Lie Algebras

Karl-Hermann Neeb

January 22, 2024

Contents

Part I
General Structure Theory

In this part we take a first look at the concept of a **Lie algebra**. Lie algebras arise naturally in many areas of mathematics. Their important role in mathematics is due to the fact that they are the infinitesimal counterparts, resp., "first order approximations" of Lie groups, so that they can be used to describe symmetries in algebraic terms. Accordingly, the Lie bracket of vector fields (the infinitesimal generators of flows) and the Poisson brackets from classical mechanics provide important examples of Lie algebra structures. [1]

In this course we study Lie algebras as independent algebraic structures. By the commutator bracket, every associative algebra inherits the structure of a Lie algebra and for every (not necessarily associative algebra) the space of derivations is a Lie algebra.

We start this chapter with the analysis of the algebraic structure of Lie algebras and the relevant concepts: What are the substructures? Under which condition does a substructure lead to a quotient structure? What are the simple structures? Does one have composition series? This leads to concepts like Lie subalgebras and ideals, nilpotent, solvable, and semisimple Lie algebras. Key results in this context are Engel's Theorems on nilpotent Lie algebras, Lie's Theorem for solvable Lie algebras and Cartan's criteria for solvability and semisimplicity. The latter are first instances in which one recognizes the usefulness of the Cartan–Killing form, which is a specific structural element of Lie algebras.

Throughout $\mathbb{K}$ denotes an arbitrary field if not specified otherwise. All vector spaces are vector spaces over $\mathbb{K}$.

1 Basic Concepts

In this section we provide the basic definitions and concepts concerning Lie algebras. In particular, we discuss ideals, quotients, homomorphisms and the elementary connections between these concepts.

1.1 Definitions and Examples

We start with the definition of a Lie algebra. [2] [3]

Definition 1.1. Let $\mathfrak{g}$ be a vector space. A *Lie bracket* on $\mathfrak{g}$ is a bilinear map $[\cdot\,,\cdot]\colon \mathfrak{g}\times\mathfrak{g} \to \mathfrak{g}$ satisfying

(L1) $[x, x] = 0$ for $x, y \in \mathfrak{g}$ (it is alternating), and

(L2) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for $x, y, z \in \mathfrak{g}$ (*Jacobi identity*). [4]

If $[\cdot, \cdot]$ is a Lie bracket on $\mathfrak{g}$, then the pair $(\mathfrak{g}, [\cdot\,, \cdot])$ is called a *Lie algebra*.

[1] In the context of dynamical systems, the Lie algebra would represent the differential equation $\dot{\mathbf{x}}(t) = F(t, \mathbf{x})$ encoding the evolution of a system and the Lie group the corresponding time evolution on the state space $\Phi_t(\mathbf{x}) = \mathbf{x}(t)$, $\mathbf{x}(0) = \mathbf{x}$.

[2] Marius Sophus Lie (1842–1899), Norwegian mathematician in Kristiania (Oslo) and Leipzig. Founder of the theory of transformation groups, which later lead to the modern concept of a Lie group.

[3] The term Lie algebra was introduced in the 1920s by Hermann Weyl, following a suggestion of N. Jacobson. Lie himself was dealing mainly with Lie algebras of vector fields, which he called (infinitesimal) transformation groups. The term Lie group was introduced by E. Cartan.

[4] Carl Gustav Jacob Jacobi (1804–1851), Mathematician in Berlin and Königsberg (Kaliningrad). He found the famous identity about 1830 in the context of Poisson brackets, occuring in Hamiltonian mechanics.

Remark 1.2. (a) From (L1) we immediately derive for $x, y \in \mathfrak{g}$ the relation

$$0 = [x + y, x + y] = [x, x] + [y, x] + [x, y] + [y, y] = [y, x] + [x, y],$$

hence

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

which means that any Lie bracket is skew-symmetric. If, conversely, a bracket $[\cdot, \cdot]$ is skew-symmetric, then we obtain for $x = y$ the relation $2[x, x] = 0$. Therefore (L1) follows, provided $\operatorname{char} \mathbb{K} \neq 2$, i.e., $2 = 1 + 1 \neq 0$ in the field $\mathbb{K}$.

Example 1.3. A vector space $\mathcal{A}$ together with a bilinear map $\cdot \colon \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ is called an *(associative) algebra* if

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \text{for} \quad a, b, c \in \mathcal{A}.$$

Then the *commutator*

$$[a, b] := a \cdot b - b \cdot a$$

defines a Lie bracket on $\mathcal{A}$. In fact, (L1) is obvious. For the Jacobi identity, we first observe that

$$\big[x, [y, z]\big] = x(yz - zy) - (yz - zy)x = xyz - xzy - yzx + zyx,$$

which leads to[5]

$$\begin{aligned}\sum_{\text{cyc.}} \big[x, [y, z]\big] &= \sum_{\text{cyc.}} xyz - xzy - yzx + zyx \\ &= \sum_{\text{cyc.}} xyz - xzy - xyz + xzy = 0.\end{aligned}$$

We write $\mathcal{A}_L := (\mathcal{A}, [\cdot, \cdot])$ for this Lie algebra.

Example 1.4. (a) Let V be a vector space and $\operatorname{End}(V)$ be the set of linear endomorphisms of V. Then $\operatorname{End}(V)$ is an associative algebra w.r.t. composition. We write $\mathfrak{gl}(V) := \operatorname{End}(V)_L$ for the corresponding Lie algebra with the bracket

$$[\varphi, \psi] := \varphi \circ \psi - \psi \circ \varphi.$$

(b) The space $M_n(\mathbb{K})$ of $(n \times n)$-matrices with entries in $\mathbb{K}$ is an associative algebra with respect to matrix multiplication. We write $\mathfrak{gl}_n(\mathbb{K}) := M_n(\mathbb{K})_L$ for the corresponding Lie algebra with the bracket

$$[A, B] := AB - BA.$$

Definition 1.5. (a) Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras. A linear map $\alpha \colon \mathfrak{g} \to \mathfrak{h}$ is called a *homomorphism (of Lie algebras)* [6] if

$$\alpha([x, y]) = [\alpha(x), \alpha(y)] \quad \text{for} \quad x, y \in \mathfrak{g}.$$

An *isomorphism* of Lie algebras is a homomorphism α for which there exists a homomorphism $\beta \colon \mathfrak{h} \to \mathfrak{g}$ with $\alpha \circ \beta = \mathrm{id}_{\mathfrak{h}}$ and $\beta \circ \alpha = \mathrm{id}_{\mathfrak{g}}$. It is easy to see that this condition is equivalent to α being bijective (Exercise). If an isomorphism $\varphi \colon \mathfrak{g} \to \mathfrak{h}$ exists, we call the Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ isomorphic.

[5] Here we use the notation $\sum_{\text{cyc.}}$ for the sum over all expressions obtained from a cyclic permutation of the variable.

[6] Based on the terminology of category theory, one also speaks of *morphisms of Lie algebras*.

(b) A *representation of a Lie algebra* $\mathfrak{g}$ *on the vector space* V is a homomorphism $\alpha\colon \mathfrak{g} \to \mathfrak{gl}(V)$. We also write (α, V) for a representation α of $\mathfrak{g}$ on V.

(c) Let $\mathfrak{g}$ be a Lie algebra and E, F be subsets of $\mathfrak{g}$. We write

$$[E, F] := \operatorname{span}\{[e, f]\colon e \in E, f \in F\}$$

for the smallest subspace containing all brackets $[e, f]$ with $e \in E$ and $f \in F$.

(d) A linear subspace $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is called a *Lie subalgebra* if $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$. Obviously, every Lie subalgebra $\mathfrak{h}$ is a Lie algebra with respect to the restriction of the Lie bracket to a map $\mathfrak{h} \times \mathfrak{h} \to \mathfrak{h}$. We then write $\mathfrak{h} < \mathfrak{g}$.

If the stronger condition $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$ is satisfied, then we call $\mathfrak{h}$ an *ideal* of $\mathfrak{g}$ and write $\mathfrak{h} \trianglelefteq \mathfrak{g}$.

(e) The Lie algebra $\mathfrak{g}$ is called *abelian* if $[\mathfrak{g}, \mathfrak{g}] = \{0\}$, which means that all brackets vanish.

Remark 1.6. From the definitions it is clear that the image of a homomorphism $\alpha\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ of Lie algebras is a subalgebra of $\mathfrak{g}_2$. Moreover, $\alpha^{-1}(\mathfrak{h})$ is an ideal in $\mathfrak{g}_1$ if $\mathfrak{h} \trianglelefteq \mathfrak{g}_2$, and $\alpha^{-1}(\mathfrak{h})$ is a subalgebra if $\mathfrak{h} < \mathfrak{g}_2$. In particular, the *kernel* $\ker \alpha$ of a Lie algebra homomorphism α is always an ideal.

Examples 1.7. (i) Let $\mathfrak{g}$ be Lie algebra. Then the *center*

$$\mathfrak{z}(\mathfrak{g}) := \{x \in \mathfrak{g} \mid (\forall y \in \mathfrak{g})\, [x, y] = 0\}$$

of $\mathfrak{g}$ is an ideal in $\mathfrak{g}$.

(ii) For each Lie algebra $\mathfrak{g}$, the subspace $[\mathfrak{g}, \mathfrak{g}]$ is an ideal (Exercise 1.11), called the *commutator algebra* of $\mathfrak{g}$.

(iii) Every one-dimensional subspace of a Lie algebra is a subalgebra since the Lie bracket is alternating.

(iv) The set

$$\mathfrak{sl}_n(\mathbb{K}) := \{x \in \mathfrak{gl}_n(\mathbb{K}) \mid \operatorname{tr}(x) = 0\}$$

is an ideal in $\mathfrak{gl}_n(\mathbb{K})$, where $\operatorname{tr}(x)$ denotes the *trace* of X. It is called the *special linear Lie algebra*. That $\mathfrak{sl}_n(\mathbb{K})$ is a Lie algebra follows from

$$\operatorname{tr}([X, Y]) = \operatorname{tr}(XY - YX) = \operatorname{tr}(XY) - \operatorname{tr}(YX) = 0$$

for $X, Y \in M_n(\mathbb{K})$. Since $\mathfrak{sl}_n(\mathbb{K})$ is a hyperplane in $M_n(\mathbb{K})$, we have $\dim \mathfrak{sl}_n(\mathbb{K}) = n^2 - 1$.

Writing E_{jk} for the matrix with entry 1 in position (j, k) and zeros elsewhere, we obtain the commutator brackets

$$[E_{jk}, E_{\ell m}] = \delta_{k\ell} E_{jm} - \delta_{jm} E_{\ell k}. \tag{1}$$

This easily implies that

$$[\mathfrak{gl}_n(\mathbb{K}), \mathfrak{gl}_n(\mathbb{K})] = \mathfrak{sl}_n(\mathbb{K})$$

(Exercise 1.9).

(v) The set

$$\mathfrak{o}_n(\mathbb{K}) := \{x \in \mathfrak{gl}_n(\mathbb{K}) \mid x = -x^\top\}$$

is a subalgebra of $\mathfrak{gl}_n(\mathbb{K})$: For $x, y \in \mathfrak{o}_n(\mathbb{K})$, we have

$$[x, y]^\top = [y^\top, x^\top] = [-y, -x] = [y, x] = -[x, y].$$

This Lie algebra is called the *orthogonal Lie algebra*. We have

$$\dim \mathfrak{o}_n(\mathbb{K}) = \frac{n(n-1)}{2}.$$

Note that

$$\mathfrak{so}_n(\mathbb{K}) := \mathfrak{o}_n(\mathbb{K}) \cap \mathfrak{sl}_n(\mathbb{K}) = \mathfrak{o}_n(\mathbb{K})$$

follows from the fact that $\operatorname{tr} x = 0$ for $x^\top = -x$.

(vi) The set

$$\mathfrak{u}_n(\mathbb{C}) := \{x \in \mathfrak{gl}_n(\mathbb{C}) \mid x = -x^*\}$$

is a real subalgebra of the complex Lie algebra $\mathfrak{gl}_n(\mathbb{C})$, the *unitary Lie algebra*. We define the *special unitary Lie algebra* by

$$\mathfrak{su}_n(\mathbb{C}) := \mathfrak{u}_n(\mathbb{C}) \cap \mathfrak{sl}_n(\mathbb{C}).$$

(vii) Let $J_n := \begin{pmatrix} \mathbf{0} & \mathbf{1}_n \\ -\mathbf{1}_n & \mathbf{0} \end{pmatrix} \in \mathfrak{gl}_{2n}(\mathbb{K})$ and note that $J_n^\top = -J_n$. Then the set

$$\mathfrak{sp}_{2n}(\mathbb{K}) := \{x \in \mathfrak{gl}_{2n}(\mathbb{K}) \colon x^\top J_n + J_n x = 0\}$$

is a Lie subalgebra of $\mathfrak{gl}_{2n}(\mathbb{K})$, called the *symplectic Lie algebra* (Exercise 1.10). Writing elements of $\mathfrak{gl}_{2n}(\mathbb{K})$ as (2×2)-block matrices, one easily verifies that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathfrak{sp}_{2n}(\mathbb{K}) \quad \Longleftrightarrow \quad B = B^\top, C = C^\top, A^\top = -D.$$

For the dimension we thus obtain

$$\dim \mathfrak{sp}_{2n}(\mathbb{K}) = n^2 + 2\frac{n(n+1)}{2} = 2n^2 + n.$$

(viii) The subspace

$$\mathfrak{n} = \{x = (x_{ij}) \in \mathfrak{gl}_n(\mathbb{K}) \mid (\forall i \geq j)\, x_{ij} = 0\}$$

of *strictly upper triangular matrices* and the subspace

$$\mathfrak{b} = \{x = (x_{ij}) \in \mathfrak{gl}_n(\mathbb{K}) \mid (\forall i > j)\, x_{ij} = 0\}$$

of *upper triangular matrices* are Lie subalgebras of $\mathfrak{gl}_n(\mathbb{K})$.

(ix) Let V be a subspace of a Lie algebra $\mathfrak{g}$. The *normalizer*

$$\mathfrak{n}_{\mathfrak{g}}(V) = \{x \in \mathfrak{g} \mid [x, V] \subseteq V\}$$

of V in $\mathfrak{g}$ is a subalgebra of $\mathfrak{g}$ (Exercise).

Example 1.8. Let V be a vector space. A tuple $\mathcal{F} = (V_0, \ldots, V_n)$ of subspaces with

$$\{0\} = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V$$

is called a *flag* in V. Then

$$\mathfrak{g}(\mathcal{F}) := \{x \in \mathfrak{gl}(V) \colon (\forall j)\ xV_j \subseteq V_j\}$$

is a Lie subalgebra of $\mathfrak{gl}(V) = \mathrm{End}(V)_L$ (it is even closed under composition).

To visualize this Lie algebra, we shall describe linear maps by suitable block matrices. If V is a vector space which is a direct sum $V = W_1 \oplus \ldots \oplus W_n$ of subspaces W_j, $j = 1, \ldots, n$, then we write an endomorphism $A \in \mathrm{End}(V)$ as an $(n \times n)$-block matrix

$$A = (A_{jk})_{j,k=1,\ldots,n} = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ A_{21} & \cdots & A_{2n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix},$$

where $A_{jk} \in \mathrm{Hom}(W_k, W_j)$ is uniquely determined by the requirement that the image of $v = (v_1, \ldots, v_n) \in V$ is

$$Av = \Big(\sum_{k=1}^{n} A_{jk} v_k\Big)_{j=1,\ldots,n}.$$

Here we simply write Aw for $A(w)$ to simplify notation.

Applying this kind of visualization to the Lie algebra $\mathfrak{g}(\mathcal{F})$, we choose in V_j a subspace W_j with $V_j \cong V_{j-1} \oplus W_j$. For each j, we then have $V_j \cong W_1 \oplus \ldots \oplus W_j$ and in particular $V \cong W_1 \oplus \ldots \oplus W_n$. Now the elements of $\mathfrak{g}(\mathcal{F})$ are those endomorphisms of V corresponding to upper triangular matrices

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ 0 & A_{22} & \cdots & A_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & A_{nn} \end{pmatrix}.$$

1.2 Derivations

Definition 1.9. Let $(\mathcal{A}, \cdot)$ be an algebra (not necessarily associative), i.e., $\mathcal{A} \times \mathcal{A} \to \mathcal{A}, (a, b) \mapsto a \cdot b$ is a bilinear map. Then $D \in \mathrm{End}(\mathcal{A})$ is called a *Derivation* if

$$D(x \cdot y) = D(x) \cdot y + x \cdot D(y) \quad \text{for} \quad x, y \in A.$$

We write $\mathrm{der}(\mathcal{A})$ for the subset of derivations in $\mathrm{End}(\mathcal{A})$.

Examples 1.10. (a) If $\mathcal{A} = C^\infty(\mathbb{R})$ is the space of smooth real-valued functions $f \colon \mathbb{R} \to \mathbb{R}$, endowed with the pointwise multiplication, then

$$D \colon \mathcal{A} \to \mathcal{A}, \quad f \mapsto f'$$

is a derivation by the Product Rule. This is the paradigmatic example of a derivation.

(b) For a general field $\mathbb{K}$, the polynomial ring $\mathcal{A} = \mathbb{K}[X]$ in one indeterminate X is an algebra and the linear map $D \colon \mathcal{A} \to \mathcal{A}$ specified by $D(X^n) = nX^{n-1}$ for $n \geq 1$ and $D(1) = 0$ is a derivation.

The notion of a derivation plays a central role in Lie algebra theory.

Lemma 1.11. *For every algebra $(\mathcal{A}, \cdot)$, the subset* $\mathrm{der}(\mathcal{A})$ *is a Lie subalgebra of* $\mathfrak{gl}(\mathcal{A}) = \mathrm{End}(\mathcal{A})_L$.

Proof. Since $\mathrm{der}(\mathcal{A})$ clearly is a linear subspace of $\mathfrak{gl}(\mathcal{A})$, we have to show that it is closed under the commutator bracket. For $D_1, D_2 \in \mathrm{der}(\mathcal{A})$ and $x, y \in \mathcal{A}$ we have

$$\begin{aligned}
&[D_1, D_2](x \cdot y) \\
&= D_1 D_2(x \cdot y) - D_2 D_1(x \cdot y) \\
&= D_1\big(D_2(x) \cdot y + x \cdot D_2(y)\big) - D_2\big(D_1(x) \cdot y + x \cdot D_1(y)\big) \\
&= \big(D_1 D_2(x)\big) \cdot y + D_2(x) \cdot D_1(y) + D_1(x) \cdot D_2(y) + x \cdot \big(D_1 D_2(y)\big) \\
&\quad - \big(D_2 D_1(x)\big) \cdot y - D_1(x) \cdot D_2(y) - D_2(x) \cdot D_1(y) - x \cdot \big(D_2 D_1(y)\big) \\
&= [D_1, D_2](x) \cdot y + x \cdot [D_1, D_2](y).
\end{aligned}$$

Therefore $[D_1, D_2] \in \mathrm{der}(\mathcal{A})$. $\square$

Specializing to Lie algebras, we obtain:

Definition 1.12. Let $\mathfrak{g}$ be a Lie algebra. A linear map $\delta\colon \mathfrak{g} \to \mathfrak{g}$ is a derivation if

$$\delta([x, y]) = [\delta(x), y] + [x, \delta(y)] \quad \text{for} \quad x, y \in \mathfrak{g}.$$

The following lemma shows that the Jacobi identity is closely linked to Lie brackets defining derivations:

Lemma 1.13. *Let $[\cdot,\cdot]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ be a bilinear map which is alternating, i.e., $[x, x] = 0$ for $x \in \mathfrak{g}$. Then $[\cdot,\cdot]$ is a Lie bracket if and only if, for every $x \in \mathfrak{g}$, the map*

$$\mathrm{ad}\, x\colon \mathfrak{g} \to \mathfrak{g}, \quad y \mapsto [x, y]$$

defines a derivation of $(\mathfrak{g}, [\cdot,\cdot])$.

Proof. Since $[x, x] = 0$ implies the skew-symmetry of the bracket, we have

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

so that we can immediately read off the assertion of the lemma. $\square$

Definition 1.14. Let $\mathfrak{g}$ be a Lie algebra and $x \in \mathfrak{g}$. We have seen above that the linear map

$$\mathrm{ad}\, x\colon \mathfrak{g} \to \mathfrak{g}, \quad y \mapsto [x, y]$$

is a derivation. Derivations of this form are called *inner derivations*. The map $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is called the *adjoint representation*. That it is a representation, i.e.,

$$\mathrm{ad}[x, y] = [\mathrm{ad}\, x, \mathrm{ad}\, y] \quad \text{for} \quad x, y \in \mathfrak{g} \tag{2}$$

follows directly from the Jacobi identity (Exercise).

Proposition 1.15. (Range and kernel of adjoint representation) *For any Lie algebra* $\mathfrak{g}$,

(i) $\mathrm{der}(\mathfrak{g}) < \mathfrak{gl}(\mathfrak{g})$ *and* $\mathrm{ad}(\mathfrak{g}) \trianglelefteq \mathrm{der}(\mathfrak{g})$ *is an ideal. In particular,*

$$[D, \mathrm{ad}\, x] = \mathrm{ad}(Dx) \quad \textit{for} \quad D \in \mathrm{der}(\mathfrak{g}), x \in \mathfrak{g}. \tag{3}$$

(ii) $\ker(\mathrm{ad}) = \mathfrak{z}(\mathfrak{g})$

Proof. (i) The first part is a special case of Lemma 1.11 and for the second one verifies (3) by direct calculation.

(ii) is trivial. $\square$

1.3 Representations and Modules

In this short subsection we introduce some terminology concerning representations of Lie algebras and the corresponding concept of a Lie algebra module.

Definition 1.16. Let $\mathfrak{g}$ be a Lie algebra and V be a vector space. Suppose that

$$\mathfrak{g} \times V \to V, \quad (x, v) \mapsto x \cdot v$$

is a bilinear map. If

$$[x, y] \cdot v = x \cdot (y \cdot v) - y \cdot (x \cdot v) \quad \text{for} \quad x, y \in \mathfrak{g}, v \in V,$$

then V is called a $\mathfrak{g}$-*module*.

Definition 1.17. (a) Let $\mathfrak{g}$ be a Lie algebra and V a $\mathfrak{g}$-module. A subspace $W \subseteq V$ is called a $\mathfrak{g}$-*submodule* if $\mathfrak{g} \cdot W \subseteq W$.

(b) A $\mathfrak{g}$-module V is called *simple* if it is nonzero and there are no submodules except $\{0\}$ and V. It is called *semisimple*, if V is the direct sum of simple submodules.

(c) If V and W are $\mathfrak{g}$-modules, then a linear map $\varphi \colon V \to W$ is called a *homomorphism (morphism) of* $\mathfrak{g}$-*modules* if

$$\varphi(x \cdot v) = x \cdot \varphi(v) \quad \text{for} \quad x \in \mathfrak{g}, v \in V.$$

We write $\mathrm{Hom}_{\mathfrak{g}}(V, W)$ for the vector space of all $\mathfrak{g}$-module homomorphisms from V to W and note that the set $\mathrm{End}_{\mathfrak{g}}(V) := \mathrm{Hom}_{\mathfrak{g}}(V, V)$ of module endomorphisms of V is an associative subalgebra of $\mathrm{End}(V)$.

If $\varphi \in \mathrm{Hom}_{\mathfrak{g}}(V, W)$ is bijective, then the inverse map $\varphi^{-1} \colon W \to V$ is also a homomorphism of $\mathfrak{g}$-modules (Exercise). Therefore φ is called an *isomorphism of* $\mathfrak{g}$-*modules*. The set of isomorphisms $V \to V$ is the group $\mathrm{Aut}_{\mathfrak{g}}(V) := \mathrm{End}_{\mathfrak{g}}(V)^{\times}$ of invertible elements in the algebra $\mathrm{End}_{\mathfrak{g}}(V)$.

Example 1.18. (a) Any Lie algebra $\mathfrak{g}$ carries a natural $\mathfrak{g}$-module structure defined by the adjoint representation $x \cdot y := [x, y]$. The $\mathfrak{g}$-submodules of $\mathfrak{g}$ are precisely the ideals (cf. Definition 1.14).

(b) If $\mathfrak{g} = \mathbb{K}$ is the one-dimensional Lie algebra and V a $\mathbb{K}$-vector space, then any endomorphism $D \in \mathrm{End}(V)$ determines a $\mathfrak{g}$-module structure on V defined by $t \cdot v := tD(v)$. Clearly, each $\mathfrak{g}$-module structure on V is of this form for $D(v) = 1 \cdot v$.

Remark 1.19. (Module structures versus representations)

(a) If $\pi \colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation, then a $\mathfrak{g}$-module structure on V is defined by

$$x \cdot v = \pi(x)v.$$

Conversely, for every $\mathfrak{g}$-module V, the map $\pi \colon \mathfrak{g} \to \mathfrak{gl}(V)$ defined by $\pi(x)v = x \cdot v$ is a representation. Thus representations of $\mathfrak{g}$ and $\mathfrak{g}$-modules are equivalent concepts.

(b) In the sense of (a), we call two representations *equivalent* if the corresponding modules are isomorphic.

Definition 1.20. A representation (π, V) of a Lie algebra $\mathfrak{g}$ is called *irreducible* if V is a simple $\mathfrak{g}$-module. It is called *completely reducible* if V is a semisimple $\mathfrak{g}$-module.

1.4 Quotients and Semidirect Sums

We have already seen that the kernel of a homomorphism of Lie algebras is an ideal. The following proposition implies in particular that each ideal is the kernel of a surjective homomorphism of Lie algebras.

Proposition 1.21. *Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{n}$ be an ideal in $\mathfrak{g}$. Then the quotient space $\mathfrak{g}/\mathfrak{n} = \{x + \mathfrak{n} \colon x \in \mathfrak{g}\}$ is a Lie algebra with respect to the bracket*

$$[x + \mathfrak{n}, y + \mathfrak{n}] := [x, y] + \mathfrak{n}.$$

The quotient map $\pi \colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{n}$ is a surjective homomorphism of Lie algebras with kernel $\mathfrak{n}$.

Proof. Since $\mathfrak{n}$ is an ideal, the bracket on $\mathfrak{g}/\mathfrak{n}$ is well-defined because $z, z' \in \mathfrak{n}$ implies

$$[x + z, y + z'] - [x, y] \in \mathfrak{n}.$$

The bracket on $\mathfrak{g}/\mathfrak{n}$ is clearly bilinear and satisfies

$$[\pi(x), \pi(y)] = \pi([x, y]) \quad \text{for} \quad x, y \in \mathfrak{g}. \tag{4}$$

Since π is surjective, the validity of (L1/2) in $\mathfrak{g}$ implies (L1/2) for the bracket on $\mathfrak{g}/\mathfrak{n}$: We have

$$[\pi(x), \pi(x)] = \pi([x, x]) = 0$$

and

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

In view of (4), π is a homomorphism of Lie algebras. □

Lemma 1.22. *A linear subspace $\mathfrak{n}$ of a Lie algebra $\mathfrak{g}$ is an ideal if and only if it is the kernel of a homomorphism of Lie algebras.*

Proof. Proposition 1.21 implies that every ideal $\mathfrak{n}$ of $\mathfrak{g}$ is the kernel of a homomorphism. If, conversely, $\varphi \colon \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras, then $\varphi([\mathfrak{g}, \ker\varphi]) \subseteq [\varphi(\mathfrak{g}), \{0\}] = \{0\}$ implies that its kernel is an ideal. □

Theorem 1.23. (Factorization Theorem) *Let $\varphi \colon \mathfrak{g}_1 \to \mathfrak{g}_2$ be a homomorphism of Lie $\mathfrak{n} \trianglelefteq \mathfrak{g}_1$ be an ideal and $\pi \colon \mathfrak{g}_1 \to \mathfrak{g}_1/\mathfrak{n}$ the quotient homomorphism. Then there exists a uniquely determined homomorphism $\overline{\varphi} \colon \mathfrak{g}_1/\mathfrak{n} \to \mathfrak{g}_2$ satisfying $\overline{\varphi} \circ \pi = \varphi$ if and only if $\mathfrak{n} \subseteq \ker\varphi$.*

Proof. If $\overline{\varphi}$ exists, then

$$\ker\varphi = \ker(\overline{\varphi} \circ \pi) \supseteq \ker\pi = \mathfrak{n}.$$

If, conversely, $\mathfrak{n} \subseteq \ker\varphi$, then we obtain a well-defined linear map

$$\overline{\varphi} \colon \mathfrak{g}_1/\mathfrak{n} \to \mathfrak{g}_2, \qquad \overline{\varphi}(x + \mathfrak{n}) := \varphi(x).$$

That $\overline{\varphi}$ is a homomorphism of Lie algebras follows from

$$\begin{aligned}\overline{\varphi}\big([\pi(x), \pi(y)]\big) &= \overline{\varphi}\big(\pi([x, y])\big) = \varphi([x, y]) = [\varphi(x), \varphi(y)]\\ &= [\overline{\varphi}\big(\pi(x)\big), \overline{\varphi}\big(\pi(y)\big)].\end{aligned}$$

That $\overline{\varphi}$ is uniquely determined by the relation $\overline{\varphi} \circ \pi = \varphi$ is due to the surjectivity of π. □

We leave the easy proof of the following proposition to the reader.[7]

Proposition 1.24. *Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras.*

(i) *If $\alpha\colon \mathfrak{g} \to \mathfrak{h}$ is a homomorphism, then $\alpha(\mathfrak{g}) \cong \mathfrak{g}/\ker\alpha$.*

(ii) *If $\mathfrak{i}, \mathfrak{j} \trianglelefteq \mathfrak{g}$ are ideals with $\mathfrak{i} \subseteq \mathfrak{j}$, then $\mathfrak{j}/\mathfrak{i} \trianglelefteq \mathfrak{g}/\mathfrak{i}$, and $(\mathfrak{g}/\mathfrak{i})/(\mathfrak{j}/\mathfrak{i}) \cong \mathfrak{g}/\mathfrak{j}$.*

(iii) *If $\mathfrak{i}, \mathfrak{j} \trianglelefteq \mathfrak{g}$ are two ideals, then $\mathfrak{i} + \mathfrak{j}$ and $\mathfrak{i} \cap \mathfrak{j}$ are ideals of $\mathfrak{g}$, and*

$$\mathfrak{i}/(\mathfrak{i} \cap \mathfrak{j}) \cong (\mathfrak{i} + \mathfrak{j})/\mathfrak{j}.$$

We have already seen that we obtain for each ideal $\mathfrak{n} \trianglelefteq \mathfrak{g}$ a quotient algebra $\mathfrak{g}/\mathfrak{n}$, so that we may consider the two Lie algebras $\mathfrak{n}$ and $\mathfrak{g}/\mathfrak{n}$ as two pieces into which $\mathfrak{g}$ is decomposed. It is therefore a natural question how we may build a Lie algebra $\mathfrak{g}$ from two Lie algebras $\mathfrak{n}$ and $\mathfrak{h}$ in such a way that $\mathfrak{n} \trianglelefteq \mathfrak{g}$ and $\mathfrak{g}/\mathfrak{n} \cong \mathfrak{h}$. The following definition describes one such construction.

Definition 1.25. (a) If $\mathfrak{g}$ is a Lie algebra, $\mathfrak{n} \trianglelefteq \mathfrak{g}$ an ideal and $\mathfrak{h} \leq \mathfrak{g}$ a subalgebra, so that $\mathfrak{g}$ is the direct vector space sum of $\mathfrak{n}$ and $\mathfrak{h}$. Then $\mathfrak{g}$ is called a *semidirect sum of* $\mathfrak{n}$ *and* $\mathfrak{h}$.

Since the ideal $\mathfrak{n} \subseteq \mathfrak{g}$ is a invariant under the adjoint representation,

$$\delta\colon \mathfrak{h} \to \operatorname{der}(\mathfrak{n}), \quad x \mapsto \operatorname{ad} x|_{\mathfrak{n}}$$

defines a homomorphism of Lie algebras. In these terms, the bracket in $\mathfrak{g} \cong \mathfrak{n} \oplus \mathfrak{h}$ is given by

$$[(n,h),(n',h')] = ([n,n'] + \delta(h)n' - \delta(h')n, [h,h']).$$

We therefore write $\mathfrak{g} = \mathfrak{n} \rtimes_\delta \mathfrak{h}$ and call $\mathfrak{g}$ the semidirect sum of $\mathfrak{n}$ and $\mathfrak{h}$ with respect to δ.

(b) If $\delta = 0$, then we write $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{h}$ and say that $\mathfrak{g}$ is the *direct sum of* $\mathfrak{n}$ *and* $\mathfrak{h}$.

Remark 1.26. A semidirect sum $\mathfrak{g} = \mathfrak{n} \rtimes_\delta \mathfrak{h}$ is a direct sum if and only if the subalgebra $\mathfrak{h} \cong \{0\} \times \mathfrak{h}$ is an ideal.

The following lemma shows that every triples $(\mathfrak{n}, \mathfrak{h}, \delta)$ as in Definition 1.25 actually corresponds to a Lie algebra.

Lemma 1.27. *Let $\mathfrak{n}$ and $\mathfrak{h}$ be Lie algebras and $\delta\colon \mathfrak{h} \to \operatorname{der}(\mathfrak{n})$ be a homomorphism of Lie algebras. Then*

$$[(n,h),(n',h')] = ([n,n'] + \delta(h)n' - \delta(h')n, [h,h'])$$

defines on the vector space $\mathfrak{g} := \mathfrak{n} \times \mathfrak{h}$ a Lie bracket such that $\mathfrak{n} \cong \mathfrak{n} \times \{0\}$ is an ideal and $\mathfrak{h} \cong \{0\} \times \mathfrak{h}$ a subalgebra with $\mathfrak{g} \cong \mathfrak{n} \rtimes_\delta \mathfrak{h}$.

Proof. Clearly,

$$[(n,h),(n,h)] := ([n,n] + \delta(h)n - \delta(h)n, [h,h])) = (0,0),$$

so that (L1) is satisfied. To verify the Jacobi identity, we put

$$J(x,y,z) := \big[x,[y,z]\big] + \big[y,[z,x]\big] + \big[z,[x,y]\big]$$

for $x, y, z \in \mathfrak{g}$ and observe that $J(x,y,z) = J(y,z,x) = J(z,x,y)$. We further have

$$J(x,x,z) = \big[x,[x,z]\big] + \big[x,[z,x]\big] + \big[z,[x,x]\big] = 0,$$

[7] These three statements are Lie algebra versions of Emmy Noether's Homomorphism Theorems for modules over rings in [EN27, §4, p. 40]).

so that J is alternating, i.e.,

$$J(x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)}) = \mathrm{sgn}(\sigma) J(x_1, x_2, x_3) \quad \text{for} \quad \sigma \in S_3, x_j \in \mathfrak{g}.$$

Therefore the verification of the Jacobi identity $J = 0$ reduces to the following four special cases:
(1) $x, y, z \in \mathfrak{n}$: $J(x, y, z) = 0$ follows from the Jacobi identity in $\mathfrak{n}$.
(2) $x, y \in \mathfrak{n}$, $z \in \mathfrak{h}$: $J(x, y, z) = -[x, \delta(z)y] - [\delta(z)x, y] + \delta(z)[x, y] = 0$ follows from the fact that $\delta(z)$ is a derivation of $\mathfrak{n}$.
(3) $x \in \mathfrak{n}$, $y, z \in \mathfrak{h}$: $J(x, y, z) = -\delta([y, z])x + \delta(y)\delta(z)x - \delta(z)\delta(y)x = 0$ follows from δ being a homomorphism.
(4) $x, y, z \in \mathfrak{h}$: $J(x, y, z) = 0$ is a consequence of the Jacobi identity in $\mathfrak{h}$. □

Example 1.28. (a) If $\delta\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation of $\mathfrak{g}$ and we consider V as an abelian Lie algebra, then $\mathrm{der}(V) = \mathfrak{gl}(V)$ and $V \rtimes_\delta \mathfrak{g}$ carries a Lie algebra structure given by

$$[(x, y), (x', y')] := (\delta(y)x' - \delta(y')x, [y, y']).$$

(b) If V is a vector space, then we write $\mathfrak{aff}(V)$ for the space of affine maps

$$L_{A,v}\colon V \to V, \quad x \mapsto Ax + v, \quad A \in \mathfrak{gl}(V), v \in V.$$

Note that $\mathfrak{aff}(V)$ carries a Lie algebra structure defined by

$$[L_{A,v}, L_{A',v'}] = L_{[A,A'],Av'-A'v}.$$

Then $\mathfrak{aff}(V) \cong V \rtimes_\delta \mathfrak{gl}(V)$ for $\delta(A)v = Av$.

(c) In $\mathfrak{gl}_{n+1}(\mathbb{K})$ we consider the Lie subalgebra

$$\mathfrak{g} := \left\{ \widetilde{L}_{A,v} = \begin{pmatrix} A & v \\ 0 & 0 \end{pmatrix} : A \in \mathfrak{gl}_n(\mathbb{K}), v \in \mathbb{K}^n \right\}.$$

Then $\widetilde{L}_{A,v} \circ \widetilde{L}_{A',v'} = \widetilde{L}_{AA',Av'}$ and therefore

$$[\widetilde{L}_{A,v}, \widetilde{L}_{A',v'}] = \widetilde{L}_{[A,A'],Av'-A'v}.$$

Hence $\mathfrak{g} \cong \mathfrak{aff}(\mathbb{K}^n) \cong \mathbb{K}^n \rtimes \mathfrak{gl}_n(\mathbb{K})$.

Example 1.29. For a derivation D of the Lie algebra $\mathfrak{n}$, we obtain a homomorphism $\delta\colon \mathbb{K} \to \mathrm{der}(\mathfrak{n})$ by $\delta(t) := tD$, and this leads to the semidirect sum $\mathfrak{n} \rtimes_D \mathbb{K} := \mathfrak{n} \rtimes_\delta \mathbb{K}$ with the bracket

$$[(n, t), (n', t')] = ([n, n'] + tDn' - t'Dn, 0)$$

(cf. Example 1.18(b)).

Example 1.30. Let $\mathfrak{h}_3(\mathbb{K})$ be the 3-dimensional vector space with the basis p, q, z equipped with the alternating bracket determined by

$$[p, q] = z, \quad [p, z] = [q, z] = 0.$$

Then $\mathfrak{h}_3(\mathbb{K})$ is a Lie algebras called the three dimensional *Heisenberg algebra.* It is isomorphic to the algebra $\mathfrak{n}$ in Example 1.7(viii) for $n = 3$. The linear endomorphism of $\mathfrak{h}_3(\mathbb{K})$ defined by

$$Dz = 0, \quad Dp = q \quad \text{and} \quad Dq = -p$$

then is a derivation of $\mathfrak{h}_3(\mathbb{K})$, so that we obtain a Lie algebra $\mathfrak{osc}(\mathbb{K}) := \mathfrak{h}_3(\mathbb{K}) \rtimes_D \mathbb{K}$, called the *oscillator algebra.* Writing $h := (0, 1)$ for the additional basis element in $\mathfrak{osc}(\mathbb{K})$, the nonzero brackets of basis elements are

$$[p, q] = z, \quad [h, p] = q \quad \text{and} \quad [h, q] = -p.$$

(cf. Exercise 1.12)

Example 1.31. If $\mathcal{F} = (V_0, \dots, V_n)$ is a flag in the vector space V (Example 1.8), then we know already the associated Lie algebra

$$\mathfrak{g}(\mathcal{F}) = \{x \in \mathfrak{gl}(V) \colon (\forall j)\ xV_j \subseteq V_j\}.$$

It is easy to see that

$$\mathfrak{g}_n(\mathcal{F}) := \{x \in \mathfrak{gl}(V) \colon (\forall j > 0)\ xV_j \subseteq V_{j-1}\}$$

is an ideal of $\mathfrak{g}(\mathcal{F})$. Here the n in $\mathfrak{g}_n(\mathcal{F})$ stands for "nilpotent".

To find a subalgebra complementary to this ideal, we choose subspaces $W_0, \dots, W_{n-1}$ of V with $V_{j+1} \cong V_j \oplus W_j$ for $j = 0, \dots, n-1$. Then

$$\mathfrak{g}_s(\mathcal{F}) := \{X \in \mathfrak{gl}(V) \colon (\forall j)\ XW_j \subseteq W_j\} \subseteq \mathfrak{g}(\mathcal{F})$$

is a subalgebra with

$$\mathfrak{g}(\mathcal{F}) \cong \mathfrak{g}_n(\mathcal{F}) \rtimes \mathfrak{g}_s(\mathcal{F}) \quad \text{and} \quad \mathfrak{g}_s(\mathcal{F}) \cong \bigoplus_{j=1}^{n} \mathfrak{gl}(W_j).$$

The s in $\mathfrak{g}_s(\mathcal{F})$ stands for "semisimple". Describing the elements of $\mathfrak{g}(\mathcal{F})$ as in Example 1.8 by block matrices, the semidirect decomposition of the Lie algebra $\mathfrak{g}(\mathcal{F})$ corresponds to the decomposition of an upper triangular matrix as a sum of a strictly upper triangular matrix and a diagonal matrix. For $n = 3$ we have in particular:

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{pmatrix} = \underbrace{\begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & 0 & A_{33} \end{pmatrix}}_{\in \mathfrak{g}_s(\mathcal{F})} + \underbrace{\begin{pmatrix} 0 & A_{12} & A_{13} \\ 0 & 0 & A_{23} \\ 0 & 0 & 0 \end{pmatrix}}_{\in \mathfrak{g}_n(\mathcal{F})}.$$

1.5 Complexification and Real Forms

Up to now, the base field did not really play a role in our considerations. But we shall see later on, that for some structure theoretic arguments, it is important for the operators $\operatorname{ad} x$ to have eigenvalues. This is certainly the case if the groundfile $\mathbb{K}$ is algebraically closed. Therefore, we also consider the *complexification* of a real Lie algebra. For this, we briefly recall how to calculate with the complexification of a vector space.

Definition 1.32. Let V be an $\mathbb{R}$-vector space. The *complexification* $V_{\mathbb{C}}$ *of* V is the vector space $V \oplus V$, endowed with the $\mathbb{C}$-vector space structure defined by

$$(x + iy)(v, w) := (xv - yw, xw + yv).$$

Identifying V with the subspace $V \times \{0\}$ of $V_{\mathbb{C}}$, we then have $iV = \{0\} \times V$ and $V_{\mathbb{C}} = V \oplus iV$ as real vector spaces. Accordingly, we write elements of $V_{\mathbb{C}}$ as $x + iy$, $x, y \in V$.[8]

The real linear map $\sigma \colon V_{\mathbb{C}} \to V_{\mathbb{C}}, v + iw \mapsto v - iw$, for $v, w \in V$, is called *complex conjugation.* Its real points consist of the subspace $V \subseteq V_{\mathbb{C}}$.

The proof of the following proposition is an elementary calculation.

Proposition 1.33. *Let* $\mathfrak{g}$ *be a real Lie algebra.*

[8]For readers familiar with tensor products: For every real vector space, we may identify the complexification $V_{\mathbb{C}}$ with $\mathbb{C} \otimes_{\mathbb{R}} V$, endowed with the complex scalar multiplication by $\lambda(z \otimes v) = \lambda z \otimes v$.

(i) $\mathfrak{g}_{\mathbb{C}}$ *is a complex Lie algebra with respect to the complex bilinear Lie bracket, defined by*

$$[x+iy, x'+iy'] := ([x,x'] - [y,y']) + i([x,y'] + [y,x']).$$

(ii) $[\mathfrak{g}_{\mathbb{C}}, \mathfrak{g}_{\mathbb{C}}] \cong [\mathfrak{g}, \mathfrak{g}]_{\mathbb{C}}$ *as complex Lie algebras.*

(iii) *The map* $\sigma(x+iy) := x - iy$ *for* $x, y \in \mathfrak{g}$ *defines an antilinear involutive automorphism of* $\mathfrak{g}_{\mathbb{C}}$ *whose fixed point set* $(\mathfrak{g}_{\mathbb{C}})^{\sigma} = \{z \in \mathfrak{g}_{\mathbb{C}} \colon \sigma(z) = z\}$ *is the real Lie algebra* $\mathfrak{g}$*.*

Definition 1.34. Let $\mathfrak{g}$ be a complex Lie algebra. A real Lie algebra $\mathfrak{h}$, for which $\mathfrak{g} = \mathfrak{h} + i\mathfrak{h}$ is a direct sum of real vector spaces is called a *real form* of $\mathfrak{g}$. Note that this implies that $\mathfrak{g} \cong \mathfrak{h}_{\mathbb{C}}$.

We have seen that to every real Lie algebra, we can assign a natural complexification (cf. Exercise 1.16). However, nonisomorphic real algebras can have isomorphic complexifications, resp., complex Lie algebras can have nonisomorphic real forms, as the following example shows.

Example 1.35. (Lie algebras with non-isomorphic real forms) In view of Exercise 1.15, we have

$$\mathfrak{su}_2(\mathbb{C})_{\mathbb{C}} \cong \mathfrak{sl}_2(\mathbb{C}) \cong \mathfrak{sl}_2(\mathbb{R})_{\mathbb{C}},$$

so that $\mathfrak{su}_2(\mathbb{C})$ and $\mathfrak{sl}_2(\mathbb{R})$ are both real forms of the complex Lie algebra $\mathfrak{sl}_2(\mathbb{C})$. We now show that $\mathfrak{sl}_2(\mathbb{R})$ and $\mathfrak{su}_2(\mathbb{C})$ are not isomorphic. To this end, we show that $\mathfrak{sl}_2(\mathbb{R})$ contains a 2-dimensional Lie subalgebra and that $\mathfrak{su}_2(\mathbb{C})$ does not. Clearly,

$$\mathfrak{b} := \left\{ \begin{pmatrix} x & y \\ 0 & -x \end{pmatrix} : x, y \in \mathbb{R} \right\}$$

is a 2-dimensional Lie subalgebra of $\mathfrak{sl}_2(\mathbb{C})$ (cf. Example 1.7(viii)). So it remains to see that $\mathfrak{su}_2(\mathbb{C})$ contains no 2-dimensional subalgebra. This will be done by first showing that $\mathfrak{su}_2(\mathbb{C}) \cong \mathfrak{so}_3(\mathbb{R})$. We consider the bases

$$a = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad b = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad c = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

of $\mathfrak{su}_2(\mathbb{C})$, and

$$x = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

of $\mathfrak{so}_3(\mathbb{R})$. Then

$$[a,b] = c, \quad [b,c] = a, \quad [c,a] = b$$

and

$$[x,y] = z, \quad [y,z] = x, \quad [z,x] = y.$$

Therefore the linear isomorphism $\mathfrak{su}_2(\mathbb{C}) \to \mathfrak{so}_3(\mathbb{R})$ with $a \mapsto x, b \mapsto y$ and $c \mapsto z$ is an isomorphism of Lie algebras. The Lie algebra $\mathfrak{so}_3(\mathbb{R})$ has no two-dimensional subalgebra because it is isomorphic to $(\mathbb{R}^3, \times)$, where $\times$ denotes the vector product (Exercise 1.8). Here we use that the vector product $x \times y$ of two linearly independent vectors $x, y \in \mathbb{R}^3$ is orthogonal to both, so that the plane $\mathbb{R}x + \mathbb{R}y$ is not a subalgebra.

Example 1.36. (A complex Lie algebra with no real form) On the abelian Lie algebra $V := \mathbb{C}^2$ we consider the linear operator D, defined by $De_1 = 2e_1$ and $De_2 = ie_2$ with respect to the canonical basis. Then we form the three dimensional complex Lie algebra $\mathfrak{g} := V \rtimes_D \mathbb{C}$ and note that $V = [\mathfrak{g}, \mathfrak{g}]$ is a 2-dimensional ideal of $\mathfrak{g}$.

Suppose that $\mathfrak{g}$ has a real form. Let $\sigma \in \mathrm{Aut}(\mathfrak{g})$ be the corresponding complex conjugation, which is an involutive antilinear automorphism of $\mathfrak{g}$ (Proposition 1.33(iii)). Then

$$\sigma(V) = \sigma([\mathfrak{g}, \mathfrak{g}]) = [\sigma(\mathfrak{g}), \sigma(\mathfrak{g})] = [\mathfrak{g}, \mathfrak{g}] = V,$$

so that σ induces an antilinear involution σ_V on V. Let $\sigma(0,1) = (v_0, \lambda)$ and note that $\sigma(V) = V$ implies that $\lambda \neq 0$. Applying σ again, we see that

$$(0,1) = \sigma^2(0,1) = \sigma_V(v_0) + \overline{\lambda}\sigma(0,1) = (\sigma_V(v_0) + \overline{\lambda}v_0, \overline{\lambda} \cdot \lambda).$$

We conclude that $|\lambda| = 1$. Further, $\sigma \circ \mathrm{ad}(0,1) \circ \sigma = \mathrm{ad}(\sigma(0,1)) = \mathrm{ad}(v_0, \lambda)$ implies by restricting to V that

$$\sigma_V \circ D \circ \sigma_V = \lambda D.$$

If $v \in V$ is a D-eigenvector with $Dv = \alpha v$, then

$$D(\sigma_V v) = \sigma_V(\lambda D v) = \overline{\lambda}\overline{\alpha}\sigma_V(v).$$

This means that $\sigma_V(V_\alpha(D)) = V_{\overline{\lambda}\overline{\alpha}}(D)$.[9] In particular, σ_V permutes the D-eigenspaces. Now $|\lambda| = 1$ and $|i| \neq 2$ show that σ_V preserves both eigenspaces. For $\alpha = 2$, this leads to $\overline{\lambda} = 1$, so that $\lambda = 1$. For $\alpha = i$ we now arrive at the contradiction $-i = \overline{\lambda}\overline{\alpha} = \alpha = i$.

This example is minimal because each complex Lie algebra of dimension 2 has a real form (cf. Example 4.2).

Exercises for Section 1

Exercise 1.1. Let $\mathcal{A}$ be an associative algebra and $\mathcal{A}_L$ be the associated Lie algebra (cf. Example 1.3).

(i) $\mathrm{der}(\mathcal{A}) \subseteq \mathrm{der}(\mathcal{A}_L)$, i.e., every derivation of the associative algebra $\mathcal{A}$ is a derivation of the Lie algebra $\mathcal{A}_L$, too.

(ii) $[a, bc] = [a, b]c + b[a, c]$ for $a, b, c \in \mathcal{A}$.

(iii) In general $\mathrm{der}(\mathcal{A}) \neq \mathrm{der}(\mathcal{A}_L)$.

(iv) If $\mathcal{A}$ is commutative, then $\mathcal{A} \cdot \mathrm{der}(\mathcal{A}) \subseteq \mathrm{der}(\mathcal{A})$.

Exercise 1.2. Let U be an open subset of $\mathbb{R}^{2n}$ and $\mathfrak{g} = C^\infty(U, \mathbb{R})$ be the set of smooth functions on U and write $q_1, \ldots, q_n, p_1, \ldots, p_n$ for the coordinates with respect to a basis. Then $\mathfrak{g}$ is a Lie algebra with respect to the *Poisson bracket*

$$\{f, g\} := \sum_{i=1}^n \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}.$$

The functions p_j and q_k satisfy the *canonical commutation relations (CCR)*:

$$\{p_j, p_k\} = \{q_j, q_k\} = 0 \quad \text{and} \quad \{q_j, p_k\} = \delta_{jk}.$$

[9] We use the following notation for eigenspaces: Let V be a $\mathbb{K}$-vector space and $A \in \mathrm{End}(V)$. We write $V_\lambda(A) := \ker(A - \lambda\mathbf{1})$ for the *eigenspace of* A corresponding to the eigenvalue λ and $V^\lambda(A) := \bigcup_{n\in\mathbb{N}} \ker(A - \lambda\mathbf{1})^n$ for the *generalized eigenspace of* A corresponding to λ.

Exercise 1.3. Let U be an open subset of $\mathbb{R}^n$, $\mathcal{A} = C^\infty(U, \mathbb{R})$, and $\mathfrak{g} = C^\infty(U, \mathbb{R}^n)$. For $f \in \mathcal{A}$ and $X \in \mathfrak{g}$, we define

$$\mathcal{L}_X f := Xf := \sum_{i=1}^n X_i \frac{\partial f}{\partial x_i}.$$

(i) The maps $\mathcal{L}_X$ are derivations of the algebra $\mathcal{A}$.

(ii) If $\mathcal{L}_X = 0$, then $X = 0$.

(iii) The commutator of two such operators has the form $[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_{[X,Y]}$, where the bracket on $\mathfrak{g}$ is defined by

$$[X, Y](p) := \mathtt{d}Y(p)X(p) - \mathtt{d}X(p)Y(p),$$

resp.,

$$[X, Y]_i = \sum_{j=1}^n X_j \frac{\partial Y_i}{\partial x_j} - Y_j \frac{\partial X_i}{\partial x_j}.$$

(iv) $(\mathfrak{g}, [\cdot, \cdot])$ is a Lie algebra.

(v) To each $A \in \mathfrak{gl}_n(\mathbb{R})$, we associate the linear vector field $X_A(x) := Ax$. Show that, for $A, B \in M_n(\mathbb{R})$, we have $X_{[A,B]} = -[X_A, X_B]$.

Exercise 1.4. Show that every 2-dimensional nonabelian Lie algebra contains a basis x, y with $[x, y] = y$. State the full classification of 2-dimensional Lie algebras. A natural matrix realization of this Lie algebra is

$$\mathfrak{aff}_1(\mathbb{K}) = \begin{pmatrix} \mathbb{K} & \mathbb{K} \\ 0 & 0 \end{pmatrix} \quad \text{with the basis} \quad x := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad y := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Exercise 1.5. (a) Show that all derivations of the 2-dimensional non-abelian Lie algebra $\mathfrak{n} = \mathfrak{aff}_1(\mathbb{K})$ are inner.
(b)* Conclude that, whenever a 3-dimensional Lie algebra $\mathfrak{g}$ has an ideal isomorphic to $\mathfrak{n}$, then $\mathfrak{g} \cong \mathfrak{n} \oplus \mathbb{K}$. Here $\mathfrak{n} \oplus \mathbb{K}$ denotes the direct sum of Lie algebras with the bracket

$$[(x, t), (x', t')] = ([x, x'], 0).$$

Exercise 1.6. Show that, every 3-dimensional Lie algebra $\mathfrak{g}$ with a 2-dimensional ideal is isomorphic to one of the following types:

(a) $\mathfrak{aff}_1(\mathbb{K}) \oplus \mathbb{K}$ or

(b) $\mathbb{K}^2 \rtimes_D \mathbb{K}$, $D \in \mathfrak{gl}_2(\mathbb{K})$ with $[(x, t), (x', t')] = (tDx' - t'Dx, 0)]$.

Hint: Use Exercise 1.5(b).

(c)* Can you determine when two Lie algebras $\mathbb{K}^2 \rtimes_D \mathbb{K}$ and $\mathbb{K}^2 \rtimes_E \mathbb{K}$, $D, E \in \mathfrak{gl}_2(\mathbb{K})$, are isomorphic? Find necessary and sufficient conditions.

Exercise 1.7. Let $\mathfrak{g}$ be a Lie algebra, $\mathfrak{n} \trianglelefteq \mathfrak{g}$ an ideal and $\mathfrak{h} < \mathfrak{g}$ a Lie subalgebra with $\mathfrak{g} = \mathfrak{n} + \mathfrak{h}$ and $\mathfrak{n} \cap \mathfrak{h} = \{0\}$. Then

$$\delta \colon \mathfrak{h} \to \operatorname{der} \mathfrak{n}, \quad \delta(x) := \operatorname{ad} x|_{\mathfrak{n}}$$

defines a homomorphism of Lie algebras and the map

$$\Phi \colon \mathfrak{n} \rtimes_\delta \mathfrak{h} \to \mathfrak{g}, \quad (x, y) \mapsto x + y$$

is an isomorphism of Lie algebras.

Exercise 1.8. (a) On $\mathbb{K}^3$ we define the vector product by

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} := \begin{pmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{pmatrix}.$$

Show that $(\mathbb{K}^3, \times)$ is a Lie algebra and that the map

$$\Phi\colon (\mathbb{K}^3, \times) \to \mathfrak{so}_3(\mathbb{K}), \quad \Phi(v)w := v \times w$$

is an isomorphism of Lie algebras.

(b) For $(\mathbf{v}, \mathbf{w}) = \sum_{j=1}^3 v_j w_j$ we have $(\mathbf{v} \times \mathbf{w}, \mathbf{v}) = 0$ and $(\mathbf{v} \times \mathbf{w}, \mathbf{w}) = 0$.

(c)* When does the Lie algebra $(\mathbb{K}^3, \times)$ contain 2-dimensional subalgebras? Express this as a condition on the field $\mathbb{K}$. Do $\mathbb{R}$, $\mathbb{C}$ satisfy this condition?

Exercise 1.9. (a) Show that $[\mathfrak{gl}_n(\mathbb{K}), \mathfrak{gl}_n(\mathbb{K})] = \mathfrak{sl}_n(\mathbb{K})$ and $\mathfrak{z}(\mathfrak{gl}_n(\mathbb{K})) = \mathbb{K}\mathbf{1}$.
(b) For which fields is the intersection of these two ideals $\{0\}$?

Exercise 1.10. Show that:

(i) For every matrix $B \in M_n(\mathbb{K})$, the subspace

$$\mathfrak{g}_B := \{x \in \mathfrak{gl}_n(\mathbb{K})\colon x^\top B + Bx = 0\}$$

is a Lie subalgebra of $\mathfrak{gl}_n(\mathbb{K})$.
Here is a more conceptual argument:

(ii) By $\pi(x)A := xA + Ax^\top$, we obtain a representation of $\mathfrak{gl}_n(\mathbb{K})$ on $M_n(\mathbb{K})$.

(iii) If V is a $\mathfrak{g}$-module and $v \in V$, then

$$\mathfrak{g}_v := \{x \in \mathfrak{g}\colon x \cdot v = 0\}$$

is a subalgebra.

Exercise 1.11. Show that, for two ideals $\mathfrak{a}$ and $\mathfrak{b}$ of the Lie algebra $\mathfrak{g}$, the subspace $[\mathfrak{a}, \mathfrak{b}]$ also is an ideal.

Exercise 1.12. On the algebra $\mathcal{A} := C^\infty(\mathbb{R}, \mathbb{C})$, consider the operators

$$Pf(x) := if'(x), \quad Qf(x) := xf(x) \quad \text{and} \quad Zf(x) = if(x).$$

Then the Lie subalgebra of $\mathfrak{gl}(\mathcal{A})$ generated by P, Q and Z is isomorphic to the Heisenberg algebra $\mathfrak{h}_3(\mathbb{R})$, i.e.,

$$[P, Q] = Z \quad \text{and} \quad [P, Z] = [Q, Z] = 0.$$

In Quantum mechanices, Q is the *position operator* and P the *momentum operator*. Adding also the operator (the Hamiltonian of the harmonic oscillator)

$$Hf(x) := \frac{i}{2}\Big(- \frac{d^2 f}{dx^2}(x) + x^2 f(x)\Big), \quad H = \frac{i}{2}(P^2 + Q^2),$$

we obtain a four-dimensional Lie subalgebra, isomorphic to the oscillator algebra (Example 1.30).

Exercise 1.13. Let (π, V) be a representation of the Lie algebra $\mathfrak{g}$ on V and $W \subseteq V$ a $\mathfrak{g}$-invariant subspace, i.e., $\pi(\mathfrak{g})W \subseteq W$. Then

$$\overline{\pi}\colon \mathfrak{g} \to \mathfrak{gl}(V/W), \quad \overline{\pi}(x)(v + W) := \pi(x)v + W$$

defines a representation of $\mathfrak{g}$ on the quotient space V/W.

Exercise 1.14. For the following Lie algebras, find a *faithful*, i.e., injective, finite dimensional representation: $\mathfrak{sl}_2(\mathbb{K})$, the Heisenberg algebra, the oscillator algebra, and the abelian Lie algebra $\mathbb{R}^n$.

Exercise 1.15. (Complexifications) Show that, as complex Lie algebras,

$$\mathfrak{gl}_n(\mathbb{R})_{\mathbb{C}} \cong \mathfrak{gl}_n(\mathbb{C}), \quad \mathfrak{sl}_n(\mathbb{R})_{\mathbb{C}} \cong \mathfrak{sl}_n(\mathbb{C}), \quad \mathfrak{o}_n(\mathbb{R})_{\mathbb{C}} \cong \mathfrak{o}_n(\mathbb{C}), \quad \mathfrak{su}_n(\mathbb{C})_{\mathbb{C}} \cong \mathfrak{sl}_n(\mathbb{C}).$$

Exercise 1.16. Show that the complexification $\mathfrak{g}_{\mathbb{C}}$ of a real Lie algebra $\mathfrak{g}$ has the following universal property. For every real linear homomorphism $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ of $\mathfrak{g}$ to a complex Lie algebra $\mathfrak{h}$, there exists a unique complex linear homomorphism $\varphi_{\mathbb{C}}\colon \mathfrak{g}_{\mathbb{C}} \to \mathfrak{h}$ with $\varphi_{\mathbb{C}}|_{\mathfrak{g}} = \varphi$.

2 Lie Algebras of Matrix Groups

Many concepts in the theory of Lie algebras correspond to the concepts in group theory:

(1) The adjoint representation $\mathrm{ad}\colon \mathfrak{g} \to \mathrm{der}(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$ corresponds to the conjugation action $c\colon G \to \mathrm{Aut}(G), c(g)(x) := gxg^{-1}$ of a group G on itself by inner automorphisms. Accordingly, the automorphisms c_g of G are called *inner* and the derivations $\mathrm{ad}\, x$, $x \in \mathfrak{g}$, of a Lie algebra $\mathfrak{g}$ are called inner.

(2) Ideals are subalgebras invariant under inner derivations. Likewise normal subgroups are subgroup invariant under all inner automorphisms.

(3) The Lie bracket $[x, y]$ in $\mathfrak{g}$ corresponds to the commutator bracket $(x, y) := xyx^{-1}y^{-1}$ for elements of a group. Groups are abelian if all commutators are trivial and abelian Lie algebras are defined likewise.

(4) The *commutator group* (G, G) of a group G is the subgroup generated by all commutators $xyx^{-1}y^{-1}$, it is automatically normal. Likewise, the commutator algebra $[\mathfrak{g}, \mathfrak{g}]$ of a Lie algebra $\mathfrak{g}$ is an ideal.

(5) For a group G, the *center* is

$$Z(G) := \{x \in G\colon (\forall y \in G) xy = yx\},$$

and the relation $xy = yx$ can also be written as $(x, y) = 1$. Accordingly,

$$\mathfrak{z}(\mathfrak{g}) = \{x \in \mathfrak{g}\colon [x, \mathfrak{g}] = \{0\}\}$$

for a Lie algebra $\mathfrak{g}$.

Later we shall also encounter the notion of a nilpotent and a solvable Lie algebra which are analogous to the corresponding notions for groups.

The correspondence between Lie algebras and groups is particularly direct for groups $G \subseteq \mathrm{GL}_n(\mathbb{R})$ of real matrices. For such a group G, we write $C^1_*([0,1], G)$ for the set of differentiable paths $\alpha\colon [0,1] \to M_n(\mathbb{R})$ with $\alpha([0,1]) \subseteq G$ and $\alpha(0) = \mathbf{1}$ (the identity matrix).

Definition 2.1. Let $G \subseteq \mathrm{GL}_n(\mathbb{R})$ be a subgroup. We define

$$\mathbf{L}(G) := \{x \in M_n(\mathbb{R}) \colon (\exists \alpha \in C^1_*([0,1], G))\, x = \alpha'(0)\}.$$

Proposition 2.2. *For every subgroup $G \subseteq \mathrm{GL}_n(\mathbb{R})$, the set $\mathbf{L}(G)$ is a Lie subalgebra of $\mathfrak{gl}_n(\mathbb{R})$.*

Proof. (a) First we observe that, for $\alpha \in C^1_*([0,1], G)$, the pointwise inverse α^{-1} is also contained in $C^1_*([0,1], G)$ and that its derivative is given by

$$(\alpha^{-1})'(t) = -\alpha(t)^{-1}\alpha'(t)\alpha(t)^{-1}.$$

This follows from

$$\lim_{h\to 0} \frac{1}{h}(\alpha^{-1}(t+h) - \alpha^{-1}(t)) = \lim_{h\to 0} \alpha^{-1}(t+h)\Big(\frac{\alpha(t) - \alpha(t+h)}{h}\Big)\alpha^{-1}(t) = -\alpha^{-1}(t)\alpha'(t)\alpha^{-1}(t).$$

We also note that the Product Rule implies that, for $\alpha, \beta \in C^1_*([0,1], G)$, the curve $\alpha \cdot \beta$ is contained in $C^1_*([0,1], G)$. It satisfies

$$(\alpha\beta)'(t) = \alpha'(t)\beta(t) + \alpha(t)\beta'(t).$$

(b) If $\alpha, \beta \in C^1_*([0,1], G)$, then

$$(\alpha\beta)'(0) = \alpha'(0) + \beta'(0), \quad (\alpha^{-1})'(0) = -\alpha(0),$$

and, for $0 \leq \lambda \leq 1$, the curve $\alpha_\lambda(t) := \alpha(\lambda t)$ satisfies $\alpha'_\lambda(0) = \lambda\alpha'(0)$. Therefore $\mathbf{L}(G)$ is a real linear subspace of $\mathfrak{gl}_n(\mathbb{R})$.

(c) For $g \in G$ and $x = \gamma'(0) \in \mathbf{L}(G)$, the curve defined by $\eta(t) := g\gamma(t)g^{-1}$ is also contained in $C^1_*([0,1], G)$ and satisfies $\mathbf{L}(G) \ni \eta'(0) = gxg^{-1}$.

(d) For $x, y \in \mathbf{L}(G)$ and $x = \gamma'(0)$ we know from (c) that $\beta(t) := \alpha(t)y\alpha(t)^{-1}$ defines a curve in $\mathbf{L}(G)$ which is differentiable by (a). Therefore

$$\mathbf{L}(G) \ni \beta'(0) = \alpha'(0)y\alpha(0)^{-1} - \alpha(0)y\alpha'(0) = xy - yx = [x, y].$$

This completes the proof. □

Example 2.3. (a) Clearly, $\mathbf{L}(\mathrm{GL}_n(\mathbb{R})) = \mathfrak{gl}_n(\mathbb{R})$ because

$$\mathrm{GL}_n(\mathbb{R}) = \{g \in M_n(\mathbb{R}) \colon \det g \neq -0\}$$

is an open neighborhood of $\mathbf{1}$.

(b) To see that

$$\mathbf{L}(\mathrm{SL}_n(\mathbb{R})) = \mathfrak{sl}_n(\mathbb{R}),$$

we observe that $\mathtt{d}\det(\mathbf{1}) = \mathrm{tr}$ (consider the linear terms in the Leibniz formula). For every $\gamma \in C^1_*([0,1], \mathrm{SL}_n(\mathbb{R}))$, we therefore obtain

$$0 = \frac{d}{dt}\Big|_{t=0} \det(\gamma(t)) = \mathrm{tr}(\gamma'(0))$$

by the Chain Rule, and thus $\gamma'(0) \in \mathfrak{sl}_n(\mathbb{R})$. If, conversely, $X \in \mathfrak{sl}_n(\mathbb{R})$, then

$$\det(e^{tX}) = e^{\mathrm{tr}(tX)} = 1 \quad \text{for every} \quad t \in \mathbb{R},$$

so that $X = \gamma'(0)$ for the curve $\gamma(t) = e^{tX}$ in $\mathrm{SL}_n(\mathbb{R})$.

Lemma 2.4. *Let V and W be finite dimensional real vector spaces and $\beta\colon V\times V\to W$ a bilinear map. For $(x,y)\in \mathrm{End}(V)\times\mathrm{End}(W)$ the following are equivalent:*

(i) $e^{ty}\beta(v,v') = \beta(e^{tx}v, e^{tx}v')$ *for all* $t\in\mathbb{R}$ *and all* $v,v'\in V$.

(ii) $y\beta(v,v') = \beta(xv,v') + \beta(v,xv')$ *for all* $v,v'\in V$.

Proof. (i) $\Rightarrow$ (ii): Taking the derivative in $t=0$, the relation (i) leads to

$$y\beta(v,v') = \beta(xv,v') + \beta(v,xv'),$$

where we use the Product and the Chain Rule (Exercise 2.1(c)).

(ii) $\Rightarrow$ (i): From (ii) we obtain inductively

$$y^n\beta(v,v') = \sum_{k=0}^{n}\binom{n}{k}\beta(x^kv, x^{n-k}v').$$

For the absolutely convergent exponential series, this leads with the general Cauchy Product Formula (Exercise 2.2) to

$$\begin{aligned} e^y\beta(v,v') &= \sum_{n=0}^{\infty}\frac{1}{n!}y^n\beta(v,v') = \sum_{n=0}^{\infty}\frac{1}{n!}\Big(\sum_{k=0}^{n}\binom{n}{k}\beta(x^kv,x^{n-k}v)\Big)\\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{n}\beta\left(\frac{1}{k!}x^kv, \frac{1}{(n-k)!}x^{n-k}v'\right)\\ &= \beta\left(\sum_{k=0}^{\infty}\frac{1}{k!}x^kv, \sum_{m=0}^{\infty}\frac{1}{m!}x^mv'\right) = \beta\left(e^xv, e^xv'\right). \end{aligned}$$

Since (ii) also holds for the pair (tx,ty) for all $t\in\mathbb{R}$, this completes the proof. □

Proposition 2.5. *Let V and W be finite dimensional vector spaces and $\beta\colon V\times V\to W$ a bilinear map. For the group*

$$\mathrm{Aut}(V,\beta) = \{g\in \mathrm{GL}(V)\colon (\forall v,v'\in V)\ \beta(gv,gv') = \beta(v,v')\},$$

we then have

$$\begin{aligned}\mathfrak{aut}(V,\beta) &:= \mathbf{L}(\mathrm{Aut}(V,\beta))\\ &= \{x\in\mathfrak{gl}(V)\colon (\forall v,v'\in V)\ \beta(xv,v') + \beta(v,xv') = 0\}.\end{aligned}$$

Proof. First we observe that $e^{\mathbb{R}X}\subseteq \mathrm{Aut}(V,\beta)$ is equivalent to the pair $(X,0)$ satisfying condition (i) in Lemma 2.4. This proves $\supseteq$. The converse is obtained by taking for $\gamma\in C^1_*([0,1],\mathrm{Aut}(V,\beta))$ the derivative of the relation

$$\beta(\gamma(t)v,\gamma(t)v') = \beta(v,v')$$

in $t=0$. □

Example 2.6. Let $\mathfrak{g}$ be a finite dimensional $\mathbb{K}$-Lie algebra ($\mathbb{K}=\mathbb{R},\mathbb{C}$) and

$$\mathrm{Aut}(\mathfrak{g}) := \{g\in\mathrm{GL}(\mathfrak{g})\colon (\forall x,y\in\mathfrak{g})\ g[x,y] = [gx,gy]\}.$$

To calculate the Lie algebra of $\mathrm{Aut}(\mathfrak{g})$, first observe that, for $\gamma\in C^1_*([0,1],\mathrm{Aut}(\mathfrak{g}))$, taking derivatives in $t=0$ of the relation

$$\gamma(t)([x,y]) = [\gamma(t)(x),\gamma(t)(y)],$$

we obtain for $D:=\gamma'(0)$ the relation:

$$D[x,y] = [Dx,y] + [x,Dy],$$

i.e., $D\in\mathrm{der}(\mathfrak{g})$. If, conversely, $D\in\mathrm{der}(\mathfrak{g})$, then we use Lemma 2.4 with $V=W=\mathfrak{g}$ and $\beta(x,y)=[x,y]$ to see that $e^{\mathbb{R}D}\subseteq\mathrm{Aut}(\mathfrak{g})$. This shows that

$$\mathfrak{aut}(\mathfrak{g}) = \mathbf{L}(\mathrm{Aut}(\mathfrak{g})) = \mathrm{der}(\mathfrak{g}).$$

Exercises for Section 2

Exercise 2.1. Let $X_1, \ldots, X_n$ be finite dimensional normed spaces and $\beta\colon X_1 \times \ldots \times X_n \to Y$ an n-linear map.

(a) Show that there exists a constant $C \geq 0$ with

$$\|\beta(x_1, \ldots, x_n)\| \leq C\|x_1\| \cdots \|x_n\| \quad \text{for} \quad x_i \in X_i.$$

(b) Show that β is continuous.

(c) Show that β is differentiable with

$$\mathtt{d}\beta(x_1, \ldots, x_n)(h_1, \ldots, h_n) = \sum_{j=1}^{n} \beta(x_1, \ldots, x_{j-1}, h_j, x_{j+1}, \ldots, x_n).$$

Exercise 2.2. [Cauchy Product Formula] Let X, Y, Z be Banach spaces and $\beta\colon X \times Y \to Z$ a continuous bilinear map. Suppose that $x := \sum_{n=0}^{\infty} x_n$ is absolutely convergent in X and that $y := \sum_{n=0}^{\infty} y_n$ is absolutely convergent in Y. Then

$$\beta(x, y) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \beta(x_k, y_{n-k}).$$

Exercise 2.3. Let $v \in \mathbb{R}^n$ and $G := \{g \in \mathrm{GL}_n(\mathbb{R})\colon gv = v\}$ be its stabilizer subgroup. Show that

$$\mathbf{L}(G_v) = \{X \in \mathfrak{gl}_n(\mathbb{R})\colon Xv = 0\}.$$

Exercise 2.4. Show that:

(i) The Lie algebra of the orthogonal group $\mathrm{O}_n(\mathbb{R}) := \{g \in \mathrm{GL}_n(\mathbb{R})\colon g^\top = g^{-1}\}$ is $\mathfrak{o}_n(\mathbb{R})$.

(ii) The Lie algebra of the symplectic group $\mathrm{Sp}_{2n}(\mathbb{R}) := \{g \in \mathrm{GL}_{2n}(\mathbb{R})\colon g^\top J_n g = J_n\}$ is $\mathfrak{sp}_{2n}(\mathbb{R})$ (cf. Examples 1.7(vii)).

