Joachim Wehler

Lie Algebras

DRAFT, Release 2.0

April 1, 2023

I prepared these notes for the students of my lectures and my seminar on Lie algebras. The lecture has been repeated several times since the summer semester 2017 at the mathematical department of LMU (Ludwig-Maximilians-Universität) at Munich. I thank the participants of the seminar and the lectures, notably J. Bartenschlager and W. Hensgen, for making some simplifications and pointing out gaps and errors. The lecture notes include some ideas from former lecture notes of H. Abels at the university of Bielefeld.

Compared to the oral lecture in class these written notes contain some additional material.

Please report any errors or typos to

wehler@math.lmu.de

Release notes:
Release 2.0: All chapters revised.
Release 1.1: General revision Chapter 1.

Contents

Topics from Lie algebras and Lie groups

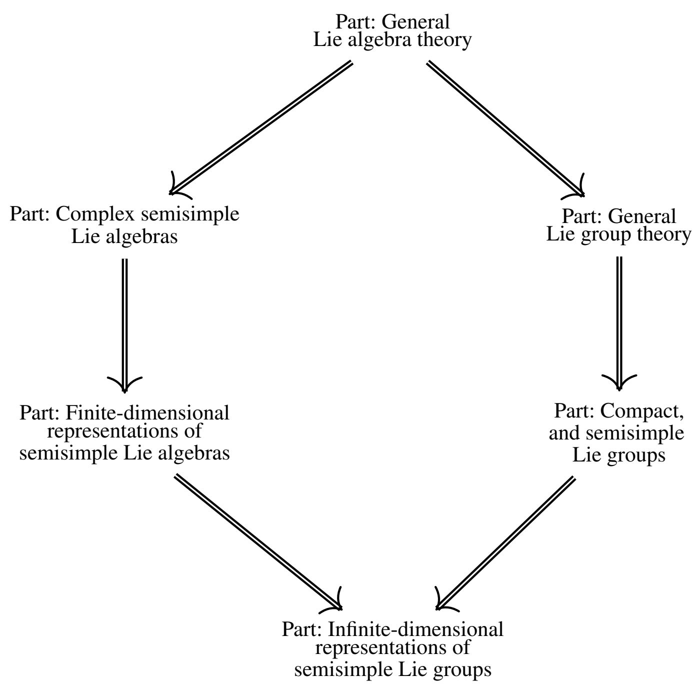

Fig. 0.1 Logical dependencies in the theory of Lie algebras and Lie groups

Part I
General Lie algebra theory

Chapter 1
Matrix functions

The paradigm of a Lie algebra is the vector space of matrices with the commutator of two matrices as Lie bracket. These concrete examples even cover all abstract finite dimensional Lie algebras. They are the focus of these notes. Nevertheless it is useful to consider Lie algebras from an abstract viewpoint as a separate algebraic structure like groups or rings.

If not stated otherwise, we denote by $\mathbb{K}$ the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers. Both fields have characteristic 0. The relevant difference is the fact that $\mathbb{C}$ is algebraically closed, i.e. each polynomial with coefficients from $\mathbb{K}$ of degree n has exactly n complex roots. This result allows to transform matrices over $\mathbb{C}$ to certain standard forms by transformations which make use of the eigenvalues of the matrix.

If not stated otherwise all vector spaces in this chapter are assumed finite-dimensional $\mathbb{K}$-vector spaces.

1.1 Power series of matrices

At high school every student learns the functional equation of the exponential function

$$exp(x) \cdot exp(y) = exp(x+y)$$

This formula holds for all real numbers $x,\ y \in \mathbb{R}$. Then exponentiation defines a group morphism

$$exp : (\mathbb{R}, +) \to (\mathbb{R}^*, \cdot)$$

Those students who attended a class on complex analysis will remember that exponentiation is defined also for complex numbers. Hence exponentiation extends to a map

$$exp : (\mathbb{C}, +) \to (\mathbb{C}^*, \cdot)$$

which also satisfies the *functional equation* above.

The seamless transition from the real field $\mathbb{R}$ to the complex field $\mathbb{C}$ is due to the fact that the exponential map is defined by a power series

$$exp(z) = \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \cdot z^{\nu}$$

The series converges not only for real numbers but for all complex numbers too.

In order to generalize the exponential map one step further we now exponentiate strictly upper triangular matrices.

Definition 1.1 (Exponentiation of strictly upper triangular matrices). For $n \in \mathbb{N}^*$ denote by

$$\mathfrak{n}(n,\mathbb{K}) := \left\{(a_{ij})_{1 \le i,j \le n} : a_{ij} \in \mathbb{K} \text{ and } a_{ij} = 0 \text{ if } j \le i\right\}$$

the $\mathbb{K}$-algebra of strictly upper triangular matrices with a typical element of the form

$$\begin{pmatrix} 0 & * & * \dots & * \\ 0 & 0 & * \dots & * \\ & \dots & & \\ 0 & 0 & 0 \dots & 0 \end{pmatrix}$$

For $A \in \mathfrak{n}(n,\mathbb{K})$ the *exponential* of A is defined as

$$exp\, A := \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \cdot A^{\nu}$$

Matrices from $\mathfrak{n}(n,\mathbb{K})$ can be added and multiplied with each other, and they can be multiplied by scalars from the field $\mathbb{K}$. Note that the series in Definition 1.1 reduces to a finite sum because $A^n = 0$. Hence there arises no question of convergence.

The goal of the present section is to extend the exponential map to all matrices from $M(n \times n,\mathbb{K})$. Now we need a concept of convergence for sequences and series of matrices. The basic ingredience is the *operator norm* of a matrix, a concept which applies to each linear map between normed vector spaces.

Definition 1.2 (Operator norm). We consider the $\mathbb{K}$-vector space $\mathbb{K}^n,\ n \in \mathbb{N}$, with the Euclidean norm

$$\|x\| := \sqrt{\sum_{i=1}^{n} |x_i|^2} \; for\; x = (x_1,...,x_n) \in \mathbb{K}^n.$$

For matrices $A \in M(n \times n,\mathbb{K})$ we define the *operator norm*

$$\|A\| := \sup\{\|Ax\| : x \in \mathbb{K}^n \textit{ and } \|x\| \le 1\} = \sup\{\|Ax\| : x \in \mathbb{K}^n \textit{ and } \|x\| = 1\}$$

as the supremum on the unit ball of $\mathbb{K}^n$ of the linear map represented by A with respect to the canonical basis.

Note that $\|A\| < \infty$ due to compactness of the unit ball

$$\{x \in \mathbb{K}^n : \|x\| \le 1\}.$$

Intuitively, the operator norm of A measures how the linear map determined by A with respect to the canonical basis of $\mathbb{K}^n$ blows-up or blows-down the unit ball of $\mathbb{K}^n$.

The $\mathbb{K}$-vector space $M(n \times n, \mathbb{K})$ of all matrices with components from $\mathbb{K}$ is an associative $\mathbb{K}$-algebra with respect to the matrix product

$$A \cdot B \in M(n \times n, \mathbb{K})$$

because

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

for matrices $A, B, C \in M(n \times n, \mathbb{K})$.

Proposition 1.3 (Normed associative matrix algebra). *The matrix algebra* $(M(n \times n, \mathbb{K}), \|\|)$ *is a normed associative algebra:*

1. $\|A\| = 0$ *iff* $A = 0$
2. $\|A + B\| \le \|A\| + \|B\|$ *(Triangle inequality)*
3. $\|\lambda \cdot A\| = |\lambda| \cdot \|A\|$ *with* $\lambda \in \mathbb{K}$
4. $\|A \cdot B\| \le \|A\| \cdot \|B\|$ *(Product estimation)*
5. $\|\mathbb{1}\| = 1$ *with* $\mathbb{1} \in M(n \times n, \mathbb{K})$ *the unit matrix.*
6. *For a matrix* $A = (a_{ij})_{i,j} \in M(n \times n, \mathbb{K})$ *holds*
$$\|A\|_{sup} \le \|A\| \le n \cdot \|A\|_{sup}$$
with the supremum norm *of the matrix components*
$$\|A\|_{sup} := \sup\{|a_{ij}| : 1 \le i, j \le n\}.$$
7. *Each eigenvalue* $\lambda \in \mathbb{C}$ *of a matrix* $A \in M(n \times n, \mathbb{K})$ *satisfies the estimate*
$$|\lambda| \le \|A\|$$

Proof. ad 6) For $j = 1, ..., n$ denote by e_j the j-th canonical basis vector of $\mathbb{K}^n$. Then for all $i = 1, ..., n$

$$\|A\| \geq \|Ae_j\| = \sqrt{\sum_{k=1}^{n} |a_{kj}|^2} \geq |a_{ij}|,$$

hence

$$\|A\| \geq \|A\|_{sup}.$$

Concerning the other estimate, the inequality of Cauchy Schwarz

$$| < x, y > |^2 \leq \|x\|^2 \cdot \|y\|^2$$

implies for all $x = (x_1, ..., x_n) \in \mathbb{K}^n$, in particular for $|x| \leq 1$

$$\|Ax\|^2 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} \cdot x_j \right|^2 \leq$$

$$\leq \sum_{i=1}^{n} \left(\sum_{j=1}^{n} |a_{ij} \cdot x_j|^2 \right) \leq \sum_{i=1}^{n} \left(\sum_{j=1}^{n} |a_{ij}|^2 \cdot \sum_{j=1}^{n} |x_j|^2 \right)$$

$$\leq \|x\|^2 \cdot \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \sup_{i,j} |a_{ij}|^2 \right) = n^2 \cdot \|x\|^2 \|A\|_{sup}^2,$$

hence

$$\|A\| \leq n \cdot \|A\|_{sup}$$

ad 7) Consider an eigenvector $\mathrm{v} \in \mathbb{K}^n$ with eigenvalue λ. If

$$\lambda \cdot \mathrm{v} = A \cdot \mathrm{v}$$

then

$$|\lambda| \cdot \|\mathrm{v}\| = \|\lambda \cdot \mathrm{v} = \|A \cdot \mathrm{v}\| \leq \|A\| \cdot \|\mathrm{v}\|,$$

and cancelling $\|\mathrm{v}\| \neq 0$ proves the claim. □

Remark 1.4 (Equivalence of norms).

1. Because the operator norm $\|A\|$ and the sup-norm refering to the entries of A dominate each other up to a constant, the following two convergence concepts are equivalent for a sequence $(A_\nu)_{\nu \in \mathbb{N}}$ of matrices

$$A_\nu \in M(n \times n, \mathbb{K}), \nu \in \mathbb{N},$$

and a matrix $A \in M(n \times n, \mathbb{K})$:

- $\lim_{\nu\to\infty} \|A_\nu - A\| = 0$, i.e. $\lim_{\nu\to\infty} A_\nu = A$ (norm convergence).
- The sequence $(A_\nu)_{\nu\in\mathbb{N}}$ converges to A componentwise.

2. In particular each Cauchy sequence of matrices wih respect to the operator norm is convergent: The matrix algebra

$$(M(n \times n, \mathbb{K}), \| \|)$$

is complete, i.e. it is a *Banach algebra*.

3. The vector space $M(n \times n, \mathbb{K})$ is finite dimensional, its dimension is n^2. Hence any two norms are equivalent in the sense that they dominate each other. Therefore the structure of a topological vector space on $M(n \times n, \mathbb{K})$ does not depend on the choice of the norm.

Because $(M(n \times n, \mathbb{K}), \| \|)$ is a normed algebra according to Proposition 1.3, concepts from analysis like convergence, Cauchy sequence, continuous function, and power series also apply to matrices. In particular, the commutator is a continuous map: For $A,\ B \in M(n \times n, \mathbb{K})$

$$\|[A,B]\| = \|AB - BA\| \leq \|AB\| + \|BA\| \leq 2\|A\|\|B\|.$$

We recall the fundamental properties of a complex power series

$$\sum_{\nu=0}^{\infty} c_\nu \cdot z^\nu$$

with radius of convergence $R > 0$. Set

$$\Delta(R) := \{z \in \mathbb{C} : |z| < R\},$$

the open disc in $\mathbb{C}$ with radius R. Then

- The series is absolutely convergent in $\Delta(R)$, i.e. $\sum_{\nu=0}^{\infty} |c_\nu| \cdot |z|^\nu$ is convergent for $z \in \Delta(R)$.
- The series converges compact in $\Delta(R)$, i.e. the convergence is uniform on each compact subset of $\Delta(R)$.
- The series is infinitely often differentiable in $\Delta(R)$, its derivation is obtained by termwise differentiation.

Lemma 1.5 (Power series of matrices). *Consider a power series*

$$f(z) = \sum_{\nu=0}^{\infty} c_\nu \cdot z^\nu$$

with coefficients $c_\nu \in \mathbb{K}, \nu \in \mathbb{N}$, and radius of convergence $R > 0$. Let

$$B(R) := \{A \in M(n \times n, \mathbb{K}) : \|A\| < R\}$$

be the open ball in $M(n \times n, \mathbb{K})$ around zero with radius R. Then:

- *The series*

$$f(A) := \sum_{\nu=0}^{\infty} c_\nu \cdot A^\nu := \lim_{n\to\infty} \left(\sum_{\nu=0}^{n} c_\nu \cdot A^\nu \right) \in M(n \times n, \mathbb{K})$$

  is absolutely convergent and compact convergent in $B(R)$.

- *For each matrix $A \in B(R)$ the series $f(A)$ satisfies*

$$[f(A), A] = 0$$

  with the commutator

$$[f(A), A] := f(A) \cdot A - A \cdot f(A).$$

- *The function*

$$f : B(R) \to M(n \times n, \mathbb{K}),\ A \mapsto f(A),$$

  is continuous.

