon_j)\,. \end{aligned}$$

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. □

**(8.28).** 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 its action on $V$, $L$ has 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 (8.2)(c) we have $V = V^w$ for all choices of Cartan subalgebra $C$.

Let $b$ be the nondegenerate $(\mathrm{Id}, \eta)$-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 $(V_{C,\lambda}, V_{C,-\lambda}) \neq 0$.

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-(1.4) (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. □

**(8.29).** 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 $(\mathrm{Id}, \eta)$-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 $1 \leq i \leq l$.*

(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 $H^* \simeq \mathbb{R}^l$. 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 $1 \leq i \leq l-1$ 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 $1 \leq i \leq l$.*

(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* |
|---|---|
| $e_{i,j} - e_{-j,-i}$ | $\varepsilon_i - \varepsilon_j$ |
| $e_{-i,-j} - e_{j,i}$ | $-(\varepsilon_i - \varepsilon_j)$ |
| $e_{i,-j} + e_{j,-i}$ | $\varepsilon_i + \varepsilon_j$ |
| $e_{-i,j} + e_{-j,i}$ | $-(\varepsilon_i + \varepsilon_j)$ |
| $e_{i,-i}$ | $2\varepsilon_i$ |
| $e_{-i,i}$ | $-2\varepsilon_i$ |

(d) *The simple roots of $\Delta$ are $\delta_i = \varepsilon_i - \varepsilon_{i+1} = \delta_i^\vee$ for $1 \leq i \leq l-1$ 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 $1 \le i \le j \le l$:

$$\begin{aligned} a_{i,j} &= -a_{-j,-i} \\ a_{-i,-j} &= -a_{j,i} \\ b_{i,-j} &= -\eta b_{j,-i} \\ b_{-i,j} &= -\eta b_{-j,i}\,. \end{aligned}$$

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 $1 \le i < j \le l$,

$$e_{i,j} - e_{-j,-i}\,, \quad e_{-i,-j} - e_{j,i}\,, \quad e_{i,-j} - \eta e_{j,-i}\,, \quad e_{-i,j} - \eta e_{-j,i}\,.$$

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 $1 \le i \le l$. 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 basis elements being $e_{i,-i}$ and $e_{-i,i}$ for $1 \le i \le l$.

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 (8.26).

(b) The calculations of (a) reveal $L_\eta$ to be irreducible on $V$, so by Theorem (8.28) 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 (8.27)(e), and the reflection in $2\varepsilon_l$ is clear. All that needs checking is

$$r_{\varepsilon_{l-1}+\varepsilon_l}(\varepsilon_{l-1}) = \varepsilon_{l-1} - (\varepsilon_{l-1}, \varepsilon_{l-1} + \varepsilon_l)(\varepsilon_{l-1} + \varepsilon_l) = \varepsilon_{l-1} - (\varepsilon_{l-1} + \varepsilon_l) = -\varepsilon_l\,. \qquad \square$$

**(8.30).** Example. *Let the $\mathbb{K}$-space $V = V_\eta = \mathbb{K}^{2l+1}$ 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 $2^l{:}\,\mathrm{Sym}(l)$.*

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 (8.29)(i).

Thus a basis for $L$ is that for $\mathfrak{so}_{2l}(\mathbb{K})$ from Example (8.29)(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 (8.26) as before.

(b) Again by Theorem (8.28) 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 (8.29)(ii). □

In the classical examples above we have seen the following useful observation in action.

**(8.31).** THEOREM. *If $(E^L, \Phi^L)$ has rank $l$, then* $\dim_{\mathbb{K}}(L) = l + |\Phi|$. □

For instance, a Lie algebra of type $\mathfrak{g}_2$ must have dimension $2 + 12 = 14$.

**(8.32).** 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. SKETCH: The circular symmetry of the Dynkin diagram $D_4$ shows, with Corollary (8.18), that $L = \mathfrak{d}_4(\mathbb{K})$ has an automorphism of order 3. Its fixed point algebra $M$ contains a proper $\mathfrak{sl}_2(\mathbb{K})$ subalgebra corresponding to the central node of the diagram, so $M$ has a nontrivial semisimple section of dimension greater than three and less than $28 = \dim_{\mathbb{K}}(L)$. There are few possibilities, and in the end it must be $\mathfrak{g}_2(\mathbb{K})$. This should be verified by examination of the action of the element of order three and in particular its fixed weights $\alpha_1 = \varepsilon_2 - \varepsilon_3$ and

$$\alpha_2 = (\varepsilon_1 - \varepsilon_2) + (\varepsilon_3 - \varepsilon_4) + (\varepsilon_3 + \varepsilon_4) = \varepsilon_1 - \varepsilon_2 + 2\varepsilon_3\,,$$

which form a simple basis for a root system of type $G_2$ in $\mathbb{R}^3 \cap (-1,1,1)^\perp$. □

**(8.33).** PROPOSITION.

(a) *If $\mathfrak{e}_8(\mathbb{K})$ exists, then it has a proper subalgebra $\mathfrak{e}_7(\mathbb{K})$.*

(b) *If $\mathfrak{e}_7(\mathbb{K})$ exists, then it has a proper subalgebra $\mathfrak{e}_6(\mathbb{K})$.*

(c) *If $\mathfrak{e}_6(\mathbb{K})$ exists, then it has a proper subalgebra $\mathfrak{f}_4(\mathbb{K})$.*

PROOF. The first two parts are clear from the Dynkin diagram.

SKETCH: The last part follows as the Dynkin diagram $E_6$ has an automorphism of order 2 which by Corollary (8.18) extends to an automorphism of $\mathfrak{e}_6(\mathbb{K})$. Its fixed point subalgebra contains a subalgebra $\mathfrak{f}_4(\mathbb{K})$. (Compare with Theorem (8.32).) □

We leave unproven:

**(8.34).** THEOREM.

(a) *The Lie algebra* $\mathfrak{e}_8(\mathbb{K})$ *exists and has dimension* 248.