Exercise 2.5. Let $E \subseteq \mathbb{R}^n$ be a linear subspace and $G := \{g \in \mathrm{GL}_n(\mathbb{R})\colon gE = E\}$ be its stabilizer subgroup. Show that

$$\mathbf{L}(G_E) = \{X \in \mathfrak{gl}_n(\mathbb{R})\colon XE \subseteq E\}.$$

3 Nilpotent Lie Algebras

In the following, we shall encounter several important classes of Lie algebras that play a central role in the structure theory of finite dimensional Lie algebras. The first of these two classes, nilpotent Lie algebras, are those for which iterated brackets $[x_1, [x_2, [x_3, [x_4, \cdots]]]]$ of sufficiently large order vanish. The most important result on nilpotent Lie algebras is Engel's Theorem which translates nilpotency of a Lie algebra into the elementwise condition that all operators $\operatorname{ad} x$ are nilpotent. Typical examples of nilpotent Lie algebras are Lie algebras of strictly upper triangular (block) matrices.

Definition 3.1. Let $\mathfrak{g}$ be a Lie algebra. We define its *descending (lower) central series* inductively by

$$C^1(\mathfrak{g}) := \mathfrak{g} \quad \text{and} \quad C^{n+1}(\mathfrak{g}) := [\mathfrak{g}, C^n(\mathfrak{g})].$$

In particular, $C^2(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$ is the commutator algebra. The Lie algebra $\mathfrak{g}$ is called *nilpotent*, if there is a $d \in \mathbb{N}_0$ with $C^{d+1}(\mathfrak{g}) = \{0\}$. If d is minimal with this property, then it is called the *nilpotence degree* of $\mathfrak{g}$. By induction, one immediately sees that each $C^n(\mathfrak{g})$ is an ideal of $\mathfrak{g}$, so that $C^{n+1}(\mathfrak{g}) \subseteq C^n(\mathfrak{g})$. Hence, for finite dimensional Lie algebras, the nilpotency of $\mathfrak{g}$ is equivalent to the vanishing of the ideal $C^\infty(\mathfrak{g}) := \bigcap_{n\in\mathbb{N}} C^n(\mathfrak{g})$.

Example 3.2. (i) The Heisenberg algebra $\mathfrak{h}_3(\mathbb{K})$ (Example 1.30) is nilpotent of degree 2 because $C^2(\mathfrak{h}_3(\mathbb{K})) = \mathbb{K}z$ is central.

(ii) A Lie algebra if nilpotent of degree 1 if and only if it is abelian.

(iii) If $\mathcal{F} = (V_0, \ldots, V_d)$ is a flag in the vector space V and we put $V_i := \{0\}$ for $i < 0$, then $\mathfrak{g}_n(\mathcal{F})$ (Example 1.31) is a nilpotent Lie algebra. In fact, an easy induction leads to

$$C^m\big(\mathfrak{g}_n(\mathcal{F})\big)V_j \subseteq V_{j-m}$$

and therefore to $C^d\big(\mathfrak{g}_n(\mathcal{F})\big) = \{0\}$.

Proposition 3.3. *Let $\mathfrak{g}$ be a Lie algebra.*

(i) *If $\mathfrak{g}$ is nilpotent, then all subalgebras and all homomorphic images of $\mathfrak{g}$ are nilpotent.*

(ii) *If $\mathfrak{z} < \mathfrak{z}(\mathfrak{g})$ and $\mathfrak{g}/\mathfrak{z}$ is nilpotent, then $\mathfrak{g}$ is nilpotent.*

(iii) *If $\mathfrak{g} \neq \{0\}$ is nilpotent, then $\mathfrak{z}(\mathfrak{g}) \neq \{0\}$.*

(iv) *If $\mathfrak{g}$ is nilpotent, then there exists an $n \in \mathbb{N}$ with $\mathrm{ad}(x)^n = 0$ for all $x \in \mathfrak{g}$, i.e., the $\mathrm{ad}(x)$ are nilpotent as linear maps.*

(v) *If $\mathfrak{i} \trianglelefteq \mathfrak{g}$, then all the spaces $C^n(\mathfrak{i})$ are ideals of $\mathfrak{g}$.*

Proof. (i) If $\mathfrak{h} < \mathfrak{g}$, then $[\mathfrak{h}, \mathfrak{h}] \subseteq [\mathfrak{g}, \mathfrak{g}]$ and $C^n(\mathfrak{h}) \subseteq C^n(\mathfrak{g})$ follows by induction. Therefore each subalgebra of a nilpotent Lie algebra is nilpotent.

For a homomorphism $\alpha\colon \mathfrak{g} \to \mathfrak{h}$, we obtain inductively

$$C^n(\alpha(\mathfrak{g})) = \alpha(C^n(\mathfrak{g})) \qquad \text{for each } n \in \mathbb{N}. \tag{5}$$

Thus, if $C^n(\mathfrak{g}) = \{0\}$, then $C^n(\operatorname{im}\alpha) = \{0\}$.

(ii) If $\mathfrak{g}/\mathfrak{z}$ is nilpotent, then there exists an $n \in \mathbb{N}$ with $C^n(\mathfrak{g}/\mathfrak{z}) = \{0\}$, so that (5), applied to the quotient homomorphism $q\colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{z}$, leads to $C^n(\mathfrak{g}) \subseteq \mathfrak{z} \subseteq \mathfrak{z}(\mathfrak{g})$ and thus to $C^{n+1}(\mathfrak{g}) \subseteq [\mathfrak{g}, \mathfrak{z}(\mathfrak{g})] = \{0\}$.

(iii) If $\mathfrak{g} \neq \{0\}$ is nilpotent, for some $d \in \mathbb{N}_0$, we have $C^{d+1}(\mathfrak{g}) = \{0\}$ and $C^d(\mathfrak{g}) \neq \{0\}$. Then $[\mathfrak{g}, C^d(\mathfrak{g})] = \{0\}$ implies that the non-zero ideal $C^d(\mathfrak{g})$ is contained in the center.

(iv) If $C^{d+1}(\mathfrak{g}) = \{0\}$, then $(\operatorname{ad} x)^d\mathfrak{g} \subseteq C^{d+1}(\mathfrak{g}) = \{0\}$.

(v) In view of Exercise 1.11, this follows by induction. □

In Proposition 3.3, we have seen that for every nilpotent Lie algebra, all the endomorphisms $\mathrm{ad}(x)$, $x \in \mathfrak{g}$, are nilpotent. Now our aim is to show that a finite dimensional Lie algebra, for which every $\operatorname{ad} x$ is nilpotent, is nilpotent itself. We start with a simple lemma, the proof of which we leave to the reader as an exercise (cf. Exercises 3.1 and 1.13).

Lemma 3.4. (i) *Let V be a finite dimensional vector space, $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ a Lie subalgebra and $x \in \mathfrak{g}$. If $x \in \mathfrak{gl}(V)$ is nilpotent, then $\mathrm{ad}(x)\colon \mathfrak{g} \to \mathfrak{g}$ is also nilpotent.*

(ii) *Let* $\mathfrak{g}$ *be a Lie algebra and* $\mathfrak{h} < \mathfrak{g}$. *Then*

$$\mathrm{ad}_{\mathfrak{g}/\mathfrak{h}} : \mathfrak{h} \to \mathfrak{gl}(\mathfrak{g}/\mathfrak{h}), \quad \mathrm{ad}_{\mathfrak{g}/\mathfrak{h}}(x)(y+\mathfrak{h}) := [x,y] + \mathfrak{h}$$

defines a representation of $\mathfrak{h}$ *on the vector space* $\mathfrak{g}/\mathfrak{h}$.

Theorem 3.5. (Engel's Theorem on linear Lie algebras) *Let* $V \neq \{0\}$ *be a finite dimensional vector space and* $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ *a Lie subalgebra. If all* $x \in \mathfrak{g}$ *are nilpotent, i.e.,* $x^n = 0$ *for some* $n \in \mathbb{N}$, *then there exists a nonzero* $v_o \in V$ *with* $\mathfrak{g}(v_o) = \{0\}$.

Proof. We proceed by induction on $\dim \mathfrak{g}$. For $\dim \mathfrak{g} = 0$ the assertion holds trivially for each nonzero $v_o \in V$.

Next we assume that $\dim \mathfrak{g} > 0$ and pick a proper subalgebra $\mathfrak{h} < \mathfrak{g}$ of maximal dimension. According to Lemma 3.4, for each $x \in \mathfrak{h}$ the operators $\mathrm{ad}_{\mathfrak{g}/\mathfrak{h}}(x)$ are nilpotent. Now our induction hypothesis implies the existence of some $x_o \in \mathfrak{g}\backslash\mathfrak{h}$ with $\mathrm{ad}_{\mathfrak{g}/\mathfrak{h}}(\mathfrak{h})(x_o + \mathfrak{h}) = \{0\}$, i.e., $[\mathfrak{h}, x_o] \subseteq \mathfrak{h}$. This implies that $\mathbb{K}x_o + \mathfrak{h}$ is a subalgebra of $\mathfrak{g}$ and, again by maximality of $\mathfrak{h}$, it follows that $\mathbb{K}x_o + \mathfrak{h} = \mathfrak{g}$. The induction hypothesis also implies that the space $V_o := \{v \in V : \mathfrak{h}(v) = \{0\}\}$ is nonzero. Moreover,

$$yx(w) = xy(w) - [x,y](w) \in x\mathfrak{h}(w) + \mathfrak{h}(w) = \{0\} \quad \text{for} \quad x \in \mathfrak{g}, y \in \mathfrak{h}, w \in V_o,$$

implies that $\mathfrak{g}(V_o) \subseteq V_o$. Since $x_o|_{V_o}$ is also nilpotent, there exists a nonzero $v_o \in V_o$ with $x_o(v_o) = 0$. Putting all this together, we arrive at $\mathfrak{g}(v_o) = \mathfrak{h}(v_o) + \mathbb{K}x_o(v_o) = \{0\}$. □

Exercise 3.8 discusses an interesting Lie algebra of nilpotent endomorphisms of an infinite dimensional space, showing in particular that Engel's Theorem does not generalize to infinite dimensional spaces.

Definition 3.6. Let V be an n-dimensional vector space. A *complete flag* in V is a flag $(V_0, \ldots, V_n)$ with $\dim V_k = k$ for each k.

Corollary 3.7. *Let* V *be a finite dimensional vector space and* $\mathfrak{g} < \mathfrak{gl}(V)$ *such that all elements of* $\mathfrak{g}$ *are nilpotent. Then there exists a complete flag* $\mathcal{F}$ *in* V *with* $\mathfrak{g} \subseteq \mathfrak{g}_n(\mathcal{F})$. *In particular, there exists a basis for* V *with respect to which the elements of* $\mathfrak{g}$ *correspond to strictly upper triangular matrices. In particular,* $\mathfrak{g}$ *is nilpotent.*

Proof. In view of Theorem 3.5, there exists a nonzero $v_1 \in V$ with $\mathfrak{g}(v_1) = \{0\}$. We set $V_1 := \mathbb{K}v_1$. Then

$$\alpha : \mathfrak{g} \ \to \ \mathfrak{gl}(V/V_1), \quad \alpha(x)(v + V_1) := x(v) + V_1$$

is a representation of $\mathfrak{g}$ on V/V_1 (Exercise 1.13), and $\alpha(\mathfrak{g})$ consists of nilpotent endomorphisms. We now proceed by induction on $\dim V$, so that the induction hypothesis implies that V/V_1 possesses a complete flag $\mathcal{F}_1 = (W_1, \ldots, W_k)$ with $\alpha(\mathfrak{g}) \subseteq \mathfrak{g}_n(\mathcal{F}_1)$. Then $\{0\}$, together with the preimage of the flag $\mathcal{F}_1$ in V is a complete flag $\mathcal{F}$ in V with $\mathfrak{g} \subseteq \mathfrak{g}_n(\mathcal{F})$. Since $\mathfrak{g}_n(\mathcal{F})$ is nilpotent (Example 3.2(iii)), the subalgebra $\mathfrak{g}$ is also nilpotent. □

Now we are able to prove the announced criterion for the nilpotency of a Lie algebra.

Theorem 3.8. (Engel's Characterization Theorem for nilpotent Lie algebras) *Let* $\mathfrak{g}$ *be a finite dimensional Lie algebra. Then* $\mathfrak{g}$ *is nilpotent if and only if for each* $x \in \mathfrak{g}$ *the operator* $\mathrm{ad}\, x$ *is nilpotent.*

Proof. We have already seen in Proposition 3.3 that for each $x \in \mathfrak{g}$ the operator $\operatorname{ad} x$ is nilpotent. It remains to show the converse.

If $\operatorname{ad} x$ is nilpotent for each $x \in \mathfrak{g}$, then Corollary 3.7 implies that the Lie algebra $\mathfrak{g}/\mathfrak{z}(\mathfrak{g}) \cong \operatorname{ad}(\mathfrak{g}) \subseteq \mathfrak{gl}(\mathfrak{g})$ is nilpotent. Now Proposition 3.3(ii) shows that $\mathfrak{g}$ is also nilpotent. □

Lemma 3.9. *If $\mathfrak{a}$ and $\mathfrak{b}$ are nilpotent ideals of the Lie algebra $\mathfrak{g}$, then so is their sum $\mathfrak{a}+\mathfrak{b}$.*

Proof. We claim that

$$C^{2m}(\mathfrak{a}+\mathfrak{b}) \subseteq C^m(\mathfrak{a}) + C^m(\mathfrak{b}) \quad \text{for} \quad m \in \mathbb{N}. \tag{6}$$

This implies the assertion because $C^m(\mathfrak{a}) = C^m(\mathfrak{b}) = \{0\}$ holds if m is sufficiently large.

The space $C^{2m}(\mathfrak{a}+\mathfrak{b})$ is spanned by elements of the form

$$y := [x_1, [x_2, [x_3, \cdots [x_{2m-1}, x_{2m}] \cdots]]]$$

with $x_j \in \mathfrak{a} \cup \mathfrak{b}$. If at least m of the x_j are contained in $\mathfrak{a}$, then $y \in C^m(\mathfrak{a})$. If this is not the case, then m of the x_j are contained in $\mathfrak{b}$, which leads to $y \in C^m(\mathfrak{b})$. This proves our claim and hence the lemma. □

Definition 3.10. The main consequence of Lemma 3.9 is that every finite dimensional Lie algebra $\mathfrak{g}$ contains a largest nilpotent ideal. In fact, if $\mathfrak{n} \trianglelefteq \mathfrak{g}$ is a nilpotent ideal of maximal dimension and $\mathfrak{m} \trianglelefteq \mathfrak{g}$ any other nilpotent ideal, then the nilpotency of $\mathfrak{n}+\mathfrak{m}$ implies that $\mathfrak{m} \subseteq \mathfrak{n}$. Therefore the ideal $\mathfrak{n}$ contains all other nilpotent ideals.

The maximal nilpotent ideal is called the *nilradical* of $\mathfrak{g}$, and is denoted by $\operatorname{nil}(\mathfrak{g})$.

Remark 3.11. One should be aware of the fact that some authors use the term "nilradical" in a different meaning, namely for the intersection of the kernels of all irreducible finite dimensional representations. Since for an abelian Lie algebra $\mathfrak{g}$, the one-dimensional representations separate the points, the intersection of the kernels of irreducible finite dimensional representations is $\{0\}$, but we have $\operatorname{nil}(\mathfrak{g}) = \mathfrak{g}$. For more details on this ideal, see Proposition 7.5.

Remark 3.12. If $\mathfrak{g}$ is not finite dimensional, then the preceding lemma implies that for each finite sequence $\mathfrak{n}_1, \ldots, \mathfrak{n}_k \trianglelefteq \mathfrak{g}$ of nilpotent ideals, their sum $\mathfrak{n}_1 + \cdots + \mathfrak{n}_k$ is a nilpotent ideal. Therefore the sum $\mathfrak{n}$ of all nilpotent ideals of $\mathfrak{g}$ coincides with the union of all nilpotent ideals. However, this ideal need not be nilpotent because there may be nilpotent ideals of an arbitrary high nilpotence degree.

Exercises for Section 3

Exercise 3.1. If $X \in \operatorname{End}(V)$ is nilpotent, then $\operatorname{ad} X \in \operatorname{End}(\operatorname{End}(V))$ is also nilpotent.

Exercise 3.2. Let V be a finite dimensional complex vector space and $x \in \operatorname{End}(V)$ diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $\operatorname{ad} x$ is diagonalizable with eigenvalues

$$\lambda_i - \lambda_j, \quad i, j = 1, \ldots, n.$$

Exercise 3.3. Let $\mathfrak{g}$ be a Lie algebra, $\mathfrak{h}$ a subalgebra and $x \in \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}) \setminus \mathfrak{h}$. Then $\mathfrak{h} + \mathbb{K}x \cong \mathfrak{h} \rtimes_\alpha \mathbb{K}x$ for $\alpha(tx) = \operatorname{ad}(tx)|_{\mathfrak{h}}$.

Exercise 3.4. Let $\mathfrak{g} = \mathfrak{h}_3(\mathbb{K})$ be the 3-dimensional Heisenberg algebra. Determine a basis for $\mathfrak{g}$ with respect to which $\operatorname{ad}\mathfrak{g}$ consists of upper triangular matrices.

Exercise 3.5. Let $\mathfrak{g}$ be a nilpotent Lie algebra and $\mathfrak{h}$ be a finite dimensional nonzero ideal in $\mathfrak{g}$. Show that the intersection of $\mathfrak{h}$ with the center of $\mathfrak{g}$ is not trivial.

Exercise 3.6. Give an example of a Lie algebra $\mathfrak{g}$ which contains a nilpotent ideal $\mathfrak{n}$ for which $\mathfrak{g}/\mathfrak{n}$ is nilpotent and $\mathfrak{g}$ is not nilpotent.

Exercise 3.7. For each Lie algebra $\mathfrak{g}$, we have

$$[C^n(\mathfrak{g}), C^m(\mathfrak{g})] \subseteq C^{n+m}(\mathfrak{g}) \quad \text{for} \quad n, m \in \mathbb{N}.$$

Exercise 3.8. This exercise shows why Engel's Theorem does not generalize to infinite dimensional spaces. We consider the vector space $V = \mathbb{K}^{(\mathbb{N})}$ with the basis $\{e_i : i \in \mathbb{N}\}$. In terms of the rank-one-operators $E_{ij} \in \mathrm{End}(V)$, defined by $E_{ij}e_k = \delta_{jk}e_i$, we consider the Lie algebra

$$\mathfrak{g} := \mathrm{span}\{E_{ij} : i > j\}$$

(strictly lower triangular matrices). Show that:

(a) $C^n(\mathfrak{g}) = \mathrm{span}\{E_{ij} : i \geq j + n\}$, $n \in \mathbb{N}$. In particular, we have $C^\infty(\mathfrak{g}) = \bigcap_{n\in\mathbb{N}} C^n(\mathfrak{g}) = \{0\}$, i.e., $\mathfrak{g}$ is *residually nilpotent.* (cf. Exercise 3.9).

(b) $\mathfrak{g}$ consists of nilpotent endomorphisms of finite rank.

(c) $\mathfrak{z}(\mathfrak{g}) = \{0\}$.

(d) $V^{\mathfrak{g}} = \{v \in V : \mathfrak{g} \cdot v = \{0\}\} = \{0\}$ (compare with Engel's Theorem).

Exercise 3.9. Show that, for a Lie algebra $\mathfrak{g}$, the following assertions hold:

(a) $C^\infty(\mathfrak{g}) = \bigcap_n C^n(\mathfrak{g})$ is the intersection of all kernels of homomorphisms $\varphi : \mathfrak{g} \to \mathfrak{n}$, $\mathfrak{n}$ nilpotent. Hint: Show that all quotients $\mathfrak{g}/C^n(\mathfrak{g})$ are nilpotent.

(b) $C^\infty(\mathfrak{g}) = \{0\}$ is equivalent to $\mathfrak{g}$ being residually nilpotent in the sense that all homomorphisms $\varphi : \mathfrak{g} \to \mathfrak{n}$, $\mathfrak{n}$ nilpotent, separate the points of $\mathfrak{g}$, i.e., for every non-zero $x \in \mathfrak{g}$, there exists a homomorphism into a nilpotent Lie algebra with $\varphi(x) \neq 0$.

Exercise 3.10. Let $\mathfrak{g}$ be a Lie algebra and $D \in \mathrm{der}(\mathfrak{g})$. Show that the semidirect sum $\widehat{\mathfrak{g}} := \mathfrak{g} \rtimes_D \mathbb{K}$ is nilpotent if and only if $\mathfrak{g}$ is nilpotent and D is nilpotent.

Exercise 3.11. Let $\mathfrak{g}$ be a Lie algebra. We define its *upper central series* inductively by

$$C_0(\mathfrak{g}) := \{0\} \quad \text{and} \quad C_{n+1}(\mathfrak{g}) := \{x \in \mathfrak{g} : [x, \mathfrak{g}] \subseteq C_n(\mathfrak{g})\}, n \in \mathbb{N}_0.$$

Show that:

(i) Every $C_n(\mathfrak{g})$ is an ideal of $\mathfrak{g}$.

(ii) $C_n(\mathfrak{g}) \subseteq C_{n+1}(\mathfrak{g})$.

(iii) $\mathfrak{g}$ is nilpotent if and only if there exists an $N \in \mathbb{N}$ with $C_N(\mathfrak{g}) = \mathfrak{g}$.

4 Solvable Lie Algebras

In this section we turn to the class of solvable Lie algebras. They are defined in a similar fashion as nilpotent ones and indeed every nilpotent Lie algebra is solvable. The central results on solvable Lie algebras are Lie's Theorem on representations of solvable Lie algebras (they preserve complete flags) and Cartan's Solvability Criterion in terms of vanishing of

$$\mathrm{tr}(\mathrm{ad}[x, y]\, \mathrm{ad}\, z) \quad \text{for} \quad x, y, z \in \mathfrak{g}.$$

As we shall see later on, similar techniques apply to semisimple Lie algebras.

4.1 Basic Properties

Definition 4.1. Let $\mathfrak{g}$ be a Lie algebra. The *derived series* of $\mathfrak{g}$ is defined by

$$D^0(\mathfrak{g}) := \mathfrak{g} \quad \text{and} \quad D^n(\mathfrak{g}) := [D^{n-1}(\mathfrak{g}), D^{n-1}(\mathfrak{g})] \quad \text{for} \quad n \in \mathbb{N}.$$

The Lie algebra $\mathfrak{g}$ is said to be *solvable*, if there exists an $n \in \mathbb{N}$ with $D^n(\mathfrak{g}) = \{0\}$.

From $D^1(\mathfrak{g}) \subseteq \mathfrak{g}$ we inductively see that $D^n(\mathfrak{g}) \subseteq D^{n-1}(\mathfrak{g})$. Further, an easy induction shows that all $D^n(\mathfrak{g})$ are ideals of $\mathfrak{g}$ (Exercise 1.11). The derived series is a descending series of ideals.

Example 4.2.

(i) The oscillator algebra $\mathfrak{osc}(\mathbb{K}) \cong \mathfrak{h}_3(\mathbb{K}) \rtimes_D \mathbb{K}$ from Example 1.30 is solvable, but not nilpotent:

$$D^1(\mathfrak{osc}(\mathbb{K})) = \mathfrak{h}_3(\mathbb{K}), \quad D^2(\mathfrak{osc}(\mathbb{K})) = D^1(\mathfrak{h}_3(\mathbb{K})) = \mathbb{K}z \quad \text{and} \quad D^3(\mathfrak{osc}(\mathbb{K})) = \{0\}.$$

(ii) Every nilpotent Lie algebra is solvable because $D^n(\mathfrak{g}) \subseteq C^{n+1}(\mathfrak{g})$ follows easily by induction.

(iii) Consider $\mathbb{R}$ and $\mathbb{C}$ as abelian real Lie algebras and write $I \in \mathrm{End}_{\mathbb{R}}(\mathbb{C})$ for the multiplication with i. Then the Lie algebra $\mathbb{C} \rtimes_I \mathbb{R}$ is solvable, but not nilpotent. It is isomorphic to $\mathfrak{osc}(\mathbb{K})/\mathbb{K}z$.

(iv) Let $\mathfrak{g}$ be a 2-dimensional nonabelian Lie algebra with basis x, y satisfying $[x, y] = y$ (Exercise 1.4). Then $D^1(\mathfrak{g}) = \mathbb{K}y$ and $D^2(\mathfrak{g}) = \{0\}$, so that $\mathfrak{g}$ is solvable. On the other hand $C^n(\mathfrak{g}) = \mathbb{K}y$ for each $n > 1$, so that $\mathfrak{g}$ is not nilpotent.

Proposition 4.3. *For a Lie algebra $\mathfrak{g}$, the following assertions hold:*

(i) *If $\mathfrak{g}$ is solvable, then all subalgebras and homomorphic images of $\mathfrak{g}$ are solvable.*

(ii) *Solvability is an* **extension property***: If $\mathfrak{i}$ is a solvable ideal of $\mathfrak{g}$ and $\mathfrak{g}/\mathfrak{i}$ is solvable, then $\mathfrak{g}$ is solvable.*

(iii) *If $\mathfrak{i}$ and $\mathfrak{j}$ are solvable ideals of $\mathfrak{g}$, then the ideal $\mathfrak{i} + \mathfrak{j}$ is solvable.*

(iv) *If $\mathfrak{i} \trianglelefteq \mathfrak{g}$ is an ideal, then the $D^n(\mathfrak{i})$ are ideals in $\mathfrak{g}$.*

Proof.

(i) If $\mathfrak{h} \subseteq \mathfrak{g}$ is a subalgebra, then $D^n(\mathfrak{h}) \subseteq D^n(\mathfrak{g})$ follows by induction. If $\alpha\colon \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras, then we obtain

$$D^n(\alpha(\mathfrak{g})) = \alpha(D^n(\mathfrak{g})) \tag{7}$$

by induction. This implies (i).

(ii) Let $\pi\colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{i}$ be the quotient map. We have already seen in (i) that $\pi(D^n(\mathfrak{g})) = D^n(\pi(\mathfrak{g}))$ for each n. If $\mathfrak{g}/\mathfrak{i}$ is solvable, then $\pi(D^n(\mathfrak{g}))$ vanishes for some $n \in \mathbb{N}$. Now $D^n(\mathfrak{g}) \subseteq \ker \pi = \mathfrak{i}$, so that $D^{n+k}(\mathfrak{g}) \subseteq D^k(\mathfrak{i})$ for each $k \in \mathbb{N}$. If $\mathfrak{i}$ is also solvable, we immediately derive that $\mathfrak{g}$ is solvable.

(iii) The ideal $\mathfrak{j}$ of $\mathfrak{i} + \mathfrak{j}$ is solvable and $(\mathfrak{i} + \mathfrak{j})/\mathfrak{j} \cong \mathfrak{i}/(\mathfrak{i} \cap \mathfrak{j})$ (Proposition 1.24(iii)) is solvable by (i). Hence (ii) implies that $\mathfrak{i} + \mathfrak{j}$ is solvable.

(iv) We only have to observe that for each ideal $\mathfrak{i}$, its commutator algebra $[\mathfrak{i}, \mathfrak{i}]$ also is an ideal (Exercise 1.11). Then (iv) follows by induction. ☐

Example 4.4. If $\mathcal{F} = (V_0, \dots, V_n)$ is a complete flag in the n-dimensional vector space V, then $\mathfrak{g}(\mathcal{F})$ is a solvable Lie algebra. In fact,

$$\mathfrak{g}(\mathcal{F}) \cong \mathfrak{g}_n(\mathcal{F}) \rtimes \mathfrak{gl}_1(\mathbb{K})^n \cong \mathfrak{g}_n(\mathcal{F}) \rtimes \mathbb{K}^n$$

(Example 1.31).

Since $\mathbb{K}^n \cong \mathfrak{g}(\mathcal{F})/\mathfrak{g}_n(\mathcal{F})$ is abelian and $\mathfrak{g}_n(\mathcal{F})$ nilpotent (Example 3.2(iii)), the solvability of $\mathfrak{g}(\mathcal{F})$ follows from Proposition 4.3(ii). Below we shall see that Lie's Theorem provides a converse for solvable subalgebras of $\mathfrak{gl}(V)$, provided the field $\mathbb{K}$ is algebraically closed and of characteristic 0 (such as $\mathbb{K} = \mathbb{C}$); they are always contained in $\mathfrak{g}(\mathcal{F})$ for some complete flag $\mathcal{F}$.

Definition 4.5. Proposition 4.3(iii) shows that every finite dimensional Lie algebra $\mathfrak{g}$ contains a maximal solvable ideal containing all other solvable ideals. This ideal is called the *radical* of $\mathfrak{g}$, and is denoted by $\mathrm{rad}(\mathfrak{g})$.

Remark 4.6. Nilpotency is not an extension property, i.e., the analog of Proposition 4.3(ii) is false for nilpotent Lie algebras: If $\mathfrak{g} = \mathbb{K}x + \mathbb{K}y$ is the 2-dimensional Lie algebra with $[x, y] = y$, then the ideal $\mathfrak{n} := \mathbb{K}y$ and the one-dimensional quotitent algebra $\mathfrak{g}/\mathfrak{n}$ are abelian, hence nilpotent, but $\mathfrak{g}$ is not (cf. Example 4.2(iv)).

4.2 Lie's Theorem

Now we turn to solvable Lie subalgebras $\mathfrak{g}$ of $\mathfrak{gl}(V)$. In this context we do not want to make any assumption on the elements of $\mathfrak{g}$, as in Corollary 3.7.

Theorem 4.7. *Suppose that $\mathbb{K}$ is algebraically closed of characteristic zero. Let V be a nonzero finite dimensional $\mathbb{K}$-vector space and $\mathfrak{g}$ be a solvable subalgebra of $\mathfrak{gl}(V)$. Then there exists a nonzero common eigenvector v for $\mathfrak{g}$, i.e., $\mathfrak{g}(v) \subseteq \mathbb{K}v$.*

Proof. We may w.l.o.g. assume that $\mathfrak{g} \neq \{0\}$. We proceed by induction on the dimension of $\mathfrak{g}$. If $\mathfrak{g} = \mathbb{K}x$, then every eigenvector of x (and such an eigenvector always exists because $\mathbb{K}$ is algebraically closed) satisfies the requirement of the theorem. So let $\dim \mathfrak{g} > 1$ and $\mathfrak{h}$ be a hyperplane in $\mathfrak{g}$ which contains $[\mathfrak{g}, \mathfrak{g}] = D^1(\mathfrak{g})$. Here we use that $D^1(\mathfrak{g})$ is a proper subspace because $\mathfrak{g}$ is solvable. In view of $[\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{h}$, the subspace $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. Now the induction hypothesis provides a nonzero common eigenvector v for $\mathfrak{h}$. If $x(v) = \lambda(x)v$ for $x \in \mathfrak{h}$, then $\lambda \colon \mathfrak{h} \to \mathbb{K}$ is a linear functional and

$$v \in V_\lambda(\mathfrak{h}) := \{w \in V \mid (\forall x \in \mathfrak{h})\, x(w) = \lambda(x)w\}.$$

Suppose that $V_\lambda(\mathfrak{h})$ is $\mathfrak{g}$-invariant and pick $y \in \mathfrak{g}\backslash\mathfrak{h}$. Then there exists a nonzero eigenvector $v_o \in V_\lambda(\mathfrak{h})$ for y. Then v_o is a common eigenvector for $\mathfrak{g} = \mathfrak{h} + \mathbb{K}y$ and the proof is complete.

It remains to show that $V_\lambda(\mathfrak{h})$ is $\mathfrak{g}$-invariant. For this, we calculate as in the proof of Theorem 3.5:

$$yx(w) = xy(w) - [x, y](w) = \lambda(y)x(w) - \lambda([x, y])(w) \quad \text{for} \quad w \in V_\lambda(\mathfrak{h}), x \in \mathfrak{g}, y \in \mathfrak{h}.$$

Hence it suffices to show that $[\mathfrak{g}, \mathfrak{h}] \subseteq \ker \lambda$. For fixed $w \in V_\lambda(\mathfrak{h})$, $x \in \mathfrak{g}$ and $k \in \mathbb{N}$, we consider the space

$$W^k = \mathbb{K}w + \mathbb{K}x(w) + \ldots + \mathbb{K}x^k(w).$$

Since

$$yx^k(w) = xy(x^{k-1}w) - [x, y](x^{k-1}w) \tag{8}$$

and $y(w) = \lambda(y)w$ for $y \in \mathfrak{h}$, we see by induction on k that $\mathfrak{h}(W^k) \subseteq W^k$ for each $k \in \mathbb{N}$: It holds trivially for $k = 0$, and if $k > 0$ and $\mathfrak{h}(W^{k-1}) \subseteq W^{k-1}$, then

$$\mathfrak{h}(x^k(w)) \subseteq x\mathfrak{h}(W^{k-1}) + \mathfrak{h}(W^{k-1}) \subseteq xW^{k-1} + W^{k-1} \subseteq W^k.$$

Now we choose $k_o \in \mathbb{N}$ maximal with respect to the property that

$$\{w, x(w), \ldots, x^{k_o}(w)\}$$

is a basis for W^{k_o}. Then $W^{k_o+m} = W^{k_o}$ for all $m \in \mathbb{N}$, and

$$\mathcal{F} = (\{0\}, W^1, \ldots, W^{k_o})$$

is a complete flag in W^{k_o} which is invariant under $\mathfrak{h}$. Thus, every $y \in \mathfrak{h}$ corresponds to an upper triangular matrix (y_{ij}) with respect to the above basis for W^{k_o}. The diagonal entries y_{ii} of this matrix are all equal to $\lambda(y)$ since $y(w) = \lambda(y)w$ and (8) imply by induction that

$$yx^k(w) \in \lambda(y)x^k(w) + W^{k-1}.$$

In fact, for $k = 1$ this is clear, and if the relation holds for $k - 1$, then

$$\begin{aligned}\big(y - \lambda(y)\mathbf{1}\big)x^k(w) &= xy(x^{k-1}w) - [x,y](x^{k-1}w) - x\big(\lambda(y)x^{k-1}(w)\big)\\ &= x\big((y - \lambda(y)\mathbf{1})(x^{k-1}w)\big) - [x,y](x^{k-1}w) \in xW^{k-2} + \mathfrak{h}(W^{k-1}) \subseteq W^{k-1}.\end{aligned}$$

Since x and y leave the space W^{k_o} invariant, we have

$$[x,y]|_{W^{k_o}} = [x|_{W^{k_o}}, y|_{W^{k_o}}].$$

In particular, $[x,y]|_{W^{k_o}}$ is a commutator of two endomorphisms so that its trace vanishes. Finally $[x,y] \in \mathfrak{h}$ leads to

$$0 = \operatorname{tr}([x,y]|_{W^{k_o}}) = (k_o + 1)\lambda([x,y]),$$

so that $\lambda([x,y]) = 0$ (here we use that $\operatorname{char}\mathbb{K} = 0$). $\square$

For the proof of the preceding theorem, we used that the field $\mathbb{K}$ is algebraically closed and of characteristic zero. The following example shows that both assumptions are crucial.

Example 4.8. (a) If $A \in \operatorname{End}(V)$, then $\mathfrak{g} = \mathbb{K}A$ is a one-dimensional Lie algebra, hence in particular solvable. The existence of an eigenvector of A is equivalent to the existence of a root of the characteristic polynomial, whose existence can only be assured by the algebraic closedness of $\mathbb{K}$. In fact, if $\mathbb{K} \subseteq \mathbb{L}$ is a proper finite algebraic field extension and $x \in \mathbb{L} \setminus \mathbb{K}$, then the multiplication operator $Ay := xy$ has no eigenvectors: If $xy = \lambda y$ for $\lambda \in \mathbb{K}$ and $y \neq 0$, then $(x - \lambda)y = 0$ leads to the contradiction $x = \lambda$.

(b) If $\operatorname{char}\mathbb{K} = 2$, then $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{K})$ is a nilpotent Lie algebra: The basis elements

$$h := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{9}$$

satisfy

$$[h,e] = 2e = 0, \quad [h,f] = -2f = 0 \quad \text{and} \quad [e,f] = h, \tag{10}$$

so that $\mathfrak{sl}_2(\mathbb{K}) \cong \mathfrak{h}_3(\mathbb{K})$ (the 3-dimensional Heisenberg algebra). Nevertheless, $\mathfrak{sl}_2(\mathbb{K})$ has no common eigenvector. Therefore Theorem 4.7 fails.

Theorem 4.9. (Lie's Theorem) *Assume that $\mathbb{K}$ is algebraically closed of characteristic zero. Let V be a finite dimensional vector space and $\mathfrak{g}$ be a solvable subalgebra of $\mathfrak{gl}(V)$. Then there exists a complete $\mathfrak{g}$-invariant flag in V.*

Proof. We may assume that V is nonzero. By Theorem 4.7, there exists a nonzero common $\mathfrak{g}$-eigenvector $v_1 \in V$. Put $V_1 := \mathbb{K}v_1$. Then

$$\alpha\colon \mathfrak{g} \to \mathfrak{gl}(V/V_1), \quad \alpha(x)(v + V_1) := x(v) + V_1$$

defines a representation of $\mathfrak{g}$ on the quotient space V/V_1 (Exercise 1.13) and $\alpha(\mathfrak{g})$ is solvable. Proceeding by induction on $\dim V$, we may assume that there exists an $\alpha(\mathfrak{g})$-invariant complete flag in V/V_1, and the preimage in V, together with $\{0\}$, is a complete $\mathfrak{g}$-invariant flag in V. $\square$

Remark 4.10. If we apply Lie's Theorem 4.9 to $V = \mathfrak{g}$ and $\mathrm{ad}(\mathfrak{g})$, where $\mathfrak{g}$ is solvable, we get a complete flag of ideals

$$\{0\} = \mathfrak{g}_0 < \mathfrak{g}_1 < \ldots < \mathfrak{g}_n = \mathfrak{g}$$

of $\mathfrak{g}$ with $\dim \mathfrak{g}_k = k$. Such a chain is called a *Hölder series* for $\mathfrak{g}$.

Definition 4.11. We call a representation (π, V) of the Lie algebra $\mathfrak{g}$, resp., the corresponding $\mathfrak{g}$-module, *nilpotent* if there exists an $n \in \mathbb{N}$ with $\rho(\mathfrak{g})^n = \{0\}$.

Corollary 4.12. *Suppose that* $\mathrm{char}\,\mathbb{K} = 0$. *Let $\pi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a finite dimensional representation of the solvable Lie algebra $\mathfrak{g}$. Then the restriction to $[\mathfrak{g}, \mathfrak{g}]$ is a nilpotent representation.*

Proof. We may w.l.o.g. assume that $V = \mathbb{K}^n$, so that $\pi(\mathfrak{g}) \subseteq \mathfrak{gl}(V) \cong \mathfrak{gl}_n(\mathbb{K})$ consists of $(n \times n)$-matrices. Let $\overline{\mathbb{K}}$ be an algebraically closed extension of $\mathbb{K}$ (such as $\overline{\mathbb{K}} = \mathbb{C}$ for $\mathbb{K} = \mathbb{R}$). Then the relation $\pi([\mathfrak{g}, \mathfrak{g}])^N = \{0\}$ will follow if we prove that the solvable $\overline{\mathbb{K}}$-Lie algebra $\overline{\mathfrak{g}} := \mathrm{span}_{\overline{\mathbb{K}}}\, \pi(\mathfrak{g}) \subseteq \mathfrak{gl}_n(\overline{\mathbb{K}})$ satisfies $[\overline{\mathfrak{g}}, \overline{\mathfrak{g}}]^N = \{0\}$ for some $N \in \mathbb{N}$. We may therefore assume that $\mathbb{K}$ is algebraically closed, i.e., $\mathbb{K} = \overline{\mathbb{K}}$.

Applying Lie's Theorem to the solvable subalgebra $\pi(\mathfrak{g})$ of $\mathfrak{gl}(V)$, we obtain a complete flag $\mathcal{F}$ with $\pi(\mathfrak{g}) \subseteq \mathfrak{g}(\mathcal{F})$. Then

$$\pi([\mathfrak{g}, \mathfrak{g}]) \subseteq [\mathfrak{g}(\mathcal{F}), \mathfrak{g}(\mathcal{F})] \subseteq \mathfrak{g}_n(\mathcal{F})$$

(cf. Example 1.31) implies the assertion. $\square$

Corollary 4.13. *A Lie algebra $\mathfrak{g}$ over a field of characteristic zero is solvable if and only if its commutator algebra $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent.*

Proof. If $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent, then $\mathfrak{g}$ is solvable because $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is abelian and solvability is an extension property (Proposition 4.3(ii)).

If, conversely, $\mathfrak{g}$ is solvable, then Corollary 4.12 implies that the adjoint representation of $[\mathfrak{g}, \mathfrak{g}]$ on $\mathfrak{g}$, and hence on $[\mathfrak{g}, \mathfrak{g}]$, is nilpotent. From that we derive in particular that $C^N([\mathfrak{g}, \mathfrak{g}]) = \{0\}$ for some $N \in \mathbb{N}$, so that $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent. $\square$

4.3 The Ideal $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]$

In this subsection we assume that $\mathrm{char}\,\mathbb{K} = 0$ and $\mathfrak{g}$ denotes a finite dimensional Lie algebra.

The following lemma will be useful for inductive arguments assertion that certain representations of Lie algebras are nilpotent.

Lemma 4.14. *Let (ρ, V) a representation of $\mathfrak{g}$.*

(a) *Let $\mathfrak{a} \subseteq \mathfrak{g}$ be a subspace for which there exists an $n \in \mathbb{N}$ with $\rho(\mathfrak{a})^n = \{0\}$, and*

(b) *$x \in \mathfrak{g}$ with $[x, \mathfrak{a}] \subseteq \mathfrak{a}$ such that $\rho(x)$ is nilpotent.*

Then there exists an $N \in \mathbb{N}$ with $\rho(\mathfrak{a} + \mathbb{K}x)^N = \{0\}$.

Proof. Replacing $\mathfrak{g}$ by $\rho(\mathfrak{g})$, we may w.l.o.g. assume that $\mathfrak{g} \subseteq \mathfrak{gl}(V)$. Let $m \in \mathbb{N}$ with $x^m = 0$. We claim that $(\mathbb{K}x + \mathfrak{a})^{nm} = \{0\}$.

Let $u = u_1 \cdots u_{nm}$ be a product of elements of $\{x\} \cup \mathfrak{a}$. We have to show that all such products vanish. For $a \in \mathfrak{a}$ we have

$$ax = xa + [a, x] \in xa + \mathfrak{a}.$$

This leads to

$$u_1 \cdots u_{nm} \in \sum_{r=0}^{nm} x^r \mathfrak{a}^t,$$

where t is the number of indices j with $u_j \in \mathfrak{a}$. Hence this product vanishes for $t \geq n$. If $t < n$, then there exists a j with $u_{j+1} \cdots u_{j+m} = x^m = 0$ because in this case at least $nm - t > n(m-1)$ factors are not contained in $\mathfrak{a}$, so that we always find a consecutive product of m such elements. We therefore have in all cases $u_1 \cdots u_{nm} = 0$. $\square$

Proposition 4.15. *For any finite dimensional representation (ρ, V) of the Lie algebra $\mathfrak{g}$, the restriction to the ideal $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]$ is nilpotent, i.e., there exists an $m \in \mathbb{N}$ with $\rho([\mathfrak{g}, \mathrm{rad}(\mathfrak{g})])^m = \{0\}$.*

Proof. Let $\mathfrak{r} := \mathrm{rad}(\mathfrak{g})$ and $\mathfrak{a} := [\mathfrak{g}, \mathfrak{r}]$. According to Corollary 4.12, the representation is nilpotent on the ideal $[\mathfrak{r}, \mathfrak{r}]$. Now let $\mathfrak{t} \subseteq [\mathfrak{g}, \mathfrak{r}]$ be a subspace containing $[\mathfrak{r}, \mathfrak{r}]$, which is maximal with respect to the property that the representation of $\mathfrak{t}$ on V is nilpotent. Note that $\mathfrak{t}$ always is an ideal of $\mathfrak{r}$, hence in particular a subalgebra, because it contains the commutator algebra.

Assume that $\mathfrak{t} \neq [\mathfrak{g}, \mathfrak{r}]$. Then there exists an $x \in \mathfrak{g}$ and $y \in \mathfrak{r}$ with $[x, y] \notin \mathfrak{t}$. The subspace $\mathfrak{b} := \mathfrak{r} + \mathbb{K}x$ is a subalgebra of $\mathfrak{g}$, $\mathfrak{r}$ is a solvable ideal of $\mathfrak{b}$, and $\mathfrak{b}/\mathfrak{r} \cong \mathbb{K}$ is abelian. Therefore $\mathfrak{b}$ is solvable (Proposition 4.3).

We use Corollary 4.12 to see that the representation is nilpotent on $[\mathfrak{b}, \mathfrak{b}]$ and hence that $\rho([x, y])$ is nilpotent. Since $\mathfrak{t} \subseteq \mathfrak{r}$ and $[x, y] \in [\mathfrak{g}, \mathfrak{r}] \subseteq \mathfrak{r}$, we have $\big[[x, y], \mathfrak{t}\big] \subseteq [\mathfrak{r}, \mathfrak{t}] \subseteq \mathfrak{t}$. Finally, the preceding Lemma 4.14 shows that the representation is nilpotent on the subspace $\mathbb{K}[x, y] + \mathfrak{t}$. This contradicts the maximality of $\mathfrak{t}$. We conclude that $\mathfrak{t} = [\mathfrak{g}, \mathfrak{r}]$, so that the representation is nilpotent on $[\mathfrak{g}, \mathfrak{r}]$. $\square$

Applying the preceding proposition to the adjoint representation and using Engel's Theorem 3.8 we get:

Corollary 4.16. *The ideal $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]$ is nilpotent. In particular, $\mathrm{ad}\, x$ is nilpotent on $\mathfrak{g}$ for each $x \in [\mathfrak{g}, \mathrm{rad}\, \mathfrak{g}]$.*

Remark 4.17. Since the ideal $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]$ is nilpotent, it is contained in the nilradical $\mathrm{nil}(\mathfrak{g})$. That it may be strictly smaller follows from the case where $\mathfrak{g}$ is abelian. Then $\mathrm{nil}(\mathfrak{g}) = \mathfrak{g}$ and $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})] = \{0\}$.

4.4 Cartan's Solvability Criterion

This subsection is devoted to a characterization of solvable Lie algebras by properties of their elements. The result will be that $\mathfrak{g}$ is solvable if and only if $\mathrm{tr}(\mathrm{ad}\, x\, \mathrm{ad}\, y) = 0$ for $x \in [\mathfrak{g}, \mathfrak{g}]$ and $y \in \mathfrak{g}$ (Cartan's criterion). Thus, we have to study the linear maps $\mathrm{ad}(x)\colon \mathfrak{g} \to \mathfrak{g}$.

In this subsection we assume that $\mathrm{char}\, \mathbb{K} = 0$.

Lemma 4.18. *For two commuting endomorphisms M, N of a vector space V, the following assertions hold:*

(a) *If M and N are diagonalizable, then $M + N$ is diagonalizable.*

(b) *If M and N are nilpotent, then $M + N$ is nilpotent.*

Proof. (a) Since M and N commute, they are simultaneously diagonalizable, and this implies in particular that $M + N$ is diagonalizable (cf. Exercise 4.1(a)-(c)).

(b) Suppose that $M^m = N^n = 0$. Then $[M, N] = 0$ implies that

$$(M + N)^k = \sum_{i+j=k} \binom{k}{i} M^i N^j.$$

If $k \geq n + m - 1$, then either $i \geq m$ or $j \geq n$, so that all summands vanish. Hence $(M + N)^k = 0$. □

For the following proof we recall the Jordan decomposition of an endomorphism $A \in \mathrm{End}(V)$ of a finite dimensional vector space V which is split in the sense that $f(A) = 0$ for some $f \in \mathbb{K}[X]$ which is a product of linear factors (cf. Theorem A.2). The Jordan decomposition is the uniquely determined additive decomposition $A = A_s + A_n$, where A_s is diagonalizable, A_n is nilpotent and $[A_s, A_n] = 0$.

Proposition 4.19. *Let V be finite dimensional and $x \in \mathfrak{gl}(V)$. If x is nilpotent (diagonalizable), then so is $\mathrm{ad}\, x$.*

This proposition can be obtained by combining Lemma 3.4(i) with Exercise 3.2. We give an alternative proof using the Jordan decomposition.

Proof. Put $L_x\colon \mathfrak{gl}(V) \to \mathfrak{gl}(V), y \mapsto xy$ and $R_x\colon \mathfrak{gl}(V) \to \mathfrak{gl}(V), y \mapsto yx$. Then $\mathrm{ad}\, x = L_x - R_x$ and $[L_x, R_x] = 0$. In view of Lemma 4.18, it suffices to see that L_x and R_x are nilpotent, resp., diagonalizable whenever x has this property.

If $x^n = 0$, then $L_x^n = L_{x^n} = 0 = R_x^n$. If x is diagonalizable, then we represent elements of $\mathfrak{gl}(V)$ as matrices with respect to a basis of eigenvectors of x. We may therefore assume that $x = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Then $L_x E_{jk} = \lambda_j E_{jk}$ and $R_x E_{jk} = \lambda_k E_{jk}$ imply that L_x and R_x are diagonalizable on $\mathfrak{gl}_n(\mathbb{K}) \cong \mathfrak{gl}(V)$. □

Corollary 4.20. *For each endomorphism $x \in \mathfrak{gl}(V)$ of the finite dimensional vector space V over the algebraically closed field $\mathbb{K}$ with Jordan decomposition $x = x_s + x_n$, the Jordan decomposition of $\mathrm{ad}\, x$ is given by*

$$\mathrm{ad}\, x = \mathrm{ad}(x_s) + \mathrm{ad}(x_n).$$

Proof. Proposition 4.19 implies that $\mathrm{ad}(x_s)$ is diagonalizable. Further, $\mathrm{ad}(x_n)$ is nilpotent, and $[\mathrm{ad}(x_s), \mathrm{ad}(x_n)] = \mathrm{ad}[x_s, x_n] = 0$, so that the assertion follows from the uniqueness of the Jordan decomposition of $\mathrm{ad}\, x$. □

Lemma 4.21. *Let V be a finite dimensional vector space over the algebraically closed field $\mathbb{K}$ of characteristic zero and $E \subseteq F$ be subspaces of $\mathfrak{gl}(V)$. Further, let*

$$x \in M := \{y \in \mathfrak{gl}(V) \mid [y, F] \subseteq E\}.$$

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

Proof. Since $\mathbb{K}$ is algebraically closed, x has a Jordan decomposition $x = x_s + x_n$ (Theorem A.2). Representing elements of $\mathfrak{gl}(V)$ as matrices with respect to a basis of eigenvectors of x_s, we may w.l.o.g. assume that $V = \mathbb{K}^n$ and that

$$x_s = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Let Q be the $\mathbb{Q}$-vector space in $\mathbb{K}$ which is spanned by the λ_j. We have to show that $Q = \{0\}$. To do this, we consider an $f \in Q^* := \operatorname{Hom}_{\mathbb{Q}}(Q, \mathbb{Q})$, the dual space (over $\mathbb{Q}$) of Q. We consider the diagonal matrix

$$y := \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_n)).$$

As in the proof of Proposition 4.19, we see that

$$\operatorname{ad}(x_s)E_{ij} = (\lambda_i - \lambda_j)E_{ij} \quad \text{and} \quad \operatorname{ad}(y)E_{ij} = \big(f(\lambda_i) - f(\lambda_j)\big)E_{ij} = f(\lambda_i - \lambda_j)E_{ij}.$$

Now, choose a polynomial $P \in \mathbb{K}[t]$ with

$$P(0) = 0 \quad \text{and} \quad P(\lambda_i - \lambda_j) = f(\lambda_i - \lambda_j)$$

for all pairs (i, j) (Exercise 4.8). Then

$$P\big(\operatorname{ad}(x_s)\big)E_{ij} = f(\lambda_i - \lambda_j)E_{ij} = \operatorname{ad}(y)E_{ij},$$

i.e., $P\big(\operatorname{ad}(x_s)\big) = \operatorname{ad}(y)$. Since $\operatorname{ad}(x_s)$ is the diagonalizable part of $\operatorname{ad}(x)$ by Corollary 4.20, it follows from $x \in M$ and Proposition A.6(iii) that $\operatorname{ad}(x_s)F \subseteq E$. But then $P(0) = 0$ implies $\operatorname{ad}(y)F \subseteq E$, i.e., $y \in M$. Since $x_s y = \operatorname{diag}(\lambda_1 f(\lambda_1), \ldots, \lambda_n f(\lambda_n))$ and $x_n y$ is nilpotent because $(x_n y)^N = x_n^N y^N = 0$ for N sufficiently large, our assumption and $y \in M$ leads to

$$\sum_{k=1}^{n} \lambda_k f(\lambda_k) = \operatorname{tr}(x_s y) = \operatorname{tr}(xy) = 0,$$

and therefore

$$\sum_{k=1}^{n} f(\lambda_k)^2 = f\Big(\sum_{k=1}^{n} \lambda_k f(\lambda_k)\Big) = 0.$$

Hence $f(\lambda_k) = 0$ for all λ_k which yields $f = 0$. Since $f \in Q^*$ was arbitrary, it follows that $Q = \{0\}$. □

Theorem 4.22. (Cartan's Solvability Criterion–linear case) *Let V be a finite dimensional vector space and $\mathfrak{g} < \mathfrak{gl}(V)$. Then the following are equivalent*

(i) $\mathfrak{g}$ *is solvable.*

(ii) $\operatorname{tr}(xy) = 0$ *for all $x \in [\mathfrak{g}, \mathfrak{g}]$ and $y \in \mathfrak{g}$.*

Proof. We may w.l.o.g. assume that $V = \mathbb{K}^n \subseteq \overline{V} := \overline{\mathbb{K}}^n$, where $\overline{\mathbb{K}}$ is an algebraically closed extension of $\mathbb{K}$. Then $\mathfrak{g}$ is solvable if and only if $\overline{\mathfrak{g}} := \operatorname{span}_{\overline{\mathbb{K}}} \mathfrak{g}$ is solvable and (ii) holds for $\mathfrak{g}$ over $\mathbb{K}$ if and only if it holds for $\overline{\mathfrak{g}}$ over $\overline{\mathbb{K}}$. To verify these claims, we only have to observe that $D^j(\overline{\mathfrak{g}}) = \overline{D^j(\mathfrak{g})}$ for every j, which follows from an easy induction. We may therefore assume that $\mathbb{K}$ is algebraically closed.

(i) $\Rightarrow$ (ii): By Lie's Theorem 4.9, there exists a basis for V with respect to which all $x \in \mathfrak{g}$ are upper triangular matrices. In particular, all elements of $[\mathfrak{g}, \mathfrak{g}]$ are given by strictly upper triangular matrices. Multiplying an upper triangular matrix with a strictly upper triangular matrix yields a strictly upper triangular matrix which has zero trace.

(ii) $\Rightarrow$ (i): By Corollary 4.13, it suffices to show that $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent. But to show that, by Corollary 3.7, we only have to prove that every element of $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent. We want to apply Lemma 4.21 with $E = [\mathfrak{g}, \mathfrak{g}]$ and $F = \mathfrak{g}$, i.e., we set

$$M := \{y \in \mathfrak{gl}(V) \mid [y, \mathfrak{g}] \subseteq [\mathfrak{g}, \mathfrak{g}]\} \supseteq \mathfrak{g}.$$

Since the trace is linear, it is enough to show that $\operatorname{tr}([x, x']y) = 0$ for $x, x' \in \mathfrak{g}$ and $y \in M$. But this follows from $[x', y] \subseteq [\mathfrak{g}, \mathfrak{g}]$ and (ii):

$$\operatorname{tr}([x, x']y) = \operatorname{tr}(x[x', y]) \in \operatorname{tr}(\mathfrak{g}[\mathfrak{g}, \mathfrak{g}]) = \{0\}$$

(cf. Exercise 4.7). $\square$

Corollary 4.23. (Cartan's Solvability Criterion–general case) *For a Lie algebra $\mathfrak{g}$ over a field of characteristic zero, the following statements are equivalent*

(i) $\mathfrak{g}$ *is solvable.*

(ii) $\operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$ *for all* $x \in [\mathfrak{g}, \mathfrak{g}]$ *and all* $y \in \mathfrak{g}$.

Proof. (i) $\Rightarrow$ (ii): Proposition 4.3(i) shows that $\operatorname{ad}(\mathfrak{g})$ is solvable, so that (ii) is an immediate consequence of Theorem 4.22.

(ii) $\Rightarrow$ (i): By the Cartan Criterion (Theorem 4.22), $\operatorname{ad}(\mathfrak{g})$ is solvable. Then it follows from $\operatorname{ad}(\mathfrak{g}) \cong \mathfrak{g}/\mathfrak{z}(\mathfrak{g})$ and Proposition 4.3(ii) that $\mathfrak{g}$ is solvable. $\square$

Exercises for Section 4

Exercise 4.1. Let V be a $\mathbb{K}$-vector space and $A \in \operatorname{End}(V)$. We write $V_\lambda(A) := \ker(A - \lambda\mathbf{1})$ for the *eigenspace of* A corresponding to the eigenvalue λ and $V^\lambda(A) := \bigcup_{n\in\mathbb{N}} \ker(A - \lambda\mathbf{1})^n$ for the *generalized eigenspace of* A corresponding to λ.

(a) If $A, B \in \operatorname{End}(V)$ commute, then

$$BV^\lambda(A) \subseteq V^\lambda(A) \quad \text{and} \quad BV_\lambda(A) \subseteq V_\lambda(A)$$

holds for each $\lambda \in \mathbb{K}$.

(b) If $A \in \operatorname{End}(V)$ is diagonalizable and $W \subseteq V$ is an A-invariant subspace, then $A|_W \in \operatorname{End}(W)$ is diagonalizable.

(c) If $A, B \in \operatorname{End}(V)$ commute and both are diagonalizable, then they are simultaneously diagonalizable, i.e., there exists a basis for V which consists of eigenvectors of A and B.

(d) If $\dim V < \infty$ and $\mathcal{A} \subseteq \operatorname{End}(V)$ is a commuting set of diagonalizable endomorphisms, then $\mathcal{A}$ can be simultaneously diagonalized, i.e., V is a direct sum of simultaneous eigenspaces of $\mathcal{A}$.

(e) For any function $\lambda\colon \mathcal{A} \to V$, we write $V_\lambda(\mathcal{A}) = \bigcap_{a\in\mathcal{A}} V_{\lambda(a)}(a)$ for the corresponding simultaneous eigenspace. Show that the sum $\sum_\lambda V_\lambda(\mathcal{A})$ is direct.

(f) If $\mathcal{A} \subseteq \mathrm{End}(V)$ is a finite commuting set of diagonalizable endomorphisms, then $\mathcal{A}$ can be simultaneously diagonalized.

(g)* Find a commuting set of diagonalizable endomorphisms of a vector space V which cannot be diagonalized simultaneously.

Exercise 4.2. Let $\mathfrak{g}$ be a Lie algebra and $\alpha\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of $\mathfrak{g}$ on V. Then $V \rtimes_\alpha \mathfrak{g}$ is a Lie algebra which contains V as an abelian ideal.

Exercise 4.3. (i) For a real Lie algebra $\mathfrak{g}$, we have

$$C^n(\mathfrak{g}_\mathbb{C}) = C^n(\mathfrak{g})_\mathbb{C} \quad \text{and} \quad D^n(\mathfrak{g}_\mathbb{C}) = D^n(\mathfrak{g})_\mathbb{C}.$$

(ii) A finite dimensional Lie algebra $\mathfrak{g}$ is nilpotent (solvable) if and only if $\mathfrak{g}_\mathbb{C}$ is nilpotent (solvable).

Exercise 4.4. Show that for the Heisenberg algebra $\mathfrak{h}_3$, the derivation algebra $\mathrm{der}(\mathfrak{h}_3)$ is isomorphic to $\mathbb{K}^2 \rtimes \mathfrak{gl}_2(\mathbb{K})$, where $\mathrm{ad}(\mathfrak{h}_3) \cong \mathbb{K}^2$ is an abelian ideal. Show that this Lie algebra is neither nilpotent nor solvable.

Exercise 4.5. Show that a representation (π, V) of a Lie algebra $\mathfrak{g}$ on a vector space V is nilpotent if and only if there exists a flag $\mathcal{F}$ in V with $\pi(\mathfrak{g}) \subseteq \mathfrak{g}_n(\mathcal{F})$.

Exercise 4.6. Show that an ideal $\mathfrak{n} \trianglelefteq \mathfrak{g}$ of the Lie algebra $\mathfrak{g}$ is nilpotent if and only if $\mathfrak{g}$ is a nilpotent $\mathfrak{n}$-module with respect to the adjoint representation of $\mathfrak{n}$ on $\mathfrak{g}$.

Exercise 4.7. A symmetric bilinear form $\kappa\colon \mathfrak{g}\times\mathfrak{g} \to \mathbb{K}$ on a Lie algebra $\mathfrak{g}$ is called *invariant* if

$$\kappa([x,y],z) = \kappa(x,[y,z]) \quad \text{for} \quad x,y,z \in \mathfrak{g}.$$

Show that:

(i) The form $\kappa(x,y) := \mathrm{tr}(xy)$ on $\mathfrak{gl}(V)$ is invariant for each finite dimensional vector space V.

(ii) For each representation (π, V) of the Lie algebra $\mathfrak{g}$, the form $\kappa_\pi(x,y) := \mathrm{tr}(\pi(x)\pi(y))$ is invariant.

(iii) For each Lie algebra $\mathfrak{g}$, the Cartan–Killing form $\kappa_\mathfrak{g}(x,y) := \mathrm{tr}(\mathrm{ad}\,x\,\mathrm{ad}\,y)$ is invariant.

(iv) For each invariant symmetric bilinear form κ on $\mathfrak{g}$, its radical

$$\mathrm{rad}(\kappa) = \{x \in \mathfrak{g}\colon \kappa(x,\mathfrak{g}) = \{0\}\}$$

is an ideal.

(v) For any invariant symmetric bilinear form κ on $\mathfrak{g}$, the trilinear map $\Gamma(\kappa)(x,y,z) := \kappa([x,y],z)$ is alternating, i.e.,

$$\Gamma(\kappa)(x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)}) = \mathrm{sgn}(\sigma)\Gamma(\kappa)(x_1,x_2,x_3)$$

for $\sigma \in S_3$ and $x_1, x_2, x_3 \in \mathfrak{g}$.

(vi) $\mathfrak{g}$ is solvable if and only if $\Gamma(\kappa_\mathfrak{g}) = 0$.

Exercise 4.8. (Interpolation polynomials) Let $\mathbb{K}$ be a field, $x_1, \ldots, x_n \in \mathbb{K}$ pairwise different, and $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$. Then there exists a polynomial $f \in \mathbb{K}[t]$ with $f(x_i) = \lambda_i$ for $i = 1, \ldots, n$. Hint: Consider the polynomials $f_i(t) := \prod_{j \neq i} \frac{t - x_j}{x_i - x_j}$ of degree $n - 1$.

Exercise 4.9. Show that the multiplication operator

$$\lambda_X \colon \mathbb{K}[X] \to \mathbb{K}[X], \quad f(X) \mapsto X f(X)$$

has no eigenvector. Conclude in particular that Lie's Theorem fails for infinite dimensional spaces V.

Exercise 4.10. We consider the vector space $V = \mathbb{K}^{(\mathbb{N})}$ with the basis $\{e_i \colon i \in \mathbb{N}\}$. In terms of the rank-one-operators $E_{ij} \in \mathrm{End}(V)$, defined by $E_{ij} e_k = \delta_{jk} e_i$, we consider the Lie algebra

$$\mathfrak{g} := \mathrm{span}\{E_{ij} \colon i \geq j\}$$

(lower triangular matrices). Show that:

(a) $D^n(\mathfrak{g}) = \mathrm{span}\{E_{ij} \colon i \geq j + 2^{n-1}\}$, $n \in \mathbb{N}$. In particular, we have $D^\infty(\mathfrak{g}) := \bigcap_{n \in \mathbb{N}} D^n(\mathfrak{g}) = \{0\}$, i.e., $\mathfrak{g}$ is *residually solvable.* (Exercise 4.11).

(b) $\mathfrak{g} = \bigcup_n \mathfrak{g}_n$ for an increasing sequence of finite dimensional solvable subalgebras $\mathfrak{g}_n$ ($\mathfrak{g}$ is *locally solvable*).

(c) $\mathfrak{g}$ has no common eigenvector in V (compare with Lie's Theorem).

Exercise 4.11. Show that, for a Lie algebra $\mathfrak{g}$, the following are equivalent:

(a) $D^\infty(\mathfrak{g}) := \bigcap_{n \in \mathbb{N}} D^n(\mathfrak{g}) = \{0\}$.

(b) $\mathfrak{g}$ is residually solvable in the sense that all homomorphisms $\varphi \colon \mathfrak{g} \to \mathfrak{s}$, $\mathfrak{s}$ solvable, separate the points of $\mathfrak{g}$, i.e., for every non-zero $x \in \mathfrak{g}$, there exists a homomorphism into a solvable Lie algebra with $\varphi(x) \neq 0$.

Hint: All quotients $\mathfrak{g}/D^n(\mathfrak{g})$ are solvable.

Exercise 4.12. Show that:

(a) A finite dimensional Lie algebra $\mathfrak{g}$ is solvable if and only if there exists a sequence

$$\{0\} = \mathfrak{g}_0 \subseteq \mathfrak{g}_1 \subseteq \ldots \subseteq \mathfrak{g}_n = \mathfrak{g}$$

of subalgebras with $[\mathfrak{g}_i, \mathfrak{g}_i] \subseteq \mathfrak{g}_{i-1}$ for $i = 1, \ldots, n$.

(b) If $\mathfrak{g}$ is solvable, then there exists a sequence as in (a), satisfying, in addition, $\dim \mathfrak{g}_i = i$. Conclude that

$$\mathfrak{g}_{i+1} \cong \mathfrak{g}_i \rtimes_{D_i} \mathbb{K} \quad \text{for some} \quad D_i \in \mathrm{der}(\mathfrak{g}_i), \quad i = 1, \ldots, n-1.$$

This means that

$$\mathfrak{g} \cong \Big(\cdots ((\mathbb{K} \rtimes_{D_1} \mathbb{K}) \rtimes_{D_2} \mathbb{K}) \cdots \rtimes_{D_{n-1}} \mathbb{K} \Big).$$

Exercise 4.13. (Is there a Cartan Criterion for nilpotent Lie algebras?)

(a) If $\mathfrak{g}$ is nilpotent, then $\kappa_{\mathfrak{g}} = 0$.

(b) Consider the Lie algebra $\mathfrak{g} = \mathbb{C}^2 \rtimes_D \mathbb{C}$, where $\mathbb{C}^2$ is considered as an abelian Lie algebra and $D = \mathrm{diag}(1, i)$. This Lie algebra is not nilpotent, but $\kappa_{\mathfrak{g}} = 0$. If we consider $\mathfrak{g}$ as a 6-dimensional real Lie algebra, then its Cartan–Killing form also vanishes. Conclude that it is NOT true that a Lie algebra is nilpotent if and only if its Cartan–Killing form vanishes.

5 Semisimple Lie Algebras

In this section we encounter a third class of Lie algebras. Semisimple Lie algebras are a counterpart to the solvable and nilpotent Lie algebras because their ideal structure is quite simple. They can be decomposed as a direct sum of simple ideals. On the other hand, they have a rich geometric structure which even makes a complete classification of finite dimensional semisimple Lie algebras over $\mathbb{R}$ or any algebraically closed field $\mathbb{K}$ is possible. We later show that every finite dimensional Lie algebra is a semidirect sum of its solvable radical and a semisimple subalgebra (cf. Levi's Theorem 6.6).

Throughout, $\mathbb{K}$ will be a field of characteristic zero and all Lie algebras are finite dimensional.

Definition 5.1. Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Then $\mathfrak{g}$ is called *semisimple* if its radical is trivial, i.e., $\mathrm{rad}(\mathfrak{g}) = \{0\}$. The Lie algebra $\mathfrak{g}$ is called *simple* if $\{0\}$ and $\mathfrak{g}$ are the only ideals of $\mathfrak{g}$ and it is non-abelian (which excludes the one-dimensional algebras).

Lemma 5.2. *Every simple Lie algebra is semisimple.*

Proof. Let $\mathfrak{g}$ be a simple Lie algebra. Since the commutator algebra $[\mathfrak{g}, \mathfrak{g}]$ is a nonzero ideal of $\mathfrak{g}$, it coincides with $\mathfrak{g}$. Hence $\mathfrak{g}$ is not solvable. Therefore $\mathrm{rad}(\mathfrak{g})$ is a proper ideal and therefore $\mathrm{rad}(\mathfrak{g}) = \{0\}$. □

We shall see in Proposition 5.14 that a Lie algebra is semisimple if and only if it is a direct sum of simple ideals.

5.1 Cartan's Semisimplicity Criterion

In this subsection we obtain the characterization of semisimple Lie algebras in terms of the Cartan–Killing form which can be defined for any Lie algebra.

Definition 5.3. In connection with the Cartan criterion for solvable Lie algebras, we have seen that the bilinear form

$$\kappa_{\mathfrak{g}} \colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{K}, \quad \kappa_{\mathfrak{g}}(x, y) := \mathrm{tr}(\mathrm{ad}\, x\, \mathrm{ad}\, y)$$

on a finite dimensional Lie algebra is of interest. It is called the *Cartan–Killing form* of $\mathfrak{g}$. Its compatibility with the Lie algebra structure is expressed by its invariance

$$\kappa_{\mathfrak{g}}([x, y], z) = \kappa_{\mathfrak{g}}(x, [y, z]) \quad \text{for} \quad x, y, z, \in \mathfrak{g}$$

(Exercise 4.7). If $\mathfrak{g}$ is clear from the context, we sometimes write κ instead of $\kappa_{\mathfrak{g}}$.

Definition 5.4. Let V be a vector space and $\beta \colon V \times V \to \mathbb{K}$ be a symmetric bilinear form. For a subset $W \subseteq V$, we then write

$$W^{\perp,\beta} := \{v \in V \mid (\forall w \in W)\, \beta(v, w) = 0\}$$

for the the orthogonal subspace of W with respect to β by $W^{\perp,\beta}$. The set $\mathrm{rad}(\beta) := V^{\perp,\beta}$ is called the *radical* of β. The form is called *degenerate* if $\mathrm{rad}(\beta) \neq \{0\}$.

Using this notation, we can reformulate the Cartan Criterion 4.22 as follows:

Remark 5.5. In terms of the Cartan–Killing form, Cartan's Solvability Criterion asserts that $\mathfrak{g}$ is solvable if and only if $[\mathfrak{g}, \mathfrak{g}] \subseteq \mathrm{rad}(\kappa_{\mathfrak{g}})$ (cf. Exercise 5.8 for the fact that $\mathrm{rad}(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]^{\perp}$ holds for every finite dimensional Lie algebra $\mathfrak{g}$).

Example 5.6. (i) With respect to the basis (h, e, f) of $\mathfrak{sl}_2(\mathbb{K})$ with

$$[h, e] = 2e, \quad [h, f] = -2f \quad \text{and} \quad [e, f] = h$$

from (10), an easy calculation leads to the following matrix for the Cartan–Killing form:

$$\kappa = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{pmatrix}.$$

(ii) With respect to the basis (x, y, z) for $\mathfrak{so}_3(\mathbb{R})$ with

$$[x, y] = z, \quad [y, z] = x, \quad [z, x] = y$$

(Example 1.35), the Cartan–Killing form has the matrix

$$\kappa = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

(iii) With respect to the basis (h, p, q, z) for the oscillator algebra $\mathfrak{osc}(\mathbb{K})$ from Example 1.30, the Cartan–Killing form has the matrix

$$\kappa = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Here we see explicitly that $\kappa(\mathfrak{osc}(\mathbb{K}), [\mathfrak{osc}(\mathbb{K}), \mathfrak{osc}(\mathbb{K})]) = \kappa(\mathfrak{osc}(\mathbb{K}), \mathfrak{h}_3(\mathbb{K})) = \{0\}$ which also follows from Cartan's criterion for solvability (Corollary 4.23).

In general, the Cartan–Killing form of a subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ cannot be calculated in terms of the Cartan–Killing form of $\mathfrak{g}$, but for ideals we have:

Lemma 5.7. *For any ideal* $\mathfrak{i} \trianglelefteq \mathfrak{g}$, $\kappa_{\mathfrak{i}} = \kappa_{\mathfrak{g}}|_{\mathfrak{i} \times \mathfrak{i}}$.

Proof. If the image of $A \in \operatorname{End}(\mathfrak{g})$ is contained in $\mathfrak{i}$, then we pick a basis for $\mathfrak{g}$ which starts with a basis for $\mathfrak{i}$. With respect to this basis, we can write A as a block matrix

$$A = \begin{pmatrix} A|_{\mathfrak{i}} & * \\ 0 & 0 \end{pmatrix},$$

and this shows $\operatorname{tr}(A) = \operatorname{tr}(A|_{\mathfrak{i}})$. We apply this to $A = \operatorname{ad}(x)\operatorname{ad}(y)$ for $x, y \in \mathfrak{i}$ to obtain $< \operatorname{tr}\big(\operatorname{ad}(x)\operatorname{ad}(y)\big) = \operatorname{tr}\big(\operatorname{ad}(x)|_{\mathfrak{i}}\operatorname{ad}(y)|_{\mathfrak{i}}\big) = \kappa_{\mathfrak{i}}(x, y)$. ☐

Remark 5.8. Let $\mathfrak{g}$ be a finite dimensional real Lie algebra. Since a basis for $\mathfrak{g}$ is also a (complex) basis for $\mathfrak{g}_{\mathbb{C}}$, one immediately sees that

$$\kappa_{\mathfrak{g}} = \kappa_{\mathfrak{g}_{\mathbb{C}}}|_{\mathfrak{g} \times \mathfrak{g}}$$

(cf. Exercise 5.3).