Proof. i) We apply the Cauchy criterion: For $N > M$ holds

$$\left\| \sum_{\nu=0}^{N} c_\nu \cdot A^\nu - \sum_{\nu=0}^{M} c_\nu \cdot A^\nu \right\| \leq \sum_{\nu=M+1}^{N} |c_\nu| \cdot \|A\|^\nu.$$

If

$$r := \|A\| < R$$

then

$$\sum_{\nu=0}^{\infty} |c_\nu| \cdot r^\nu$$

converges. Hence the Cauchy criterion is satisfied with respect to the operator norm $\|...\|$, hence a posteriori with respect to the sup-norm $\|...\|_{sup}$. The limit

$$\lim_{n\to\infty} \left(\sum_{\nu=0}^{n} c_\nu \cdot A^\nu \right)$$

exists with respect to $\|...\|_{sup}$, hence also with respect to the operator norm. The remaining statements about convergence follow from the estimate

$$\|f(A)\| \leq \sum_{\nu=0}^{\infty} |c_\nu| \cdot \|A\|^\nu$$

and the corresponding properties of complex power series.

ii) The equation $[A, f(A)] = 0$ about the commutator follows: We take the limit $\lim_{n\to\infty}$ and employ the continuity of the commutator

$$[f(A), A] = \left[A, \sum_{\nu=0}^{\infty} c_\nu \cdot A^\nu\right] = \lim_{n\to\infty} \left[A, \sum_{\nu=0}^{n} c_\nu \cdot A^\nu\right] = \lim_{n\to\infty} \sum_{\nu=0}^{n} c_\nu \cdot [A, \cdot A^\nu] = 0$$

iii) Continuity of f is a consequence of the compact convergence. $\square$

Proposition 1.6 (Transposition and base change for power series of matrices). *Consider a complex power series*

$$f(z) = \sum_{\nu=0}^{\infty} a_\nu \cdot z^\nu$$

with radius of covergence $R > 0$ and a matrix $A \in M(n \times n, \mathbb{C})$ with $\|A\| < R$. Then:

i) Transposition*:*

$$f(A^\top) = f(A)^\top$$

ii) Similarity*:For each invertible matrix $S \in GL(n, \mathbb{C})$*

$$f(S \cdot A \cdot S^{-1}) = S \cdot f(A) \cdot S^{-1}$$

Proof. i) The proof follows from the fact that transposition is a continuous map: Proposition 1.3 implies

$$\|A^\top\| \leq n \cdot \|A\|$$

Then taking the limit $\lim_{N\to\infty}$ of

$$\sum_{\nu=0}^{N} a_\nu \cdot (A^\top)^\nu = \left(\sum_{\nu=0}^{N} a_\nu \cdot A^\nu\right)^\top$$

ii) For each $\nu \in \mathbb{C}$

$$(S \cdot A \cdot S^{-1})^\nu = S \cdot A^\nu \cdot S^{-1}.$$

The claim of the proposition follows from the continuity of the matrix multiplication by taking the limit $\lim_{N\to\infty}$ of

$$S \cdot \left(\sum_{\nu=0}^{N} a_\nu \cdot A^\nu \right) \cdot S^{-1} = \sum_{\nu=0}^{N} a_\nu \cdot S \cdot A^\nu \cdot S^{-1} = \sum_{\nu=0}^{N} a_\nu \cdot (S \cdot A \cdot S^{-1})^\nu$$

□

One can even show for $A \in M(n \times n, \mathbb{C})$ the strict equality

$$\|A^\top\| = \|A\|$$

Definition 1.7 (Derivation of matrix functions). Let $I \subset \mathbb{R}$ be an open interval. A matrix function

$$A : I \to M(n \times n, \mathbb{K})$$

is *differentiable* at a point $t_0 \in I$ iff the limit

$$\lim_{h \to 0} \frac{A(t_0 + h) - A(t_0)}{h} =: A'(t_0) \in M(n \times n, \mathbb{K})$$

exists. In this case we employ the notation

$$\frac{dA}{dt}(t_0) := A'(t_0).$$

Lemma 1.8 (Derivation of power series of matrices depending on a parameter). *Let $I \subset \mathbb{R}$ be an open interval.*

i) Consider two differentiable maps

$$A, B : I \to B(R).$$

Then for all $t \in I$ holds the product rule

$$\frac{d}{dt} A(t) \cdot B(t) = A'(t) \cdot B(t) + A(t) \cdot B'(t).$$

ii) Consider a complex power series

$$f(z) = \sum_{\nu=0}^{\infty} c_\nu \cdot z^\nu$$

with radius of convergence $R > 0$, and

$$A : I \to B(R) \subset M(n \times n, \mathbb{K})$$

a differentiable function, which satisfies for all $t \in I$

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

Then also the function

$$f \circ A : I \to M(n \times n, \mathbb{K}),\ t \mapsto f(A(t)) := \sum_{\nu=0}^{\infty} c_\nu \cdot A^\nu(t),$$

is differentiable for all $t \in I$ and satisfies the chain rule

$$\frac{d}{dt} f(A(t)) = f'(A(t)) \cdot A'(t) = A'(t) \cdot f'(A(t)).$$

Here f' denotes the derivation of the complex power series f term by term.

Proof. i) The proof follows the proof of the product rule from calculus: One inserts a suitable additional term.

$$\lim_{h\to 0} \frac{A(t+h) \cdot B(t+h) - A(t) \cdot B(t)}{h} =$$

$$= \lim_{h\to 0} \frac{A(t+h) \cdot B(t+h) - A(t) \cdot B(t+h) + A(t) \cdot B(t+h) - A(t)B(t)}{h} =$$

$$\lim_{h\to 0} \frac{A(t+h) - A(t)}{h} \cdot \lim_{h\to 0} B(t+h) + A(t) \cdot \lim_{h\to 0} \frac{B(t+h) - B(t)}{h}$$

ii) i) After differentiating the series

$$f(A(t)) = \sum_{\nu=0}^{\infty} c_\nu \cdot A^\nu(t)$$

term by term, the resulting power series

$$\sum_{\nu=1}^{\infty} \nu \cdot c_\nu \cdot A^{\nu-1}(t)$$

is again compactly convergent. Hence

$$f'(A(t)) = \sum_{\nu=1}^{\infty} \nu \cdot c_\nu \cdot A^{\nu-1}(t)$$

The proof of the lemma reduces to proving

$$\frac{d}{dt} A^\nu(t) = \nu \cdot A^{\nu-1}(t) \cdot A'(t) = \nu \cdot A'(t) \cdot A^{\nu-1}(t).$$

Here the proof goes by induction on $\nu \in \mathbb{N}$. The induction step uses the product formula from part i)

$$\frac{d}{dt}A^{\nu+1}(t) = \frac{d}{dt}\left(A^{\nu}(t)\cdot A(t)\right) = \left(\frac{d}{dt}A^{\nu}(t)\right)\cdot A(t) + A^{\nu}(t)\cdot A'(t)$$

and the induction assumption

$$\frac{d}{dt}A^{\nu}(t) = \nu \cdot A^{\nu-1}(t)\cdot A'(t).$$

As a consequence

$$\frac{d}{dt}A^{\nu+1}(t) = \nu \cdot A^{\nu-1}\cdot A'(t)\cdot A(t) + A^{\nu}(t)\cdot A'(t) =$$

$$= (\nu+1)\cdot A^{\nu}(t)\cdot A'(t) = (\nu+1)\cdot A'(t)\cdot A^{\nu-1}(t)$$

□

1.2 Jordan decomposition

The central aim of Jordan decomposition:

- Decompose a complex vector space V with respect to a given endomorphism $f \in End(V)$ into eigenspaces or generalized eigenspaces of f,

- and decompose f as the sum of two endomorphisms of special type, one semisimple while the other nilpotent.

Definition 1.9 (Eigenspaces and generalized eigenspaces). Consider a $\mathbb{K}$-vector space V, a fixed endomorphism $f \in End(V)$, and $\lambda \in \mathbb{K}$.

- If
$$V_{\lambda}(f) := ker(f-\lambda) \neq \{0\}$$
then $V_{\lambda}(f)$ is the *eigenspace* of f with respect to λ, which is named an *eigenvalue* of f.

- If
$$V^{\lambda}(f) := \bigcup_{n\in\mathbb{N}} ker(f-\lambda)^n \neq \{0\}$$
then $V^{\lambda}(f)$ is the *generalized eigenspace* of f with respect to λ.

Remark 1.10 (Generalized eigenspace).

1. All generalized eigenspaces $V^\lambda(f)$ are f-invariant, i.e.
$$f(V^\lambda(f)) \subset V^\lambda(f):$$
For all $k \in \mathbb{N}$
$$[(f-\lambda)^k, f] = 0.$$
If
$$x \in V^\lambda(f), \text{ i.e. } (f-\lambda)^k(x) = 0$$
for suitable $k \in \mathbb{N}$, then
$$(f-\lambda)^k(f(x)) = ((f-\lambda)^k \circ f)(x) = (f \circ (f-\lambda)^k)(x) = f(0) = 0.$$
2. Every non-zero generalized eigenspace $V^\lambda(f)$ contains at least one eigenvector $\mathrm{v} \in V$ of f with eigenvalue λ: Take a non-zero vector $\mathrm{v}_0 \in V^\lambda(f)$ and choose $n \in \mathbb{N}$ maximal with
$$\mathrm{v} := (f-\lambda)^n(\mathrm{v}_0) \neq 0.$$
Hence knowing a generalized eigenspace of f allows to find an eigenvector of f.

We recall the following types of matrix representation of an endomorphisms.

Definition 1.11 (Diagonalizable, triangularizable, nilpotent). Consider an n-dimensional $\mathbb{K}$-vector space V.

1. An endomorphism $f \in End(V)$ is *diagonalizable* iff it can be represented by a diagonal matrix from $M(n \times n, \mathbb{K})$.

2. An endomorphism $f \in End(V)$ is *triangularizable* iff V has a *flag*, i.e. a sequence $(V_i)_{i=0,...,n}$ of subspaces of V with
$$dim\ V_i = i \text{ and } V_i \subset V_{i+1},$$
which is *f-stable*, i.e. satisfying
$$f(V_i) \subset V_i,\ i = 1,...,n.$$
3. An endomorphism $f \in End(V)$ is *nilpotent* iff $f^k(V) = 0$ for a suitable $k \in \mathbb{N}$.

A matrix $A = (a_{ij}) \in M(n \times n, \mathbb{K})$ is an *upper triangular matrix* iff $a_{ij} = 0$ for all $j < i$. Apparently an endomorphism is triangularizable iff it can be represented by an upper triangular matrix from $M(n \times n, \mathbb{K})$.

For a vector space $V \neq \{0\}$ each nilpotent endomorphism $f \in End\ V$ has an eigenvector $\mathrm{v} \in V$ with eigenvalue zero: For the proof one chooses a non-zero element $w \in V$ and considers the greatest index $n \in \mathbb{N}^*$ with

$$f^n(w) \neq 0.$$

By nilpotency of f such an index exists. Then

$$\mathrm{v} := f^n(x)$$

is an eigenvector of f with eigenvalue zero.

Note: The only endomorphism $f \in End(V)$, which is both diagonalizable and nilpotent, is $f = 0$.

It is well-known from Linear Algebra that the sum of two diagonalizable endomorphisms, which commute with each other, is diagonalizable. A similar result holds for nilpotent endomorphisms.

Proposition 1.12. *Consider a finite-dimensional $\mathbb{K}$-vector space V and two nilpotent endomorphims $f,\ g \in End(V)$ with commutator $[f,g] = 0$. Then*

$$f + g \in End(V)$$

is nilpotent.

Proof. We choose an index $N \in \mathbb{N}$ with

$$f^N = g^N = 0.$$

Because $[f,g] = 0$ the binomial theorem applies and shows

$$(f+g)^{2N} = \sum_{\nu=0}^{2N} \binom{2N}{\nu} f^\nu \cdot g^{2N-\nu} = 0$$

because each summand has at least one factor equal to zero. □

The question whether an endomorphism f is triangularizable or even diagonalizable depends on the roots of its *characteristic polynomial*. These roots are the eigenvalues of f. The corresponding crtieria are stated in Propositions 1.14 and 1.15.

Definition 1.13 (Characteristic polynomial of an endomorphism). Denote by V a $\mathbb{K}$-vector space and by $f \in End(V)$ an endomorphism.

1. The *characteristic polynomial* of f is the polynomial

$$p_{char}(T) := det(T \cdot \mathbb{1} - A) \in \mathbb{K}[T]$$

with $\mathbb{1} \in M(n \times n, \mathbb{K})$ the unit matrix and $A \in M(n \times n, \mathbb{K})$ an arbitrary matrix representing f. The polynomial is independent from the representing matrix.

2. It is well-known that the roots $\lambda \in \mathbb{C}$ of p_{char} are the *eigenvalues* of f. We denote by

$$\mu(p_{char};\lambda) \in \mathbb{N}$$

the multiplicity of the root, i.e. the *algebraic multiplicity* of λ.

For later aplication we prove the following lemma about diagonal approximation.

Proposition 1.14 (Triangular form). *For an endomorphism $f \in End(V)$ with a finite dimensional $\mathbb{K}$-vector space V are equivalent:*

1. *The endomorphism f is triangularizable.*

2. *The characteristic polynomial p_{char} splits over $\mathbb{K}$ into a product of - not necessarily pairwise distinct - linear factors.*

For the proof see [12].

In particular, over $\mathbb{K} = \mathbb{C}$ every endomorphism is triangularizable.

Proposition 1.15 (Diagonal form). *For an endomorphism $f \in End(V)$ with a finite dimensional $\mathbb{K}$-vector space V are equivalent:*

1. *The endomorphism f is diagonalizable.*

2. *The characteristic polynomial p_{char} splits over $\mathbb{K}$ into linear factors, and for all eigenvalues λ of f the algebraic multiplicity equals the geometric multiplicity, i.e.*

$$\mu(p_{char};\lambda) = dim\, V_\lambda(f)(\text{Geometric multiplicity}).$$

3. *The vector space V splits as direct sum of eigenspaces*

$$V = \bigoplus_{\lambda\ eigenvalue} V_\lambda(f).$$

For the proof see [12].

Lemma 1.16 (Diagonal approximation). *For each matrix $B \in M(n \times n, \mathbb{C})$ exists a sequence $(B_\nu)_{\nu \in \mathbb{N}}$ of diagonalizable matrices $B_\nu \in M(n \times n, \mathbb{C})$ with*

$$B = \lim_{\nu \to \infty} B_\nu.$$

Here the limit of matrices is to be understood componentwise.

Proof. Over the algebraically closed base field $\mathbb{C}$ the given matrix B is triangularizable: There exists an invertible matrix $S \in GL(n, \mathbb{C})$ such that

$$A := S \cdot B \cdot S^{-1}$$

is an upper triangular matrix of the form

$$A = \Delta + N$$

with a diagonal matrix

$$\Delta = diag(\lambda_1, ..., \lambda_n)$$

and a strictly upper triangular matrix

$$N = (a_{ij}), a_{ij} = 0 \text{ if } i \geq j.$$

For all $i = 1, ..., n$ one defines successively sequences

$$a^i = (a^i_\nu)_{\nu \in \mathbb{N}}$$

of complex numbers converging to zero, such that for each fixed $\nu \in \mathbb{N}$ the numbers

$$\lambda_i + a^i_\nu,\ i = 1, .., n,$$

are pairwise distinct. Then one defines for each $\nu \in \mathbb{N}$

$$A_\nu := diag(\lambda_1 + a^1_\nu, ..., \lambda_n + a^n_\nu) + N.$$

Each matrix A_ν has the n pairwise distinct eigenvalues

$$\lambda_i + a^i_\nu, i = 1, ..., n,$$

and is therefore diagonalizable due to Proposition 1.15. By construction

$$A = \lim_{\nu \to \infty} A_\nu,$$

which implies

$$B = S^{-1} \cdot A \cdot S = S^{-1} \cdot (\lim_{\nu \to \infty} A_\nu) \cdot S = \lim_{\nu \to \infty} (S^{-1} \cdot A_\nu \cdot S)$$

The matrices

$$B_\nu := S^{-1} \cdot A_\nu \cdot S,\ \nu \in \mathbb{N},$$

are diagonalizable because the corresponding matrices A_v are diagonalizable. $\square$

A second polynomial from $\mathbb{K}[T]$ which encodes important properties of an endomorphism f is the *minimal polynomial* of f.

Definition 1.17 (Minimal polynomial and semisimpleness). Let V be an n-dimensional $\mathbb{K}$-vector space. The $\mathbb{K}$-vector space $End(V)$ of endomorphisms has dimension $= n^2$. Hence for each endomorphism $f \in End(V)$ the family $(f^k)_{k\in\mathbb{N}}$ is linearly dependent and each endomorphism $f \in End(V)$ satisfies a polynomial equation

$$f^k = \sum_{i=0}^{k-1} \alpha_i \cdot f^i,\ \alpha_i \in \mathbb{C}.$$

with suitable $k \in \mathbb{N}^*$.