(b) *The Lie algebras* $\mathfrak{e}_7(\mathbb{K})$, $\mathfrak{e}_6(\mathbb{K})$, *and* $\mathfrak{f}_4(\mathbb{K})$ *have respective dimensions* 133, 78, *and* 52. □□

## 8.5 Semisimple algebras V: Classification

We now can prove almost all of:

**(8.35).** 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 (8.15) and (8.26). (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 (8.17).

In the four classical cases, each exists by Examples (8.27), (8.29), and (8.30) 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 (by dimension considerations) or from any other case, since all Cartan subalgebras are conjugate under the corresponding automorphism groups.

The rank 2 algebra $\mathfrak{g}_2(\mathbb{K})$ exists and has dimension 14 by Theorem (8.32).

Leaving aside existence and dimension for the moment, 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 can not in any case be isomorphic to each other by Proposition (8.33). Furthermore, none is isomorphic to a classical algebra (as mentioned above) or to $\mathfrak{g}_2(\mathbb{K})$ of dimension 14, since the smallest, namely $\mathfrak{f}_4(\mathbb{K})$, contains two disjoint subalgebras $\mathfrak{sl}_3(\mathbb{K}) = \mathfrak{a}_2(\mathbb{K})$ and so has dimension at least $8 + 8 = 16$.

The actual existence and dimensions for the algebras $\mathfrak{e}_8(\mathbb{K})$, $\mathfrak{e}_7(\mathbb{K})$, $\mathfrak{e}_6(\mathbb{K})$, and $\mathfrak{f}_4(\mathbb{K})$ are contained in (our only unproven result) Theorem (8.34). □

As mentioned in the proof, the only parts of the theorem that we have not proven are those from Theorem (8.34). For the following corollary that theorem is not necessary as Proposition (8.33) suffices.

**(8.36).** 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 (8.17) 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 (8.16).

Directly after the statement of the theorem we observed that its parts (2), (3), and (4) are equivalent. Theorem (8.17) was then devoted to proving 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

**(8.37).** Problem. *Prove: $\Phi^\vee$ is a root system with simple basis $\Delta^\vee$ and $\mathrm{W}(\Phi^\vee) = \mathrm{W}(\Phi)$.*

**(8.38).** 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 (8.2).*

**(8.39).** PROBLEM. *Totally positive word or totally negative word is the same as minimal word.*

**(8.40).** PROBLEM. *Highest root.* $\Phi^+ \longrightarrow \Phi^-$.

Chapter 9

# Representations of semisimple algebras

Again $L\,(\neq\, 0)$ will be a finite dimensional, semisimple Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic 0. Now we study the representation theory of $L$. We take our lead from Chapter 7 by looking at properties satisfied by finite dimensional irreducible modules and then studying cyclic modules that are not necessarily finite dimensional or irreducible but do possess some of those properties.

Starting in Section 9.2 we continue the notation and terminology detailed in the introduction to Chapter 8.

## 9.1 Universal enveloping algebras

In Section 4.2 we introduced representation of Lie algebras in extrinsic and intrinsic form. Starting from that we introduce another point of view.

Let $V$ be an $\mathbb{E}$-space $V$. For $f \in \mathbb{N}$, let $V^{\otimes f}$ be the $f^{\mathrm{th}}$ tensor power of the module $V$ (with $V^{\otimes 0} = \mathbb{K}$). The *tensor algebra* $\mathrm{T}(V)$ is the associative $\mathbb{E}$-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 $M$ over $\mathbb{E}$, then its *universal enveloping algebra* $\mathrm{U}(M)$ is the quotient $\mathrm{T}(M)/I$ where $I$ is the ideal in $\mathrm{T}(M)$ generated by all the elements $x \otimes y - y \otimes x - [x, y]1$ for $x, y \in M$. The construction gives us a natural representation $\upsilon_M \colon M \longrightarrow \mathrm{U}(M)^-$. Especially $\mathrm{U}(M)$-modules are

$M$-modules. This correspondence is readily seen to be universal in at least two senses.

**(9.1).** THEOREM.

(a) *If* $\varphi\colon M \longrightarrow \mathrm{End}^-_{\mathbb{E}}(V)$ *is a representation of* $M$*, then there is a unique associative algebra morphism* $\varphi_V\colon \mathrm{U}(M) \longrightarrow \mathrm{End}_{\mathbb{E}}(V)$ *with* $\varphi = \varphi_V v_M$*.*

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

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 $M$- and $\mathrm{U}(M)$-module is a quotient of $\mathrm{U}(M)$. This is an improvement. For instance in Chapter 7 we found that 3-dimensional $\mathfrak{sl}_2(\mathbb{K})$ has irreducible modules of arbitrarily large finite dimension as well as infinite dimensional irreducibles. (Among other things, this implies that the universal enveloping algebra for $\mathfrak{sl}_2(\mathbb{K})$ is infinite dimensional.)

Therefore to study $M$-modules, we begin with a more detailed study of $\mathrm{U}(M)$.