Lemma 5.9. *For any ideal* $\mathfrak{j}$ *of a Lie algebra* $\mathfrak{g}$, *the following assertions hold:*

(i) *Its orthogonal space* $\mathfrak{j}^{\perp}$ *with respect to* $\kappa_{\mathfrak{g}}$ *also is an ideal.*

(ii) $\mathfrak{j} \cap \mathfrak{j}^{\perp}$ *is a solvable ideal.*

(iii) *If* $\mathfrak{j}$ *or* $\mathfrak{g}$ *is semisimple, then* $\mathfrak{g}$ *decomposes as a direct sum* $\mathfrak{g} = \mathfrak{j} \oplus \mathfrak{j}^\perp$ *of Lie algebras.*

Proof. (i) For $x \in \mathfrak{j}^\perp$, $y \in \mathfrak{g}$ and $z \in \mathfrak{j}$, we find $\kappa_\mathfrak{g}([x,y],z) = \kappa_\mathfrak{g}(x,[y,z]) = 0$, so that $\mathfrak{j}^\perp$ is an ideal of $\mathfrak{g}$.

(ii) For $\mathfrak{i} := \mathfrak{j} \cap \mathfrak{j}^\perp$, the Cartan–Killing form $\kappa_\mathfrak{g}$ vanishes on $\mathfrak{i} \times \mathfrak{i}$. Hence $\mathrm{rad}(\kappa_\mathfrak{i}) = \mathfrak{i}$ by Lemma 5.7. In particular, $\mathfrak{i}$ is solvable by Remark 5.5.

(iii) If $\mathfrak{j}$ is semisimple, then (ii) implies that $\mathfrak{j} \cap \mathfrak{j}^\perp \subseteq \mathrm{rad}(\mathfrak{j}) = \{0\}$. If $\mathfrak{g}$ is semisimple, we likewise obtain $\mathfrak{j} \cap \mathfrak{j}^\perp \subseteq \mathrm{rad}(\mathfrak{g}) = \{0\}$. Since $\mathfrak{j}^\perp$ is the kernel of the linear map $\mathfrak{g} \to \mathfrak{j}^*, x \mapsto \kappa_\mathfrak{g}(x,\cdot)$, we have $\dim \mathfrak{j}^\perp \geq \dim \mathfrak{g} - \dim \mathfrak{j}$, which implies $\mathfrak{j} + \mathfrak{j}^\perp = \mathfrak{g}$, so that $\mathfrak{g}$ is a direct sum of the vector subspaces $\mathfrak{j}$ and $\mathfrak{j}^\perp$. As both are ideals by (i), $[\mathfrak{j}, \mathfrak{j}^\perp] \subseteq \mathfrak{j} \cap \mathfrak{j}^\perp = \{0\}$, and we obtain a direct sum of Lie algebras. □

Proposition 5.10. *For a semisimple Lie algebra* $\mathfrak{g}$*, the following assertions hold:*

(i) $\mathfrak{g}$ *is perfect, i.e.,* $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$.

(ii) *Each ideal* $\mathfrak{n} \trianglelefteq \mathfrak{g}$ *is semisimple and there exists a semisimple ideal* $\mathfrak{c}$ *with* $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{c}$.

(iii) *All homomorphic images of* $\mathfrak{g}$ *are semisimple.*

Proof. (i) In view of Lemma 5.9(iii), $\mathfrak{g}$ decomposes as $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}] \oplus [\mathfrak{g},\mathfrak{g}]^\perp$. Then the ideal $\mathfrak{a} := [\mathfrak{g},\mathfrak{g}]^\perp$ satisfies $[\mathfrak{a},\mathfrak{a}] \subseteq \mathfrak{a} \cap [\mathfrak{g},\mathfrak{g}] = \{0\}$, so that $\mathfrak{a}$ is an abelian ideal, hence trivial.

(ii) Using Lemma 5.9(iii) again, we write $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{c}$ with $\mathfrak{c} := \mathfrak{n}^\perp$. Then $\mathrm{rad}(\mathfrak{n})$ commutes with $\mathfrak{n}^\perp$, hence is a solvable ideal of $\mathfrak{g}$ and therefore trivial. This shows that $\mathfrak{n}$ is semisimple. The same argument shows that $\mathfrak{c}$ is semisimple.

(iii) follows from (ii) because $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{c}$ implies $\mathfrak{g}/\mathfrak{n} \cong \mathfrak{c}$. □

We can also characterize semisimplicity in terms of the Cartan–Killing form.

Theorem 5.11. (Cartan's Semisimplicity Criterion) *A Lie algebra* $\mathfrak{g}$ *is semisimple if and only if* $\kappa_\mathfrak{g}$ *is nondegenerate, i.e.,* $\mathrm{rad}(\kappa_\mathfrak{g}) = \{0\}$.

Proof. With Lemma 5.9(ii), we see that $\mathfrak{g} \cap \mathfrak{g}^\perp = \mathrm{rad}(\kappa_\mathfrak{g})$ is a solvable ideal, so that $\mathrm{rad}(\kappa_\mathfrak{g}) \subseteq \mathrm{rad}(\mathfrak{g})$. In particular, $\kappa_\mathfrak{g}$ is nondegenerate if $\mathfrak{g}$ is semisimple.

Suppose, conversely, that $\mathfrak{g}$ is not semisimple and put $\mathfrak{r} := \mathrm{rad}(\mathfrak{g}) \neq \{0\}$. Let $n \in \mathbb{N}_0$ be maximal with $\mathfrak{h} := D^n(\mathfrak{r}) \neq \{0\}$. Then $\mathfrak{h}$ is an abelian ideal of $\mathfrak{g}$. For $x \in \mathfrak{h}$ and $y \in \mathfrak{g}$, we then have $(\mathrm{ad}\, x\, \mathrm{ad}\, y)\mathfrak{g} \subseteq \mathfrak{h}$ and therefore $(\mathrm{ad}\, x\, \mathrm{ad}\, y)^2 = 0$. This implies that $\kappa_\mathfrak{g}(x,y) = \mathrm{tr}(\mathrm{ad}\, x\, \mathrm{ad}\, y) = 0$. Since $y \in \mathfrak{g}$ was arbitrary, this means that $x \in \mathrm{rad}(\kappa_\mathfrak{g})$, i.e., $\kappa_\mathfrak{g}$ is degenerate. □

Remark 5.12. In view of $\mathrm{rad}(\mathfrak{g})_\mathbb{C} = \mathrm{rad}(\mathfrak{g}_\mathbb{C})$ (cf. Exercise 5.4(i)), a real Lie algebra $\mathfrak{g}$ is semisimple if and only if its complexification $\mathfrak{g}_\mathbb{C}$ is semisimple.

For simplicity the situation is a litlle more complicated, as the following proposition shows. [10]

Proposition 5.13. *Let* $\mathfrak{g}$ *be a simple real Lie algebra. Then either*

(i) $\mathfrak{g}_\mathbb{C}$ *is simple, or*

[10]Recall that, for a real vector space V, an endomorphism $I \in \mathrm{End}(V)$ is called a *complex structure* if $I^2 = -\mathbf{1}$. Then $(x+iy)v := xv + yIv$ defines on V the structure of a complex vector space, denoted (V, I). The *opposite complex vector space* is $(V, -I)$, where the scalar multiplication is defined by $(x+iy)v := xv - yIv$.

(ii) $\mathfrak{g}$ *carries a complex structure* $I \in \mathrm{End}(\mathfrak{g})$ *turning it into a complex Lie algebra* $(\mathfrak{g}, I)$. *In this case* $\mathfrak{g}_{\mathbb{C}} \cong (\mathfrak{g}, I) \oplus (\mathfrak{g}, -I)$.

If, conversely, $\mathfrak{g}$ *is a real Lie algebra for which* $\mathfrak{g}_{\mathbb{C}}$ *is simple, then* $\mathfrak{g}$ *is simple.*

Proof. Suppose that $\mathfrak{g}_{\mathbb{C}}$ is not simple and let $\mathfrak{a} \trianglelefteq \mathfrak{g}_{\mathbb{C}}$ be a proper complex ideal. Let σ denote the complex conjugation on $\mathfrak{g}_{\mathbb{C}}$. We claim that $\mathfrak{g}_{\mathbb{C}} \cong \mathfrak{a} \oplus \sigma(\mathfrak{a})$ is a direct sum of Lie algebras. First we observe that $\sigma(\mathfrak{a})$ is also a complex ideal, so that the two subspaces $\mathfrak{a} \cap \sigma(\mathfrak{a})$ and $\mathfrak{a} + \sigma(\mathfrak{a})$ are both σ-invariant complex ideals of $\mathfrak{g}_{\mathbb{C}}$. For any σ-invariant complex subspace $\mathfrak{b} \subseteq \mathfrak{g}_{\mathbb{C}}$, we obtain

$$\mathfrak{b} = \{x \in \mathfrak{b} \colon \sigma(x) = x\} \oplus \{x \in \mathfrak{b} \colon \sigma(x) = -x\} \cong (\mathfrak{g} \cap \mathfrak{b})_{\mathbb{C}}$$

from the eigenspace decomposition of $\sigma|_{\mathfrak{b}}$ (Exercise A.1(ii)).

As $\mathfrak{a}$ is a proper ideal, $\mathfrak{g} \not\subseteq \mathfrak{a}$, and this implies that the ideal $\mathfrak{a} \cap \mathfrak{g}$ of $\mathfrak{g}$ is $\{0\}$. Therefore $\mathfrak{a} \cap \sigma(\mathfrak{a}) = \{0\}$. Further, the ideal $\mathfrak{g} \cap (\mathfrak{a} + \sigma(\mathfrak{a}))$ of $\mathfrak{g}$ is non-zero, hence all of $\mathfrak{g}$, and thus $\mathfrak{a} + \sigma(\mathfrak{a}) = \mathfrak{g}_{\mathbb{C}}$. This proves our claim. Note that the ideal property of $\mathfrak{a}$ and $\sigma(\mathfrak{a})$ implies $[\mathfrak{a}, \sigma(\mathfrak{a})] = \{0\}$.

It remains to show that $\mathfrak{a} \cong \mathfrak{g}$ as real Lie algebras. To this end, we consider the real linear map

$$\varphi \colon \mathfrak{a} \to \mathfrak{g}, \quad z \mapsto z + \sigma(z).$$

Then

$$[\varphi(z), \varphi(w)] = [z + \sigma(z), w + \sigma(w)] = [z, w] + [\sigma(z), \sigma(w)] = [z, w] + \sigma([z, w]) = \varphi([z, w])$$

shows that φ is a homomorphism of Lie algebras. As $\ker \varphi = \mathfrak{a} \cap i\mathfrak{g} = \{0\}$ and $\mathfrak{g}_{\mathbb{C}} = \mathfrak{a} \oplus \sigma(\mathfrak{a})$, the map φ is bijective, hence an isomorphism of Lie algebras.

We finally assume that $\mathfrak{g}$ is a real Lie algebra for which $\mathfrak{g}_{\mathbb{C}}$ is simple. Remark 5.12 implies that $\mathfrak{g}$ is semisimple. If $\mathfrak{a} \trianglelefteq \mathfrak{g}$ is a proper non-zero ideal, then there exists a proper decomposition $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ (Proposition 5.10(ii)) and then $\mathfrak{g}_{\mathbb{C}} \cong \mathfrak{a}_{\mathbb{C}} \oplus \mathfrak{b}_{\mathbb{C}}$ is not simple. Therefore $\mathfrak{g}$ must be simple. □

Proposition 5.14. *Let* $\mathfrak{g}$ *be a semisimple Lie algebra. Then there exist simple ideals* $\mathfrak{g}_1, \ldots, \mathfrak{g}_k$ *of* $\mathfrak{g}$ *with*

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_k.$$

Every ideal $\mathfrak{i} \trianglelefteq \mathfrak{g}$ *is semisimple and a direct sum* $\mathfrak{i} = \bigoplus_{j \in I} \mathfrak{g}_j$ *for some subset* $I \subseteq \{1, \ldots, k\}$. *Conversely, each direct sum of simple Lie algebras is semisimple.*

Proof. Let $\mathfrak{g}_1 \trianglelefteq \mathfrak{g}$ be a minimal non-zero ideal and write $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{c}_1$ (Proposition 5.10(ii)). Then every ideal of $\mathfrak{g}_1$ is also an ideal of $\mathfrak{g}$, and therefore $\mathfrak{g}_1$ is a simple Lie algebra. As $\mathfrak{c}_1$ is also semisimple by Proposition 5.10(ii), an easy induction on $\dim \mathfrak{g}$ now implies that $\mathfrak{g}$ decomposes as a direct sum

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_k$$

of simple ideals.

Finally, let $\mathfrak{i} \neq \{0\}$ be an ideal of $\mathfrak{g}$. Let $\pi_j \colon \mathfrak{g} \to \mathfrak{g}_j$ be the projections. Then we have $\pi_j(\mathfrak{i}) \neq \{0\}$ for at least one j. But since π_j is surjective $\pi_j(\mathfrak{i})$ is an ideal of $\mathfrak{g}_j$ and therefore equal to $\mathfrak{g}_j$ by simplicity. Thus

$$\mathfrak{g}_j = [\mathfrak{g}_j, \mathfrak{g}_j] = [\mathfrak{g}_j, \pi_j(\mathfrak{i})] = [\mathfrak{g}_j, \mathfrak{i}] \subseteq \mathfrak{i}$$

because $[\mathfrak{g}_j, \pi_\ell(\mathfrak{i})] = \{0\}$ for $\ell \neq j$. The argument shows that every $\mathfrak{g}_j$ with $\pi_j(\mathfrak{i}) \neq \{0\}$ is contained in $\mathfrak{i}$. But then $\mathfrak{i}$ is the direct sum of these $\mathfrak{g}_j$.

The preceding argument shows in particular that every nonzero ideal $\mathfrak{r}$ of a direct sum $\mathfrak{g} := \oplus_j \mathfrak{g}_j$ of simple Lie algebras contains a simple ideal, hence cannot be solvable because simple Lie algebras are not solvable (Lemma 5.2). We conclude that $\mathrm{rad}(\mathfrak{g}) = \{0\}$, so that $\mathfrak{g}$ is semisimple. □

Proposition 5.15. *If $\mathfrak{g}$ is a complex simple Lie algebra, then $\mathfrak{g}$ is also simple as a real Lie algebra.*

Proof. Suppose that $\{0\} \neq \mathfrak{a} \trianglelefteq \mathfrak{g}$ is a minimal non-zero ideal of the underlying real Lie algebra. As $i[\mathfrak{g}, \mathfrak{a}] = [i\mathfrak{g}, \mathfrak{a}] = [\mathfrak{g}, \mathfrak{a}] \subseteq \mathfrak{a}$, $[\mathfrak{g}, \mathfrak{a}]$ is a complex subspace of $\mathfrak{a}$. As the center of $\mathfrak{g}$ is trivial, $[\mathfrak{g}, \mathfrak{a}] \neq \{0\}$, and thus the minimality of $\mathfrak{a}$ leads to $\mathfrak{a} = [\mathfrak{g}, \mathfrak{a}]$, showing that $\mathfrak{a}$ is a complex subspace of $\mathfrak{g}$. Finally the simplicity of $\mathfrak{g}$ as a complex Lie algebra shows that $\mathfrak{a} = \mathfrak{g}$. □

Example 5.16. (a) We claim that the Lie algebras $\mathfrak{sl}_2(\mathbb{K})$, $\mathfrak{so}_3(\mathbb{K})$ and $\mathfrak{su}_2(\mathbb{C})$ are simple: We have seen in Example 5.6 that the Cartan–Killing forms of $\mathfrak{sl}_2(\mathbb{K})$ and $\mathfrak{so}_3(\mathbb{K})$ are nondegenerate, so that they are semisimple, hence simple because they are 3-dimensional (Exercise 5.1). Further, $\mathfrak{su}_2(\mathbb{C})$ is a real form of the complex simple Lie algebra $\mathfrak{sl}_2(\mathbb{C})$, hence simple by the second half of Proposition 5.13.

(b) We have just seen that $\mathfrak{sl}_2(\mathbb{C})$ is a complex simple Lie algebra, so that Proposition 5.15 implies that $\mathfrak{sl}_2(\mathbb{C})$ is also simple as a real 6-dimensional Lie algebra.

Later we shall see how one can use root decomposition to verify the simplicity of larger classes of matric Lie algebras.

In Example 1.14, we have seen that the adjoint representation provides derivations on the Lie algebra. In the case of semisimple Lie algebras, this representation in fact gives *all* derivations.

Theorem 5.17. *For a semisimple Lie algebra $\mathfrak{g}$ all derivations are inner, i.e.,*

$$\mathrm{ad}(\mathfrak{g}) = \mathrm{der}(\mathfrak{g}).$$

Proof. By Proposition 1.15(i), $\mathrm{ad}\,\mathfrak{g} \trianglelefteq \mathrm{der}(\mathfrak{g})$ is an ideal, and since $\mathfrak{z}(\mathfrak{g}) = \{0\}$, the ideal $\mathrm{ad}(\mathfrak{g}) \cong \mathfrak{g}$ is semisimple. Therefore $\mathrm{der}\,\mathfrak{g}$ decomposes as a direct sum $\mathfrak{j} \oplus \mathrm{ad}\,\mathfrak{g}$ for the orthogonal complement $\mathfrak{j}$ of $\mathrm{ad}(\mathfrak{g})$ with respect to the Cartan–Killing form of $\mathrm{der}(\mathfrak{g})$ (Lemma 5.9(iii)). For $\delta \in \mathfrak{j}$ and $x \in \mathfrak{g}$ we then have

$$0 = [\delta, \mathrm{ad}\,x](y) = \delta([x, y]) - [x, \delta(y)] = [\delta(x), y] = \mathrm{ad}\,\big(\delta(x)\big).$$

This means that $\delta(x) \in \mathfrak{z}(\mathfrak{g}) = \{0\}$, i.e., $\delta = 0$. We conclude that $\mathfrak{j} = \{0\}$, so that $\mathrm{der}(\mathfrak{g}) = \mathrm{ad}\,\mathfrak{g}$. □

5.2 Weyl's Theorem on Complete Reducibility

We have already seen how Engel's Theorem and Lie's Theorem provide important information on representations of nilpotent, resp., solvable Lie algebras. For semisimple Lie algebras, Weyl's Theorem, which asserts that each representation of a semisimple Lie algebra is completely reducible, plays a similar role. The crucial tool needed for the proof of Weyl's Theorem is the Casimir element.

Definition 5.18. (cf. Exercise 4.7) A symmetric bilinear form $\beta\colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{K}$ on a Lie algebra $\mathfrak{g}$ is called *invariant* if

$$\beta([x, y], z) = \beta(x, [y, z]) \quad \text{for} \quad x, y, z \in \mathfrak{g}.$$

Definition 5.19. Let β be a nondegenerate invariant symmetric bilinear form on the Lie algebra $\mathfrak{g}$, $x_1, \ldots, x_k$ a basis for $\mathfrak{g}$ and $x^1, \ldots, x^k$ the dual basis with respect to β, i.e., $\beta(x_i, x^j) = \delta_{ij}$ (Kronecker delta). For any Lie algebra homomorphism $\rho\colon \mathfrak{g} \to \mathcal{A}_L$, $\mathcal{A}$ an associative algebra, we define the *Casimir element*

$$\Omega(\beta, \rho) := \sum_{i=1}^{k} \rho(x_i)\rho(x^i).$$

The same argument that shows the independence of the trace of an operator (defined as the sum of its diagonal matrix entries) of the choice of the basis, shows that $\Omega(\beta, \rho)$ does not depend on the choice of the basis $x_1, \ldots, x_k$ (cf. Exercise 5.10).

The Casimir element $\Omega(\beta, \rho)$ is a useful tool for the study of representations since it commutes with $\rho(\mathfrak{g})$:

Lemma 5.20. *For each nondegenerate invariant symmetric bilinear form β on $\mathfrak{g}$ and each homomorphism $\rho\colon \mathfrak{g} \to \mathcal{A}_L$, the Casimir element $\Omega(\beta, \rho) \in \mathcal{A}$ commutes with $\rho(\mathfrak{g})$.*

Proof. Let $z \in \mathfrak{g}$. Then we have

$$\operatorname{ad} z(x_j) = \sum_{k=1}^{n} a_{kj} x_k \quad \text{and} \quad \operatorname{ad} z(x^j) = \sum_{k=1}^{n} a^{kj} x^k$$

with two matrices (a_{ij}) and (a^{ij}) in $M_k(\mathbb{K})$. Then

$$a_{kj} = \beta([z, x_j], x^k) = -\beta(x_j, [z, x^k]) = -a^{jk},$$

and with this relation we obtain

$$\begin{aligned}
[\rho(z), \Omega] &= \sum_{j=1}^{n} [\rho(z), \rho(x_j)\rho(x^j)] = \sum_{j=1}^{n} [\rho(z), \rho(x_j)]\rho(x^j) + \rho(x_j)[\rho(z), \rho(x^j)] \\
&= \sum_{j=1}^{n} \rho([z, x_j])\rho(x^j) + \rho(x_j)\rho([z, x^j]) = \sum_{j,k=1}^{n} a_{kj}\rho(x_k)\rho(x^j) + a^{kj}\rho(x_j)\rho(x^k) \\
&= \sum_{j,k=1}^{n} a_{kj}\rho(x_k)\rho(x^j) - a_{jk}\rho(x_j)\rho(x^k) = 0.
\end{aligned}$$

□

Lemma 5.21. (Fitting decomposition) *Let V be a finite dimensional vector space and $T \in \operatorname{End}(V)$. If $V^+(T) := \bigcap_{n \in \mathbb{N}} T^n(V)$ and $V^0(T) = \bigcup_n \ker(T^n)$, then*

$$V = V^0(T) \oplus V^+(T).$$

The space $V^+(T)$ is called the *Fitting one component* of T. In this context the generalized eigenspace $V^0(T)$ is called the *Fitting null component* of T.

Proof. The sequence $T^n(V)$ is decreasing and $\dim V < \infty$ implies that there exists some n with $T^{n+1}(V) = T^n(V)$, so that $T^n(V) = T^+(V)$. As

$$\dim T^n(V) + \dim \ker(T^n) = \dim V$$

is independent of n, it follows $\ker(T^n) = \ker(T^{n+1}) = V^0(T)$. Then $T|_{T^n(V)}\colon T^n(V) \to T^n(V)$ is surjective, hence bijective and on the intersection $V^0(T) \cap V^+(T)$, the restriction of T is nilpotent and bijective at the same time, which leads to $V^0(T) \cap V^+(T) = \{0\}$. Finally

$$\dim V^+(T) + \dim V^0(T) = \dim T^n(V) + \dim \ker(T^n) = \dim V$$

implies that $V^+(T) + V^0(T) = V$, and this proves the lemma. □

Proposition 5.22. *Let $\mathfrak{g}$ be a semisimple Lie algebra and (ρ, V) a finite dimensional representation. Then V is the direct sum of the $\mathfrak{g}$-modules*

$$V^{\mathfrak{g}} := \bigcap_{x\in\mathfrak{g}} \ker\rho(x) \qquad \textit{and} \qquad V_{\text{eff}} := \sum_{x\in\mathfrak{g}} \rho(x)(V).$$

Proof. Note that $\rho(\mathfrak{g})(V^{\mathfrak{g}}) = \{0\}$ and $\rho(\mathfrak{g})(V_{\text{eff}}) \subseteq V_{\text{eff}}$, so that $V^{\mathfrak{g}}$ and V_{eff} are indeed $\mathfrak{g}$-invariant. We argue by induction on $\dim V$. The case $\dim V = \{0\}$ is trivial. Since the statement of the proposition is obvious for $\rho = 0$, we assume that $\rho \neq 0$.

Step 1: Let $\beta_\rho(x, y) = \operatorname{tr}\big(\rho(x)\rho(y)\big)$ denote the trace form on $\mathfrak{g}$ and $\mathfrak{a} := \operatorname{rad}(\beta_\rho)$ denote the radical of β_ρ, which is an ideal because β_ρ is invariant (Exercise 4.7). Let $\mathfrak{b} \trianglelefteq \mathfrak{g}$ be a complementary ideal, so that $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ is a direct sum of Lie algebras (Proposition 5.14). In view of Cartan's Solvability Criterion (Theorem 4.22), the Lie algebra $\rho(\mathfrak{a})$ is a solvable ideal of $\rho(\mathfrak{g})$ because the trace form vanishes on this Lie algebra. Since $\rho(\mathfrak{g})$ is semisimple, $\rho(\mathfrak{a}) \subseteq \operatorname{rad}(\rho(\mathfrak{g})) = \{0\}$, so that $\mathfrak{a} \subseteq \ker\rho$. Conversely, the ideal $\ker\rho$ is contained in $\operatorname{rad}(\beta_\rho)$, which leads to $\mathfrak{a} = \ker\rho$. It follows in particular that $\beta := \beta_\rho|_{\mathfrak{b}\times\mathfrak{b}}$ is nondegenerate on the semisimple Lie algebra $\mathfrak{b}$.

Step 2: Let

$$\Omega := \Omega(\beta, \rho|_{\mathfrak{b}}) := \sum_j \rho(x_j)\rho(x^j) \in \operatorname{End}(V)$$

be the associated Casimir element (Definition 5.19). Then Lemma 5.20 implies that

$$\Omega \in \operatorname{End}_{\mathfrak{b}}(V) := \{A \in \operatorname{End}(V) \colon (\forall x \in \mathfrak{b})\ A\rho(x) = \rho(x)A\}.$$

Since $\mathfrak{a} = \ker\rho$, this implies

$$\Omega \in \operatorname{End}_{\mathfrak{g}}(V) := \{A \in \operatorname{End}(V) \colon (\forall x \in \mathfrak{g})\ A\rho(x) = \rho(x)A\}.$$

Finally we note that

$$\operatorname{tr}\Omega = \sum_j \operatorname{tr}(\rho(x_j)\rho(x^i)) = \sum_j \beta(x_j, x^j) = \dim\mathfrak{b}.$$

Step 3: If V is the direct sum of two nonzero $\mathfrak{g}$-invariant subspaces, then $V^{\mathfrak{g}}$ and V_{eff} decompose accordingly, and we can use our induction hypothesis. Let $V = V^0(\Omega) \oplus V^+(\Omega)$ be the Fitting decomposition of V with respect to Ω (Lemma 5.21). Since Ω commutes with $\mathfrak{g}$, both summands are $\mathfrak{g}$-invariant, so that we may assume that one of these summands is trivial.

Since we assume that $\mathfrak{b} \cong \rho(\mathfrak{g}) \neq \{0\}$, we have $\operatorname{tr}\Omega > 0$, so that Ω is not nilpotent and thus $V^+(\Omega)$ is nonzero. Hence $V^0(\Omega) = \{0\}$ and, consequently, $V = V^+(\Omega)$. Then Ω is invertible, so that $V = V^+(\Omega) \subseteq V_{\text{eff}}$ and $V^{\mathfrak{g}} \subseteq V^0(\Omega) = \{0\}$. This completes the proof. □

For the following proposition recall Definitions 1.17 and 1.20.

Proposition 5.23. (Characterization of semisimple modules) *For a finite dimensional representation (ρ, V) of the Lie algebra $\mathfrak{g}$, the following are equivalent:*

(i) *Each $\mathfrak{g}$-invariant subspace of V possesses a $\mathfrak{g}$-invariant complement (each submodule has a module complement).*

(ii) *(ρ, V) is completely reducible (V is a semisimple $\mathfrak{g}$-module).*

(iii) *V is a sum of simple submodules.*

Proof. (i) ⇒ (ii): For dimensional reasons, each V contains a nonzero $\mathfrak{g}$-submodule V_1 of minimal dimension. Then there exists a module complement W, so that $V = V_1 \oplus W$. Then W also satisfies (i): If $W_1 \subseteq W$ is a submodule and $V' \subseteq V$ is a module complement for W_1 in V, i.e., $V = W_1 \oplus V'$, then $W = W_1 \oplus (V' \cap W)$. We can thereore argue by induction on $\dim V$ and apply the induction hypothesis to the representation of $\mathfrak{g}$ on W.

(ii) ⇒ (iii) is trivial.

(iii) ⇒ (ii): Let $V_1, \ldots, V_n$ be a maximal set of simple submodules whose sum $W := \sum_{i=1}^n V_i$ is direct. We claim that $W = V$, which implies (ii). If W is a proper subspace, then (iii) implies the existence of a minimal nonzero submodule U not contained in W. Then $W \cap U = \{0\}$ follows from the minimality of U, so that the sum $U + \sum_i V_i$ is direct, contradicting the maximality of the set $\{V_1, \ldots, V_n\}$. This proves $W = V$.

(ii) ⇒ (i): Let $V = \bigoplus_{i=1}^n V_i$ be a direct sum of simple submodules and $W \subseteq V$ a $\mathfrak{g}$-invariant subspace. Further, let $J \subseteq \{1, \ldots, n\}$ be maximal with

$$W \cap \Big(\sum_{i \in J} V_i\Big) = \{0\}.$$

Then $W' := \sum_{i \in J} V_i$ satisfies $W \cap W' = \{0\}$ and it remains to see that $W + W' = V$.

Pick $i \in I$. If $i \in J$, then $V_i \subseteq W' \subseteq W + W'$. If $i \notin J$, then the maximality of J implies that $(W' + V_i) \cap W \neq \{0\}$ and hence $(W + W') \cap V_i \neq \{0\}$. Hence the simplicity of V_i implies that $V_i \subseteq W + W'$, and this proves $V = W + W'$. □

Proposition 5.24. *If $\mathfrak{g}$ is a real Lie algebra and V a finite dimensional $\mathfrak{g}$-module, then the following are equivalent:*

(i) *V is semisimple.*

(ii) *$V_{\mathbb{C}}$ is a semisimple complex $\mathfrak{g}$-module.*

Proof. (i) ⇒ (ii): If V is semisimple, then V is a direct sum of simple submodules V_i, then $V_{\mathbb{C}}$ is the direct sum of the submodules $(V_i)_{\mathbb{C}}$. Hence it suffices to show that the complexification $W_{\mathbb{C}}$ of a simple real $\mathfrak{g}$-module W is semisimple. In fact, if $W_{\mathbb{C}}$ is not simple, then let $U \subseteq W_{\mathbb{C}}$ be a nonzero minimal complex submodule. This implies in particular that U is simple. Let $\sigma\colon W_{\mathbb{C}} \to W_{\mathbb{C}}$ be the complex conjugation defined by $\sigma(x + iy) = x - iy$ for $x, y \in W$. Then σ commutes with the action of $\mathfrak{g}$ on $W_{\mathbb{C}}$, and therefore $\sigma(U)$ also is a simple complex submodule. Now $U + \sigma(U)$ is a complex σ-invariant submodule of $W_{\mathbb{C}}$, hence of the form $X_{\mathbb{C}}$ for $X := W \cap (U + \sigma(U))$ (Exercise A.1(ii)). Then X is a nonzero $\mathfrak{g}$-submodule of W, so that the simplicity of W yields $X = W$ and thus $U + \sigma(U) = W_{\mathbb{C}}$. Now Proposition 5.23 shows that $W_{\mathbb{C}}$ is semisimple because it is the sum of two simple submodules.

(ii) ⇒ (i): Let $W \subseteq V$ be a submodule. We have to show that there exists a module complement U (Proposition 5.23). Since $V_{\mathbb{C}}$ is semisimple, there exists a module complement X of $W_{\mathbb{C}}$ in $V_{\mathbb{C}}$, i.e., a complex linear projection $p\colon V_{\mathbb{C}} \to W_{\mathbb{C}}$ commuting with $\mathfrak{g}$. Let $q_W\colon W_{\mathbb{C}} \to W, x + iy \mapsto x$ $(x, y \in W)$, be the projection onto the "real part". Then q_W is a real linear projection commuting with $\mathfrak{g}$. Hence

$$P := q_W \circ p|_V \colon V \to W$$

is a $\mathfrak{g}$-equivariant real linear map with $P|_W = \mathrm{id}_W$. Therefore $\ker P$ is a submodule of V complementing W. □

Corollary 5.25. *Let V be a finite dimensional real vector space and $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ an abelian subalgebra consisting of elements whose complex linear extension to $V_{\mathbb{C}}$ is diagonalizable. Then V is a semisimple $\mathfrak{g}$-module.*

Proof. In view of Proposition 5.24, it suffices to show that $V_{\mathbb{C}}$ is a semisimple complex module of $\mathfrak{g}$, resp., $\mathfrak{g}_{\mathbb{C}}$. On $V_{\mathbb{C}}$, each $x \in \mathfrak{g}$ is diagonalizable, and since $\mathfrak{g}$ is abelian, $\mathfrak{g}$ is simultaneously diagonalizable (Exercise 4.1(d)), so that $V_{\mathbb{C}}$ is a direct sum of one-dimensional submodules, hence semisimple. $\square$

Theorem 5.26. (Weyl's Theorem on Complete Reducibility) *Each finite dimensional representation of a semisimple Lie algebra is completely reducible.*

Proof. Let (ρ, V) be a finite dimensional representation of the semisimple Lie algebra $\mathfrak{g}$. In view of Proposition 5.23, it suffices to show that each $\mathfrak{g}$-invariant subspace $W \subseteq V$ possesses a $\mathfrak{g}$-invariant complement U.
Step 1: Let $W \subseteq V$ be a $\mathfrak{g}$-invariant subspace of codimension 1. Then the representation $(\overline{\rho}, V/W)$, defined by $\overline{\rho}(x)(v + W) := \rho(x)v + W$ is one-dimensional. Since $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ is perfect and $\mathfrak{gl}_1(\mathbb{K}) \cong \mathbb{K}$ is abelian, $\overline{\rho} = 0$, so that $\rho(\mathfrak{g})V \subseteq W$. In view of Proposition 5.22, $V = V^{\mathfrak{g}} \oplus V_{\text{eff}}$, and since V_{eff} is contained in W, there exists some $v_o \in V^{\mathfrak{g}} \setminus W$. Then $\mathbb{K}v_o$ is a $\mathfrak{g}$-invariant complement of W.
Step 2: Now let $W \subseteq V$ be an arbitrary $\mathfrak{g}$-invariant subspace. We define a representation of $\mathfrak{g}$ on $\mathrm{Hom}(V, W)$ by

$$\pi(x)\varphi := \rho(x)|_W \circ \varphi - \varphi \circ \rho(x)$$

(Show as an exercise that this is a representation). Then the subspace

$$U := \{\varphi \in \mathrm{Hom}(V, W) \colon \varphi|_W \in \mathbb{K}\,\mathrm{id}_W\}$$

is $\mathfrak{g}$-invariant because we have for $\varphi \in U$ the relation $(\pi(x)\varphi)(W) = \{0\}$: For $\varphi|_W = \lambda\,\mathrm{id}_W$ and $w \in W$ we have

$$(\pi(x)\varphi)(w) = \rho(x)\varphi(w) - \varphi(\rho(x)w) = \rho(x)(\lambda w) - \lambda\rho(x)w = 0.$$

Therefore

$$U_0 := \{\varphi \in U \colon \varphi(W) = \{0\}\}$$

is a $\mathfrak{g}$-invariant subspace of U of codimension 1. Step 1 now implies the existence of a $\mathfrak{g}$-invariant $\varphi_0 \in U \setminus U_0$. The $\mathfrak{g}$-invariance of φ_0 means that $\varphi_0 \in \mathrm{Hom}_{\mathfrak{g}}(V, W)$ and since $\varphi_0|_W \in \mathbb{K}^\times\,\mathrm{id}_W$ is invertible, $\ker\varphi_0$ is a $\mathfrak{g}$-invariant subspace complementing W. $\square$

Exercises for Section 5

Exercise 5.1. Show that the dimension of a simple Lie algebra is at least 3. Conclude that every semisimple Lie algebra of dimension ≤ 5 is simple.

Exercise 5.2. Let $\mathfrak{g}$ be a finite dimensional Lie algebra and $\varphi \in \mathrm{Aut}(\mathfrak{g})$. Show that

(i) $\kappa(\varphi(x), \varphi(y)) = \kappa(x, y)$ for $x, y \in \mathfrak{g}$..

(ii) If $\mathfrak{g}$ is complex and $\varphi \colon \mathfrak{g} \to \mathfrak{g}$ is an antilinear automorphism, then

$$\kappa(\varphi(x), \varphi(y)) = \overline{\kappa(x, y)} \quad \text{for} \quad x, y \in \mathfrak{g}.$$

Hint: Use/verify the following LA fact: If V is a finite dimensional complex vector space, $\psi \in \mathrm{End}(V)$ and $\varphi \colon V \to V$ antilinear and bijective, then

$$\mathrm{tr}(\varphi \circ \psi \circ \varphi^{-1}) = \overline{\mathrm{tr}(\psi)}.$$

Exercise 5.3. Let $\mathfrak{g}$ be a real Lie algebra and $\mathfrak{g}_{\mathbb{C}}$ its complexification. Show that the Cartan–Killing forms of $\mathfrak{g}$ and $\mathfrak{g}_{\mathbb{C}}$ are related by

$$\kappa_{\mathfrak{g}}(x,y) = \kappa_{\mathfrak{g}_{\mathbb{C}}}(x,y) \quad \text{for} \quad x,y \in \mathfrak{g}.$$

Hint: If $A\colon V \to V$ is a real linear endomorphism of the real vector space V and $A_{\mathbb{C}}\colon V_{\mathbb{C}} \to V_{\mathbb{C}}$ its complex linear extension, then $\mathrm{tr}(A_{\mathbb{C}}) = \mathrm{tr}(A)$.

Exercise 5.4. For a real Lie algebra $\mathfrak{g}$, we have:

(i) $\mathrm{rad}(\mathfrak{g}_{\mathbb{C}}) = \mathrm{rad}(\mathfrak{g})_{\mathbb{C}}$. Hint: Show that the radical of $\mathfrak{g}_{\mathbb{C}}$ is invariant under complex conjugation $\sigma\colon \mathfrak{g}_{\mathbb{C}} \to \mathfrak{g}_{\mathbb{C}}$.

(ii) $\mathrm{rad}(\kappa_{\mathfrak{g}})_{\mathbb{C}} = \mathrm{rad}(\kappa_{\mathfrak{g}_{\mathbb{C}}})$.

(iii) $\mathfrak{g}$ is semisimple if and only if $\mathfrak{g}_{\mathbb{C}}$ is semisimple.

Exercise 5.5. Verify the computations of the Cartan–Killing forms of $\mathfrak{sl}_2(\mathbb{K})$, $\mathfrak{so}_3(\mathbb{R})$ and of the oscillator algebra in Example 5.6.

Exercise 5.6. (i) Let $\alpha : \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of the Lie algebra $\mathfrak{g}$ on V and $\mathfrak{n} \trianglelefteq \mathfrak{g}$ be an ideal. Then the space

$$V_0(\mathfrak{n}) := \{v \in V \mid (\forall x \in \mathfrak{n})\ \alpha(x)v = 0\}$$

is $\mathfrak{g}$-invariant.

(ii) Let

$$\mathfrak{a}_0 = \{0\} \subseteq \mathfrak{a}_1 \subseteq \ldots \subseteq \mathfrak{a}_n = \mathfrak{g}$$

be a maximal chain of ideals of $\mathfrak{g}$ and $\mathfrak{n} \trianglelefteq \mathfrak{g}$ a nilpotent ideal. Then $[\mathfrak{n}, \mathfrak{a}_j] \subseteq \mathfrak{a}_{j-1}$ for $j > 0$.

Exercise 5.7. Let $\mathfrak{g}$ be a finite dimensional Lie algebra. Every nilpotent ideal $\mathfrak{n}$ of $\mathfrak{g}$ is orthogonal to $\mathfrak{g}$ with respect to the Cartan–Killing form.

Exercise 5.8. Show that $[\mathfrak{g}, \mathfrak{g}]^{\perp} = \mathrm{rad}(\mathfrak{g})$ for every finite dimensional Lie algebra $\mathfrak{g}$ over a field of characteristic zero. Here $\perp$ refers to the Cartan–Killing form $\kappa_{\mathfrak{g}}$.
Hint: Use Exercise 5.7 and Corollary 4.16 to show that $\kappa(\mathfrak{g}, [\mathrm{rad}(\mathfrak{g}), \mathfrak{g}]) = \{0\}$ and the Cartan Criterion for the solvability of $[\mathfrak{g}, \mathfrak{g}]^{\perp}$.

Exercise 5.9. Each one-dimensional representation (π, V) of a perfect Lie algebra is trivial.

Exercise 5.10. Let β be a nondegenerate invariant symmetric bilinear form on the Lie algebra $\mathfrak{g}$, $x_1, \ldots, x_k$ a basis for $\mathfrak{g}$ and $x^1, \ldots, x^k$ the dual basis with respect to β, i.e., $\beta(x_i, x^j) = \delta_{ij}$. For any Lie algebra homomorphism $\rho\colon \mathfrak{g} \to \mathcal{A}_L$, $\mathcal{A}$ an associative algebra, we define the Casimir element

$$\Omega(\beta, \rho) := \sum_{i=1}^{k} \rho(x_i)\rho(x^i).$$

Show that $\Omega(\beta, \rho)$ does not depend on the choice of the basis $x_1, \ldots, x_k$.

Exercise 5.11. Let V be a finite dimensional vector space and V^* its dual space. Show that:

(a) The map $\gamma\colon V \otimes V^* \to \operatorname{End}(V)$ specified by $\gamma(v \otimes \alpha)(w) := \alpha(w)v$, is a linear isomorphism.

(b) If $v_1, \ldots, v_n$ is a basis for V and $v_1^*, \ldots, v_n^*$ the dual basis for V^*, defined by $v_j^*(v_i) = \delta_{ij}$, then $\gamma\big(\sum_{i=1}^n v_i \otimes v_i^*\big) = \mathrm{id}_V$.

(c) If $\beta\colon V \times V \to \mathbb{K}$ is a nondegenerate symmetric bilinear form, then

$$\widetilde{\gamma}\colon V \otimes V \to \operatorname{End}(V), \quad \widetilde{\gamma}(v \otimes w)(x) := \beta(x, w)v$$

is a linear isomorphism. If $v_1, \ldots, v_n$ is a basis for V and $v^1, \ldots, v^n \in V$ with $\beta(v^i, \cdot) = v_i^*$, $i = 1, \ldots, n$, then $\widetilde{\gamma}(\sum_{i=1}^n v_i \otimes v^i) = \mathrm{id}_V$.

Exercise 5.12. (i) Let $\mathfrak{g}$ be a simple Lie algebra over $\mathbb{C}$. Show that each invariant bilinear form κ on $\mathfrak{g}$ is a scalar multiple of the Cartan–Killing form $\kappa_{\mathfrak{g}}$.
Hint: Show first that every bilinear form $\beta\colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ can be written as $\beta(x, y) = \kappa_{\mathfrak{g}}(Ax, y)$ for some $A \in \operatorname{End}(\mathfrak{g})$.

(ii) Show that the result (i) does not hold for all simple real Lie algebras.
Hint: Consider $\mathfrak{sl}_2(\mathbb{C})$ (cf. Exampl 5.16).

Exercise 5.13. Show that semisimplicity is an extension property: If $\mathfrak{g}$ is a finite dimensional Lie algebra and $\mathfrak{n} \trianglelefteq \mathfrak{g}$ an ideal, then $\mathfrak{g}$ is semisimple if and only if $\mathfrak{n}$ and $\mathfrak{g}/\mathfrak{n}$ are semisimple.

Exercise 5.14. Let $\mathfrak{g}$ be a real Lie algebra and $I \in \operatorname{End}(\mathfrak{g})$ be a complex structure. Show that the complex vector space (V, I), endowed with the scalar multiplication $(\alpha + i\beta)z := \alpha z + \beta I z$ is a complex Lie algebra if and only if

$$[I, \operatorname{ad} x] = 0 \quad \text{for every} \quad x \in \mathfrak{g}.$$

Exercise 5.15. Let V and W be module of the Lie algebra $\mathfrak{g}$. Show that

$$(X\varphi)(v) := X\varphi(v) - \varphi(Xv), \quad v \in V, \varphi \in \operatorname{Hom}(V, W), X \in \mathfrak{g}$$

defines the structure of a $\mathfrak{g}$-module on $\operatorname{Hom}(V, W)$. Note that the submodule $\operatorname{Hom}(V, W)^{\mathfrak{g}}$ of invariant elements are precisely the module morphisms.

6 Levi's Splitting Theorem

In the preceding sections we dealt in particular with solvable and semisimple Lie algebras separately. Now we shall address the question how a finite dimensional Lie algebra $\mathfrak{g}$ decomposes into its maximal solvable ideal $\operatorname{rad}(\mathfrak{g})$ and the semisimple quotient $\mathfrak{g}/\operatorname{rad}(\mathfrak{g})$. Levi's Theorem is fundamental for the structure theory of finite dimensional Lie algebras. It asserts the existence of a semisimple subalgebra $\mathfrak{s}$ of $\mathfrak{g}$ complementing the radical $\operatorname{rad}(\mathfrak{g})$, also called a Levi complement. As a consequence, $\mathfrak{g} \cong \operatorname{rad}(\mathfrak{g}) \rtimes \mathfrak{s}$ is a semidirect sum.

Throughout, $\mathbb{K}$ will be a field of characteristic zero.

Lemma 6.1. *The quotient Lie algebra* $\mathfrak{g}/\operatorname{rad}(\mathfrak{g})$ *is semisimple.*

Proof. Let $q\colon \mathfrak{g} \to \mathfrak{g}/\operatorname{rad}(\mathfrak{g})$ be the quotient homomorphism and $\mathfrak{a} \trianglelefteq \mathfrak{g}/\operatorname{rad}(\mathfrak{g})$ a solvable ideal. Then $\mathfrak{b} := q^{-1}(\mathfrak{a}) \trianglelefteq \mathfrak{g}$ is an ideal containing $\operatorname{rad}(\mathfrak{g})$, for which $\mathfrak{a} \cong \mathfrak{b}/\operatorname{rad}(\mathfrak{g})$ is solvable. Since solvability is an extension property, $\mathfrak{b}$ is solvable, hence $\mathfrak{b} \subseteq \operatorname{rad}(\mathfrak{g})$, and thus $\mathfrak{a} = \{0\}$. This proves that $\operatorname{rad}(\mathfrak{g}/\operatorname{rad}(\mathfrak{g})) = \{0\}$, i.e., $\mathfrak{g}/\operatorname{rad}(\mathfrak{g})$ is semisimple. $\square$

Proposition 6.2. *If $\alpha\colon \mathfrak{g} \to \mathfrak{h}$ is a surjective homomorphism of Lie algebras, then $\alpha(\operatorname{rad}\mathfrak{g}) = \operatorname{rad}\mathfrak{h}$.*

Proof. Let $\mathfrak{r} := \operatorname{rad}\mathfrak{g}$. First we note that $\alpha(\mathfrak{r})$ is a solvable ideal of $\mathfrak{h}$, hence contained in $\operatorname{rad}(\mathfrak{h})$. Here we use that images of ideals under surjective homomorphisms are ideals: $[\mathfrak{h}, \alpha(\mathfrak{r})] = [\alpha(\mathfrak{g}), \alpha(\mathfrak{r})] = \alpha([\mathfrak{g}, \mathfrak{r}]) \subseteq \alpha(\mathfrak{r})$.

Let $\pi\colon \mathfrak{h} \to \mathfrak{h}/\alpha(\mathfrak{r})$ be the quotient homomorphism. The homomorphism $\alpha\colon \mathfrak{g} \to \mathfrak{h}$ induces a surjective homomorphism $\widetilde{\alpha}\colon \mathfrak{g}/\mathfrak{r} \to \mathfrak{h}/\alpha(\mathfrak{r})$. Since $\mathfrak{g}/\mathfrak{r}$ is semisimple (Lemma 6.1), the homomorphic image $\mathfrak{h}/\alpha(\mathfrak{r})$ is also semisimple (Proposition 5.10). Consequently $\pi(\operatorname{rad}\mathfrak{h}) \subseteq \operatorname{rad}\big(\mathfrak{h}/\alpha(\mathfrak{r})\big) = \{0\}$, i.e., $\operatorname{rad}\mathfrak{h} \subseteq \alpha(\mathfrak{r})$. We thus obtain $\operatorname{rad}\mathfrak{h} = \alpha(\mathfrak{r})$. $\square$

Definition 6.3. An ideal $\mathfrak{a} \trianglelefteq \mathfrak{g}$ is called *characteristic* if it is invariant under all derivations of $\mathfrak{g}$.

Lemma 6.4. *For the radical of the Lie algebra $\mathfrak{g}$, the following assertions hold:*

(i) $\operatorname{rad}(\mathfrak{g})$ *is a characteristic ideal.*

(ii) *If $\mathfrak{a} \subseteq \mathfrak{g}$ is an ideal, then* $\operatorname{rad}(\mathfrak{a}) = \operatorname{rad}(\mathfrak{g}) \cap \mathfrak{a}$.

Proof. (i) First we note that $[\mathfrak{g}, \mathfrak{g}]$ is a characteristic ideal of $\mathfrak{g}$ because for each derivation $D \in \operatorname{der}\mathfrak{g}$ and $x, y \in \mathfrak{g}$ we have $D([x, y]) = [Dx, y] + [x, Dy] \in [\mathfrak{g}, \mathfrak{g}]$. Next we note that the Cartan–Killing form is invariant under $\operatorname{der}(\mathfrak{g})$ (cf. Exercise 4.7) :

$$\begin{aligned}\kappa_{\mathfrak{g}}(Dx, y) &= \operatorname{tr}\big(\operatorname{ad}(Dx)\operatorname{ad}y\big) = \operatorname{tr}\big([D, \operatorname{ad}x]\operatorname{ad}y\big)\\ &= -\operatorname{tr}\big(\operatorname{ad}x[D, \operatorname{ad}y]\big) = -\operatorname{tr}\big(\operatorname{ad}x\operatorname{ad}(Dy)\big) = -\kappa_{\mathfrak{g}}(x, Dy).\end{aligned}$$

Therefore $\operatorname{rad}(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]^{\perp, \kappa_{\mathfrak{g}}}$ (Exercise 5.8) is also invariant under $\operatorname{der}(\mathfrak{g})$.

(ii) Clearly, $\operatorname{rad}(\mathfrak{g}) \cap \mathfrak{a}$ is a solvable ideal of $\mathfrak{a}$, hence contained in $\operatorname{rad}(\mathfrak{a})$. Since $\operatorname{rad}(\mathfrak{a})$ is a characteristic ideal of $\mathfrak{a}$, it is invariant under the adjoint representation of $\mathfrak{g}$ on $\mathfrak{a}$, hence a solvable ideal of $\mathfrak{g}$. This proves that $\operatorname{rad}(\mathfrak{a}) \subseteq \operatorname{rad}(\mathfrak{g})$. $\square$

We will need the following technical lemma.

Lemma 6.5. *Let (ρ, V) be a representation of $\mathfrak{g}$ and $\mathfrak{n} \trianglelefteq \mathfrak{g}$ an ideal. For $v \in V$, let*

$$\mathfrak{z}_{\mathfrak{g}}(v) := \{x \in \mathfrak{g} \colon \rho(x)v = 0\}$$

be the stabilizer of v. If $v \in V$ satisfies

$$\rho(\mathfrak{g})v = \rho(\mathfrak{n})v \quad \text{and} \quad \mathfrak{z}_{\mathfrak{g}}(v) \cap \mathfrak{n} = \mathfrak{z}_{\mathfrak{n}}(v) = \{0\},$$

then $\mathfrak{g} \cong \mathfrak{n} \rtimes \mathfrak{z}_{\mathfrak{g}}(v)$.

Proof. The linear map $\varphi\colon \mathfrak{g} \to V, x \mapsto \rho(x)v$ satisfies $\varphi(\mathfrak{g}) = \varphi(\mathfrak{n})$, hence $\mathfrak{g} = \mathfrak{n} + \ker\varphi = \mathfrak{n} + \mathfrak{z}_{\mathfrak{g}}(v)$. Since $\mathfrak{z}_{\mathfrak{g}}(v)$ is a subalgebra, the assertion follows. $\square$

Theorem 6.6. (Levi's Splitting Theorem) *If $\alpha\colon \mathfrak{g} \to \mathfrak{s}$ is a surjective homomorphism of Lie algebras and $\mathfrak{s}$ is semisimple, then there exists a homomorphism $\beta\colon \mathfrak{s} \to \mathfrak{g}$ with $\alpha \circ \beta = \operatorname{id}_{\mathfrak{s}}$.*

$$\begin{tikzcd} & \mathfrak{g} \arrow[d, "\forall\alpha"] \\ \mathfrak{s} \arrow[ur, "\exists\beta"] \arrow[r, "\operatorname{id}_{\mathfrak{s}}"'] & \mathfrak{s} \end{tikzcd}$$

Proof. Let $\mathfrak{n} := \ker \alpha$. We have to show the existence of a subalgebra $\widetilde{\mathfrak{s}}$ of $\mathfrak{g}$ with $\mathfrak{g} \cong \mathfrak{n} \rtimes \widetilde{\mathfrak{s}}$. Then $\alpha|_{\widetilde{\mathfrak{s}}} : \widetilde{\mathfrak{s}} \to \mathfrak{s}$ is an isomorphism and we may put $\beta := (\alpha|_{\widetilde{\mathfrak{s}}})^{-1}$. We argue by induction on the dimension of $\mathfrak{n}$. For $\mathfrak{n} = \{0\}$, there is nothing to show. So we assume that $\mathfrak{n} \neq \{0\}$.

Case 1: The ideal $\mathfrak{n} \trianglelefteq \mathfrak{g}$ is not minimal, i.e., there exists a nonzero ideal $\mathfrak{n}_1$ of $\mathfrak{g}$ which is a proper subspace of $\mathfrak{n}$. Now α factors through a surjective homomorphism $\alpha_1 : \mathfrak{g}/\mathfrak{n}_1 \to \mathfrak{s}$ with

$$\dim(\ker \alpha_1) = \dim \mathfrak{n} - \dim \mathfrak{n}_1 < \dim \mathfrak{n}.$$

Therefore our induction hypothesis implies the existence of a homomorphism $\beta_1 : \mathfrak{s} \to \mathfrak{g}/\mathfrak{n}_1$ with $\alpha_1 \circ \beta_1 = \mathrm{id}_{\mathfrak{s}}$. Let $q : \mathfrak{g} \to \mathfrak{g}/\mathfrak{n}_1$ be the quotient map and $\mathfrak{b} := q^{-1}\big(\beta_1(\mathfrak{s})\big)$. Then $\mathfrak{b}$ is a subalgebra of $\mathfrak{g}$ and the homomorphism

$$\alpha_2 := q|_{\mathfrak{b}} : \mathfrak{b} \to \beta_1(\mathfrak{s}) \cong \mathfrak{s}, \quad x \mapsto x + \mathfrak{n}_1$$

is surjective. In view of $\dim(\ker \alpha_2) = \dim \mathfrak{n}_1 < \dim \mathfrak{n}$, the induction hypothesis implies the existence of a homomorphism $\beta_2 : \beta_1(\mathfrak{s}) \to \mathfrak{b}$ with $\alpha_2 \circ \beta_2 = \mathrm{id}_{\beta_1(\mathfrak{s})}$. Now $\beta := \beta_2 \circ \beta_1 : \mathfrak{s} \to \mathfrak{g}$ is a homomorphism satisfying

$$\alpha \circ \beta = \alpha_1 \circ q \circ \beta_2 \circ \beta_1 = \alpha_1 \circ \alpha_2 \circ \beta_2 \circ \beta_1 = \alpha_1 \circ \beta_1 = \mathrm{id}_{\mathfrak{s}}.$$

Case 2: The ideal $\mathfrak{n}$ is minimal. Since $\mathfrak{s}$ is semisimple, the radical $\mathfrak{r} := \mathrm{rad}(\mathfrak{g})$ of $\mathfrak{g}$ is contained in $\mathfrak{n}$ (Proposition 6.2). If $\mathfrak{r} = \{0\}$, then $\mathfrak{g}$ is semisimple, and the assertion follows from Proposition 5.14 because $\mathfrak{g}$ contains an ideal $\widetilde{\mathfrak{s}}$ complementing $\mathfrak{n}$. So let us assume that $\mathfrak{r} \neq \{0\}$. Then the minimality of $\mathfrak{n}$ shows that $\mathfrak{n} = \mathfrak{r}$ is abelian.

The representation $\rho : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{n}), x \mapsto \mathrm{ad}\, x|_{\mathfrak{n}}$ satisfies $\mathfrak{n} \subseteq \ker \rho$ ($\mathfrak{n}$ is abelian), hence factors through a representation $\overline{\rho}$ of $\mathfrak{s}$ on $\mathfrak{n}$, determined by $\overline{\rho} \circ \alpha = \rho$. Since $\mathfrak{n}$ is a minimal ideal of $\mathfrak{g}$, we thus obtain on $\mathfrak{n}$ an irreducible representation of $\mathfrak{s}$. If $\overline{\rho} = 0$, then $\mathfrak{n}$ is central in $\mathfrak{g}$, and the adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathrm{der}(\mathfrak{g})$ factors through a representation of $\mathfrak{s}$ on $\mathfrak{g}$. According to Weyl's Theorem, there exists an ideal of $\mathfrak{g}$ complementing $\mathfrak{n}$ (Proposition 5.23) and the proof is complete. We may therefore assume that $\overline{\rho}$ is nonzero.

We are now at the point where we can use Lemma 6.5. On $V := \mathrm{End}(\mathfrak{g})$, we consider the representation

$$\pi(x)\varphi := \mathrm{ad}\, x \circ \varphi - \varphi \circ \mathrm{ad}\, x = [\mathrm{ad}\, x, \varphi]$$

(cf. Exercise 5.15). We consider the following three subspaces of $V = \mathrm{End}(\mathfrak{g})$:

$$\begin{aligned} P := \mathrm{ad}\, \mathfrak{n} \subseteq Q &:= \{\varphi \in V : \varphi(\mathfrak{g}) \subseteq \mathfrak{n}, \varphi(\mathfrak{n}) = \{0\}\} \\ \subseteq R &:= \{\varphi \in V : \varphi(\mathfrak{g}) \subseteq \mathfrak{n}, \varphi|_{\mathfrak{n}} \in \mathbb{K}\, \mathrm{id}_{\mathfrak{n}}\}. \end{aligned}$$

Since $Q \subseteq R$ is the kernel of the linear map $\chi : R \to \mathbb{K}$, defined by $\varphi|_{\mathfrak{n}} = \chi(\varphi)\, \mathrm{id}_{\mathfrak{n}}$, we see that $\dim(R/Q) = 1$.

We claim that P, Q and R are $\mathfrak{g}$-invariant. To this end, let $y \in \mathfrak{g}$. For $x \in \mathfrak{n}$ we have $[\mathrm{ad}\, y, \mathrm{ad}\, x] = \mathrm{ad}[y, x] \in P$, so that P is $\mathfrak{g}$-invariant. To see that R and Q are $\mathfrak{g}$-invariant, we show that $\pi(\mathfrak{g})R \subseteq Q$. So let $x \in \mathfrak{g}$, $\varphi \in R$ and $\varphi|_{\mathfrak{n}} = \lambda\, \mathrm{id}_{\mathfrak{n}}$. For $n \in \mathfrak{n}$ we then have

$$(\pi(x)\varphi)(n) = [x, \varphi(n)] - \varphi([x, n]) = [x, \lambda n] - \lambda[x, n] = 0,$$

hence $\pi(x)\varphi \in Q$. For $y \in \mathfrak{n}$ we get

$$[\mathrm{ad}\, y, \varphi] = \mathrm{ad}\, y \circ \varphi - \varphi \circ \mathrm{ad}\, y = -\lambda\, \mathrm{ad}\, y \in P. \tag{11}$$

This proves that $\pi(\mathfrak{n})R \subseteq P$. The ideal $\mathfrak{n}$ acts trivially on the quotient space R/P, which therefore inherits a representation of $\mathfrak{s} \cong \mathfrak{g}/\mathfrak{n}$.

According to Weyl's Theorem 5.26, there exists an $\mathfrak{s}$-invariant subspace of R/P complementing Q/P. This complement is one-dimensional, hence generated by the image $\overline{\varphi}$ of one element $\varphi \in R \setminus Q$, of which we may assume that $\varphi|_{\mathfrak{n}} = \mathrm{id}_{\mathfrak{n}}$. As the one-dimensional representation of $\mathfrak{s}$ on $\mathbb{R}\overline{\varphi}$ is trivial because $\mathfrak{s}$ is perfect (Exercise 5.9), we see that $\pi(\mathfrak{g})\varphi \subseteq P$. Next we verify the assumptions of Lemma 6.5.

For $x \in \mathfrak{n}$, we have already seen in (11) that $\pi(x)\varphi = [\operatorname{ad} x, \varphi] = -\operatorname{ad} x$. If $\pi(x)\varphi = 0$, then $\operatorname{ad} x = 0$, i.e., $x \in \mathfrak{z}(\mathfrak{g})$. Since $\mathfrak{n}$ is a minimal ideal of $\mathfrak{g}$ which is not central, we derive that $x = 0$. This leads to $\mathfrak{z}_{\mathfrak{n}}(\varphi) = \{0\}$ and $\pi(\mathfrak{n})\varphi = \operatorname{ad} \mathfrak{n} = P \supseteq \pi(\mathfrak{g})\varphi$. Finally, we apply Lemma 6.5 to complete the proof. $\square$

Definition 6.7. If $\mathfrak{g}$ is a finite dimensional Lie algebra, then we call a subalgebra $\mathfrak{s} \leq \mathfrak{g}$ complementing the solvable radical $\operatorname{rad}(\mathfrak{g})$ a *Levi complement*. Note that $\mathfrak{g} \cong \operatorname{rad}(\mathfrak{g}) \rtimes \mathfrak{s}$ holds for any Levi complement.

Corollary 6.8. *Each finite dimensional Lie algebra $\mathfrak{g}$ contains a semisimple Levi complement.*

Proof. Let $\mathfrak{s} := \mathfrak{g}/\operatorname{rad}(\mathfrak{g})$ and $\alpha \colon \mathfrak{g} \to \mathfrak{s}$ be the quotient map. According to Lemma 6.1, $\mathfrak{s}$ is semisimple. Hence Theorem 6.6 provides a homomorphism $\beta \colon \mathfrak{s} \to \mathfrak{g}$ with $\alpha \circ \beta = \mathrm{id}_{\mathfrak{s}}$. Then β is injective, so that $\beta(\mathfrak{s}) \cap \operatorname{rad}(\mathfrak{g}) = \{0\}$ as well as $\beta(\mathfrak{s}) + \operatorname{rad}(\mathfrak{g}) = \mathfrak{g}$. Thus $\beta(\mathfrak{s})$ is a semisimple Levi complement. $\square$

Corollary 6.9. *If $\mathfrak{s}$ is a Levi complement in $\mathfrak{g}$, then*

$$[\mathfrak{g}, \mathfrak{g}] = [\mathfrak{g}, \operatorname{rad}(\mathfrak{g})] \rtimes \mathfrak{s}.$$

If $\operatorname{rad}(\mathfrak{g}) = \mathfrak{z}(\mathfrak{g})$*, then* $[\mathfrak{g}, \mathfrak{g}]$ *is a Levi complement.*

Proof. The second assertion immediately follows from the first and the fact that $\operatorname{rad}(\mathfrak{g}) = \mathfrak{z}(\mathfrak{g})$ is equivalent to $[\mathfrak{g}, \operatorname{rad}(\mathfrak{g})] = \{0\}$.

For the first assertion, we note that $[\mathfrak{s}, \mathfrak{s}] = \mathfrak{s}$ leads to

$$[\mathfrak{g}, \mathfrak{g}] = [\mathfrak{g}, \operatorname{rad}(\mathfrak{g})] + [\mathfrak{g}, \mathfrak{s}] = [\mathfrak{g}, \operatorname{rad}(\mathfrak{g})] + [\operatorname{rad}(\mathfrak{g}), \mathfrak{s}] + [\mathfrak{s}, \mathfrak{s}] = [\mathfrak{g}, \operatorname{rad}(\mathfrak{g})] + \mathfrak{s}. \qquad \square$$

Corollary 6.10. (Lifting homomorphisms) *If $q \colon \widehat{\mathfrak{g}} \to \mathfrak{g}$ is a surjective homomorphism of finite dimensional Lie algebras, $\mathfrak{s}$ is semisimple and $\alpha \colon \mathfrak{s} \to \mathfrak{g}$ is a homomorphism, then there exists a homomorphism $\widehat{\alpha} \colon \mathfrak{s} \to \widehat{\mathfrak{g}}$ with $q \circ \widehat{\alpha} = \alpha$.*

Proof. Apply Levi's Theorem 6.6 to the surjective homomorphism $q \colon q^{-1}(\alpha(\mathfrak{s})) \to \alpha(\mathfrak{s})$ and note that the homomorphic image $\alpha(\mathfrak{s})$ of $\mathfrak{s}$ is semisimple. $\square$

Remark 6.11. If $\mathfrak{g}$ is a solvable Lie algebra, then $\mathfrak{g}$ is isomorphic to a nested semidirect sum

$$\big(\ldots\big((\mathfrak{g}_1 \rtimes_{\alpha_1} \mathfrak{g}_2) \rtimes_{\alpha_2} \mathfrak{g}_3\big) \ldots \rtimes_{\alpha_{n-1}} \mathfrak{g}_n\big)$$

of one-dimensional Lie algebras (cf. Exercise 4.12).

Composing this with Levi's Theorem and using Proposition 5.14, we obtain a similar factorization for arbitrary finite dimensional Lie algebras $\mathfrak{g}$, the only difference is that the $\mathfrak{g}_j$ are either one-dimensional or simple. In fact, we can start with a maximal chain

$$\mathfrak{a}_0 = \{0\} \subseteq \mathfrak{a}_1 \subseteq \ldots \subseteq \mathfrak{a}_k = \mathfrak{g}$$

of subalgebras of $\mathfrak{g}$ for which $\mathfrak{a}_{j-1}$ is an ideal in $\mathfrak{a}_j$. Such a series is called a *Jordan–Hölderseries* of $\mathfrak{g}$. Then the quotient $\mathfrak{g}_j := \mathfrak{a}_j/\mathfrak{a}_{j-1}$ is either one-dimensional or simple so that Levi's Theorem implies that

$$\mathfrak{a}_j \cong \mathfrak{a}_{j-1} \rtimes \mathfrak{g}_j.$$

Exercises for Section 6

Exercise 6.1. Show that, for every Jordan–Hölder series

$$\mathfrak{a}_0 = \{0\} \subseteq \mathfrak{a}_1 \subseteq \ldots \subseteq \mathfrak{a}_k = \mathfrak{g}$$

of subalgebras of $\mathfrak{g}$ (i.e., $\mathfrak{a}_{j-1}$ is an ideal in $\mathfrak{a}_j$ and the quotient $\mathfrak{a}_j/\mathfrak{a}_{j-1}$ is either one-dimensional or simple), the set of quotients $\{\mathfrak{a}_j/\mathfrak{a}_{j-1} \colon j = 1, \ldots, k\}$ does not depend on the Jordan–Hölder series (cf. Remark 6.11).

7 Reductive Lie Algebras

We conclude this chapter with a brief discussion of reductive Lie algebras. This class of Lie algebras is only slightly larger than the class of semisimple Lie algebras and it contains the abelian Lie algebras. Reductive Lie algebras often occur as stabilizer subalgebras inside semisimple Lie algebras. Thus they appear frequently in proofs by induction on the dimension.

Definition 7.1. We call a finite dimensional Lie algebra $\mathfrak{g}$ *reductive* if $\mathfrak{g}$ is a semisimple module with respect to the adjoint representation, i.e., for each ideal $\mathfrak{a} \trianglelefteq \mathfrak{g}$, there exists an ideal $\mathfrak{b} \trianglelefteq \mathfrak{g}$ with $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$.

Remark 7.2. By Lemma 5.9(iii), every semisimple Lie algebra is reductive. It is also clear that every abelian Lie algebra is reductive.

Lemma 7.3. *For a reductive Lie algebra $\mathfrak{g}$, the following assertions hold:*

(i) *If $\mathfrak{n} \trianglelefteq \mathfrak{g}$ is an ideal, then $\mathfrak{n}$ and $\mathfrak{g}/\mathfrak{n}$ are reductive.*

(ii) $\mathfrak{g} = \mathfrak{z}(\mathfrak{g}) \oplus [\mathfrak{g}, \mathfrak{g}]$ *and* $[\mathfrak{g}, \mathfrak{g}]$ *is semisimple.*

(iii) $\mathfrak{g}$ *is semisimple if and only if* $\mathfrak{z}(\mathfrak{g}) = \{0\}$.

Proof. (i) Since $\mathfrak{g}$ is reductive, there exists an ideal $\mathfrak{b} \trianglelefteq \mathfrak{g}$ with $\mathfrak{g} = \mathfrak{n} \oplus \mathfrak{b}$. Then $[\mathfrak{b}, \mathfrak{n}] = \{0\}$, so that $\mathfrak{g}$ is a direct sum of Lie algebras. As submodules of the semisimple $\mathfrak{g}$-module $\mathfrak{g}$, the ideals $\mathfrak{n}$ and $\mathfrak{b}$ are semisimple $\mathfrak{g}$-modules, and since the complementary ideals do not act on each other, it follows that $\mathfrak{n}$ and $\mathfrak{b} \cong \mathfrak{g}/\mathfrak{n}$ are reductive Lie algebras.

(ii) Let $\mathfrak{a} \subseteq \mathfrak{g}$ be an ideal complement of $[\mathfrak{g}, \mathfrak{g}]$. Then $\mathfrak{g} = \mathfrak{a} \oplus [\mathfrak{g}, \mathfrak{g}]$, and $[\mathfrak{g}, \mathfrak{a}] \subseteq \mathfrak{a} \cap [\mathfrak{g}, \mathfrak{g}] = \{0\}$ implies that $\mathfrak{a}$ is central. Further, (i) implies that $[\mathfrak{g}, \mathfrak{g}]$ is reductive. To see that $\mathfrak{z}(\mathfrak{g})$ is not larger than $\mathfrak{a}$, we choose an ideal $\mathfrak{b}$ of $[\mathfrak{g}, \mathfrak{g}]$ complementing $\mathfrak{z}(\mathfrak{g}) \cap [\mathfrak{g}, \mathfrak{g}]$ in $[\mathfrak{g}, \mathfrak{g}]$. Then $[\mathfrak{g}, \mathfrak{g}] = [\mathfrak{b}, \mathfrak{b}] \subseteq \mathfrak{b}$ yields $\mathfrak{z}(\mathfrak{g}) \cap [\mathfrak{g}, \mathfrak{g}] = \{0\}$, and hence $\mathfrak{z}(\mathfrak{g}) = \mathfrak{a}$.

Since $[\mathfrak{g}, \mathfrak{g}]$ is reductive, it is a direct sum of simple modules $\mathfrak{g}_1, \ldots, \mathfrak{g}_m$ for the adjoint representation. The preceding argument implies that none of these ideals is abelian, hence they are simple Lie algebras and thus $[\mathfrak{g}, \mathfrak{g}]$ is semisimple.

(iii) If $\mathfrak{z}(\mathfrak{g}) = \{0\}$, then (ii) implies that $\mathfrak{g}$ is semisimple. If, conversely, $\mathfrak{g}$ is semisimple, then $\mathfrak{z}(\mathfrak{g}) \subseteq \mathrm{rad}(\mathfrak{g}) = \{0\}$. □

Proposition 7.4. *For a finite dimensional Lie algebra $\mathfrak{g}$, the following are equivalent:*

(i) $\mathfrak{g}$ *is reductive.*

(ii) $[\mathfrak{g}, \mathfrak{g}]$ *is semisimple.*

(iii) $\mathrm{rad}(\mathfrak{g})$ *is central in* $\mathfrak{g}$, *i.e.,* $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})] = \{0\}$.

Proof. (i) $\Rightarrow$ (ii) follows from Lemma 7.3.
(ii) $\Rightarrow$ (iii): Let $\mathfrak{r} := \mathrm{rad}(\mathfrak{g})$ and $\mathfrak{g} = \mathfrak{r} \rtimes \mathfrak{s}$ be a Levi decomposition. Then $[\mathfrak{g}, \mathfrak{g}] = [\mathfrak{g}, \mathfrak{r}] \rtimes \mathfrak{s}$ by Corollary 6.9, so that the semisimplicity of $[\mathfrak{g}, \mathfrak{g}]$ implies that $[\mathfrak{g}, \mathfrak{r}] = \{0\}$.
(iii) $\Rightarrow$ (i): If $\mathfrak{r}$ is central in $\mathfrak{g}$, then any Levi decomposition $\mathfrak{g} = \mathfrak{r} \rtimes \mathfrak{s}$ is a direct sum $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{s}$, where $\mathfrak{r}$ is a central ideal. Since $\mathfrak{z}(\mathfrak{s}) = \{0\}$, we immediately get $\mathfrak{z}(\mathfrak{g}) = \mathfrak{r}$, so that $\mathfrak{g} = \mathfrak{z}(\mathfrak{g}) \oplus \mathfrak{s}$. Thus $\mathfrak{g} \cong \mathbb{K}^n \oplus \mathfrak{s}$ is a direct sum of simple submodules with respect to the adjoint representation, hence a semisimple $\mathfrak{g}$-module (Proposition 5.23), and this means that $\mathfrak{g}$ is reductive. $\square$

Proposition 7.5. *The ideal* $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]$ *coincides with the intersection of the kernels of finite dimensional irreducible representations.*

Proof. If $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a finite dimensional irreducible representation and $\mathfrak{n} := [\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]$, then each subspace $V_k := \rho(\mathfrak{n})^k(V)$ is $\mathfrak{g}$-invariant because

$$\begin{aligned}&\rho(x)\rho(x_1)\cdots\rho(x_k)v\\&= \Big(\sum_{j=1}^k \rho(x_1)\cdots \underbrace{[\rho(x), \rho(x_j)]}_{\in \rho(\mathfrak{n})} \cdots \rho(x_k)v\Big) + \rho(x_1)\cdots\rho(x_k)\rho(x)v\end{aligned}$$

for each $v \in V$. Since V is a nilpotent $\mathfrak{n}$-module by Proposition 4.15, there exists a minimal m with $V_m = \{0\}$. Then $V_{m-1} \neq \{0\}$, so that the irreducibility of the representation implies that $V_{m-1} = V$. Hence $\rho(\mathfrak{n})V = \rho(\mathfrak{n})V_{m-1} \subseteq V_m = \{0\}$, i.e., $\mathfrak{n} \subseteq \ker \rho$.