1. Because the ring $R := \mathbb{K}[T]$ is a principal ideal domain, the ideal of all polynomials which annihilate f

$$< p \in R :\ p(f) = 0 \in End(V) >$$

has a unique generator of positive degree with leading coefficient $= 1$. It is named the *minimal polynomial of f*

$$p_{min}(T) \in R.$$

2. The endomorphism f is named *semisimple* if its minimal polynomial splits as

$$p_{min}(T) = \prod_{j=1,...,k} g_j(T)$$

with irreducible polynomials

$$g_j(T) \in \mathbb{K}[T],\ j = 1,...,k,$$

which are pairwise distinct up to scalars.

If V is a complex vector space and $f \in End(V)$, then *semisimpleness* of f reduces to the property, that the roots of the minimal polynomial of f are pairwise distinct, i.e. $p_{min}(T)$ is a product of pairwise distinct linear factors.

Lemma 1.18 (Restriction of semisimple endomorphisms). *Consider a complex vector space V, an endomorphism $f \in End(V)$, and an f-invariant subspace $W \subset V$. If f is semisimple, then also the restriction*

$$f|W : W \to W$$

is semisimple.

Proof. The minimal polynomial

$$p_{(min,f|W)}(T)$$

of the restriction $f|W$ divides the minimal polynomial $p_{min}(T)$ of f. Hence also

$$p_{(min,f|W)}(T)$$

splits into pairwise distinct linear factors. □

Note that Lemma 1.18 also holds in the real context.

Theorem 1.19 (Jordan decomposition). *Let $f \in End(V)$ be an endomorphism of a finite dimensional complex vector space V.*

1. *A unique decomposition*

$$f = f_s + f_n \text{ (Jordan decomposition)}$$

exists with a semisimple endomorphism $f_s \in End(V)$ and a nilpotent endomorphism $f_n \in End(V)$ such that both satisfy

$$[f_s, f_n] = 0.$$

2. *The two summands f_s and f_n depend on f in a polynomial way, i.e. polynomials*

$$p_s(T), p_n(T) \in \mathbb{C}[T]$$

exist with $p_s(0) = p_n(0) = 0$ such that

$$f_s = p_s(f) \text{ and } f_n = p_n(f).$$

In particular, if $[f, g] = 0$ for an endomorphism $g \in End(V)$ then

$$[f_s, g] = [f_n, g] = 0.$$

3. *The vector space V splits as direct sum of the generalized eigenspaces of f*

$$V = \bigoplus_{\lambda \text{ eigenvalue}} V^{\lambda}(f).$$

For each eigenvalue λ the generalized eigenspace of f equals the eigenspace of f_s:

$$V^{\lambda}(f) = V_{\lambda}(f_s).$$

4. *The minimal polynomial $p_{min}(T)$ of f and the characteristic polynomial $p_{char}(T)$ of f have the same roots.*

The proof uses the fact that the field $\mathbb{C}$ is algebraically closed. Hence the minimal polynomial p_{min} of f splits completely into linear factors. If p_{min} has the roots

$$\lambda_i,\ i = 1,...,r,$$

then the linear factors of p_{min} induce a family of polynomials without a common factor. Because

$$R := \mathbb{C}[T]$$

is a principal domain, the factors generate a partition of unity in R. It induces a partition of the identity $id \in End(V)$. The corresponding summands form a family of pairwise commuting projectors

$$E_i : V \to V,\ i = 1,...,r$$

Setting

$$V_i := im\ E_i \subset V,\ i = 1,...,r,$$

decomposes V as the direct sum

$$V = \bigoplus_{i=1}^{r} V_i,$$

of f-stable subspaces V_i. One checks that

$$f_s := \sum_{i=1}^{r} \lambda_i \cdot E_i$$

is semisimple and that

$$f - f_s =: f_n$$

is the searched nilpotent summand of f. One checks that the direct summands V_i equal the generalized eigenspaces of f, and that these are also the eigenspaces of f_s.

Proof (of Theorem 1.19).

i) *Splitting id_V as a sum of pairwise commuting, non-zero projectors*: Because the field $\mathbb{C}$ is algebraically closed, the minimal polynomial $p_{min}(T)$ of f splits into linear factors with positive exponents $m_i \in \mathbb{N}^*$

$$p_{min}(T) = \prod_{i=1}^{r} (T - \lambda_i)^{m_i}$$

with $\lambda_i \neq \lambda_j$ if $i \neq j$. For $i = 1,...,r$ consider the polynomials

$$p_i := \frac{p_{min}(T)}{(T - \lambda_i)^{m_i}} \in R,$$

obtained by cancelling the corresponding factor from the minimal poylnomial. By construction these polynomials are coprime, their greatest common divisor is $\pm 1 \in R$. Because R is a principal ideal domain there exist polynomials $r_i \in R,\ i = 1, ..., r$, with

$$1 = \sum_{i=1}^{r} r_i \cdot p_i \in R.$$

Applying the polynomial equation to the endomorphism f creates the endomorphisms

$$E_i := r_i(f) \cdot p_i(f) \in End(V),\ i = 1, ..., r.$$

They satisfy:

- By construction
$$id_V = \sum_{i=1}^{r} E_i$$
- For each $i = 1, ..., r$
$$E_i \neq \{0\}.$$
Otherwise assume the existence of an index $i \in \{1, ..., r\}$ with $E_i = \{0\}$, w.l.o.g. $i = 1$. Then
$$r_1(f) \cdot p_1(f) = 0$$
The minimality of p_{min} implies
$$p_{min} \text{ divides } r_1 \cdot p_1 \text{ in } R,$$
and by definition of p_1
$$p_{min} = (T - \lambda_1)^{m_1} \cdot p_1$$
Therefore
$$(T - \lambda_1)^{m_1} \cdot p_1 \text{ divides } r_1 \cdot p_1 \text{ in } R,$$
and because R is a domain of integrity
$$(T - \lambda_1)^{m_1} \text{ divides } r_1.$$
By definition $(T - \lambda_1)^{m_1}$ also divides all $p_j,\ j = 2, ..., r$. Hence
$$(T - \lambda_1)^{m_1} \text{ divides } 1 = \sum_{i=1}^{r} r_i \cdot p_i,$$
a contradiction.
- If $i \neq j$ then p_{min} divides
$$(r_i \cdot p_i) \cdot (r_j \cdot p_j) \in R.$$

Hence $p_{min}(f) = 0 \in End(V)$ and the minimality of p_{min} imply

$$E_i \cdot E_j = 0 \in End(V).$$

- The family $(E_k)_{k=1,...,r}$ is a family of projectors:

$$E_k^2 = E_k \circ \sum_{i=1}^{r} E_i = E_k \circ id_V = E_k.$$

Hence the family $(E_i)_{i=1,...,r}$ splits the identity id_V as a sum of pairwise commuting, non-zero projectors. We define the ranges

$$V_i := im\ E_i \subset V,\ i = 1, ..., r$$

of the projectors, and obtain the direct sum decomposition

$$V = \bigoplus_{i=1}^{r} V_i.$$

For all $i = 1, ..., r$ the projector E_i is as a polynomial in f by definition, hence it satisfies

$$[f, E_i] = 0,$$

which implies that the subspace V_i is f-stable.

ii) *The semisimple summand and* $V_i \subset V^{\lambda_i}(f)$: On each subspace

$$V_i, i = 1, ..., r,$$

the corresponding projector E_i acts as identity, hence the endomorphism

$$\lambda_i \cdot (E_i|V_i) \in End(V_i)$$

acts as multiplication by λ_i. We define the polynomial

$$p_s(T) := \sum_{i=1}^{r} \lambda_i \cdot r_i(T) \cdot p_i(T) \in R$$

and the endomorphism

$$f_s := p_s(f) = \sum_{i=1}^{r} \lambda_i \cdot E_i \in End(V).$$

Recall from part i) for all $i = 1, ..., r$

$$V_i \neq \{0\}.$$

Due to Proposition 1.15 the decomposition

$$V = \bigoplus_{i=1}^{r} V_i$$

implies that f_s is diagonalizable with eigenspaces

$$V_i = V_{\lambda_i}(f_s),\ i = 1, ..., r.$$

In particular, f_s is semisimple with minimal polynomial

$$p_{min,f_s} = \prod_{i=1}^{r} (T - \lambda_i)$$

For each $i = 1, ..., r$

$$V_i \subset V^{\lambda_i}(f)$$

because for each $\mathrm{v} \in V_i$ holds

$$(f-\lambda_i)^{m_i}(\mathrm{v}) = (f-\lambda_i)^{m_i}(E_i(\mathrm{v})) = (f-\lambda_i)^{m_i}(r_i(f) \circ p_i(f))(\mathrm{v}) = (r_i(f) \circ p_{min}(f))(\mathrm{v}) = 0,$$

the penultimate equality is due to the definition

$$p_i := \frac{p_{min}}{(T - \lambda_i)^{m_i}}$$

iii) *The nilpotent summand*: To obtain the nilpotent summand of f we consider the polynomial

$$p_n(T) := T - p_s(T) \in \mathbb{C}[T]$$

and define the corresponding endomorphism

$$f_n := p_n(f) = f - f_s \in End(V).$$

Then

$$f = f_s + f_n.$$

The two definitions

$$f_s := p_s(f) \text{ and } f_n := p_n(f)$$

imply

$$[f, f_s] = [f, f_n] = 0.$$

In order to prove the nilpotency of f_n we consider an arbitrary index $i \in 1, ..., r$ and the restriction of f_n to V_i. Set

$$m := dim\, V_i$$

On V_i

$$f_n = f - f_s = (f - \lambda_i) - (f_s - \lambda_i).$$

Using $[f_s, f] = 0$ the binomial theorem implies

$$f_n{}^m = \sum_{\mu=0}^{m} \binom{m}{\mu} (-1)^{m-\mu} (f-\lambda_i)^\mu \circ (f_s-\lambda_i)^{m-\mu}.$$

Here the second factor

$$(f_s-\lambda_i)^{m-\mu}$$

of each summand with index $0 \le \mu < m$ vanishes, because λ_i is an eigenvalue of the restriction $f_s|V_i$. The first factor

$$(f-\lambda_i)^\mu$$

vanishes for the summand with index $\mu = m$ as proved in part ii). As a consequence of the binomial theorem

$$f_n{}^m(V_i) = 0.$$

Hence the restriction $f_n|V_i$ is nilpotent for every $i = 1,...,r$, which implies the nilpotency of $f_n \in End(V)$.

iv) *Inclusion* $V^{\lambda_i}(f) \subset V_i$: Consider an arbitrary, but fixed $i = 1,...,r$. We use the direct sum decomposition from part ii)

$$V = \bigoplus_{j=1}^{r} V_j.$$

Consider an arbitrary vector $\mathrm{v} \in V^{\lambda_i}(f) \subset V$ and decompose

$$\mathrm{v} = \sum_{j=1}^{r} \mathrm{v}_j, \mathrm{v}_j \in V_j \text{ for } j = 1,...,r.$$

We have to show that v reduces to v_i, i.e.$\mathrm{v}_j = 0$ for $j \ne i$.

For large $n \in \mathbb{N}$ by assumption

$$0 = (f-\lambda_i)^n(\mathrm{v}) = \sum_{j=1}^{r} (f-\lambda_i)^n(\mathrm{v}_j).$$

The f-stableness of each V_j implies

$$(f-\lambda_i)^n(\mathrm{v}_j) \in V_j.$$

Assume the existence of an index $j = 1,...,r$ such that $\mathrm{v}_j \ne 0$. Then

$$(f-\lambda_i)^n(\mathrm{v}_j) = 0,$$

i.e. f has the generalized eigenvector v_j belonging to the eigenvalue λ_i:

$$\mathrm{v}_j \in V^{\lambda_i}(f).$$

Due to Remark 1.10 the endomorphism f has also an eigenvector $w \in V_{\lambda_i}(f)$. The f-stableness of V_j implies $w \in V_j$. We obtain for large m

$$0 = (f - \lambda_j)^m(w) = (\lambda_i - \lambda_j)^m \cdot w.$$

Here the left equality is due to

$$w \in V_j \subset V^{\lambda_j}(f)$$

with the last inclusion proven in part i). And the right equality is due to

$$w \in V_{\lambda_i}(f).$$

The equality

$$0 = (\lambda_i - \lambda_j)^m \cdot w$$

implies $j = i$. As a consequence $\mathrm{v} = \mathrm{v}_i \in V_i$, which finishes the proof of the inclusion

$$V^{\lambda_i}(f) \subset V_i$$

Together with the opposite inclusion from part iii) we obtain

$$V_i = V^{\lambda_i}(f)$$

and

$$V = \bigoplus_{i=1}^{r} V_i = \bigoplus_{i=1}^{r} V^{\lambda_i}(f) = \bigoplus_{i=1}^{r} V_{\lambda_i}(f_s).$$

In particular, the vector space V splits as the direct sum of the generalized eigenspaces of f. The corresponding generalized eigenvalues of f are the roots of the minimal polynomial p_{min} of f.

v) *Both polynomials p_{min} and p_{char} have the same roots*:

- Due to part iv) each root λ of the minimal polynomial of f defines a generalized eigenspace $V^{\lambda}(f)$. Due to Remark 1.10 the latter contains an eigenvector of f with eigenvalue λ. The eigenvalue λ is a root of the characteristic polynomial of f. Hence all roots of the minimal polynomial are also roots of the characteristic polynomial.

- For the oppposite direction we consider an eigenvalue λ of f with corresponding eigenvector $\mathrm{v} \in V$:

$$f(\mathrm{v}) = \lambda \cdot \mathrm{v}.$$

Because $p_{min}(f) = 0 \in End(V)$ we have in particular

$$p_{min}(f)(\mathrm{v}) = 0 \in V.$$

Hence
$$0 = p_{min}(f)(\mathrm{v}) = p_{min}(\lambda) \cdot \mathrm{v} \in V.$$

Because $\mathrm{v} \neq 0$ we conclude
$$p_{min}(\lambda) = 0 \in \mathbb{C},$$

i.e. the eigenvalue λ is a root of the minimal polynomial.

vi) *Semisimplicity implies diagonalizability*: If f is semisimple then each λ_j is a simple root of the minimal polynomial of f, i.e.
$$p_{min}(T) = \prod_{j=1}^{r} (T - \lambda_j)$$

We show for each arbitrary but fixed index $i = 1, ..., r$
$$V^{\lambda_i}(f) = V_{\lambda_i}(f)$$

To prove the non-trivial inclusion
$$V^{\lambda_i}(f) \subset V_{\lambda_i}(f),$$

i.e. that generalized eigenvectors are eigenvectors, we consider a generalized eigenvector $\mathrm{v} \in V^{\lambda_i}(f)$. We have
$$0 = p_{min}(f)(\mathrm{v}) = \prod_{j=1}^{r} (f - \lambda_j)(\mathrm{v}) = \left((\prod_{j \neq i} (f - \lambda_j)) \circ (f - \lambda_i) \right)(\mathrm{v})$$

In case
$$(f - \lambda_i)(\mathrm{v}) \neq 0$$

we obtain due to the f-stableness of $V^{\lambda_i}(f)$ by iteration a non-zero vector $w \in V^{\lambda_i}(f)$ and an index $j \neq i$ with
$$(f - \lambda_j)(w) = 0$$

In particular
$$\mathrm{w} \in V_{\lambda_j}(f) \subset V^{\lambda_j}(f).$$

Hence
$$0 \neq w \in V^{\lambda_j}(f) \cap V^{\lambda_i}(f),$$

a contradiction to
$$V^{\lambda_j}(f) \cap V^{\lambda_i}(f) = \emptyset$$

according to part iv). Therefore
$$(f - \lambda_i)(\mathrm{v}) = 0, \ i.e. \ \mathrm{v} \in V_{\lambda_i}(f).$$

vii) *Uniqueness of the Jordan decomposition*: Assume a second decomposition

$$f = f'_s + f'_n$$

with the properties stated in part 1) of the theorem. From

$$[f'_s, f'_n] = 0$$

follows

$$[f, f'_n] = [f'_s, f'_n] + [f'_n, f'_n] = [f'_n, f'_n] = 0.$$

Part iii) with

$$f_n = p_n(f)$$

shows

$$[f_n, f'_n] = 0.$$

Analogously, one proves

$$[f_s, f'_s] = 0.$$

As a consequence

$$f_s - f'_s = f'_n - f_n \in End(V)$$

is an endomorphism which is both

- semisimple, as the sum of two commuting semisimple, i.e. diagonalizable according to part vi), endomorphisms,
- and nilpotent - being the sum of two commuting nilpotent endomorphisms, see Proposition 1.12.