**(9.2).** THEOREM. (POINCARÉ-BIRKHOFF-WITT THEOREM) *Let the Lie algebra* $M$ *have the* $\mathbb{E}$*-basis* $\{\, v_i \mid i \in I \,\}$ *for some totally ordered set* $(I, \leq)$*.*

(a) (WEAK PBW) *The universal enveloping algebra* $\mathrm{U}(M)$ *has as* $\mathbb{E}$*-spanning set the collection of all monomials* $v_{i_1} \cdots v_{i_n}$ *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* $\mathrm{U}(M)$ *has as* $\mathbb{E}$*-basis the collection of all monomials* $v_{i_1} \cdots v_{i_n}$ *for* $n \in \mathbb{N}$ *and* $i_i \leq \cdots \leq i_n$ *(where* $n = 0$ *corresponds to the monomial* $1$*)*

PROOF. Weak PBW follows easily by induction from the fact that

$$v_i v_j = v_j v_i - [v_i, v_j]\,.$$

Strong PBW is much harder. There are (at least) two standard proofs. We prefer that of Serre [Ser06]. (But we do not give it here.) □

For many applications Weak PBW is enough, but there are places where Strong PBW is unavoidable.

**(9.3).** COROLLARY.

(a) *The representation* $v_M\colon M \longrightarrow \mathrm{U}(M)^-$ *is faithful.*

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

PROOF. The first part follows from the Strong PBW Theorem, and the second part follows from the first. □

We encountered the second part early in these notes as Theorem (1.6)(a). As mentioned above, the faithful representation guaranteed by (a) may well be infinite dimensional, even when $M$ has finite dimension. For semisimple $M$,the adjoint representation suffices for (b) (and so, by universality, for (a) as well), since the kernel of the adjoint representation is the solvable ideal $\mathrm{Z}(M)$, which is 0 for semisimple $M$. This was also noted earlier in Theorem (1.5).

## 9.2 Finite dimensional modules, highest weights

We return to the notation and terminology detailed in the introduction to Chapter 8.

Recall that a *weight module* for $L$ is a module $V$ that is spanned by its weight vectors $V_\mu^w$ or, equivalently by Proposition (8.1)(a), is generated as $L$-module by its weight vectors.

**(9.4).** PROPOSITION. *For the finite dimensional cyclic $L$-module $V = Lv$ with $0 \neq v \in V_\mu^w$, the following are equivalent:*

(1) $L^+v = 0$.

(2) $\mu + \alpha \notin \Phi_V$ *for all* $\alpha \in \Phi^+$.

(3) $\mu + \delta \notin \Phi_V$ *for all* $\delta \in \Delta$.

PROOF. Statement (1) implies (2) by Proposition (8.1)(a), and (2) certainly implies (3). Finally, (3) implies (1) since the subalgebra $L^+$ is generated by the weight spaces $L_\delta$ for $\delta \in \Delta$ by Theorem (8.2) and Proposition (8.19). □

For the cyclic $L$-module $V = Lv$ with $0 \neq v \in V_\lambda^w$, if $L^+v = 0$, then $\lambda$ is a *highest weight*, $v$ is a *highest weight vector*, and the module $V$ is a *highest weight module*. In general, for nonzero $w$ in the $L$-module $W$, if $B^+w \leq \mathbb{K}w$, then the vector $w$ is a *maximal vector*, so it is a highest weight vector in the cyclic submodule of $W$ that it generates.

If $(\lambda, \delta^\vee) \in \mathbb{Z}$ for all $\delta \in \Delta$, then $\lambda$ is an *integral weight*. By Theorem (8.3)(a) all roots are integral weights, but there are others. If for all $\delta \in \Delta$ we have $(\lambda, \delta^\vee) \geq 0$ then $\lambda$ is a *dominant weight*.

**(9.5).** THEOREM. *Every finite dimensional $L$-module $V$ contains a maximal vector $v$ for some dominant integral weight $\lambda$. Especially, if finite dimensional $V$ is irreducible then $V$ is a weight module generated by a maximal vector $v$ with dominant integral highest weight $\lambda$.*

PROOF. Every weight in $\Phi_V$ is integral by (8.1)(d). Choose among the finitely many members of $\Phi_V$ a weight $\lambda$ that maximizes $\sum_{\delta\in\Delta}(\lambda, \delta^\vee)$. Then $\lambda$ is dominant integral, and a weight vector $v$ for it is maximal. If $V$ is irreducible, then it is the cyclic module generated by $v$, so $\lambda$ is a highest weight. □

**(9.6).** PROPOSITION. *Let $\Phi^+ = \{\alpha_1, \ldots, \alpha_N\}$ for $N = |\Phi^+|$. Let $V$ be a highest weight module generated by the nonzero maximal vector $v^+$ for the weight $\lambda$.*

(a) *$V$ is spanned by the various vectors $f_{i_1}\cdots f_{i_j}\cdots f_{i_n}v^+$ for $n\in\mathbb{N}$ and $i_1\le i_2\le\cdots\le i_n$. This vector belongs to $V_{\lambda-\sum_{j=1}^n \alpha_{i_j}}$.*

(b) *For every weight $\mu\in\Phi_V$ we have $\dim_{\mathbb{K}}(V_\mu)<\infty$. Especially $V_\lambda=\mathbb{K}v^+$ of dimension 1.*

(c) *Every quotient of $V$ is a highest weight module for $\lambda$. Every submodule of $V$ is a weight module.*

(d) *$V$ has a unique maximal submodule and a unique irreducible quotient.*

PROOF.

(a) SKETCH: This follows from the PBW Theorem (9.2) and Proposition (9.7). (Weak PBW is actually enough for this.)

(b) For a fixed $\lambda$ and $\mu$ the number of solutions to

$$\mu=\lambda-\sum_{j=1}^n \alpha_{i_j}$$

is finite, so this follows from (a). Furthermore $\sum_{j=1}^n \alpha_{i_j}=0$ gives the only solution when $\mu=\lambda$.

(c) A quotient of a highest weight module is a highest weight module. A submodule of a weight module is a weight module.