Next we consider the quotient $\mathfrak{q} := \mathfrak{g}/[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]$ and the quotient homomorphism $q\colon \mathfrak{g} \to \mathfrak{q}$. In view of Proposition 6.2, $q(\mathrm{rad}(\mathfrak{g})) = \mathrm{rad}(\mathfrak{q})$, so that

$$[\mathfrak{q}, \mathrm{rad}(\mathfrak{q})] = [q(\mathfrak{g}), q(\mathrm{rad}(\mathfrak{g}))] = q([\mathfrak{g}, \mathrm{rad}(\mathfrak{g})]) = \{0\}$$

implies that $\mathrm{rad}(\mathfrak{q})$ is central, so that $\mathfrak{q}$ is reductive by Proposition 7.4.

It remains to observe that, for each non-zero $x \in \mathfrak{q}$, there exists an irreducible finite dimensional representations $\rho\colon \mathfrak{q} \to \mathfrak{gl}(V)$ with $\rho(x) \neq 0$. Since $\rho \circ q\colon \mathfrak{g} \to \mathfrak{gl}(V)$ then is an irreducible representation of $\mathfrak{g}$, this implies the assertion.

We know that $\mathfrak{q} = \mathfrak{z}(\mathfrak{q}) \oplus [\mathfrak{q}, \mathfrak{q}]$ and that $[\mathfrak{q}, \mathfrak{q}]$ is semisimple, hence a sum of simple ideals $\mathfrak{q}_1, \ldots, \mathfrak{q}_k$. If $x \notin \mathfrak{z}(\mathfrak{q})$, then x projects to a non-zero element of some $\mathfrak{q}_j$, and then $\rho(y) := \mathrm{ad}\, y|_{\mathfrak{q}_j}$ defines an irreducible representation with $\rho(x) \neq 0$. If $x \in \mathfrak{z}(\mathfrak{g})$, then there exists a linear functional $\lambda\colon \mathfrak{q} \to \mathbb{K}$ with $[\mathfrak{q}, \mathfrak{q}] \subseteq \ker \lambda$ and $\lambda(x) \neq 0$. Now $\rho(y) := \lambda(y)\mathbf{1}$ defines a one-dimensional representation of $\mathfrak{q}$ with $\rho(x) \neq 0$. $\square$

Example 7.6. Let V be a finite dimensional vector space.

(a) Let $\mathfrak{g} = \mathfrak{gl}(V)$. Then the identical representation of $\mathfrak{g}$ on V is irreducible and faithful, so that the preceding proposition implies that $[\mathfrak{g}, \mathrm{rad}(\mathfrak{g})] = \{0\}$, so that $\mathfrak{g}$ is reductive by Propostion 7.4. This implies that

$$\mathrm{rad}(\mathfrak{gl}(V)) = \mathfrak{z}(\mathfrak{gl}(V)) = \mathbb{K}\mathbf{1}$$

and that $\mathfrak{sl}(V) = [\mathfrak{gl}(V), \mathfrak{gl}(V)]$ is a Levi complement in $\mathfrak{gl}(V)$ (Example 1.7). Note that we use $\mathrm{char}\, \mathbb{K} = 0$ to see that $\mathrm{tr}\, \mathbf{1} \neq 0$.
(b) Let V be a finite dimensional vector space and $\mathcal{F} = (V_0, \ldots, V_n)$ a flag in V. Then

$$\mathfrak{r} := \{\varphi \in \mathfrak{g}(\mathcal{F})\colon (\forall i)(\exists \lambda_i \in \mathbb{K})\ (\varphi - \lambda_i \mathbf{1})(V_i) \subseteq V_{i-1}\} \cong \mathfrak{g}_n(\mathcal{F}) \rtimes \mathbb{K}^n$$

is a solvable ideal of $\mathfrak{g}(\mathcal{F})$ because $[\mathfrak{g}, \mathfrak{r}] \subseteq \mathfrak{g}_n(\mathcal{F})$. Choosing subspaces $W_1, \ldots, W_n \subseteq V$ with $V_i = W_1 \oplus \ldots \oplus W_i$, we have

$$\mathfrak{g}(\mathcal{F}) \cong \mathfrak{g}_n(\mathcal{F}) \rtimes \bigoplus_{i=1}^n \mathfrak{gl}(W_i),$$

(Example 1.31), and this leads to

$$\mathfrak{g}(\mathcal{F}) \cong \mathfrak{r} \rtimes \bigoplus_{i=1}^n \mathfrak{sl}(W_i).$$

Since we know from (a) that the Lie algebras $\mathfrak{sl}(W_i)$ are semisimple, it follows that $\mathfrak{r} = \mathrm{rad}(\mathfrak{g}(\mathcal{F}))$ and that $\bigoplus_{i=1}^n \mathfrak{sl}(W_i)$ is a Levi complement.

Remark 7.7. Let $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ be a Lie subalgebra, where V is finite dimensional. Let $\mathcal{F} = (V_0, \ldots, V_k)$ be a maximal flag of $\mathfrak{g}$-invariant subspaces. We fix a Levi decomposition $\mathfrak{g} = \mathfrak{r} \rtimes \mathfrak{s}$, where $\mathfrak{r} = \mathrm{rad}(\mathfrak{g})$. Then V is a semisimple $\mathfrak{s}$-module by Weyl's Theorem, so that there exist subspaces W_i with $V_i = V_{i-1} \oplus W_i$, $i = 1, \ldots, k$. We then have

$$\mathfrak{g} \subseteq \mathfrak{g}(\mathcal{F}) \quad \text{and} \quad \mathfrak{s} \subseteq \mathfrak{g}_s(\mathcal{F}) \cong \oplus_{j=1}^k \mathfrak{gl}(W_j)$$

(cf. Example 1.31). Since the $\mathfrak{g}$-modules V_i/V_{i-1} are simple, the ideal $[\mathfrak{g}, \mathfrak{r}]$ acts trivially on this quotient (Proposition 7.5), and we further obtain

$$[\mathfrak{g}, \mathfrak{r}] \subseteq \mathfrak{g}_u(\mathcal{F}).$$

This shows that

$$[\mathfrak{g}, \mathfrak{g}] = [\mathfrak{g}, \mathfrak{r}] \rtimes \mathfrak{s} \subseteq \mathfrak{g}_u(\mathcal{F}) \rtimes \mathfrak{g}_s(\mathcal{F}) = \mathfrak{g}(\mathcal{F})$$

is adapted to the semidirect decomposition $\mathfrak{g}(\mathcal{F}) \cong \mathfrak{g}_u(\mathcal{F}) \rtimes \mathfrak{g}_s(\mathcal{F})$.

Assume, in addition, that $\mathbb{K}$ is algebraically closed. Then $\mathfrak{r}/[\mathfrak{g}, \mathfrak{r}]$ is central in the reductive quotient algebra $\mathfrak{g}/[\mathfrak{g}, \mathfrak{r}]$ which acts on the simple $\mathfrak{g}$-modules $U = V_i/V_{i-1}$. Since $\rho_U(\mathfrak{r})$ has a simultaneous eigenvector by Lie's Theorem and the corresponding eigenspace

$$U_\lambda(\mathfrak{r}) := \{u \in U \mid (\forall x \in \mathfrak{r}) \;\; \rho_U(x)u = \lambda(x)u\}$$

is $\mathfrak{g}$-invariant because $[\rho_U(\mathfrak{r}), \rho_U(\mathfrak{g})] = \{0\}$, the simplicity of U implies that $\rho_U(\mathfrak{r}) \subseteq \mathbb{C}\,\mathrm{id}_U$. This means that

$$\mathfrak{r} \subseteq \mathfrak{g}_u(\mathcal{F}) \oplus \bigoplus_{j=1}^k \mathbb{C}\,\mathrm{id}_{W_j} = \mathrm{rad}(\mathfrak{g}(\mathcal{F})).$$

Further, the perfectness of the Levi complement $\mathfrak{s}$ shows that $\rho_{W_j}(\mathfrak{s}) \subseteq \mathfrak{sl}(W_j)$. This finally shows that the Levi decomposition of $\mathfrak{g}$ is fully adapted to the Levi decomposition

$$\mathfrak{g}(\mathcal{F}) = \mathrm{rad}(\mathfrak{g}(\mathcal{F})) \rtimes \oplus_{j=1}^k \mathfrak{sl}(W_j).$$

Notes on Part I

The Jacobi identity was discovered around 1830 by Carl Gustav Jacob Jacobi (1804–1851) as an identity for the Poisson bracket $\{\cdot,\cdot\}$ on smooth functions on $\mathbb{R}^{2n}$ (Exercise 1.2).

The term Lie algebra was introduced in the 1920s by Hermann Weyl, following a suggestion of Nathan Jacobson. Sophus Lie himself was dealing mainly with Lie algebras of vector fields (Exercise 1.3), which he called (infinitesimal) transformation groups. The term "Lie group" was introduced later by Élie Cartan.

The Jordan decompositions and the Jordan normal form are due to Camille Jordan (1838–1922). In the 1870s he wrote a text book on Galois theory of polynomial equations, thus making the ideas of Évariste Galois' (1811–1832), developped shortly before his untimely death, available to the mathematical world. This promoted group theoretical ideas considerably. In particular, it inspired Sophus Lie to work on a "Galois theory" for differential equations, using symmetries of differential equations to understand the structure of their solutions.

In the original proof of his theorem, Weyl used the famous "unitary trick". For $\mathbb{K} = \mathbb{C}$ one can derive Weyl's theorem on complete reducibility from the representation theory of compact groups (cf. [HiNe12, Ch. 15]). For $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{R})$ this works roughly as follows. One shows that the complex representations of $\mathfrak{g}$ are in one-to-one correspondence with the complex representations of $\mathfrak{sl}_n(\mathbb{C})$, resp., its real form $\mathfrak{su}_n(\mathbb{C})$, hence further with unitary representations of the compact simply connected Lie group $\mathrm{SU}_n(\mathbb{C})$. For unitary representations complete reducibility is a simple consequence of the existence of an invariant inner product. A purely algebraic proof was found later in the 1935 by Hendrik Brugt Gerhard Casimir (1909–2000) and Bartel Leendert van der Waerden (1903–1996) [CW35], after Casimir had dealt with the case $\mathfrak{sl}_2(\mathbb{C})$ using the operator named after him. Another algebraic proof was found in 1935 by Richard Dagobert Brauer (1911–1977) [Br36]. A completely different approach based on Lie algebra cohomology has been developed by John Henry Constantine Whitehead (1904–1960).

The original proof of Levi's Theorem for complex Lie algebras [Le05] was based on the classification of simple Lie algebras. The classification free proof for real Lie algebras given here goes back to Whitehead [Wh36]. The conjugacy of the Levi complements was shown by Anatoly Ivanovich Malcev (1909–1967) in [Ma42].

For a detailed account of the early history of Lie theory up to 1926 we refer to the book [Haw00] of Thomas Hawkins.

Part II
Root Decomposition

Since a simple Lie algebra $\mathfrak{g}$ has no other ideals than $\mathfrak{g}$ and $\{0\}$, we cannot analyze its structure by breaking it up into an ideal $\mathfrak{n}$ and the corresponding quotient algebra $\mathfrak{g}/\mathfrak{n}$. We therefore need refined tools to understand the internal structure of simple Lie algebras. It turns out that toral Cartan subalgebras and the corresponding root decompositions provide such a tool.

Roots and root spaces have remarkable properties some of which one turns into a system of axioms for *abstract root systems*. We derive a number of additional properties from these axioms. Moreover, we define certain objects associated with abstract roots systems like Weyl groups and Weyl chambers. Using these structural elements one could proceed rather easily to a complete classification of complex simple Lie algebras, but we refrain from doing this since our emphasis is on structure rather than classification.

In this part of these lecture notes, we first develop the concept of a toral Cartan subalgebra and root decompositions for general Lie algebras. Then we turn to semisimple Lie algebras and we finally discuss the geometry of the root system.

8 Weight and Root Space Decompositions

For a better understanding of the structure of a Lie algebra $\mathfrak{g}$, one decomposes it into simultaneous eigenspaces of operator sets $\operatorname{ad}\mathfrak{h}$ for a subalgebra $\mathfrak{h}$. Subalgebras for which this is possible are called *toral*.

8.1 Weights and Roots

Root decompositions are the simultaneous eigenspace decompositions of the type mentioned above. They are special cases of weight decompositions.

Definition 8.1. (a) Let (π, V) be a representation of the Lie algebra $\mathfrak{h}$. For a function $\lambda\colon \mathfrak{h} \to \mathbb{K}$, we define the corresponding *weight space* and the corresponding *generalized weight space* by

$$V_\lambda(\mathfrak{h}) := \bigcap_{x\in\mathfrak{h}} V_{\lambda(x)}(\pi(x)) \quad \text{and} \quad V^\lambda(\mathfrak{h}) := \bigcap_{x\in\mathfrak{h}} V^{\lambda(x)}(\pi(x)).$$

Any function $\lambda\colon \mathfrak{h} \to \mathbb{K}$ for which $V^\lambda(\mathfrak{h}) \neq \{0\}$ is called a *weight of the representation* (π, V). We write $\mathcal{P}_\mathfrak{h}(V)$ for the set of weights of (π, V).

(b) A module V of the Lie algebra $\mathfrak{h}$ is called *diagonalizable* if V is the sum of all weight spaces $V_\alpha(\mathfrak{h})$. Recall from Exercise 4.1 that the sum $\sum_{\mu\in\mathcal{P}_\mathfrak{h}(V)} V_\mu(\mathfrak{h})$ of weight spaces is direct, so that

$$V = \bigoplus_{\mu\in\mathcal{P}_\mathfrak{h}(V)} V_\mu(\mathfrak{h}).$$

If $\pi\colon \mathfrak{h} \to \mathfrak{gl}(V)$ is the corresponding representation, then this means that the subset $\pi(\mathfrak{h}) \subseteq \mathfrak{gl}(V)$ is simultaneously diagonalizable. Note that this implies that $\pi(\mathfrak{h})$ is abelian (Exercise 8.1).

(c) A subalgebra $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is called *toral* or *splitting*, if $\operatorname{ad}\mathfrak{h}$ is simultaneously diagonalizable. Then the weights of the representation $\pi = \operatorname{ad}|_\mathfrak{h}$ which are different from zero are called *roots* of $\mathfrak{g}$ with respect to $\mathfrak{h}$. The set of all roots is denoted $\Delta(\mathfrak{g}, \mathfrak{h})$. The weight spaces $\mathfrak{g}^\lambda(\mathfrak{h}) = \mathfrak{g}_\lambda(\mathfrak{h})$ are called *root spaces*. Sometimes we write $\mathfrak{g}^\lambda$ instead of $\mathfrak{g}^\lambda(\mathfrak{h})$. If $0 \neq \mu \in \mathfrak{h}^*$ is not a root, we put $\mathfrak{g}^\mu := \{0\}$.

The decomposition

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \Delta(\mathfrak{g},\mathfrak{h})} \mathfrak{g}_\alpha \tag{12}$$

is called the *root decomposition of* $\mathfrak{g}$ *w.r.t.* $\mathfrak{h}$.

Lemma 8.2. *Let* $\mathfrak{h} \subseteq \mathfrak{g}$ *be a toral subalgebra. Then*

(i) $\mathfrak{h}$ *is abelian.*

(ii) $\mathfrak{h} = \mathfrak{g}_0(\mathfrak{h})$ *if and only if* $\mathfrak{h}$ *is maximal abelian in* $\mathfrak{g}$.

(iii) *If* $\mathfrak{h} = \mathfrak{g}_0(\mathfrak{h})$, *then* $\mathfrak{z}(\mathfrak{g}) = \{x \in \mathfrak{h} \colon (\forall \alpha \in \Delta)\alpha(x) = 0\}$.

Proof. (i) Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a toral subalgebra and $x, y \in \mathfrak{h}$. Since $\operatorname{ad} \mathfrak{h}$ is abelian, we have $\operatorname{ad}[x, y] = [\operatorname{ad} x, \operatorname{ad} y] = 0$, hence $[x, y] \in \mathfrak{z}(\mathfrak{g})$. Therefore $(\operatorname{ad} x)^2(y) = 0$. Now the diagonalizability of $\operatorname{ad} x$ implies that $\operatorname{ad} x(y) = [x, y] = 0$.

(ii) Since $\mathfrak{g}_0 = \mathfrak{z}_{\mathfrak{g}}(\mathfrak{h})$ is the centralizer of $\mathfrak{h}$ in $\mathfrak{g}$, the subalgebra $\mathfrak{h}$ is maximal abelian if and only if $\mathfrak{h} = \mathfrak{g}_0$.

(iii) Clearly $\mathfrak{z}(\mathfrak{g}) \subseteq \mathfrak{g}_0(\mathfrak{h}) \subseteq \mathfrak{h}$. An element $x \in \mathfrak{h}$ is central if and only if all roots vanish on x. $\square$

Definition 8.3. A toral subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ which is maximal abelian in $\mathfrak{g}$ is called a *toral Cartan subalgebra of* $\mathfrak{g}$ [11]

Remark 8.4. (a) Toral Cartan subalgebras do not always exist: If $\mathfrak{g}$ is nilpotent, then all operators $\operatorname{ad} x$ are nilpotent. Therefore the diagonalizability of $\operatorname{ad} x$ is equivalent to $\operatorname{ad} x = 0$. We conclude that $\mathfrak{z}(\mathfrak{g})$ is maximal toral. But if $\mathfrak{g}$ is not abelian, then $\mathfrak{z}(\mathfrak{g})$ is not maximal abelian.

(b) In the structure theory of finite dimensional Lie algebras, a subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is called a *Cartan subalgebra* if $\mathfrak{h}$ is nilpotent and self-normalizing, i.e.,

$$\mathfrak{h} = \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}) := \{x \in \mathfrak{g} \colon [x, \mathfrak{h}] \subseteq \mathfrak{h}\}.$$

If $\mathfrak{h}$ is toral, then $\mathfrak{h}$ is abelian and therefore nilpotent. Further, $\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{z}_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{g}_0$ (Exercise!). Hence a toral subalgebra is a Cartan subalgebra if $\mathfrak{h} = \mathfrak{g}_0$.

8.2 Examples of Root Decompositions

Example 8.5. (a) (The Lie algebra $\mathfrak{sl}_2(\mathbb{K})$) We recall the basis (h, e, f) of $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{K})$ given by

$$h := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{13}$$

satisfying the commutator relations

$$[h, e] = 2e, \quad [h, f] = -2f \quad \text{and} \quad [e, f] = h. \tag{14}$$

[11] Cartan subalgebras of complex simple Lie algebras were first used by W. Killing in his classification, before É. Cartan used them about 10 years later. Conversely, the Cartan–Killing form (often called the Killing form) was much more in the focus of Cartan's work than in Killing's. Note that Cartan's criteria for solvability and semisimplicity both refer to properties of this form.

We conclude that $\mathfrak{h} := \mathbb{K}h$ is a toral Cartan subalgebra. With $\alpha(h) = 2$ we thus obtain the root decomposition

$$\mathfrak{g} = \mathfrak{h} + \mathfrak{g}_\alpha + \mathfrak{g}_{-\alpha} \quad \text{with} \quad \mathfrak{g}_\alpha = \mathbb{K}e \quad \text{and} \quad \mathfrak{g}_{-\alpha} = \mathbb{K}f.$$

(b) Let $\mathfrak{b}$ be the 2-dimensional non-abelian Lie algebra with the basis (h, x) satisfying $[h, x] = x$ (Exercise 1.4). Then $\mathfrak{h} := \mathbb{K}h$ is a toral Cartan subalgebra and with $\alpha(h) = 1$ we obtain the root decomposition

$$\mathfrak{g} = \mathfrak{h} + \mathfrak{g}_\alpha \quad \text{with} \quad \mathfrak{g}_\alpha = \mathbb{K}y.$$

In this case there exists only one root.

(c) In $\mathfrak{g} = \mathfrak{so}_3(\mathbb{C})$, we consider a basis (x, y, z) with

$$[x, y] = z, \quad [y, z] = x, \quad [z, x] = y.$$

(Example 1.35). Then $\mathfrak{h} = \mathbb{C}x$ is a toral Cartan subalgebra and

$$[x, y \mp iz] = z \pm iy = \pm i(y \mp iz)$$

implies that we have a root decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \quad \text{with} \quad \alpha(x) = i, \quad \mathfrak{g}_{\pm\alpha} = \mathbb{C}(y \mp iz).$$

Example 8.6. (The Lie algebra $\mathfrak{gl}_n(\mathbb{K})$) In $\mathfrak{g} := \mathfrak{gl}_n(\mathbb{K})$, we consider the basis $(E_{jk})_{1 \leq j,k \leq n}$ defined by $E_{jk}e_m = \delta_{km}e_j$ for the canonical basis $(e_j)_{1 \leq j \leq n}$ of $\mathbb{K}^n$. From Example 1.7 we recall the commutation relations

$$[E_{jk}, E_{\ell m}] = \delta_{k\ell}E_{jm} - \delta_{jm}E_{\ell k}. \tag{15}$$

The subalgebra $\mathfrak{h} := \operatorname{span}\{E_{jj} : j = 1, \ldots, n\}$ of diagonal matrices is abelian and (15) implies

$$[E_{mm}, E_{jk}] = \delta_{jm}E_{mk} - \delta_{km}E_{jm} = (\delta_{jm} - \delta_{km})E_{jk}.$$

Defining linear functionals $\varepsilon_j : \mathfrak{h} \to \mathbb{K}$ by $\varepsilon_j(E_{kk}) = \delta_{jk}$, it follows that $E_{jk} \in \mathfrak{g}_{\varepsilon_j - \varepsilon_k}$. This leads to the root decomposition

$$\mathfrak{h} = \mathfrak{g}_0 \quad \text{and} \quad \mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{j \neq k} \mathfrak{g}_{\varepsilon_j - \varepsilon_k} = \mathfrak{h} \oplus \bigoplus_{j \neq k} \mathbb{K}E_{jk}.$$

In particular, $\mathfrak{h}$ is a toral Cartan subalgebra and

$$\Delta = \{\varepsilon_j - \varepsilon_k : j \neq k \in \{1, \ldots, n\}\}$$

is the corresponding set of roots.

Example 8.7. (The special linear Lie algebra) For $\mathfrak{g} := \mathfrak{sl}_n(\mathbb{K})$ we obtain as in Example 8.6 that the subalgebra $\mathfrak{h} = \{\operatorname{diag}(h) : \sum_j h_j = 0\}$ of diagonal matrices in $\mathfrak{g}$ is a toral Cartan subalgebra, and with $\varepsilon_j(\operatorname{diag}(h)) := h_j$, we obtain a root decomposition $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha$ with $\mathfrak{g}_{\varepsilon_j - \varepsilon_k} = \mathbb{K}E_{jk}$ and the root system

$$A_{n-1} := \{\varepsilon_j - \varepsilon_k : 1 \leq j \neq k \leq n\}.$$

Example 8.8. (The orthogonal Lie algebras) (a) Let

$$I_{n,n} := \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \in M_{2n}(\mathbb{K}) \cong M_2(M_n(\mathbb{K}))$$

and consider the Lie algebra

$$\mathfrak{o}_{n,n}(\mathbb{K}) = \{x \in \mathfrak{gl}_{2n}(\mathbb{K}) \colon x^\top I_{n,n} + I_{n,n} x = 0\}.$$

(cf. Exercise 1.10 for the fact that this is indeed a Lie algebra). In terms of block matrices, we then have

$$\mathfrak{o}_{n,n}(\mathbb{K}) = \Big\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(M_n(\mathbb{K})) \colon a^\top = -a, d^\top = -d, b^\top = c \Big\}.$$

In this matrix presentation, it is quite inconvenient to describe a root decomposition of this Lie algebra. It is much simpler to use an equivalent description, based on the following observation. For

$$g := \begin{pmatrix} \mathbf{1} & \frac{1}{2}\mathbf{1} \\ \mathbf{1} & -\frac{1}{2}\mathbf{1} \end{pmatrix}$$

we find

$$g^\top I_{n,n} g = \begin{pmatrix} \mathbf{1} & \mathbf{1} \\ \frac{1}{2}\mathbf{1} & -\frac{1}{2}\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{1} & \frac{1}{2}\mathbf{1} \\ \mathbf{1} & -\frac{1}{2}\mathbf{1} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \mathbf{1} \\ \frac{1}{2}\mathbf{1} & -\frac{1}{2}\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{1} & \frac{1}{2}\mathbf{1} \\ -\mathbf{1} & \frac{1}{2}\mathbf{1} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} =: S.$$

Hence $x \in \mathfrak{o}_{n,n}(\mathbb{K})$ is equivalent to $g^{-1}xg$ being contained in

$$\mathfrak{g} := \mathfrak{o}_{2n}(\mathbb{K}, S) := \{x \in \mathfrak{gl}_{2n}(\mathbb{K}) \colon x^\top S + Sx = 0\} = g^{-1}\mathfrak{o}_{n,n}(\mathbb{K})g$$

and $\varphi \colon \mathfrak{o}_{n,n}(\mathbb{K}) \to \mathfrak{o}_{2n}(\mathbb{K}, S), x \mapsto g^{-1}xg$ is an isomorphism of Lie algebras. From

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathfrak{o}_{2n}(\mathbb{K}, S) \quad \Leftrightarrow \quad d = -a^\top, b^\top = -b, c^\top = -c,$$

we immediately derive that $\mathfrak{g}$ has a root decomposition with respect to the maximal abelian subalgebra

$$\mathfrak{h} = \operatorname{span}\{E_{jj} - E_{n+j,n+j} : j = 1, \ldots, n\} = \{\operatorname{diag}(h, -h) \colon h \in \mathbb{K}^n\}.$$

The corresponding root system is

$$D_n := \{\pm\varepsilon_j \pm \varepsilon_k : j, k = 1, \ldots, n, j \neq k\},$$

where $\varepsilon_j \colon \mathfrak{h} \to \mathbb{K}$ is the linear functional defined by $\varepsilon_k(\operatorname{diag}(h, -h)) := h_k$. Here the roots in the subsystem A_{n-1} of D_n correspond to the root spaces in the image of the embedding

$$\mathfrak{gl}_n(\mathbb{K}) \to \mathfrak{o}_{2n}(\mathbb{K}, S), \quad x \mapsto \begin{pmatrix} x & 0 \\ 0 & -x^\top \end{pmatrix}.$$

Further, $\mathfrak{g}_{\varepsilon_j+\varepsilon_k} = \mathbb{K}(E_{j,n+k} - E_{k,n+j})$ and $\mathfrak{g}_{-\varepsilon_j-\varepsilon_k} = \mathbb{K}(E_{n+k,j} - E_{n+j,k})$.

(b) For the symmetric matrix

$$T := \begin{pmatrix} \mathbf{0} & \mathbf{1} & 0 \\ \mathbf{1} & \mathbf{0} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} S & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} \in M_{2n+1}(\mathbb{K}),$$

we also obtain a Lie algebra

$$\mathfrak{o}_{2n+1}(\mathbb{K}, T) := \{x \in \mathfrak{gl}_{2n+1}(\mathbb{K}) \colon x^\top T + Tx = 0\}.$$

Then

$$\begin{pmatrix} a & b & x \\ c & d & y \\ \widetilde{y} & \widetilde{x} & z \end{pmatrix} \in \mathfrak{o}_{2n+1}(\mathbb{K}, T) \quad \Leftrightarrow \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathfrak{o}_{2n}(\mathbb{K}, S),\ \widetilde{x} = -x^\top,\ \widetilde{y} = -y^\top, z = 0$$

implies that this Lie algebra has a root decomposition with respect to the maximal abelian subalgebra

$$\mathfrak{h} = \mathrm{span}\{E_{jj} - E_{n+j,n+j} : j = 1, \ldots, n\} = \{\mathrm{diag}(h, -h, 0) \colon h \in \mathbb{K}^n\}.$$

The corresponding root system is

$$B_n := \{\pm\varepsilon_j, \pm\varepsilon_j \pm \varepsilon_k : j, k = 1, \ldots, n, j \neq k\},$$

where $\varepsilon_j \colon \mathfrak{h} \to \mathbb{K}$ is the linear functional defined by $\varepsilon_k(\mathrm{diag}(h, -h, 0)) := h_k$. Here the root spaces corresponding to roots in the subsystem D_n of B_n correspond to root spaces in the subalgebra $\mathfrak{o}_{2n}(\mathbb{K}, S)$ (corresponding to $x = y = 0$), and

$$\mathfrak{g}_{\varepsilon_j} = \mathbb{K}(E_{j,2n+1} - E_{2n+1,n+j}) \quad \text{and} \quad \mathfrak{g}_{-\varepsilon_j} = \mathbb{K}(E_{j+n,2n+1} - E_{2n+1,j}).$$

Example 8.9. (The symplectic Lie algebra) For the skew-symmetric matrix

$$J := \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \in M_{2n}(\mathbb{K}),$$

we obtain the symplectic Lie algebra

$$\mathfrak{sp}_{2n}(\mathbb{K}) = \{x \in \mathfrak{gl}_{2n}(\mathbb{K}) \colon x^\top J + Jx = 0\}.$$

Using

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathfrak{sp}_{2n}(\mathbb{K}) \quad \Leftrightarrow \quad d = -a^\top, b^\top = b, c^\top = c,$$

(Example 1.7(vii)) we see that $\mathfrak{g} := \mathfrak{sp}_{2n}(\mathbb{K})$ has a root decomposition with respect to the maximal abelian subalgebra

$$\mathfrak{h} = \mathrm{span}\{E_{jj} - E_{n+j,n+j} : j = 1, \ldots, n\} = \{\mathrm{diag}(h, -h) \colon h \in \mathbb{K}^n\}.$$

The corresponding root system is

$$C_n := \{\pm 2\varepsilon_j, \pm\varepsilon_j \pm \varepsilon_k : j, k = 1, \ldots, n, j \neq k\},$$

where $\varepsilon_j \colon \mathfrak{h} \to \mathbb{K}$ is the linear functional defined by $\varepsilon_k(\mathrm{diag}(h, -h)) := h_k$. Again, the roots in the subsystem A_{n-1} of C_n correspond to the root spaces in the image of the embedding

$$\mathfrak{gl}_n(\mathbb{K}) \to \mathfrak{sp}_{2n}(\mathbb{K}), \quad x \mapsto \begin{pmatrix} x & 0 \\ 0 & -x^\top \end{pmatrix}.$$

Further,

$$\mathfrak{g}_{\varepsilon_j + \varepsilon_k} = \mathbb{K}(E_{j,n+k} + E_{k,n+j}) \quad \text{and} \quad \mathfrak{g}_{-\varepsilon_j - \varepsilon_k} = \mathbb{K}(E_{n+k,j} + E_{n+j,k}).$$

8.3 General Facts on Weights and Roots

Having discussed a bunch of examples, we proceed with some general observations concerning weight and root decompositions.

Lemma 8.10. *Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a toral subalgebra and $\mathfrak{g} = \mathfrak{g}_0 + \sum_{\alpha\in\Delta} \mathfrak{g}_\alpha$ be the corresponding root space decomposition. Then the following assertions hold:*

(i) *If V is a $\mathfrak{g}$-module, $\alpha \in \Delta \cup \{0\}$ and $\beta \in \mathcal{P}_V(\mathfrak{h})$, then*

$$\mathfrak{g}_\alpha.V_\beta \subseteq V_{\alpha+\beta}. \tag{16}$$

(ii) *For $\alpha, \beta \in \Delta \cup \{0\}$, we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.*

(iii) *$\mathfrak{g}^0$ is a subalgebra of $\mathfrak{g}$.*

Proof. (i) For $v_\beta \in V_\beta$, $x \in \mathfrak{h}$ and $y \in \mathfrak{g}_\alpha$, we obtain

$$x.(y.v_\beta) = y.(x.v_\beta) + [x, y].v_\beta = \beta(x)y.v_\beta + \alpha(x)y.v_\beta = (\alpha + \beta)(x)y.v_\beta.$$

(ii) follows from (i), applied to the adjoint module $V = \mathfrak{g}$ with $x \cdot y := [x, y]$. Note that the corresponding weight set is $\Delta \cup \{0\}$.

(iii) is a direct consequence of (ii). □

Example 8.11. In the $\mathfrak{gl}_n(\mathbb{K})$-module $V := \mathbb{K}^n$, every basis vector e_j, $j = 1, \ldots, n$, is an $\mathfrak{h}$-weight vector of weight ε_j. The $\mathfrak{h}$-module V is diagonalizable with weight set

$$\mathcal{P}_V = \{\varepsilon_1, \ldots, \varepsilon_n\}.$$

Then $E_{jk}V_{\varepsilon_\ell} \subseteq \delta_{k,\ell}V_{\varepsilon_j}$ is a special case of Lemma 8.10(i).

The following observation will become useful later on.

Proposition 8.12. *If V is a diagonalizable module of the abelian Lie algebra $\mathfrak{h}$ and $W \subseteq V$ a submodule, then W is adapted to the weight decomposition, i.e.,*

$$W = \bigoplus_{\alpha\in\mathcal{P}_V} (W \cap V_\alpha) = \bigoplus_{\alpha\in\mathcal{P}_W} W_\alpha.$$

Proof. Clearly, $W_0 := \bigoplus_{\alpha\in\mathcal{P}_V}(W \cap V_\alpha)$ is an $\mathfrak{h}$-submodule of W. It therefore remains to show that, if we write $w \in W$ as $w = \sum_\alpha w_\alpha$ with $w_\alpha \in V_\alpha$, then all components w_α are contained in W. We prove this assertion by induction on the number N of non-zero summands w_α. For $N = 1$, there is nothing to show. So we assume $N > 1$. Let $w_\alpha \neq 0$. Then there exists $\beta \in \mathcal{P}_V \setminus \{\alpha\}$ with $w_\beta \neq 0$. We choose $x \in \mathfrak{h}$ with $\alpha(x) \neq \beta(x)$. Then

$$x.w - \beta(x)w = \sum_{\gamma\neq\beta}(\gamma(x) - \beta(x))w_\gamma$$

is a sum of at most $N-1$ nonzero summands. Since $\alpha(x)-\beta(x) \neq 0$, we obtain $w_\alpha \in W$. □

Corollary 8.13. *If $\mathfrak{h} \subseteq \mathfrak{g}$ is a toral subalgebra and $\mathfrak{a} \subseteq \mathfrak{g}$ an $\mathfrak{h}$-invariant subalgebra, i.e., $[\mathfrak{h}, \mathfrak{a}] \subseteq \mathfrak{a}$, then $\mathfrak{a}$ is adapated to the root decomposition:*

$$\mathfrak{a} = \mathfrak{a}_0 \oplus \bigoplus_{\alpha\in\Delta} \mathfrak{a}_\alpha.$$

In particular, all ideals $\mathfrak{a} \trianglelefteq \mathfrak{g}$ are adapted to the root decomposition.

Proof. We apply Proposition 8.12 to the diagonalizable $\mathfrak{h}$-module $\mathfrak{g}$. $\square$

The following lemma is a useful tool to see that the examples discussed below are indeed semisimple Lie algebras.

Lemma 8.14. *Suppose that* $\mathfrak{g} = \mathfrak{h} + \sum_{\alpha\in\Delta} \mathfrak{g}_\alpha$ *is a Lie algebra and* $\mathfrak{h}$ *a toral Cartan subalgebra, such that*

(i) $\mathfrak{g}(\alpha) := \mathfrak{g}_\alpha + \mathfrak{g}_{-\alpha} + [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \cong \mathfrak{sl}_2(\mathbb{K})$ *for each root* α*, and*

(ii) $\mathfrak{z}(\mathfrak{g}) = \{0\}$.

Then $\mathfrak{g}$ *is semisimple.*

Proof. Let $\mathfrak{r} := \mathrm{rad}(\mathfrak{g})$ be the solvable radical of $\mathfrak{g}$. As an ideal, it is $\mathfrak{h}$-invariant, hence adapted to the root space decomposition: $\mathfrak{r} = \mathfrak{r}_0 + \sum_\alpha \mathfrak{r}_\alpha$ (Proposition 8.12). Since all semisimple subalgebras $\mathfrak{s}$ of $\mathfrak{g}$ intersect $\mathfrak{r}$ trivially (otherwise $\mathfrak{s}\cap\mathfrak{r}$ would be a nontrivial solvable ideal of $\mathfrak{s}$), $\mathfrak{r}_\alpha \subseteq \mathfrak{g}(\alpha)\cap\mathfrak{r} = \{0\}$. Hence $\mathfrak{r}\subseteq\mathfrak{h}$, and, in view of $[\mathfrak{r},\mathfrak{g}_\alpha]\subseteq\mathfrak{r}\cap\mathfrak{g}_\alpha = \{0\}$, we get $\mathfrak{r}\subseteq\bigcap_{\alpha\in\Delta}\ker\alpha = \mathfrak{z}(\mathfrak{g}) = \{0\}$. $\square$

Exercises for Section 8

Exercise 8.1. Show that, if $\mathfrak{h}\subseteq\mathfrak{gl}(V)$ is a simultaneously diagonalizable Lie subalgebra, then $\mathfrak{h}$ is abelian.

Exercise 8.2. (a) Let $\mathfrak{h}$ be an abelian Lie algebra and V a diagonalizable $\mathfrak{h}$-module. Then $\mathfrak{h}$ is a toral subalgebra of $\mathfrak{g} := V \rtimes \mathfrak{h}$ (Example 1.28). The root decomposition is given by

$$\mathfrak{g} = (\mathfrak{h} + V_0) + \sum_{\alpha\in\mathcal{P}_V\setminus\{0\}} V_\alpha.$$

The subalgebra $\mathfrak{h}$ is a Cartan subalgebra if and only if $V_0 = \{0\}$.

(b) Let $\mathfrak{h}$ be an abelian Lie algebra and $\Delta\subseteq\mathfrak{h}^*\setminus\{0\}$ be a subset. Construct a Lie algebra $\mathfrak{g}$ containing $\mathfrak{h}$ as a toral Cartan subalgebra such that Δ is the corresponding set of roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$.

Exercise 8.3. For $Q_1, Q_2 \in M_n(\mathbb{K})$, we put

$$\mathfrak{o}(\mathbb{K}, Q_i) := \{x\in\mathfrak{gl}_n(\mathbb{K}) : x^\top Q_i = -Q_i x\}.$$

Show that: If $Q_2 = S^\top Q_1 S$ holds for some $S\in\mathrm{GL}_n(\mathbb{K})$, then the map

$$\Phi_S\colon \mathfrak{o}(\mathbb{K},Q_1)\to\mathfrak{o}(\mathbb{K},Q_2), \Phi_S(x) := S^{-1}xS$$

is an isomorphism of Lie algebras.

Exercise 8.4. Let $Q = Q^\top \in \mathrm{GL}_n(\mathbb{K})$ and suppose that $\mathbb{K}$ is algebraically closed. Show that:

(i) The corresponding symmetric bilinear form $\beta(x,y) := x^\top Q y$ on $\mathbb{K}^n$ is non-degenerate.

(ii) There exists a β-orthonormal basis. Conclude that there exists an $S\in\mathrm{GL}_n(\mathbb{K})$ with $S^\top QS = \mathbf{1}$ and that $\mathfrak{o}_n(\mathbb{K},Q) = \{x\in\mathfrak{gl}_n(\mathbb{K})\colon x^\top Q + Qx = 0\}\cong\mathfrak{o}_n(\mathbb{K})$.

(iii) If n is even and $k := n/2$, then there exists a basis $b_1, \ldots, b_n$ of $\mathbb{K}^n$ with

$$\beta(b_i, b_{i+k}) = \beta(b_{i+k}, b_i) = 1 \quad \text{and} \quad 0 \text{ otherwise.}$$

Conclude that there exists an $S \in \mathrm{GL}_n(\mathbb{K})$ with $S^\top QS = Q' := \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}$ and that $\mathfrak{o}_{2k}(\mathbb{K}, Q) \cong \mathfrak{o}_{2k}(\mathbb{K}, Q')$ (cf. Example 8.8(a)).

(iv) If n is odd and $k := (n-1)/2$, then there exists a basis $b_1, \ldots, b_{2k+1}$ of $\mathbb{K}^n$ such that $\beta(b_i, b_j)$ is as under (iii) for $1 \leq i, j \leq 2k$, $\beta(b_n, b_n) = 1$ and $\beta(b_j, b_n) = 0$ for $j < n$. Conclude that there exists an $S \in \mathrm{GL}_n(\mathbb{K})$ with

$$S^\top QS = Q' := \begin{pmatrix} \mathbf{0} & \mathbf{1} & 0 \\ \mathbf{1} & \mathbf{0} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and that $\mathfrak{o}_{2k+1}(\mathbb{K}, Q) \cong \mathfrak{o}_{2k+1}(\mathbb{K}, Q')$ (cf. Example 8.8(b)).

Exercise 8.5. Let $\mathfrak{g} = \mathbb{R}h + \mathbb{R}p + \mathbb{R}q + \mathbb{R}z$ be the oscillator algebra with the bracket relations

$$[p, q] = z, \quad [h, p] = q \quad \text{and} \quad [h, q] = -p$$

(z is central). Determine a toral Cartan subalgebra of the complexification $\mathfrak{g}_\mathbb{C}$ and the corresponding root decomposition.

Exercise 8.6. Suppose that $\mathbb{K}$ is algebraically closed of characteristic zero and that $\mathfrak{h}$ is a nilpotent Lie algebra. We consider a finite dimensional $\mathfrak{h}$-module V with the corresponding representation $\rho \colon \mathfrak{h} \to \mathfrak{gl}(V)$. Show that:

(i) $\rho(\mathfrak{h})$ commutes with all diagonalizable Jordan components $\rho(x)_s$, $x \in \mathfrak{h}$.
Hint: $(\mathrm{ad}\,\rho(x))_s = \mathrm{ad}(\rho(x)_s)$ on $\mathfrak{gl}(V)$.

(ii) $[\rho(x)_s, \rho(y)_s] = 0$ for $x, y \in \mathfrak{h}$. Hint: $\rho(x)_s$ is a polynomial in $\rho(x)$.

(iii) There exists a direct module decompositions $V = \oplus_{j \in J} V_j$ such that $\rho(x)_s|_{V_j} \in \mathbb{K}\mathbf{1}$ for $x \in \mathfrak{h}$.

(iv) $\lambda_j(x) := \frac{1}{\dim V_j} \mathrm{tr}(\rho(x)|_{V_j})$ defines a linear functional on $\mathfrak{h}$ with $V_j \subseteq V^{\lambda_j}(\mathfrak{h})$.

(v) The module V has a generalized weight space decomposition $V = \oplus_{j=1}^n V^{\lambda_j}(\mathfrak{h})$ and the generalized weight spaces are submodules.

9 Finite Dimensional $\mathfrak{sl}_2(\mathbb{K})$-Modules

As we shall see in Section 10 below, the Lie algebra $\mathfrak{sl}_2(\mathbb{K})$ is of particular importance because semisimple Lie algebras with toral Cartan subalgebras contain many subalgebras isomorphic to $\mathfrak{sl}_2(\mathbb{K})$ and the collection of these subalgebras essentially determines the structure of the whole Lie algebra. Therefore the representation theory of $\mathfrak{sl}_2(\mathbb{K})$ plays a key role in the structure theory of these Lie algebras.

In the following, we shall use the basis

$$h := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{17}$$

of $\mathfrak{sl}_2(\mathbb{K})$. It satisfies

$$[h, e] = 2e, \quad [h, f] = -2f \quad \text{and} \quad [e, f] = h. \tag{18}$$

9.1 A family of $\mathfrak{sl}_2$-modules

We start with a discussion of a concrete family of representations of $\mathfrak{sl}_2(\mathbb{K})$. It will turn out later that the study of this family already provides all irreducible finite dimensional representations of $\mathfrak{sl}_2(\mathbb{K})$.

Example 9.1. Let $\mathcal{A} := \mathbb{K}[Z, Z^{-1}]$ be the algebra of Laurent polynomials in Z. For any $f \in \mathcal{A}$ the operator $D := f\frac{d}{dZ}$ is a derivation of $\mathcal{A}$ (Product Rule) and any derivation of the algebra $\mathcal{A}$ is of this kind (Exercise 9.2). The Lie bracket on $\mathrm{der}(\mathcal{A})$ satisfies

$$\Big[f\frac{d}{dZ}, g\frac{d}{dZ}\Big] = (fg' - f'g)\frac{d}{dZ}. \tag{19}$$

For $\lambda \in \mathbb{K}$, we consider the operators

$$E := \frac{d}{dZ}, \quad F := -Z^2\frac{d}{dZ} + \lambda Z\mathbf{1}, \quad H := \lambda\mathbf{1} - 2Z\frac{d}{dZ}.$$

With (19) we obtain

$$\Big[Z^n\frac{d}{dZ}, Z^m\frac{d}{dZ}\Big] = (m-n)Z^{n+m-1}\frac{d}{dZ}$$

and hence the commutator relations

$$\begin{aligned}
[H,E] &= & -2\big[Z\tfrac{d}{dZ}, \tfrac{d}{dZ}\big] &= 2\tfrac{d}{dZ} &&= 2E,\\
[H,F] &= & -2\big[Z\tfrac{d}{dZ}, -Z^2\tfrac{d}{dZ} + \lambda Z\mathbf{1}\big] &= 2Z^2\tfrac{d}{dZ} - 2\lambda Z\mathbf{1} &&= -2F,\\
[E,F] &= & \big[\tfrac{d}{dZ}, -Z^2\tfrac{d}{dZ} + \lambda Z\mathbf{1}\big] &= -2Z\tfrac{d}{dZ} + \lambda\mathbf{1} &&= H.
\end{aligned}$$

Since these are precisely the commutator relations of $\mathfrak{g} := \mathfrak{sl}_2(\mathbb{K})$, we obtain by

$$e \mapsto E, \quad f \mapsto F, \quad h \mapsto H$$

a representation $\rho_\lambda \colon \mathfrak{sl}_2(\mathbb{K}) \to \mathrm{End}(\mathcal{A})$, resp., an $\mathfrak{sl}_2(\mathbb{K})$-module structure on $\mathcal{A}$.

To understand the structure of this module, we consider the action of the operators H, E and F on the canonical basis:

$$H \cdot Z^n = (\lambda - 2n)Z^n, \quad E \cdot Z^n = nZ^{n-1}, \quad F \cdot Z^n = (\lambda - n)Z^{n+1}. \tag{20}$$

In particular, we see that H is diagonalizable with one-dimensional eigenspaces. With this information, it is easy to determine all submodules. Any submodule is adapted to the eigenspace decomposition of H (Proposition 8.12). Hence each submodule is of the form

$$\mathcal{A}_J := \mathrm{span}\{Z^n \colon n \in J\} \quad \text{for some subset} \quad J \subseteq \mathbb{Z}.$$

From (20), we see that $\mathcal{A}_J$ is a submodule if and only if J satisfies the following conditions:
(i) If $n \in J$ and $n \neq 0$, then $n - 1 \in J$.
(ii) If $n \in J$ and $\lambda \neq n$, then $n + 1 \in J$.

If $\lambda \notin \mathbb{Z}$, then $\mathbb{K}[Z] = \mathcal{A}_{\mathbb{N}_0}$ is the only nontrivial submodule of $\mathcal{A}$. If $\lambda \in \mathbb{Z}$, then there are two possibilities. For $\lambda < 0$, the only proper subsets of $\mathbb{Z}$ satisfying (i) and (ii) are

$$\{\ldots, \lambda - 1, \lambda\}, \quad \mathbb{N}_0, \quad \text{and} \quad \{\ldots, \lambda - 1, \lambda\} \cup \mathbb{N}_0.$$

$$\circ \quad \circ \quad \circ \quad] \quad \circ \quad \circ \quad \cdots \quad \circ \quad [\quad \circ \quad \circ \quad \cdots.$$

For $\lambda \geq 0$, the subsets of $\mathbb{Z}$ defining submodules are

$$\mathbb{N}_0, \quad \{\ldots, \lambda - 1, \lambda\}, \quad \{0, 1, \ldots, \lambda - 1, \lambda\}.$$

$$\circ \quad \circ \quad \circ \quad [\quad \circ \quad \circ \quad \cdots \quad \circ \quad] \quad \circ \quad \circ \quad \cdots.$$

In this case we obtain in particular a finite dimensional submodule

$$L(\lambda) := \mathrm{span}\{1, Z, \ldots, Z^\lambda\}. \tag{21}$$

Since $L(\lambda)$ contains no nontrivial proper submodule, it is simple.

We have seen in the preceding example that, for each $\lambda \in \mathbb{N}_0$, there exists a simple $\mathfrak{sl}_2(\mathbb{K})$-module of dimension $\lambda+1$. Our next goal is to show that all simple finite dimensional modules are isomorphic to some $L(\lambda)$.

9.2 The classification

The following lemma specializes for $\mu = 1$ to an assertion on the Lie algebra $\mathfrak{sl}_2(\mathbb{K})$.

Lemma 9.2. *Let (e, h, f) be a triple of elements of an associative algebra $\mathcal{A}$, satisfying the commutator relations*

$$[h, e] = 2e, \quad [h, f] = -2f \quad \text{and} \quad [e, f] = h.$$

Then the following assertions hold:

(i) $[h, e^n] = 2ne^n$ *and* $[h, f^n] = -2nf^n$ *for* $n \in \mathbb{N}_0$.

(ii) *For* $n > 0$,

$$[f, e^n] = -ne^{n-1}\big(h + (n-1)\big) = -n\big(h - (n-1)\big)e^{n-1}$$

and

$$[e, f^n] = nf^{n-1}\big(h - (n-1)\big) = n\big(h + (n-1)\big)f^{n-1}.$$

Proof. (i) Since $\operatorname{ad} h(a) := ha - ah$ is a derivation of $\mathcal{A}$ and $[h, e] = 2e$ commutes with e, we obtain inductively $[h, e^n] = n[h, e]e^{n-1} = 2ne^n$. The second part of (i) is obtained similarly.
(ii) We calculate

$$\begin{aligned}
[f, e^n] &= \sum_{j=0}^{n-1} e^j[f, e]e^{n-j-1} = \sum_{j=0}^{n-1} e^j(-h)e^{n-j-1} \\
&= -\sum_{j=0}^{n-1} e^j[h, e^{n-j-1}] - \sum_{j=0}^{n-1} e^{n-1}h = -\Big(\sum_{j=0}^{n-1} 2(n-j-1)e^{n-1}\Big) - ne^{n-1}h \\
&= -\Big(\sum_{j=0}^{n-1} 2je^{n-1}\Big) - ne^{n-1}h = -n(n-1)e^{n-1} - ne^{n-1}h.
\end{aligned}$$

In view of (i), this equals

$$-n(n-1)e^{n-1} - nhe^{n-1} + n[h, e^{n-1}] = n(n-1)e^{n-1} - nhe^{n-1}.$$

This is the first part of (ii). The second part is reduced to the first one by considering the triple $(f, -h, e)$, satisfying the same commutation relations as (e, h, f). $\square$

Proposition 9.3. *Let V be a finite dimensional $\mathfrak{sl}_2(\mathbb{K})$-module and $v_0 \in V$ an element with $e \cdot v_0 = 0$ and $h \cdot v_0 = \lambda v_0$. Then*

(i) $\lambda \in \mathbb{N}_0$.

(ii) *v_0 generates a submodule isomorphic to $L(\lambda)$.*

Proof. (i) Let $V_\alpha := V_\alpha(h)$ be the h-eigenspace corresponding to the eigenvalue α on V, which is a weight space for the representation of the subalgebra $\mathfrak{h} = \mathbb{K}h$. From $v_0 \in V_\lambda$ and $[h, f] = -2f$, we obtain with Lemma 8.10 the relation $h \cdot (f^n \cdot v_0) = (\lambda - 2n)(f^n \cdot v_0)$.

We further obtain with Lemma 8.10:

$$e \cdot (f^n \cdot v_0) = [e, f^n] \cdot v_0 + f^n \cdot \underbrace{(e \cdot v_0)}_{=0} = nf^{n-1}(h - n + 1) \cdot v_0 = n(\lambda - n + 1)f^{n-1} \cdot v_0. \quad (22)$$

This shows that the submodule W generated by v_0 is

$$W = \mathrm{span}\{f^n \cdot v_0 \colon n \in \mathbb{N}_0\}.$$

Since V is finite dimensional, h has only finitely many eigenvalues on V. Hence there is a minimal $N \in \mathbb{N}_0$ with $f^{N+1} \cdot v_0 = 0$. From $e \cdot (f^{N+1} \cdot v_0) = 0$ we derive that $\lambda = N \in \mathbb{N}_0$.
(ii) To see that $W \cong L(\lambda)$, we consider the basis

$$v_k := \frac{f^k \cdot v_0}{\lambda(\lambda - 1) \cdots (\lambda - k + 1)}, \quad k = 0, \ldots, \lambda,$$

for W (note that the denominator never vanishes.). For this basis, we have

$$h \cdot v_k = (\lambda - 2k)v_k, \quad f \cdot v_k = (\lambda - k)v_{k+1}, \qquad e \cdot v_0 = 0$$

and, for $k > 0$ by (22),

$$e \cdot v_k = \frac{k(\lambda - k + 1)}{\lambda(\lambda - 1) \cdots (\lambda - k + 1)} f^{k-1} \cdot v_0 = \frac{k}{\lambda(\lambda - 1) \cdots (\lambda - k + 2)} f^{k-1} \cdot v_0 = kv_{k-1}.$$

With respect to this basis, e, f and h are represented by the same matrices as on $L(\lambda)$ ((20)), and this shows that $W \cong L(\lambda)$. □

Lemma 9.4. *If V is a finite dimensional real vector space and $a, b \in \mathfrak{gl}(V)$ with $[a, b] = b$, then b is nilpotent.*

Proof. We apply Proposition 4.15 to the solvable subalgebra $\mathbb{K}a + \mathbb{K}b \subseteq \mathfrak{gl}(V)$. Then $[a, b] = b$ implies that b is nilpotent. □

Theorem 9.5. (Classification of finite dimensional simple $\mathfrak{sl}_2(\mathbb{K})$-Modules) *Each finite dimensional simple $\mathfrak{sl}_2(\mathbb{K})$-module is isomorphic to some $L(\lambda)$, $\lambda \in \mathbb{N}_0$. For each $n \in \mathbb{N}$, there exists a simple $\mathfrak{sl}_2(\mathbb{K})$-module of dimension n which is unique up to isomorphism.*

Proof. Let (ρ, V) be a simple $\mathfrak{sl}_2(\mathbb{K})$-module. We consider the solvable subalgebra $\mathfrak{b} := \mathrm{span}\{e, h\}$. We apply Lemma 9.4 to $a := \frac{1}{2}\rho(h)$ and $b := \rho(e)$, to see that $\rho(e)$ is nilpotent. Let $d \in \mathbb{N}$ be minimal with $\rho(e)^d = 0$. Then Lemma 9.2(ii) yields

$$0 = [\rho(f), \rho(e)^d] = -d(\rho(h) - (d - 1)\mathbf{1})\rho(e)^{d-1},$$

so that each nonzero $v_0 \in \rho(e)^{d-1}(V)$ is an eigenvector of $\rho(h)$. In view of the simplicity of the module V, it is generated by v_0, and Proposition 9.3 shows that $V \cong L(\lambda)$. The remaining assertions are immediate from Example 9.1. □

Example 9.6. A particular interesting infinite dimensional representation of $\mathfrak{sl}_2(\mathbb{R})$ is the *oscillator representation*. Here we consider the space

$$\mathcal{P} = \mathbb{C}[x_1, \ldots, x_n]$$

of complex-valued polynomials on $\mathbb{R}^n$. Let $\Delta = \sum_j \frac{\partial^2}{\partial x_j^2}$ be the *Laplacian*. We put $f := \frac{1}{2}\Delta$ and $e = -\frac{1}{2}m_{r^2}$ (multiplication operator with $r^2 := \sum_j x_j^2$), and $h := E + \frac{n}{2}\mathbf{1}$, where

$E = \sum_j x_j \frac{\partial}{\partial x_j}$ is the *Euler operator*, for which a homogeneous polynomial of degree d is an eigenvector of degree d.

It is easily verified that $(h, e, f) \in \mathrm{End}(\mathcal{P})$ satisfies the commutation relations of $\mathfrak{sl}_2(\mathbb{R})$ ((14)), so that $\mathcal{P}$ is an $\mathfrak{sl}_2(\mathbb{K})$-module (Exercise 9.3). This module plays an important role in quantum mechanics of systems on $\mathbb{R}^n$ with full rotational symmetry. An important example is the spherical harmonic oscillator on $\mathbb{R}^3$, corresponding to the hydrogen atom ([St94]).

Note that the operator e is injective, and that this implies that $\mathcal{P}$ contains no non-zero $\mathfrak{sl}_2(\mathbb{R})$ submodules.

Proposition 9.7. *For a finite dimensional $\mathfrak{sl}_2(\mathbb{K})$-representation (ρ, V), the following assertions hold:*

(i) *$\rho(h)$ is diagonalizable and the set $\mathcal{P}_V$ of all eigenvalues is contained in $\mathbb{Z}$.*

(ii) *$\mathcal{P}_V = -\mathcal{P}_V$ and and $\dim V_\alpha(\rho(h)) = \dim V_{-\alpha}(\rho(h))$ for every $\alpha \in \mathcal{P}_V$.*

(iii) *If $\alpha, \alpha + 2k \in \mathcal{P}_V$ for some $k \in \mathbb{N}_0$, then $\alpha + 2j \in \mathcal{P}_V$ for $j = 0, 1, \ldots, k$* (String property)*.*

Proof. In view of Weyl's Theorem 5.26, V is a direct sum of simple submodules $V_1, \ldots, V_m$, and Theorem 9.5 implies that $V_i \cong L(\lambda_i)$ for some $\lambda_i \in \mathbb{N}_0$.

(i) and (ii) now follow immediately from the corresponding property of the modules $L(\lambda)$ (Example 9.1).

(iii) In view of (ii), we may w.l.o.g. assume that $\beta := \alpha + 2k$ satisfies $|\beta| \geq |\alpha|$. Then we pick some simple submodule $V_i \cong L(\lambda_i)$ of V such that β is an eigenvalue $\rho(h)|_{V_i}$. Then $\lambda_i - \beta \in 2\mathbb{N}_0$ and all integers $n \in \beta + 2\mathbb{Z}$ with $|n| \leq |\beta|$ are eigenvalues of $\rho(h)|_{V_i}$. This contains in particular the set of all integers of the form $\alpha + 2j$, $j = 0, 1, \ldots, k$, between α and β. $\square$

Exercises for Section 9

Exercise 9.1. We consider the 2-dimensional nonabelian complex Lie algebra $\mathfrak{b}$ in which we choose a basis (h, e) satisfying $[h, e] = e$. In the following V denotes a $\mathfrak{b}$-module and $\rho \colon \mathfrak{b} \to \mathfrak{gl}(V)$ the corresponding representation. Classify all finite-dimensional $\mathfrak{b}$-modules V for which $\rho(h)$ is diagonalizable. Hint: Proceed along the following steps:

(i) If V is generated by $v_0 \in V_\lambda(h)$, then there exists a basis $(v_0, \ldots, v_n)$ of V with

$$h \cdot v_k = (\lambda + k) v_k \quad \text{and} \quad e \cdot v_k = \begin{cases} v_{k+1} & \text{if } k < n, \\ 0 & \text{if } k = n \end{cases}.$$

We write $V(\lambda, n)$ for the $(n+1)$-dimensional $\mathfrak{b}$-module, defined by these relations.

(ii) If $k \leq n$, then $V(\lambda + k, n - k)$ is a submodule of $V(\lambda, n)$.

(iii) Each simple finite-dimensional $\mathfrak{b}$-module is isomorphic to some $V(\lambda, 0)$. Hint: Use Lie's Theorem.

(iv) For each finite-dimensional representation (ρ, V) of $\mathfrak{b}$, the operator $\rho(e)$ is nilpotent and for each n the subspaces $\ker\big(\rho(e)^n\big)$ and $\mathrm{im}\big(\rho(e)^n\big)$ are invariant under $\rho(h)$, hence $\mathfrak{b}$-submodules.

(v)* Show that each finite-dimensional representation (ρ, V) for which $\rho(h)$ is diagonalizable is a direct sum of modules of the form $V(\lambda, n)$. Hint: Derive a Jordan normal form of $\rho(e)$, adapted to the eigenspace decomposition of $\rho(h)$.

Exercise 9.2. Let $\mathcal{A} = \mathbb{K}[Z, Z^{-1}]$ be the algebra of Laurent polynomials with coefficients in the field $\mathbb{K}$. Show that every derivation of $\mathcal{A}$ is of the form $D := f\frac{d}{dZ}$ for some $f \in \mathcal{A}$.

Exercise 9.3. On the space $V = C^\infty(\mathbb{R}^n)$ we consider the operators

$$\Delta := \sum_j \frac{\partial^2}{\partial x_j^2}, \quad \text{and} \quad (Mf)(x) := \Big(\sum_j x_j^2\Big) f(x),$$

and the *Euler operator*

$$E := \sum_j x_j \frac{\partial}{\partial x_j}.$$

Verify the commutator relations

$$[E, M] = 2M, \qquad [E, \Delta] = -2\Delta \qquad \text{and} \qquad [\Delta, M] = 4E + 2n\mathbf{1}.$$

Conclude that $h := E + \frac{n}{2}\mathbf{1}$, $e := -M/2$ and $f := \Delta/2$ satisfy the $\mathfrak{sl}_2$-relations

$$[h, e] = 2e, \quad [h, f] = -2f \quad \text{and} \quad [e, f] = h.$$

10 Root Decompositions of Semisimple Lie Algebras

The technique of root decompositions is particularly fruitful for semisimple Lie algebras $\mathfrak{g}$ because, for this class of Lie algebras, over an algebraically closed field, maximal toral subalgebras turn out to be Cartan subalgebras. For complex Lie algebras we thus obtain a root space decomposition diagonalizing $\operatorname{ad}\mathfrak{h}$. In the following we write $\kappa = \kappa_{\mathfrak{g}}$ for the Cartan–Killing form of $\mathfrak{g}$.

10.1 Existence of Toral Cartan Subalgebras

We start with the root decomposition with respect to an arbitrary toral subalgebra and show later that, if $\mathbb{K}$ is algebrically closed, toral Cartan subagebras exist.

Proposition 10.1. *Let $\mathfrak{g}$ be a semisimple Lie algebra, $\mathfrak{h}$ a toral subalgebra of $\mathfrak{g}$ and $m_\lambda := \dim \mathfrak{g}_\lambda$.*

(i) $\kappa(h, h') = \sum_{\lambda \in \Delta(\mathfrak{g},\mathfrak{h})} m_\lambda \lambda(h)\lambda(h')$ *for* $h, h' \in \mathfrak{h}$.

(ii) *If $\lambda + \mu \neq 0$, then $\mathfrak{g}_\lambda$ and $\mathfrak{g}_\mu$ are orthogonal with respect to the Cartan–Killing form.*

(iii) *The Cartan–Killing form κ induces a nondegenerate pairing of $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$, i.e., for $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_{-\alpha}$,*

$$\kappa(x, \mathfrak{g}_{-\alpha}) = \{0\} \Rightarrow x = 0 \quad \text{and} \quad \kappa(\mathfrak{g}_\alpha, y) = \{0\} \Rightarrow y = 0.$$

In particular, $m_\alpha = m_{-\alpha}$ and $\kappa|_{\mathfrak{g}_0 \times \mathfrak{g}_0}$ is nondegenerate.

(iv) $\Delta(\mathfrak{g}, \mathfrak{h})$ *spans* $\mathfrak{h}^*$.

Proof. (i) The product $\operatorname{ad} h \operatorname{ad} h'$ preserves each root space $\mathfrak{g}_\lambda$, and acts on this m_λ-dimensional space by multiplication with $\lambda(h)\lambda(h')$. This implies (i).
(ii) From the invariance of κ, we obtain for $x \in \mathfrak{g}_\lambda$, $y \in \mathfrak{g}_\mu$ and $h \in \mathfrak{h}$ the relation

$$\lambda(h)\kappa(x, y) = \kappa([h, x], y) = -\kappa(x, [h, y]) = -\mu(h)\kappa(x, y)$$

and therefore $(\lambda + \mu)(h)\kappa(x, y) = 0$. This implies (ii).
(iii) follows from (ii) because the Cartan–Killing form is nondegenerate.
(iv) As a consequence of the injectivity of the adjoint representation, $\Delta(\mathfrak{g}, \mathfrak{h}) \subseteq \mathfrak{h}^*$ separates the points of $\mathfrak{h}$, and this is equivalent to (iv). $\square$

Example 10.2. With Proposition 10.1(i) we can calculate the Cartan–Killing form of the Lie algebra $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ with the root decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{j \neq k} \mathfrak{g}_{\varepsilon_j - \varepsilon_k},$$

where $\mathfrak{h}$ consists of diagonal matrices with trace 0.

For diagonal matrices $x = \sum_j x_j E_{jj}$ and $y = \sum_j y_j E_{jj}$ we obtain, taking $\sum_i x_i = \sum_j y_j = 0$ into account:

$$\begin{aligned}
\kappa(x, y) &= \sum_{j \neq k} (x_j - x_k)(y_j - y_k) = \sum_{j,k=1}^{n} (x_j - x_k)(y_j - y_k) \\
&= \sum_{j,k=1}^{n} x_j y_j - \sum_{j,k=1}^{n} x_j y_k - \sum_{j,k=1}^{n} x_k y_j + \sum_{j,k=1}^{n} x_k y_k \\
&= 2n \sum_{j}^{n} x_j y_j = 2n \operatorname{tr}(xy).
\end{aligned}$$

We likewise have $\kappa(E_{ij}, E_{k\ell}) = 0$ for $(k, \ell) \neq (j, i)$ and

$$\begin{aligned}
\kappa(E_{ij}, E_{ji}) &= \frac{1}{2}\kappa([E_{ii} - E_{jj}, E_{ij}], E_{ji}) = \frac{1}{2}\kappa(E_{ii} - E_{jj}, [E_{ij}, E_{ji}]) \\
&= \frac{1}{2}\kappa(E_{ii} - E_{jj}, E_{ii} - E_{jj}) = n \operatorname{tr}((E_{ii} - E_{jj})^2) = 2n = 2n \operatorname{tr}(E_{ij}E_{ji}).
\end{aligned}$$

In view of Proposition 10.1(i), we thus obtain

$$\kappa(x, y) = 2n \operatorname{tr}(xy) \quad \text{for} \quad x, y \in \mathfrak{sl}_n(\mathbb{K}).$$

Definition 10.3. (Jordan decomposition in semisimple Lie algebras) Suppose that $\mathbb{K}$ is algebraically closed. Let $\mathfrak{g}$ be asemisimple Lie algebra and $x \in \mathfrak{g}$. According to Theorem 5.17 and the fact that $\ker \operatorname{ad} = \mathfrak{z}(\mathfrak{g}) = \{0\}$, $\operatorname{ad}\colon \mathfrak{g} \to \operatorname{der} \mathfrak{g}$ is an isomorphism of Lie algebras. Proposition A.7 shows that for every derivation $D \in \operatorname{der}(\mathfrak{g})$ with Jordan decomposition $D = D_s + D_n$, its diagonalizable Jordan component D_s and its nilpotent Jordan components D_n are derivations of $\mathfrak{g}$. We can therefore define

$$x_s := \operatorname{ad}^{-1}\big((\operatorname{ad} x)_s\big) \quad \text{and} \quad x_n := \operatorname{ad}^{-1}\big((\operatorname{ad} x)_n\big)$$

and call $x = x_s + x_n$ the *Jordan decomposition of x in $\mathfrak{g}$*. An element $x \in \mathfrak{g}$ is called *semisimple/diagonalizable* if $x = x_s$ and *nilpotent* if $x = x_n$.

Lemma 10.4. *Let $\varphi, \psi \in \operatorname{End}(V)$ be commuting endomorphisms of the finite dimensional vector space V. If φ is nilpotent, then so is $\varphi\psi$ and, in particular, $\operatorname{tr}(\varphi\psi) = 0$.*

Proof. If $\varphi^n = 0$, then $(\varphi\psi)^n = \varphi^n\psi^n = 0$, so that $\varphi\psi$ is nilpotent, and this implies that $\mathrm{tr}(\varphi\psi) = 0$. [12] $\square$

Proposition 10.5. (Existence of toral Cartan subalgebras) *If $\mathfrak{h}$ is a maximal toral subalgebra of the finite dimensional semisimple Lie algebra $\mathfrak{g}$ over the algebraically closed field $\mathbb{K}$, then $\mathfrak{h} = \mathfrak{g}_0$. In particular $\mathfrak{h}$ is a toral Cartan subalgebra.*

Since maximal toral subalgebras exist for dimensional reasons, this result implies the existence of toral Cartan subalgebras.

Proof. We divide the proof into several steps.

Step 1: If $x \in \mathfrak{g}_0$, then so are the semisimple and nilpotent component x_s and x_n. The condition $x \in \mathfrak{g}_0$ means that $\mathrm{ad}\, x(\mathfrak{h}) = \{0\}$. Since $\mathrm{ad}\, x_s = (\mathrm{ad}\, x)_s$ and $\mathrm{ad}\, x_n = (\mathrm{ad}\, x)_n$ are polynomials in $\mathrm{ad}\, x$ without constant term (Theorem A.2), we also have $\mathrm{ad}\, x_s(\mathfrak{h}) = \{0\}$, i.e., $x_s \in \mathfrak{g}_0$, and likewise $x_n \in \mathfrak{g}_0$.

Step 2: If $x \in \mathfrak{g}_0$ is semisimple, then $x \in \mathfrak{h}$.
Since x is semisimple, $\mathrm{ad}\, x$ is diagonalizable and commutes with $\mathrm{ad}\,\mathfrak{h}$. Therefore $\mathrm{ad}\,\mathfrak{h}$ and $\mathrm{ad}\, x$ can be diagonalized simultaneously. This means that $\mathfrak{h} + \mathbb{K}x$ is a toral subalgebra. Now the maximality of $\mathfrak{h}$ implies that $x \in \mathfrak{h}$.

Step 3: $\mathfrak{g}_0$ is nilpotent.
Let $x \in \mathfrak{g}_0$. Then $x_s \in \mathfrak{g}_0$ by Step 1 and $x_s \in \mathfrak{h}$ by Step 2, so that $\mathrm{ad}_{\mathfrak{g}_0} x_s = 0$. Therefore $\mathrm{ad}_{\mathfrak{g}_0} x = \mathrm{ad}_{\mathfrak{g}_0} x_n$ is nilpotent because it is a restriction of a nilpotent endomorphism of $\mathfrak{g}$. Now Engel's Theorem implies that $\mathfrak{g}_0$ is nilpotent.

Step 4: $\mathfrak{g}_0$ is abelian.
In view of Step 3 and Corollary 4.12, $\mathfrak{g}$ is a nilpotent module of the Lie algebra $[\mathfrak{g}_0, \mathfrak{g}_0]$, so that $\kappa(\mathfrak{g}_0, [\mathfrak{g}_0, \mathfrak{g}_0]) = \{0\}$ follows from Exercise 10.1. Since κ is non-degenerate on $\mathfrak{g}_0$ (Proposition 10.1(iii)), it follows that $[\mathfrak{g}_0, \mathfrak{g}_0] = \{0\}$.

Step 5: $\mathfrak{g}_0 = \mathfrak{h}$.
If this is not the case, then Steps 1 and 2 imply that $\mathfrak{g}_0 \setminus \mathfrak{h}$ contains a nilpotent element x. In view of Step 4 and Lemma 10.4, we then have $\kappa(x, \mathfrak{g}_0) = \{0\}$ so that $x = 0$ follows from Proposition 10.1(iii). $\square$

Definition 10.6. Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a toral Cartan subalgebra of the semisimple Lie algebra $\mathfrak{g}$. Since the Cartan–Killing form is nondegenerate on $\mathfrak{h}$, we can assign to every root α a uniquely determined element $t_\alpha \in \mathfrak{h}$ via the equation

$$\kappa(h, t_\alpha) = \alpha(h). \tag{23}$$

Further, we can introduce a bilinear form on $\mathfrak{h}^*$ via

$$(\alpha, \beta) := \kappa(t_\alpha, t_\beta) = \alpha(t_\beta) = \beta(t_\alpha). \tag{24}$$

Lemma 10.7. *For $\alpha \in \Delta(\mathfrak{g}, \mathfrak{h})$,*

$$[x, y] = \kappa(x, y)t_\alpha \quad \text{for} \quad x \in \mathfrak{g}_\alpha,\ y \in \mathfrak{g}_{-\alpha}. \tag{25}$$

Proof. Both sides of the equation are in $\mathfrak{h}$ (Lemma 8.10(ii)), hence (25) follows from

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

since the Cartan–Killing form is nondegenerate on $\mathfrak{h}$ by Proposition 10.1. $\square$

[12] The most direct way to argue that $\mathrm{tr}(\varphi) = 0$ for a nilpotent endomorphism is that its characteristic polynomial is of the form $\det(\varphi - t\mathbf{1}) = (-1)^n t^n$, so that the coefficient $(-1)^{n-1}\mathrm{tr}(\varphi)$ of t^{n-1} vanishes.

10.2 $\mathfrak{sl}_2$-Triples in Semisimple Lie Algebras

The following theorem is the starting point of a complete classification of simple Lie algebras. It emphasizes the special role of the Lie algebra $\mathfrak{sl}_2(\mathbb{K})$.

Definition 10.8. Let $\mathfrak{g}$ be a Lie algebra. A triple (h, e, f) of elements of $\mathfrak{g}$ is called an $\mathfrak{sl}_2$-*triple* if

$$[h, e] = 2e, \quad [h, f] = -2f \quad \text{and} \quad [e, f] = h.$$

Then h, e, f are eigenvectors of $\operatorname{ad} h$ for different eigenvalues, so that h, e, f are linearly independent and $\operatorname{span}\{h, e, f\} \cong \mathfrak{sl}_2(\mathbb{K})$.