Hence the endomorphism is zero, i.e.

$$f_s = f'_s \text{ and } f_n = f'_n.$$

viii) *Killing the constant terms*: We prove that the polynomials p_s and p_n can be choosen with

$$p_s(0) = p_n(0) = 0:$$

- Either $\lambda_i = 0$ for one index $i = 1, ..., r$. Then

  $$V^0(f) \neq \{0\}$$

  and the generalized eigenspace of f contains also an eigenvector of f with eigenvalue 0, i.e.

  $$\{0\} \neq ker\ f.$$

  The representation

  $$p_s(T) = a_0 + a_1 T + ... + a_k T^k \in \mathbb{C}[T]$$

  implies

$$f_s = p_s(f) = a_0 \cdot id + a_1 \cdot f + ... + a_k \cdot f^k \; End(V).$$

The last equation applied to an eigenvector

$$v \in ker\ f = ker\ f_s$$

shows $a_0 = 0$. Hence $p_s(T) \in \mathbb{C}[T]$ has no constant term, i.e. $p_s(0) = 0$. From the definition

$$p_n(T) := T - p_s(T)$$

follows

$$p_n(0) = -p_s(0) = 0.$$

- Or $\lambda_i \neq 0$ for all indices $i = 1, ..., r$: Then

$$p_{min}(0) = \pm \prod_{i=1}^{r} \lambda_i \neq 0$$

Now one replaces the polynomials p_s and p_n respectively by

$$\tilde{p}_s := p_s - \frac{p_s(0)}{p_{min}(0)} \cdot p_{min}$$

and

$$\tilde{p}_n := p_n - \frac{p_n(0)}{p_{min}(0)} \cdot p_{min}.$$

Then

$$\tilde{p}_s(0) = \tilde{p}_n(0) = 0$$

without changing the polynomial representations

$$f_s = \tilde{p}_s(f) \text{ and } f_n = \tilde{p}_n(f),$$

because $p_{min}(f) = 0$.

□

The decomposition of a complex endomorphism into the sum of its semisimple and its nilpotent part points to the two fundamental classes of Lie algebras. Nilpotent - and slightly more general - solvable Lie algebras are the subject of Chapter 3. While the study of semisimple Lie algebras will start in Chapter 4 and continue as the subject of Part II.

The Cayleigh-Hamilton theorem relates the minimal polynomial of an endomorphism to its characteristic polynomial. The theorem can be obtained as a corollary of the Jordan decomposition.

Theorem 1.20 (Cayleigh-Hamilton). *Consider a complex, finite-dimensional vector space V and an endomorphism $f \in End(V)$.*

- *The characteristic polynomial* $p_{char}(T) \in \mathbb{C}$ *annihilates* f*, i.e.*

$$p_{char}(f) = 0 \in End(V).$$

- *The minimal polynomial* $p_{min}(T)$ *of* f *divides the characteristic polynomial* $p_{char}(T)$ *in the ring* $\mathbb{C}[T]$.

Proof. We denote by

$$f = f_s + f_n$$

the Jordan decomposition of f according to Theorem 1.19 and take over the notation from its proof. The minimal polynomial of f has the form

$$p_{min}(T) = \prod_{i=1}^{r} (T - \lambda_i)^{m_i}$$

with pairwise distinct λ_i.

i) We choose an arbitrary but fixed index $j \in \{1, ..., r\}$, set

$$\lambda := \lambda_j,\ W := V^{\lambda}(f),\ \text{and}\ k := dim\, W,$$

and denote the restriction of endomorphisms to the f-stable subspace W by priming. Then the restrictions

$$f'_s = \lambda \cdot id_W \text{ and } f'_n$$

are semisimple respectively nilpotent, and

$$f' = f'_s + f'_n$$

The nilpotency of f'_n implies

$$(f' - \lambda)^k = (f' - f'_s)^k = (f'_n)^k = 0$$

ii) For each $i = 1, ..., r$ we use the shorthand

$$V_i := V^{\lambda_i}(f)$$

The characteristic polynomial of f is

$$p_{char}(T) = \prod_{i=1}^{r} (T - \lambda_i)^{dim\ V_i}$$

Part i) shows for each $i = 1, ..., r$

$$(f - \lambda_i)^{dim\ V_i} | V_i = 0$$

As a consequence of the Jordan decomposition

$$V = \bigoplus_{i=1}^{r} V_i$$

holds

$$p_{char}(f) = \prod_{i=1}^{r} (f - \lambda_i)^{dim\ V_i} = 0$$

Due to the minimality of $p_{min}(T)$ the minimal polynomial $p_{min}(T)$ divides the characteristic polynomial $p_{char}(T)$. $\square$

Apparently the statement $p_{char}(f) = 0$ of the Cayleigh-Hamilton theorem also holds for an endomorphism of a real vector space V. Because any real endomorphism f defines the complex endomomorphism

$$f \otimes id$$

of the complexification $V \otimes_{\mathbb{R}} \mathbb{C}$.

1.3 The exponential map of matrices

The complex *exponential series*

$$exp(z) = \sum_{\nu=0}^{\infty} \frac{z^{\nu}}{\nu!}$$

has convergence radius $R = \infty$. We now generalize the exponential series from complex numbers $z \in \mathbb{C}$ as argument to matrices $A \in M(n \times n, \mathbb{K})$ as argument.

Definition 1.21 (Exponential of matrices). For any matrix $A \in M(n \times n, \mathbb{K})$ one defines the *exponential*

$$exp(A) := \sum_{\nu=0}^{\infty} \frac{A^{\nu}}{\nu!} \in M(n \times n, \mathbb{K}).$$

In Definition 1.21 the defining power series of matrices is absolutely and compactly convergent due to Lemma 1.5.

Proposition 1.22 (Derivation of the exponential with respect to a parameter). *Consider an open interval $I \subset \mathbb{R}$ and a differentiable function*

$$A : I \to M(n \times n, \mathbb{K})$$

with

$$[A'(t), A(t)] = 0$$

for all $t \in I$. Then for all $t \in I$

$$\frac{d}{dt}(exp\ A(t)) = A'(t) \cdot exp\ A(t) = (exp\ A(t)) \cdot A'(t).$$

Proof. We apply Lemma 1.8 with the power series

$$f(z) = \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \cdot z^{\nu}$$

and its derivative

$$f'(z) = f(z).$$

We obtain

$$\frac{d}{dt}(exp\ A(t)) = \left(\sum_{\nu=0}^{\infty} \frac{1}{\nu!} \cdot A^{\nu}(t) \right) \cdot A'(t) = (exp\ A(t)) \cdot A'(t) = A'(t) \cdot exp\ A(t).$$

□

Theorem 1.23 (Exponential of commuting matrices). *The exponential of matrices $A, B \in M(n \times n, \mathbb{C})$ satisfies the following rules:*

1. Functional equation in the commutative case*: If $[A,B] = 0$ then*

$$exp(A+B) = exp(A) \cdot exp(B).$$

2. Determinant and trace*: $det(exp\ A) = exp(tr\ A)$.*

Proof. 1. First we apply the binomial theorem making use of the assumption $[A,B] = 0$:

$$exp(A+B) = \sum_{\nu=0}^{\infty} \frac{(A+B)^{\nu}}{\nu!} = \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \left(\sum_{\mu=0}^{\nu} \binom{\nu}{\mu} A^{\nu-\mu} \cdot B^{\mu} \right) =$$

$$= \sum_{\nu=0}^{\infty} \left(\sum_{\mu=0}^{\nu} \frac{A^{\nu-\mu}}{(\nu-\mu)!} \cdot \frac{B^{\mu}}{\mu!} \right)$$

Secondly we invoke the Cauchy product formula. It rests on the fact that any rearrangement of absolutely convergent series is admissible:

$$exp(A)\cdot exp(B)=\sum_{\nu=0}^{\infty}\frac{A^{\nu}}{\nu!}\cdot\sum_{\nu=0}^{\infty}\frac{B^{\nu}}{\nu!}=\sum_{\nu=0}^{\infty}\left(\sum_{\mu=0}^{\nu}\frac{A^{\nu-\mu}}{(\nu-\mu)!}\cdot\frac{B^{\mu}}{\mu!}\right).$$

2. According to Proposition 1.14 and 1.6 w.l.o.g.

$$A=\begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix},\ \lambda_i\in\mathbb{C},\ i=1,...,n,$$

is an upper triangular matrix. We obtain

$$A^{\nu}=\begin{pmatrix}\lambda_1^{\nu} & & * \\ & \ddots & \\ 0 & & \lambda_n^{\nu}\end{pmatrix},\nu\in\mathbb{N},$$

and conclude

$$det(exp\,A)=det\left(\sum_{\nu=0}^{\infty}\frac{A^{\nu}}{\nu!}\right)=det\begin{pmatrix}e^{\lambda_1} & & * \\ & \ddots & \\ 0 & & e^{\lambda_n}\end{pmatrix}.$$

Hence

$$det(exp\,A)=\prod_{i=1}^{n}e^{\lambda_i}=e^{\lambda_1+...+\lambda_n}=exp(tr\,A).$$

□

Corollary 1.24 (Exponential map). *The exponential defines a map*

$$exp:M(n\times n,\mathbb{K})\to GL(n,\mathbb{K}),\ A\mapsto exp(A),$$

i.e. the matrix $exp(A)$ is invertible, and $exp(A)^{-1}=exp(-A)$.

The inverse of exponentiation is taking the logarithm. But even in the case of complex numbers the exponential map is not injective because

$$exp(z+2\pi i)=exp(z)$$

Anyhow the exponential map of complex numbers is locally invertible. And this property carries over to the exponential of matrices around the unit $\mathbb{1}\in GL(n,\mathbb{C})$. We first recall the complex power series of the *logarithm*

$$log(1+z)=\sum_{\nu=1}^{\infty}\frac{(-1)^{\nu+1}}{\nu}\cdot z^{\nu},\ z\in\mathbb{C},$$

and the complex *geometric series*

$$\sum_{\nu=0}^{\infty} z^{\nu},\ z \in \mathbb{C}.$$

Both power series have radius of convergence $R = 1$.

Definition 1.25 (Logarithm and geometric series of matrices). For a matrix $A \in M(n \times n, \mathbb{K})$ with $\|A\| < 1$ one defines its *logarithm*

$$log(\mathbb{1} + A) := \sum_{\nu=1}^{\infty} (-1)^{\nu+1} \cdot \frac{A^{\nu}}{\nu} \in M(n \times n, \mathbb{K})$$

and its *geometric series*

$$\sum_{\nu=0}^{\infty} A^{\nu} \in M(n \times n, \mathbb{K}).$$

Proposition 1.26 (Inverse matrix and derivation of logarithm).

1. *For a matrix $A \in M(n \times n, \mathbb{K})$ with $\|A\| < 1$ the matrix $\mathbb{1} - A$ is invertible with inverse the geometric series*

$$(\mathbb{1} - A)^{-1} = \sum_{\nu=0}^{\infty} A^{\nu}.$$

2. *Consider an open interval $I \subset \mathbb{R}$ and a differentiable function*

$$B : I \to M(n \times n, \mathbb{K})$$

with $\|B(t) - \mathbb{1}\| < 1$ and $[B'(t), B(t)] = 0$ for all $t \in I$.

Then for all $t \in I$ the inverse $B(t)^{-1}$ exists and

$$\frac{d}{dt}(log\ B(t)) = B(t)^{-1} \cdot B'(t) = B'(t) \cdot B(t)^{-1}.$$

Proof. ad 1) For each $n \in \mathbb{N}$

$$(\mathbb{1} - A) \cdot \sum_{\nu=0}^{n} A^{\nu} = \sum_{\nu=0}^{n} A^{\nu} - \sum_{\nu=1}^{n+1} A^{\nu} = \mathbb{1} - A^{n+1}.$$

Because $\|A\| < 1$ it follows

$$(\mathbb{1} - A) \cdot \sum_{\nu=0}^{\infty} A^{\nu} = \mathbb{1} - \lim_{n\to\infty} \|A\|^{n+1} = \mathbb{1}.$$

ad 2) For $t \in I$ define

$$A := \mathbb{1} - B(t).$$

Because $\|A\| = \|\mathbb{1} - B(t)\| < 1$ apply part 1) to A:

$$B(t) = \mathbb{1} - A$$

has the inverse

$$B(t)^{-1} = \sum_{\nu=0}^{\infty} A^{\nu} = \sum_{\nu=0}^{\infty} (\mathbb{1} - B(t))^{\nu}.$$

In addition,

$$log(B(t)) = log(\mathbb{1} + (B(t) - \mathbb{1}))$$

is well-defined. The chain rule from Lemma 1.8 with the power series of the logarithm

$$f(1+z) = \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu+1}}{\nu} \cdot z^{\nu}$$

and its derivative

$$f'(1+z) = \sum_{\nu=0}^{\infty} (-z)^{\nu}$$

implies

$$\frac{d}{dt}(log\ B(t)) = \frac{d}{dt} log(\mathbb{1} + (B(t) - \mathbb{1})) = \left(\sum_{\nu=0}^{\infty} (\mathbb{1} - B(t))^{\nu} \right) \cdot B'(t) =$$

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

□

Proposition 1.27 (Exponential and logarithm of matrices as locally inverse maps). *For a matrix $A \in M(n \times n, \mathbb{C})$ holds:*

1. *If $\|A - \mathbb{1}\| < 1$ or $A - \mathbb{1}$ nilpotent, then*

$$exp(log\ A) = A$$

2. *If $\|A\| < log\ 2$ or A nilpotent, then*

$$log(exp\ A) = A$$

Proof. 1. i) Assume $\|A - \mathbb{1}\| < 1$: For the proof cf. [18, Theor. 2.8].