(d) $V$ is cyclic, generated by the weight vector $v$. As $V$ is a weight module, every proper submodule is contained in $\bigoplus_{\lambda\neq\mu\in\Phi_V} V_\mu$. Thus they generate the unique maximal proper submodule (still contained in this subspace), which is then the kernel for the unique irreducible quotient. □

## 9.3 Verma modules and weight lattices

Set $L^-$ be $\bigoplus_{\alpha\in\Phi^+} L_{-\alpha}$, further $L^+=\bigoplus_{\alpha\in\Phi^+} L_\alpha$, and $B^+=H\oplus L^+$, all three subalgebras of $L$ by Theorem (5.11).

The Strong PBW Theorem (9.2)(b) and the Cartan decomposition $L=L^-\oplus B^+$ provide a useful tensor factorization of the universal algebra $\mathrm{U}(L)$:

**(9.7).** PROPOSITION. *Then $\mathrm{U}(L)=\mathrm{U}(L^-)\otimes_{\mathbb{K}}\mathrm{U}(B^+)$.* □

For each $\lambda\in H^*$, let the associated 1-dimensional $B^+$- and $\mathrm{U}(B^+)$-module be $\mathbb{K}v_\lambda$ with $(h+u)v_\lambda=\lambda(h)v_\lambda$ for $h\in H$ and $u\in L^+$. The *Verma module* $M(\lambda)$ is then $\mathrm{U}(B^+)$-module $\mathbb{K}v^+$ induced up to the $\mathrm{U}(L)$- and $L$-module

$$M(\lambda)=\mathrm{U}(L)\otimes_{\mathrm{U}(B^+)}\mathbb{K}v_\lambda=\mathrm{U}(L^-)\otimes_{\mathbb{K}}\mathbb{K}v_\lambda\,,$$

where we have applied Proposition (9.7).

**(9.8).** THEOREM.

(a) *For* $\lambda \in H^*$*, the Verma module* $M(\lambda)$ *and its unique irreducible quotient* $L(\lambda)$ *are nonzero highest weight modules for* $\lambda$*. If* $\lambda \neq \mu$ *then* $L(\lambda)$ *and* $L(\mu)$ *are not isomorphic.*

(b) *Let* $V$ *be an* $L$*-module generated by the highest weight vector* $v^+$ *for* $\lambda$*. Then the map* $v_\lambda \mapsto v^+$ *extends to a surjective map from* $M(\lambda)$ *to* $V$*.* □

A *lattice* in the Euclidean space $E$ is the $\mathbb{Z}$-module spanned by some basis. The *root lattice* associated with the root system $(E^L, \Phi^L) = (E, L)$ is

$$\Lambda_R = \sum_{\alpha \in \Phi} \mathbb{Z}\alpha = \bigoplus_{i=1}^{l} \mathbb{Z}\delta_i \,.$$

The associated *integral weight lattice* or just *weight lattice* is

$$\Lambda_W = \{\, \mu \in E \mid (\mu, \delta_i^\vee) \in \mathbb{Z}\,,\ 1 \le i \le l \,\}\,.$$

It gets its name from the fact that every weight of a finite dimensional $L$-module belongs to the weight lattice by Proposition (8.1)(d). Especially $\Lambda_R \le \Lambda_W$.

A $\mathbb{Z}$-basis for $\Lambda_W$ is provided by the *fundamental weights* $\omega_i$, for $1 \le i \le l$, defined by

$$(\omega_i, \delta_i^\vee) = 1 \quad \text{and} \quad (\omega_i, \delta_j^\vee) = 0 \text{ for } i \neq j$$

so that $\Lambda_W = \oplus_{i=1}^{l} \mathbb{Z}\omega_i$. The *integral dominant weights* or just *dominant weights* are then those of $\Lambda_W^+ = \oplus_{i=1}^{l} \mathbb{N}\omega_i$ .

It is initially of some concern that the root lattice's natural home is $H^* (\simeq \mathbb{K}^l)$ while the weight lattice is defined within $E \simeq \mathbb{R}^l$. The next result obviates that worry by showing that $\Lambda_W \le E_{\mathbb{Q}} = \oplus_{i=1}^{l} \mathbb{Q}\delta_i$, where $E_{\mathbb{Q}}$ is a rational subspace naturally contained in $H^*$ and in $E = \mathbb{K} \otimes_{\mathbb{Q}} E_{\mathbb{Q}}$, both by definition.

**(9.9).** PROPOSITION.

(a) $\Lambda_W \le \mathbb{Z}[d^{-1}] \otimes_{\mathbb{Z}} \Lambda_R \le E_{\mathbb{Q}}$ *where* $d = \det(\mathrm{Cart}\,\Delta)$.

(b) $\delta_i = \sum_{j=1}^{l} (\delta_i, \delta_j^\vee)\omega_j$.

PROOF. Write each $\omega_i \in E$ as a linear combination of the simple basis $\Delta = \{\ldots, \delta_k, \ldots\}$: $\omega_i = \sum_{k=1}^{l} a_{i,k}\delta_k$ . Then

$$(\omega_i, \delta_j^\vee) = \left( \sum_{k=1}^{l} a_{i,k}\delta_k, \delta_j^\vee \right) = \sum_{k=1}^{l} a_{i,k}(\delta_k, \delta_j^\vee)\,.$$

Let the matrix $A = (a_{i,k})_{i,k}$ and the Cartan matrix $C = \mathrm{Cart}\,\Delta = (\delta_k, \delta_j^\vee)_{k,j}$. From the definition of the fundamental weights we get $I = AC$ hence $A = C^{-1}$. Thus by Cramer's Rule each $a_{i,k}$ belongs to $\mathbb{Z}[d^{-1}]$, as claimed.