Theorem 10.9. ($\mathfrak{sl}_2$-Theorem) *Let $\mathfrak{g}$ be a semisimple Lie algebra, $\mathfrak{h} \subseteq \mathfrak{g}$ a toral Cartan subalgebra and $\alpha \in \Delta(\mathfrak{g}, \mathfrak{h})$.*

(i) *For every root $\alpha \in \Delta(\mathfrak{g}, \mathfrak{h})$, we have $(\alpha, \alpha) \neq 0$ and there are elements $e_\alpha \in \mathfrak{g}_\alpha$, $f_\alpha \in \mathfrak{g}_{-\alpha}$ and $h_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ such that $(h_\alpha, e_\alpha, f_\alpha)$ is an $\mathfrak{sl}_2$-triple.*

(ii) *$m_\alpha = \dim \mathfrak{g}_\alpha = \dim([\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]) = 1$ and $\mathbb{Z}\alpha \cap \Delta = \{\pm\alpha\}$.*

(iii) *$\alpha(h_\alpha) = 2$.*

In the following we write

$$\mathfrak{g}(\alpha) := \operatorname{span}\{h_\alpha, e_\alpha, f_\alpha\} \subseteq \mathfrak{h} \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$$

for the $\mathfrak{sl}_2$-subalgebra defined by the triple $(h_\alpha, e_\alpha, f_\alpha)$. In view of (ii), it does not depend on the choice of $e_\alpha \in \mathfrak{g}_\alpha$ and $f_\alpha \in \mathfrak{g}_{-\alpha}$.

Proof. (i),(iii) From $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}) \neq \{0\}$ and Lemma 10.7 we obtain elements $e_{\pm\alpha} \in \mathfrak{g}_{\pm\alpha}$ with $[e_\alpha, e_{-\alpha}] = t_\alpha$. To see that $(\alpha, \alpha) = \alpha(t_\alpha)$ is nonzero, let us assume the contrary and consider some $\beta \in \Delta$. Then the subspace

$$V := \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_{\beta + k\alpha}$$

of $\mathfrak{g}$ is invariant under $\operatorname{ad}(e_{\pm\alpha})$, so that

$$0 = \operatorname{tr}([\operatorname{ad} e_\alpha|_V, \operatorname{ad} e_{-\alpha}|_V]) = \operatorname{tr}(\operatorname{ad} t_\alpha|_V) = \sum_{k \in \mathbb{Z}} (\beta + k\alpha)(t_\alpha) \cdot m_{\beta + k\alpha} = \beta(t_\alpha) \cdot \sum_{k \in \mathbb{Z}} m_{\beta + k\alpha}.$$

Since $\sum_{k \in \mathbb{Z}} m_{\beta + k\alpha} \geq m_\beta > 0$, we get $\beta(t_\alpha) = 0$ for all roots β. But since the roots span $\mathfrak{h}^*$ (Proposition 10.1(iv)), this contradicts $t_\alpha \neq 0$. We conclude that $(\alpha, \alpha) = \alpha(t_\alpha) \neq 0$. The element $h_\alpha := 2\frac{t_\alpha}{\alpha(t_\alpha)}$ satisfies (iii). From Proposition 10.1(iii) we get an element $f_\alpha \in \mathfrak{g}_{-\alpha}$ with $\kappa(e_\alpha, f_\alpha) = \frac{2}{\alpha(t_\alpha)}$, so that Lemma 10.7 implies that $[e_\alpha, f_\alpha] = h_\alpha$. Now (i) follow from (iii).

(ii) We consider the subspace

$$V := \mathbb{K} f_\alpha + \mathfrak{h} + \sum_{n=1}^{\infty} \mathfrak{g}_{n\alpha}$$

of $\mathfrak{g}$. One verifies easily that this subspace is invariant under $\operatorname{ad}\big(\mathfrak{g}(\alpha)\big)$ because it is invariant under $\operatorname{ad} \mathfrak{h}$, $[f_\alpha, \mathfrak{h}] = \mathbb{K} f_\alpha$, and $[e_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\beta+\alpha}$. According to Lemma D.10, we therefore have

$$\dim V_m(\operatorname{ad} h_\alpha) = \dim V_{-m}(\operatorname{ad} h_\alpha)$$

for all $m \in \mathbb{Z}$. This leads to

$$\dim \mathfrak{g}_\alpha = \dim V_2(\operatorname{ad} h_\alpha) = \dim V_{-2}(\operatorname{ad} h_\alpha) = 1$$

and $\dim \mathfrak{g}_{n\alpha} = 0$ for $n > 1$. Now (ii) follows from (i). $\square$

10.3 Coroots and Root Strings

In the following $\mathfrak{g}$ is a semisimple Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ is a toral Cartan subalgebra. Then we obtain a root decomposition of $\mathfrak{g}$ with respect to $\mathfrak{h}$. For brevity we put $\Delta := \Delta(\mathfrak{g}, \mathfrak{h})$.

Definition 10.10. (Coroots) For a root $\alpha \in \Delta$, we have already seen that $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is a one-dimensional subspace of $\mathfrak{h}$ on which α does not vanish. Hence there is a unique element

$$\check{\alpha} = h_\alpha \in [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \quad \text{with} \quad \alpha(\check{\alpha}) = 2,$$

called the *coroot* corresponding to α (cf. Theorem 10.9).

Lemma 10.11. (Root String Lemma) *Let $\alpha, \beta \in \Delta$.*

(i) *For $\beta \in \Delta \setminus \{\pm\alpha\}$ the set $\{k \in \mathbb{Z} \colon \beta + k\alpha \in \Delta\}$ is an interval in $\mathbb{Z}$. If it is of the form $[-p, q] \cap \mathbb{Z}$ with $p, q \in \mathbb{Z}$, then $p - q = \beta(\check{\alpha})$. In particular, $\beta(\check{\alpha}) \in \mathbb{Z}$.*

(ii) *If $\beta(\check{\alpha}) < 0$, then $\beta + \alpha \in \Delta$ and if $\beta(\check{\alpha}) > 0$, then $\beta - \alpha \in \Delta$.*

(iii) *If $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \{0\}$, then $\beta(\check{\alpha}) \geq 0$.*

(iv) *If $\alpha + \beta \neq 0$, then $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$.*

Proof. (i) We consider the subspace $V := \sum_{k \in \mathbb{Z}} \mathfrak{g}_{\beta + k\alpha}$. Note that $\beta \neq \{\pm\alpha\}$ implies that 0 is not contained in $\beta + \mathbb{Z}\alpha$ (Theorem 10.9(ii)). From $[\mathfrak{g}_\gamma, \mathfrak{g}_\delta] \subseteq \mathfrak{g}_{\gamma+\delta}$, we derive that V is a $\mathfrak{g}(\alpha)$-submodule of $\mathfrak{g}$ (cf. Theorem 10.9). The eigenvalues of $\check{\alpha} = h_\alpha$ on V are given by

$$\mathcal{P}_V := \{(\beta + k\alpha)(\check{\alpha}) \colon \beta + k\alpha \in \Delta\} = \beta(\check{\alpha}) + 2\{k \colon \beta + k\alpha \in \Delta\}.$$

Hence the string property of $\mathfrak{sl}_2$-modules (Proposition 9.7) implies the string property of the root system.

Next we note that $\beta \in \Delta$ leads to $p \geq 0$. In view of Proposition 9.7, we have $\mathcal{P}_V = -\mathcal{P}_V$. Therefore

$$\beta(\check{\alpha}) - 2p = (\beta - p\alpha)(\check{\alpha}) = -(\beta + q\alpha)(\check{\alpha}) = -\beta(\check{\alpha}) - 2q.$$

(ii) If $\beta(\check{\alpha}) < 0$, then (i) leads to $q > 0$ and hence to $\beta + \alpha \in \Delta$. The second assertion follows similarly.

(iii) As all multiplicities of the eigenvalues of $\check{\alpha}$ on V are 1 (Theorem 10.9(ii)), the $\mathfrak{sl}_2(\mathbb{K})$-module V is simple and isomorphic to $L(\beta(\check{\alpha}) + 2q)$ (apply Proposition 9.3 to a nonzero element of $\mathfrak{g}_{\beta+q\alpha}$). This immediately shows that

$$[\mathfrak{g}_\alpha, \mathfrak{g}_{\beta+k\alpha}] = \mathfrak{g}_{\beta+(k+1)\alpha} \quad \text{for} \quad k = -p, -p+1, \ldots, q-1. \tag{26}$$

If $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \{0\}$, then $V := \sum_{k \leq 0} \mathfrak{g}_{\beta+k\alpha}$ is invariant under $\mathfrak{g}(\alpha) \cong \mathfrak{sl}_2(\mathbb{K})$ and $\beta(\check{\alpha})$ is the maximal eigenvalue of $\mathrm{ad}(\check{\alpha})$ on V. Hence Proposition 9.7 shows that $\beta(\check{\alpha}) \geq 0$.

(iv) We may assume that $\beta + \alpha \in \Delta$ (otherwise $\mathfrak{g}_{\alpha+\beta} = \{0\}$), so that $q \geq 1$. Then (iv) follows from (26). $\square$

Lemma 10.12. *The subspace $\mathfrak{h}_\mathbb{Q} := \mathrm{span}_\mathbb{Q}\{\check{\alpha} \colon \alpha \in \Delta\}$ of $\mathfrak{h}$ has the following properties:*

(i) *$\alpha(\mathfrak{h}_\mathbb{Q}) \subseteq \mathbb{Q}$ for every $\alpha \in \Delta$.*

(ii) *κ restricts to a $\mathbb{Q}$-valued positive definite form on $\mathfrak{h}_\mathbb{Q}$.*

(iii) *$\mathrm{span}_\mathbb{K}\, \mathfrak{h}_\mathbb{Q} = \mathfrak{h}$.*

(iv) *The $\mathbb{Q}$-valued symmetric bilinear form on $\mathfrak{h}_{\mathbb{Q}}^*$, specified by*

$$(\alpha, \beta) := \kappa(t_\alpha, t_\beta) = \alpha(t_\beta)$$

is positive definite.

Proof. (i) This follows from $\beta(\check{\alpha}) \in \mathbb{Z}$ for $\alpha, \beta \in \Delta$, which is a consequence of the Root String Lemma 10.11.

(ii) That κ is $\mathbb{Q}$-valued on $\mathfrak{h}_{\mathbb{Q}}$ follows from

$$\kappa(\check{\alpha}, \check{\beta}) = \sum_{\gamma \in \Delta} \gamma(\check{\alpha})\gamma(\check{\beta}) \in \mathbb{Z} \quad \text{for} \quad \alpha, \beta \in \Delta.$$

For $h \in \mathfrak{h}_{\mathbb{Q}}$, we have $\alpha(h) \in \mathbb{Q}$ for every $\alpha \in \Delta$ by (i). This implies that

$$\kappa(h, h) = \sum_{\alpha \in \Delta} \alpha(h)^2 \geq 0.$$

If $\kappa(h, h) = 0$, then $\alpha(h) = 0$ for every $\alpha \in \Delta$, and therefore $h \in \mathfrak{z}(\mathfrak{g}) = \{0\}$ (Lemma 8.2).

(iii) The non-degenerate form κ on $\mathfrak{h}$ defines a linear isomorphism

$$\Gamma\colon \mathfrak{h} \to \mathfrak{h}^*, \quad \Gamma(x) := \kappa(x, \cdot).$$

It satisfies $\Gamma(t_\alpha) = \alpha$. From $\check{\alpha} = \frac{2t_\alpha}{(\alpha,\alpha)}$, we further derive that

$$\kappa(\check{\alpha}, \check{\alpha}) = \frac{4}{(\alpha, \alpha)^2}\kappa(t_\alpha, t_\alpha) = \frac{4}{(\alpha, \alpha)^2}(\alpha, \alpha) = \frac{4}{(\alpha, \alpha)},$$

which implies that $(\alpha, \alpha) \in \mathbb{Q}$. Therefore $\mathbb{Q}\check{\alpha} = \mathbb{Q}t_\alpha$, and thus

$$\Gamma(\mathfrak{h}_{\mathbb{Q}}) = \operatorname{span}_{\mathbb{Q}} \Delta.$$

This further leads to

$$\Gamma(\operatorname{span}_{\mathbb{K}} \mathfrak{h}_{\mathbb{Q}}) = \operatorname{span}_{\mathbb{K}} \Delta = \mathfrak{h}^*.$$

(Proposition 10.1), and since Γ is a linear isomorphism $\mathfrak{h}_{\mathbb{Q}}$ spans $\mathfrak{h}$.

(iv) As κ is $\mathbb{Q}$-valued on $\mathfrak{h}_{\mathbb{Q}}$, we also obtain a linear isomorphism

$$\Gamma_{\mathbb{Q}}\colon \mathfrak{h}_{\mathbb{Q}} \to \mathfrak{h}_{\mathbb{Q}}^*, \quad \Gamma(x) = \kappa(x, \cdot)$$

with $\Gamma(t_\alpha) = \alpha$. The natural scalar product $(\cdot, \cdot)$ on $\mathfrak{h}_{\mathbb{Q}}^*$ is now defined in such a way that $(\Gamma(x), \Gamma(y)) = \kappa(x, y)$, so that (ii) implies that it is positive definite. □

Exercises for Section 10

Exercise 10.1. Let (ρ, V) be a finite dimensional representation of the Lie algebra $\mathfrak{g}$ and $\kappa_\rho(x, y) := \operatorname{tr}(\rho(x)\rho(y))$ be the corresponding invariant symmetric bilinear form. Show that: If $\mathfrak{n} \trianglelefteq \mathfrak{g}$ is an ideal and V is a nilpotent $\mathfrak{n}$-module, then $\kappa_\rho(\mathfrak{n}, \mathfrak{g}) = \{0\}$. Hint: Consider a maximal flag $\mathcal{F}$ of submodules and show that $\rho(\mathfrak{n}) \subseteq \mathfrak{g}_n(\mathcal{F})$.

11 Abstract Root Systems and their Weyl Groups

In the previous section we proved a number of results on the root systems $\Delta(\mathfrak{g}, \mathfrak{h})$ associated with a given toral Cartan subalgebra $\mathfrak{h}$ of a semisimple Lie algebra $\mathfrak{h}$. In this section we distill some of the properties of these root systems into the concept of an abstract (finite) root system in a euclidean vector space and show how to derive further properties using this abstract level.

11.1 Abstract Root Systems

Definition 11.1. Let E be a *euclidean space*, i.e., a finite dimensional real vector space with an inner product $(\cdot, \cdot)$, i.e., a positive definite symmetric bilinear form. A *reflection* in E is a linear map σ for which there exists a non-zero vector α with $\sigma = \sigma_\alpha$, where

$$\sigma_\alpha(\beta) = \beta - 2\frac{(\beta, \alpha)}{(\alpha, \alpha)}\alpha \quad \text{for} \quad \beta \in E.$$

Note that

$$\sigma_\alpha(\alpha) = -\alpha \quad \text{and} \quad \sigma_\alpha(\beta) = \beta \quad \text{for} \quad \beta \perp \alpha,$$

i.e., σ_α is the orthogonal reflection in the hyperplane $\alpha^\perp$.

Definition 11.2. Let E be a euclidean space and $\Delta \subseteq E \setminus \{0\}$ be a finite subset which spans E. Then Δ is called a *reduced root system* if it satisfies the following conditions

(R1) $\Delta \cap \mathbb{R}\alpha = \{\pm\alpha\}$ for all $\alpha \in \Delta$. In particular, $\Delta = -\Delta$.

(R2) $\sigma_\alpha(\Delta) \subseteq \Delta$ for all $\alpha \in \Delta$,

(R3) For $\alpha \in \Delta$ the *coroot*

$$\check{\alpha} := \frac{2\alpha}{(\alpha, \alpha)}$$

satisfies $(\beta, \check{\alpha}) \in \mathbb{Z}$ for all $\beta \in \Delta$. Note that the reflection σ_α can be expressed by the coroot as

$$\sigma_\alpha(\beta) = \beta - (\beta, \check{\alpha})\alpha. \tag{27}$$

It is called *root system* if it only satisfies (R2) and (R3). If Δ is a root system, then we call the group $\mathcal{W} = \mathcal{W}(\Delta)$ generated by the reflections $(\sigma_\alpha)_{\alpha\in\Delta}$, the *Weyl group* of the root system.

Remark 11.3. If Δ is a (nonreduced) root system and $\alpha, c\alpha \in \Delta$ for some $c > 1$, then

$$c \cdot (\alpha, \check{\alpha}) = 2c \in \mathbb{Z}$$

implies that $c \in \frac{1}{2}\mathbb{Z}$. Further, $(c\alpha)\check{} = \frac{1}{c}\check{\alpha}$ leads to $(\alpha, c^{-1}\check{\alpha}) = \frac{2}{c} \in \mathbb{Z}$, so that $c = 2$. We therefore get

$$\Delta \cap \mathbb{R}\alpha = \{\pm\alpha, \pm 2\alpha\}.$$

In the root decomposition of Lie algebras, the root system is a finite subset of the dual $\mathfrak{h}^*$ of a toral Cartan subalgebra, which is a vector space over $\mathbb{K}$. To find a euclidean space E containing Δ, we cannot simply take the $\mathbb{R}$-span of Δ because $\mathbb{R}$ need not be a subfield of $\mathbb{K}$. However, the rational vector space $\mathfrak{h}^*_{\mathbb{Q}} = \mathrm{Hom}_{\mathbb{Q}}(\mathfrak{h}_{\mathbb{Q}}, \mathbb{Q})$, which also contains Δ, is contained in the real space $\mathrm{Hom}_{\mathbb{Q}}(\mathfrak{h}_{\mathbb{Q}}, \mathbb{R})$ and we shall see below that this space carries a natural inner product.

Proposition 11.4. (The root systems $\Delta(\mathfrak{g}, \mathfrak{h})$) *Let $\mathfrak{g}$ be a semisimple Lie algebra, $\mathfrak{h} \subseteq \mathfrak{g}$ a toral Cartan subalgebra and $\mathfrak{h}_{\mathbb{Q}} := \operatorname{span} \check{\Delta}$. Then the scalar product on $\operatorname{span}_{\mathbb{Q}} \Delta$ extends to an inner product on*

$$E := \mathrm{Hom}_{\mathbb{Q}}(\mathfrak{h}_{\mathbb{Q}}, \mathbb{R})$$

defining the structure of a euclidean space on E. Then $\Delta(\mathfrak{g}, \mathfrak{h}) \subseteq E$ is a reduced root system.

Proof. We identify Δ with a subspace of $\mathfrak{h}_{\mathbb{Q}}^*$ (Lemma 10.12(i)). First we show that $E = \operatorname{span}_{\mathbb{R}} \Delta$. Let $b_1, \ldots, b_r \in \mathfrak{h}_{\mathbb{Q}}$ be a $\mathbb{Q}$-basis. Then

$$\Phi\colon E \to \mathbb{R}^r, \quad \Phi(\gamma) = (\gamma(b_1), \ldots, \gamma(b_r))$$

is a linear isomorphism and $\Phi(\mathfrak{h}_{\mathbb{Q}}^*) = \mathbb{Q}^r$. Since Δ spans $\mathfrak{h}_{\mathbb{Q}}^*$ (it separates the points of $\mathfrak{h}_{\mathbb{Q}}$ by Lemma 8.2), it follows that $\Phi(\Delta)$ spans $\Phi(E)$, i.e., Δ spans E.

If a real linear combination $\sum_{\alpha \in \Delta} x_\alpha \alpha$ vanishes, then $\sum_{\alpha,\beta\in\Delta} x_\alpha y_\beta (\alpha, \beta)$ vanishes for any other linear combintation $\sum_\beta y_\beta \beta$. This implies that

$$\Big(\sum_\alpha x_\alpha \alpha, \sum_\beta y_\beta \beta\Big) := \sum_{\alpha,\beta\in\Delta} x_\alpha y_\alpha (\alpha, \beta) \tag{28}$$

is a well-defined symmetric bilinear form $E \times E \to \mathbb{R}$.

If $B \subseteq \Delta$ is an $\mathbb{R}$-basis of E, then it also is a $\mathbb{Q}$-basis of $\mathfrak{h}_{\mathbb{Q}}^*$, then Lemma 10.12(ii) implies that the matrix $\big((\alpha, \beta)\big)_{\alpha,\beta\in B}$ is positive definite as a real matrix. [13] As

$$\Big(\sum_{\alpha\in B} x_\alpha \alpha, \sum_{\beta\in B} y_\beta \beta\Big) := \sum_{\alpha,\beta\in\Delta} x_\alpha y_\alpha (\alpha, \beta),$$

it now follows that the scalar product on E is positive definite. We now verify (R1)-(R3).

(R3) For $\alpha, \beta \in \Delta$, Lemma 10.11 yields $(\beta, \check{\alpha}) = p - q \in \mathbb{Z}$, where $p, q \in \mathbb{N}_0$ are such that $\beta - p\alpha$ and $\beta + q\alpha$ are the ends of the α-string through β.

(R2) The same lemma also implies $\beta - (\beta, \check{\alpha})\alpha \in \Delta$ because $-p \leq -(\beta, \check{\alpha}) = q - p \leq q$.

(R1) Since (R2) and (R3) are satisfies, Remark 11.3 implies that it suffices to show that $\mathbb{Z}\alpha \cap \Delta = \{\pm\alpha\}$, but this follows from the $\mathfrak{sl}_2$-Theorem 10.9.

This shows that Δ is a reduced root system. $\square$

Definition 11.5. (Weyl group) Let $\mathfrak{h}$ be a toral Cartan subalgebra of the semisimple Lie algebra $\mathfrak{g}$. In view of Proposition 11.4, this data defines a *Weyl group*

$$\mathcal{W}(\mathfrak{g}, \mathfrak{h}) := \mathcal{W}(\Delta(\mathfrak{g}, \mathfrak{h})).$$

Examples 11.6. (Weyl groups of the classical root systems) We now discuss the four series of root systems A_n, B_n, C_n and D_n in the euclidean space $\mathbb{R}^n$, where $(x, y) = \sum_j x_j y_j$ is the canonical inner product. We have already seen in Examples 8.7, 8.8 and 8.9 how these root systems arise from Lie algebras. However, the normalization of the scalar product induced from the Cartan–Killing form may be different, as we have seen in Example 10.2.

(a) Let us write $(\varepsilon_j)_{1\leq j\leq n}$ for the canonical basis of $\mathbb{R}^n$. We consider in the euclidean space $E := \{x \in \mathbb{R}^n \colon \sum_i x_i = 0\}$ the root system

$$A_{n-1} := \{\varepsilon_j - \varepsilon_k \colon j \neq k\} = \Delta(\mathfrak{sl}_n(\mathbb{K}), \mathfrak{h})$$

(cf. Example 8.7). For the root $\alpha_{jk} := \varepsilon_j - \varepsilon_k$, we then have $\check{\alpha}_{jk} = \alpha_{jk}$, so that

$$\sigma_{\alpha_{jk}}(\varepsilon_\ell) = \varepsilon_\ell - (\delta_{\ell,j} - \delta_{\ell,k})(\varepsilon_j - \varepsilon_k) = \begin{cases} \varepsilon_\ell & \text{for } \ell \neq j, k \\ \varepsilon_k & \text{for } \ell = j \\ \varepsilon_j & \text{for } \ell = k. \end{cases}$$

[13] Note that the Hurwitz criterion for positive definiteness of a matrix in terms of its minors can be applied to any subfield of $\mathbb{R}$ containing all entries of the matrix.

A more direct argument works as follows: If A is positive definite over $\mathbb{Q}$, then the density of $\mathbb{Q}$ in $\mathbb{R}$ implies $x^\top Ax \geq 0$ for every $x \in \mathbb{R}^n$. If $x^\top Ax = 0$, then $y^\top Ax = 0$ for every $y \in \mathbb{R}^n$ by Cauchy–Schwarz. Now $Ax = 0$, but $\det_{\mathbb{R}}(A) = \det_{\mathbb{Q}}(A) \neq 0$ yields $x = 0$.

Therefore $\sigma_{\alpha_{jk}}$ acts on the orthonormal basis $(\varepsilon_\ell)_{1\leq\ell\leq n}$ as the transposition $(jk) \in S_n$. Since these transpositions generate the symmetric group, it follows that

$$\mathcal{W}(A_{n-1}) \cong S_n.$$

It follows in particular that the action of this group preserves A_{n-1}, so that (R1-3) are satisfied, i.e., A_{n-1} is a reduced root system in the sense of Definition 11.2.

(b) For the root system

$$B_n := \{\pm\varepsilon_j, \pm\varepsilon_j \pm \varepsilon_k : j,k = 1,\ldots,n, j \neq k\} \subseteq \mathbb{R}^n$$

of the Lie algebras $\mathfrak{o}_{n+1,n}(\mathbb{K})$ (Example 8.8(b)), we obtain from $\check{\varepsilon}_j = 2\varepsilon_j$ the relation

$$\sigma_{\varepsilon_j}(\varepsilon_\ell) = \varepsilon_\ell - 2\delta_{\ell,j}\varepsilon_j = \begin{cases} \varepsilon_\ell & \text{for } \ell \neq j \\ -\varepsilon_\ell & \text{for } \ell = j. \end{cases}$$

and from $(\varepsilon_j + \varepsilon_j)\check{} = \varepsilon_j + \varepsilon_k$ further

$$\sigma_{\varepsilon_j+\varepsilon_k}(\varepsilon_\ell) = \varepsilon_\ell - (\delta_{\ell,j} + \delta_{\ell,k})(\varepsilon_j + \varepsilon_k) = \begin{cases} \varepsilon_\ell & \text{for } \ell \neq j,k \\ -\varepsilon_k & \text{for } \ell = j \\ -\varepsilon_j & \text{for } \ell = k. \end{cases}$$

In particular, $\sigma_{\varepsilon_j+\varepsilon_k} = \sigma_{\varepsilon_j}\sigma_{\varepsilon_k}$. We conclude that

$$\mathcal{W}(B_n) \cong \{\pm 1\}^n \rtimes S_n$$

is the group of *signed permutations*.

Again, the description of the group generated by the reflections implies that (R1-3) are satisfied for B_n.

(c) For the root system

$$C_n := \{\pm 2\varepsilon_j, \pm\varepsilon_j \pm \varepsilon_k : j,k = 1,\ldots,n, j \neq k\} \subseteq \mathbb{R}^n$$

of the Lie algebras $\mathfrak{sp}_{2n}(\mathbb{K})$ (Example 8.9), we obtain from $(2\varepsilon_j)\check{} = \varepsilon_j$ the relation $\sigma_{2\varepsilon_j} = \sigma_{\varepsilon_j}$. Therefore

$$\mathcal{W}(C_n) \cong \mathcal{W}(B_n) \cong \{\pm 1\}^n \rtimes S_n$$

is also the group of signed permutations.

(d) For the root system

$$D_n := \{\pm\varepsilon_j \pm \varepsilon_k : j,k = 1,\ldots,n, j \neq k\} \subseteq \mathbb{R}^n$$

of the Lie algebras $\mathfrak{o}_{n,n}(\mathbb{K})$ (Example 8.8(a)), the calculations under (b) show that the Weyl group contains all permutations and all signed permutations with an even number of sign changes:

$$\mathcal{W}(D_n) \cong \{(\eta_1,\ldots,\eta_n) \in \{\pm 1\}^n \colon |\{j : \eta_j = -1\}| \in 2\mathbb{N}_0\} \rtimes S_n.$$

This is a subgroup of index 2 in $\mathcal{W}(B_n) \cong \{\pm 1\}^n \rtimes S_n$. It is the kernel of the homomorphism

$$\gamma \colon \mathcal{W}(B_n) \to \{\pm 1\}, \quad \gamma((\eta_j)_{1\leq j\leq n}, \sigma) := \prod_{j=1}^n \eta_j.$$

Remark 11.7. The angle $\theta \in [0, \pi]$ between α and β is defined by the identity

$$\|\alpha\| \, \|\beta\| \cos\theta = (\alpha, \beta),$$

where $\|\alpha\| = \sqrt{(\alpha, \alpha)}$ is the norm of the euclidean space E. We have

$$(\beta, \check{\alpha}) = 2\frac{(\beta, \alpha)}{(\alpha, \alpha)} = 2\frac{\|\beta\|}{\|\alpha\|}\cos\theta \quad \text{and} \quad (\alpha, \check{\beta})(\beta, \check{\alpha}) = 4\cos^2\theta.$$

Hence $(\beta, \check{\alpha}), (\alpha, \check{\beta}) \in \mathbb{Z}$ leads to $4\cos^2\theta \in \mathbb{Z}$, and there are only the following possibilities for $\|\alpha\| \leq \|\beta\|$ and $\beta \notin \mathbb{R}\alpha$:

$(\alpha, \check{\beta})$	0	1	-1	1	-1	1	-1
$(\beta, \check{\alpha})$	0	1	-1	2	-2	3	-3
θ	$\frac{\pi}{2}$	$\frac{\pi}{3}$	$\frac{2\pi}{3}$	$\frac{\pi}{4}$	$\frac{3\pi}{4}$	$\frac{\pi}{6}$	$\frac{5\pi}{6}$
$\frac{\|\beta\|^2}{\|\alpha\|^2}$	arb.	1	1	2	2	3	3

Note that exchanging β by $-\beta$ replaces θ by $\pi - \theta$ and that $(\alpha, \check{\beta}) > 0$ is equivalent to $0 < \theta < \frac{\pi}{2}$.

Lemma 11.8. *Let $\Phi \subseteq E$ be a finite generating subset of the euclidean space E which is invariant under all reflections $(\sigma_\alpha)_{\alpha \in \Phi}$. If $\sigma \in \mathrm{GL}(E)$ fixes a hyperplane H pointwise, $\sigma(\alpha) = -\alpha$ for some $\alpha \in \Phi$ and $\sigma(\Phi) = \Phi$, then $\sigma = \sigma_\alpha$.*

Proof. Let $\tau := \sigma\sigma_\alpha$. Then $\tau(\Phi) = \Phi$ because both factors have this property. Then $\tau(\alpha) = \alpha$ and the linear automorphism $\widetilde{\tau} \in \mathrm{GL}(E/\mathbb{R}\alpha)$ induced by τ coincides with the linear automorphism $\widetilde{\sigma} \in \mathrm{GL}(E/\mathbb{R}\alpha)$ induced by σ. As $E = H + \mathbb{R}\alpha$, we have $\widetilde{\sigma} = \mathrm{id}$ and therefore $\widetilde{\tau} = \mathrm{id}$. We conclude that $\tau - \mathbf{1}$ is nilpotent, i.e., τ is unipotent. The τ-invariance of the finite set Φ shows that there has to be a power τ^k which keeps Φ pointwise fixed. But Φ spans E, so that $\tau^k = \mathrm{id}$. As τ is unipotent, it follows that $\tau = \mathrm{id}$ (Exercise 11.1). □

Proposition 11.9. *Let $\Delta \subseteq E$ be a root system with Weyl group $\mathcal{W}$. If $\tau \in \mathrm{GL}(E)$ satisfies $\tau(\Delta) = \Delta$, then*

(i) $\tau\sigma_\alpha\tau^{-1} = \sigma_{\tau\alpha}$ *for all* $\alpha \in \Delta$.

(ii) $(\beta, \check{\alpha}) = \big(\tau(\beta), \tau(\alpha)\check{}\big)$ *for all* $\alpha, \beta \in \Delta$.

Proof. (i) First we note that $\sigma := \tau\sigma_\alpha\tau^{-1}$ satisfies $\sigma(\Delta) = \Delta$ because all factors have this property. Further,σ keeps the hyperplane $\tau(\alpha^\perp)$ pointwise fixed, and it maps $\tau\alpha$ to $-\tau\alpha$. Hence Lemma 11.8 shows that $\tau\sigma_\alpha\tau^{-1} = \sigma_{\tau\alpha}$.
(ii) In view of (i), this follows by comparison of the formulas

$$\sigma_{\tau(\alpha)}(\tau(\beta)) = \tau\sigma_\alpha\tau^{-1}(\tau(\beta)) = \tau(\sigma_\alpha(\beta)) = \tau(\beta - (\beta, \check{\alpha})\alpha) = \tau(\beta) - (\beta, \check{\alpha})\tau(\alpha)$$

and

$$\sigma_{\tau(\alpha)}(\tau(\beta)) = \tau(\beta) - \big(\tau(\beta), \tau(\alpha)\check{}\big)\tau(\alpha).$$ □

Lemma 11.10. *Let Δ be a root system, and suppose that $\alpha, \beta \in \Delta$ are not proportional.*

- *If $(\alpha, \beta) > 0$, then $\alpha - \beta \in \Delta$.*
- *If $(\alpha, \beta) < 0$, then $\alpha + \beta \in \Delta$.*

Proof. Exchanging β by $-\beta$, we see that it suffices to verify the first assertion.

Since (α, β) is positive if and only if $(\alpha, \check{\beta})$ is positive, Remark 11.7 shows that we have $(\alpha, \check{\beta}) = 1$ or $(\beta, \check{\alpha}) = 1$. If $(\alpha, \check{\beta}) = 1$, then $\alpha - \beta = \sigma_\beta(\alpha) \in \Delta$. Similarly, for $(\beta, \check{\alpha}) = 1$, we have $\beta - \alpha = \sigma_\alpha(\beta) \in \Delta$, hence $\alpha - \beta \in \Delta$. □

The following lemma asserts that in abstract root systems root strings are "unbroken".

Lemma 11.11. (Root String Lemma) *If $\alpha, \beta \in \Delta$ are not proportional, then there exist $p, q \in \mathbb{N}_0$ such that*

$$(\beta + \mathbb{Z}\alpha) \cap \Delta = \{\beta - j\alpha \colon j = -p, -p+1, \ldots, q-1, q\} \quad \textit{and} \quad (\beta, \check{\alpha}) = p - q.$$

Proof. Let $p, q \in \mathbb{N}_0$ be maximal with $\beta + q\alpha, \beta - p\alpha \in \Delta$. Since

$$\sigma_\alpha(\beta + j\alpha) = \beta + j\alpha - (\beta + j\alpha, \check{\alpha})\alpha \in \beta + \mathbb{Z}\alpha,$$

the α-string through β is invariant under σ_α. We conclude that

$$\sigma_\alpha(\beta - p\alpha) = \beta + q\alpha,$$

and this leads to

$$(\beta, \check{\alpha}) - 2p = (\beta - p\alpha, \check{\alpha}) = -q - p,$$

so that $(\beta, \check{\alpha}) = p - q$.

Suppose that there exists a $j \in \mathbb{Z}$ with $-p < j < q$ and $\gamma := \beta + j\alpha \notin \Delta$. We may w.l.o.g. assume that j is maximal with this property. Then $\sigma_\alpha(\gamma) = \beta + k\alpha \notin \Delta$ implies $k \leq j$, so that $(\gamma, \check{\alpha}) \geq 0$ and thus $(\gamma, \alpha) \geq 0$. The maximality of j further implies $\gamma + \alpha \in \Delta$. As $(\gamma + \alpha, \alpha) > (\gamma, \alpha) \geq 0$, Lemma 11.10 leads to the contradiction $\gamma = (\gamma + \alpha) - \alpha \in \Delta$. □

11.2 Root Bases

Definition 11.12. Let $\Delta \subseteq E$ be a root system and $\Pi \subseteq \Delta$. We write

$$\mathbb{N}_0[\Pi] := \Big\{ \sum_{\alpha \in \Pi} k_\alpha \alpha \colon k_\alpha \in \mathbb{N}_0 \Big\} \subseteq \mathbb{Z}[\Pi] = \operatorname{span}_{\mathbb{Z}} \Pi \subseteq E.$$

We call Π a *basis* for Δ if Π is a basis for the vector space E, and if

$$\Delta \subseteq \mathbb{N}_0[\Pi] \cup \mathbb{N}_0[-\Pi].$$

Then

$$\Delta^+ := \mathbb{N}_0[\Pi]$$

is called the corresponding *positive system*. It satisfies

$$\Delta = \Delta^+ \dot{\cup} - \Delta^+.$$

The elements of Π are called *simple roots* or *base roots*. The *height* of the root $\beta = \sum_{\alpha \in \Pi} k_\alpha \alpha$ is the number

$$\operatorname{ht}(\beta) := \sum_{\alpha \in \Pi} k_\alpha.$$

Lemma 11.13. *Let Δ be a root system and Π be a basis for Δ. Suppose that $\alpha, \beta \in \Pi$ with $\alpha \neq \beta$. Then $(\alpha, \beta) \leq 0$ and $\alpha - \beta$ is not a root.*

Proof. Since Π is a basis for E, α and β cannot be proportional. If $(\alpha, \beta) > 0$, then Lemma 11.10 shows that $\alpha - \beta \in \Delta$. But this contradicts the definition of a basis for Δ. □

Lemma 11.14. (Criterion for linear independence) *Let $M \subseteq E$ be contained in an open half space of E, i.e., there is a $\lambda \in E$ with $(\lambda, \alpha) > 0$ for all $\alpha \in M$, and $(\alpha, \beta) \leq 0$ for all $\alpha, \beta \in M$ with $\alpha \neq \beta$. Then M is linearly independent.*

Proof. Suppose that $\sum_{\alpha \in M} r_\alpha \alpha = 0$ with $r_\alpha \in \mathbb{R}$, and set

$$M_\pm := \{\alpha \in M \mid \pm r_\alpha > 0\}.$$

Then

$$\nu := \sum_{\alpha \in M_+} r_\alpha \alpha = \sum_{\beta \in M_-} (-r_\beta)\beta,$$

and therefore

$$(\nu, \nu) = \sum_{\alpha \in M_+, \beta \in M_-} r_\alpha(-r_\beta)(\alpha, \beta) \leq 0$$

which leads to $\nu = 0$. But we then have

$$0 = (\lambda, \nu) = \sum_{\alpha \in M_\pm} |r_\alpha|(\lambda, \alpha),$$

which implies $M_\pm = \emptyset$ since we otherwise arrive at a contradiction. $\square$

Definition 11.15. (a) Let $\Delta \subseteq E$ be a root system and $\lambda \in E$. Then λ is called *regular* if $\lambda \notin \alpha^\perp$ holds for all $\alpha \in \Delta$. Otherwise, λ is called *singular*.

(b) For any regular element $\lambda \in E$, the set

$$\Delta^+(\lambda) := \{\alpha \in \Delta \mid (\lambda, \alpha) > 0\}$$

is called the corresponding *positive system*. Note that

$$\Delta = \Delta^+(\lambda) \dot\cup - \Delta^+(\lambda).$$

An element $\alpha \in \Delta^+(\lambda)$ is called *decomposable* if there are $\beta_1, \beta_2 \in \Delta^+(\lambda)$ with $\alpha = \beta_1 + \beta_2$, otherwise, it is called *indecomposable*.

Theorem 11.16. *For each regular element $\lambda \in E$, the set $\Pi := \Pi(\lambda)$ of indecomposable elements in $\Delta^+(\lambda)$ is a basis for Δ. Conversely, every basis of Δ is of this form.*

Proof. **Claim 1**: $\Pi := \Pi(\lambda)$ is a basis for Δ.

First, we show that $\Delta^+(\lambda) \subseteq \mathbb{N}_0[\Pi]$. For this, we suppose that $\alpha \in \Delta^+(\lambda)$ cannot be written in this form, and that (α, λ) is minimal among all positive roots with this property. Then there exist $\beta_1, \beta_2 \in \Delta^+(\lambda)$ with $\alpha = \beta_1 + \beta_2$, and then $0 < (\beta_j, \lambda) < (\alpha, \lambda)$ for $j = 1, 2$. Now the minimality of (α, λ) yields $\beta_j \in \mathbb{N}_0[\Pi]$, which leads to the contradiction $\alpha \in \mathbb{N}_0[\Pi]$.

As a consequence, we see that $\Delta \subseteq \mathbb{N}_0[\Pi] \cup -\mathbb{N}_0[\Pi]$. Since Δ spans the space E, so does Π. It therefore remains to show that Π is linearly independent.

Next we show that $(\alpha, \beta) \leq 0$ for all $\alpha, \beta \in \Pi$ with $\alpha \neq \beta$. If $\alpha \in \mathbb{R}\beta$, then Remark 11.3 implies that $\alpha = 2\beta$ or $\beta = 2\alpha$, contradicting the definition of Π. We may therefore assume that $\beta \notin \mathbb{R}\alpha$. If $(\alpha, \beta) > 0$, then Lemma 11.10 implies $\alpha - \beta \in \Delta = \Delta^+(\lambda) \cup -\Delta^+(\lambda)$. If $\alpha - \beta \in \Delta^+(\lambda)$, then $\alpha = \beta + (\alpha - \beta)$ which contradicts the assumption that α is indecomposable. Similarly, $\beta - \alpha \in \Delta^+(\lambda)$ gives a contradiction by $\beta = \alpha + (\beta - \alpha)$. Now Claim 1 follows by Lemma 11.14, applied to $M = \Pi$.

Claim 2: Every basis Π for Δ is of the form $\Pi(\lambda)$ for some regular element $\lambda \in E$.

For $\alpha \in \Pi$, define $\widetilde{\alpha} \in E$ by $(\widetilde{\alpha}, \beta) = \delta_{\alpha,\beta}$ for $\beta \in \Pi$, so that we obtain a dual basis $\widetilde{\Pi} = \{\widetilde{\alpha} \colon \alpha \in \Pi\}$. Let $\lambda := \sum_{\alpha \in \Pi} \widetilde{\alpha}$. Then $(\lambda, \alpha) = 1$ for every $\alpha \in \Pi$, so that

$(\lambda, \alpha) > 0$ for all $\alpha \in \Pi$. Since every $\beta \in \Delta$ can be written as a linear combination of the $\alpha \in \Pi$ with coefficients of the same sign, λ is regular. Then the set Δ^+ of positive roots defined by Π satisfies $\Delta^+ \subseteq \Delta^+(\lambda)$ which leads to $\Delta^+ = \Delta^+(\lambda)$ because of $\Delta^+ \cup -\Delta^+ = \Delta = \Delta^+(\lambda) \cup -\Delta^+(\lambda)$. From the definition of a basis for Δ, we see that Π consists of indecomposable elements of $\Delta^+ = \Delta^+(\lambda)$, and therefore it is contained in $\Pi(\lambda)$. On the other hand, the cardinalities of Π and $\Pi(\lambda)$ are both equal to $n = \dim E$, since both sets are bases for E. This proves that $\Pi = \Pi(\lambda)$. $\square$

Examples 11.17. We describe bases of the root systems of types A-D. To this end we consider the element $\lambda = (n, n-1, \ldots, 1) \in \mathbb{R}^n$, which is regular for the root systems A_{n-1}, B_n, C_n and D_n. The corresponding bases and positive systems can be described as follows:

(a) For
$$A_{n-1} = \{\varepsilon_j - \varepsilon_k : j \neq k\} = \Delta(\mathfrak{sl}_n(\mathbb{K}), \mathfrak{h}),$$
the corresponding positive system is
$$\Delta^+(\lambda) = \{\varepsilon_j - \varepsilon_k : j < k\}$$
(strictly speaking, this is a positive system defined by the orthogonal projection $\widetilde{\lambda}$ of λ onto $\operatorname{span} A_{n-1}$) and
$$\Pi(\lambda) = \{\varepsilon_j - \varepsilon_{j+1} : j = 1, \ldots, n-1\}$$
is the corresponding basis.

(b) In
$$B_n = \{\pm\varepsilon_j, \pm\varepsilon_j \pm \varepsilon_k : j, k = 1, \ldots, n, j \neq k\} \subseteq \mathbb{R}^n$$
the corresponding positive system is
$$\Delta^+(\lambda) = \{\varepsilon_j, \varepsilon_j \pm \varepsilon_k : j < k\}$$
and
$$\Pi(\lambda) = \{\varepsilon_j - \varepsilon_{j+1} : j = 1, \ldots, n-1\} \cup \{\varepsilon_n\}$$
is the corresponding basis.

(c) In
$$C_n = \{\pm 2\varepsilon_j, \pm\varepsilon_j \pm \varepsilon_k : j, k = 1, \ldots, n, j \neq k\} \subseteq \mathbb{R}^n$$
the corresponding positive system is
$$\Delta^+(\lambda) = \{2\varepsilon_j, \varepsilon_j \pm \varepsilon_k : j < k\}$$
and
$$\Pi(\lambda) = \{\varepsilon_j - \varepsilon_{j+1} : j = 1, \ldots, n-1\} \cup \{2\varepsilon_n\}$$
is the corresponding basis.

(d) In
$$D_n = \{\pm\varepsilon_j \pm \varepsilon_k : j, k = 1, \ldots, n, j \neq k\} \subseteq \mathbb{R}^n$$
the corresponding positive system is
$$\Delta^+(\lambda) = \{\varepsilon_j \pm \varepsilon_k : j < k\}$$
and
$$\Pi(\lambda) = \{\varepsilon_j - \varepsilon_{j+1} : j = 1, \ldots, n-1\} \cup \{\varepsilon_{n-1} + \varepsilon_n\}$$
is the corresponding basis.

11.3 Weyl Chambers

Definition 11.18. The connected components of the set

$$E_{\text{reg}} := E \setminus \bigcup_{\alpha\in\Delta} \alpha^{\perp}$$

of the regular elements are called *Weyl chambers.* The Weyl chamber which contains the regular element $\lambda \in E$ is denoted by $\mathcal{C}(\lambda)$. It coincides with the set

$$\{\mu \in E\colon (\forall \alpha \in \Delta)\, (\lambda, \alpha) > 0 \Rightarrow (\mu, \alpha) > 0\}.$$

Remark 11.19. Let $\Delta \subseteq E$ be a root system and $\lambda, \lambda' \in E$ be regular elements.

(i) $\mathcal{C}(\lambda) = \mathcal{C}(\lambda') \quad\Leftrightarrow\quad \Delta^+(\lambda) = \Delta^+(\lambda') \quad\Leftrightarrow\quad \Pi(\lambda) = \Pi(\lambda')$.

(ii) By (i) and Theorem 11.16, there is a bijection between the set of Weyl chambers and the set of bases for Δ.

(iii) If $\Pi = \Pi(\lambda)$, then we call $\mathcal{C}(\Pi) := \mathcal{C}(\lambda)$ the *fundamental chamber* associated with the basis Π. It is given by

$$\begin{aligned}\mathcal{C}(\Pi) &= \{\beta \in E \mid (\beta, \alpha) > 0 \text{ for all } \alpha \in \Pi\}\\ &= \{\beta \in E \mid (\beta, \alpha) > 0 \text{ for all } \alpha \in \Delta^+\}.\end{aligned}$$

Definition 11.20. Let $\Pi \subseteq \Delta \subseteq E$ be a root basis. We define a partial order $\prec$ on E by

$$\alpha \prec \beta \quad :\Longleftrightarrow \quad \beta - \alpha \in \mathbb{N}_0[\Delta^+] = \mathbb{N}_0[\Pi].$$

Lemma 11.21. *Let $\Delta \subseteq E$ be a root system and Π be a basis for Δ.*

(i) *For $\alpha \in \Delta^+ \setminus \Pi$, there exists a $\beta \in \Pi$ with $\alpha - \beta \in \Delta^+$.*

(ii) *If $\alpha \in \Pi$, then σ_α permutes the set $\Delta^+ \setminus \mathbb{Z}\alpha$ (which coincides with $\Delta^+ \setminus \{\alpha\}$ is Δ is reduced).*

(iii) *Let $\alpha_1, \ldots, \alpha_r \in \Pi$ and set $\sigma_i := \sigma_{\alpha_i}$. If $\sigma_1 \cdots \sigma_{r-1}\sigma_r(\alpha_r) \in \Delta^+$, then there is an $s \in \{1, \ldots, r-1\}$ such that*

$$\sigma_1 \cdots \sigma_r = \sigma_1 \cdots \sigma_{s-1}\sigma_{s+1} \cdots \sigma_{r-1}.$$

Proof. (i) Suppose for all $\beta \in \Pi$, we have $(\alpha, \beta) \leq 0$. Writing $\Pi = \Pi(\lambda)$ for some regular element $\lambda \in E$ (Theorem 11.16), the set $\Pi \cup \{\alpha\}$ now satisfies the assumptions of Lemma 11.14, hence is linearly independent. Since Π is a basis for E, this cannot be the case, i.e., there is a $\beta \in \Pi$ with $(\alpha, \beta) > 0$.

Case 1: α and β are not proportional. Then Lemma 11.10 shows that $\alpha - \beta$ is a root. Since $\alpha \in \Delta^+ \setminus \Pi$, it is a linear combination of elements in Π with at least two positive (integral) coefficients. Subtracting β leaves at least one positive coefficient, so $\alpha - \beta$, being a root, has to be positive.

Case 2: α and β are proportional. Then Remark 11.3 shows that $\alpha = 2\beta$ since β is indecomposable. But then $\alpha - \beta = \beta$ is a root.

(ii) Let $\beta \in \Delta^+ \setminus \mathbb{Z}\alpha$ and $\beta = \sum_{\gamma\in\Pi} k_\gamma \gamma$ with $k_\gamma \in \mathbb{N}_0$. Then there is a $\gamma \neq \alpha$ with $k_\gamma > 0$. Since

$$\sigma_\alpha(\beta) = \beta - (\beta, \check{\alpha})\alpha = (k_\alpha - (\beta, \check{\alpha}))\alpha + \sum_{\gamma\in\Pi\setminus\{\alpha\}} k_\gamma \gamma,$$

$\sigma_\alpha(\beta)$ has a positive coefficient k_γ in its representation as a linear combination of simple roots. Thus, all coefficients are nonnegative, and $\sigma_\alpha(\beta) \in \Delta^+$. By $\sigma_\alpha(\alpha) = -\alpha$, we also have $\sigma_\alpha(\beta) \neq \alpha$, i.e., $\sigma_\alpha(\beta) \in \Delta^+ \setminus \{\alpha\}$. Since the latter set is finite, the claim follows.

(iii) Set

$$\beta_i := \begin{cases} \sigma_{i+1} \cdots \sigma_{r-1}(\alpha_r) & \text{for } i = 0, \ldots, r-2, \\ \alpha_r & \text{for } i = r-1. \end{cases}$$

Then $\beta_0 \notin \Delta^+$ and $\beta_{r-1} = \alpha_r \in \Delta^+$, so that there exists a minimal $s \in \{1, \ldots, r-1\}$ with $\beta_s \in \Delta^+$. For this s, we have $\sigma_s(\beta_s) = \beta_{s-1} \notin \Delta^+$. In view of (ii), this shows $\beta_s = \alpha_s$ because $\frac{1}{2}\beta_s$ is not a root. By Proposition 11.9, for $w := \sigma_{s+1} \cdots \sigma_{r-1}$, we have

$$\sigma_s = \sigma_{\beta_s} = \sigma_{w\alpha_r} = w\sigma_r w^{-1} = (\sigma_{s+1} \cdots \sigma_{r-1})\sigma_r(\sigma_{r-1} \cdots \sigma_{s+1}),$$

which shows that

$$\sigma_s \cdots \sigma_r = \sigma_{s+1} \cdots \sigma_{r-1},$$

and this implies the claim. □

Corollary 11.22. *Let $\Delta \subseteq E$ be a root system and Π be a basis for Δ.*

(i) *Every $\beta \in \Delta^+$ can be written in the form $\alpha_1 + \ldots + \alpha_m$ with $\alpha_j \in \Pi$ such that $\sum_{j=1}^k \alpha_j \in \Delta^+$ for each $k \in \{1, \ldots, m\}$.*

(ii) *Let $\sigma = \sigma_1 \cdots \sigma_r$, where the $\sigma_j = \sigma_{\alpha_j}$ are reflections associated with the simple roots $\alpha_j \in \Pi$, and where r is the minimal number of factors needed to represent σ as such a product. Then $\sigma(\alpha_r) \in -\Delta^+$.*

Proof. (i) follows by induction with Lemma 11.21(i).

(ii) If the assertion is not true, then $\sigma(\alpha_r) \in \Delta^+$, so that Lemma 11.21(iii) contradicts the minimality of the number r of factors. □

Theorem 11.23. *Let Δ be a reduced root system, $\mathcal{W}$ be the corresponding Weyl group, and Π a basis for Δ.*

(i) *For every regular element $\lambda \in E$, there is a $\sigma \in \mathcal{W}$ such that*

$$(\sigma\lambda, \alpha) > 0 \quad \textit{for} \quad \alpha \in \Pi,$$

i.e., $\sigma(\mathcal{C}(\lambda)) = \mathcal{C}(\Pi)$. In particular, $\mathcal{W}$ acts transitively on the set of the Weyl chambers.

(ii) *Let Π' be another basis for Δ. Then there is a $\sigma \in \mathcal{W}$ with $\sigma(\Pi') = \Pi$, i.e., the Weyl group also acts transitively on the set of the bases.*

(iii) *For every root $\alpha \in \Delta$, there is a $\sigma \in \mathcal{W}$ with $\sigma(\alpha) \in \Pi$.*

(iv) *$\mathcal{W}$ is generated by the σ_α with $\alpha \in \Pi$.*

(v) *If $\sigma(\Pi) = \Pi$ for $\sigma \in \mathcal{W}$, then $\sigma = \mathbf{1}$.*

Proof. Let $\mathcal{W}'$ be the subgroup of $\mathcal{W}$, generated by the σ_α with $\alpha \in \Pi$. Suppose, (iii) holds for $\mathcal{W}'$ instead of $\mathcal{W}$. Then for $\alpha \in \Delta$, we can find a $w \in \mathcal{W}'$ with $w\alpha \in \Pi$. Then by $\sigma_{w\alpha} = w\sigma_\alpha w^{-1}$ (cf. Proposition 11.9), we obtain that $\sigma_\alpha = w^{-1}\sigma_{w\alpha}w \in \mathcal{W}'$. Since $\mathcal{W}$ is generated by the σ_α with $\alpha \in \Delta$, we get $\mathcal{W} = \mathcal{W}'$, which is (iv). Further (v) is an immediate consequence of Corollary 11.22(ii) applied to w because it implies that the minimal number of factors σ_α, $\alpha \in \Pi$, in w must be zero. Therefore, we see that it suffices to show (i)-(iii) for $\mathcal{W}'$ instead of $\mathcal{W}$.

(i) Let $\mu \in \mathcal{C}(\Pi)$ and choose $w \in \mathcal{W}'$ such that the number $(w\lambda, \mu)$ is maximal for the given λ. For $\alpha \in \Pi$, we then obtain

$$(w\lambda, \mu) \geq (\sigma_\alpha w\lambda, \mu) = (w\lambda, \sigma_\alpha\mu) = (w\lambda, \mu) - (w\lambda, (\mu, \check{\alpha})\alpha) = (w\lambda, \mu) - (w\lambda, \alpha)(\mu, \check{\alpha}).$$

This gives

$$(w\lambda, \alpha)(\mu, \alpha) \geq 0,$$

and since $(\mu, \alpha) > 0$ for every $\alpha \in \Pi$, we obtain $(w\lambda, \alpha) \geq 0$ and $(w\lambda, \alpha) > 0$ by regularity. Therefore $w\lambda \in \mathcal{C}(\Pi)$ and thus $w\mathcal{C}(\lambda) = \mathcal{C}(\Pi)$.

(ii) By (i), this is an immediate consequence of Remark 11.19.

(iii) Because of (ii), it suffices to show that α is an element of *some* basis. If $\beta \neq \pm\alpha$ is a root, then α and β are not proportional since Δ is reduced. Thus, $\alpha^\perp \neq \beta^\perp$, and we can find a

$$\lambda \in \alpha^\perp \setminus \bigcup_{\beta \in \Delta \setminus \{\pm\alpha\}} \beta^\perp$$

because we have to avoid finitely many hyperplanes in $\alpha^\perp$. Let $\lambda' := \lambda + \varepsilon\alpha$, where $\varepsilon > 0$ is chosen such that

$$0 < \varepsilon(\alpha, \alpha) = (\lambda', \alpha) < \min\{|(\lambda', \beta)| \colon \beta \in \Delta \setminus \{\pm\alpha\}\}.$$

Then λ' is regular with $\alpha \in \Pi(\lambda')$ (cf. Theorem 11.16). $\square$

Proposition 11.24. (The dual root system) *If $\Delta \subseteq E$ is a root system, then*

$$\check{\Delta} := \{\check{\alpha} \colon \alpha \in \Delta\}$$

also is a root system. It is reduced if and only if Δ is reduced. Moreover, if Δ^+ is a positive system of Δ, then

$$\check{\Delta}^+ := \{\check{\alpha} \colon \alpha \in \Delta^+\}$$

is a positive system of $\check{\Delta}$, and if $\Pi \subseteq \Delta^+$ is a root basis, then

$$\check{\Pi} := \{\check{\alpha} \colon \alpha \in \Pi, 2\alpha \notin \Delta\} \cup \{\tfrac{1}{2}\check{\alpha} \colon \alpha \in \Pi, 2\alpha \in \Delta\}$$

is a root basis of $\check{\Delta}$.

The root system $\check{\Delta}$ is called the *root system dual to* Δ.

Proof. To verify (R1) for $\check{\Delta}$ (if Δ is reduced), we note that $\check{\beta} \in \mathbb{R}\check{\alpha}$ implies $\beta \in \mathbb{R}\alpha$ and hence $\beta = \pm\alpha$, which in turn leads to $\check{\beta} = \pm\check{\alpha}$. If Δ is not reduced, then $(2\alpha)\check{} = \frac{1}{2}\check{\alpha}$ shows that $\check{\Delta}$ is also not reduced.

Since $\sigma_{\check{\alpha}}$ is the orthogonal reflection in $\alpha^\perp = \check{\alpha}^\perp$, we have $\sigma_\alpha = \sigma_{\check{\alpha}}$. As σ_α is an isometry, it satisfies $\sigma_\alpha(\check{\beta}) = \sigma_\alpha(\beta)\check{}$, so that $\check{\Delta}$ satisfies (R2). Finally we note that for $\alpha, \beta \in \Delta$, we have

$$(\check{\alpha}, \check{\alpha}) = \frac{4}{(\alpha, \alpha)},$$

so that $(\check{\alpha})\check{} = \alpha$. Therefore

$$(\check{\alpha}, (\check{\beta})\check{}) = (\beta, \check{\alpha}) \in \mathbb{Z},$$

and we conclude that $\check{\Delta}$ also is a root system.

Now let

$$\Delta^+ = \Delta^+(\lambda) = \{\alpha \in \Delta \colon (\lambda, \alpha) > 0\}$$

be a positive system of Δ. From the definition of the dual root system, it follows that Δ and $\check{\Delta}$ define the same set of regular elements. Therefore

$$\check{\Delta}^+ = \{\check{\alpha} \in \check{\Delta} \colon (\lambda, \check{\alpha}) > 0\}$$

is a positive system of the dual root system $\check{\Delta}$.

To see that $\check{\Pi}$ is a root basis, we argue that $\check{\Pi}$ is contained in the root basis Π' of indecomposable elements of $\check{\Delta}^+$. Then $|\check{\Pi}| = |\Pi| = \dim E = |\Pi'|$ implies that $\check{\Pi} = \Pi'$ is a root basis. So let $\check{\alpha} \in \check{\Pi}$ and assume that it is decompasable: $\check{\alpha} = \check{\beta}_1 + \check{\beta}_2$ with $\beta_j \in \Delta^+$. Write $\beta_j = \sum_{\gamma \in \Pi} c_\gamma^{(j)} \gamma$ with $c_\gamma^{(j)} \geq 0$. As $\check{\beta}_j$ is a positive multiple of β_j and $\check{\alpha} \in \mathbb{R}^+\alpha$, it follows that all coefficients $c_\gamma^{(j)}$ vanish for $\gamma \neq \alpha$. Hence $\beta_j \in \mathbb{R}\alpha$ and thus $\beta_j = 2\alpha$. But then the definition of $\check{\Pi}$ implies that $\check{\alpha} \notin \check{\Pi}$. □

Examples 11.25. (a) In the root systems A_n and D_n, all roots have the square-lenght 2, so that they are self-dual: $\check{A}_n = A_n$ and $\check{D}_n = D_n$.

(b) In the root sustem B_n, we have two different root lengths, the roots $\pm\varepsilon_i \pm \varepsilon_j$ have square lenght 2, and the roots $\pm\varepsilon_j$ have square lenght 1. This leads to $\check{\varepsilon}_j = 2\varepsilon_j$. From that we derive that $\check{B}_n = C_n$ and $\check{C}_n = B_n$. In this sense the root systems B_n and C_n are dual to each other.

Exercises for Section 11

Exercise 11.1. Let V be a finite dimensional vector space over the field $\mathbb{K}$ and let $U \in \mathrm{GL}(V)$ be unipotent, i.e., $U - \mathbf{1}$ is nilpotent. Show that

(a) If $\operatorname{char} \mathbb{K} = 0$ and U is of finite order in $\mathrm{GL}(V)$, then $U = \mathbf{1}$. Hint: Describe U by an upper triangular matrix with respect to a suitable basis.

(b) Suppose that $\operatorname{char} \mathbb{K} > 0$. Find a non-trivial unipotent matrix of finite order. Which orders can occur?

Exercise 11.2. (An irreducible non-reduced root system) Let $n \in \mathbb{N}$. Show that

$$BC_n := B_n \cup C_n = \{\pm\varepsilon_j, \pm 2\varepsilon_j, \pm\varepsilon_j \pm \varepsilon_k : j,k = 1, \ldots, n, j \neq k\}$$

is an irreducible, non-reduced root system in the euclidean space $\mathbb{R}^n$, endowed with the standard scalar product. Sketch the root system for $n = 2$. What is the corresponding Weyl group?

Exercise 11.3. In the euclidean plane

$$E := \Big\{x \in \mathbb{R}^3 \colon \sum_j x_j = 0\Big\}$$

we consider the subset

$$G_2 := A_2 \cup \big\{ \pm(2\varepsilon_1 - \varepsilon_2 - \varepsilon_3), \pm(2\varepsilon_2 - \varepsilon_1 - \varepsilon_3), \pm(2\varepsilon_3 - \varepsilon_1 - \varepsilon_2)\big\}.$$

1. Show that G_2 is a root system.
2. Determine a root basis.
3. Determine the number of Weyl chambers. Hint: They all have the same angle.
4. Draw a picture.
5. Determine its Weyl group.

Exercise 11.4. Let V be a real vector space. A element $\sigma \in \mathrm{GL}(V)$ is called a *reflection* if it is of the form

$$\sigma_v(w) = w - v^*(w)v \quad \text{with} \quad v^* \in V^*, v^*(v) = 2.$$

We call a set $(\sigma_v)_{v\in S}$ of reflections *irreducible* if, for $v, w \in S$, there exist $v_0 = v, v_1, \ldots, v_n = w \in S$ with

$$v_{j+1}^*(v_j) \neq 0 \quad \text{for} \quad j = 1, \ldots, n-1.$$

We write $\mathcal{W} := \langle \sigma_v \colon v \in S\rangle \subseteq \mathrm{GL}(V)$ for the subgroup generated by the reflections σ_v, $v \in S$. Show that:

(i) $V^{\mathcal{W}} := \{v \in V \colon (\forall w \in \mathcal{W}) w(v) = v\} = \bigcap_{s\in S} \ker s^*$.

(ii) $V_{\text{eff}} := \mathrm{span}\{w(v) - v \colon w \in \mathcal{W}, v \in V\}$ is a $\mathcal{W}$-invariant subspace containing S and minimal with this property.

(iii) Assume that $(\sigma_v)_{v\in S}$ is irreducible, that $\mathrm{span}\, S = V$ and $\bigcap_{s\in S} \ker s^* = \{0\}$. Then the representation of $\mathcal{W}$ on V is irreducible.
Hint: Show that every non-zero $\mathcal{W}$-invariant subspace $V_0 \subseteq V$ contains S.

(iv) If $\varphi \in \mathrm{End}(V)$ commutes with $\mathcal{W}$, then $\varphi \in \mathbb{R}\,\mathrm{id}_V$.

Exercise 11.5. (Uniqueness of scalar product for irreducible root systems) Let $(E, (\cdot,\cdot))$ be a euclidean space and $\Delta \subseteq E$ be an irreducible root system. Further, let $\langle\cdot,\cdot\rangle$ be a second inner product on E for which $\Delta \subseteq (E, \langle\cdot,\cdot\rangle)$ also is a root system. Show that there exists a $c > 0$ with $\langle v, w\rangle = c(v, w)$ for $v, w \in E$. One may proceed along the following steps:

(a) For $\alpha, \beta \in \Delta$, we have

$$\frac{\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} = \frac{(\beta,\alpha)}{(\alpha,\alpha)}.$$

Hint: This can be derived from the Root String Lemma.

(b) The reflections σ_α, $\alpha \in \Delta$, are the same for both scalar products.

(c) For $\alpha \in \Delta$, put $c_\alpha := \frac{\langle\alpha,\alpha\rangle}{(\alpha,\alpha)}$. Show that

(i) $c_\alpha = c_\beta$ for $(\beta, \alpha) \neq 0$.
(ii) $c := c_\alpha$ is independet of α.
(iii) $\langle\beta,\alpha\rangle = c(\beta,\alpha)$ for $\alpha, \beta \in \Delta$.

12 The Classification of Simple Split Lie Algebras

In this section we eventually come to the classification of those simple finite dimensional Lie algebras over a field $\mathbb{K}$ of characteristic zero which are *split* in the sense that they contain a toral Cartan subalgebra $\mathfrak{h}$. We won't go through the full proof of the classification but we shall explain the strategy and the result. The main idea is to associate to a basis Π of the root system $\Delta(\mathfrak{g}, \mathfrak{h})$ a (weighted directed) graph, called the *Dynkin diagram.* This translates the classification of finite reduced root systems into the classification problem for Dynkin diagrams, which can be carried out by elementary means (but we won't go into details of the proof of the classification).

12.1 Cartan Matrices

Definition 12.1. Let Δ be a roots system and $\Pi \subseteq \Delta$ be a root basis. The matrix

$$CM(\Pi) = (\alpha(\check{\beta}))_{\alpha,\beta\in\Pi}$$

is called the *Cartan matrix of* Π. [14] The integers $\alpha(\check{\beta})$ are called *Cartan integers.*

Lemma 12.2. *The Cartan matrix $CM(\Pi)$ has the following properties (for $\alpha, \beta \in \Pi$):*

(CM1) $\alpha(\check{\beta}) \in \mathbb{Z}$.

(CM2) $\alpha(\check{\alpha}) = 2$.

(CM3) $\alpha(\check{\beta}) \leq 0$ *for* $\alpha \neq \beta$.

(CM4) $\alpha(\check{\beta}) < 0 \Rightarrow \beta(\check{\alpha}) < 0$.

(CM5) $\alpha(\check{\beta})\beta(\check{\alpha}) \in \{0, 1, 2, 3\}$.

Proof. (CM1) follows from Lemma 11.11.
(CM2) follows from the Definition of $\check{\alpha}$.
(CM3) follows from Lemma 11.10.
(CM4) follows from

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

(CM5) follows from Remark 11.7 which asserts that $\alpha(\check{\beta})\beta(\check{\alpha}) = 4\cos^2(\theta)$, where θ is the angle between α and β. □

There is a subtlety involved in the definition of Cartan matrices that deserves to be noticed because it is relevant for classification purposes. A Cartan matrix, as we defined it, is a function $\Pi \times \Pi \to \mathbb{Z}$ and as such an "abstract matrix" (which could encode a linear map with respect to a basis without reference to any enumeration). To write this as an element of $M_r(\mathbb{Z})$, $r := |\Pi|$, one first has to introduce a linear order on Π:

$$\Pi = \{\alpha_1, \dots, \alpha_\ell\}.$$

Once this enumeration is fixed, we can identify the Cartan matrix with a "concrete matrix" $A = (a_{ij})_{1\leq i,j\leq r} \in M_r(\mathbb{Z})$, where $a_{ij} = \alpha_i(\check{\alpha}_j)$. In this representation, the matrix A has the following properties (for all i, j):

[14] Élie Cartan used these matrices 1894 in his classification of the complex simple Lie algebras.

(CM1)' $a_{ij} \in \mathbb{Z}$.

(CM2)' $a_{ii} = 2$.

(CM3)' $a_{ij} \leq 0$.

(CM4)' $a_{ij} < 0 \Rightarrow a_{ji} < 0$.

(CM5)' $a_{ij}a_{ji} \in \{0, 1, 2, 3\}$.

The two concrete matrices Cartan matrices

$$C_1 = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \quad \text{und} \quad C_2 = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

are two concretizations of the same abstract Cartan matrix of the root system A_3. More generally, for the ordered root basis $\alpha_i = \varepsilon_i - \varepsilon_{i+1}$, $i = 1, \ldots, n-1$, of the root system A_{n-1}, we obtain the concrete Cartan matrix

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

Proposition 12.3. *A reduced root system Δ is completely determined by the Cartan matrix $CM(\Pi)$ of a root basis Π.*

Proof. We show by induction on the height n, that the set

$$\Delta^+(n) := \{\beta \in \Delta^+ \colon |\beta| = n\}$$

is determined by the Cartan matrix. For $n = 1$, we have $\Delta^+(1) = \Pi$. So let us assume that $n > 1$. We assume inductively that $\Delta^+(k)$, $k < n$, is determined by the Cartan matrix $CM(\Pi)$. For every $\beta \in \Delta^+(n)$, there exists by Lemma 11.21(i) a simple root $\alpha \in \Pi$ with $\beta - \alpha \in \Delta$, and thus $\beta - \alpha \in \Delta^+(n-1)$. Therefore we only have to determine for which roots $\gamma \in \Delta^+(n-1)$ and which $\alpha \in \Pi$ the sum $\alpha + \gamma$ is a root.

If $\gamma = \alpha$, then $\gamma + \alpha = 2\alpha \notin \Delta$. We may therefore assume that $\gamma \neq \alpha$, i.e., in the representation of γ with respect to the basis Π at least one other simple root $\alpha' \neq \alpha$ occurs. Then all roots of the form $\gamma - j\alpha$, $j \in \mathbb{Z}$, are positive, because there α'-coefficient is positive. Since we know already that the set $\bigcup_{j \leq n-1} \Delta^+(j)$ is determined by $CM(\Pi)$, we can determine the maximal $p \in \mathbb{N}_0$ with $\gamma - p\alpha \in \Delta$. Then the Root String Lemma 11.11 implies that $q = p - \gamma(\check{\alpha})$ is maximal with the property that $\gamma + q\alpha \in \Delta$. Hence $\gamma + \alpha$ is equivalent to $\gamma(\check{\alpha}) < p$. Note that $\gamma(\check{\alpha})$ can be calculated from the representation of γ w.r.t. the basis Π and the Cartan matrix. This completes the proof. $\square$

The proof of Proposition 12.3 contains in particular an algorithm how to construct the root system from its Cartan matrix by constructing inductively the layers $\Delta^+(n)$ from $\Delta^+(1) = \Pi$.

Example 12.4. We explain this procedure for the Cartan matrix

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \alpha_1(\check{\alpha}_1) & \alpha_1(\check{\alpha}_2) \\ \alpha_2(\check{\alpha}_1) & \alpha_2(\check{\alpha}_2) \end{pmatrix}.$$

We start with

$$\Pi = \Delta^+(1) = \{\alpha_1, \alpha_2\}.$$

From $\alpha_2 - \alpha_1 \notin \Delta$ we find for the α_1-string through α_2 and the α_2-string through α_1:

$$\alpha_2, \alpha_2 + \alpha_1; \quad \alpha_1, \alpha_1 + \alpha_2, \alpha_1 + 2\alpha_2, \alpha_1 + 3\alpha_2$$

(Root String Lemma 11.11). We successively obtain

$$\Delta^+(2) = \{\alpha_1 + \alpha_2\} \quad \text{and} \quad \Delta^+(3) = \{\alpha_1 + 2\alpha_2\}.$$

To $\gamma = \alpha_1 + 2\alpha_2$ we may add α_2 and obtain the root $\gamma + \alpha_2$. The α_1-string through γ contains no root of smaller height because $2\alpha_2$ is not a root. In view of $\gamma(\check{\alpha}_1) = 2 - 2 = 0$, the α_1-string through γ is trivial. We thus obtain

$$\Delta^+(4) = \{\alpha_1 + 3\alpha_2\}.$$

Here we are not allowed to add α_2, but since $(\alpha_1 + 3\alpha_2)(\check{\alpha}_1) = 2 - 3 = -1$, we get

$$\Delta^+(5) = \{2\alpha_1 + 3\alpha_2\}.$$

As the α_1-string through $\alpha_1 + 3\alpha_2$ ends in $2\alpha_1 + 3\alpha_2$ (because $(\alpha_1 + 3\alpha_2)(\check{\alpha}_1) = -1$), the element $2\alpha_1 + 2\alpha_2$ is not a root, and $(2\alpha_1 + 3\alpha_2)(\check{\alpha}_2) = -6 + 6 = 0$, we cannot add any other root. We thus obtain 6 positive roots

$$\Delta^+ = \{\alpha_1, \alpha_2, \alpha_1 + \alpha_2, \alpha_1 + 2\alpha_2, \alpha_1 + 3\alpha_2, 2\alpha_1 + 3\alpha_2\}.$$

The so constructed root system $\Delta = \Delta^+ \cup -\Delta^+$ is called G_2. One can show that it belongs to a 14-dimensional simple Lie algebra $\mathfrak{g}$ with a 2-dimensional toral Cartan subalgebra. The dimension of $\mathfrak{g}$ is obtained from the root decomposition by the formula

$$\dim \mathfrak{g} = \dim \mathfrak{h} + |\Delta| = 2 + 12 = 14.$$

Presently, we have not shown that a Lie algebra with this root system exists.

The next step to the classification is to classify all Cartan matrices of finite root systems. This can be achieved by an analysis of the corresponding Dynkin diagram, introduced below.

12.2 Irreducibility and simplicity

First we reduce the classification of Cartan matrices to the case of irreducible root systems which corresponds to the case of simple Lie algebras.

Definition 12.5. A subset $F \subseteq \Delta$ is called *indecomposable* if it cannot be written as a disjoint sum $F = F_1 \dot\cup F_2$ of two non-empty subsets with $F_1 \perp F_2$. Note that the orthogonality of F_1 and F_2 is equivalent to $\alpha(\check{\beta}) = 0$ for $\alpha \in F_1$ and $\beta \in F_2$.