First we consider the case that the matrix A is diagonalizable: Assume the existence of matrix $S \in GL(n, \mathbb{C})$ with

$$S \cdot A \cdot S^{-1} = \Delta$$

a diagonal matrix

$$diag\ (\lambda_1, ..., \lambda_n).$$

Then for all $\nu \in \mathbb{N}$

$$(A-\mathbb{1})^\nu = S^{-1} \cdot \begin{pmatrix} (\lambda_1 - 1)^\nu & 0 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & (\lambda_n - 1)^\nu \end{pmatrix} \cdot S$$

The complex numbers $\lambda_j,\ j = 1, ..., n$ are the eigenvalues of A, while

$$\lambda_j - 1,\ j = 1, ..., n$$

are the eigenvalues of $A - \mathbb{1}$. Hence according to Proposition 1.3

$$|\lambda_j - 1| \leq \|A - \mathbb{1}\| < 1,\ j = 1, ..., n.$$

We obtain

$$log\ A = log(\mathbb{1} + (A - \mathbb{1})) = \sum_{\nu=1}^{\infty} (-1)^{\nu+1} \cdot \frac{(A-\mathbb{1})^\nu}{\nu} =$$

$$= S^{-1} \cdot \begin{pmatrix} log\ \lambda_1 & 0 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & log\ \lambda_n \end{pmatrix} \cdot S$$

Then by Theorem 1.23

$$exp(log\ A) = S^{-1} \cdot \begin{pmatrix} exp(log\ \lambda_1) & 0 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & exp(log\ \lambda_n) \end{pmatrix} \cdot S = A$$

Secondly, for a general matrix $A \in M(n \times n, \mathbb{C})$ Lemma 1.16 provides a sequence $(A_\nu)_\nu$ of diagonalizable matrices with

$$A = \lim_{\nu \to \infty} A_\nu$$

Due to the continuity of the matrix functions exp and log, see Lemma 1.5,

$$exp(log\ A) = exp(log(\lim_{\nu\to\infty} A_\nu)) = \lim_{\nu\to\infty}(exp(log\ A_\nu)) = \lim_{\nu\to\infty} A_\nu = A$$

ii) Assume $A - \mathbb{1}$ nilpotent, but not necessarily

$$\|A - \mathbb{1}\| < 1,$$

Consider for each $t \in \mathbb{R}$ the matrix

$$A(t) := \mathbb{1} + t \cdot (A - \mathbb{1})$$

For all $t \in \mathbb{R}$ the matrix

$$A(t) - \mathbb{1} = t \cdot (A - \mathbb{1})$$

is nilpotent, and for small $|t|$ holds

$$\|A(t) - \mathbb{1}\| = |t| \cdot \|A - \mathbb{1}\| < 1$$

For all $t \in \mathbb{R}$ the power series

$$log\, A(t) = log(\mathbb{1} + (A(t) - \mathbb{1}))$$

reduces to a finite series, and for small $|t|$ holds due to part i)

$$exp(log\, A(t)) = A(t)$$

Both sides of the last equality are power series in $t \in \mathbb{R}$. The identity theorem for real power series implies that the equality holds for all $t \in \mathbb{R}$. For $t = 1$ one has $A(t) = A$, hence

$$exp(log\, A) = A$$

2. i) Assume $\|A\| < log\, 2$. Then

$$\|exp(A) - \mathbb{1}\| = \|\sum_{\nu=1}^{\infty} \frac{A^\nu}{\nu!}\| \leq \sum_{\nu=1}^{\infty} \frac{\|A\|^\nu}{\nu!} = exp(\|A\|) - 1 < 2 - 1 = 1.$$

Hence

$$log(exp\, A) = log(\mathbb{1} + (exp(A) - \mathbb{1}))$$

is a well-defined convergent power series.

For a diagonal matrix

$$\Delta = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

holds

$$max\{|\lambda_j| : 1 \leq j \leq n\} = \|\Delta\|.$$

If $\|\Delta\| < log\, 2$ then

$$log(exp(\Delta)) = \begin{pmatrix} log(e^{\lambda_1}) & & 0 \\ & \ddots & \\ 0 & & log(e^{\lambda_n}) \end{pmatrix} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \Delta$$

Here we use that the power series of the logarithm computes the principal value of the logarithm.

According to Lemma 1.16 an arbitrary complex matrix $A \in M(n \times n, \mathbb{C})$ can be approximated by a series of diagonalizable matrices $(A_\nu)_{\nu \in \mathbb{N}}$

$$A = \lim_{\nu \to \infty} A_\nu.$$

If $\|A\| < log\ 2$ then for sufficiently large index ν also

$$\|A_\nu\| < 2$$

For each sufficiently large fixed $\nu \in \mathbb{N}$ there exists an invertible matrix $S \in GL(n, \mathbb{C})$ such that

$$\Delta_\nu := S \cdot A_\nu \cdot S^{-1}$$

is a diagonal matrix, and estimating the operator norm against the modulus of the largest eigenvalue shows

$$\|\Delta_\nu\| = \|A_\nu\| < log\ 2.$$

The previous step implies

$$log(exp\ S \cdot A_\nu \cdot S^{-1}) = log(exp\ \Delta_\nu) = \Delta_\nu$$

Hence

$$S \cdot log(exp\ A_\nu) \cdot S^{-1} = log(exp\ S \cdot A_\nu \cdot S^{-1}) = \Delta_\nu$$

implies

$$log(exp\ A_\nu) = S^{-1} \cdot \Delta_\nu \cdot S = A_\nu$$

Taking the limit of the outer terms of the last equation and using the continuity of the functions *log* and *exp* gives

$$log(exp(A)) = A.$$

ii) Assume A nilpotent, but not necessarily

$$\|A\| < log\ 2$$

Then the reduction is similar: For each $t \in \mathbb{R}$ the matrix

$$A(t) := t \cdot A$$

is nilpotent, and satisfies for small $|t|$

$$\|A(t)\| < log\ 2$$

For all $t \in \mathbb{R}$ the power series

$$exp(A(t)) - \mathbb{1} = \sum_{\nu=1}^{\infty} \frac{A(t)^\nu}{\nu!}$$

reduces to a finite series, and its value

$$exp(A(t)) - \mathbb{1}$$

is nilpotent as finite sum of pairwise commuting nilpotent matrices. Hence

$$log(\exp\ A(t)) = log(\mathbb{1} + (exp(A(t) - \mathbb{1})) = \sum_{\nu=1}^{\infty} (-1)^{\nu+1} \frac{((exp\ A(t)) - \mathbb{1})^\nu}{\nu}$$

also reduces to a finite sum and is therefore well-defined. For small $|t|$ holds due to part i)

$$log(exp\ A(t)) = A(t)$$

Both sides of the last equality are power series in $t \in \mathbb{R}$. The identity theorem for real power series implies that the equality holds for all $t \in \mathbb{R}$. For $t = 1$ one has $A(t) = A$, hence

$$log(exp\ A) = A$$

□

Remark 1.28 (Counter example against invertibility). In Proposition 1.27, part 2 the assumption $\|A\| < log\ 2$ cannot be dropped: Consider the matrix

$$A = 2\pi\ i \cdot \mathbb{1} \in M(n \times n, \mathbb{C})$$

with $\|A\| = 2\pi > log\ 2$. We have

$$exp(A) = e^{2\pi\ i} \cdot \mathbb{1} = \mathbb{1}$$

hence

$$log(exp(A)) = log(\mathbb{1}) = 0 \neq A.$$

Theorem 1.29 (Surjectivity of the complex exponential map). *The exponential map of complex matrices*

$$exp : M(n \times n, \mathbb{C}) \to GL(n, \mathbb{C}),\ A \mapsto exp\ A,$$

is surjective.

To obtain an inverse image of the endomorphism $f \in End(\mathbb{C}^n)$ represented by a given matrix $B \in GL(n, \mathbb{C})$ the Jordan decomposition of f allows to focus on a generalized eigenspace $V^\lambda(f)$ of f. Therefore we have to find an inverse image of the restriction

$$f_\lambda := f|V^\lambda(f).$$

The Jordan decomposition decomposes

$$f_\lambda = \lambda \cdot \mathbb{1} + f_n = \lambda \cdot \left(\mathbb{1} + \frac{1}{\lambda} \cdot f_n\right)$$

with a nilpotent endomorphism f_n and $\lambda \neq 0$ because f is invertible. We find inverse images separately for each factor of the induced multiplicative decomposition by using the logarithm for nilpotent matrices and for complex numbers. Then the sum of both inverse images maps to f_λ and solves the problem.

A given matrix $B \in GL(n,\mathbb{C})$ is similar to a block matrix with respect to the Jordan decomposition and the exponential map is compatible with conjugation due to Proposition 1.6. Hence we may assume B as a block matrix.

Proof. Consider a fixed but arbitrary matrix $B \in GL(n,\mathbb{C})$. According to Theorem 1.19 the vector space $\mathbb{C}^n$ splits up to conjugation into the sum of the generalized eigenspaces with respect to the generalized eigenvalues λ of B

$$\mathbb{C}^n \cong \bigoplus_\lambda V^\lambda(B)$$

and for each generalized eigenvalue λ the restriction

$$B_\lambda := B|V^\lambda(B)$$

is an endomorphism of the B-stable subspace $V^\lambda(B)$.

W.l.o.g. we may assume $B = B_\lambda$ with a fixed $\lambda \in \mathbb{C}$. Note $\lambda \neq 0$ because the matrix B is invertible. According to the Jordan decomposition of B the matrix

$$N := B_\lambda - \lambda \cdot \mathbb{1}$$

is nilpotent. We obtain an additive and a multiplicative decomposition

$$B_\lambda = \lambda \cdot \mathbb{1} + N = \lambda \cdot \left(\mathbb{1} + \frac{1}{\lambda} \cdot N\right)$$

Then

$$A_\lambda := log\left(\mathbb{1} + \frac{1}{\lambda} \cdot N\right) = \sum_{\nu=1}^{\infty} \frac{(-1)^{\nu+1}}{\nu} \cdot \frac{N^\nu}{\lambda^\nu}$$

is well-defined because the sum is finite. We have

$$exp\, A_\lambda = exp\left(log\left(\mathbb{1} + \frac{1}{\lambda} \cdot N\right)\right) = \mathbb{1} + \frac{1}{\lambda} \cdot N$$

according to Proposition 1.27. In order to deal with the numerical factor $\lambda \in \mathbb{C}$ we choose a complex logarithm $\mu \in \mathbb{C}$ with

$$e^{\mu} = \lambda.$$

Combining both steps we set

$$A := \mu \cdot \mathbb{1} + A_{\lambda},$$

Theorem 1.23 implies

$$exp\,A = exp(\mu \cdot \mathbb{1} + A_{\lambda}) = exp(\mu \cdot \mathbb{1}) \cdot exp\,A_{\lambda} = (\lambda \cdot \mathbb{1}) \cdot exp\,A_{\lambda} = \lambda \left(\mathbb{1} + \frac{1}{\lambda} \cdot N\right) = B.$$

□

Remark 1.30 (Counter examples against surjectivity). In general, the exponential map of matrices is not surjective on domains where one possibly expects it to be.

1. *Complex case*: Set

$$sl(2,\mathbb{C}) := \{A \in M(2 \times 2, \mathbb{C}) : tr\,A = 0\}$$

and

$$SL(2,\mathbb{C}) := \{B \in GL(2,\mathbb{C}) : det\,B = 1\}$$

and consider

$$exp : sl(2,\mathbb{C}) \to SL(2,\mathbb{C}).$$

The map is well-defined due to Theorem 1.23. Each matrix

$$B := \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} \in SL(2,\mathbb{C}),\ b \in \mathbb{C}^*,$$

has no inverse image.

2. *Real case*: Set

$$GL^+(2,\mathbb{R}) := \{B \in GL(2,\mathbb{R}) : det\,B > 0\}$$

and consider

$$exp : gl(2,\mathbb{R}) \to GL^+(2,\mathbb{R}).$$

Each matrix

$$B := \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} \in GL^+(2,\mathbb{R}),\ b \in \mathbb{R}^*,$$

has no inverse image.

Proof. 1. Assume the existence of a matrix $A \in sl(2,\mathbb{C})$ with

$$exp\,A = B.$$

The two complex eigenvalues

$$\lambda_i,\ i = 1,2,$$

of A satisfy

$$0 = tr\,A = \lambda_1 + \lambda_2.$$

- The case $\lambda_1 = \lambda_2$, i.e. $0 = \lambda_1 = \lambda_2$ is excluded, because then $exp\,A$ has the eigenvalue $e^0 = 1$. But B has the only eigenvalue -1.

- Hence $\lambda_1 \neq \lambda_2$. Then A is diagonalizable: A matrix $S \in GL(2,\mathbb{C})$ exists with

  $$S \cdot A \cdot S^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

  We obtain

  $$S \cdot B \cdot S^{-1} = S \cdot (exp\,A) \cdot S^{-1} = exp(S \cdot A \cdot S^{-1}) = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix}$$

  Hence B is diagonalizable. Then its minimal polynomial is

  $$p_{min}(T) = T + 1 \in \mathbb{C}[T],$$

  but

  $$p_{min}(B) \neq 0 \text{ because } b \neq 0.$$

  We obtain a contradiction.

2. Assume the existence of a matrix $A \in M(2 \times 2,\mathbb{R})$ with

$$exp\,A = B.$$

The real matrix A has two complex eigenvalues

$$\lambda_i,\ i = 1,2,$$

which are conjugate to each other.

- If both eigenvalues are real, then

  $$\lambda_1 = \lambda_2 =: \lambda \in \mathbb{R}$$

  and $exp\,A$ has the eigenvalue $e^{\lambda} > 0$, while B has the single eigenvalue -1, a contradiction.

- If λ_1 is not real, then

  $$\lambda_2 = \overline{\lambda_1} \neq \lambda_1.$$

Hence A has two distinct eigenvalues. Therefore A is diagonalizable. Hence also B is diagonalizable, a contradiction according to the end of part 1.

□

For later application in Proposition 2.13 we provide the following useful formula:

Proposition 1.31 (Lie-Trotter product formula). *For each pair of matrices $A, B \in M(n \times n, \mathbb{C})$ the exponential map satisfies*

$$exp(A+B) = \lim_{\nu \to \infty} \left(exp\ \frac{A}{\nu} \cdot exp\ \frac{B}{\nu} \right)^{\nu}.$$

In the proof we will use the standard notation: A function f belongs to the class

$$O(1/\nu^k)$$

iff there exists a positive constant C such that $|f|$ remains bounded by $\dfrac{C}{\nu^k}$ for $\lim_{\nu \to \infty}$. Note

$$f \in O(1/\nu^2) \implies \nu \cdot f \in O(1/\nu),$$

in particular

$$\lim_{\nu \to \infty} \nu \cdot f = 0$$

Proof. We conside the Taylor series

$$exp\ \frac{A}{\nu} = \mathbb{1} + \frac{A}{\nu} + O(1/\nu^2)$$

$$exp\ \frac{B}{\nu} = \mathbb{1} + \frac{B}{\nu} + O(1/\nu^2)$$

and

$$exp\ \frac{A}{\nu} \cdot exp\ \frac{B}{\nu} = \mathbb{1} + \frac{A}{\nu} + \frac{B}{\nu} + O(1/\nu^2).$$

For large $\nu \in \mathbb{N}$ we have

$$\left\| \left(exp\ \frac{A}{\nu} \right) \cdot \left(exp\ \frac{B}{\nu} \right) - \mathbb{1} \right\| < log\ 2.$$

Therefore the logarithm is well-defined:

$$log \left(exp\ \frac{A}{\nu} \cdot exp\ \frac{B}{\nu} \right) = log \left(\mathbb{1} + \frac{A}{\nu} + \frac{B}{\nu} + o(1/\nu^2) \right) = \frac{A}{\nu} + \frac{B}{\nu} + O(1/\nu^2).$$

- Then on one hand, applying *exp* to the last equality

$$(exp \circ log)\left(exp\ \frac{A}{\nu} \cdot exp\ \frac{B}{\nu}\right) = exp\left(\frac{A}{\nu} + \frac{B}{\nu} + O(1/\nu^2)\right).$$

- While on the other hand, Proposition 1.27 implies

$$(exp \circ log)\left(exp\ \frac{A}{\nu} \cdot exp\ \frac{B}{\nu}\right) = exp\ \frac{A}{\nu} \cdot exp\ \frac{B}{\nu}.$$

We get

$$exp\ \frac{A}{\nu} \cdot exp\ \frac{B}{\nu} = exp\left(\frac{A}{\nu} + \frac{B}{\nu} + O(1/\nu^2)\right)$$

and

$$\left(exp\ \frac{A}{\nu} \cdot exp\ \frac{B}{\nu}\right)^{\nu} = \left(exp\left(\frac{A}{\nu} + \frac{B}{\nu} + o(1/\nu^2)\right)\right)^{\nu} = exp(A + B + O(1/\nu)).$$

Here the last equality follows by expanding the exponential on both sides up to linear terms, and expanding on the left-hand side the ν-th power. Now for $\lim_{\nu\to\infty}$ the continuity of the exponential proves the claim

$$\lim_{\nu\to\infty}\left(exp\ \frac{A}{\nu} \cdot exp\ \frac{B}{\nu}\right)^{\nu} = exp(A+B)$$

□

How does the functional equation of the exponential map from the commutative case of Theorem 1.23 generalize to the non-commutative case?