Alternatively, when we write $\delta_i = \sum_{k=1}^{l} c_{i,k}\omega_k$ we find

$$(\delta_i, \delta_j^\vee) = \left( \sum_{k=1}^{l} c_{i,k}\omega_k, \delta_j^\vee \right) = \sum_{k=1}^{l} c_{i,k}(\omega_k, \delta_j^\vee) = c_{i,j}\,. \qquad \square$$

As the $c_{i,j}$ are all nonpositive, this implies that simple roots are rarely fundamental weights.

**(9.10).** THEOREM.
(CLASSIFICATION OF IRREDUCIBLE HIGHEST WEIGHT MODULES)

(a) *For every $\lambda \in H^*$, up to isomorphism there is a unique irreducible $L$-module $L(\lambda)$ with highest weight $\lambda$. These are nonzero and pairwise nonisomorphic.*

(b) *If $L(\lambda)$ is of finite dimension then $\lambda \in \Lambda_W^+$.*

(c) *If $\lambda \in \Lambda_W^+$ then $L(\lambda)$ is of finite dimension.*

PROOF. (a) This follows directly from Theorem (9.8).
(b) This is contained in Theorem (9.5).
(c) We postpone discussion of this to the next section. $\square$

So, for instance, when $L = \mathfrak{sl}_2(\mathbb{K}) = \mathfrak{a}_1(\mathbb{K})$ where $l = 1$, we have $\Delta = \{\delta_1\} = \{2\}$ and $\Lambda_R = 2\mathbb{Z}$. Thus $\delta_1^\vee = 1 = \omega_1$ and $\Lambda_W = \mathbb{Z} \leq \mathbb{Q} \leq \mathbb{K}$ with $\Lambda_W^+ = \mathbb{N}$. Now compare the current theorem with Theorem (7.21), where we catalogued the highest weight irreducible $\mathfrak{sl}_2(\mathbb{K})$-modules—there is (up to isomorphism) exactly one $L_+(\lambda) = L(\lambda)$ for every $\lambda \in \mathbb{K} \simeq H^*$ and it has finite dimension if and only if $\lambda \in \mathbb{N} = \Lambda_W^+$.

## 9.4 Tensor products of modules

It is well-known, and easy to check, that if $A$ and $B$ are associative $\mathbb{E}$-algebras and $V$ and $W$ are, respectively, unital $A$- and $B$-modules, then $V \otimes_{\mathbb{E}} W$ is naturally a unital $A \otimes_{\mathbb{E}} B$-module via

$$(a \otimes b)(v \otimes w) = av \otimes bw\,.$$

Furthermore, if $V$ and $W$ are irreducible, then so is $V \otimes_{\mathbb{E}} W$. (Exercise.)

In the special case $A = B$, we get representations of $A \otimes_{\mathbb{E}} A$ from representations of $A$. We would hope then to use this to get new representations of $A$ itself. As $A$ is an algebra, we already have the natural multiplication map

$$\mu\colon A \otimes_{\mathbb{E}} A \longrightarrow A$$

but the arrow goes in the wrong direction for us to get $A$-modules from $A \otimes_{\mathbb{E}} A$-modules. Instead we need a *comultiplication*

$$\nu\colon A \longrightarrow A \otimes_{\mathbb{E}} A\,;$$

that is, an $\mathbb{E}$-algebra map from $A$ into its tensor square $A^{\otimes 2}$. Then the $A \otimes_{\mathbb{E}} A$-module $V \otimes W$ is a $\nu(A)$-module by restriction and so an $A$-module. This is unlikely to be irreducible even if $V$ and $W$ are, and to some extent this is the point: decomposition of the reducible $A$-module $V \otimes W$ will often give rise to new irreducible $A$-modules.[1]

This is the case with groups, where the natural embedding of $G$ on the diagonal of $G \times G$ leads via $g \mapsto g \otimes g$ to the group algebra comultiplication

$$\mathbb{E}G \longrightarrow \mathbb{E}(G \times G) \simeq \mathbb{E}G \otimes_{\mathbb{E}} \mathbb{E}G\,.$$

Therefore, when $V$ and $W$ are $G$-modules, $V \otimes_{\mathbb{E}} W$ is also naturally a $G$-module. Abstraction from these observations leads to the study of *Hopf algebras* and related classes of algebras where the representation theory is very rich. Group algebras are the most fundamental examples of Hopf algebras.

Luckily for us, the universal enveloping algebra of a Lie algebras is also a Hopf algebra. In particular it has a suitable comultiplication.[2]

**(9.11).** THEOREM. *Let $A$ be a Lie $\mathbb{E}$-algebra. Then the map*

$$A \longrightarrow \mathrm{U}(A) \otimes_{\mathbb{E}} \mathrm{U}(A)$$

*given by $x \mapsto x \otimes 1 + 1 \otimes x$ is an injective Lie algebra mapping of $A$ into $(\mathrm{U}(A) \otimes_{\mathbb{E}} \mathrm{U}(A))^-$.*

PROOF. This is clearly an linear transformation, injective by the PBW Theorem (9.2). We then check

$$\begin{aligned}
&[x\otimes 1 + 1 \otimes x, y \otimes 1 + 1 \otimes y] = \\
&\quad = (x \otimes 1 + 1 \otimes x)(y \otimes 1 + 1 \otimes y) - (y \otimes 1 + 1 \otimes y)(x \otimes 1 + 1 \otimes x) \\
&\quad = (xy \otimes 1 + y \otimes x + x \otimes y + 1 \otimes xy) - (yx \otimes 1 + x \otimes y + y \otimes x + 1 \otimes yx) \\
&\quad = (xy \otimes 1 + 1 \otimes xy) - (yx \otimes 1 + 1 \otimes yx) \\
&\quad = (xy - yx) \otimes 1 + 1 \otimes (xy - yx)\,. \qquad \square
\end{aligned}$$

The construction can be motivated by first considering the diagonal comultiplication for the group algebra of a Lie group and then translating that into the Lie algebra context, using derivatives:

$$\begin{aligned}
\exp(tx) \mapsto \varphi(t) &= \exp(tx) \otimes \exp(tx) \\
&= (1 + tx + \cdots) \otimes (1 + tx + \cdots) \\
&= 1 \otimes 1 \,+\, tx \otimes 1 \,+\, 1 \otimes tx \,+\, tx \otimes tx \,+ \cdots \\
&= 1 \otimes 1 + t(x \otimes 1 + 1 \otimes x) + t^2(\cdots)
\end{aligned}$$

[1] The trivial comultiplications $a \mapsto a \otimes 1$ and $a \mapsto 1 \otimes a$ merely recover $V$ and $W$.