Theorem 12.6. (Generation Theorem) *Let* $\Pi \subseteq \Delta$ *be a root basis. Then* $\mathfrak{g}$ *is generated by the* 3*-dimensional subalgebras* $\mathfrak{g}(\alpha), \alpha \in \Pi$.

Proof. Let $\mathfrak{q} \subseteq \mathfrak{g}$ be the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{g}(\alpha)$, $\alpha \in \Pi$.

In view of Corollary 11.22(i), every $\beta \in \Delta^+$ can be written in the form $\alpha_1 + \ldots + \alpha_m$ with $\alpha_j \in \Pi$ such that $\sum_{j=1}^k \alpha_j \in \Delta^+$ for each $k \in \{1, \ldots, m\}$. Then Lemma 10.11, which asserts that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$ if $\alpha, \beta, \alpha + \beta \in \Delta$, implies that

$$\mathfrak{g}_\beta = [\mathfrak{g}_{\alpha_m}, [\mathfrak{g}_{\alpha_{m-1}}, [\cdots, [\mathfrak{g}_{\alpha_2}, \mathfrak{g}_{\alpha_1}] \cdots]]] \subseteq \mathfrak{q}.$$

We likewise see that $\mathfrak{g}_\beta \subseteq \mathfrak{q}$ for $\beta \in -\Delta^+$. Therefore $\mathfrak{q}$ contains all root spaces and

$$\operatorname{span}\{h_\alpha \colon \alpha \in \Pi\} = \operatorname{span}\{t_\alpha \colon \alpha \in \Pi\} = \mathfrak{h}$$

(Lemma 10.12(iii)). $\square$

Proposition 12.7. *A semisimple Lie algebra $\mathfrak{g}$ with toral Cartan subalgebra $\mathfrak{h}$ and root system $\Delta = \Delta(\mathfrak{g}, \mathfrak{h})$ is simple if and only if Δ has an indecomposable root basis.*

Proof. Suppose first that $\mathfrak{g}$ is not simple and let $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$, where $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are proper semisimple ideals. Let $\mathfrak{h}_j \subseteq \mathfrak{g}_j$ be toral Cartan subalgebras. Then $\mathfrak{h} := \mathfrak{h}_1 \oplus \mathfrak{h}_2$ is a toral Cartan subalgebra of $\mathfrak{g}$ and the corresponding root system is given by

$$\Delta = \Delta_1 \dot{\cup} \Delta_2, \quad \text{where} \quad \Delta_1 = \Delta(\mathfrak{g}_1, \mathfrak{h}_1) \times \{0\} \quad \text{and} \quad \Delta_2 = \{0\} \times \Delta(\mathfrak{g}_2, \mathfrak{h}_2).$$

The corresponding root spaces are

$$\mathfrak{g}^{(\alpha,0)} = \mathfrak{g}_1^\alpha \times \{0\} \quad \text{and} \quad \mathfrak{g}^{(0,\beta)} = \{0\} \times \mathfrak{g}_2^\beta.$$

If $\Pi_j \subseteq \Delta_j$, $j = 1, 2$, are root bases, then

$$\Pi = (\Pi_1 \times \{0\}) \cup (\{0\} \times \Pi_2)$$

is a partition of the root basis Π of Δ into subsets with $(\alpha, 0)\big((0, \check{\beta})\big) = 0$ for $\alpha \in \Pi_1$ and $\beta \in \Pi_2$. Therefore Π is decomposable.

Now we assume that $\Pi = \Pi_1 \dot{\cup} \Pi_2$ with $\alpha(\check{\beta}) = 0$ for $\alpha \in \Pi_1$ und $\beta \in \Pi_2$. Let $\mathfrak{g}_j \subseteq \mathfrak{g}$ denote the subagebras generated by the subalgebras $\mathfrak{g}(\alpha)$, $\alpha \in \Pi_j$. For $\alpha \in \Pi_1$, $\beta \in \Pi_2$ we then have $\alpha - \beta \notin \Delta$ because both are base roots, so that $\alpha(\check{\beta}) = 0$ also implies $\alpha + \beta \notin \Delta$ (Root String Lemma 11.11). We conclude that $[\mathfrak{g}_{\pm\alpha}, \mathfrak{g}_{\pm\beta}] = \{0\}$ and thus $[\mathfrak{g}_{\pm\alpha}, \mathfrak{g}(\beta)] = \{0\}$, so that we further obtain $[\mathfrak{g}(\alpha), \mathfrak{g}(\beta)] = \{0\}$. This in turn implies that $[\mathfrak{g}(\beta), \mathfrak{g}_1] = \{0\}$ because $\alpha \in \Pi_1$ was arbitrary and the centralizer $\mathfrak{z}_{\mathfrak{g}}(\mathfrak{g}(\beta))$ of $\mathfrak{g}(\beta)$ in $\mathfrak{g}$ is a subalgebra. With a similar argument we see that $[\mathfrak{g}_2, \mathfrak{g}_1] = \{0\}$. From the Generation Theorem 12.6, it now follows that $\mathfrak{g} = \mathfrak{g}_1 + \mathfrak{g}_2$, so that $\mathfrak{g}_1 \trianglelefteq \mathfrak{g}$ is a proper ideal. In particular $\mathfrak{g}$ is not simple. $\square$

From the preceding proof one can easily derive that, in general, the irreducible components of Π correspond to the simple ideals of $\mathfrak{g}$.

12.3 Simplicity of the Classical Series

In this subsection we use the preceding results to verify that the Lie algebras $\mathfrak{sl}_n(\mathbb{K})$, $\mathfrak{o}_{n,n}(\mathbb{K})$ ($n \neq 1, 2$), $\mathfrak{o}_{n,n+1}(\mathbb{K})$ and $\mathfrak{sp}_{2n}(\mathbb{K})$ are simple. We have seen in Examples 8.7, 8.8 and 8.9 that these Lie algebra contain a toral Cartan subalgebra for which the root systems are of type

- A_{n-1} for $\mathfrak{sl}_n(\mathbb{K})$,
- B_n for $\mathfrak{o}_{n,n+1}(\mathbb{K})$,
- C_n for $\mathfrak{sp}_{2n}(\mathbb{K})$, and
- D_n for $\mathfrak{o}_{n,n}(\mathbb{K})$.

With the explicit description of the root spaces, it is easy to verify that all subalgebras $\mathfrak{g}(\alpha) = \mathfrak{g}_\alpha + \mathfrak{g}_{-\alpha} + [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ are isomorphic to $\mathfrak{sl}_2(\mathbb{K})$, so that the semisimplicity criterion from Lemma 8.14 implies that these four Lie algebras are semisimple if we can show that $\mathfrak{z}(\mathfrak{g}) = \{0\}$. If, in addition, Π is indecomposable, then it is simple by Proposition 12.7.

Example 12.8. For $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ we have the roots $\alpha_{ij} = \varepsilon_i - \varepsilon_j$. Then $\mathfrak{g}_{\alpha_{ij}} = \mathbb{K}E_{ij}$ and

$$\alpha_{ij}([E_{ij}, E_{ji}]) = (\varepsilon_i - \varepsilon_j)(E_{ii} - E_{jj}) = 2,$$

so that $\check{\alpha}_{ij} = E_{ii} - E_{jj}$. A diagonal matrix $h = \mathrm{diag}(h_1, \ldots, h_n)$ is central if and only if $h_i - h_j = \alpha_{ij}(h) = 0$ for $1 \leq i < j \leq n$. This is equivalent to $h_1 = \ldots = h_n$, so that $\sum_j h_j = 0$ leads to $h = 0$. Therefore $\mathfrak{z}(\mathfrak{sl}_n(\mathbb{K})) = \{0\}$. Now Lemma 8.14 implies the semisimplicity of $\mathfrak{g}$. To verify the simplicity, we consider the root basis

$$\Pi = \{\varepsilon_1 - \varepsilon_2, \ldots, \varepsilon_{n-1} - \varepsilon_n\}$$

from Example 11.17 which satisfies $\alpha_{i,i+1}(\check{\alpha}_{i+1,i+2}) = -1 \neq 0$. Therefore Π is indecomposable and therefore $\mathfrak{g}$ is simple. For $\alpha_j := \varepsilon_j - \varepsilon_{j+1}$, the Cartan matrix is

$$(\alpha_j(\check{\alpha}_k))_{1\leq j,k\leq n-1} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & & \ddots & & \vdots \\ \vdots & & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

Example 12.9. We consider the Lie algebra $\mathfrak{g} = \mathfrak{sp}_{2n}(\mathbb{K})$, whose root decomposition we know from Example 8.9. We have seen that

$$\mathfrak{h} := \mathrm{span}\{E_{jj} - E_{j+n,j+n} \colon j = 1, \ldots, n\} \subseteq \mathfrak{g} \subseteq \mathfrak{gl}_{2n}(\mathbb{K})$$

is a toral Cartan subalgebra and that the root system is given by

$$\Delta = \{\pm 2\varepsilon_j, \pm(\varepsilon_j \pm \varepsilon_k) \colon j \neq k, j, k = 1, \ldots, n\} = C_n.$$

From this explicit description of the root system, we immediately derive that $\mathfrak{z}(\mathfrak{g}) = \bigcap_{\alpha\in\Delta} \ker\alpha = \{0\}$. For the coroots we obtain

$$\begin{aligned} (\varepsilon_j - \varepsilon_k)\check{} &= E_{jj} - E_{kk} - E_{j+n,j+n} + E_{k+n,k+n} \\ (\varepsilon_j + \varepsilon_k)\check{} &= E_{jj} + E_{kk} - E_{j+n,j+n} - E_{k+n,k+n}, \quad j \neq k, \\ (2\varepsilon_j)\check{} &= E_{jj} - E_{j+n,j+n}, \end{aligned}$$

because

$$\mathfrak{g}_{\varepsilon_j+\varepsilon_k} = \mathbb{K}(E_{j,n+k} + E_{k,n+j}), \qquad \mathfrak{g}_{-\varepsilon_j-\varepsilon_k} = \mathbb{K}(E_{n+k,j} + E_{n+j,k})$$

and

$$[E_{j,k+n} + E_{k,j+n}, E_{j+n,k} + E_{k+n,j}] = (1 + \delta_{jk})(E_{jj} - E_{j+n,j+n}) + (1 + \delta_{jk})(E_{kk} - E_{k+n,k+n}).$$

With this explicit information on the brackets, it is easy to see that all subalgebras $\mathfrak{g}(\alpha) = \mathfrak{g}_\alpha + \mathfrak{g}_{-\alpha} + [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ are isomorphic to $\mathfrak{sl}_2(\mathbb{K})$, so that Lemma 8.14 implies that $\mathfrak{g}$ is semisimple. The subset

$$\Delta^+ := \{2\varepsilon_j, \varepsilon_j \pm \varepsilon_k \colon j < k, j, k = 1, \ldots, n\}$$

is a positive system with the root basis

$$\Pi = \{\varepsilon_1 - \varepsilon_2, \ldots, \varepsilon_{n-1} - \varepsilon_n, 2\varepsilon_n\}$$

(Example 11.17(c)). For $\alpha_j := \varepsilon_j - \varepsilon_{j+1}, j = 1, \ldots, n-1$ and $\alpha_n = 2\varepsilon_n$, the Cartan matrix is

$$(\alpha_j(\check{\alpha}_k))_{1\leq j,k\leq n} = \begin{pmatrix} 2 & -1 & 0 & \cdots & & & 0 \\ -1 & 2 & -1 & \cdots & & & 0 \\ 0 & & \ddots & & & & \vdots \\ \vdots & & -1 & 2 & -1 & 0 \\ & & & -1 & 2 & -1 \\ 0 & \cdots & 0 & 0 & -2 & 2 \end{pmatrix}.$$

Since the root basis for A_{n-1} is irreducible and ε_n is not orthogonal to $\varepsilon_{n-1} - \varepsilon_n$, we see that Π is irreducible and hence that $\mathfrak{sp}_{2n}(\mathbb{K})$ is simple.

Example 12.10. We consider the Lie algebra $\mathfrak{so}_{n,n+1}(\mathbb{K}) \subseteq \mathfrak{gl}_{2n+1}(\mathbb{K})$. In Example 8.8 we have seen that

$$\mathfrak{h} := \operatorname{span}\{E_{jj} - E_{j+n,j+n} \colon j = 1, \ldots, n\}$$

is a toral Cartan subalgebra and that

$$\Delta = \{\pm\varepsilon_j, \pm(\varepsilon_j \pm \varepsilon_k) \colon j \neq k, j,k = 1, \ldots, n\} = B_n$$

is the corresponding root system. For the coroots we obtain

$$\begin{aligned} (\varepsilon_j - \varepsilon_k)\check{} &= E_{jj} - E_{kk} - E_{j+n,j+n} + E_{k+n,k+n} \\ (\varepsilon_j + \varepsilon_k)\check{} &= E_{jj} + E_{kk} - E_{j+n,j+n} - E_{k+n,k+n}, \quad j \neq k. \end{aligned}$$

The subset

$$\Delta^+ := \{\varepsilon_j, \varepsilon_j \pm \varepsilon_k \colon j < k, j,k = 1, \ldots, n\}$$

is a positive system with root basis

$$\Pi = \{\varepsilon_1 - \varepsilon_2, \ldots, \varepsilon_{n-1} - \varepsilon_n, \varepsilon_n\}$$

(Example 11.17(b)). For $\alpha_j := \varepsilon_j - \varepsilon_{j+1}, j = 1, \ldots, n-1$ and $\alpha_n = \varepsilon_n$, the Cartan matrix is

$$(\alpha_j(\check{\alpha}_k))_{1\leq j,k\leq n} = \begin{pmatrix} 2 & -1 & 0 & \cdots & & & 0 \\ -1 & 2 & -1 & \cdots & & & 0 \\ 0 & & \ddots & & & & \vdots \\ \vdots & & -1 & 2 & -1 & 0 \\ & & & -1 & 2 & -2 \\ 0 & \cdots & 0 & 0 & -1 & 2 \end{pmatrix}.$$

With the same argument as for C_n, we see that Π is irreducible, so that $\mathfrak{o}_{n,n+1}(\mathbb{K})$ is a simple Lie algebra. For $n = 1$, we obtain

$$\mathfrak{o}_{1,2}(\mathbb{K}) = \mathfrak{g}(\varepsilon_1) \cong \mathfrak{sl}_2(\mathbb{K}).$$

If $\mathbb{K}$ is algebraically closed, then $\mathfrak{o}_{2n+1}(\mathbb{K}) \cong \mathfrak{o}_{n,n+1}(\mathbb{K})$ (Exercise 8.4), so that we also obtain the simplicity of the Lie algebra $\mathfrak{o}_{2n+1}(\mathbb{K})$ of skew-symmetric matrices of size $2n+1$.

Example 12.11. Finally we consider the Lie algebra $\mathfrak{o}_{n,n}(\mathbb{K})$ (Example 8.8) whose root decomposition can be obtained from $\mathfrak{o}_{n,n+1}(\mathbb{K})$ by restriction. We have the root set

$$\Delta = D_n = \{\pm(\varepsilon_j \pm \varepsilon_k) \colon j \neq k\}$$

and the coroots

$$(\varepsilon_j - \varepsilon_k)\check{} = E_{jj} - E_{kk} - E_{j+n,j+n} + E_{k+n,k+n}$$
$$(\varepsilon_j + \varepsilon_k)\check{} = E_{jj} + E_{kk} - E_{j+n,j+n} - E_{k+n,k+n}, \quad j \neq k.$$

The subset

$$\Delta^+ := \{\varepsilon_j \pm \varepsilon_k \colon j < k, j, k = 1, \ldots, n\}$$

is a positive system with the root basis

$$\Pi = \{\varepsilon_1 - \varepsilon_2, \ldots, \varepsilon_{n-1} - \varepsilon_n, \varepsilon_{n-1} + \varepsilon_n\}$$

(Example 11.17(d)). For $\alpha_j := \varepsilon_j - \varepsilon_{j+1}, j = 1, \ldots, n-1$ and $\alpha_n = \varepsilon_{n-1} + \varepsilon_n$, the Cartan matrix is

$$(\alpha_j(\check{\alpha}_k))_{1 \leq j,k \leq n} = \begin{pmatrix} 2 & -1 & 0 & & \cdots & & 0 \\ -1 & 2 & -1 & & \cdots & & \\ 0 & & \ddots & & & & \vdots \\ & & -1 & 2 & -1 & 0 & 0 \\ \vdots & & & -1 & 2 & -1 & -1 \\ & & & & -1 & 2 & 0 \\ 0 & & \cdots & & -1 & 0 & 2 \end{pmatrix}.$$

For $n \geq 3$, the root basis Π is indecomposable, but for $n = 2$ this is not the case. It follows that $\mathfrak{o}_{n,n}(\mathbb{K})$ is simple for $n \geq 3$, but not for $n = 2$. For $n = 1$ it is one-dimensional, hence also not simple. From the root decomposition one can further see that, for $n = 2$ and $\Pi = \{\alpha, \beta\}$, we have $|\Delta^+| = 2$ and

$$\mathfrak{g} \cong \mathfrak{g}(\alpha) \oplus \mathfrak{g}(\beta) \cong \mathfrak{sl}_2(\mathbb{K})^2 \cong \mathfrak{o}_{1,2}(\mathbb{K})^2.$$

If $\mathbb{K}$ is algebraically closed, then $\mathfrak{o}_{2n}(\mathbb{K}) \cong \mathfrak{o}_{n,n}(\mathbb{K})$ (Exercise 8.4), so that we obtain the simplicity of $\mathfrak{o}_{2n}(\mathbb{K})$ for $n \geq 3$.

We collect the results of this inspection in the following theorem:

Theorem 12.12. *For every field $\mathbb{K}$ of characteristic zero, the Lie algebras*

$$\mathfrak{sl}_{n+1}(\mathbb{K}), n \geq 1, \quad \mathfrak{sp}_{2n}(\mathbb{K}), n \geq 1, \quad \mathfrak{o}_{n,n+1}(\mathbb{K}), n \geq 1, \quad \mathfrak{o}_{n,n}(\mathbb{K}), n \geq 3$$

are simple and split. Moreover,

$$\mathfrak{sp}_2(\mathbb{K}) \cong \mathfrak{o}_{1,2}(\mathbb{K}) \cong \mathfrak{sl}_2(\mathbb{K}), \qquad \mathfrak{o}_{2,2}(\mathbb{K}) \cong \mathfrak{sl}_2(\mathbb{K})^2 \quad \text{and} \quad \mathfrak{o}_{1,1}(\mathbb{K}) \cong \mathbb{K}.$$

12.4 The Classification of Cartan Matrices

In the preceding subsection, we have seen that the root bases of the simple Lie algebras with a toral Cartan subalgebra $\mathfrak{h}$ are precisely the indecomposable ones. If $A = CM(\Pi)$ is the Cartan matrix of the root basis Π, then any orthogonal partition $\Pi = \Pi_1 \dot{\cup} \Pi_2$ corresponds to a description of A as a (2×2)-block matrix

$$A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}.$$

In this sense the irreducibility of Π corresponds to the "irreducibility" of the corresponding Cartan matrix in the sense that there is no reordering of the basis which turns A into a (2×2)-block matrix. With respect to this order Π_1 should come before Π_2 or vice versa.

Now let $\Pi = \{\alpha_1, \ldots, \alpha_\ell\}$ be a root basis and $A = (\alpha_i(\check{\alpha}_j))_{i,j=1,\ldots,\ell}$ the corresponding Cartan matrix. If $\alpha(\check{\beta})\beta(\check{\alpha}) \neq 0$, then

$$\frac{\alpha(\check{\beta})}{\beta(\check{\alpha})} = \frac{(\alpha, \alpha)}{(\beta, \beta)} = \frac{\|\alpha\|^2}{\|\beta\|^2}$$

is the ratio of the square lengths (Remark 11.7). If $\alpha(\check{\beta})\beta(\check{\alpha}) = 1$, then

$$\alpha(\check{\beta}) = -1 = \beta(\check{\alpha}),$$

because $\alpha(\check{\beta}) < 0$ and both roots have the same lenght (Remark 11.7). If $\alpha(\check{\beta})$ and $\beta(\check{\alpha})$ are different, then both roots have different lengths. In view of $\alpha(\check{\beta})\beta(\check{\alpha}) \in \{0, 1, 2, 3\}$ (Remark 11.7), at least one of the two factors is -1. For $\beta(\check{\alpha}) = -1$, we obtain

$$-\alpha(\check{\beta}) = \frac{\|\alpha\|^2}{\|\beta\|^2} \in \{1, 2, 3\}. \tag{29}$$

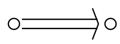
Definition 12.13. Now we associate a *Dynkin diagram* to the Cartan matrix. We consider the elements of the root basis as *vertices* of the diagram and connect two vertices α and β by $\alpha(\check{\beta})\beta(\check{\alpha}) \in \{0, 1, 2, 3\}$ *edges*. If α and β are connected and not of the same length, then we also add an arrow pointing from the longer to the shorter root. Recall that the ratio of the root lengths is determined via (29) by the number of connecting edges.

Here are some examples which, in view of Remark 11.7, cover all the 2-vertex configurations:

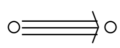
$$A_1 \times A_1 : \quad \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

$$A_2 : \quad \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$

$$B_2 : \quad \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix},$$



$$G_2 : \quad \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix},$$



Remark 12.14. Conversely, we can reconstruct from every Dynkin diagram the corresponding Cartan matrix:

(0) If α and β are not connected, then $\alpha(\check{\beta}) = \beta(\check{\alpha}) = 0$.

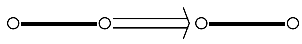
(1) If α and β are connected by a simple edge, then $\alpha(\check{\beta}) = \beta(\check{\alpha}) = -1$.

(2) If α and β are connected by a double edge with an arrow from α to β, then $\alpha(\check{\beta}) = -2$ and $\beta(\check{\alpha}) = -1$.

(3) If α and β are connected by a triple edge with an arrow from α to β, then $\alpha(\check{\beta}) = -3$ and $\beta(\check{\alpha}) = -1$.

Example 12.15. From the diagram

F_4

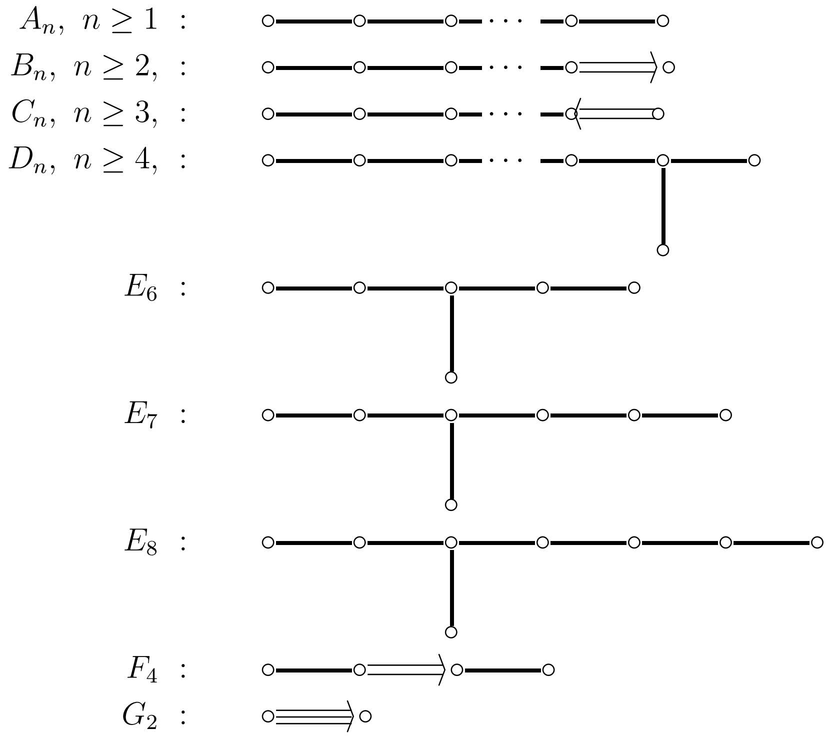


we thus obtain the Cartan matrix

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

The root basis Π is reducible if and only if the corresponding diagram is not *connected*. Then the maximal irreducible subsets of Π correspond to the connected components of the diagram. Since we know already that irreducible root bases correspond to simple Lie algebras (Proposition 12.7), we restrict our attention to connected Dynkin diagrams.

Theorem 12.16. *If $\mathfrak{g}$ is a finite dimensional simple Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ a toral Cartan subalgebra, then the Dynkin diagram with respect to any root basis Π of $\Delta(\mathfrak{g}, \mathfrak{h})$ is contained in the following list:*



The restriction of n for the different series are necessary to avoid overlaps. For instance, B_2 and C_2 correspond to the same diagram. Likewise D_3 and A_3 correspond to the same diagram and one could also put $E_5 := D_5$.

Theorem 12.17. (Existence of root systems) *Every Dynkin diagram corresponds to a root system.*

Proof. It suffices to verify the assertion for connected diagrams, for which the possibilities are listed in Theorem 12.16. For the classical series *A*-*D* we have already seen how to realize the corresponding root systems in some $\mathbb{R}^n$, which leads to the root bases described in Example 11.17:

$$\begin{aligned}
A_n \quad &: \quad \Pi = \Big\{\varepsilon_i - \varepsilon_{i+1} \colon i = 1, \ldots, n\Big\} \subseteq \mathbb{R}^{n+1}, \\
B_n \quad &: \quad \Pi = \Big\{\varepsilon_i - \varepsilon_{i+1} \colon i = 1, \ldots, n-1\Big\} \cup \{\varepsilon_n\} \subseteq \mathbb{R}^n, \\
C_n \quad &: \quad \Pi = \Big\{\varepsilon_i - \varepsilon_{i+1} \colon i = 1, \ldots, n-1\Big\} \cup \{2\varepsilon_n\} \subseteq \mathbb{R}^n, \\
D_n \quad &: \quad \Pi = \Big\{\varepsilon_i - \varepsilon_{i+1} \colon i = 1, \ldots, n-1\Big\} \cup \Big\{\varepsilon_{n-1} + \varepsilon_n\Big\} \subseteq \mathbb{R}^n.
\end{aligned}$$

This leaves only the diagrams of the expectional types *E*-*G*. Since E_6 and E_7 are subdiagrams of E_8, it suffices to find a realization of E_8, F_4 and G_2. This can be done by verifying that the following sets are root systems (in $\mathbb{R}^n$ with respect to the standard scalar product):

$$
\begin{array}{lcl}
E_8 & : & D_8 \cup \left\{ \frac{1}{2} \sum_{j=1}^{8} (-1)^{\eta_j} \varepsilon_j \colon \sum_{j=1}^{8} \eta_j \in 2\mathbb{Z} \right\} \subseteq \mathbb{R}^8 \\
F_4 & : & B_4 \cup \left\{ \frac{1}{2}(\pm\varepsilon_1 \pm \varepsilon_2 \pm \varepsilon_3 \pm \varepsilon_4) \right\} \subseteq \mathbb{R}^4 \text{ (signs independent)} \\
G_2 & : & A_2 \cup \left\{ \pm (2\varepsilon_1 - \varepsilon_2 - \varepsilon_3), \pm(2\varepsilon_2 - \varepsilon_1 - \varepsilon_3), \pm(2\varepsilon_3 - \varepsilon_1 - \varepsilon_2) \right\} \subseteq \mathbb{R}^3
\end{array}
$$

(see Exercise 11.3 for G_2). Accordingly, we obtain the root bases:

$$
\begin{array}{lcl}
E_8 & : & \left\{ \varepsilon_2 - \varepsilon_3, \ldots, \varepsilon_7 - \varepsilon_8, \varepsilon_7 + \varepsilon_8, \frac{1}{2}(\varepsilon_1 + \varepsilon_8 - \varepsilon_2 - \varepsilon_3 - \ldots - \varepsilon_7) \right\} \\
F_4 & : & \left\{ \varepsilon_2 - \varepsilon_3, \varepsilon_3 - \varepsilon_4, \varepsilon_4, \frac{1}{2}(\varepsilon_1 - \varepsilon_2 - \varepsilon_3 - \varepsilon_4) \right\} \\
G_2 & : & \left\{ \varepsilon_1 - \varepsilon_2, -2\varepsilon_1 + \varepsilon_2 + \varepsilon_3 \right\},
\end{array}
$$

and it is easy to verify that they belong to the listed diagrams. □

To complete the classification, one still has to prove that:

- That every pair $(\mathfrak{g}, \mathfrak{h})$ of a simple Lie algebra $\mathfrak{g}$ and a toral Cartan subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ determines a unique Dynkin diagram. To this end one first shows that for any other pair $(\mathfrak{g}, \mathfrak{h}')$, there exists an automorphism $\psi \in \mathrm{Aut}(\mathfrak{g})$ with $\varphi(\mathfrak{h}) = \mathfrak{h}'$ (Conjugacy of toral Cartan subalgebras). Next one shows that every element of the Weyl group $\mathcal{W}(\mathfrak{g}, \mathfrak{h})$ is induced by the action of an automorphism of $\mathfrak{g}$ preserving $\mathfrak{h}$ on the set of roots, and now the transitivity of the action of the Weyl group on the set of root bases (Theorem 11.23) shows that all root bases associated to a pair $(\mathfrak{g}, \mathfrak{h})$ lead to the same Dynkin diagrams (up to enumeration of their vertices).

- That every Dynking diagram actually arises from some Lie algebra. This is clear for the infinite series A-D, but not for types E-G. We have already seen above that all Dynking diagrams arise from a root system, but this is much less information than what is contained in the corresponding Lie algebra. There are two natural methods to achieve this goal. The geometric method consists in realizing them as Lie algebras of automorphism groups of suitable geometric structures. This is very interesting but has the disadvantage that it proceeds very much by case-by-case analysis. There is another approach to obtain the Lie algebras by defining them in terms of generators and relations (the *Serre relations*) (cf. Proposition 14.1). This approach has the advantage of being universal and it even points to the more general class of Kac–Moody Lie algebras which form an interesting class of infinite dimensional Lie algebras.

Notes on Part II

Cartan subalgebras actually occur first in the work of W. Killing who classified the finite dimensional simple complex Lie algebras (cf. [Kil89]). Unfortunately, Killing's work contained some serious gaps, concerning the basic properties of Cartan subalgebras. These were cleaned up later by Élie Cartan in his thesis [Ca94], and this is why they nowadays carry his name.

Serre's Theorem on the presentation of semisimple Lie algebras with a toral Cartan subalgebra can be extended to a construction of a semisimple Lie algebra from an abstract root system Δ with a basis $\Pi = \{\alpha_1, \ldots, \alpha_r\}$. Then we put $a_{ij} := \alpha_j(\check{\alpha}_i)$ and consider the Lie algebra $L(X, R)$ defined by the generators h_i, e_i, f_i, $i = 1, \ldots, r$ and the relations

$$[h_i, h_j] = 0, \quad [h_i, e_j] = a_{ij} e_j, \quad [h_i, f_j] = -a_{ij} f_j, \quad [e_i, f_j] = \delta_{ij} h_i$$

and

$$(\operatorname{ad} e_i)^{1-a_{ij}} e_j = 0, \quad (\operatorname{ad} f_i)^{1-a_{ij}} f_j = 0 \quad \text{for} \quad i \neq j.$$

In this context the main point is to show that $L(X, R)$ is a semisimple Lie algebra with the Cartan subalgebra $\mathfrak{h} = \operatorname{span}\{h_1, \ldots, h_r\}$ and a root system isomorphic to Δ. In the 1960s this description of the finite dimensional semisimple Lie algebras was the starting point for the theory of Kac–Moody–Lie algebras, which are defined by the same set of generators and relations for more general matrices $(a_{ij}) \in M_r(\mathbb{Z})$, called *generalized Cartan matrices.*

Part III
Representation Theory of Lie Algebras

In this final part of the lecture we address the classification of finite dimensional representations of a split semisimple Lie algebra $\mathfrak{g}$ of characteristic zero. We first introduce the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$.[15] It has the universal property that each representation of $\mathfrak{g}$ defines a unique algebra representation of $\mathcal{U}(\mathfrak{g})$, so that any $\mathfrak{g}$-module becomes a $\mathcal{U}(\mathfrak{g})$-module. We may thus translate freely between $\mathfrak{g}$-modules and $\mathcal{U}(\mathfrak{g})$-modules, which is often convenient. A key point is that, if a $\mathfrak{g}$-module V is generated by a vector $v \in V$, then $\mathfrak{g} \cdot v \neq V$ in general, but then v is also a generator of the corresponding $\mathcal{U}(\mathfrak{g})$-module, and this implies $V = \mathcal{U}(\mathfrak{g}) \cdot v$, which can be used to obtained finer information on V.

The Poincaré–Birkhoff–Witt (PBW) Theorem 13.7 that we discuss here provides crucial information on linear generating subsets of $\mathcal{U}(\mathfrak{g})$. We then proceed with the Highest Weight Theorem 15.18 (also called the Cartan–Weyl Theorem) which provides a classification of irreducible finite dimensional representations of split semisimple Lie algebras. This is the main result of the Cartan–Weyl Theory of simple modules of split semisimple Lie algebras. In view of Weyl's Theorem that any module of a Lie algebra is semisimple, the classification of the simple modules provides a complete picture of the finite dimensional representations.

13 The Universal Enveloping Algebra

Representing a Lie algebra by linear maps leads to a mapping of the Lie algebra into an associative algebra such that the Lie bracket turns into the commutator bracket. The main point of the enveloping algebra $\mathcal{U}(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$ is that every representation of $\mathfrak{g}$ on V factors through a homomorphism $\mathcal{U}(\mathfrak{g}) \to \mathrm{End}(V)$ of associative algebras.

Definition 13.1. Let $\mathfrak{g}$ be a Lie algebra. A pair $(\mathcal{U}(\mathfrak{g}), \sigma)$, consisting of a unital associative algebra $\mathcal{U}(\mathfrak{g})$ and a homomorphism $\sigma\colon \mathfrak{g} \to \mathcal{U}(\mathfrak{g})_L$ of Lie algebras, is called a (universal) *enveloping algebra* of $\mathfrak{g}$ if it has the following universal property. For each homomorphism $f\colon \mathfrak{g} \to \mathcal{A}_L$ of $\mathfrak{g}$ into the Lie algebra $\mathcal{A}_L$, where $\mathcal{A}$ is a unital associative algebra, there exists a unique homomorphism $\widetilde{f}\colon \mathcal{U}(\mathfrak{g}) \to \mathcal{A}$ of unital associative algebras with $\widetilde{f} \circ \sigma = f$.

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\forall f"] \arrow[d, "\sigma"'] & \mathcal{A} \\ \mathcal{U}(\mathfrak{g}) \arrow[ur, "\exists \widetilde{f}"'] & \end{tikzcd}$$

The universal property determines a universal enveloping algebra uniquely in the following sense:

Lemma 13.2. (Uniqueness of the enveloping algebra) *If $(\mathcal{U}(\mathfrak{g}), \sigma)$ and $(\widetilde{\mathcal{U}}(\mathfrak{g}), \widetilde{\sigma})$ are two enveloping algebras of the Lie algebra $\mathfrak{g}$, then there exists an isomorphism $f\colon \mathcal{U}(\mathfrak{g}) \to \widetilde{\mathcal{U}}(\mathfrak{g})$ of unital associative algebras satisfying $f \circ \sigma = \widetilde{\sigma}$.*

[15]In categorical terms: The enveloping algebra defines a functor from the category of Lie algebras to the category of unital associative algebras which is the adjoint of the forgetful functor $\mathcal{A} \mapsto \mathcal{A}_L = (\mathcal{A}, [\cdot, \cdot])$.

Proof. Since $\widetilde{\sigma}\colon \mathfrak{g} \to \widetilde{\mathcal{U}}(\mathfrak{g})_L$ is a homomorphism of Lie algebras, the universal property of the pair $(\mathcal{U}(\mathfrak{g}), \sigma)$ implies the existence of a unique algebra homomorphism

$$f\colon \mathcal{U}(\mathfrak{g}) \to \widetilde{\mathcal{U}}(\mathfrak{g}) \quad \text{with} \quad f \circ \sigma = \widetilde{\sigma}.$$

Similarly, the universal property of $(\widetilde{\mathcal{U}}(\mathfrak{g}), \widetilde{\sigma})$ implies the existence of an algebra homomorphism

$$g\colon \widetilde{\mathcal{U}}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g}) \quad \text{with} \quad g \circ \widetilde{\sigma} = \sigma.$$

Then $g \circ f\colon \mathcal{U}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g})$ is an algebra homomorphism with $(g \circ f) \circ \sigma = \sigma$, so that the uniqueness part of the universal property of $(\mathcal{U}(\mathfrak{g}), \sigma)$ yields $g \circ f = \mathrm{id}_{\mathcal{U}(\mathfrak{g})}$. We likewise get $f \circ g = \mathrm{id}_{\widetilde{\mathcal{U}}(\mathfrak{g})}$, showing that f is an isomorphism of unital algebras. $\square$

To prove the existence of an enveloping algebra, we recall some basic algebraic concepts. Let $\mathcal{A}$ be an associative algebra. A subspace J of $\mathcal{A}$ is called an *ideal* if

$$\mathcal{A}J \cup J\mathcal{A} \subseteq J.$$

Let M be a subset of $\mathcal{A}$. Since the intersection of a family of ideals is again an ideal, the intersection J_M of all ideals of $\mathcal{A}$ containing M is the smallest ideal of $\mathcal{A}$ containing M. It is called the *ideal generated by* M. If J is an ideal of $\mathcal{A}$, then the *factor algebra* $\mathcal{A}/J$ is the quotient vector space, endowed with the associative multiplication

$$(a_1 + J)(a_2 + J) := a_1 a_2 + J \quad \text{for} \quad a_1, a_2 \in \mathcal{A}.$$

13.1 Existence

Proposition 13.3. (Existence of an enveloping algebra) *Each Lie algebra* $\mathfrak{g}$ *has an enveloping algebra* $(\mathcal{U}(\mathfrak{g}), \sigma)$.

Proof. Let $\mathcal{T}(\mathfrak{g})$ be the tensor algebra of $\mathfrak{g}$ (Definition B.7) and consider the subset

$$M := \{x \otimes y - y \otimes x - [x, y] \in \mathcal{T}(\mathfrak{g})\colon x, y \in \mathfrak{g}\}.$$

Then

$$\mathcal{U}(\mathfrak{g}) := \mathcal{T}(\mathfrak{g})/J_M$$

is a unital associative algebra and

$$\sigma\colon \mathfrak{g} \to \mathcal{U}(\mathfrak{g}), \quad \sigma(x) := x + J_M,$$

is a linear map, satisfying

$$\sigma([x, y]) = [x, y] + J_M = x \otimes y - y \otimes x + J_M = \sigma(x)\sigma(y) - \sigma(y)\sigma(x),$$

so that σ is a homomorphism of Lie algebras $\mathfrak{g} \to \mathcal{U}(\mathfrak{g})_L$.

To verify the universal property for $(\mathcal{U}(\mathfrak{g}), \sigma)$, let $f\colon \mathfrak{g} \to \mathcal{A}_L$ be a homomorphism of Lie algebras, where $\mathcal{A}$ is a unital associative algebra. In view of the universal property of $\mathcal{T}(\mathfrak{g})$ (Lemma B.8), there exists an algebra homomorphism $\widehat{f}\colon \mathcal{T}(\mathfrak{g}) \to \mathcal{A}$ with $\widehat{f}(x) = f(x)$ for all $x \in \mathfrak{g}$. Then $M \subseteq \ker \widehat{f}$, and since $\ker \widehat{f}$ is an ideal of $\mathcal{T}(\mathfrak{g})$, we also have $J_M \subseteq \ker \widehat{f}$, so that $\widehat{f}$ factors through an algebra homomorphism

$$\widetilde{f}\colon \mathcal{U}(\mathfrak{g}) \to \mathcal{A} \quad \text{with} \quad \widetilde{f} \circ \sigma = f.$$

To see that $\widetilde{f}$ is unique, it suffices to note that $\sigma(\mathfrak{g})$ and $\mathbf{1}$ generate $\mathcal{U}(\mathfrak{g})$ as an associative algebra because $\mathfrak{g}$ and $\mathbf{1}$ generate $\mathcal{T}(\mathfrak{g})$ as an associative algebra. $\square$

Remark 13.4. The universal property of $(\mathcal{U}(\mathfrak{g}), \sigma)$ implies that each representation (π, V) of $\mathfrak{g}$ defined a representation $\widetilde{\pi}\colon \mathcal{U}(\mathfrak{g}) \to \mathrm{End}(V)$, which is uniquely determined by $\widetilde{\pi} \circ \sigma = \pi$. From the construction of $\mathcal{U}(\mathfrak{g})$ we also know that the algebra $\mathcal{U}(\mathfrak{g})$ is generated by $\sigma(\mathfrak{g})$. This implies that, for each $v \in V$, the subspace

$$\mathcal{U}(\mathfrak{g}) \cdot v \subseteq V$$

is the smallest subspace containing v and invariant under $\mathfrak{g}$, i.e., the $\mathfrak{g}$-submodule of V generated by v. Hence the enveloping algebra provides a tool to understand $\mathfrak{g}$-submodules of a $\mathfrak{g}$-module. But before we are able to use this tool effectively, we need some more information on the structure of $\mathcal{U}(\mathfrak{g})$.

13.2 The Poincaré–Birkhoff–Witt Theorem

Our next goal is a convenient description of a linearly generating subset of $\mathcal{U}(\mathfrak{g})$. Here $\mathfrak{g}$ needs not be finite dimensional. Let $(x_j)_{j \in J}$ be a linear basis of $\mathfrak{g}$ and assume that the index set I carries a total order $\leq$. For a k-tuple $I = (i_1, \ldots, i_k) \subseteq J^k$, we put $\xi_I := \xi_{i_1} \cdots \xi_{i_k}$. We write

$$\mathcal{U}_p(\mathfrak{g}) := \sum_{k \leq p} \sigma(\mathfrak{g})^k = \sum_{k \leq p} \mathrm{span}\{\xi_I \colon I \in J^k\}.$$

These subspaces satisfy

$$\mathcal{U}_p(\mathfrak{g})\mathcal{U}_q(\mathfrak{g}) \subseteq \mathcal{U}_{p+q}(\mathfrak{g}) \quad \text{for} \quad p, q \in \mathbb{N}_0.$$

Lemma 13.5. *Let $y_1, \ldots, y_p \in \mathfrak{g}$ and π be a permutation of $\{1, \ldots, p\}$, then*

$$\sigma(y_1) \cdots \sigma(y_p) - \sigma(y_{\pi(1)}) \cdots \sigma(y_{\pi(p)}) \in \mathcal{U}_{p-1}(\mathfrak{g}).$$

Proof. Since every permutation is a composition of transpositions of neighboring elements, it suffices to prove the claim for $\pi(j) = j$ for $j \not\in \{i, i+1\}$ and $\pi(i) = i + 1$. But then we have

$$\begin{aligned}
&\sigma(y_1) \cdots \sigma(y_p) - \sigma(y_{\pi(1)}) \cdots \sigma(y_{\pi(p)}) \\
&= \sigma(y_1) \cdots \sigma(y_{i-1}) \big(\sigma(y_i)\sigma(y_{i+1}) - \sigma(y_{i+1})\sigma(y_i)\big) \sigma(y_{i+2}) \cdots \sigma(y_p) \\
&= \sigma(y_1) \cdots \sigma(y_{i-1}) \sigma([y_i, y_{i+1}]) \sigma(y_{i+2}) \cdots \sigma(y_p) \in \mathcal{U}_{p-1}(\mathfrak{g}).
\end{aligned}$$

□

Lemma 13.6. *The vector space $\mathcal{U}_p(\mathfrak{g})$ is spanned by the ξ_I with increasing sequences I of length less than or equal to p. In particular, the elements of the form ξ_I with arbitrary finite increasing sequences I generate $\mathcal{U}(\mathfrak{g})$.*

Proof. It is clear that $\mathcal{U}_p(\mathfrak{g})$ is spanned by the elements ξ_I with $I \in J^k$, $k \leq p$, i.e., I is an arbitrary sequences of length less than or equal to p. By induction on p, the claim holds for $\mathcal{U}_{p-1}(\mathfrak{g})$. But since for an increasing rearrangement I' of the sequence I, we have $\xi_I - \xi_{I'} \in \mathcal{U}_{p-1}(\mathfrak{g})$ by Lemma 13.5, we also obtain the claim for $\mathcal{U}_p(\mathfrak{g})$. □

In the following we shall only need Lemma 13.6, but one can actually show much more:

Theorem 13.7. (Poincaré–Birkhoff–Witt Theorem (PBW)) *Let $\mathfrak{g}$ be a finite dimensional Lie algebra and $\{x_1, \ldots, x_n\}$ be a basis for $\mathfrak{g}$. Then*

$$\{\xi_1^{\mu_1} \cdots \xi_n^{\mu_n} \in \mathcal{U}(\mathfrak{g}) \mid \mu_k \in \mathbb{N} \cup \{0\}\}$$

is a basis for $\mathcal{U}(\mathfrak{g})$.

Exercises for Section 13

Exercise 13.1. Let $\mathfrak{g}$ be a finite dimensional Lie algebra and β be a nondegenerate symmetric bilinear form on $\mathfrak{g}$. Suppose that $x_1, \ldots, x_n$ is a basis for $\mathfrak{g}$ and let $x^1, \ldots, x^n \in \mathfrak{g}$ be the dual basis w.r.t. β, i.e., $\beta(x_i, x^j) = \delta_{ij}$.

(i) Show that the *Casimir element* $\Omega := \sum_{i=1}^n x^i x_i$ lies in the center of $\mathcal{U}(\mathfrak{g})$. [16]

(ii) Let $\mathfrak{g} = \mathfrak{so}_n(\mathbb{R})$. Show that:

 (a) $\beta(x, y) = -\frac{1}{2}\operatorname{tr}(xy) = \frac{1}{2}\operatorname{tr}(xy^\top)$ defines an invariant scalar product on $\mathfrak{so}_n(\mathbb{R})$.

 (b) For a β-orthonormal basis $x_1, \ldots, x_N$, we have $\Omega := \sum_{i=1}^N x_i^2 \in Z\big(\mathcal{U}(\mathfrak{so}_n(\mathbb{R}))\big)$.

 (c) The matrices $L_{ij} := E_{ij} - E_{ji}$, $i < j$, form an orthonormal basis of $\mathfrak{so}_n(\mathbb{R})$ w.r.t. β.

(iii) Show that the *operators of angular momentum*

$$x_i \frac{\partial}{\partial x_j} - x_j \frac{\partial}{\partial x_i} : C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n), \qquad i, j = 1, \ldots, n,$$

generate a Lie algebra which is isomorphic to $\mathfrak{so}_n(\mathbb{R})$. Hint: See Exercise 1.3(v) for the Lie brackets of these operators.

(iv) The *Laplace operator* $\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$ commutes with the angular momentum operators. [17]

(v) Show that there is a nondegenerate symmetric invariant bilinear form on the oscillator algebra. Hence such forms do not only exist on semisimple Lie algebras.

Exercise 13.2. A function $f \in C^\infty(\mathbb{R}^n)$ is called *harmonic* if $\Delta(f) = 0$ for the Laplace operator $\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$. Show that the subspace $H \subseteq C^\infty(\mathbb{R}^n)$ of the harmonic functions is invariant under the angular momentum operators (cf. Exercise 13.1).

14 Generators and Relations for Semisimple Lie Algebras

In this section we shall use the root decomposition of a semisimple Lie algebra to find a description by generators and relations.

14.1 A Generating Set for Semisimple Lie Algebras

Proposition 14.1. *Let $\mathfrak{g}$ be a semisimple Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ a toral Cartan subalgebra. Fix a positive system $\Delta^+ \subseteq \Delta$ and let $\Pi \subseteq \Delta^+$ be the set of simple roots. For each $\alpha \in \Pi$, we fix a corresponding $\mathfrak{sl}_2$-triple $(h_\alpha, e_\alpha, f_\alpha)$. Then the following assertions hold:*

(i) *The subspace $\mathfrak{n} := \sum_{\beta \in \Delta^+} \mathfrak{g}_\beta$ is a nilpotent subalgebra generated by $\{e_\alpha \colon \alpha \in \Pi\}$.*

---

[16] Here we identify x with $\eta(x)$, so that we consider $\mathfrak{g}$ as a subset of $\mathcal{U}(\mathfrak{g})$.

[17] The Laplacian is NOT the Casimir operator of the Lie algebra generated by the angular momentum operators; it is the Casimir operator of the abelian Lie algebra generated by the operators $\frac{\partial}{\partial x_j}$.

(ii) *The Lie algebra* $\mathfrak{g}$ *is generated by* $\{h_\alpha, e_\alpha, f_\alpha \colon \alpha \in \Pi\}$. *These elements satisfy the relations*

$$[h_\alpha, h_\beta] = 0, \quad [h_\alpha, e_\beta] = \beta(\check{\alpha})e_\beta, \quad [h_\alpha, f_\beta] = -\beta(\check{\alpha})f_\beta, \quad [e_\alpha, f_\beta] = \delta_{\alpha,\beta}h_\alpha \tag{30}$$

and we further have

$$(\operatorname{ad} e_\alpha)^{1-\beta(\check{\alpha})}e_\beta = 0, \quad (\operatorname{ad} f_\alpha)^{1-\beta(\check{\alpha})}f_\beta = 0 \quad \textit{for} \quad \alpha \neq \beta \ \textit{ in } \ \Pi. \tag{31}$$

The relations (30) and (31) are called the *Serre relations.*

Proof. (i) If $\beta, \gamma \in \Delta^+$, then either $\beta + \gamma \in \Delta^+$ or $\beta + \gamma$ is not a root. Hence $\mathfrak{n}$ is a subalgebra of $\mathfrak{g}$. Pick $x_0 \in \mathfrak{h}_{\mathbb{Q}}$ with $\Delta^+ = \{\beta \in \Delta \colon \beta(x_0) > 0\}$ and let

$$m := \min\{\beta(x_0) \colon \beta \in \Delta^+(x_0)\} \quad \text{ and } \quad M := \max\{\beta(x_0) \colon \beta \in \Delta^+(x_0)\},$$

For any $N \in \mathbb{N}$ with $Nm > M$ we then have $C^N(\mathfrak{n}) = \{0\}$, showing that $\mathfrak{n}$ is nilpotent.

In the proof of the Generation Theorem 12.6, we have already seen that $\mathfrak{n}$ is generated by the root spaces $\mathfrak{g}_\alpha = \mathbb{K}e_\alpha$, $\alpha \in \Pi$, and that $\mathfrak{g}$ is generated by the subalgebras $\mathfrak{g}(\alpha)$, $\alpha \in \Pi$.

(ii) It remains to verify the Serre relations. Since $h_\alpha = \check{\alpha}$, the first three relations are trivial, and the fact that $\alpha - \beta \not\in \Delta$ for $\alpha \neq \beta$ in Π implies that $[e_\alpha, f_\beta] = \{0\}$ in this case.

If $\alpha \neq \beta$, then we consider the α-string through β. Since $\beta - \alpha \not\in \Delta$, it is of the form $\{\beta + j\alpha \colon 0 \leq j \leq q\}$, where $q = -\beta(\check{\alpha})$. This implies that $\beta + (1 - \beta(\check{\alpha}))\alpha \not\in \Delta$. As a consequence, $(\operatorname{ad} e_\alpha)^{1-\beta(\check{\alpha})}e_\beta = (\operatorname{ad} f_\alpha)^{1-\beta(\check{\alpha})}f_\beta = 0$.

The first relation in (31) is obtained with similar arguments, applied to the $\mathfrak{sl}_2$-triple $(-h_\alpha, f_\alpha, e_\alpha)$ and the $\mathfrak{g}(\alpha)$-submodule generated by e_β. $\square$

Example 14.2. We have seen in Example 8.7 how to find a natural root decomposition of the Lie algebra $\mathfrak{sl}_n(\mathbb{K})$ with respect to the Cartan subalgebra $\mathfrak{h}$ of diagonal matrices. In the root system

$$\Delta = \{\varepsilon_j - \varepsilon_k : 1 \leq j \neq k \leq n\},$$

the subset

$$\Delta^+ = \{\varepsilon_j - \varepsilon_k : 1 \leq j < k \leq n\}$$

is a natural positive system with root basis $\Pi = \{\varepsilon_1 - \varepsilon_2, \ldots, \varepsilon_{n-1} - \varepsilon_n\}$. Then

$$\mathfrak{n} = \sum_{\alpha \in \Delta^+} \mathfrak{g}_\alpha = \operatorname{span}\{E_{jk} \colon j < k\}$$

is the Lie algebra of strictly upper triangular matrices. It is generated by the root vectors $E_{j,j+1}$, $j = 1, \ldots, n-1$. For each pair of indices $j \neq k$, we have

$$\mathfrak{g}(\varepsilon_j - \varepsilon_k) = \operatorname{span}\{E_{jk}, E_{kj}, E_{jj} - E_{kk}\} \cong \mathfrak{sl}_2(\mathbb{K}),$$

and the subalgebras

$$\mathfrak{g}(\varepsilon_1 - \varepsilon_2), \quad \ldots, \quad \mathfrak{g}(\varepsilon_{n-1} - \varepsilon_n)$$

sitting on the diagonal, generate $\mathfrak{sl}_n(\mathbb{K})$. Moreover, $\mathfrak{sl}_n(\mathbb{K})$ is also generated by the $2(n-1)$ element: $E_{j,j+1}, E_{j+1,j}$, $j = 1, \ldots, n-1$.

15 Highest Weight Representations

We know already from Weyl's Theorem 5.26 on Complete Reducibility that any finite dimensional module over a semisimple Lie algebra $\mathfrak{g}$ is semisimple. This reduces the classification of finite dimensional modules to the classification of simple ones. In this section, we address this problem for the class of those semisimple Lie algebras which are *split*, i.e., contain a toral Cartan subalgebra. Note that $\mathfrak{sl}_n(\mathbb{K})$ and in particular any semisimple Lie algebra over an algebraically closed field is split.

Throughout this section, $\mathfrak{g}$ denotes a semisimple Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ a toral Cartan subalgebra. For $\lambda \in \mathfrak{h}^*$ and a representation (π, V) of $\mathfrak{g}$, we write

$$V_\lambda = V_\lambda(\mathfrak{h}) = \{v \in V \colon (\forall h \in \mathfrak{h})\ \pi(h)(v) = \lambda(h)v\}$$

for the corresponding weight space in V and $\mathcal{P}(V) := \{\lambda \in \mathfrak{h}^* \colon V_\lambda \neq \{0\}\}$ for the set of $\mathfrak{h}$-weights of V. We simply write $\Delta := \Delta(\mathfrak{g}, \mathfrak{h})$ for the set of roots and $\mathfrak{g}_\alpha := \mathfrak{g}_\alpha(\mathfrak{h})$, $\alpha \in \Delta$, for the root spaces.

Proposition 15.1. *If* $\dim V < \infty$, *then* $\mathfrak{h}$ *acts by diagonalizable operators on* V *and* V *is the direct sum of its weight spaces. All weights take rational values on* $\mathfrak{h}_{\mathbb{Q}}$.

Proof. In view of Lemma 10.12, $\mathfrak{h}$ is spanned by the coroots $h_\alpha = \check{\alpha}$. Since $\mathfrak{h}$ is abelian, it therefore suffices to see that for each root $\alpha \in \Delta$, the element $h_\alpha \in \mathfrak{h}$ is diagonalizable on V. Since the $\mathfrak{g}$-representation on V restricts to a representation of

$$\mathfrak{g}(\alpha) = \mathfrak{g}_\alpha + \mathfrak{g}_{-\alpha} + \mathbb{K}h_\alpha \cong \mathfrak{sl}_2(\mathbb{K})$$

on V, it suffices to apply Proposition 9.7. Moreover, we see that all eigenvalues of h_α are integral, which implies that each weight takes only rational values on the real subspace $\mathfrak{h}_{\mathbb{Q}}$. $\square$

Definition 15.2. Let $\Delta^+ \subseteq \Delta$ be a positive system of roots. (cf. Theorem 11.16). Then

$$\mathfrak{b} := \mathfrak{h} + \mathfrak{n} \cong \mathfrak{h} \ltimes \mathfrak{n}, \qquad \mathfrak{n} := \sum_{\beta \in \Delta^+} \mathfrak{g}_\beta$$

is a solvable subalgebra of $\mathfrak{g}$ because it is of the form $\mathfrak{b} = \mathfrak{n} \rtimes \mathfrak{h}$ for a nilpotent Lie algebra $\mathfrak{n}$ (Proposition 14.1). Subalgebras of this type are called *standard Borel subalgebras* with respect to $\mathfrak{h}$.

15.1 Highest Weights

Definition 15.3. A $\mathfrak{g}$-module V is called a *module with highest weight* $\lambda \in \mathfrak{h}^*$ if there is a $\mathfrak{b}$-invariant line $\mathbb{K}v \in V$ with

$$h \cdot v = \lambda(h)v \quad \text{for} \quad h \in \mathfrak{h},$$

and v generates the $\mathfrak{g}$-module V (i.e., V is the smallest submodule containing v). Then λ is called the *highest weight* and the nonzero elements of the generating line $\mathbb{K}v$ are called *highest weight vectors*.

Remark 15.4. Since $[\mathfrak{b}, \mathfrak{b}] = \mathfrak{n}$, all one-dimensional representations of $\mathfrak{b}$ vanish on $\mathfrak{n}$. This implies that an element $v \in V_\lambda(\mathfrak{h})$ is a $\mathfrak{b}$-eigenvector if and only if

$$(\forall \alpha \in \Delta^+)\ \pi(\mathfrak{g}_\alpha)v = \mathfrak{g}_\alpha \cdot v = \{0\}.$$

Proposition 15.5. *Let $\Delta^+ \subseteq \Delta$ be a positive system. Then each finite dimensional simple $\mathfrak{g}$-module is a highest weight module.*

Proof. Let V be a simple $\mathfrak{g}$-module. In view of Proposition 15.1, V is a direct sum of its weight spaces $V = \bigoplus_{\alpha\in\mathcal{P}(V)} V_\alpha$. Since V is finite dimensional, the set $\mathcal{P}(V)$ of weights is finite.

Pick $x_0 \in \mathfrak{h}_\mathbb{Q}$ with

$$\Delta^+ = \Delta^+(x_0) = \{\alpha \in \Delta\colon \alpha(x_0) > 0\}.$$

Then $\mathcal{P}(V)(x_0) \subseteq \mathbb{Q}$ (Proposition 15.1) and we can pick a $\lambda \in \mathcal{P}(V)$ such that $\lambda(x_0)$ is maximal. Let $v \in V_\lambda \setminus \{0\}$. For each $\alpha \in \Delta^+$, we then have $\mathfrak{g}_\alpha \cdot v \subseteq V_{\lambda+\alpha}$, but the choice of λ implies that $V_{\lambda+\alpha} = \{0\}$. Hence v is a $\mathfrak{b}$-eigenvector of weight λ. Since V is simple, it is generated by v. □

Remark 15.6. If $\mathbb{K}$ is algebraically closed, then we can also use Lie's Theorem 4.9 to see that a simple $\mathfrak{g}$-module V contains an eigenvector for the solvable Lie algebra $\mathfrak{b}$, hence is a highest weight module.

For $\beta \in \Delta^+ = \{\beta_1, \ldots, \beta_m\}$, we choose an $\mathfrak{sl}_2$-triple $(h_\beta, e_\beta, f_\beta)$ as in the $\mathfrak{sl}_2$-Theorem 10.9. As for abstract root systems (Definition 11.20), we define a partial order $\prec$ on $\mathfrak{h}^*$ by

$$\lambda \prec \mu \quad :\iff \quad \mu - \lambda \in \mathbb{N}_0[\Delta^+] := \sum_{\beta\in\Delta^+} \mathbb{N}_0\beta.$$

Let $\Pi \subseteq \Delta^+$ be the corresponding set of simple roots (Theorem 11.16).

The following theorem describes some properties of highest weight modules which are not necessarily finite dimensional.

Theorem 15.7. *Let V be a $\mathfrak{g}$-module with highest weight λ and $0 \neq v_\lambda \in V_\lambda$ a highest weight vector. Then*

(i) $V = \operatorname{span}\{f_{\beta_1}^{i_1} \cdots f_{\beta_m}^{i_m} \cdot v_\lambda \mid i_j \in \mathbb{N}_0\}$ *for $\Delta^+ = \{\beta_1, \ldots, \beta_m\}$. In particular, V is the direct sum of its weight spaces.*

(ii) $\mathcal{P}(V) \subseteq \lambda - \mathbb{N}_0[\Delta^+] = \lambda - \mathbb{N}_0[\Pi]$.

(iii) $\dim V_\mu < \infty$ *for all $\mu \in \mathcal{P}(V)$.*

(iv) $\dim V_\lambda = 1$.

(v) *V contains exactly one maximal proper $\mathfrak{g}$-submodule $V_{\max}$ and $V/V_{\max}$ is the unique simple quotient module of V.*

(vi) *Every nonzero module quotient of V is a module with highest weight λ.*

Proof. (i) Let $\Pi = \{\alpha_1, \ldots, \alpha_r\}$ and $\Delta^+ = \{\beta_1, \ldots, \beta_m\}$. Then

$$f_{\beta_1}, \ldots, f_{\beta_m}, h_{\alpha_1}, \ldots, h_{\alpha_r}, e_{\beta_1}, \ldots, e_{\beta_m}$$

is a basis for $\mathfrak{g}$, to which we apply Lemma 13.6. Then the claim follows from

$$h_{\alpha_1}^{j_1} \cdots h_{\alpha_r}^{j_r} e_{\beta_1}^{\ell_1} \cdots e_{\beta_m}^{\ell_m} \cdot v_\lambda \subseteq \mathcal{U}(\mathfrak{b})v_\lambda \subseteq \mathbb{K}v_\lambda$$

and

$$V = \mathcal{U}(\mathfrak{g}) \cdot v_\lambda = \mathcal{U}(\overline{\mathfrak{n}})\mathcal{U}(\mathfrak{b}) \cdot v_\lambda = \mathcal{U}(\overline{\mathfrak{n}}) \cdot v_\lambda \quad \text{for} \quad \overline{\mathfrak{n}} = \sum_{\beta\in\Delta^+} \mathfrak{g}_{-\beta}$$

(cf. Remark 13.4).

(ii) In view of (i), V is spanned by vectors of the form $f_{\beta_1}^{i_1} \cdots f_{\beta_m}^{i_m} \cdot v_\lambda$, which are weight vectors of weight $\lambda - \sum_{\ell=1}^m i_\ell \beta_\ell$ by Lemma 8.10. Assertion (ii) now follows, since the positive roots can be written as sums of simple roots.

(iii) For $\mu \in \mathfrak{h}^*$, there are only finitely many vectors of the form $f_{\beta_1}^{i_1} \cdots f_{\beta_m}^{i_m} \cdot v_\lambda$ for which $\lambda - \sum_{\ell=1}^m i_\ell \beta_\ell$ equals μ. In fact, if $x_0 \in \mathfrak{h}_{\mathbb{Q}}$ satisfies $\beta_j(x_0) > 0$ for each j, then $\mu = \lambda - \sum_{l=1}^m i_\ell \beta_\ell$ yields

$$\sum_{\ell=1}^m i_\ell \beta_\ell(x_0) = (\lambda - \mu)(x_0),$$

and there are only finitely many solutions $(i_1, \ldots, i_m) \in \mathbb{N}_0^m$ of this equation because

$$i_\ell \leq \frac{(\lambda - \mu)(x_0)}{\beta_\ell(x_0)}, \qquad \ell = 1, \ldots, m.$$

(iv) The equality $\lambda = \lambda - \sum_{\ell=1}^m i_\ell \beta_\ell$ is only possible for $i_1 = \ldots = i_m = 0$.

(v) By Proposition 8.12, every submodule $W \subseteq V$ is adapted to the weight decomposition. Since V_λ is one-dimensional and generates V, it follows that every proper submodule $W \subseteq V$ is contained in $\sum_{\mu \neq \lambda} V_\mu$. Therefore the union of all proper submodules is still proper, and hence, a maximal proper submodule $V_{\max}$ exists. The quotient module $V/V_{\max}$ is simple, since for every nontrivial submodule $W \subseteq V/V_{\max}$, its inverse image W' in V would be a proper submodule of V, strictly containing $V_{\max}$. Conversely, every submodule W of V, for which V/W is simple (and nonzero), is a maximal submodule, hence equal to $V_{\max}$.

(vi) This is obvious. □

Corollary 15.8. *If V is a simple highest weight module, then V contains only one $\mathfrak{b}$-invariant line.*

Proof. Let $\mathbb{K}v_\lambda$ be a $\mathfrak{b}$-invariant line. Then v_λ is a weight vector for some weight μ and v_λ generates the simple module V (each simple module is generated by each nonzero element). Hence $\mathcal{P}(V) \subseteq \mu - \mathbb{N}_0[\Delta^+]$. If λ is the highest weight of V, we also have $\mathcal{P}(V) \subseteq \lambda - \mathbb{N}_0[\Delta^+]$, which leads to

$$\lambda \prec \mu \prec \lambda,$$

and hence to $\lambda = \mu$. Finally, Theorem 15.7(iv) implies that V_λ is one-dimensional, which completes the proof. □

Proposition 15.9. *Two simple $\mathfrak{g}$-modules with the same highest weight λ are isomorphic.*

Proof. Let V and W be two such modules. We choose nonzero elements $v_\lambda \in V_\lambda$ and $w_\lambda \in W_\lambda$. Set $M := V \oplus W$ and $m := v_\lambda + w_\lambda$. Then $\mathbb{K}m$ is a $\mathfrak{b}$-invariant line and the submodule $M' := \mathcal{U}(\mathfrak{g}) \cdot m$ of M generated by m is a module with highest weight λ. Let $\mathrm{pr}_V \colon M' \to V$ and $\mathrm{pr}_W \colon M' \to W$ be the canonical projections with respect to the direct sum $V \oplus W$. Then both, pr_V and pr_W, are homomorphisms of $\mathfrak{g}$-modules. From $\mathrm{pr}_V(m) = v_\lambda$ and $\mathrm{pr}_W(m) = w_\lambda$ we derive that pr_V and pr_W are surjective. Therefore, we must have $\ker \mathrm{pr}_V = M'_{\max} = \ker \mathrm{pr}_W$ by Theorem 15.7(v), and this implies $V \cong M'/M'_{\max} \cong W$. □

Definition 15.10. (Verma modules) Let $\mathfrak{g}$ be a semisimple Lie algebra, $\mathfrak{h} \subseteq \mathfrak{g}$ a toral Cartan subalgebra, and $\mathfrak{b} = \mathfrak{h} + \sum_{\beta \in \Delta^+} \mathfrak{g}_\beta$ the Borel subalgebra of $\mathfrak{g}$ corresponding to a positive system Δ^+ of Δ. For $\lambda \in \mathfrak{h}^*$, we extend λ to a linear functional $\widehat{\lambda}$ on $\mathfrak{b}$ vanishing on all root spaces $\mathfrak{g}_\alpha$, $\alpha \in \Delta^+$. Then

$$[\mathfrak{b}, \mathfrak{b}] = \sum_{\beta \in \Delta^+} \mathfrak{g}_\beta \subseteq \ker \widehat{\lambda}$$

implies that $\widehat{\lambda}\colon \mathfrak{b} \to \mathbb{K} \cong \mathfrak{gl}_1(\mathbb{K})$ is a homomorphism of Lie algebras, hence defines a one-dimensional representation of $\mathfrak{b}$. The Lie algebra homomorphism λ further extends to a unital algebra homomorphism $\widetilde{\lambda}\colon \mathcal{U}(\mathfrak{b}) \to \mathbb{K}$.

In the following we shall use the notation

$$AB := \operatorname{span}\{ab\colon a \in A, n \in B\}$$

for subsets A, B of an associative algebra. In this sense, we define

$$M(\lambda) := M(\lambda, \Delta^+) := \mathcal{U}(\mathfrak{g})/L_\lambda,$$

where

$$L_\lambda := \mathcal{U}(\mathfrak{g})\{b - \widehat{\lambda}(b)\mathbf{1}\colon b \in \mathfrak{b}\}$$

is the left ideal of $\mathcal{U}(\mathfrak{g})$ generated by the elements of the form $b - \widehat{\lambda}(b)\mathbf{1}$, $b \in \mathfrak{b}$. The module $M(\lambda)$ is called the *Verma module* of highest weight λ. We write $[D], D \in \mathcal{U}(\mathfrak{g})$, for its elements. Since $M(\lambda)$ is a quotient by a left ideal of $\mathcal{U}(\mathfrak{g})$, it carries a natural $\mathcal{U}(\mathfrak{g})$-module structure, hence in particular a $\mathfrak{g}$-module structure.

To see that $M(\lambda)$ is non-zero, we use the Poincaré–Birkhoff–Witt Theorem 13.7 to see that, for $\overline{\mathfrak{n}} = \sum_{\alpha \in \Delta^+} \mathfrak{g}_{-\alpha}$, the multiplication map

$$\mathcal{U}(\overline{\mathfrak{n}}) \otimes \mathcal{U}(\mathfrak{b}) \to \mathcal{U}(\mathfrak{g}), \quad d_1 \otimes d_2 \mapsto d_1 d_2$$

is a linear bijection. Therefore $L_\lambda = \mathcal{U}(\mathfrak{n}) \cdot \ker \widetilde{\lambda}$ is a proper subspace of $\mathcal{U}(\mathfrak{g})$ and this means that $M(\lambda) \neq \{0\}$.

The $\mathfrak{g}$-module $M(\lambda)$ is indeed a highest weight module of highest weight λ because $v_\lambda := [\mathbf{1}]$ satisfies for $b \in \mathfrak{b}$:

$$b \cdot v_\lambda = b \cdot [\mathbf{1}] = [b] = [\lambda(b)\mathbf{1}] = \lambda(b)[\mathbf{1}] = \lambda(b)v_\lambda,$$

and

$$\mathcal{U}(\mathfrak{g}) \cdot [\mathbf{1}] = [\mathcal{U}(\mathfrak{g})] = M(\lambda).$$

Using Theorem 15.7, we obtain a simple highest weight module with highest weight λ as

$$L(\lambda) := L(\lambda, \Delta^+) := M(\lambda, \Delta^+)/M(\lambda, \Delta^+)_{\max}.$$

Combining Proposition 15.9 with the preceding lemma, we obtain:

Proposition 15.11. *Fix a positive system Δ^+ of Δ. Then, for every $\lambda \in \mathfrak{h}^*$, there exists a simple highest weight module $L(\lambda, \Delta^+)$ with highest weight λ which is unique up to isomorphism.*

15.2 Classification of Finite Dimensional Simple Modules

We want to characterize the linear functionals on $\mathfrak{h}$ which occur as highest weights of simple $\mathfrak{g}$-modules.

Definition 15.12. A linear functional $\lambda \in \mathfrak{h}^*$ is said to be *integral* if

$$\lambda(\check{\alpha}) \in \mathbb{Z} \quad \text{for} \quad \alpha \in \Delta,$$

and it is called *dominant* with respect to the positive system Δ^+ if

$$\lambda(\check{\alpha}) \geq 0 \quad \text{for} \quad \alpha \in \Delta^+.$$

We denote the set of all integral functionals on $\mathfrak{h}$ by $\mathcal{P}$ (the *weight lattice*) , and the set of all dominant integral functionals by $\mathcal{P}^+$.

Let $\Pi \subseteq \Delta^+$ be the set of simple roots, which is a basis for $\mathfrak{h}^*$. Since the set $\check{\Pi} = \{\check{\alpha} \colon \alpha \in \Pi\}$ is a basis for the dual root system $\check{\Delta} \subseteq \mathfrak{h}_{\mathbb{R}}$ (Proposition 11.24),

$$\mathcal{P} = \{\lambda \in \mathfrak{h}^* \colon (\forall \alpha \in \Pi)\ \lambda(\check{\alpha}) \in \mathbb{Z}\} \cong \mathbb{Z}^r$$

and

$$\mathcal{P}^+ = \{\lambda \in \mathfrak{h}^* \colon (\forall \alpha \in \Pi)\ \lambda(\check{\alpha}) \in \mathbb{N}_0\} \cong \mathbb{N}_0^r.$$

Remark 15.13. The Weyl group of the root system Δ can be identified with the subgroup $\mathcal{W} \subseteq \mathrm{GL}(\mathfrak{h}^*)$ generated by the reflections

$$\sigma_\alpha(\lambda) = \lambda - \lambda(\check{\alpha})\alpha,$$

and this formula immediately implies that the weight lattice $\mathcal{P}$ is invariant under $\mathcal{W}$. Theorem 11.23 implies that for each $\nu \in \mathcal{P}$, there exists a $w \in \mathcal{W}$ with $w(\nu) \in \mathcal{P}^+$, i.e.,

$$\mathcal{P} = \mathcal{W}\mathcal{P}^+.$$

Proposition 15.14. *Let V be a finite dimensional $\mathfrak{g}$-module. Then $\mathcal{P}_V \subseteq \mathcal{P}$ and if λ is a highest weight of V with respect to Δ^+, then $\lambda \in \mathcal{P}^+$.*

Proof. Let $\alpha \in \Delta^+$ and $\mathfrak{g}(\alpha) \cong \mathfrak{sl}_2(\mathbb{K})$ be the corresponding 3-dimensional subalgebra of $\mathfrak{g}$. If $\mu \in \mathcal{P}_V$ is a weight of V, then Proposition 9.7 implies that $\mu(\check{\alpha}) \in \mathbb{Z}$, so that $\mu \in \mathcal{P}$. If $v_\lambda \in V_\lambda(\mathfrak{h})$ is a highest weight vector, then $\mathfrak{g}_\alpha \cdot v_\lambda = \{0\}$ implies that $\lambda(\check{\alpha}) \in \mathbb{N}_0$ (Proposition 9.3). □

Lemma 15.15. *Let V be a $\mathfrak{g}$-module and $\mathcal{W} = \mathcal{W}(\Delta)$ the Weyl group of Δ. If, for each $\alpha \in \Pi$, V is a locally finite $\mathfrak{g}(\alpha)$-module, i.e., a union of finite dimensional submodules, then each weight μ of V satisfies*

$$\dim V_\mu = \dim V_{w(\mu)} \quad \textit{for} \quad w \in \mathcal{W}.$$

In particular, $\mathcal{W}\mathcal{P}_V = \mathcal{P}_V$.