Example 1.32 (Counter example against the expected functional equation). We consider two strictly upper triangular matrices of the form

$$P = \begin{pmatrix} 0 & p & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & q \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{n}(3, \mathbb{K}),\ p, q \neq 0.$$

Note that

$$[P, Q] = \begin{pmatrix} 0 & 0 & pq \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$$

We compute

$$exp(P) := \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \cdot P^{\nu} = \mathbb{1} + P = \begin{pmatrix} 1 & p & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$exp(Q) := \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \cdot Q^\nu = \mathbb{1} + Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & q \\ 0 & 0 & 1 \end{pmatrix}.$$

Note: For all $\nu \geq 2$

$$P^\nu = Q^\nu = 0$$

Hence we do not need to care about questions of convergence.

i) *Failure of the expected functional equation*: To test the functional equation we compute on one hand

$$exp(P) \cdot exp(Q) = \begin{pmatrix} 1 & p & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & q \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & p & pq \\ 0 & 1 & q \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{1} + P + Q + PQ.$$

On the other hand we compute

$$exp(P+Q).$$

For the computation of matrix products it is often helpful to introduce the basis $(E_{ij})_{1 \leq i,j \leq n}$ of the $\mathbb{K}$-vector space $M(n \times n, \mathbb{K})$ with the matrices

$$E_{ij} \in M(n \times n, \mathbb{K})$$

having value $= 1$ at the position with index ij and value $= 0$ at all other positions. Then

$$E_{ij} \cdot E_{km} = \delta_{jk} \cdot E_{im}$$

We have

$$P = p \cdot E_{12} \text{ and } Q = q \cdot E_{23}$$

Then

$$(P+Q)^2 = (pE_{12} + qE_{23}) \cdot (pE_{12} + qE_{23}) = pqE_{13} = PQ$$

and

$$(P+Q)^\nu = 0, \ \nu \geq 3,$$

which implies

$$exp(P+Q) = \mathbb{1} + P + Q + (1/2) \cdot (P+Q)^2 = \mathbb{1} + P + Q + (1/2) \cdot PQ.$$

As a consequence

$$exp(P+Q) \neq exp(P) \cdot exp(Q)$$

and the expected functional equation is not satisfied.

ii) *Correction term*: Fortunately, the defect can be fixed by introducing a correcting term C. Because the expected functional equation holds for arbitrary commuting matrices according to Theorem 1.23, the most simple ansatz for the correction term is

$$C := \alpha \cdot [P,Q] \text{ with suitable } \alpha \in \mathbb{K}$$

and the commutator

$$[P,Q] := PQ - QP \in \mathfrak{n}(3,\mathbb{K}).$$

With

$$PQ = pq \cdot E_{13} \text{ and } QP = q \cdot E_{23} \cdot E_{13} = 0$$

we obtain

$$PQ = [P,Q] = \frac{1}{\alpha} \cdot C, \ [P,Q] = pq \cdot E_{13}, \ \text{ and } C = \alpha \cdot pq \cdot E_{13} = \alpha \cdot [P,Q]$$

$$(P+Q+C)^2 = (pE_{12}+qE_{23}+(\alpha \cdot pq) \cdot E_{13}) \cdot (pE_{12}+qE_{23}+(\alpha \cdot pq) \cdot E_{13}) = pq \cdot E_{13}$$

and

$$(P+Q+C)^\nu = 0, \ \nu \geq 3.$$

Hence

$$exp(P+Q+C) = exp(P) \cdot exp(Q)$$

$$\iff$$

$$\mathbb{1} + (P+Q+C) + (1/2)(P+Q+C)^2 = \mathbb{1} + P + Q + C \cdot (1 + \frac{1}{2\alpha}) = \mathbb{1} + P + Q + \frac{C}{\alpha}$$

$$\iff$$

$$\alpha = 1/2.$$

The correction term is

$$C = \frac{1}{2} \cdot [P,Q],$$

proportional to the commutator:

$$exp\left(P + Q + \frac{1}{2} \cdot [P,Q]\right) = exp(P) \cdot exp(Q)$$

iii) *Generalization*: More general one can show: The exponential

$$exp : \mathfrak{n}(3,\mathbb{K}) \to GL(3,\mathbb{K}).$$

satisfies the functional equation in the generalized form: For $A, \ B \in \mathfrak{n}(3 \times 3, \mathbb{K})$

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

The exponential map of matrices

$$exp : M(n \times n, \mathbb{K}) \to GL(n, \mathbb{K})$$

is a well-defined map according to Corollary 1.24. But the map is not a group morphism

$$exp : (M(n \times n, \mathbb{K}), +) \to (GL(n, \mathbb{K}), \cdot)$$

as the counter example from Example 1.32, part i) shows. Instead, the functional equation of the exponential map depends in a certain way on the commutator of the matrices in question. The correct statement of the functional equation in the case of two arbitrary matrices is the special case of a deep theorem from general Lie group theory, see also [47]:

Remark 1.33 (Baker-Campbell-Hausdorff formula for matrices). There exists a sequence of polynomials in two matrix-valued indeterminates X and Y

$$H_\nu(X,Y)_{\nu \in \mathbb{N}^*}$$

with values in $M(n \times n, \mathbb{K})$ and homogeneous with respect to commutators of degree ν, and an open zero-neighbourhood

$$U \subset M(n \times n, \mathbb{K}),$$

which together satisfy the following properties:

- The *Hausdorff polynomials* of low order are

$$\begin{aligned}
H_1(X,Y) &= X + Y \\
H_2(X,Y) &= (1/2)[X,Y] \\
H_3(X,Y) &= (1/12)[[X,Y],Y] - [[X,Y],X]) \\
H_4(X,Y) &= -(1/24)[Y,[X,[X,Y]]]
\end{aligned}$$

- The Baker-Campbell-Hausdorff series

$$H(X,Y) := \sum_{\nu=1}^{\infty} H_\nu(X,Y)$$

  is absolute and compact convergent in $U \times U$. It satisfies for all $X, Y \in U$ the functional equation

$$exp\, X \cdot exp\, Y = exp(H(X,Y))$$

Chapter 2
Fundamentals of Lie algebra theory

In this chapter the base field is either $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

2.1 Definitions and first examples

In general, in the associative algebra $M(n \times n, \mathbb{K})$ the product of two matrices is not commutative:

$$AB \neq BA, \ A, B \in M(n \times n, \mathbb{K}).$$

The commutator

$$[A, B] := AB - BA \in M(n \times n, \mathbb{K})$$

measures the degree of non-commutativity. The commutator depends $\mathbb{K}$-bilinearly on the matrices A and B. In addition, it satisfies the following rules:

1. *Permutation of two matrices*: $[A, B] + [B, A] = 0$.
2. *Cyclic permutation of three matrices*: $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$.

These properties make

$$gl(n, \mathbb{K}) := (M(n \times n, \mathbb{K}), [-, -])$$

the prototype of a Lie algebra.

Definition 2.1 (Lie algebra).

1. A $\mathbb{K}$-*Lie algebra* is a $\mathbb{K}$-vector space L together with a $\mathbb{K}$-bilinear map

$$[-, -] : L \times L \to L \ (\textit{Lie bracket})$$

such that

- $[x, x] = 0$ for all $x \in L$ and

- $[x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0$ for all $x,y,z \in L$ (*Jacobi identity*).

2. A *morphism of* $\mathbb{K}$-*Lie algebras* is a $\mathbb{K}$-linear map $f: L_1 \to L_2$ between two $\mathbb{K}$-Lie algebras satisfying

$$f([x_1,x_2]) = [(f(x_1),f(x_2))],\ x_1,x_2 \in L_1.$$

Note that the Lie bracket satisfies

$$[x,y]+[y,x]=0 \ for\ all\ x,y \in L.$$

For the proof one computes $[x+y,x+y] \in L$.

As a consequence

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

We have just seen that the Lie algebra $gl(n,\mathbb{K})$ derives from the associative algebra $M(n\times n,\mathbb{K})$. Actually, any finite dimensional $\mathbb{K}$-Lie algebra derives from a matrix algebra. But the theory becomes more transparent when considering Lie algebras and their morphisms as abstract mathematical objects.

A Lie algebra is a vector space with the Lie bracket as additional operator. This additional algebraic structure resembles the multiplication within a group or a ring. The Lie algebra concept of the commutator is taken from group theory while the concept of an ideal comes from ring theory.

Definition 2.2 (Basic algebraic concepts). Consider a $\mathbb{K}$-Lie algebra L. One defines:

- A vector subspace $M \subset L$ is a *Lie subalgebra* of L iff

$$[M,M] \subset M,$$

i.e. iff M is closed with respect to the Lie bracket of L.

- A vector subspace $I \subset L$ is an *ideal* of L iff

$$[L,I] \subset I,$$

i.e. if I is L-invariant.

- The *normalizer* of a subalgebra $M \subset L$ is the subalgebra

$$N_L(M) := \{x \in L : [x,M] \subset M\} \subset L,$$

i.e. the largest subalgebra of L which includes M as an ideal.

- The *center* of L is the ideal

$$Z(L) := \{x \in L : [x,L] = 0\}.$$

The center of L collects those elements from L which commute with all elements from L.

- The *centralizer* $C_L(Y)$ of a subset $Y \subset L$ is the largest subalgebra of L which commutes with all elements from Y

$$C_L(Y) := \{x \in L : [x,Y] = 0\}$$

- The *derived algebra* or *commutator algebra* DL of L is the subalgebra generated by all commutators

$$DL := [L,L] := span_{\mathbb{K}}\{[x,y] : x,y \in L\}.$$

Iff $[L,L] = 0$ then L is named *Abelian* because the condition is equivalent to

$$[x,y] = [y,x] \text{ for all } x,y \in L.$$

For a morphism
$$f : L_1 \to L_2$$
between to $\mathbb{K}$-Lie algebras the kernel
$$ker(f) \in L_1$$
is an ideal. Apparently, for each Lie algebra L the derived algebra $[L,L] \subset L$ is an ideal.

If one compares these basic concepts from Lie algebra theory with concepts from group theory then subalgebras correspond to subgroups, while ideals are an analogue to normal subgroups.

Definition 2.3 (Specific matrix Lie algebras).

1. For a finite dimensional $\mathbb{K}$-vector space V we define the $\mathbb{K}$-Lie algebra

$$gl(V) := (End(V), [-,-])$$

with
$$[f,g] := f \circ g - g \circ f.$$

In particular
$$gl(n,\mathbb{K}) := gl(\mathbb{K}^n).$$

Lie subalgebras of the Lie algebra $gl(n,\mathbb{K})$ are named *matrix Lie algebras* or *embedded* Lie algebras.

2. The subalgebra of $gl(n,\mathbb{K})$ of *strictly upper triangular* matrices is

$$\mathfrak{n}(n,\mathbb{K}) := \{A = (a_{ij}) \in gl(n,\mathbb{K}) : a_{ij} = 0 \text{ if } i \geq j\}.$$

Each strictly upper triangular matrix has the form

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

3. The subalgebra of $gl(n,\mathbb{K})$ of *upper triangular* matrices is

$$\mathfrak{t}(n,\mathbb{K}) := \{A = (a_{ij}) \in gl(n,\mathbb{K}) : a_{ij} = 0 \text{ if } i > j\}.$$

Each upper triangular matrix has the form

$$\begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix}$$

4. The subalgebra of $gl(n,\mathbb{K})$ of *diagonal* matrices is

$$\mathfrak{d}(n,\mathbb{K}) := \{A = (a_{ij}) \in gl(n,\mathbb{K}) : a_{ij} = 0 \text{ if } i \neq j\}.$$

Each diagonal matrix has the form

$$\begin{pmatrix} * & & 0 \\ & \ddots & \\ 0 & & * \end{pmatrix}$$

Moreover for $n \geq 2$

$$\mathfrak{n}(n,\mathbb{K}) \subsetneq \mathfrak{t}(n,\mathbb{K}) \subsetneq gl(n,\mathbb{K})$$

and

$$\mathfrak{d}(n,\mathbb{K}) \subsetneq \mathfrak{t}(n,\mathbb{K}).$$

We will see in Chapter 3.1 and Chapter 3.2 that $\mathfrak{n}(n,\mathbb{K})$ and $\mathfrak{t}(n,\mathbb{K})$ are the prototype of the important classes of respectively nilpotent and solvable Lie algebras.

The fundamental tool for studying Lie algebras are their *representations*. To represent an abstract Lie algebra L means to define a map from L to a matrix Lie algebra. The concept of a representation has also many applications in physics. Some examples we will see in later chapters. The scope of the concept of a representation is not restricted to Lie algebra theory.

Definition 2.4 (Representation of a Lie algebra). Consider a $\mathbb{K}$-Lie algebra L.

1. A *representation* of L on a finite dimensional $\mathbb{K}$-vector space V is a morphism of Lie algebras
$$\rho : L \to gl(V).$$
In particular,
$$\rho([x,y]) = [\rho(x), \rho(y)] = \rho(x) \circ \rho(y) - \rho(y) \circ \rho(x).$$
The vector space V is named an *L-module* with respect to the multiplication
$$L \times V \to V, (x, \mathrm{v}) \mapsto x.\mathrm{v} := \rho(x)(\mathrm{v}).$$
It satisfies
$$[x,y].\mathrm{v} = \rho([x,y])(\mathrm{v}) = [\rho(x), \rho(y)](\mathrm{v}) = \rho(x)(\rho(y)(\mathrm{v})) - \rho(y)(\rho(x)(\mathrm{v})) =$$
$$x.(y.\mathrm{v}) - y.(x.\mathrm{v}).$$
The representation ρ is *faithful* iff ρ is injective, i.e. ρ embeds L into a Lie algebra of matrices.