[2] A Hopf algebra is equipped not only with an associative multiplication and unit but also with an associative comultiplication (so that several representations can be tensored together) and counit. It is this last that precludes the comultiplication from being trivial in the sense of the previous footnote.

so that

$$\frac{d}{dt}\varphi(t)|_{t=0} = x \otimes 1 + 1 \otimes x\,.$$

This construction is particularly useful in the context of highest weight modules.

**(9.12).** PROPOSITION. *Let $V$ be a highest weight $L$-module generated by the maximal vector $v^+$ for $\lambda$, and let $W$ be a highest weight $L$-module generated by the maximal vector $w^+$ for $\mu$. Then in the $L$-module $V \otimes_{\mathbb{K}} W$, the vector $v^+ \otimes w^+$ is a maximal vector for the weight $\lambda + \mu$.*

PROOF. For $h \in H$ we have

$$\begin{aligned}(h \otimes 1 + 1 \otimes h)(v^+ \otimes w^+) &= (h \otimes 1)(v^+ \otimes w^+) + (1 \otimes h)(v^+ \otimes w^+)\\ &= (\lambda(h)v^+ \otimes w^+) + (v^+ \otimes \mu(h)w^+)\\ &= \lambda(h)(v^+ \otimes w^+) + \mu(h)(v^+ \otimes w^+)\\ &= (\lambda(h) + \mu(h))(v^+ \otimes w^+)\\ &= (\lambda + \mu)(h)\,(v^+ \otimes w^+)\,.\end{aligned}$$

For $u \in L^+$ we have

$$\begin{aligned}(u \otimes 1 + 1 \otimes u)(v^+ \otimes w^+) &= (u \otimes 1)(v^+ \otimes w^+) + (1 \otimes u)(v^+ \otimes w^+)\\ &= (0 \otimes w^+) + (v^+ \otimes 0) = 0 \qquad \Box\end{aligned}$$

The construction can then be iterated with as (finitely) many tensor products as one desires.

**(9.13).** COROLLARY. *Let the dominant weight $\lambda = \sum_{i=1}^{l} f_i\omega_i \in \Lambda_W^+$ (with $f_i \in \mathbb{N}$). Then the $L$-module*

$$V(\lambda) = L(\omega_1)^{\otimes f_1} \otimes \cdots \otimes L(\omega_i)^{\otimes f_i} \otimes \cdots \otimes L(\omega_l)^{\otimes f_l}$$

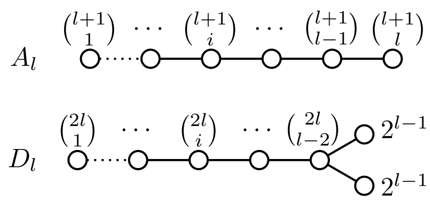
*contains a maximal vector for $\lambda$. In particular, if the $L(\omega_i)$ are finite dimensional for $1 \le i \le l$, then $L(\lambda)$ is finite dimensional.* $\Box$

Therefore to prove part (c) of Theorem (9.10) it suffices to construct, for every simple Lie algebra $L$ of Theorem (8.35) and for each of its fundamental weights $\omega_i$, a finite dimensional irreducible $L$-module with highest weight $\omega_i$. (Identification with $L(\omega_i)$ then comes from the uniqueness of Theorem (9.10)(a).) Such constructions we address (at least for $\mathfrak{sl}_{l+1}(\mathbb{K})$ and $\mathfrak{so}_{2l}(\mathbb{K})$) in the next section.

It must be noted that there are more elegant (and shorter) ways of proving Theorem (9.10)(c). By Proposition (8.1)(f) the set of weights $\Phi_{L(\lambda)}$ for the irreducible module $L(\lambda)$ is invariant under the Weyl group $\mathrm{W}(\Phi)$. As all multiplicities are finite (by Proposition (9.6)(b)), the module $L(\lambda)$ is finite dimensional if and only if the weight set $\Phi_{L(\lambda)}$ is finite. This is the beginning of an argument, carried out almost entirely within the root system $\Phi$, that shows the number of weights "under" the integral highest weight $\lambda$ is finite if and only if $\lambda$ is dominant integral. See [Eld15].

## 9.5 Fundamental modules

For each simple root $\delta_i$ there is a corresponding fundamental weight $\omega_i$ with $(\omega_i, \delta_i^\vee) = 1$ and $(\omega_i, \delta_j^\vee) = 0$ for $i \neq j$. The usual convention is to write the dimension of the *fundamental irreducible module* $L(\omega_i)$ next to the node of the Dynkin diagram corresponding to $\delta_i$. Thus:




As we shall now see, the fundamental irreducible module $L(\omega_i)$ for $L = \mathfrak{sl}_{l+1}(\mathbb{K}) = \mathfrak{a}_l(\mathbb{K})$ is the $i^{\text{th}}$-exterior power $\wedge^i V$ of the natural module $V = \mathbb{K}^{l+1}$. The dimension of $\wedge^i V$ is thus $\binom{l+1}{i}$. These irreducible modules for $\mathfrak{sl}_{2l}(\mathbb{K}) = \mathfrak{a}_{2l-1}(\mathbb{K})$ remain irreducible and fundamental upon restriction to the split orthogonal algebra $L = \mathfrak{so}_{2l}(\mathbb{K}) = \mathfrak{d}_l(\mathbb{K})$ provided $1 \leq i \leq l-2$, but the two remaining fundamental modules for $\mathfrak{d}_l(\mathbb{K})$ come from the associated *Clifford algebra*, a generalization of the exterior algebra for the natural module $\mathbb{K}^{2l}$. This pattern maintains—most of the fundamental modules for the simple algebras $L$ come by restriction from the exterior powers of the associated "natural" module, usually $L(\omega_1)$.