Proof. Since $\mathcal{W}$ is generated by the reflections σ_α, $\alpha \in \Pi$, it suffices to prove the assertion for $w = \sigma_\alpha$. Let $\rho \colon \mathfrak{g} \to \mathfrak{gl}(V)$ denote the representation of $\mathfrak{g}$ on V and $(h_\alpha, e_\alpha, f_\alpha)$ an $\mathfrak{sl}_2$-triple corresponding to $\alpha \in \Pi$. According to our hypothesis, each $v \in V$ is contained in a finite dimensional $\mathfrak{g}(\alpha)$-submodule, so that $\rho(h_\alpha)$ is diagonalizable by Proposition 9.7. Let $\mathfrak{h}_\alpha := \ker \alpha \subseteq \mathfrak{h}$, so that we have the direct sum decomposition $\mathfrak{h} = \mathfrak{h}_\alpha \oplus \mathbb{K}\check{\alpha}$. Then the weight space V_μ is contained in the subspace

$$W := V_{\mu|_{\mathfrak{h}_\alpha}}(\mathfrak{h}_\alpha) = \sum_{\nu - \mu \in \mathfrak{h}_\alpha^\perp = \mathbb{K}\alpha} V_\nu = \sum_{c \in \mathbb{K}} V_{\mu + c\alpha}$$

which is invariant under $\mathfrak{g}(\alpha)$ because $\mathfrak{g}(\alpha)$ commutes with $\mathfrak{h}_\alpha$ (cf. Exercise 4.1). Next we note that

$$\sigma_\alpha(\mu) = \mu - \mu(\check{\alpha})\alpha \in \mu + \mathbb{K}\alpha,$$

so that $V_{\sigma_\alpha(\mu)} \subseteq W$. As $V_{\mu+c\alpha}$ is the $\mu(h_\alpha) + 2c$-eigenspace of $\rho(h_\alpha)$ in W and $(\sigma_\alpha\mu)(h_\alpha) = -\mu(h_\alpha)$, the assertion follows from Proposition 9.7(ii). Note that this proposition can be used here because W is a union of finite dimensional $\mathfrak{g}(\alpha)$-submodules. □

Lemma 15.16. *If $\lambda \in \mathcal{P}^+$, then the set $\{\mu \in \mathcal{P}^+ \colon \mu \prec \lambda\}$ is finite.*

Proof. We consider Δ as a root system realized in a euclidean vector space $E \cong \mathrm{Hom}(\mathfrak{h}_{\mathbb{Q}}, \mathbb{R})$ (cf. Proposition 11.4). Then

$$\begin{aligned}\mathcal{P} &= \{\mu \in E \colon (\forall \alpha \in \Pi)\, (\mu, \check{\alpha}) \in \mathbb{Z}\} \\ \supseteq \mathcal{P}^+ &= \{\mu \in E \colon (\forall \alpha \in \Pi)\, (\mu, \check{\alpha}) \in \mathbb{N}_0\} \subseteq \{\mu \in E \colon (\forall \alpha \in \Pi)\, (\mu, \alpha) \geq 0\}.\end{aligned}$$

For $\mu \prec \lambda$, we have $\lambda - \mu \in \mathbb{N}_0[\Pi]$, so that we obtain

$$0 \leq (\lambda + \mu, \lambda - \mu) = \|\lambda\|^2 - \|\mu\|^2.$$

Therefore μ is contained in the closed ball B of radius $\|\lambda\|$, which is compact. On the other hand, $\mathcal{P}$ is a discrete subset of E, and therefore $B \cap \mathcal{P}$ is finite. $\square$

Proposition 15.17. *Let $\mathfrak{g}$ be a semisimple Lie algebra and $\mathfrak{h}$ be a toral Cartan subalgebra of $\mathfrak{g}$. If $\lambda \in \mathfrak{h}^*$ is dominant integral with respect to Δ^+, then the simple highest weight module $L(\lambda, \Delta^+)$ of highest weight λ is finite dimensional.*

Proof. For each simple root $\alpha \in \Pi$, using Theorem 10.9, we choose an $\mathfrak{sl}_2$-triple $(h_\alpha, e_\alpha, f_\alpha)$, so that $\mathfrak{g}(\alpha) = \mathrm{span}\{h_\alpha, e_\alpha, f_\alpha\} \cong \mathfrak{sl}_2(\mathbb{K})$ and $h_\alpha = \check{\alpha}$. Let $v_\lambda \in V := L(\lambda, \Delta^+)$ be a highest weight vector. For $\alpha, \beta \in \Pi$, we put $m_\alpha := \lambda(h_\alpha) = \lambda(\check{\alpha}) \in \mathbb{N}_0$ and observe that

$$e_\beta f_\alpha^{m_\alpha+1} v_\lambda = \begin{cases} f_\alpha^{m_\alpha+1} e_\beta v_\lambda = 0 & \text{if } \alpha \neq \beta, \\ f_\alpha^{m_\alpha+1} e_\beta v_\lambda + (m_\alpha + 1) f_\alpha^{m_\alpha}(h_\alpha - m_\alpha \mathbf{1}) v_\lambda = 0 & \text{if } \alpha = \beta. \end{cases}$$

Here we use that $[e_\beta, f_\alpha] \in \mathfrak{g}_{\beta - \alpha} = \{0\}$ for $\alpha \neq \beta$, and for $\alpha = \beta$ we use the formula

$$[e, f^{n+1}] = (n+1) f^n (h - n\mathbf{1})$$

from Lemma 9.2(ii) and $e_\beta \cdot v_\lambda = 0$. Since the e_α, $\alpha \in \Pi$, generate the subalgebra $\mathfrak{n} = \sum_{\beta \in \Delta^+} \mathfrak{g}_\beta$ (Proposition 14.1), the vector $f_\alpha^{m_\alpha+1} v_\lambda$ is a $\mathfrak{b}$-eigenvector, hence 0 because V is simple (Corollary 15.8). Therefore

$$\mathrm{span}\{v_\lambda, f_\alpha v_\lambda, \ldots, f_\alpha^{m_\alpha} v_\lambda\}$$

is a finite dimensional $\mathfrak{g}(\alpha)$-module with highest weight $\lambda(\check{\alpha}) = m_\alpha$. Therefore the subspace V_{fin} of V spanned by all finite dimensional $\mathfrak{g}(\alpha)$-submodules is nonzero.

Let $E \subseteq V$ be a finite dimensional $\mathfrak{g}(\alpha)$-submodule. Then $\mathrm{span}(\mathfrak{g} \cdot E)$ is finite dimensional and $\mathfrak{g}(\alpha)$-stable because we have for $x \in \mathfrak{g}(\alpha)$, $y \in \mathfrak{g}$ and $v \in E$ the relation

$$x \cdot (y \cdot v) = [x, y] \cdot v + y \cdot (x \cdot v) \in \mathrm{span}(\mathfrak{g} \cdot E).$$

Therefore $\mathfrak{g} \cdot E \subseteq V_{\mathrm{fin}}$. This implies that V_{fin} is a $\mathfrak{g}$-submodule. Since V is simple, we obtain $V = V_{\mathrm{fin}}$, i.e., V is a locally finite $\mathfrak{g}(\alpha)$-module.

Since all weights of V are integral (Proposition 15.14) and the weight set $\mathcal{P}_V$ of V is invariant under the Weyl group $\mathcal{W}$ (Lemma 15.15), Remark 15.13 shows that

$$\mathcal{P}_V \subseteq \mathcal{W}(\mathcal{P}_V \cap \mathcal{P}^+) \subseteq \mathcal{W}\big(\{\mu \in \mathcal{P}^+ \mid \mu \prec \lambda\}\big).$$

This set is finite because $\mathcal{W}$ is finite and the set $\{\mu \in \mathcal{P}^+ \mid \mu \prec \lambda\}$ is finite by Lemma 15.16. As all weight spaces V_μ are finite dimensional by Theorem 15.7, this concludes the proof. $\square$

Theorem 15.18. (Highest Weight Theorem) *Let $\mathfrak{g}$ be a split semisimple Lie algebra, $\mathfrak{h}$ be a toral Cartan subalgebra of $\mathfrak{g}$ and $\Delta^+ \subseteq \Delta$ be a positive system. The assignment $\lambda \mapsto L(\lambda, \Delta^+)$ defines a bijection between the set $\mathcal{P}^+$ of dominant integral functionals and the set of isomorphism classes of finite dimensional simple $\mathfrak{g}$-modules.*

Proof. To see that the assignment is defined, we first use Proposition 15.17 to see that, for $\lambda \in \mathcal{P}^+$, the simple $\mathfrak{g}$-module $L(\lambda, \Delta^+)$ is finite dimensional.

Next we recall from Proposition 15.5 that each finite dimensional simple $\mathfrak{g}$-module V is a highest weight module with some highest weight λ. In view of Proposition 15.14, $\lambda \in \mathcal{P}^+$, so that $V \cong L(\lambda, \Delta^+)$. Hence the assignment is surjective. That it is also injective follows from the fact that, for $\lambda \neq \lambda' \in \mathcal{P}^+$, we have $L(\lambda, \Delta^+) \not\cong L(\lambda', \Delta^+)$ by Proposition 15.9. $\square$

15.3 The Eigenvalue of the Casimir Operator

In this section we construct a special element $C_{\mathfrak{g}}$ in the center of the enveloping algebra of $\mathfrak{g}$ and calculate its (scalar) action in a highest weight module. In special cases this allows to identify a given simple $\mathfrak{g}$-module.

Definition 15.19. (Universal Casimir element) Let $\mathfrak{g}$ be a finite dimensional split semisimple Lie algebra with Cartan–Killing form κ. As before, we choose for each $\beta \in \Delta^+$ an $\mathfrak{sl}_2$-triple $(h_\beta, e_\beta, f_\beta)$ and $e_\beta^* \in \mathfrak{g}_{-\beta}$, $f_\beta^* \in \mathfrak{g}_\beta$ with

$$\kappa(e_\beta, e_\beta^*) = 1 = \kappa(f_\beta, f_\beta^*).$$

We further choose a basis $h_1, \ldots, h_r$ for $\mathfrak{h}$, and we write $h^1, \ldots, h^r$ for the dual basis with respect to the nondegenerate restriction of κ to $\mathfrak{h} \times \mathfrak{h}$. Then

$$\{h_i, e_\beta, f_\beta \colon i = 1, \ldots, r; \beta \in \Delta^+\}$$

is a basis for $\mathfrak{g}$ and

$$\{h^i, e_\beta^*, f_\beta^* \colon i = 1, \ldots, r; \beta \in \Delta^+\}$$

is the dual basis with respect to κ. We therefore obtain a central element of $\mathcal{U}(\mathfrak{g})$ by

$$C_{\mathfrak{g}} = \sum_{i=1}^{k} h_i h^i + \sum_{\beta \in \Delta^+} e_\beta e_\beta^* + f_\beta f_\beta^* \tag{32}$$

(Lemma 5.20). It is called the *universal Casimir element.*

Lemma 15.20. *For a positive system Δ^+, we put $\rho = \frac{1}{2}\sum_{\alpha \in \Delta^+} \alpha$. If (ρ_V, V) is a highest weight module with highest weight λ, also considered as a $\mathcal{U}(\mathfrak{g})$-module, then*

$$\rho_V(C_{\mathfrak{g}}) = (\lambda, \lambda + 2\rho)\mathbf{1} = \big(\|\lambda + \rho\|^2 - \|\rho\|^2\big)\mathbf{1}.$$

If λ is dominant and nonzero, then $\rho_V(C_{\mathfrak{g}}) \neq 0$.

Proof. We write the Casimir element $C_{\mathfrak{g}}$ in the form (32) as described in Definition 15.19. We compute the action of $C_{\mathfrak{g}}$ on V. Let v_λ be a highest weight vector in V. Then $e_\beta \cdot v_\lambda = f_\beta^* \cdot v_\lambda = 0$ for each $\beta \in \Delta^+$, and $[e_\beta, e_\beta^*] = t_\beta$ (cf. Lemma 10.7) implies

$$e_\beta e_\beta^* \cdot v_\lambda = [e_\beta, e_\beta^*] \cdot v_\lambda + e_\beta^* e_\beta \cdot v_\lambda = \lambda(t_\beta) v_\lambda = (\lambda, \beta) v_\lambda,$$

so that $\sum_{\beta \in \Delta^+}(e_\beta e_\beta^* + f_\beta f_\beta^*) \cdot v_\lambda = 2(\lambda, \rho) v_\lambda$. On the other hand, we calculate

$$\sum_{i=1}^{k} \lambda(h_i)\lambda(h^i) \;=\; \lambda\Big(\sum_{i=1}^{k} \kappa(h'_\lambda, h^i) h_i\Big) = \lambda(h'_\lambda) = (\lambda, \lambda).$$

Putting these facts together yields

$$C_{\mathfrak{g}} v_\lambda = (\lambda, \lambda + 2\rho) v_\lambda = \big(\|\lambda + \rho\|^2 - \|\rho\|^2\big) v_\lambda.$$

Since $C_{\mathfrak{g}}$ is central in $\mathcal{U}(\mathfrak{g})$ (Exercise 13.1), $C_{\mathfrak{g}}$ acts by the same scalar on the entire $\mathcal{U}(\mathfrak{g})$-module $V = \mathcal{U}(\mathfrak{g}) v_\lambda$.

Finally, we assume that λ is dominant and nonzero. Then $\lambda(\check{\alpha}) \geq 0$ for all $\alpha \in \Delta^+$ implies that $(\lambda, \alpha) \geq 0$, and hence that $(\lambda, \rho) \geq 0$. This leads to $(\lambda, \lambda + 2\rho) \geq (\lambda, \lambda) > 0$. □

Exercises for Section 15

Exercise 15.1. A $\mathfrak{g}$-module V is said to be *cyclic* if it is generated by some element $v \in V$. If $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is the module structure and $\widetilde{\rho}\colon \mathcal{U}(\mathfrak{g}) \to \mathrm{End}(V)$ the canonical extension, then the annihilator

$$I := \mathrm{Ann}_{\mathcal{U}(\mathfrak{g})}(v) := \{D \in \mathcal{U}(\mathfrak{g})\colon \widetilde{\rho}(D)v = 0\}$$

of v is a left ideal. Show that:

(a) If $I \subseteq \mathcal{U}(\mathfrak{g})$ is a left ideal, then the quotient $\mathcal{U}(\mathfrak{g})/I$ carries a natural $\mathfrak{g}$-module structure, defined by $x \cdot (D + I) := \sigma(x)D + I$, and this $\mathfrak{g}$-module is cyclic.

(b) Every cyclic $\mathfrak{g}$-module is isomorphic to one of the form $\mathcal{U}(\mathfrak{g})/I$, as in (a).

Exercise 15.2. Simple $\mathfrak{g}$-modules are particular examples of cyclic $\mathfrak{g}$-modules. Show that:

(a) If $I \subseteq \mathcal{U}(\mathfrak{g})$ is a maximal (proper) left ideal, then the quotient $\mathcal{U}(\mathfrak{g})/I$ is a simple $\mathfrak{g}$-module.

(b) Every simple $\mathfrak{g}$-module is isomorphic to one of the form $\mathcal{U}(\mathfrak{g})/I$, where I is a maximal left ideal in $\mathcal{U}(\mathfrak{g})$.

Exercise 15.3. Let $\mathfrak{g}$ be a reductive Lie algebra and $\mathfrak{h} \subseteq \mathfrak{g}$ a toral Cartan subalgebra. Let V be a simple $\mathfrak{g}$-module on which $\mathfrak{h}$ acts by diagonalizable operators (such modules are called *weight modules*. Identifying the root system of $\mathfrak{g}$ with that of its semisimple commutator algebra, the notion of a positive system makes also sense for $\mathfrak{g}$. Show that, in this sense, for each positive system Δ^+ of Δ, V is a highest weight module.

16 Applications to elementary particles

In this brief section we describe the bridge between the representation theory of Lie algebras and its application in the theory of elementary particles.

The fundamental idea underlying this connection is that the *state space* of a quantum mechanical system is the *projective space*

$$\mathbb{P}(\mathcal{H}) := \{[v] = \mathbb{C}v\colon 0 \neq v \in \mathcal{H}\}$$

of one-dimensional subspaces of a complex Hilbert space $\mathcal{H}$. In the discussion below we assume that $\mathcal{H}$ is finite dimensional which is sufficiently general for the applications we want to describe ([BH09]). States will always be represented by unit vectors: $\langle v, v\rangle = 1$.

In the following we shall always assume that the scalar product $\langle\cdot|\cdot\rangle$ on $\mathcal{H}$ is linear in the second and antilinear in the first argument. As we shall see below, this is natural in the quantum mechanical context. In particular, it is consistent with *Dirac's bra-ket notation*, where vectors in the Hilbert space $\mathcal{H}$ are written $|v\rangle$ and elements in the dual space of continuous linear functionals on $\mathcal{H}$ as $\langle v|$. Applying the linear functional $\langle v|$ to the vector $|w\rangle$ results in the complex number $\langle v|w\rangle$. Applying operators $A\colon \mathcal{H} \to \mathcal{H}$ to vectors then looks like $A|v\rangle$.

The *observables* of the quantum system correspond in this context to symmetric operators $A = A^*$ on $\mathcal{H}$. The real number

$$H_A([v]) = \langle v|A|v\rangle \in \mathbb{R}$$

is interpreted as the *expectation value of the observable A in the states* $[v]$. Since A is diagonalizable, we can always write $v = \sum_\alpha v_\alpha$, where v_α is an A-eigenvector corresponding to the eigenvalue $\alpha \in \mathbb{R}$. Then

$$H_A([v]) = \sum_\alpha \alpha \|v_\alpha\|^2 \in \mathrm{conv}(\mathrm{Spec}(A))$$

is a convex combination of the eigenvalues α because $1 = \|v\|^2 = \sum_\alpha \|v_\alpha\|^2$. So the value of the observable A in $[v]$ is a superposition of the different eigenvalues and one says that A has in $[v]$ the "sharp value" α if $v = v_\alpha$, i.e., if the variance vanishes: $\langle v|(A - \alpha\mathbf{1})^2|v\rangle = \|Av - \alpha v\|^2 = 0$.

In the theory of elementary particles, the finite dimensional space $\mathcal{H}$ represents a system with $\dim \mathcal{H}$ different particles, a so-called *multiplet* which corresponds to an orthonormal basis of $\mathcal{H}$. One further requires that the space $\mathcal{H}$ carries an irreducible unitary representation of a compact Lie group G. In terms of Lie algebras, this means that we have a homomorphism of Lie algebras

$$\rho\colon \mathfrak{g} \to \mathfrak{u}(\mathcal{H}) := \{X \in \mathrm{End}(\mathcal{H})\colon X^* = -X\},$$

i.e., a *unitary representation of* $\mathfrak{g}$ *on* $\mathcal{H}$ (Section 2). Then ρ extends to a complex linear representation

$$\rho_\mathbb{C}\colon \mathfrak{g}_\mathbb{C} \to \mathrm{End}(\mathcal{H}), \quad \rho_\mathbb{C}(x+iy) := \rho(x) + i\rho(y), \quad x, y \in \mathfrak{g},$$

satisfying

$$\rho_\mathbb{C}(Z)^* = -\rho_\mathbb{C}(\overline{Z}), \quad \text{where} \quad \overline{x+iy} = x - iy.$$

In particular, the operators $\rho_\mathbb{C}(ix) = i\rho(x)$, $x \in \mathfrak{g}$, are symmetric and correspond to observables of the corresponding quantum system. We are thus lead to finite dimensional simple modules of complex (semisimple) Lie algebras which can be classified in terms of their highest weights.

In the theory of elementary particles one also considers composed particles. This is modelled by tensor products of representations. Consider two unitary representations $(\rho_j, \mathcal{H}_j)$, $j = 1, 2$, corresponding to two families $\mathcal{F}_1$ and $\mathcal{F}_2$ of particles. Then we can form the tensor product representation

$$\rho := \rho_1 \otimes \rho_2 \quad \text{with} \quad \rho(z)(v \otimes w) := \rho_1(z)v \otimes w + v \otimes \rho_2(z)w.$$

It is unitary with respect to the natural scalar product on $\mathcal{H}_1 \otimes \mathcal{H}_2$, specified by

$$\langle v_1 \otimes w_1 | v_2 \otimes w_2\rangle := \langle v_1|v_2\rangle\langle w_1|w_2\rangle.$$

The particles described by the representation $(\rho_1 \otimes \rho_2, \mathcal{H}_1 \otimes \mathcal{H}_2)$ are interpreted as composed from one particle in the family $\mathcal{F}_1$ and one in $\mathcal{F}_2$. More generally, the representation $\rho_1^{\otimes n} \otimes \rho_2^{\otimes m}$ on $\mathcal{H}_1^{\otimes n} \otimes \mathcal{H}_2^{\otimes m}$ describes particles composed from n particles in $\mathcal{F}_1$ and m particles in $\mathcal{F}_2$.

For a unitary representation $(\rho, \mathcal{H})$ one can also form the *dual representation* on the dual space $\mathcal{H}^*$ of continuous linear functionals $\alpha\colon \mathcal{H} \to \mathbb{C}$. Writing $\beta \in \mathcal{H}^*$ as $\beta = \langle v_\beta|$ for some $v_\beta \in \mathcal{H}$, we have

$$\langle \alpha|\beta\rangle := \langle v_\beta|v_\alpha\rangle.$$

The representation of the Lie algebra $\mathfrak{g}$ on the dual space is given by

$$\rho^*(z)\alpha := -\alpha \circ \rho(z).$$

In the context of elementary particles, the dual representation corresponds to *antiparticles*.[18]

[18] The map $\mathcal{H} \to \mathcal{H}^*, v \mapsto \langle v|$ is an antilinear isometry. One can therefore identify $\mathcal{H}^*$ with the space

16.1 Nucleons and the isospin Lie algebra

To make this dictionary for the translation between representations and elementary particles more concrete, we start with a very simple system. The underlying idea is to consider proton and neutron (particles in an atomic nucleus) as two states of the same particle, the *nucleon*. It goes back to the early days of quantum mechnics (Heisenberg; 1932) and is based on the observation that both nucleons behave almost alike under strong interactions (nuclear forces).

To distinguish the two types of nucleons, one introduces a (virtual) coordinate called *isospin* τ. Since this obervable has to take two values, it corresponds to an operator on the 2-dimensional Hilbert space $\mathbb{C}^2$, which carries the obvious irreducible unitary representation of $G = \mathrm{SU}_2(\mathbb{C})$, resp., $\mathfrak{g}_{\mathbb{C}} = \mathfrak{sl}_2(\mathbb{C})$. Accordingly, $\mathfrak{su}_2(\mathbb{C})$ is called the *isospin Lie algebra*.

We define the three *isospin operators* by

$$\widehat{I}_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \widehat{I}_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \widehat{I}_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

These symmetric operators satisfy the commutation relations

$$[\widehat{I}_1, \widehat{I}_2] = i\widehat{I}_3, \quad [\widehat{I}_2, \widehat{I}_3] = i\widehat{I}_1, \quad [\widehat{I}_3, \widehat{I}_1] = i\widehat{I}_2,$$

so that the corresponding elements $i\widehat{I}_j$ of $\mathfrak{su}_2(\mathbb{C})$ satisfy

$$[-i\widehat{I}_1, -i\widehat{I}_2] = -i\widehat{I}_3, \quad [-i\widehat{I}_2, -i\widehat{I}_3] = -i\widehat{I}_1, \quad [-i\widehat{I}_3, -i\widehat{I}_1] = -i\widehat{I}_2$$

(cf. Example 1.35).

From Theorem 9.5 we know that the finite dimensional irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$ are of the form (π_n, V_n) with $V_n = L(n)$, $n \in \mathbb{N}_0$, and $\dim V_n = n+1$. The number $I = \frac{n}{2}$ is called the *total isospin* of the particles described by V_n. Accordingly, we define an operator $\widehat{I}$ on V_n by $\frac{n}{2}\mathbf{1}$.

The eigenvalues of the operators $\widehat{I}$ and $\widehat{I}_3$ (often called the third isospin component) classify the states in all finite dimensional irreducible representations of $\mathfrak{su}_2(\mathbb{C})$, resp., $\mathfrak{sl}_2(\mathbb{C})$. The first one determines the simple module as V_{2I} and $\widehat{I}_3$ determines the state within the multiplet described by V_{2I}. In physics, this is expressed by Dirac's bra-ket notation

$$\widehat{I}|II_3\rangle = I|II_3\rangle \quad \text{and} \quad \widehat{I}_3|II_3\rangle = I_3|II_3\rangle.$$

If the total isosping I is fixed, then

$$|II_3\rangle, \quad I_3 \in \{-I, -I+1, \ldots, I-1, I\}$$

describes an orthonormal basis of $\widehat{I}_3$ eigenvectors in V_{2I}.

The doublet of *nucleons*, proton $\mathbf{p}$ and neutron $\mathbf{n}$, consists of two particles with total isospin $\frac{1}{2}$ given by the vectors

$$|\mathbf{p}\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\mathbf{n}\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{in} \quad V_1 \cong \mathbb{C}^2.$$

These are the two states of a single nucleon corresponding to the values $\pm\frac{1}{2}$ of the (third component of) isospin $\widehat{I}_3$. The terminology 'isospin' has been chosen because nuclei with

$\mathcal{H}$, endowed with the new scalar multiplication $\lambda * v := \overline{\lambda} v$. This complex vector space is denoted $\overline{\mathcal{H}}$. Then the scalar product on $\mathcal{H}^*$ corresponds to the hermitian scalar product $\langle v|w\rangle_* := \langle w|v\rangle$ on $\overline{\mathcal{H}}$. This picture has the advantage that the dual representation ρ^* of $\mathfrak{g}$ corresponds to the original representation $\rho\colon \mathfrak{g} \to \mathfrak{u}(\mathcal{H}) = \mathfrak{u}(\overline{\mathcal{H}})$.

the same number of nucleons (isotopes) can be distinguished by the eigenvalues of $\widehat{I}_3$. A nucleus with N_p protons and N_n neutrons contains $N = N_p + N_n$ nucleons and has the isospin $\frac{1}{2}(N_p - N_n)$, hence is determined by the pair $(N, \frac{1}{2}(N_p - N_n))$. The operator corresponding to the *charge of a nucleon* is $\widehat{Q} = e(\widehat{I}_3 + \frac{1}{2}\mathbf{1})$, where $-e$ is the charge of an electron. The neutron has charge 0 and the proton has charge e.

Multiplets of particles composed from N nucleons correspond to irreducible subrepresentation of the N-fold tensor product $V_1^{\otimes N}$, on which the charge operator takes the form $\widehat{Q} = e(\widehat{I}_3 + \frac{N}{2}\mathbf{1})$. We therefore need some information on the decomposition of tensor products of representations of $\mathfrak{sl}_2(\mathbb{C})$:

Proposition 16.1. *For $n, m \in \mathbb{N}_0$, we have $V_n \otimes V_m = \bigoplus_{q=0}^{\min(n,m)} V_{n+m-2q}$.*

Since tensor products are distributive and every finite dimensional representation of $\mathfrak{sl}_2(\mathbb{C})$ is a direct sum of irreducible ones, the preceding proposition can be used to determine for every tensor product of finite dimensional representations the decomposition into irreducible ones. From

$$V_1 \otimes V_1 \cong S^2(V_1) \oplus \Lambda^2(V_1) \cong V_2 \oplus V_0 \tag{33}$$

we obtain in particular

$$V_1 \otimes V_1 \otimes V_1 \cong V_1 \otimes (V_2 \oplus V_0) \cong (V_1 \otimes V_2) \oplus (V_1 \otimes V_0) \cong V_3 \oplus V_1 \oplus V_1 = 2V_1 \oplus V_3. \tag{34}$$

In terms of nucleons, (33) means that from 2 nucleons we obtain two multiplets: a triplet V_2 and a singlet V_0. One of these particles is the deuteron $\mathbf{d}$, composed of one proton and one neutron. Its charge is e and its isospin is 0:

$$\widehat{Q}|\mathbf{d}\rangle = e|\mathbf{d}\rangle \quad \text{and} \quad I_3|\mathbf{d}\rangle = 0|\mathbf{d}\rangle.$$

For the total isospin I, only the two values 0 and 2 are possible. The state with $I = 0$ is called a *ground state of the deuteron*. It is represented in $V_1^{\otimes 2}$ by the invariant unit vector

$$\frac{1}{\sqrt{2}}(\mathbf{p} \otimes \mathbf{n} - \mathbf{n} \otimes \mathbf{p}).$$

From the triplet $V_2 \cong S^2(V_1) \subseteq V_1^{\otimes 2}$ only

$$\Big[\frac{1}{\sqrt{2}}(\mathbf{p} \otimes \mathbf{n} + \mathbf{n} \otimes \mathbf{p})\Big]$$

corresponds to a deuteron. It represents an excited deuteron state of total isospin 2.

16.2 Up and down quarks and the isospin Lie algebra

If is one of the fundamental insights of nuclear physics that the nucleons are not elementary particles in the sense that they should be considered as composed from more fundamental particles called quarks.

The *up- and down-quark* $\mathbf{u}$ and $\mathbf{d}$ corresponds to a 2-dimensional representation of the isospin algebra $\mathfrak{su}_2(\mathbb{C})$:

$$|\mathbf{u}\rangle = \Big|\frac{1}{2}, \frac{1}{2}\Big\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\mathbf{d}\rangle = \Big|\frac{1}{2}, -\frac{1}{2}\Big\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{in} \quad V_1 \cong \mathbb{C}^2.$$

With respect to the isosping algebra, this doublet behaves in the same way as the doublet of nucleons. These two quarks are considered as two states of a quark particle of total isospin $\frac{1}{2}$ which are distinguished by the I_3-value.

To particles composed from quarks we associate a *baryon number* B in an additive way with $B = \frac{1}{3}$ for quarks and $B = -\frac{1}{3}$ for antiquarks. Particles composed from p quarks and q antiquarks (corresponding to subrepresentations of $V_1^{\otimes p} \otimes (V_1^*)^{\otimes q} \cong V_1^{\otimes(p+q)}$ have the baryon number $\frac{1}{3}(p-q)$. If $\widehat{B}$ is the corresponding operator, then the Gell–Mann–Nishijima formula

$$\widehat{Q} = \widehat{I}_3 + \frac{1}{2}\widehat{B}$$

describes the *normalized charge* of a particle in terms of multiples of e:

$$\widehat{Q}|\mathbf{u}\rangle = \frac{2}{3}|\mathbf{u}\rangle \quad \text{and} \quad \widehat{Q}|\mathbf{d}\rangle = -\frac{1}{3}|\mathbf{d}\rangle.$$

In terms of the $\widehat{I}_3, \widehat{Q}$-eigenvalues, we obtain the following description of the quarks

$$|\mathbf{u}\rangle = \left|\frac{1}{2}, \frac{2}{3}\right\rangle, \quad |\mathbf{d}\rangle = \left|-\frac{1}{2}, -\frac{1}{3}\right\rangle$$

and their antiparticles:

$$|\overline{\mathbf{u}}\rangle = \left|-\frac{1}{2}, -\frac{2}{3}\right\rangle, \quad |\overline{\mathbf{d}}\rangle = \left|\frac{1}{2}, \frac{1}{3}\right\rangle.$$

Particles composed from a quark and an antiquark are called *mesons*; their baryon number is $B = 0$. They correspond to states in

$$V_1 \otimes V_1^* \cong V_1^{\otimes 2} \cong V_2 \oplus V_0,$$

so they can form a triplet and a singulet. The triplet corresponds to the *pions, or π-mesons*

$$\pi^+ = |1,0\rangle = [\mathbf{u}\otimes\overline{\mathbf{d}}], \quad \pi^0 = |0,0\rangle = \Big[\frac{1}{\sqrt{2}}(\mathbf{u}\otimes\overline{\mathbf{u}}+\mathbf{d}\oplus\overline{\mathbf{d}})\Big], \quad \pi^- = |-1,0\rangle = [\mathbf{d}\otimes\overline{\mathbf{u}}]. \tag{35}$$

Particles composed from three quarks are called *baryons*. Their baryon number is $B = 1$. The baryon representation is

$$V_1^{\otimes 3} \cong 2V_1 \oplus V_3.$$

On of the two doublets in this representation corresponds to the nucleons, as particles composed from three quarks:

$$\mathbf{p} = \Big[\frac{1}{\sqrt{2}}(\mathbf{u}\otimes\mathbf{d}\otimes\mathbf{u} - \mathbf{d}\otimes\mathbf{u}\otimes\mathbf{u})\Big] \quad \text{and} \quad \mathbf{n} = \Big[\frac{1}{\sqrt{2}}(\mathbf{u}\otimes\mathbf{d}\otimes\mathbf{d} - \mathbf{d}\otimes\mathbf{u}\otimes\mathbf{d})\Big]. \tag{36}$$

This means that the proton is composed from two u-quarks and one d-quark and the neutron from two d-quarks and one u-quark. It is easy to verify that the subspace generated by these two states is actually invariant under the natural action of the isospin algebra on $V_1^{\otimes 3}$.

16.3 Strange quarks and the flavor Lie algebra

We now turn to the larger Lie algebra $\mathfrak{g} = \mathfrak{su}_3(\mathbb{C})$ with $\mathfrak{g}_{\mathbb{C}} \cong \mathfrak{sl}_3(\mathbb{C})$ that contains $\mathfrak{su}_2(\mathbb{C})$ in the obvious way as a subalgebra. We shall see below how this Lie algebra is related to the composition of particles by quarks. In this context it is called the *flavor Lie algebra*, which refers to the different flavors of the quarks.

The Cartan subalgebra $\mathfrak{h} \subseteq \mathfrak{g}_{\mathbb{C}}$ is 2-dimensional, so that, in addition to the isospin operator $\widehat{I}_3$, we need a second operator $\widehat{Y}$ called the *hypercharge operator*, to generate $\mathfrak{h}$:

$$\widehat{I}_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \widehat{Y} = \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

We know that the a positive system of roots for $\mathfrak{sl}_3(\mathbb{C})$ can be written as

$$\Delta^+ = \{\alpha, \beta, \alpha+\beta\}, \quad \alpha = \varepsilon_1 - \varepsilon_2, \beta = \varepsilon_2 - \varepsilon_3.$$

We have $\alpha(\widehat{I}_3) = 1$, $\beta(\widehat{I}_3) = -\frac{1}{2}$, $\alpha(\widehat{Y}) = 0$ and $\beta(\widehat{Y}) = 1$. This leads to

$$\widehat{I}_3 = \frac{1}{2}\check{\alpha} \quad \text{und} \quad \widehat{Y} = \frac{1}{3}(\check{\alpha} + 2\check{\beta}).$$

Let $V_{p,q}$ be the simple highest weight module with highest weight λ determined by $\lambda(\check{\alpha}) = p$ and $\lambda(\check{\beta}) = q$. Then

$$\lambda(\widehat{I}_3) = \frac{p}{2} \quad \text{und} \quad \lambda(\widehat{Y}) = \frac{1}{3}(p+2q).$$

The triplet $V_{1,0} \cong \mathbb{C}^3$ (with the identical representation) is spanned by the vectors $\mathbf{u}$, $\mathbf{d}$, und $\mathbf{s}$ called *up-, down-, and strange quark* (u-, d-, and s-quark for short). The correspond to the canonical basis vectors:

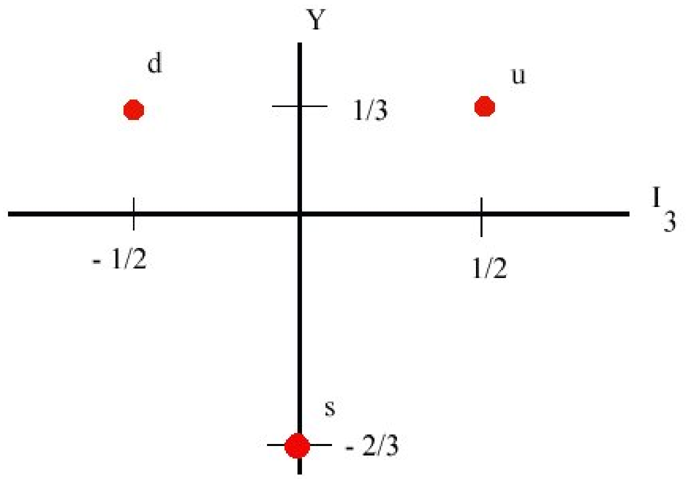
$$|\mathbf{u}\rangle = \left|\frac{1}{2}, \frac{1}{3}\right\rangle = \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right], \quad |\mathbf{d}\rangle = \left|-\frac{1}{2}, \frac{1}{3}\right\rangle = \left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right], \quad \text{and} \quad |\mathbf{s}\rangle = \left|0, -\frac{2}{3}\right\rangle = \left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right],$$

where the labels are the corresponding eigenvalues of $\widehat{I}_3$ and $\widehat{Y}$. The charge operator is

$$\widehat{Q} = \tfrac{1}{2}\widehat{Y} + \widehat{I}_3 = \frac{1}{3}\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \frac{1}{3}(2\check{\alpha} + \check{\beta}),$$

so that

$$\widehat{Q}|\mathbf{u}\rangle = \frac{2}{3}|\mathbf{u}\rangle, \quad \widehat{Q}|\mathbf{d}\rangle = -\frac{1}{3}|\mathbf{d}\rangle, \quad \text{and} \quad \widehat{Q}|\mathbf{s}\rangle = -\frac{1}{3}|\mathbf{s}\rangle.$$



The dual representation is the triplet $V_{1,0}^* \cong V_{0,1}$. The corresponding states are the *antiquarks*

$$|\bar{\mathbf{u}}\rangle = \left|-\frac{1}{2}, -\frac{1}{3}\right\rangle, \quad |\bar{\mathbf{d}}\rangle = \left|\frac{1}{2}, -\frac{1}{3}\right\rangle, \quad \text{and} \quad |\bar{\mathbf{s}}\rangle = \left|0, \frac{2}{3}\right\rangle.$$

Every particule with $\mathfrak{su}_3(\mathbb{C})$-symmetry composed from p quarks and q antiquarks is contained in a submultiplet of the tensor product $V_{1,0}^p \otimes V_{0,1}^q$ which contains in particular the simple submodule $V_{p,q}$. One can show that $V_{p,q}^* \cong V_{q,p}$. As in the preceding subsection, we define the baryon number B in an additive way with $B = \frac{1}{3}$ for quarks and $B = -\frac{1}{3}$ for antiquarks. If $\widehat{B}$ is the corresponding operator, then $\widehat{S} := \widehat{Y} - \widehat{B}$ is called the *strangeness operator*. In the fundamental representation we have

$$\widehat{S} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \text{i.e.,} \quad \widehat{S}|\mathbf{u}\rangle = 0, \quad \widehat{S}|\mathbf{d}\rangle = 0, \quad \widehat{S}|\mathbf{s}\rangle = -|\mathbf{s}\rangle,$$

so that the s-quark has strangeness -1. The charge of a particle can be computed with the *Gell-Mann–Nishijima formula*:

$$\widehat{Q} = \widehat{I}_3 + \frac{1}{2}(\widehat{B} + \widehat{S}) = \widehat{I}_3 + \widehat{Y}.$$

Example 16.2. (a) Particles composed from 2 quarks are obtained by

$$V_{1,0} \otimes V_{1,0} \cong S^2(V_{1,0}) \oplus \Lambda^2(V_{1,0}) \cong V_{2,0} \oplus V_{0,1},$$

a sextet and a triplet.

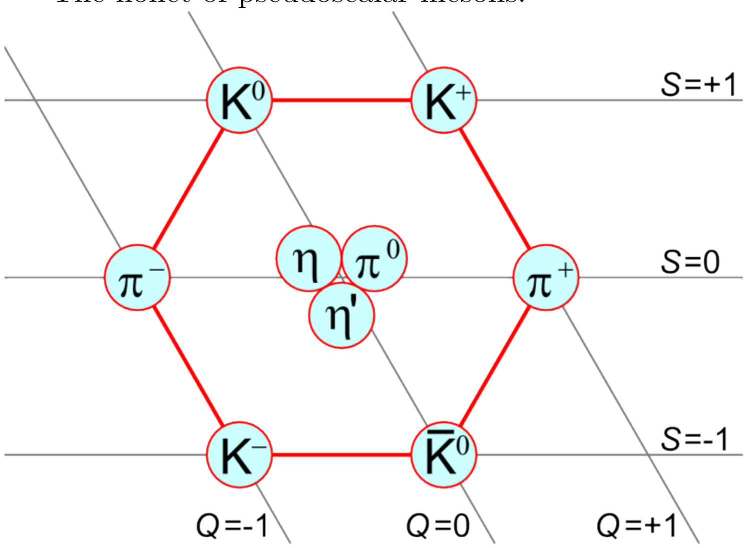
(b) Likewise, particles composed from 2 antiquarks are obtained from $V_{0,1} \otimes V_{0,1} = V_{0,2} \oplus V_{1,0}$, a sextet and a triplet.

(c) Particles composed from one quark and one antiquark, mesons, are obtained from

$$V_{1,0} \otimes V_{0,1} \cong V_{1,0} \otimes V_{1,0}^* \cong \mathrm{End}(V_{1,0}) \cong \mathfrak{sl}(V_{1,0}) \oplus \mathbb{C}\,\mathrm{id}_{V_{1,0}} \cong V_{1,1} \oplus V_0,$$

an octet and a singlet. Together, these particles form the *nonet of pseudoscalar mesons* ([GM85, p. 320]). Here $V_{1,1}$ is the 8-dimensional representation of $\mathfrak{su}_3(\mathbb{C})$ corresponding to the adjoint representation of $\mathfrak{sl}_3(\mathbb{C})$. In the following diagram $\widehat{S} = \widehat{Y}$ is the *strangeness operator* and the two coordinates are charge and strangeness.

The nonet of pseudoscalar mesons:



Since the $[\widehat{S}, \mathfrak{g}(\alpha)] = \{0\}$, the eigenspaces of $\widehat{S}$ are representations of the isospin Lie algebra $\mathfrak{g}(\alpha) \cong \mathfrak{sl}_2(\mathbb{C})$ (the horizontal lines in the diagram). As $[\widehat{Q}, \mathfrak{g}(\beta)] = \{0\}$, the eigenspaces of $\widehat{Q}$ are representations of the Lie algebra $\mathfrak{g}(\beta) \cong \mathfrak{sl}_2(\mathbb{C})$ (the slanted lines with negative slope).

The particles in this nonet are the pions (cf. (35))

$$\pi^+ = [\mathbf{u} \otimes \overline{\mathbf{d}}], \quad \pi^0 = \Big[\frac{1}{\sqrt{2}}(\mathbf{u} \otimes \overline{\mathbf{u}} + \mathbf{d} \otimes \overline{\mathbf{d}})\Big], \quad \pi^- = [\mathbf{d} \otimes \overline{\mathbf{u}}]$$

(a 3-dimensional representation of the isospin algebra), the *kaons*

$$K^0 = [\mathbf{d} \otimes \overline{\mathbf{s}}], \quad K^+ = [\mathbf{u} \otimes \overline{\mathbf{s}}], \quad K^- = [\mathbf{s} \otimes \overline{\mathbf{u}}], \quad \overline{K}^0 = [\mathbf{s} \otimes \overline{\mathbf{d}}]$$

(two 2-dimensional representations of the isospin algebra), and the particle

$$\eta' = \Big[\frac{1}{\sqrt{3}}(\mathbf{u} \otimes \overline{\mathbf{u}} + \mathbf{d} \otimes \overline{\mathbf{d}} + \mathbf{s} \otimes \overline{\mathbf{s}})\Big],$$

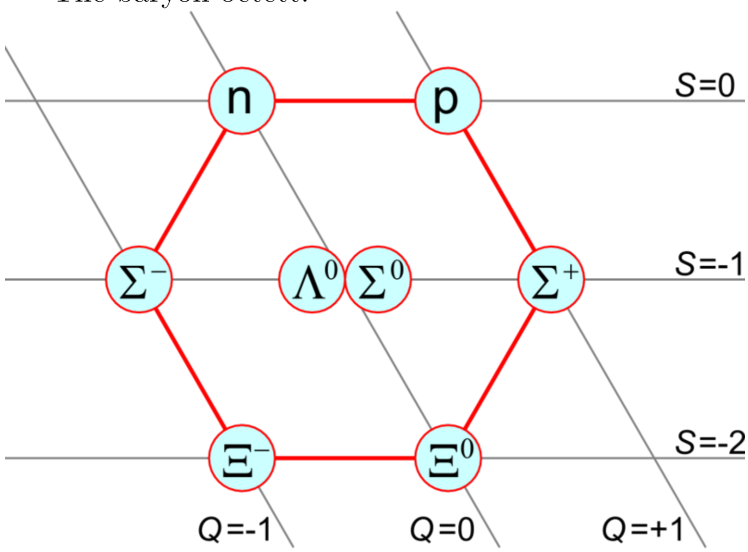
which forms a singulet because it corresponds to a trivial subrepresentation of the flavor algebra $\mathfrak{sl}_3(\mathbb{C})$.

(d) Another important octet are the baryons, composed of 3 quarks. It is obtained as a summand of type $V_{1,1}$ from the decomposition

$$V_{1,0}^{\otimes 3} \cong S^3(V_{1,0}) \oplus \Lambda^3(V_{1,0}) \oplus V_{1,1} \oplus V_{1,1} \cong V_{3,0} \oplus V_{0,0} \oplus 2V_{1,1}.$$

Here $V_{3,0}$ is a 10-dimensional space and $V_{0,0}$ is one-dimensional.

The baryon octett:



Here we find the nucleons (proton and neutron) in the $S = 0$-eigenspace of a subrepresentation of type $V_{1,1}$ (see (36))).

For further reading we recommend the survey [BH09] on grand unified theories.

A The Jordan Decomposition

In this appendix we develop a tool that will be of crucial importance throughout the structure theory of Lie algebras: the Jordan decomposition of an endomorphism of a finite dimensional vector space. Although the existence of the Jordan decomposition can be derived from the Jordan normal form, the proof of the Jordan decomposition is less involved because it does not specify the structure of the nilpotent component. Since we need various properties of the Jordan decomposition, we give a direct self-contained proof which does not require more than some elementary properties of polynomials.

Definition A.1. Let V be a vector space and $M \in \mathrm{End}(V)$.

(a) For $\lambda \in \mathbb{K}$, we define the *eigenspace* with respect to λ as

$$V_\lambda(M) := \ker(M - \lambda\mathbf{1})$$

and the *generalized eigenspace* as

$$V^\lambda(M) := \bigcup_{n\in\mathbb{N}} \ker(M - \lambda\mathbf{1})^n.$$

Note that the ascending sequence $\ker(M - \lambda\mathbf{1})^n$ is eventually constant if V is finite dimensional. We call λ an *eigenvalue* if $V_\lambda(M) \neq \{0\}$.

(b) We call M *diagonalizable* if $V = \bigoplus_{\lambda\in\mathbb{K}} V_\lambda(M)$, i.e., V is a direct sum of the eigenspaces of M.

(c) We call M *nilpotent* if there exists an $n \in \mathbb{N}$ with $M^n = 0$. If M is nilpotent, then $V = V^0(M)$.

(d) We call M *split* if there is a nonzero polynomial $f \in \mathbb{K}[X]$ with $f(M) = 0$ which decomposes as a product of linear factors. This is always the case for $\mathbb{K} = \mathbb{C}$.

(e) For $\mathbb{K} = \mathbb{R}$, we call M *semisimple* if the endomorphism $M_\mathbb{C}$ of $V_\mathbb{C}$, defined by $M_\mathbb{C}(v + iv') = MviMv'$ is diagonalizable (cf. Exercise A.5).

Theorem A.2. (Jordan Decomposition Theorem) *Let V be a finite dimensional vector space and $M \in \mathrm{End}(V)$ a split endomorphism. Then there exists a diagonalizable endomorphism M_s and a nilpotent endomorphism M_n such that*

(i) $M = M_s + M_n$.

(ii) $V^\lambda(M_s) = V_\lambda(M_s) = V^\lambda(M)$ *for each* $\lambda \in \mathbb{K}$.

(iii) *There exist polynomials $P, Q \in \mathbb{K}[X]$ with $P(0) = Q(0) = 0$ such that $M_s = P(M)$ and $M_n = Q(M)$.*

(iv) *If $L \in \mathrm{End}(V)$ commutes with M, then it also commutes with M_s and M_n.*

(v) (Uniqueness of the Jordan decomposition) *If $S, N \in \mathrm{End}(V)$ commute, S is diagonalizable and N nilpotent with $M = S + N$, then $S = M_s$ and $N = M_n$.*

Proof. Let $f \in \mathbb{K}[X]$ be the minimal polynomial of M, i.e., a generator of the ideal $I_M := \{f \in \mathbb{K}[X] \colon f(M) = 0\}$ with leading coefficient 1. By assumption, I_M contains a nonzero polynomial which is a product of linear factors, so that Exercise A.6 implies that f also has this property. Hence there exist pairwise different $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ and $k_i \in \mathbb{N}$ such that f can be written as

$$f = (X - \lambda_1)^{k_1}(X - \lambda_2)^{k_2} \cdots (X - \lambda_m)^{k_m}.$$

Put $f_i := f/(X-\lambda_i)^{k_i}$. Then the ideal

$$I := (f_1) + \ldots + (f_m) \subseteq \mathbb{K}[X]$$

is generated by some element g ($\mathbb{K}[X]$ is a principal ideal domain, a simple consequence of Euclid's Algorithm) which is the greatest common divisor of the polynomials f_i. The fact that the $f_1, \ldots, f_m$ have no nontrivial common divisor (cf. Exercise A.6) implies that g is constant, so that $I = \mathbb{K}[X]$. Hence $1 \in I$, so that there exist polynomials $r_1, \ldots, r_m \in \mathbb{K}[X]$ with

$$1 = r_1 f_1 + \ldots + r_m f_m.$$

Put $E_i := (r_i f_i)(M) \in \mathrm{End}(V)$ and note that $\sum_i E_i = \mathrm{id}_V$. If $i \neq j$, then f divides $r_i f_i r_j f_j$, so that $f(M) = 0$ leads to $E_i E_j = 0$, and thus $E_i^2 = E_i\big(\sum_{j=1}^m E_j\big) = E_i$. Therefore the E_i are pairwise commuting projections onto subspaces V_i with $V = \bigoplus_{i=1}^m V_i$ (since $\sum_{i=1}^m E_i = \mathbf{1}$). Now $M_s := \sum_{i=1}^m \lambda_i E_i$ is diagonalizable with $V_i = V_{\lambda_i}(M_s)$.

Since M commutes with each E_i, it preserves the subspaces V_i, and therefore $f_i(M)$ preserves V_i. The relation

$$\mathrm{id}_{V_i} = E_i|_{V_i} = r_i(M) f_i(M)|_{V_i}$$

shows that the restriction of $f_i(M)$ to V_i is invertible. Therefore $f(M) = 0$ leads to

$$(M - \lambda_i \mathbf{1})^{k_i}(V_i) = (M - \lambda_i \mathbf{1})^{k_i} f_i(M)(V_i) = f(M)(V_i) = \{0\},$$

i.e., $V_i \subseteq V^{\lambda_i}(M)$.

With $M_n := M - M_s$ and $k_0 := \max\{k_i \colon i = 1, \ldots, m\}$ we finally get $M_n^{k_0} = 0$, proving (i).

(ii) We have to show that $V_i = V^{\lambda_i}(M)$. We know already that $V_i \subseteq V^{\lambda_i}(M)$. So let $v \in V^{\lambda_i}(M)$ and write v as $v = \sum_{j=1}^m v_j$ with $v_j \in V_j$. Then the invariance of V_j under M implies that $v_j \in V^{\lambda_i}(M)$. If $v_j \neq 0$, then there exists a nonzero eigenvector $v_j' \in V_{\lambda_i}(M) \cap V_j$ (put $v_j' = (M - \lambda_i \mathbf{1})^k v_j$, where k is maximal with the property that this vector is nonzero). Then $(M - \lambda_j \mathbf{1})^{n_j} v_j' = (\lambda_i - \lambda_j)^{n_j} v_j' = 0$, hence $\lambda_j = \lambda_i$, i.e., $j = i$. This implies that $v = v_i \in V_i$ and therefore $V^{\lambda_i}(M) = V_i$.

(iii) By construction, $M_s = P_1(M)$ and $M_n = Q_1(M)$ for $P_1 = \sum_i \lambda_i r_i f_i$ and $Q_1 = X - P_1$. It remains to be seen that these polynomials can be chosen with trivial constant term. If one eigenvalue λ_j vanishes, then $\{0\} \neq V_0 := \ker M \subseteq V_j$ and $M_s|_{V_0} = 0$ implies that P_1 has no constant term. Then $Q_1 = X - P_1$ likewise has no constant term and (iii) holds with $P := P_1$ and $Q := Q_1$.

If all eigenvalues λ_j are nonzero, then $f(0) \neq 0$ and (iii) holds with

$$P := P_1 - \frac{P_1(0)}{f(0)} f \quad \text{and} \quad Q := Q_1 - \frac{Q_1(0)}{f(0)} f.$$

(iv) is a direct consequence of (iii).
(v) Since N and S commute with $M = N + S$, (iii) shows that they both commute with M_s and M_n. Then Lemma 4.18 shows that

$$S - M_s = M_n - N$$

is nilpotent as well as diagonalizable, which leads to $0 = S - M_s = M_n - N$. □

Definition A.3. The decomposition $M = M_s + M_n$ is called the *Jordan decomposition* of M, M_s is called the *semisimple Jordan component* and M_n the *nilpotent Jordan component* of M.

Example A.4. If M is a Jordan block $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, then the Jordan decomposition is

$$M = \underbrace{\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}}_{M_s} + \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{M_n}.$$

The matrix $M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ is diagonalizable and therefore $M = M_s$. In this case

$$M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$$

is not the Jordan decomposition, even though the first summand is diagonalizable and the second summand is nilpotent. These summands do not commute.

The preceding theorem does not apply to all endomorphisms of real vector spaces. We now explain how this problem can be overcome, so that we also obtain a Jordan decomposition for endomorphisms of real vector spaces.

Definition A.5. (Jordan decomposition in the real case) If V is a finite dimensional real vector space and $M \in \mathrm{End}(V)$, then $M_{\mathbb{C}} \in \mathrm{End}(V_{\mathbb{C}})$, defined by $M_{\mathbb{C}}(v + iw) := Mv + iMw$ has a Jordan decomposition

$$M_{\mathbb{C}} = M_{\mathbb{C},s} + M_{\mathbb{C},n}.$$

Let $\sigma \colon V_{\mathbb{C}} \to V_{\mathbb{C}}$ be the antilinear map defined by $\sigma(v + iw) := v - iw$ for $v, w \in V$ and define for any complex linear $A \in \mathrm{End}(V_{\mathbb{C}})$ the complex linear endomorphism $\overline{A} := \sigma \circ A \circ \sigma \in \mathrm{End}(V_{\mathbb{C}})$. Then $\overline{M_{\mathbb{C}}} = M_{\mathbb{C}}$ leads to

$$M_{\mathbb{C}} = \overline{M_{\mathbb{C}}} = \overline{M_{\mathbb{C},s}} + \overline{M_{\mathbb{C},n}},$$

where the summands on the right commute, the first is diagonalizable and the second is nilpotent (Exercise). Hence the uniqueness of the Jordan decomposition yields

$$\overline{M_{\mathbb{C},s}} = M_{\mathbb{C},s} \quad \text{and} \quad \overline{M_{\mathbb{C},n}} = M_{\mathbb{C},n}.$$

In view of Exercise A.1, this implies the existence of $M_s \in \mathrm{End}(V)$ and $M_n \in \mathrm{End}(V)$, with

$$(M_s)_{\mathbb{C}} = M_{\mathbb{C},s} \quad \text{and} \quad (M_n)_{\mathbb{C}} = M_{\mathbb{C},n}.$$

Then $M = M_s + M_n$, and this is called the *Jordan decomposition of* M. It is uniquely characterized by the properties that $[M_s, M_n] = 0$, M_s is semisimple and M_n is nilpotent (Exercise).

Proposition A.6. (Properties of the Jordan decomposition) *Let V be a finite dimensional vector space and $M \in \mathrm{End}(V)$.*

(i) *If $M' \in \mathrm{End}(V')$ and $f \colon V \to V'$ satisfy $f \circ M = M' \circ f$, then*

$$f \circ M_s = M'_s \circ f \quad \textit{and} \quad f \circ M_n = M'_n \circ f.$$

(ii) *If $W \subseteq V$ is an M-invariant subspace, then*

$$(M|_W)_s = M_s|_W \quad \textit{and} \quad (M|_W)_n = M_n|_W.$$

In particular, W is invariant under M_s and M_n. If $\overline{M}$ denotes the induced endomorphism of V/W, then

$$(\overline{M})_s = \overline{M_s} \quad \textit{and} \quad (\overline{M})_n = \overline{M_n}.$$

(iii) *If $U \subseteq W$ are subspaces of V with $MW \subseteq U$, then $M_s W \subseteq U$ and $M_n W \subseteq U$.*

Proof. (i) Let $W := V \oplus V'$, $L := M \oplus M'$, and consider the linear map $\varphi\colon W \to W$ defined by $\varphi(v, v') = \big(0, f(v)\big)$. Then $\varphi \circ L = L \circ \varphi$ and thus $\varphi L_s = L_s \varphi$ and $\varphi L_n = L_n \varphi$. Further, $L_s = M_s \oplus M'_s$ and $L_n = M_n \oplus M'_n$ follows from the uniqueness of the Jordan decomposition (Theorem A.2(v)) and the semisimplicity of $M_s \oplus M'_s$. This shows that

$$M'_s \circ f = f \circ M_s \quad \text{and} \quad M'_n \circ f = f \circ M_n.$$

(ii) Apply (i) to the inclusion $j\colon W \to V$ and the quotient map $p\colon V \to V/W$.

(iii) For $\mathbb{K} = \mathbb{C}$, this follows from Theorem A.2(iii) and the real case is obtained by complexification. □

Proposition A.7. *If $\mathcal{A}$ is a finite dimensional algebra and $D \in \operatorname{der}(\mathcal{A})$, then the Jordan components D_s and D_n are also derivations of $\mathcal{A}$.*

Proof. **First proof:** Let $m\colon \mathcal{A} \otimes \mathcal{A} \to \mathcal{A}$ denote the linear map defined by the algebra multiplication. Then $D \in \operatorname{der}(\mathcal{A})$ is equivalent to the relation

$$D \circ m = m \circ (D \otimes \mathrm{id}_{\mathcal{A}} + \mathrm{id}_{\mathcal{A}} \otimes D).$$

Next we observe that

$$D \otimes \mathrm{id}_{\mathcal{A}} + \mathrm{id}_{\mathcal{A}} \otimes D = (D_s \otimes \mathrm{id}_{\mathcal{A}} + \mathrm{id}_{\mathcal{A}} \otimes D_s) + (D_n \otimes \mathrm{id}_{\mathcal{A}} + \mathrm{id}_{\mathcal{A}} \otimes D_n)$$

is the Jordan decomposition (Exercise!), so that Proposition A.6 implies that

$$D_s \circ m = m \circ (D_s \otimes \mathrm{id}_{\mathcal{A}} + \mathrm{id}_{\mathcal{A}} \otimes D_s),$$

which means that $D_s \in \operatorname{der}(\mathcal{A})$, and hence that $D_n = D - D_s \in \operatorname{der}(\mathcal{A})$ because $\operatorname{der}(\mathcal{A})$ is a linear space.

Second proof: (for $\mathbb{K} = \mathbb{R}, \mathbb{C}$) Since $\operatorname{der}(\mathcal{A})$ is a vector space, it suffices to show that $D_s \in \operatorname{der}(\mathcal{A})$. Furthermore, $D \in \operatorname{der}(\mathcal{A})$ is equivalent to $D_{\mathbb{C}} \in \operatorname{der}(\mathcal{A}_{\mathbb{C}})$, so that we may assume that $\mathbb{K} = \mathbb{C}$.

For $a, b \in \mathcal{A}$ and $\lambda, \mu \in \mathbb{K}$ we have for all $n \in \mathbb{N}$ the formula

$$\big(D - (\lambda + \mu)\mathbf{1}\big)^n (ab) = \sum_{k=0}^{n} \binom{n}{k} (D - \lambda\mathbf{1})^k (a) \cdot (D - \mu\mathbf{1})^{n-k}(b)$$

(Exercise A.7). It follows that for $a \in \mathcal{A}_\lambda(D_s) = \mathcal{A}^\lambda(D)$ and $b \in \mathcal{A}_\mu(D_s) = \mathcal{A}^\mu(D)$, we have $ab \in \mathcal{A}^{\lambda+\mu}(D) = \mathcal{A}_{\lambda+\mu}(D_s)$. Furthermore

$$D_s(a)b + aD_s(b) = \lambda ab + \mu ab = (\lambda + \mu)ab = D_s(ab).$$

Since $\mathcal{A} = \sum_{\lambda \in \mathbb{K}} \mathcal{A}_\lambda(D_s)$, it follows that $D_s \in \operatorname{der}(\mathcal{A})$. □

Exercises for Appendix A

Exercise A.1. Let V be a real vector space and

$$V_{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} V = (1 \otimes V) \oplus (i \otimes V)$$

its complexification. We identify V with the real subspace $1 \otimes V$, so that

$$V_{\mathbb{C}} \cong V \oplus iV.$$

Show that:

(i) $\sigma(z \otimes v) := \overline{z} \otimes v$ defines an antilinear involution of $V_{\mathbb{C}}$ whose fixed point space is V.

(ii) A complex subspace $E \subseteq V_{\mathbb{C}}$ is of the form $W_{\mathbb{C}}$ for some real subspace $W \subseteq V$ if and only if $\sigma(E) = E$.

(iii) For each $M \in \mathrm{End}(V)$, the complexification $M_{\mathbb{C}} \in \mathrm{End}(V_{\mathbb{C}})$, defined by $M_{\mathbb{C}}(z \otimes v) := z \otimes Mv$ commutes with σ.

(iv) For $A \in \mathrm{End}(V_{\mathbb{C}})$ the following are equivalent

 (a) A commutes with σ.

 (b) A preserves the real subspace V.

 (c) $A = M_{\mathbb{C}}$ for some $M \in \mathrm{End}(V)$.

Exercise A.2. Let V be a complex vector space and $M \in \mathrm{End}(V)$. Show that M is diagonalizable if and only if each M-invariant subspace $W \subseteq V$ possesses an M-invariant complement.

Exercise A.3. Let V be a real vector space, $A \in \mathrm{End}(V)$ and $z \in V_{\mathbb{C}}$ an eigenvector of $A_{\mathbb{C}}$ with respect to the eigenvalue λ. Show that if $z = x + iy$ with $x, y \in V$ and $\lambda = a + ib$, then

$$Ax = ax - by \quad \text{and} \quad Ay = ay + bx.$$

In particular, the 2-dimensional subspace $E := \mathrm{span}\{x, y\} \subseteq V$ is invariant under A.

Exercise A.4. Let $A \in M_2(\mathbb{R})$ with no real eigenvalue. Then there exists a basis $x, y \in \mathbb{R}^2$ and $a, b \in \mathbb{R}$ with

$$Ax = ax - by \quad \text{and} \quad Ay = ay + bx.$$

Exercise A.5. Let V be a real vector space and $M \in \mathrm{End}(V)$. Show that M is semisimple if and only if each M-invariant subspace $W \subseteq V$ possesses an M-invariant complement.

Exercise A.6. Let $f \in \mathbb{K}[X]$ be a polynomial of the form

$$f = (X - \lambda_1)^{k_1}(X - \lambda_2)^{k_2} \cdots (X - \lambda_m)^{k_m}$$

and $g \in \mathbb{K}[X]$ a divisor of f with leading coefficient 1. Show that there exist $\ell_i \leq k_i$ with

$$g = (X - \lambda_1)^{\ell_1}(X - \lambda_2)^{\ell_2} \cdots (X - \lambda_m)^{\ell_m}.$$

Exercise A.7. Show that for each algebra $\mathcal{A}$, a derivation $D \in \mathrm{der}(\mathcal{A})$ and $\lambda, \mu \in \mathbb{K}$, we have for $a, b \in \mathcal{A}$:

$$\big(D - (\lambda + \mu)\mathbf{1}\big)^n(ab) = \sum_{k=0}^{n} \binom{n}{k} (D - \lambda\mathbf{1})^k(a) \cdot (D - \mu\mathbf{1})^{n-k}(b).$$

Exercise A.8. Let V be a finite dimensional vector space over $\mathbb{K}$ and $A \in \mathrm{End}(V)$. Then the multiplicity of the root 0 of its characteristic polynomial

$$\det(A - X\mathbf{1}) \in \mathbb{K}[X]$$

coincides with $\dim V^0(A)$.

B Tensor Products and Tensor Algebra

In this appendix we provide some tools from multilinear algebra. Throughout, $\mathbb{K}$ is an arbitrary field of characteristic zero.

Let V and W be vector spaces. A *tensor product* of V and W is a pair $(V \otimes W, \otimes)$ of a vector space $V \otimes W$ and a bilinear map

$$\otimes \colon V \times W \to V \otimes W, \quad (v, w) \mapsto v \otimes w$$

with the following universal property. For each bilinear map $\beta \colon V \times W \to U$ into a vector space U, there exists a unique linear map $\widetilde{\beta} \colon V \otimes W \to U$ satisfying

$$\widetilde{\beta}(v \otimes w) = \beta(v, w) \quad \text{for} \quad v \in V, w \in W.$$

Taking $(U, \beta) = (V \otimes W, \otimes)$, we conclude immediately that $\mathrm{id}_{V \otimes W}$ is the unique linear endomorphism of $V \otimes W$ fixing all elements of the form $v \otimes w$.

Before we turn to the existence of tensor products, we discuss their uniqueness. In category theory, one gives a precise meaning to the statement that objects with a universal property are determined up to isomorphism. The following lemma makes this precise for tensor products.

Lemma B.1. (Uniqueness of tensor products) *If $(V \otimes W, \otimes)$ and $(V \widetilde{\otimes} W, \widetilde{\otimes})$ are two tensor products of the vector spaces V and W, then there exists a unique linear isomorphism*

$$f \colon V \otimes W \to V \widetilde{\otimes} W \quad \textit{with} \quad f(v \otimes w) = v \widetilde{\otimes} w \quad \textit{for} \quad v \in V, w \in W.$$

Proof. Since $\widetilde{\otimes}$ is bilinear, the universal property of $(V \otimes W, \otimes)$ implies the existence of a unique linear map

$$f \colon V \otimes W \to V \widetilde{\otimes} W \quad \text{with} \quad f(v \otimes w) = v \widetilde{\otimes} w \quad \text{for} \quad v \in V, w \in W.$$

Similarly, the universal property of $(V \widetilde{\otimes} W, \widetilde{\otimes})$ implies the existence of a linear map

$$g \colon V \widetilde{\otimes} W \to V \otimes W \quad \text{with} \quad g(v \widetilde{\otimes} w) = v \otimes w \quad \text{for} \quad v \in V, w \in W.$$

Then $g \circ f \in \mathrm{End}(V \otimes W)$ is a linear map with $(g \circ f)(v \otimes w) = v \otimes w$ for $v \in V$ and $w \in W$, so that the uniqueness part of the universal property of $(V \otimes W, \otimes)$ yields $g \circ f = \mathrm{id}_{V \otimes W}$. We likewise get $f \circ g = \mathrm{id}_{V \widetilde{\otimes} W}$, showing that f is a linear isomorphism. $\square$

Now we turn to the existence of the tensor product.

Definition B.2. Let S be a set. We write $F(S) := \mathbb{K}^{(S)}$ for the *free vector space on S*. It is the subspace of the cartesian product $\mathbb{K}^S$, the set of all functions $f \colon S \to \mathbb{K}$ for which the set $\{s \in S \colon f(s) \neq 0\}$ is finite.

For $s \in S$, we define $\delta_s(t) := \delta_{st}$, which is 1 for $s = t$, and 0 otherwise. Then $(\delta_s)_{s \in S}$ is a basis for the vector space $F(S)$ and we have a map

$$\delta \colon S \to F(S), \quad s \mapsto \delta_s.$$

Now the pair $(F(S), \delta)$ has the universal property that, for each map $f \colon S \to V$ to a vector space V, there exists a unique linear map $\widetilde{f} \colon F(S) \to V$ with $\widetilde{f} \circ \delta = f$.

Proposition B.3. (Existence of tensor products) *If V and W are vector spaces, then there exists a tensor product $(V \otimes W, \otimes)$.*

Proof. In the free vector space $F(V \times W)$ over $V \times W$, we consider the subspace N, generated by elements of the form

$$\delta_{(v_1+v_2,w)} - \delta_{(v_1,w)} - \delta_{(v_2,w)}, \quad \delta_{(v,w_1+w_2)} - \delta_{(v,w_1)} - \delta_{(v,w_2)},$$

and

$$\delta_{(\lambda v,w)} - \delta_{(v,\lambda w)}, \quad \lambda\delta_{(v,w)} - \delta_{(\lambda v,w)},$$

for $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ and $\lambda \in \mathbb{K}$. We put

$$V \otimes W := F(V \times W)/N \quad \text{and} \quad v \otimes w := \delta_{(v,w)} + N.$$

The bilinearity of $\otimes$ follows from the definition of N. In particular, we have

$$(v_1 + v_2) \otimes w = \delta_{(v_1+v_2,w)} + N = \delta_{(v_1,w)} + \delta_{(v_2,w)} + N = v_1 \otimes w + v_2 \otimes w$$

and

$$(\lambda v) \otimes w = \delta_{(\lambda v,w)} + N = \lambda\delta_{(v,w)} + N = \lambda(v \otimes w).$$

The linearity in the second argument is verified similarly.

To show that $(V \otimes W, \otimes)$ has the required universal property, let $\beta\colon V \times W \to U$ be a bilinear map. We use the universal property of $(F(V \times W), \delta)$ to obtain a linear map

$$\gamma\colon F(V \times W) \to U \quad \text{with} \quad \gamma(\delta_{(v,w)}) = \beta(v, w)$$

for $v \in V, w \in W$. The bilinearity of β now implies that $N \subseteq \ker\gamma$, so that γ factors through a unique linear map

$$\widetilde{\beta}\colon V \otimes W = F(V \times W)/N \to U \quad \text{with} \quad \widetilde{\beta}(v \otimes w) = \gamma(\delta_{(v,w)}) = \beta(v, w).$$

That $\widetilde{\beta}$ is uniquely determined by this property follows from the fact that the elements of the form $v \otimes w$ generate $V \otimes W$ linearly, which in turn follows from $\delta(V \times W)$ being a linear basis for $F(V \times W)$. $\square$

Tensor products of finitely many factors are defined in a similar fashion as follows.

Definition B.4. Let $V_1, \ldots, V_k$ be vector spaces. A *tensor product* of $V_1, \ldots, V_k$ is a pair

$$(V_1 \otimes V_2 \otimes \cdots \otimes V_k, \otimes)$$

of a vector space $V_1 \otimes V_2 \otimes \cdots \otimes V_k$ and a k-linear map

$$\otimes\colon V_1 \times \cdots \times V_k \to V_1 \otimes V_2 \otimes \cdots \otimes V_k, \quad (v_1, \ldots, v_k) \mapsto v_1 \otimes \cdots \otimes v_k,$$

with the following universal property. For each k-linear map

$$\beta\colon V_1 \times \cdots \times V_k \to U$$

into a vector space U, there exists a unique linear map $\widetilde{\beta}\colon V_1 \otimes \cdots \otimes V_k \to U$ satisfying

$$\widetilde{\beta}(v_1 \otimes \cdots \otimes v_k) = \beta(v_1, \ldots, v_k) \quad \text{for} \quad v_i \in V_i.$$

For $(U, \beta) = (V_1 \otimes \cdots \otimes V_k, \otimes)$, we conclude immediately that $\mathrm{id}_{V_1 \otimes \cdots \otimes V_k}$ is the unique linear endomorphism of $V_1 \otimes \cdots \otimes V_k$ fixing all elements of the form $v_1 \otimes \cdots \otimes v_k$.

Again, the universal property determines k-fold tensor products.

Lemma B.5. (Uniqueness of k-fold tensor products) *If*

$$(V_1 \otimes \cdots \otimes V_k, \otimes) \quad \text{and} \quad (V_1 \widetilde{\otimes} \cdots \widetilde{\otimes} V_k, \widetilde{\otimes})$$

are two tensor products of the vector spaces $V_1, \ldots, V_k$, then there exists a unique linear isomorphism

$$f \colon V_1 \otimes \cdots \otimes V_k \to V_1 \widetilde{\otimes} \cdots \widetilde{\otimes} V_k \quad \text{with} \quad f(v_1 \otimes \cdots \otimes v_k) = v_1 \widetilde{\otimes} \cdots \widetilde{\otimes} v_k$$

for $v_i \in V_i$.

We omit the simple proof of the uniqueness. The existence is easily reduced to the two-fold case:

Lemma B.6. *If $V_1, \ldots, V_k$ are vector spaces and $k \geq 2$, then the iterated two-fold tensor product*

$$V_1 \otimes \cdots \otimes V_k := (V_1 \otimes \cdots \otimes V_{k-1}) \otimes V_k$$

and

$$v_1 \otimes \cdots \otimes v_k := (v_1 \otimes \cdots \otimes v_{k-1}) \otimes v_k$$

is a tensor product of $V_1, \ldots, V_k$.

Proof. Since we know already that this is true for $k = 2$, we argue by induction and assume that the assertion holds for $(k-1)$-fold iterated tensor products. In this way we immediately see that $(v_1 \otimes \cdots \otimes v_{k-1}) \otimes v_k$ is k-linear.