2. A linear map
$$f : V \to W$$
between two L-modules is a *morphism* of L-modules if for all $x \in L,\ \mathrm{v} \in V$:
$$f(x.\mathrm{v}) = x.f(\mathrm{v}).$$

3. The *adjoint representation* of L is the map
$$ad : L \to gl(L),\ x \mapsto ad\ x,$$
defined as
$$ad\ x : L \to L,\ y \mapsto (ad\ x)(y) := [x,y].$$

Note. One often speaks about an L-module V by surpressing the name of the defining map of the representation
$$\rho : L \to gl(V).$$

Lemma 2.5 (Adjoint representation of a Lie algebra). *The adjoint representation is a representation, i.e. a morphism of Lie algebras.*

Proof. For a Lie algebra L we have to show for all $x, y \in L$

$$ad\ [x,y] = [ad\ x, ad\ y] : L \to L.$$

Consider $z \in L$. On one hand,

$$ad[x,y](z) = [[x,y],z] = -[[y,z],x] - [[z,x],y]\ (Jacobi\ identity)$$

On the other hand

$$[ad\ x, ad\ y](z) := (ad\ x \circ ad\ y - ad\ y \circ ad\ x)(z) = [x,[y,z]] - [y,[x,z]].$$

Both results are equal because the Lie bracket is antisymmetric. □

The adjoint representation maps any abstract Lie algebra to a matrix algebra. But in general the adjoint representation is not injective. The kernel of the adjoint representation of a Lie algebra L is the center $Z(L)$. While the adjoint representation is not always faithful, it is theorem of Ado, [4, Chap. I, §7.3, Theor. 3], that each Lie algebra has a faithful representation, i.e. each Lie algebra is a matrix algebra.

The characteristic feature of a Lie algebra L is the Lie bracket. It refines the underlying vector space of L. In order to investigate the Lie bracket of L one studies how each given element $x \in L$ acts on L as the endomorphism $ad\ x$.

This procedure is similar to the study of number fields $\mathbb{Q} \subset K \subset \mathbb{C}$. The multiplication on K refines the underlying $\mathbb{Q}$-vector space structure. And one studies the multiplication by considering the $\mathbb{Q}$-endomorphisms which result from the multiplication by all elements $x \in K$. Norm and trace of these endomorphisms are important concepts in algebraic number theory.

For all $x \in L$ the element $ad\ x$ is not only an endomorphism of L but also a *derivation* of L: With respect to the Lie bracket it satisfies a rule similar to the Leibniz rule for the derivation of the product of two functions.

Definition 2.6 (Derivation of a Lie algebra). Let L be a Lie algebra. A *derivation* of L is an endomorphism

$$D : L \to L$$

which satisfies the product rule

$$D([y,z]) = [D(y), z] + [y, D(z)].$$

Lemma 2.7 (Adjoint representation and derivation). *Consider a Lie algebra L. For every $x \in L$ the endomorphism*

$$ad\ x : L \to L$$

is a derivation of L.

Proof. Set $D := ad\ x \in End(L)$. The Jacobi identity implies

$$D([y,z]) := (ad\ x)([y,z]) = [x,[y,z]] = -[y,[z,x]] - [z,[x,y]] = [y,[x,z]] + [[x,y],z] =$$

$$= [y,(ad\ x)(z)] + [(ad\ x)(y),z] = [y,D(z)] + [D(y),z].$$

□

Derivations arising from the adjoint representation are named *inner* derivations, all other derivations are *outer* derivations.

Lemma 2.8 (Algebra of derivations). *Let L be a Lie algebra. The set of all derivations of L*

$$Der(L) := \{D \in End(L) : D\ derivation\}$$

is a subalgebra $Der(L) \subset gl(L)$.

Proof. Apparently $Der(L) \subset End(L)$ is a subspace. In order to show that $Der(L)$ is even a subalgebra, we have to prove: If $D_1, D_2 \in Der(L)$ then $[D_1, D_2] \in Der(L)$.

$$[D_1,D_2]([x,y]) = (D_1 \circ D_2)([x,y]) - (D_2 \circ D_1)([x,y]) =$$

$$= D_1([D_2(x),y] + [x,D_2(y)]) - D_2([D_1(x),y] + [x,D_1(y)]) =$$

$$= [D_1(D_2(x)),y] + [D_2(x),D_1(y)] + [D_1(x),D_2(y)] + [x,D_1(D_2(y)]$$

$$-[D_2(D_1(x)),y] - [D_1(x),D_2(y)] - [D_2(x),D_1(y)] - [x,D_2(D_1(y)] =$$

$$= [[D_1,D_2](x),y] + [x,[D_1,D_2](y)].$$

Hence the commutator of two derivations satisfies the product rule, i.e. is again a derivation. □

2.2 Lie algebras of the classical groups

An important class of Lie algebras are the Lie algebras attached to the classical groups. The classical groups are groups of matrices, and the corresponding Lie algebras are the Lie algebras of the infinitesimal generators of their 1-parameter subgroups.

Definition 2.9 (1-parameter subgroup and infinitesimal generator). For a given matrix $X \in M(n \times n, \mathbb{K})$ the differentiable group morphism

$$f_X : (\mathbb{R}, +) \to (GL(n, \mathbb{K}), \cdot),\ t \mapsto exp(t \cdot X),$$

with derivation

$$\left(\frac{d}{dt} exp(t \cdot X)\right)(0) = X$$

according to Proposition 1.22 is named the 1-*parameter subgroup of* $GL(n, \mathbb{K})$ *with infinitesimal generator* X.

Definition 2.10 (Matrix group and 1-parameter subgroups).

1. A *matrix group* G is a closed subgroup

$$G \subset GL(n, \mathbb{K}).$$

   Here $GL(n, \mathbb{K}) \subset \mathbb{K}^{(n^2)}$ is equipped with the subspace topology of the Euclidean space.

2. Consider a matrix group G. For a matrix $X \in M(n \times n, \mathbb{K})$ the group morphism

$$f_X : \mathbb{R} \to GL(n, \mathbb{K}),\ t \mapsto exp(t \cdot X),$$

   is a 1-*parameter* subgroup of G iff for all $t \in \mathbb{R}$

$$f_X(t) \in G.$$

Note that the definition of 1-parameter subgroups refers to real parameters t.
A 1-parameter subgroup of G with infinitesimal generator X is a differentiable curve in G which passes through the unit element $e \in G$ with tangent vector X.

Remark 2.11 (Matrix group and 1*-parameter subgroups).*

1. The term *closed* subgroup of $G \subset GL(n, \mathbb{K})$ in Definition 2.10 refers to the topology of G which is induced as a subset of $GL(n, \mathbb{K})$. A subgroup

$$G \subset GL(n, \mathbb{K})$$

   is closed iff for any sequence $(A_\nu)_{\nu \in \mathbb{N}}$ of matrices $A_\nu \in G, \nu \in \mathbb{N}$, which converges in $GL(n, \mathbb{K})$, also the limit belongs to G, i.e.

$$A = \lim_{\nu \to \infty} A_\nu \in GL(n, \mathbb{K}) \implies A \in G.$$

2. For a real Lie group G each closed subgroup $H \subset G$ has a unique real Lie group structure such that the injection $H \hookrightarrow G$ becomes an embedding of real Lie groups. Therefore matrix groups according to Definition 2.10 are Lie groups.

   Note that closed subgroups of a complex Lie group are not necessarily complex Lie groups; a simple counter-example are the real Lie groups
$$SU(n) \subset GL(n, \mathbb{C}).$$

3. The general definition of a 1-parameter subgroup of an arbitrary Lie group G requires only a *continuous* group morphism
$$f : \mathbb{R} \to G.$$
   But one can show: All continuous 1-parameter subgroups of a Lie group G have the form
$$f(t) = exp(t \cdot X)$$
   with an element $X \in Lie\ G$, the Lie algebra of G, and with respect to the exponential map
$$exp : Lie\ G \to G.$$
   In particular, every continuous 1-parameter subgroup depends on the parameter t in a differentiable - even analytic - manner.

Notation 2.12 (Restricting scalars from $\mathbb{C}$ to $\mathbb{R}$).

For a complex vector space V there is a method of restricting scalars: The real vector space $V_{\mathbb{R}}$ has the same elements as V, but the elements are multiplied only by scalars from $\mathbb{R}$.

An analogous notation applies to a complex Lie algebra L. By restricting scalars from $\mathbb{C}$ to $\mathbb{R}$ the elements of L form a real Lie algebra which is denoted $L_{\mathbb{R}}$. The Lie algebra $L_{\mathbb{R}}$ has the same elements as L, but they are multiplied only by scalars from $\mathbb{R}$.

Be aware that the notation is not standard in the literature.

The infinitesimal generators of all 1-parameter subgroups of a matrix group form a real Lie algebra. Recall Notation 2.12.

Proposition 2.13 (Infinitesimal generators of a matrix group). *For a matrix group*

$$G \subset GL(n,\mathbb{C})$$

the set of infinitesimal generators of all 1*-parameter subgroups of* G

$$L := \{X \in M(n \times n, \mathbb{C}) : f_X(t) \in G \textit{ for all } t \in \mathbb{R}\}$$

with the functions f_X *from Definition 2.10, is a real Lie-subalgebra of* $gl(n,\mathbb{C})_{\mathbb{R}}$.

Proof. i) *Scalar multiplication*: For $X \in L$ and $s \in \mathbb{R}$ also

$$s \cdot X \in L$$

because for all $t \in \mathbb{R}$

$$exp(t \cdot (sX)) = exp((t \cdot s)X) \in G.$$

ii) *Additivity*: If $X,\ Y \in L$ then Proposition 1.31 implies for all $t \in \mathbb{R}$

$$exp(t(X+Y)) = \lim_{\nu \to \infty} \left(exp\ \frac{tX}{\nu} \cdot exp\ \frac{tY}{\nu} \right)^{\nu}.$$

The closedness of $G \subset GL(n,\mathbb{C})$ implies

$$exp(t(X+Y)) \in G.$$

iii) *Lie bracket*: First, for all $X \in L$ and all $A \in G$ also the conjugate

$$AXA^{-1} \in L:$$

Proposition 1.6 implies for all $t \in \mathbb{R}$

$$exp(t(AXA^{-1})) = exp(A(tX)A^{-1}) = A \cdot exp(tX) \cdot A^{-1} \in G.$$

Secondly, consider for arbitrary fixed $X,\ Y \in L$ the differentiable map

$$f : \mathbb{R} \to M(n \times n, \mathbb{C}),\ f(t) := exp(tX) \cdot Y \cdot exp(-tX).$$

Due to the first step

$$f(t) \in L \text{ for all } t \in \mathbb{R}.$$

The chain rule and the product rule from Proposition 1.8 imply:

$$\frac{d}{dt} f(t) = \frac{d}{dt}(exp(tX) \cdot Y \cdot exp(-tX)) =$$

$$= X \cdot exp(tX) \cdot Y \cdot exp(-tX) - exp(tX) \cdot Y \cdot X \cdot exp(-tX)$$

Hence

$$\frac{df}{dt}(0) = XY - YX = [X,Y].$$

On the other hand

$$\frac{df}{dt}(0) = \lim_{t \to 0} \frac{f(t) - f(0)}{t}$$

Because each fraction belongs to L, which is a subspace of the finite dimensional vector space $M(n \times n, \mathbb{C})$ due to part i) and ii), and is therefore closed. In the limit we get

$$[X,Y] = \frac{df}{dt}(0) \in L.$$

□

According to Proposition 2.13 the infinitesimal generators of all 1-parameter subgroups of a matrix group form a Lie algebra.

Definition 2.14 (Lie algebra of a matrix group). Consider a matrix group G. The Lie algebra of all infinitesimal generators of 1-parameter subgroups of G is named *Lie G*, the *Lie algebra* of G.

If for each $X \in$ *Lie G* also $i \cdot X \in$ *Lie G* then G is named a *complex* matrix group.

Not each matrix group $G \subset GL(n, \mathbb{C})$ which contains non-real, complex matrices is a complex matrix group. Counter examples are the unitary groups $U(m)$. Whether a matrix group is a *complex* matrix group depends on the possibility to multiply the elements of its Lie algebra by the imaginary unit $i \in \mathbb{C}$. It does not depend on the question whether the entries of the matrices are complex numbers.

Proposition 2.15 (The Lie algebras of the classical groups). *Consider a parameter $r \in \mathbb{N}$ and a number $m = m(r) \in \mathbb{N}$. We use the shorthand*

$$G := (GL(m, \mathbb{K}), \cdot)$$

for the matrix group and

$$L := (gl(m, \mathbb{K}), [-,-])$$

for the Lie algebra Lie G. Then the classical groups and their Lie algebras are:

i) Series $A_r,\ r \geq 1,\ m := r+1$*: The* special linear *group*

$$SL(m, \mathbb{K}) := \{g \in G : det\ g = 1\} \subset G$$

has the $\mathbb{K}$-Lie algebra

$$sl(m, \mathbb{K}) := \{X \in L : tr\, X = 0\}$$

of traceless *matrices* .

ii) Series $B_r,\ r \geq 2,\ m := 2r+1$*: The* special orthogonal *group*

$$SO(m,\mathbb{K}) := \{g \in G : g \cdot g^\top = \mathbb{1},\ det\ g = 1\}$$

has the $\mathbb{K}$*-Lie algebra*

$$so(m,\mathbb{K}) := \{X \in L : X + X^\top = 0\}$$

of skew-symmetric *matrices* .

iii) Series $C_r,\ r \geq 3,\ m = 2r$*: The* symplectic *group*

$$Sp(m,\mathbb{K}) := \{g \in G : g^\top \cdot \sigma \cdot g = \sigma\}$$

with

$$\sigma := \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix},\ \sigma^{-1} = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} = -\sigma,\ \mathbb{1} \in GL(r,\mathbb{K})\ unit\ matrix,$$

has the $\mathbb{K}$*-Lie algebra*

$$sp(m,\mathbb{K}) := \{X \in L : X^\top \cdot \sigma + \sigma \cdot X = 0\},$$

named the symplectic *algebra.*

Note that any $g \in Sp(m,\mathbb{K})$ *has* $det\ g = 1$*, see [38]. Moreover, each* $X \in sp(m,\mathbb{K})$ *has* $tr\ X = 0$*, see [24, Chap. 1.2].*

iv) Series $D_r,\ r \geq 4,\ m = 2r$*: The* special orthogonal *group*

$$SO(m,\mathbb{K}) := \{g \in G : g \cdot g^\top = \mathbb{1}, det\ g = 1\}$$

has the $\mathbb{K}$*-Lie algebra*

$$so(m,\mathbb{K}) := \{X \in L : X + X^\top = 0\}$$

of skew-symmetric *matrices* .

v) Special unitary group*: For each* $m \in \mathbb{N}$ *the* special unitary *group*

$$SU(m) := \{g \in GL(m,\mathbb{C}) : g \cdot g^* = \mathbb{1},\ det\ g = 1\},\ g^* := \overline{g}^\top\ Hermitian\ conjugate,$$

has the real Lie algebra

$$su(m) := \{X \in L : X + X^* = 0, tr\ X = 0\}$$

of traceless skew-Hermitian *matrices* .

Proof (of Proposition 2.15).