If $V$ is a $\mathbb{K}$-space and $\mathrm{T}(V)$ its tensor algebra, then the corresponding *exterior algebra* $\wedge(V)$ is the quotient of $\mathrm{T}(V)$ by its ideal generated by all $v \otimes v$ for $v \in V$. As the ideal is homogeneous, the quotient $\wedge(V)$ inherits the $\mathbb{N}$-grading of $\mathrm{T}(V)$:

$$\wedge(V) = \bigoplus_{k \in \mathbb{N}} \wedge^k(V)\,.$$

Here $\wedge^k(V)$ is the image of $V^{\otimes k}$. The arguments of the previous section (particularly those of Proposition (9.12) and its corollary) show it to be a module for $\mathrm{End}_{\mathbb{K}}(V)$ and $\mathrm{End}_{\mathbb{K}}^-(V)$.

We shall only be considering the finite dimensional case $\dim_{\mathbb{K}}(V) = n \in \mathbb{Z}^+$.

**(9.14).** THEOREM. *Let $v_1, \ldots, v_n$ be a $\mathbb{K}$-basis of $V$.*

(a) *$\wedge^k(V)$ has dimension $\binom{n}{k}$ with basis consisting of the monomials $v_{i_1} \cdots v_{i_k}$ for $i_1 < \cdots < i_k$.*

(b) *$\wedge(V)$ has dimension $2^n$ with basis consisting of the monomials $v_{i_1} \cdots v_{i_k}$ for $i_1 < \cdots < i_k$ for $0 \leq k \leq n$.*

PROOF. This follows easily since $vv = 0$ and $vw = -wv$ for $v, w \in V$. $\square$

**(9.15).** THEOREM. *Consider $\mathfrak{sl}_{l+1}(\mathbb{K}) = \mathfrak{a}_l(\mathbb{K})$. Set $V = \mathbb{K}^{l+1}$. In the root system $\Phi \subset \mathbb{R}^{l+1} \cap \mathbf{1}^\perp$ let the simple basis of roots be $\delta_k = \varepsilon_k - \varepsilon_{k+1}$ for $1 \leq k \leq l$.*

(a) $\omega_k = \sum_{j=1}^{k} \varepsilon_j + c_k\mathbf{1}$ *for* $1 \le k \le l$ *with* $c_k = -k(l+1)^{-1}$.

(b) $L(\omega_k) = \wedge^k V$ *for* $1 \le k \le l$.

Proof. (a) Easily $(\sum_{j=1}^{k} \varepsilon_j, \delta_i)$ is 1 if $k = i$ and 0 otherwise. The linear functional $\sum_{j=1}^{k} \varepsilon_j$ is not in the weight lattice $\mathbb{R}^{l+1} \cap \mathbf{1}^{\perp}$, but this can be easily fixed by adding the multiple $c_k\mathbf{1}$ for $c_k = -k(l+1)^{-1}$, which induces the trivial functional.

(b) Sketch: Calculating as in the previous section, we find that the basis vector $v_{i_1} \cdots v_{i_k}$ is a weight vector for the weight $\sum_{j=1}^{k} \varepsilon_{i_j}$.

By Proposition (8.1)(f) the Weyl group $\mathrm{Sym}(l+1)$ acts on the set of weights, so all of the possible weights $\sum_{j=1}^{k} \varepsilon_{i_j}$ occur for the irreducible quotient $L(\omega_k)$. Comparing dimensions, we conclude that $L(\omega_k) = \wedge^k V$. □

**(9.16).** Theorem. *Consider* $\mathfrak{so}_{2l}(\mathbb{K}) = \mathfrak{d}_l(\mathbb{K})$. *Set* $V = \mathbb{K}^{2l}$. *In the root system* $\Phi \subset \mathbb{R}^l$ *let the simple basis of roots be* $\delta_k = \varepsilon_k - \varepsilon_{k+1}$ *for* $1 \le k \le l-1$ *and* $\delta_l = \varepsilon_{l-1} + \varepsilon_l$.

(a) $\omega_k = \sum_{j=1}^{k} \varepsilon_j$ *for* $1 \le k \le l-2$, *but* $\omega_{l-1} = \frac{1}{2}(-\varepsilon_l + (\sum_{j=1}^{l-1} \varepsilon_j))$ *and* $\omega_{l-1} = \frac{1}{2}(\varepsilon_l + (\sum_{j=1}^{l-1} \varepsilon_j))$.

(b) $L(\omega_k) = \wedge^k V$ *for* $1 \le k \le l-2$.

(c) $L(\omega_k) = \mathrm{C}(V)^{\pm}$, *for* $k \in \{l-1, l\}$. *These are* $2^{l-1}$*-dimensional submodules of the Clifford algebra* $\mathrm{C}(V)$ *of dimension* $2^l$.

Proof. (a) This is easily checked.

(b) Sketch: For $1 \le k \le l-2$ the fundamental weight (actually, the associated linear functionals) remain the same, as do the corresponding modules.