To verify the universal property, let $\beta \colon V_1 \times \cdots \times V_k \to U$ be a k-linear map. We first use the induction hypothesis to obtain for each $v_k \in V_k$ a unique linear map $\widetilde{\beta}_{v_k} \colon V_1 \otimes \cdots \otimes V_{k-1} \to U$ with

$$\widetilde{\beta}_{v_k}(v_1 \otimes \ldots \otimes v_{k-1}) = \beta(v_1, \ldots, v_{k-1}, v_k) \quad \text{for} \quad v_i \in V_i, i \leq k-1.$$

From the uniqueness of $\widetilde{\beta}_{v_k}$ we further derive that

$$\widetilde{\beta}_{\lambda v_k + \lambda' v_k'} = \lambda \widetilde{\beta}_{v_k} + \lambda' \widetilde{\beta}_{v_k'}$$

for $\lambda, \lambda' \in \mathbb{K}$ and $v_k, v_k' \in V_k$. Hence the map

$$(V_1 \otimes \cdots \otimes V_{k-1}) \times V_k \to U, \quad (x, v_k) \mapsto \widetilde{\beta}_{v_k}(x)$$

is bilinear. Now the universal property of the two-fold tensor product provides a unique linear map

$$\widetilde{\beta} \colon (V_1 \otimes \cdots \otimes V_{k-1}) \otimes V_k \to U$$

with $\widetilde{\beta}((v_1 \otimes \cdots \otimes v_{k-1}) \otimes v_k) = \widetilde{\beta}_{v_k}(v_1 \otimes \cdots \otimes v_{k-1}) = \beta(v_1, \ldots, v_{k-1}, v_k)$. $\square$

Definition B.7. (The tensor algebra of a vector space) Let V be a $\mathbb{K}$-vector space and $V^{\otimes n}$ the n-fold tensor product of V with itself. For $n = 0, 1$, we put $V^{\otimes 0} := \mathbb{K}$ and $V^{\otimes 1} := V$.

We claim that, for $n, m \in \mathbb{N}$, there exists a bilinear map

$$\mu_{n,m} \colon V^{\otimes n} \times V^{\otimes m} \to V^{\otimes (n+m)}$$

with

$$\mu_{n,m}\big((v_1 \otimes \ldots \otimes v_n), (v_{n+1} \otimes \ldots \otimes v_{n+m})\big) = v_1 \otimes \ldots \otimes v_{n+m}$$

for $v_1, \ldots, v_{n+m} \in V$. In fact, for each $\mathbf{x} = (x_1, \ldots, x_n) \in V^n$, the map

$$\mu_{\mathbf{x}} \colon V^m \to V^{\otimes (n+m)}, \quad (w_1, \ldots, w_m) \mapsto x_1 \otimes \cdots \otimes x_n \otimes w_1 \otimes \ldots \otimes w_m$$

is m-linear, hence determines a linear map

$$\widetilde{\mu}_{\mathbf{x}} : V^{\otimes m} \to V^{\otimes(n+m)} \quad \text{with} \quad \widetilde{\mu}_{\mathbf{x}}(w_1 \otimes \cdots \otimes w_m) = \mu_{\mathbf{x}}(w_1, \ldots, w_m).$$

Since $\mu_{\mathbf{x}}$ is n-linear in $\mathbf{x}$, we obtain a uniquely determined bilinear map

$$\mu_{n,m} : V^{\otimes n} \times V^{\otimes m} \to V^{\otimes(n+m)}$$

with

$$\begin{aligned} &\mu_{n,m}\big((v_1 \otimes \ldots \otimes v_n), (v_{n+1} \otimes \ldots \otimes v_{n+m})\big) \\ &= \widetilde{\mu}_{(v_1 \otimes \ldots \otimes v_n)}(v_{n+1} \otimes \ldots \otimes v_{n+m}) = v_1 \otimes \ldots \otimes v_n \otimes v_{n+1} \otimes \ldots \otimes v_{n+m}. \end{aligned}$$

We further define bilinear maps

$$\mu_{0,n} : V^{\otimes 0} \times V^{\otimes n} = \mathbb{K} \times V^{\otimes n} \to V^{\otimes n}, \quad (\lambda, v) \mapsto \lambda v$$

and

$$\mu_{n,0} : V^{\otimes n} \otimes V^{\otimes 0} = V^{\otimes n} \times \mathbb{K} \to V^{\otimes n}, \quad (v, \lambda) \mapsto \lambda v.$$

Putting all maps $\mu_{n,k}$, $n, k \in \mathbb{N}_0$, together, we obtain a bilinear multiplication on the vector space

$$\mathcal{T}(V) := \bigoplus_{n=0}^{\infty} V^{\otimes n}.$$

It is now easy to show that this multiplication is associative and has an identity element $\mathbf{1} \in V^{\otimes 0}$ (Exercise B.5). The algebra obtained in this way is called the *tensor algebra of* V.

Lemma B.8. (Universal property of the tensor algebra) *Let V be a vector space and $\eta : V \to \mathcal{T}(V)$ the canonical embedding of V as $V^{\otimes 1}$. Then the pair $(\mathcal{T}(V), \eta)$ has the following property. For any linear map $f : V \to \mathcal{A}$ into a unital associative $\mathbb{K}$-algebra $\mathcal{A}$, there exists a unique homomorphism $\widetilde{f} : \mathcal{T}(V) \to \mathcal{A}$ of unital associative algebras with $\widetilde{f} \circ \eta = f$.*

Proof. For the uniqueness of $\widetilde{f}$ we first note that the requirement of being a homomorphism of unital algebras determines $\widetilde{f}$ on $\mathbf{1}$ via $\widetilde{f}(\mathbf{1}) = \mathbf{1}_{\mathcal{A}}$. On $\eta(V) = V^{\otimes 1}$ it is determined by $\widetilde{f} \circ \eta = f$, and on $\mathcal{T}(V)$ it is thus determined since the algebra $\mathcal{T}(V)$ is generated by the subspace $\mathbb{K}\mathbf{1} + V$.

For the existence of $\widetilde{f}$, we note that, for each $n \in \mathbb{N}$, the map

$$V^n \to \mathcal{A}, \quad (v_1, \ldots, v_n) \mapsto f(v_1) \cdots f(v_n)$$

is n-linear, so that there exists a unique linear map

$$\widetilde{f}_n : V^{\otimes n} \to \mathcal{A} \quad \text{with} \quad \widetilde{f}(v_1 \otimes \cdots \otimes v_n) = f(v_1) \cdots f(v_n)$$

for $v_i \in V$. We now combine these linear maps $\widetilde{f}_n$ to a linear map

$$\widetilde{f} : \mathcal{T}(V) \to \mathcal{A} \quad \text{with} \quad \widetilde{f}_n(\mathbf{1}) = \mathbf{1}_{\mathcal{A}}, \quad \widetilde{f}|_{V^{\otimes n}} = \widetilde{f}_n.$$

Then the construction implies that $\widetilde{f} \circ \eta = f$. That $\widetilde{f}$ is an algebra homomorphism follows from

$$\begin{aligned} \widetilde{f}\big((v_1 \otimes \cdots \otimes v_n) \cdot (w_1 \otimes \cdots \otimes w_m)\big) &= f(v_1) \cdots f(v_n) f(w_1) \cdots f(w_m) \\ &= \widetilde{f}(v_1 \otimes \cdots \otimes v_n) \widetilde{f}(w_1 \otimes \cdots \otimes w_m) \end{aligned}$$

for $v_1, \ldots, v_n, w_1, \ldots, w_m \in V$. $\square$

Exercises for Section B

Exercise B.1. Let U, V and W be finite dimensional vector spaces. Show that there are isomorphisms:

(i) $U \otimes V \cong V \otimes U$.

(ii) $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$.

Exercise B.2. The aim of this exercise is to get a more concrete picture of the tensor product of two vector spaces in terms of bases. Let V and W be vector spaces. We consider a basis $B_V = \{e_i \colon i \in I\}$ for V and a basis $B_W = \{f_j \colon j \in J\}$ for W. Show that:

(i) Each function $f \colon B_V \times B_W \to \mathbb{K}$ has a unique bilinear extension $\widetilde{f} \colon V \times W \to \mathbb{K}$.

(ii) The set $B_V \otimes B_W = \{e_i \otimes f_j \colon i \in I, j \in J\}$ is a basis for $V \otimes W$.

(iii) Each element $x \in V \otimes W$ has a unique representation as a finite sum $x = \sum_{i\in I} e_i \otimes w_i$ with $w_i \in W$.

(iv) If V_1 and V_2 are vector spaces, then $(V_1 \oplus V_2) \otimes W \cong (V_1 \otimes W) \oplus (V_2 \otimes W)$.

Exercise B.3. Let $V := \mathbb{K}^n$ and $W := \mathbb{K}^m$. Show that one can turn the space $M_{n,m}(\mathbb{K})$ of $(n \times m)$-matrices with entries in $\mathbb{K}$ into a tensor product $(\mathbb{K}^n \otimes \mathbb{K}^m, \otimes)$ satisfying

$$e_i \otimes e_j := E_{ij},$$

where $e_1, \ldots, e_n$ denotes the canonical basis vectors in $\mathbb{K}^n$ and E_{ij} is the matrix which has a single nonzero entry in the i-th row and the j-th column.

Exercise B.4. If V and W are finite dimensional, then the map

$$\Phi \colon V^* \otimes W \to \mathrm{Hom}(V, W), \qquad \Phi(\alpha \otimes w)(v) := \alpha(v)w$$

is a linear isomorphism.

Exercise B.5. Let V be a vector space and $\mathcal{T}(V) = \bigoplus_{n\in\mathbb{N}_0} V^{\otimes n}$. Show that the multiplication on $\mathcal{T}(V)$ defined by Definition B.7 yields an associative $\mathbb{K}$-algebra.

Exercise B.6. Let V_i and W_i be $\mathbb{K}$-vector spaces (for $i = 1, 2$) and $A \in \mathrm{Hom}_{\mathbb{K}}(V_1, V_2)$, $B \in \mathrm{Hom}_{\mathbb{K}}(W_1, W_2)$. Show that there exists a unique $\mathbb{K}$-linear map $C \colon V_1 \otimes V_2 \to W_1 \otimes W_2$ such that

$$C(v_1 \otimes v_2) = A(v_1) \otimes B(v_2)$$

for all $v_1 \in V_1$ and $v_2 \in V_2$. The map C is usually denoted by $A \otimes B$.

Exercise B.7. Suppose that $V_1, \ldots, V_k$ are vector spaces and that a group G acts linearly on each of them. Show that

$$g \cdot (v_1 \otimes \ldots \otimes v_k) := g \cdot v_1 \otimes \ldots \otimes g \cdot v_k$$

for $g \in G$ and $v_j \in V_j$ defines a linear action on $V_1 \otimes \ldots \otimes V_k$.

C Symmetric and Exterior Products

C.1 Symmetric and Exterior Powers

Definition C.1. Let V be a vector space and $n \geq 2$. We define

$$S^n(V) := V^{\otimes n}/U,$$

where U is the subspace spanned by all elements of the form

$$v_1 \otimes \ldots \otimes v_n - v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(n)}, \quad \sigma \in S_n.$$

The space $S^n(V)$ is called the *n-th symmetric power of V*. We put

$$v_1 \vee \cdots \vee v_n := v_1 \otimes \cdots \otimes v_n + U$$

and observe that this product is symmetric in the sense that

$$v_1 \vee \cdots \vee v_n = v_{\sigma(1)} \vee \cdots \vee v_{\sigma(n)}$$

for each $\sigma \in S_n$ and $v_1, \ldots, v_n \in V$. For $n = 1$, we put $S^1(V) := V$ and also $S^0(V) := \mathbb{K}$.

If X and Y are sets, then a map $f \colon X^n \to Y$ is said to be *symmetric* if, for each permutation $\sigma \in S_n$, we have

$$f(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \quad \text{for} \quad x \in X^n.$$

Lemma C.2. (Universal property of symmetric powers) *Let V and X be vector spaces and $f \colon V^n \to X$ be a symmetric n-linear map. Then there exists a unique linear map $\widetilde{f} \colon S^n(V) \to X$ with*

$$\widetilde{f}(v_1 \vee \cdots \vee v_n) = f(v_1, \ldots, v_n) \quad \textit{for} \quad v_1, \ldots, v_n \in V.$$

Proof. From the universal property of the n-fold tensor product $V^{\otimes n}$, we obtain a unique linear map $f_0 \colon V^{\otimes n} \to X$ with

$$f_0(v_1 \otimes \cdots \otimes v_n) = f(v_1, \ldots, v_n) \quad \text{for} \quad v_1, \ldots, v_n \in V.$$

In view of the symmetry of f, the linear map f_0 vanishes on U, hence factors through a linear map $\widetilde{f} \colon S^n(V) \to X$ with the desired property. $\square$

Definition C.3. Let V and W be $\mathbb{K}$-vector spaces, $n \in \mathbb{N}$, and

$$\operatorname{sgn} \colon S_n \to \{1, -1\}$$

be the *signature homomorphism* mapping all transpositions to -1. An n-linear map $f \colon V^n \to W$ is called *alternating* if

$$f(v_1, \ldots, v_n) = \operatorname{sgn}(\sigma) f(v_{\sigma(1)}, \ldots, v_{\sigma(n)})$$

holds for all $\sigma \in S_n$ and $v_1, \ldots, v_n \in V$.

We write $\operatorname{Alt}^n(V, W)$ for the set of alternating n-linear maps $V^n \to W$. Clearly, sums and scalar multiples of alternating maps are alternating, so that $\operatorname{Alt}^n(V, W)$ carries a natural vector space structure. For $n = 0$, we shall follow the convention that $\operatorname{Alt}^0(V, W) := W$ is the set of constant maps, which are considered to be 0-linear.

Example C.4. From linear algebra, we know the n-linear map

$$(\mathbb{K}^n)^n \to \mathbb{K}, \quad \det(v_1, \ldots, v_k) := \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) v_{1,\sigma(1)} \cdots v_{k,\sigma(k)}.$$

Here we identify the space $M_n(\mathbb{K})$ of $(n \times n)$-matrices with entries in $\mathbb{K}$ with the space $(\mathbb{K}^n)^n$ of n-tuples of (column) vectors ([La93, Sect. XIII.4]).

Definition C.5. Let V be a vector space and $n \geq 2$. We define

$$\Lambda^n(V) := V^{\otimes n}/W,$$

where W is the subspace spanned by the elements of the form

$$v_1 \otimes \cdots \otimes v_n - \operatorname{sgn}(\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}, \quad \sigma \in S_n.$$

The space $\Lambda^n(V)$ is called the *n-th exterior power of V*. We put

$$v_1 \wedge \cdots \wedge v_n := v_1 \otimes \cdots \otimes v_n + W$$

and note that this product is alternating, i.e.,

$$v_1 \wedge \cdots \wedge v_n = \operatorname{sgn}(\sigma) v_{\sigma(1)} \wedge \cdots \wedge v_{\sigma(n)}$$

for all $\sigma \in S_n$ and $(v_1, \ldots, v_n) \in V^n$. For $n = 2$, this means that

$$v_1 \wedge v_2 = -v_2 \wedge v_1.$$

We also put $\Lambda^1(V) := V$ and $\Lambda^0(V) := \mathbb{K}$.

Lemma C.6. *[Universal property of the exterior power] Let V and X be vector spaces and $f \in \operatorname{Alt}^n(V, X)$. Then there exists a unique linear map $\widetilde{f} \colon \Lambda^n(V) \to X$ with*

$$\widetilde{f}(v_1 \wedge \cdots \wedge v_n) = f(v_1, \ldots, v_n) \quad \textit{for} \quad v_1, \ldots, v_n \in V.$$

We thus obtain a linear bijection

$$\operatorname{Alt}^n(V, X) \to \operatorname{Hom}(\Lambda^n(V), X), \quad f \mapsto \widetilde{f}.$$

Proof. The proof is completely analogous to the symmetric case. □

C.2 Symmetric and Exterior Algebra

Definition C.7. Let V be a vector space and $(\mathcal{T}(V), \eta)$ the tensor algebra of V (cf. Lemma B.8). We define the *symmetric algebra* $S(V)$ over V as the quotient $\mathcal{T}(V)/I_s$, where I_s is the ideal generated by the elements $\eta(v) \otimes \eta(w) - \eta(w) \otimes \eta(v)$. We write

$$\eta_s \colon V \to S(V), \quad v \mapsto \eta(v) + I_s$$

for the canonical map induced by η. The product in $S(V)$ is denoted by $\vee$.

Likewise, we define the *exterior algebra* $\Lambda(V)$ over V as the quotient $\mathcal{T}(V)/I_a$, where I_a is the ideal generated by the elements

$$\eta(v) \otimes \eta(w) + \eta(w) \otimes \eta(v), \quad v, w \in V.$$

We write

$$\eta_a \colon V \to \Lambda(V), \quad v \mapsto \eta(v) + I_a$$

for the canonical map induced by η. The product in $\Lambda(V)$ is denoted by $\wedge$.

Lemma C.8. (Universal property of the symmetric algebra) *Let V be a vector space and $(S(V), \eta_s)$ its symmetric algebra. Then $S(V)$ is a commutative unital algebra and for any linear map $f\colon V \to A$ into a unital commutative associative algebra A, there exists a unique homomorphism $\widetilde{f}\colon S(V) \to A$ of unital associative algebras with $\widetilde{f} \circ \eta_s = f$.*

Proof. Using the universal property of the tensor algebra $\mathcal{T}(V)$, we see that there exists a unique unital algebra homomorphism $\widehat{f}\colon \mathcal{T}(V) \to A$ with $\widehat{f} \circ \eta = f$. Since A is commutative, for any $v, w \in V$, the element $\eta(v) \otimes \eta(w) - \eta(w) \otimes \eta(v)$ is contained in $\ker \widehat{f}$, and therefore $I_s \subseteq \ker \widehat{f}$ shows that $\widehat{f}$ factors through an algebra homomorphism $\widetilde{f}\colon S(V) \to A$ with $\widetilde{f} \circ \eta_s = f$. The uniqueness of $\widetilde{f}$ follows from the fact that $\mathcal{T}(V)$ is generated, as a unital algebra, by $\eta(V)$, so that $S(V)$ is generated by the image of η_s. Since the generators $\eta_s(v)$, $v \in V$, of $S(V)$ commute, the algebra $S(V)$ is commutative. $\square$

Remark C.9. (a) The structure of the symmetric algebra can be made more concrete as follows. Let $\mathcal{T}(V)_k := V^{\otimes k}$ and $U_2 \subseteq \mathcal{T}(V)_2$ the subspace spanned by the commutators $[\eta(v), \eta(w)], v, w \in V$. Then the ideal I_s is of the form

$$I_s = \mathcal{T}(V)U_2\mathcal{T}(V) = \sum_{p,q \in \mathbb{N}_0} \mathcal{T}(V)_p \otimes U_2 \otimes \mathcal{T}(V)_q = \bigoplus_{n=2}^{\infty} I_{s,n},$$

where $I_{s,n} := \sum_{p+q=n-2} \mathcal{T}(V)_p \otimes U_2 \otimes \mathcal{T}(V)_q$. This implies that the symmetric algebra $S(V)$ is a direct sum

$$S(V) = \bigoplus_{n=0}^{\infty} S(V)_n, \quad \text{where} \quad S(V)_n := \mathcal{T}(V)_n / I_{s,n}.$$

Let

$$\mu_n\colon V^n \to S(V)_n, \quad (v_1, \ldots, v_n) \mapsto \eta_s(v_1) \vee \cdots \vee \eta_s(v_n)$$

denote the n-fold multiplication map. Since $S(V)$ is commutative, this map is symmetric, hence induces a linear map

$$\widetilde{\mu}_n\colon S^n(V) \to S(V)_n,$$

determined by

$$\widetilde{\mu}_n(v_1 \vee \cdots \vee v_n) = \eta_s(v_1) \vee \cdots \vee \eta_s(v_n).$$

On the other hand, it is clear that the subspace $I_{s,n}$ of $V^{\otimes n}$ is contained in the kernel of the quotient map $V^{\otimes n} \to S^n(V)$, so that there exists a linear map $f_n\colon S(V)_n \to S^n(V)$, with

$$f_n(\eta_s(v_1) \vee \cdots \vee \eta_s(v_n)) = v_1 \vee \cdots \vee v_n.$$

Then $f_n \circ \widetilde{\mu}_n = \mathrm{id}_{S^n(V)}$ and, similarly, $\widetilde{\mu}_n \circ f_n = \mathrm{id}_{S(V)_n}$. This proves that $\widetilde{\mu}_n$ is a linear isomorphism. In the following we therefore identify $S^n(V)$ with the subspace $S(V)_n$ of the symmetric algebra and write $\eta_s(v)$ simply as v.

Note that $S^n(V) \vee S^m(V) \subseteq S^{n+m}(V)$, so that the direct sum

$$S(V) = \bigoplus_{n \in \mathbb{N}} S^n(V)$$

defines the structure of a *graded algebra* on $S(V)$ with $S^0(V) = \mathbb{K}\mathbf{1}$ containing the identity element.

(b) A similar argument applies to the exterior algebra and shows that the ideal I_a has the form $I_a = \bigoplus_{n=2}^\infty (I_a \cap V^{\otimes n})$, so that

$$\Lambda(V) = \bigoplus_{n=0}^\infty \Lambda(V)_n, \quad \text{where} \quad \Lambda(V)_n := \mathcal{T}(V)_n / I_{a,n}.$$

Let $\mu_n \colon V^n \to \Lambda(V)_n, (v_1, \ldots, v_n) \mapsto \eta_a(v_1) \wedge \cdots \wedge \eta_a(v_n)$ denote the n-fold multiplication map. Then the relation $\eta_a(v_i)\eta_a(v_j) + \eta_a(v_j)\eta_a(v_i) = 0$ and the fact that S_n is generated by transpositions imply that μ_n is alternating. Hence it induces a linear map $\widetilde{\mu}_n \colon \Lambda^n(V) \to \Lambda(V)_n$, determined by

$$\widetilde{\mu}_n(v_1 \wedge \cdots \wedge v_n) = \eta_a(v_1) \wedge \cdots \wedge \eta_a(v_n).$$

On the other hand, it is clear that the subspace $I_{a,n}$ of $V^{\otimes n}$ is contained in the kernel of the quotient map $V^{\otimes n} \to \Lambda^n(V)$, so that there is a linear map $f_n \colon \Lambda(V)_n \to \Lambda^n(V)$ with

$$f_n(\eta_a(v_1) \wedge \cdots \wedge \eta_a(v_n)) = v_1 \wedge \cdots \wedge v_n.$$

As in the symmetric case, we now see that $\widetilde{\mu}_n$ is a linear isomorphism. In the following we therefore identify $\Lambda^n(V)$ with the subspace $\Lambda(V)_n$ of the symmetric algebra and write $\eta_a(v)$ simply as v.

Each subspace $\Lambda^n(V)$ is spanned by elements of the form $v_1 \wedge \cdots \wedge v_n$, and this implies that for $\alpha \in \Lambda^n(V)$ and $\beta \in \Lambda^m(V)$ we have

$$\alpha \wedge \beta = (-1)^{mn} \beta \wedge \alpha. \tag{37}$$

In this sense the graded algebra $\Lambda(V)$ is *graded commutative.* The even part of this algebra is the subspace

$$\Lambda^{\text{even}}(V) := \bigoplus_{k=0}^\infty \Lambda^{2k}(V)$$

which is a central subalgebra, and the odd part is

$$\Lambda^{\text{odd}}(V) := \bigoplus_{k=0}^\infty \Lambda^{2k+1}(V).$$

For two elements α, β of this subspace we have $\alpha \wedge \beta = -\beta \wedge \alpha$.

Lemma C.10. (Universal property of the exterior algebra) *Let V be a vector space and $(\Lambda(V), \eta_a)$ be its exterior algebra. Then $\Lambda(V)$ is a graded commutative unital algebra and for any linear map $f \colon V \to A$ into a unital associative algebra A, satisfying*

$$f(v)f(w) = -f(w)f(v) \quad \text{for} \quad v, w \in V,$$

there exists a unique homomorphism $\widetilde{f} \colon \Lambda(V) \to A$ of unital associative algebras with $\widetilde{f} \circ \eta_a = f$.

Proof. Using the universal property of the tensor algebra $\mathcal{T}(V)$, we see that there exists a unique unital algebra homomorphism $\widehat{f} \colon \mathcal{T}(V) \to A$ with $\widehat{f} \circ \eta = f$. Then we have for $v, w \in V$

$$\widehat{f}\big(\eta(v) \otimes \eta(w) + \eta(w) \otimes \eta(v)\big) = f(v)f(w) + f(w)f(v) = 0.$$

Therefore $I_a \subseteq \ker \widehat{f}$ shows that $\widehat{f}$ factors through a unital algebra homomorphism $\widetilde{f} \colon \Lambda(V) \to A$ with $\widetilde{f} \circ \eta_a = f$. The uniqueness of $\widetilde{f}$ follows from the fact that $\mathcal{T}(V)$ is generated, as a unital algebra, by $\eta(V)$, so that $\Lambda(V)$ is generated by the image of η_a. $\square$

C.3 Exterior Algebra and Alternating Maps

Below we shall see how general alternating maps can be expressed in terms of determinants.

Proposition C.11. *For any $\omega \in \mathrm{Alt}^k(V,W)$ we have:*

(i) *For $b_1,\dots,b_k \in V$ and linear combinations $v_j = \sum_{i=1}^k a_{ij} b_i$, we have*

$$\omega(v_1,\dots,v_k) = \det(A)\omega(b_1,\dots,b_k), \qquad \text{and} \qquad A := (a_{ij}) \in M_k(\mathbb{K}).$$

(ii) *$\omega(v_1,\dots,v_k) = 0$ if $v_1,\dots,v_k$ are linearly dependent.*

(iii) *For $b_1,\dots,b_n \in V$ and linear combinations $v_j = \sum_{i=1}^n a_{ij} b_i$ we have*

$$\omega(v_1,\dots,v_k) = \sum_I \det(A_I)\omega(b_{i_1},\dots,b_{i_k}),$$

where $A := (a_{ij}) \in M_{n,k}(\mathbb{K})$, $I = \{i_1,\dots,i_k\}$ is a k-element subset of $\{1,\dots,n\}$, $1 \le i_1 < \ldots < i_k \le n$, and $A_I := (a_{ij})_{i \in I, j=1,\dots,k} \in M_k(\mathbb{K})$.

Proof. (i) For the following calculation we note that if $\sigma\colon \{1,\dots,k\} \to \{1,\dots,k\}$ is a map which is not bijective, then the alternating property implies that $\omega(v_{\sigma(1)},\dots,v_{\sigma(k)}) = 0$. We therefore get

$$\begin{aligned}
\omega(v_1,\dots,v_k) &= \omega\Big(\sum_{i=1}^k a_{i1}b_i,\dots,\sum_{i=1}^k a_{ik}b_i\Big) \\
&= \sum_{i_1,\dots,i_k=1}^k a_{i_1 1}\cdots a_{i_k k}\cdot\omega(b_{i_1},\dots,b_{i_k}) \\
&= \sum_{\sigma\in S_k} a_{\sigma(1)1}\cdots a_{\sigma(k)k}\cdot\omega(b_{\sigma(1)},\dots,b_{\sigma(k)}) \\
&= \sum_{\sigma\in S_k} \mathrm{sgn}(\sigma) a_{\sigma(1)1}\cdots a_{\sigma(k)k}\cdot\omega(b_1,\dots,b_k) = \det(A)\cdot\omega(b_1,\dots,b_k).
\end{aligned}$$

(ii) follows immediately from (i) because the linear dependence of $v_1,\dots,v_k$ implies that $\det A = 0$.

(iii) First we expand

$$\begin{aligned}
\omega(v_1,\dots,v_k) &= \omega\Big(\sum_{i=1}^n a_{i1}b_i,\dots,\sum_{i=1}^n a_{ik}b_i\Big) \\
&= \sum_{i_1,\dots,i_k=1}^n a_{i_1 1}\cdots a_{i_k k}\cdot\omega(b_{i_1},\dots,b_{i_k}).
\end{aligned}$$

If $|\{i_1,\dots,i_k\}| < k$, then the alternating property implies that $\omega(b_{i_1},\dots,b_{i_k}) = 0$ because two entries coincide. If $|\{i_1,\dots,i_k\}| = k$, there exists a permutation $\sigma \in S_k$ with $i_{\sigma(1)} < \ldots < i_{\sigma(k)}$. We therefore get

$$\begin{aligned}
\omega(v_1,\dots,v_k) &= \sum_{1\le i_1<\ldots<i_k\le n}\ \sum_{\sigma\in S_k} a_{i_{\sigma(1)}1}\cdots a_{i_{\sigma(k)}k}\cdot\omega(b_{i_{\sigma(1)}},\dots,b_{i_{\sigma(k)}}) \\
&= \sum_{1\le i_1<\ldots<i_k\le n}\ \sum_{\sigma\in S_k} \mathrm{sgn}(\sigma) a_{i_{\sigma(1)}1}\cdots a_{i_{\sigma(k)}k}\cdot\omega(b_{i_1},\dots,b_{i_k}) \\
&= \sum_I \det(A_I)\omega(b_{i_1},\dots,b_{i_k}),
\end{aligned}$$

where the sum is to be extended over all k-element subsets $I = \{i_1,\dots,i_k\}$ of $\{1,\dots,n\}$, where $i_1 < \ldots < i_k$. $\square$

Corollary C.12. (i) *If* $\dim V < k$, *then* $\mathrm{Alt}^k(V,W) = \{0\}$.

(ii) *Let* $\dim V = n$ *and* $b_1, \ldots, b_n$ *be a basis for* V. *Then the map*

$$\Phi\colon \mathrm{Alt}^k(V,W) \to W^{\binom{n}{k}}, \quad \Phi(\omega) = (\omega(b_{i_1}, \ldots, b_{i_k}))_{i_1 < \ldots < i_k}$$

is a linear isomorphism. We obtain in particular $\dim\big(\mathrm{Alt}^k(V,\mathbb{K})\big) = \binom{n}{k}$.

(iii) *If* $\dim V = k$ *and* $b_1, \ldots, b_k$ *is a basis for* V, *then the map*

$$\Phi\colon \mathrm{Alt}^k(V,W) \to W, \quad \Phi(\omega) = \omega(b_1, \ldots, b_k)$$

is a linear isomorphism.

Proof. (i) In Proposition C.11(i), we may choose $b_k = 0$.

(ii) First we show that Φ is injective. So let $\omega \in \mathrm{Alt}^k(V,W)$ with $\Phi(\omega) = 0$. We now write any k elements $v_1, \ldots, v_k \in V$ with respect to the basis elements as $v_j = \sum_{i=1}^n a_{ij} b_i$ and obtain with Proposition C.11:

$$\omega(v_1, \ldots, v_k) = \sum_{1 \leq i_1 < \ldots < i_k \leq n} \det(A_I)\omega(b_{i_1}, \ldots, b_{i_k}) = 0.$$

To see that Φ is surjective, we pick for each k-element subset $I = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ with $1 \leq i_1 < \ldots < i_k \leq n$ an element $w_I \in W$. Then the tuple (w_I) is a typical element of $W^{\binom{n}{k}}$.

Expressing k elements $v_1, \ldots, v_k$ in terms of the basis elements $b_1, \ldots, b_n$ via $v_j = \sum_{i=1}^n a_{ij} b_i$, we obtain an $(n \times k)$-matrix A. We now define an alternating k-linear map $\omega \in \mathrm{Alt}^k(V,W)$ by

$$\omega(v_1, \ldots, v_k) := \sum_I \det(A_I) w_I.$$

The k-linearity of ω follows directly from the k-linearity of the maps

$$(v_1, \ldots, v_k) \mapsto \det(A_I).$$

For $i_1 < \ldots < i_k$ we further have $\omega(b_{i_1}, \ldots, b_{i_k}) = w_I$ because in this case $A_I \in M_k(\mathbb{K})$ is the identity matrix and all other matrices $A_{I'}$ have some vanishing columns. This implies that $\Phi(\omega) = (w_I)$, and hence that Φ is surjective.

(iii) is a special case of (ii). $\square$

Definition C.13. (Alternator) Let V and W be vector spaces. For a k-linear map $\omega\colon V^k \to W$, we define a new k-linear map by

$$\mathrm{Alt}(\omega)(v_1, \ldots, v_k) := \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma)\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}).$$

Writing

$$\omega^\sigma(v_1, \ldots, v_k) := \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}),$$

we then have

$$\mathrm{Alt}(\omega) = \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma)\omega^\sigma.$$

The map $\mathrm{Alt}(\cdot)$ is called the *alternator*. We claim that it turns any k-linear map into an alternating k-linear map. To see this, we first note that for $\sigma, \pi \in S_k$, we have

$$\begin{aligned}(\omega^\sigma)^\pi(v_1, \ldots, v_k) &= (\omega^\sigma)(v_{\pi(1)}, \ldots, v_{\pi(k)}) \\ &= \omega(v_{\pi\sigma(1)}, \ldots, v_{\pi\sigma(k)}) = \omega^{\pi\sigma}(v_1, \ldots, v_k).\end{aligned}$$

This implies that

$$\begin{aligned}\operatorname{Alt}(\omega)^\pi &= \frac{1}{k!}\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)(\omega^\sigma)^\pi = \frac{1}{k!}\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\omega^{\pi\sigma} = \frac{1}{k!}\sum_{\sigma\in S_k}\operatorname{sgn}(\pi^{-1}\sigma)\omega^\sigma\\ &= \frac{1}{k!}\sum_{\sigma\in S_k}\operatorname{sgn}(\pi)\operatorname{sgn}(\sigma)\omega^\sigma = \operatorname{sgn}(\pi)\operatorname{Alt}(\omega) = \frac{1}{k!}\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma\pi^{-1})\omega^\sigma\\ &= \frac{1}{k!}\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\omega^{\sigma\pi} = \frac{1}{k!}\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)(\omega^\pi)^\sigma = \operatorname{Alt}(\omega^\pi).\end{aligned}$$

In particular, we see that $\operatorname{Alt}(\omega)$ is alternating.

Remark C.14. (a) We observe that if ω is alternating, then $\omega^\sigma = \operatorname{sgn}(\sigma)\omega$ for each permutation σ, and therefore

$$\operatorname{Alt}(\omega) = \frac{1}{k!}\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma)\omega = \frac{1}{k!}\sum_{\sigma\in S_k}\omega = \omega.$$

(b) For $k=2$ we have $\operatorname{Alt}(\omega)(v_1,v_2) = \frac{1}{2}(\omega(v_1,v_2)-\omega(v_2,v_1))$, and for $k=3$:

$$\begin{aligned}\operatorname{Alt}(\omega)(v_1,v_2,v_3) = \frac{1}{6}\big(\omega(v_1,v_2,v_3) &- \omega(v_2,v_1,v_3) + \omega(v_2,v_3,v_1)\\ &- \omega(v_3,v_2,v_1) + \omega(v_3,v_1,v_2) - \omega(v_1,v_3,v_2)\big).\end{aligned}$$

Definition C.15. Let $p,q\in\mathbb{N}_0$. For two multilinear maps

$$\omega_1\colon V_1\times\ldots\times V_p\to\mathbb{K} \quad\text{and}\quad \omega_2\colon V_{p+1}\times\ldots\times V_{p+q}\to\mathbb{K}$$

we define the *tensor product* $\omega_1\otimes\omega_2\colon V_1\times\cdots\times V_{p+q}\to\mathbb{K}$ by

$$(\omega_1\otimes\omega_2)(v_1,\ldots,v_{p+q}) := \omega_1(v_1,\ldots,v_p)\omega_2(v_{p+1},\ldots,v_{p+q}).$$

It is clear that $\omega_1\otimes\omega_2$ is a $(p+q)$-linear map.

For $\lambda\in\mathbb{K}$ (the set of 0-linear maps), and a p-linear map ω as above, we obtain in particular

$$\lambda\otimes\omega := \omega\otimes\lambda := \lambda\omega.$$

For two alternating maps $\alpha\in\operatorname{Alt}^p(V,\mathbb{K})$ and $\beta\in\operatorname{Alt}^q(V,\mathbb{K})$ we define their *exterior product*:

$$\alpha\wedge\beta := \frac{(p+q)!}{p!q!}\operatorname{Alt}(\alpha\otimes\beta) = \frac{1}{p!q!}\sum_{\sigma\in S_{p+q}}\operatorname{sgn}(\sigma)(\alpha\otimes\beta)^\sigma. \tag{38}$$

It follows from (38) that $\alpha\wedge\beta$ is alternating, so that we obtain a bilinear map

$$\wedge\colon \operatorname{Alt}^p(V,\mathbb{K})\times\operatorname{Alt}^q(V,\mathbb{K})\to\operatorname{Alt}^{p+q}(V,\mathbb{K}),\quad (\alpha,\beta)\mapsto\alpha\wedge\beta.$$

On the direct sum

$$\operatorname{Alt}(V,\mathbb{K}) := \bigoplus_{p\in\mathbb{N}_0}\operatorname{Alt}^p(V,\mathbb{K})$$

we now obtain a bilinear product by putting

$$\Big(\sum_p\alpha_p\Big)\wedge\Big(\sum_q\beta_q\Big) := \sum_{p,q}\alpha_p\wedge\beta_q.$$

As before, we identify $\operatorname{Alt}^0(V,\mathbb{K})$ with $\mathbb{K}$ and obtain

$$\lambda\alpha = \lambda\wedge\alpha = \alpha\wedge\lambda$$

for $\lambda\in\operatorname{Alt}^0(V,\mathbb{K}) = \mathbb{K}$ and $\alpha\in\operatorname{Alt}^p(V,\mathbb{K})$.

We take a closer look at the structure of the algebra $(\mathrm{Alt}(V,\mathbb{K}), \wedge)$.

Lemma C.16. *For $\alpha \in \mathrm{Alt}^p(V,\mathbb{K})$, $\beta \in \mathrm{Alt}^q(V,\mathbb{K})$ and $\gamma \in \mathrm{Alt}^r(V,\mathbb{K})$, we have*

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma).$$

In particular the algebra $(\mathrm{Alt}(V,\mathbb{K}), \wedge)$ is associative.

Proof. First we recall from Definition C.13 that for any n-linear map $\omega\colon V^n \to W$ and $\pi \in S_n$ we have

$$\mathrm{Alt}(\omega^\pi) = \mathrm{sgn}(\pi)\,\mathrm{Alt}(\omega). \tag{39}$$

We identify S_{p+q} in the natural way with the subgroup of S_{p+q+r} fixing the numbers $p+q+1,\ldots,p+q+r$. We thus obtain

$$\begin{aligned}
(\alpha \wedge \beta) \wedge \gamma &= \frac{(p+q+r)!}{(p+q)!r!}\,\mathrm{Alt}((\alpha\wedge\beta)\otimes\gamma)\\
&= \frac{(p+q+r)!}{p!q!(p+q)!r!}\sum_{\sigma\in S_{p+q}} \mathrm{sgn}(\sigma)\,\mathrm{Alt}((\alpha\otimes\beta)^\sigma\otimes\gamma)\\
&= \frac{(p+q+r)!}{p!q!(p+q)!r!}\sum_{\sigma\in S_{p+q}} \mathrm{sgn}(\sigma)\,\mathrm{Alt}((\alpha\otimes\beta\otimes\gamma)^\sigma)\\
&\overset{(39)}{=} \frac{(p+q+r)!}{p!q!(p+q)!r!}\sum_{\sigma\in S_{p+q}} \mathrm{Alt}(\alpha\otimes\beta\otimes\gamma)\\
&= \frac{(p+q+r)!}{p!q!r!}\,\mathrm{Alt}(\alpha\otimes\beta\otimes\gamma) = \frac{(p+q+r)!}{p!q!r!}\,\mathrm{Alt}(\alpha\otimes(\beta\otimes\gamma))\\
&= \ldots = \frac{(p+q+r)!}{p!(q+r)!}\,\mathrm{Alt}(\alpha\otimes(\beta\wedge\gamma)) = \alpha\wedge(\beta\wedge\gamma).
\end{aligned}$$ ☐

From the associativity asserted in the preceding lemma, it follows that the multiplication in $\mathrm{Alt}(V,\mathbb{K})$ is associative. We may therefore suppress brackets and define

$$\omega_1 \wedge \ldots \wedge \omega_n := (\ldots((\omega_1\wedge\omega_2)\wedge\omega_3)\cdots\wedge\omega_n).$$

Remark C.17. (a) From the calculation in the preceding proof we know that for three elements $\alpha_i \in \mathrm{Alt}^{p_i}(V,\mathbb{K})$, the triple product in the associative algebra $\mathrm{Alt}(V,\mathbb{K})$ satisfies

$$\alpha_1\wedge\alpha_2\wedge\alpha_3 = \frac{(p_1+p_2+p_3)!}{p_1!p_2!p_3!}\,\mathrm{Alt}(\alpha_1\otimes\alpha_2\otimes\alpha_3).$$

Inductively this leads for n elements $\alpha_i \in \mathrm{Alt}^{p_i}(V,\mathbb{K})$ to

$$\alpha_1\wedge\ldots\wedge\alpha_n = \frac{(p_1+\ldots+p_n)!}{p_1!\cdots p_n!}\,\mathrm{Alt}(\alpha_1\otimes\cdots\otimes\alpha_n)$$

(Exercise C.2).

(b) For $\alpha_i \in \mathrm{Alt}^1(V,\mathbb{K}) \cong V^*$, we in particular obtain

$$\begin{aligned}
(\alpha_1\wedge\ldots\wedge\alpha_n)(v_1,\ldots,v_n) &= n!\,\mathrm{Alt}(\alpha_1\otimes\cdots\otimes\alpha_n)(v_1,\ldots,v_n)\\
&= \sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)\alpha_1(v_{\sigma(1)})\cdots\alpha_n(v_{\sigma(n)}) = \det(\alpha_i(v_j)).
\end{aligned}$$

Proposition C.18. *The exterior algebra is graded commutative, i.e., for $\alpha \in \mathrm{Alt}^p(V,\mathbb{K})$ and $\beta \in \mathrm{Alt}^q(V,\mathbb{K})$ we have*

$$\alpha\wedge\beta = (-1)^{pq}\beta\wedge\alpha.$$

Proof. Let $\sigma \in S_{p+q}$ denote the permutation defined by

$$\sigma(i) := \begin{cases} i+p & \text{for } 1 \leq i \leq q \\ i-q & \text{for } q+1 \leq i \leq p+q \end{cases}$$

which moves the first q elements to the last q positions. Then we have

$$\begin{aligned} &(\beta \otimes \alpha)^\sigma(v_1, \ldots, v_{p+q}) = (\beta \otimes \alpha)(v_{\sigma(1)}, \ldots, v_{\sigma(p+q)}) \\ &= \beta(v_{p+1}, \ldots, v_{p+q})\alpha(v_1, \ldots, v_p) = (\alpha \otimes \beta)(v_1, \ldots, v_{p+q}). \end{aligned}$$

This leads to

$$\begin{aligned} \alpha \wedge \beta &= \frac{(p+q)!}{p!q!} \operatorname{Alt}(\alpha \otimes \beta) = \frac{(p+q)!}{p!q!} \operatorname{Alt}((\beta \otimes \alpha)^\sigma) \\ &= \operatorname{sgn}(\sigma) \frac{(p+q)!}{p!q!} \operatorname{Alt}(\beta \otimes \alpha) = \operatorname{sgn}(\sigma)(\beta \wedge \alpha). \end{aligned}$$

On the other hand $\operatorname{sgn}(\sigma) = (-1)^F$, where

$$\begin{aligned} F &:= |\{(i,j) \in \{1, \ldots, p+q\} \colon i < j, \sigma(j) < \sigma(i)\}| \\ &= |\{(i,j) \in \{1, \ldots, p+q\} \colon i \leq q, j > q\}| = pq \end{aligned}$$

is the number of inversions of σ. Putting everything together, the lemma follows. □

Corollary C.19. *If $\alpha \in \operatorname{Alt}^p(V, \mathbb{K})$ and p is odd, then $\alpha \wedge \alpha = 0$.*

Proof. In view of Proposition C.18, we have $\alpha \wedge \alpha = (-1)^{p^2} \alpha \wedge \alpha = -\alpha \wedge \alpha$, which leads to $\alpha \wedge \alpha = 0$. □

Corollary C.20. *If $\alpha_1, \ldots, \alpha_k \in V^* = \operatorname{Alt}^1(V, \mathbb{K})$ and $\beta_j = \sum_{i=1}^k a_{ij}\alpha_i$, then*

$$\beta_1 \wedge \ldots \wedge \beta_k = \det(A) \cdot \alpha_1 \wedge \ldots \wedge \alpha_k \quad \textit{for} \quad A = (a_{ij}) \in M_k(\mathbb{K}).$$

Proof. The k-fold multiplication map

$$\Phi \colon (V^*)^k \to \operatorname{Alt}^k(V, \mathbb{K}), \quad (\gamma_1, \ldots, \gamma_k) \mapsto \gamma_1 \wedge \ldots \wedge \gamma_k$$

is alternating by Proposition C.18 because S_k is generated by transpositions. Hence the assertion follows from Proposition C.11. □

Corollary C.21. *If $\dim V = n$, $b_1, \ldots, b_n$ is a basis for V, and $b_1^*, \ldots, b_n^*$ the dual basis for V^*, then the products*

$$b_I^* := b_{i_1}^* \wedge \ldots \wedge b_{i_k}^*, \quad I = (i_1, \ldots, i_k), \quad 1 \leq i_1 < \ldots < i_k \leq n,$$

form a basis for $\operatorname{Alt}^k(V, \mathbb{K})$.

Proof. For $J = (j_1, \ldots, j_k)$ with $j_1 < \ldots < j_k$, we get with Remark C.17(b)

$$b_I^*(b_{j_1}, \ldots, b_{j_k}) = \det(b_{i_l}^*(b_{j_m})_{l,m=1,\ldots,k}) = \begin{cases} 1 & \text{for } I = J \\ 0 & \text{for } I \neq J. \end{cases}$$

If follows in particular that the elements b_I are linearly independent, and since $\dim \operatorname{Alt}^k(V, \mathbb{K}) = \binom{n}{k}$ (Corollary C.12), the assertion follows. □

Remark C.22. (a) From Corollary C.12 it follows in particular that

$$\dim \mathrm{Alt}(V, \mathbb{K}) = \sum_{k=0}^{\dim V} \binom{\dim V}{k} = 2^{\dim V}$$

if V is finite dimensional.

(b) If V is infinite dimensional, then it has an infinite basis $(b_i)_{i\in I}$ (this requires Zorn's Lemma). In addition, the set I carries a linear order $\leq$ (this requires the Well Ordering Theorem), and for each k-element subset $J = \{j_1, \ldots, j_k\} \subseteq I$ with $j_1 < \ldots < j_k$, we thus obtain an element

$$b_J^* := b_{j_1}^* \wedge \ldots \wedge b_{j_k}^*.$$

Applying the b_J^* to k-tuples of basis elements shows that they are linearly independent, so that for each $k > 0$ the space $\mathrm{Alt}^k(V, \mathbb{K})$ is infinite dimensional.

Definition C.23. Let $\varphi\colon V_1 \to V_2$ be a linear map and W a vector space. For each p-linear map $\alpha\colon V_2^p \to W$ we define its *pull-back by* φ:

$$(\varphi^*\alpha)(v_1, \ldots, v_p) := \alpha(\varphi(v_1), \ldots, \varphi(v_p))$$

for $v_1, \ldots, v_p \in V_1$. It is clear that $\varphi^*\alpha$ is a p-linear map $V_1^p \to W$ and that $\varphi^*\alpha$ is alternating if α has this property.

Remark C.24. If $\varphi\colon V_1 \to V_2$ and $\psi\colon V_2 \to V_3$ are linear maps and $\alpha\colon V_3^p \to W$ is p-linear, then

$$(\psi \circ \varphi)^*\alpha = \varphi^*(\psi^*\alpha).$$

Proposition C.25. *Let $\varphi\colon V_1 \to V_2$ be a linear map. Then the pull-back map*

$$\varphi^*\colon\ \mathrm{Alt}(V_2, \mathbb{K}) \to \mathrm{Alt}(V_1, \mathbb{K})$$

is a homomorphism of algebras with unit.

Proof. For $\alpha \in \mathrm{Alt}^p(V_2, \mathbb{K})$ and $\beta \in \mathrm{Alt}^q(V_2, \mathbb{K})$ we have

$$\begin{aligned}\varphi^*(\alpha \wedge \beta) &= \frac{(p+q)!}{p!q!}\varphi^*(\mathrm{Alt}(\alpha \otimes \beta)) = \frac{(p+q)!}{p!q!}\mathrm{Alt}(\varphi^*(\alpha \otimes \beta)) \\ &= \frac{(p+q)!}{p!q!}\mathrm{Alt}(\varphi^*\alpha \otimes \varphi^*\beta) = \varphi^*\alpha \wedge \varphi^*\beta.\end{aligned}$$ □

Remark C.26. The results in this section remain valid for alternating forms with values in any commutative algebra A. Then

$$\mathrm{Alt}(V, A) = \bigoplus_{p\in\mathbb{N}_0} \mathrm{Alt}^p(V, A)$$

also carries an associative, graded commutative algebra structure defined by

$$\alpha \wedge \beta := \frac{(p+q)!}{p!q!}\mathrm{Alt}(\alpha \otimes \beta),$$

where

$$(\alpha \otimes \beta)(v_1, \ldots, v_{p+q}) := \alpha(v_1, \ldots, v_p) \cdot \beta(v_{p+1}, \ldots, v_{p+q})$$

for $\alpha \in \mathrm{Alt}^p(V, A)$, $\beta \in \mathrm{Alt}^q(V, A)$.

This applies in particular to the 2-dimensional real algebra $A = \mathbb{C}$.

Exercises for Section C

Exercise C.1. Fix $n \in \mathbb{N}$. Show that:

(1) For each matrix $A \in M_n(\mathbb{K})$, we obtain a bilinear map

$$\beta_A \colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}, \quad \beta_A(x,y) := \sum_{i,j=1}^n a_{ij} x_i y_j.$$

(2) A can be recovered from β_A via $a_{ij} = \beta_A(e_i, e_j)$.

(3) Each bilinear map $\beta \colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ is of the form $\beta = \beta_A$ for a unique matrix $A \in M_n(\mathbb{R})$.

(4) $\beta_{A^\top}(x,y) = \beta_A(y,x)$.

(5) β_A is skew-symmetric if and only if A is so.

Exercise C.2. Show that for $\alpha_i \in \mathrm{Alt}^{p_i}(V, \mathbb{K}), i = 1, \ldots, n$, the exterior product satisfies

$$\alpha_1 \wedge \ldots \wedge \alpha_n = \frac{(p_1 + \ldots + p_n)!}{p_1! \cdots p_n!} \mathrm{Alt}(\alpha_1 \otimes \cdots \otimes \alpha_n)$$

Exercise C.3. Show that $(\mathrm{Alt}(V, \mathbb{K}), \wedge)$ is an exterior algebra over V^*.

D Supplementary material

D.1 The nilradical is characteristic

We can now record the following interesting consequence of Corollary 4.16.

Proposition D.1. *If $\mathfrak{g}$ is a finite dimensional Lie algebra and $D \in \mathrm{der}(\mathfrak{g})$, then $D(\mathrm{rad}(\mathfrak{g})) \subseteq \mathrm{nil}(\mathfrak{g})$. In particular, $\mathrm{nil}(\mathfrak{g})$ and $\mathrm{rad}(\mathfrak{g})$ are invariant under $\mathrm{der}(\mathfrak{g})$, i.e., characteristic ideals.*

Proof. For $D \in \mathrm{der}(\mathfrak{g})$, consider the Lie algebra $\widetilde{\mathfrak{g}} := \mathfrak{g} \rtimes_D \mathbb{K}$ for the Lie algebra with the bracket

$$[(x,t),(x',t')] := (tDx' - t'Dx + [x,x'], 0)$$

and identify $\mathfrak{g}$ with the subalgebra $\mathfrak{g} \times \{0\}$. Then $\mathrm{rad}(\mathfrak{g})$ is a solvable ideal of $\widetilde{\mathfrak{g}}$ because $\mathrm{rad}(\mathfrak{g})$ is a characteristic ideal of $\mathfrak{g}$ (Lemma 6.4), hence contained in $\mathrm{rad}(\widetilde{\mathfrak{g}})$. We thus obtain

$$D(\mathrm{rad}(\mathfrak{g})) = [(0,1), \mathrm{rad}(\mathfrak{g})] \subseteq \mathfrak{g} \cap [\widetilde{\mathfrak{g}}, \mathrm{rad}(\widetilde{\mathfrak{g}})] \subseteq \mathfrak{g} \cap \mathrm{nil}(\widetilde{\mathfrak{g}}) \subseteq \mathrm{nil}(\mathfrak{g}). \qquad \square$$

Corollary D.2. *If $\mathfrak{a} \trianglelefteq \mathfrak{g}$ is an ideal, then $\mathrm{nil}(\mathfrak{a}) = \mathfrak{a} \cap \mathrm{nil}(\mathfrak{g})$.*

Proof. Clearly, $\mathfrak{a} \cap \mathrm{nil}(\mathfrak{g})$ is a nilpotent ideal of $\mathfrak{a}$, hence contained in $\mathrm{nil}(\mathfrak{a})$. Conversely, $\mathrm{nil}(\mathfrak{a})$ is an ideal of $\mathfrak{g}$ because it is invariant under all the derivations $\mathrm{ad}\, x|_{\mathfrak{a}}$, $x \in \mathfrak{g}$. This implies that $\mathrm{nil}(\mathfrak{a}) \subseteq \mathrm{nil}(\mathfrak{g}) \cap \mathfrak{a}$. $\square$

D.2 Malcev's Theorem

Lemma D.3. *If $x \in \mathfrak{g}$ is such that* $\operatorname{ad} x$ *is nilpotent, then*

$$e^{\operatorname{ad} x} y = \sum_{k=0}^{\infty} \frac{1}{k!} (\operatorname{ad} x)^k y$$

defines an automorphism of $\mathfrak{g}$.

Note that this series is actually finite and that it makes sense over every field of characteristic zero.

Proof. This follows from the proof of (ii) $\Rightarrow$ (i) in Lemma 2.4, applied with $V = W = \mathfrak{g}$ and $\beta(x, y) = [x, y]$. $\square$

Theorem D.4. (Malcev's Theorem) *For two Levi complements $\mathfrak{s}$ and $\mathfrak{s}'$ in $\mathfrak{g}$, there exists some $x \in [\mathfrak{g}, \operatorname{rad}(\mathfrak{g})]$ with $e^{\operatorname{ad} x}\mathfrak{s}' = \mathfrak{s}$.*

Proof. Let $\mathfrak{r} := \operatorname{rad}(\mathfrak{g})$. We first consider some special cases.

(a) If $[\mathfrak{g}, \mathfrak{r}] = \{0\}$, then $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{s}$ is a direct sum of Lie algebras and $\mathfrak{r} = \mathfrak{z}(\mathfrak{g})$ is abelian. Therefore $\mathfrak{s} = [\mathfrak{s}, \mathfrak{s}] = [\mathfrak{g}, \mathfrak{g}] = [\mathfrak{s}', \mathfrak{s}'] = \mathfrak{s}'$, and there is nothing to show.

(b) If $[\mathfrak{g}, \mathfrak{r}] \neq \{0\}$ and $\mathfrak{r}$ is a minimal nonzero ideal of $\mathfrak{g}$, then $[\mathfrak{g}, \mathfrak{r}] = \mathfrak{r}$, $[\mathfrak{r}, \mathfrak{r}] = \{0\}$ (since $D^1(\mathfrak{r}) \neq \mathfrak{r}$), and $\mathfrak{z}(\mathfrak{g}) = \{0\}$ (because $\mathfrak{r} \not\subseteq \mathfrak{z}(\mathfrak{g})$). We define a map $h\colon \mathfrak{s}' \to \mathfrak{r}$ by $x + h(x) \in \mathfrak{s}$ for $x \in \mathfrak{s}'$, i.e., $-h$ is the projection of $\mathfrak{s}'$ to $\mathfrak{s}$ along $\mathfrak{r}$. Since $\mathfrak{s}$ is a subalgebra and $\mathfrak{r}$ is abelian, we have

$$[x + h(x), y + h(y)] = [x, y] + [x, h(y)] + [h(x), y] \in \mathfrak{s}.$$

Therefore

$$h([x, y]) = [x, h(y)] + [h(x), y].$$

This implies that

$$\pi(x)(r, t) := ([x, r] + th(x), 0)$$

defines a representation of $\mathfrak{s}'$ on $\mathfrak{r} \times \mathbb{K}$. The subspace $\mathfrak{r} \cong \mathfrak{r} \times \{0\}$ is $\mathfrak{s}'$-invariant. According to Weyl's Theorem, there exists an $\mathfrak{s}'$-invariant complement $\mathbb{K}(v, 1)$ of $\mathfrak{r}$ in $\mathfrak{r} \oplus \mathbb{K}$. As $\mathfrak{s}'$ is semisimple, $\pi(\mathfrak{s}')(v, 1) = \{0\}$, and hence $h(x) + [x, v] = 0$ for $x \in \mathfrak{s}'$. Now we have

$$e^{\operatorname{ad} v} x = x + [v, x] = x + h(x) \in \mathfrak{s} \quad \text{ for } \quad x \in \mathfrak{s}'$$

and thus $e^{\operatorname{ad} v}(\mathfrak{s}') \subseteq \mathfrak{s}$. Equality follows from $\dim \mathfrak{s} = \dim \mathfrak{g}/\mathfrak{r} = \dim \mathfrak{s}'$. This proves the theorem if $[\mathfrak{g}, \mathfrak{r}]$ is nonzero and a minimal ideal.

(c) Finally, we turn to the general case. We argue by induction on $n := \dim \mathfrak{r}$. The case $n = 0$ is trivial, so that we assume $n > 0$ and that the assertion holds for all Lie algebras $\mathfrak{h}$ with $\dim \operatorname{rad}(\mathfrak{h}) < n$. In view of (a), we may assume that $[\mathfrak{g}, \mathfrak{r}] \neq \{0\}$. As the ideal $[\mathfrak{g}, \mathfrak{r}]$ is nilpotent (Corollary 4.16), its center $\mathfrak{c} := \mathfrak{z}([\mathfrak{g}, \mathfrak{r}])$ is nonzero (Proposition 3.3). Let $\mathfrak{m} \neq \{0\}$ be a minimal ideal of $\mathfrak{g}$ contained in $\mathfrak{c}$. If $\mathfrak{m} = \mathfrak{r}$, then we are in the situation of (b). We therefore assume $\mathfrak{m} \neq \mathfrak{r}$. Let $\pi\colon \mathfrak{g} \to \mathfrak{g}_1 := \mathfrak{g}/\mathfrak{m}$ be the quotient map. Then $\mathfrak{r}_1 := \pi(\mathfrak{r})$ is the radical of $\mathfrak{g}_1$ (Proposition 6.2), and $\pi(\mathfrak{s})$ and $\pi(\mathfrak{s}')$ are Levi complements in $\mathfrak{g}/\mathfrak{m}$ because both are semisimple (Proposition 5.10) and complementing $\pi(\mathfrak{r})$. Now our induction hypothesis provides an $x_1 \in [\mathfrak{g}_1, \mathfrak{r}_1]$ with $e^{\operatorname{ad} x_1}\pi(\mathfrak{s}') = \pi(\mathfrak{s})$. Using $\pi([\mathfrak{g}, \mathfrak{r}]) = [\mathfrak{g}_1, \mathfrak{r}_1]$, we find an $x \in [\mathfrak{g}, \mathfrak{r}]$ with $\pi(x) = x_1$. Then $e^{\operatorname{ad} x_1}\pi(\mathfrak{s}') = \pi(e^{\operatorname{ad} x}\mathfrak{s}') \subseteq \pi(\mathfrak{s})$, i.e.,

$$e^{\operatorname{ad} x}\mathfrak{s}' \subseteq \mathfrak{h} := \mathfrak{s} + \mathfrak{m}.$$

Now $e^{\operatorname{ad} x}\mathfrak{s}'$ and $\mathfrak{s}$ are two Levi complements in the Lie algebra $\mathfrak{h}$ with

$$\dim \operatorname{rad}(\mathfrak{h}) = \dim \mathfrak{m} < n = \dim \mathfrak{r}.$$

Hence the induction hypothesis provides a $y \in \mathfrak{m}$ with $e^{\operatorname{ad} y}e^{\operatorname{ad} x}\mathfrak{s}' \subseteq \mathfrak{s}$. Since $\mathfrak{m}$ is central in $[\mathfrak{g}, \mathfrak{r}]$, we have $[x, y] = 0$ and therefore $e^{\operatorname{ad} y}e^{\operatorname{ad} x}\mathfrak{s}' = e^{\operatorname{ad}(x+y)}\mathfrak{s}' \subseteq \mathfrak{s}$. □

Malcev's Theorem has interesting consequences:

Corollary D.5. *Each semisimple subalgebra of $\mathfrak{g}$ is contained in a Levi complement. In particular, the Levi complements are precisely the maximal semisimple subalgebras of $\mathfrak{g}$.*

Proof. Let $\mathfrak{r} := \operatorname{rad}(\mathfrak{g})$ be the radical of $\mathfrak{g}$, $\mathfrak{h} \subseteq \mathfrak{g}$ a semisimple subalgebra, and $\mathfrak{a} := \mathfrak{r} + \mathfrak{h}$. Then $\mathfrak{a}$ is a subalgebra of $\mathfrak{g}$ and $\mathfrak{r}$ is a solvable ideal of $\mathfrak{a}$. Since the solvable ideal $\operatorname{rad}(\mathfrak{a}) \cap \mathfrak{h}$ of the semisimple Lie algebra $\mathfrak{h}$ is trivial, we see that $\mathfrak{r} = \operatorname{rad}(\mathfrak{a})$. The ideal $\mathfrak{h} \cap \mathfrak{r}$ of $\mathfrak{h}$ is solvable and semisimple, hence trivial. This proves that $\mathfrak{h}$ is a Levi complement in $\mathfrak{a}$.

Let $\mathfrak{s}$ be a Levi complement in $\mathfrak{g}$. Then $\mathfrak{a} = \mathfrak{r} + (\mathfrak{a} \cap \mathfrak{s})$ is a semidirect sum and since $\mathfrak{a} \cap \mathfrak{s} \cong \mathfrak{a}/\mathfrak{r} \cong \mathfrak{h}$ is semisimple, $\mathfrak{a} \cap \mathfrak{s}$ is a Levi complement in $\mathfrak{a}$. According to Malcev's Theorem D.4, there exists an $x \in [\mathfrak{a}, \mathfrak{r}]$ with $e^{\operatorname{ad} x}(\mathfrak{a} \cap \mathfrak{s}) = \mathfrak{h}$, i.e., $\mathfrak{h}$ is a contained in the Levi complement $e^{\operatorname{ad} x}(\mathfrak{s})$ of $\mathfrak{g}$. □

Corollary D.6. *If $\mathfrak{n} \trianglelefteq \mathfrak{g}$ is an ideal of $\mathfrak{g}$ and $\mathfrak{g} = \mathfrak{r} \rtimes \mathfrak{s}$ a Levi decomposition, i.e., $\mathfrak{s}$ is a Levi complement, then $\mathfrak{n} = (\mathfrak{n} \cap \mathfrak{r}) \rtimes (\mathfrak{n} \cap \mathfrak{s})$ is a Levi decomposition of $\mathfrak{n}$.*

Proof. We have already seen in Lemma 6.4 that $\mathfrak{n} \cap \mathfrak{r} = \operatorname{rad}(\mathfrak{n})$. If $\mathfrak{s}_\mathfrak{n}$ is a Levi complement in $\mathfrak{n}$, then Corollary D.5 implies that the semisimple Lie algebra $\mathfrak{s}_\mathfrak{n}$ is contained in a Levi complement $\mathfrak{s}'$ of $\mathfrak{g}$. For $x \in [\mathfrak{g}, \mathfrak{r}]$ with $e^{\operatorname{ad} x}\mathfrak{s}' = \mathfrak{s}$ we now see that $e^{\operatorname{ad} x}\mathfrak{s}_\mathfrak{n} \subseteq \mathfrak{n} \cap \mathfrak{s}$, because

$$e^{\operatorname{ad} x}\mathfrak{n} \subseteq \mathfrak{n} + [x, \mathfrak{n}] \subseteq \mathfrak{n}.$$

Since the ideal $\mathfrak{n} \cap \mathfrak{s}$ of $\mathfrak{s}$ is semisimple (Proposition 5.10) and $\mathfrak{s}_\mathfrak{n}$ is maximal semisimple in $\mathfrak{n}$, we obtain $e^{\operatorname{ad} x}\mathfrak{s}_\mathfrak{n} = \mathfrak{n} \cap \mathfrak{s}$. This shows that $\mathfrak{n} \cap \mathfrak{s}$ is a Levi complement in $\mathfrak{n}$. □

D.3 Reflections of $\mathfrak{sl}_2$-modules

For $\mathbb{K} = \mathbb{R}, \mathbb{C}$, we have for every finite dimensional vector space V the exponential function

$$\exp\colon \mathfrak{gl}(V) \to \operatorname{GL}(V), \quad \exp(X) := \sum_{n=0}^{\infty} \frac{1}{n!} X^n$$

defined by the convergent exponential series.

Definition D.7. (Exponential function over fields of characteristic zero) If $\mathbb{K}$ is a genreal field of characteristic zero and $X \in \mathfrak{gl}(V)$ a nilpotent element, then the exponential series

$$\exp(X) := \sum_{n=0}^{\infty} \frac{1}{n!} X^n$$

still makes sense because only finitely many summands are non-zero.

Lemma D.8. *Let V be a finite dimensional $\mathbb{K}$-vector space and $x, y \in \mathfrak{gl}(V)$, where either $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ or $\operatorname{char} \mathbb{K} = 0$ and x, y are nilpotent.*

(i) *If $xy = yx$, then $\exp(x + y) = \exp x \exp y$.*

(ii) $\exp(x) \in \mathrm{GL}(V)$, $\exp(0) = \mathbf{1}$, *and* $(\exp x)^{-1} = \exp(-x)$.

(iii) *For* $g \in \mathrm{GL}(V)$, *we have the relation*

$$g \exp(x) g^{-1} = \exp(gxg^{-1}).$$

Proof. (i) Using the general form of the Cauchy Product Formula in the case $\mathbb{K} = \mathbb{R}, \mathbb{C}$ (Exercise 2.2), we obtain

$$\begin{aligned}\exp(x+y) &= \sum_{k=0}^{\infty} \frac{(x+y)^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\ell=0}^{k} \binom{k}{\ell} x^\ell y^{k-\ell} \\ &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{k} \frac{x^\ell}{\ell!} \frac{y^{k-\ell}}{(k-\ell)!} = \Big(\sum_{p=0}^{\infty} \frac{x^p}{p!}\Big)\Big(\sum_{\ell=0}^{\infty} \frac{y^\ell}{\ell!}\Big).\end{aligned}$$

If both x and y are nilpotent, all these series are finite, so that the assertion follows likewise.

(ii) From (i) we derive in particular $\exp x \exp(-x) = \exp 0 = \mathbf{1}$, which implies (ii).

(iii) is a consequence of $gx^n g^{-1} = (gxg^{-1})^n$ and the continuity of the conjugation map $c_g(x) := gxg^{-1}$ on $M_n(\mathbb{K})$ for $\mathbb{K} = \mathbb{R}, \mathbb{C}$. $\square$

Lemma D.9. *For* $x, y \in \mathfrak{gl}(V)$ *and* $\mathbb{K} = \mathbb{R}, \mathbb{C}$ *or for* x *nilpotent and* $\mathrm{char}\,\mathbb{K} = 0$, *we have*

$$(\exp x) y (\exp x)^{-1} = \exp(\mathrm{ad}\, x) y. \tag{40}$$

Proof. We define the linear maps

$$\lambda_x \colon \mathrm{End}(V) \to \mathrm{End}(V), \quad y \mapsto xy, \quad \rho_x \colon \mathrm{End}(V) \to \mathrm{End}(V), \quad y \mapsto yx.$$

Then $\lambda_x \rho_x = \rho_x \lambda_x$ and $\mathrm{ad}\, x = \lambda_x - \rho_x$, so that Lemma D.8(ii) leads to

$$(\exp x) y (\exp x)^{-1} = e^x y e^{-x} = e^{\lambda_x} e^{-\rho_x} y = e^{\lambda_x - \rho_x} y = e^{\mathrm{ad}\, x} y.$$

This proves (40).

If x is nilpotent and $\mathbb{K}$ is a general field of characteristic zero, then we recall from Proposition 4.19 that $\mathrm{ad}\, x$ is nilpotent, so that both sides make sense. Their equality follows as above. $\square$

We consider the element

$$\theta := e^{\mathrm{ad}\, e} e^{-\mathrm{ad}\, f} e^{\mathrm{ad}\, e} \in \mathrm{Aut}(\mathfrak{sl}_2(\mathbb{K}))$$

(Example 2.6) and

$$\sigma := \exp(e) \exp(-f) \exp(e) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{K}).$$

Then Lemma D.9 implies for $z \in \mathfrak{sl}_2(\mathbb{K})$ the relation $\theta(z) = \sigma z \sigma^{-1}$, hence in particular

$$\theta(h) = -h, \quad \theta(e) = -f \quad \text{and} \quad \theta(f) = -e.$$

Lemma D.10. *Let* (ρ, V) *be a finite dimensional representation of* $\mathfrak{sl}_2(\mathbb{K})$ *and* $\sigma_V := e^{\rho(e)} e^{-\rho(f)} e^{\rho(e)} \in \mathrm{GL}(V)$. *Then*

$$\begin{aligned} \sigma_V \rho(z) \sigma_V^{-1} &= \rho(\sigma z \sigma^{-1}) \quad \textit{for} \quad z \in \mathfrak{sl}_2(\mathbb{K}), \\ \sigma_V(V_\alpha(\rho(h))) &= V_{-\alpha}(\rho(h)) \quad \textit{for the eigenspaces of} \quad \rho(h). \end{aligned} \tag{41}$$

Proof. For $z \in \mathfrak{sl}_2(\mathbb{K})$, we obtain with Lemma D.9 the relation

$$\sigma_V \rho(z) \sigma_V^{-1} = \rho(\sigma z \sigma^{-1}).$$

For $v \in V_\alpha(\rho(h))$ we have

$$\rho(h)(\sigma_V(v)) = \sigma_V\big(\sigma_V^{-1} \rho(h) \sigma_V\big)(v) = \sigma_V \rho(-h)(v) = -\alpha \sigma_V(v).$$

This implies that (41). $\square$

References

[BH09] Baez, J., and J. Huerta, *The Algebra of Grand Unified Theories*, arXiv:hep-th.0904.1556

[Bou89] Bourbaki, N., “Lie Groups and Lie Algebras, Chs. 1-3”, Springer, 1989

[Br36] Brauer, R., *Eine Bedingung für vollständige Reduzibilität von Darstellungen gewöhnlicher und infinitesimaler Gruppen*, Math. Zeitschr. **61** (1936), 330–339

[Ca94] Cartan, É., “Sur la structure des groupes de transformations finis et continue”, thèse, Paris, Nony, 1894

[CW35] Casimir, H., and B. L. van der Waerden, *Algebraischer Beweis der vollständigen Reduzibilität der Darstellungen halbeinfacher Liescher Gruppen*, Math. Ann. **111** (1935), 1–12

[CE48] Chevalley, C. and S. Eilenberg, *Cohomology Theory of Lie Groups and Lie Algebras,* Transactions of the Amer. Math. Soc. **63** (1948) 85–124

[GK96] Goze, M., and Y. Khakimdjanov, “Nilpotent Lie Algebras,” Mathematics and Its Applications, Kluwer Acad. Pub., 1996

[GM85] Greiner, W., and B. Müller, “Theoretische Physik, Bd. 5, Quantenmechanik II, Symmetrien”, Verlag Henri Deutsch, 1985

[Gr78] Greub, W., “Multilinear Algebra,” Second ed., Universitext, Springer-Verlag, New York-Heidelberg, 1978

[Haw00] Hawkins, Th., “Emergence of the Theory of Lie Groups,” Springer, 2000

[HN91] Hilgert, J., and K.-H. Neeb, “Lie-Gruppen und Lie-Algebren”, Vieweg, Braunschweig, 1991

[HiNe12] Hilgert, J., and K.-H. Neeb, “Structure and Geometry of Lie Groups”, Springer, 2012

[Ho54a] Hochschild, G., *Lie algebra kernels and cohomology*, Amer. J. Math. **76** (1954), 698–716

[Ho54b] —, *Cohomology classes of finite type and finite dimensional kernels for Lie algebras*, Amer. J. Math. **76** (1954), 763–778

[Kil89] Killing, W., *Die Zusammensetzung der stetigen endlichen Transformationsgruppen II*, Math. Ann. **33** (1889), 1–48

[Kos50] J.-L. Koszul, *Homologie et cohomologie des algèbres de Lie*, Bull. Soc. Math. France **78** (1950), 65–127

[La93] Lang, S., “Algebra,” Addison Wesley Publ. Comp., London, 1993

[Le05] Levi, E. E., *Sulla structura dei gruppi finiti e continui*, Atti Accad. Torino **40** (1905), 3–17

[LV94] D. Luminet and A. Valette, *Faithful uniformly continuous representations of Lie groups*, J. London Math. Soc. (2) **49** (1994), 100–108

[Ma42] Malcev, A. I., *On the representation of an algebra as a direct sum of the radical and a semi-simple subalgebra*, Dokl. Akad. Nauk SSSR, **36:2** (1942), 42–45

[Ner02] Neretin, Y. A., *A construction of finite dimensional faithful representation of Lie algebra*, in Proceedings of the 22nd Winter School "Geometry and Physics" (Srni 2002), Rend. Circ. Mat. Palermo (2) Suppl. **71** (2003), 159–161, arXiv:math/0202190 [math.RT]

[EN27] Noether, E., *Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionenkörpern*, Math. Annalen **96:1**(1927), 26–61

[St94] Sternberg, S., "Group Theory and Physics," Cambridge University Press, 1994 (A nice book discussing many application to concrete problems from physics and chemistry.)

[Wh36] Whitehead, J. H. C., *On the decomposition of an infinitesimal group*, Proc. Camb. Philos. Soc. **32** (1936), 229–237