To treat the remaining two fundamental weights precisely requires study of the Clifford algebra at greater length than possible here. □

The Weyl group $\mathrm{W}(D_l)$ contains $\mathrm{W}(A_{l-1}) \simeq \mathrm{Sym}(l)$ as a subgroup, so the exterior power $\wedge^{l-1}V$ remains irreducible, but it is no longer fundamental as the weight $\sum_{j=1}^{l-1} \varepsilon_j = \omega_{l-1} + \omega_l$ .

The *Clifford algebra* $\mathrm{C}(V)$ is a generalization of the exterior algebra. It is the quotient of the tensor algebra $\mathrm{T}(V)$ by the ideal generated by $v \otimes v - b(v,v)1$ for all $v \in V$. Thus the exterior algebra is the Clifford algebra for the trivial orthogonal form which is identically 0 on $V \times V$. The Clifford algebra also has dimension $2^l$ with the same monomial basis as the exterior algebra. (Exercise.)

The diagram $D_l$ has an automorphism of order 2 that switches the simple roots $\delta_{l-1}$ and $\delta_l$ and so the fundamental weights $\omega_{l-1}$ and $\omega_l$. The corresponding automorphism of $\mathfrak{d}_l(\mathbb{K})$ thus switches the corresponding fundamental representations. In fact, this graph automorphism is induced by any reflection of $\mathrm{O}_{2l}^{+}(\mathbb{K})$. Such reflections have determinant $-1$ and so do not belong to the group $\mathrm{SO}_{2l}^{+}(\mathbb{K})$. Nevertheless they act on the Clifford algebra, switching the two modules $\mathrm{C}(V)^{+}$ and $\mathrm{C}(V)^{-}$.

# Appendix A

# Bilinear 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 $\mathbb{E} \le \mathbb{R}$ and $b$ is an orthogonal form, we say that $b$ is *positive definite* if it has the property

$$b(x,x) \ge 0 \text{ always and } b(x,x) = 0 \iff x = 0\,.$$

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 $v = (\dots, v_i, \dots)$. 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 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 $\{\, e_i, e_{-i} \mid 1 \leq i \leq l \,\}$ 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 $\{\, e_0, e_i, e_{-i} \mid 1 \leq i \leq l \,\}$ and is equipped with the split orthogonal form $b$ given by

$$b(e_0, e_0) = 1\,,\ b(e_i, e_{-i}) = b(e_{-i}, e_i) = 1\,,\ \text{otherwise } b(e_a, e_b) = 0\,.$$

**(1.4).** 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* $\begin{pmatrix} 0 & 1 \\ \eta & 0 \end{pmatrix}$. □

**(1.5).** 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. □

**(1.6).** 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. □

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

$$r_v : x \mapsto x - \frac{2(x,v)}{(v,v)}v$$

is the *reflection* with center $v$.

**(2.1).** LEMMA. *Let $0 \neq v \in E$.*

(a) *$r_v$ belongs to $\mathrm{O}(E)$, the orthogonal group of isometries of $E$, being the reflection in the hyperplane orthogonal to $v$. In particular $r_v = r_{av}$ for all nonzero scalars $a$.*

(b) *If $g \in \mathrm{O}(E)$ then $r_v^g = r_{g(v)}$.*

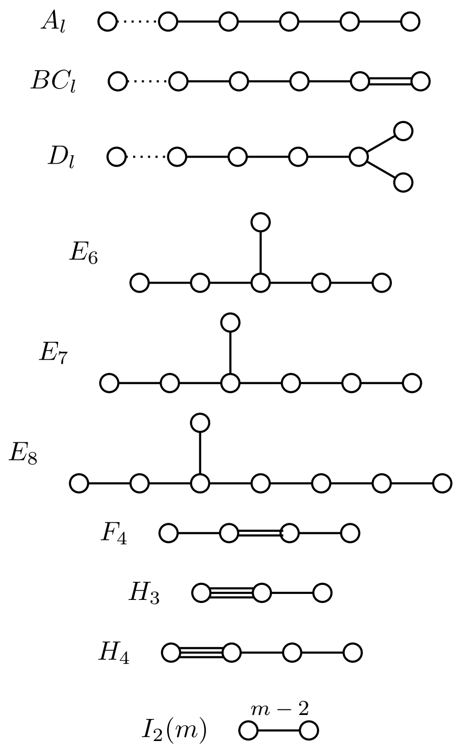
(c) *If $\mathbb{R}v^{r_x} = \mathbb{R}v$ if and only if $v \in \mathbb{R}x$ or $(v,x) = 0$.* □

**(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 $\frac{\pi}{m_{\alpha,\beta}}$.* □

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:*




PROOF. By Lemma (2.1) the graphs are all connected. We do not provide a complete proof of the theorem; but we do a proof, typical for these arguments, of one important property:

Claim: The Coxeter graph is a tree.

PROOF. Let $\mathcal{C} = (v_0, v_1, \ldots, v_{n-1}, v_n = v_0)$ be a circuit in the graph. Normalize so that $(v_i, v_i) = 1$ for all $i$, and let $G$ be the Gram matrix of $\mathcal{C}$. Then each diagonal entry of $G$ is 1 and in each row (and column) there are exactly two other nonzero entries. As the full basis is obtuse, each of these nonzero $(v_i, v_{i+1})$ is negative. Furthermore by Lemma B-(2.2) above, each of these has absolute value at least $\frac{1}{2}$. Therefore for

$$0 \neq x = \sum_{i=1}^{n} v_i = (1, 1, \ldots, 1)^\top$$

we have

$$(x, x) = x^\top G x \leq 0\,.$$

As Euclidean $(\cdot, \cdot)$ is positive definite, this is a contradiction. $\square$

Similar arguments then show that the given graphs are the only ones that are possible. □

# 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.

[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

# Index