Proposition 2.13 ensures that the infinitesimal generators of all these matrix groups form a Lie algebra. Hence we have to show

- Each infinitesimal generator of a 1-parameter subgroup of the matrix group in question belongs to the defined “small-letter”-set.
- Each element of the defined “small-letter” set generates a 1-parameter subgroup of the matrix group in question.

i) Consider an infinitesimal generator $X \in gl(m,\mathbb{K})$ of a 1-parameter subgroup of $SL(m,\mathbb{K})$. Taking the derivative $\dfrac{d}{dt}$ at $t = 0$ on both sides of the equation

$$1 = det(exp(t \cdot X)) = exp(tr\ (t \cdot X))$$

gives

$$0 = (tr\, X) \cdot exp(tr\ (t \cdot X)).$$

Hence $tr\, X = 0$ because the exponential function has no zeros.

In the opposite direction: If $tr\, X = 0$ then

$$det\ e^{tX} = e^{tr\ (tX)} = e^{t\cdot(tr\ X)} = e^0 = 1.$$

ii) and iv) Taking the derivative of

$$\mathbb{1} = exp(t \cdot X) \cdot exp(t \cdot X)^\top$$

and using the product rule gives

$$0 = X \cdot exp(t \cdot X) \cdot exp(t \cdot X^\top) + exp(t \cdot X) \cdot X^\top \cdot exp(t \cdot X^\top).$$

Hence for $t = 0$:

$$X + X^\top = 0.$$

In the opposite direction:

$$X + X^\top = 0$$

implies

$$exp(tX) \cdot exp(tX^\top) = exp(t(X + X^\top)) = e^0 = \mathbb{1}$$

because

$$[X, X^\top] = 0 \text{ due to } X^\top = -X.$$

And $tr\, X = 0$ implies

$$det(exp(tX)) = e^{tr\ X} = 1.$$

iii) Taking the derivative of

$$\sigma = exp(t \cdot X)^{\top} \cdot \sigma \cdot exp(t \cdot X)$$

gives

$$0 = X^{\top} \cdot exp(t \cdot X^{\top}) \cdot \sigma \cdot exp(t \cdot X) + exp(t \cdot X^{\top}) \cdot \sigma \cdot X \cdot exp(t \cdot X).$$

Hence for $t = 0$:

$$0 = X^{\top} \cdot \sigma + \sigma \cdot X.$$

In the opposite direction:

$$0 = X^{\top} + \sigma \cdot X \cdot \sigma^{-1}$$

implies

$$(exp(tX))^{\top} \cdot \sigma \cdot exp(tX) \cdot \sigma^{-1} = \mathbb{1}$$

because

$$[X^{\top}, \sigma \cdot X \cdot \sigma^{-1}] = [X^{\top}, -X^{\top}] = 0$$

v) Taking the derivative of

$$\mathbb{1} = exp(t \cdot X) \cdot exp(t \cdot X)^*$$

and using the product rule gives

$$0 = X \cdot exp(t \cdot X) \cdot exp(t \cdot X^*) + exp(t \cdot X) \cdot X^* \cdot exp(t \cdot X^*).$$

Hence for $t = 0$

$$X + X^* = 0.$$

For all $t \in \mathbb{R}$

$$1 = det\ exp(tX) = e^{t \cdot tr\ X} = 0$$

implies

$$t \cdot tr\ X \in \mathbb{Z}2\pi i.$$

Therefore

$$tr\ X = 0.$$

In the opposite direction:

$$X + X^* = 0$$

implies

$$[X, X^*] = -[X, X] = 0$$

and

$$exp(tX) \cdot (exp(tX))^* = exp(tX) \cdot exp(tX^*) = exp(t(X + X^*)) = \mathbb{1}.$$

And

$$tr\ X = 0 \implies det(exp\ X) = e^{tr\ X} = 1.$$

□

Note that $su(m)$ is not a complex Lie algebra:

$$X \in su(2) \iff X = \begin{pmatrix} ia & b \\ -\overline{b} & -ia \end{pmatrix},\ a \in \mathbb{R},\ b \in \mathbb{C}.$$

If $X \in su(2)$ and $X \neq 0$ then $iX \notin su(2)$. Therefore SU(2) and more general the groups SU(n) are examples of real matrix groups which are not complex matrix groups.

- Elements of the complex matrix groups of the series $B_r,\ D_r$ preserve the $\mathbb{K}$-bilinear form on $\mathbb{K}^m$
$$(z,w) = \sum_{j=1}^{m} z_j \cdot w_j.$$
- Elements of the complex matrix groups of the series C_r preserve the $\mathbb{K}$-bilinear form on $\mathbb{K}^m$
$$(x,y) = \sum_{i=1}^{r} x_i \cdot y_{r+i} - x_{r+i} \cdot y_i.$$
- Elements of the special unitary group $SU(m)$ preserve the sesquilinear (Hermitian) form on $\mathbb{C}^m$
$$(z,w) = \sum_{j=1}^{m} z_j \cdot \overline{w}_j.$$

  The sesquilinear form (z,w) is $\mathbb{C}$-linear with respect to the first component and $\mathbb{C}$-antilinear with respect to the second component.

Remark 2.16 (Classical Lie groups of low dimension).

The reason for introducing the different series in Proposition 2.15 and for distinguishing the two series B_r and D_r of special orthogonal groups will become clear later in Chapter 6.

The lower bound for the parameter $r \in \mathbb{N}$ has been choosen to avoid duplicates or product decompositions. Otherwise we would have for the base field $\mathbb{K} = \mathbb{C}$ and the corresponding Lie algebras, see [40, Chap. II,7]:

$$A_1 = B_1 = C_1,\ B_2 = C_2,\ D_1 \text{ not simple},\ D_2 = A_1 \times A_1,\ D_3 = A_3$$

When the term *classical Lie group* is taken in a narrow sense, then it applies only to the complex matrix groups of the $A - D$-series.

Proposition 2.17 (Exponential map of SU(n)). *For each $n \in \mathbb{N}$ the exponential map*

$$su(n) \to SU(n)$$

is surjective.

For the proof cf. [12, Kor. 6.4.9].

Proof. i) *Assume* $A \in U(n)$: Each unitary matrix is diagonalizable. Hence there exists an invertible matrix $S \in U(n)$ with

$$S \cdot A \cdot S^* = diag(\lambda_1, ..., \lambda_n)$$

and for $j = 1, ..., n$

$$|\lambda_j| = 1,$$

hence

$$\lambda_j = e^{i\theta_j} \text{ with } \theta \in [0, 2\pi[$$

Set

$$B := S^* \cdot diag(i\theta_1, ..., i\theta_n) \cdot S \in u(n)$$

Then

$$exp(B) = S^* \cdot diag(e^{i\theta_1}, ..., e^{i\theta_n}) \cdot S = A$$

ii) *Assume* $A \in SU(n)$: The eigenvalues

$$\lambda_j = e^{i\theta_j},\ j = 1, ..., n$$

satisfy

$$\sum_{j=1}^{n} i\theta_j = k \cdot 2\pi i,\ k \in \mathbb{Z}.$$

Replace θ_n by

$$\theta'_n := -\sum_{j=1}^{n-1} \theta_j$$

Then

$$\theta'_n - \theta_n = -k \cdot 2\pi$$

which implies

$$e^{i\theta'_n} = e^{i\theta}$$

Set

$$B' := S^* \cdot diag(i\theta_1, ..., i\theta_{n-1}, i\theta'_n) \cdot S \in su(n)$$

Then

$$exp\ B' = exp\ B = A$$

□

2.3 Topology of the classical groups

The present sections investigates topological properties of some classical matrix groups. The study of matrix groups often reduces to the study of *simply connected* matrix groups and covering maps. The simply connected matrix groups can be studied by their Lie algebras. The study of covering projections is often the starting point for a course on algebraic topology. The following proofs about the matrix groups

$$SU(2),\ SO(3,\mathbb{R}),\ SL(2,\mathbb{C}),\ O(3,1),$$

the group $O(3,1)$ beeing the Lorentz group, are special cases from general Lie group theory.

Proposition 2.18 (Topology of $SO(3,\mathbb{R}),\ SU(2),\ SL(2,\mathbb{C})$).

1. *The matrix group $SO(3,\mathbb{R})$ is connected, the matrix group $O(3,\mathbb{R})$ has two connected components.*

2. *The matrix group $SU(2)$ - as a differentiable manifold - is diffeomorphic to the 3-sphere:*
$$SU(2) \simeq S^3 := \{x \in \mathbb{R}^4 : \|x\| = 1\}.$$
In particular, $SU(2)$ is simply connected.

3. *The matrix group $SL(2,\mathbb{C})$ - as a differentiable manifold - is diffeomorphic to the product*
$$(\mathbb{C}^2 \setminus \{0\}) \times \mathbb{C},$$
and the latter is homeomorphic to
$$S^3 \times \mathbb{R}^3.$$
In particular, $SL(2,\mathbb{C})$ is simply connected.

Proof. 1. *Matrix group $SO(3,\mathbb{R})$*: We show that the matrix group $SO(3,\mathbb{R})$ is path-connected: For a given rotation matrix A we choose an orthonormal basis of $\mathbb{R}^3$ with the rotation axis of A as third basis element. Then we may assume

$$A = \begin{pmatrix} \cos\delta & \sin\delta & 0 \\ -\sin\delta & \cos\delta & 0 \\ 0 & 0 & 1 \end{pmatrix} \in SO(3,\mathbb{R}), \delta \in [0,2\pi[.$$

The path

$$\gamma : [0,1] \to SO(3,\mathbb{R}), t \mapsto \begin{pmatrix} \cos(t\cdot\delta) & \sin(t\cdot\delta) & 0 \\ -\sin(t\cdot\delta) & \cos(t\cdot\delta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \in SO(3,\mathbb{R})$$

is a continuous map and connects within $SO(3,\mathbb{R})$ the start $\gamma(0) = \mathbb{1}$ the unit matrix with $\gamma(1) = A$.

Each matrix
$$A \in O(3,\mathbb{R}) := \{A \in GL(3,\mathbb{R}) : \ A \cdot A^{\top} = \mathbb{1}\}$$
has
$$det\ A = \pm 1.$$

Hence $O(3,\mathbb{R})$ has two connected components: $SO(3,\mathbb{R})$, the connected component of the identity, and the residue class
$$SO(3,\mathbb{R}) \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

As a consequence, both Lie groups $SO(3,\mathbb{R})$ and $O(3,\mathbb{R})$ have the same Lie algebra
$$so(3,\mathbb{R}) = Lie\ SO(3,\mathbb{R}) = Lie\ O(3,\mathbb{R}).$$

2. *Matrix group $SU(2)$*: Consider a matrix
$$A = \begin{pmatrix} a & z \\ c & w \end{pmatrix} \in SU(2).$$

Then
$$A^{-1} = \begin{pmatrix} w & -z \\ -c & a \end{pmatrix} \ and\ A^* = \begin{pmatrix} \overline{a} & \overline{c} \\ \overline{z} & \overline{w} \end{pmatrix}.$$

Hence
$$A^{-1} = A^*$$
implies
$$SU(2) := \{A \in GL(2,\mathbb{C}) : A \cdot A^* = \mathbb{1},\ det\ A = 1\} =$$
$$\left\{ \begin{pmatrix} \overline{w} & z \\ -\overline{z} & w \end{pmatrix} : z,\ w \in \mathbb{C},\ |z|^2 + |w|^2 = 1 \right\} \simeq$$
$$\simeq S^3 \subset \mathbb{C}^2 \simeq \mathbb{R}^4$$
is diffeomorphic to the 3-dimensional unit sphere. The unit sphere S^3 is compact, connected and simply connected.

Simply connected means the vanishing of the fundamental group
$$\pi_1(S^3, *) = 0$$
or equivalently: Any closed path in S^3 is contractible in S^3 to one point. Intuitively, simple connectedness means that S^3 has no “holes”.

The result $\pi_1(S^3,*)=0$ is a particular case of the Seifert-van Kampen theorem, see [19, Theor. 1.20], [44, Kap. 5.3]: One decomposes S^3 as the union of its northern and southern hemispheres, which are homeomorphic to the 3-dimensional closed solid ball B^3 and intersect each other in a space homeomorphic to S^2. Then $\pi_1(S^3,*)$ is the quotient of a free product:

$$\pi_1(S^3,*) = \pi_1(B^3,*) \cdot_{\pi_1(S^2,*)} \pi_1(B^3,*) = 0$$

because $\pi_1(B^3,*)=0$.

3. *Matrix group* $SL(2,\mathbb{C})$: Set $B := \mathbb{C}^2 \setminus \{0\}$.

i) First, projecting a matrix $A \in SL(2,\mathbb{C})$ onto its last column defines the differentiable map

$$p : SL(2,\mathbb{C}) \to B,\ A = \begin{pmatrix} a & z \\ c & w \end{pmatrix} \mapsto \begin{pmatrix} z \\ w \end{pmatrix}.$$

ii) Secondly, one expands p to a map

$$f : SL(2,\mathbb{C}) \to B \times \mathbb{C}$$

To define f one considers the open covering $\mathscr{U} = (U_1,U_2)$ of the base B with

$$U_1 := \{(z,w) \in B : z \neq 0\},\ U_2 := \{(z,w) \in B : w \neq 0\}$$

Then one defines

$$f : SL(2,\mathbb{C}) \to B \times \mathbb{C}$$

by distinction of cases

$$A = \begin{pmatrix} a & z \\ c & w \end{pmatrix} \mapsto \begin{cases} (p(A), (a - \overline{w} \cdot r^{-2}) \cdot z^{-1}) & \text{if } p(A) \in U_1 \\ (p(A), (c + \overline{z} \cdot r^{-2}) \cdot w^{-1}) & \text{if } p(A) \in U_2 \end{cases}$$

with

$$r^2 := |z|^2 + |w|^2.$$

If $p(A) \in U_1 \cap U_2$ then

$$(a - \overline{w} \cdot r^{-2}) \cdot z^{-1} = (c + \overline{z} \cdot r^{-2}) \cdot w^{-1},$$

after multiplying by $z \cdot w$ and expanding both sides, employing the determinant formula

$$aw - zc = 1$$

as well as the definition of r^2. Hence the map f is well-defined.

iii) The map f is a differentiable isomorphism satisfying $pr_1 \circ f = p$: The inverse

$$g : B \times \mathbb{C} \to SL(2,\mathbb{C})$$

is obtained as

$$g|(U_1 \times \mathbb{C}) : U_1 \times \mathbb{C} \to SL(2,\mathbb{C}),\ \left(\begin{pmatrix} z \\ w \end{pmatrix}, s\right) \mapsto \begin{pmatrix} a & z \\ c & w \end{pmatrix}$$

with

$$a := s \cdot z + \frac{\overline{w}}{r^2} \text{ and } c := \frac{aw-1}{z}$$

One checks

$$A := \begin{pmatrix} a & z \\ c & w \end{pmatrix} \in SL(2,\mathbb{C}).$$

A similar calculation determines the restriction

$$g|(U_2 \times \mathbb{C})$$

There results a diffeomorphism

$$SL(2,\mathbb{C}) \simeq B \times \mathbb{C}$$

iv) Concerning the homeomorphy type of the base B one has

$$B = \mathbb{C}^2 \setminus \{0\} \simeq S^3 \times]0,\infty[\simeq S^3 \times \mathbb{R}.$$

Because S^3 is simply connected according to part 1), also B and eventually

$$SL(2,\mathbb{C}) \simeq B \times \mathbb{C}$$

are simply connected. □

The groups from Proposition 2.18 serve as an example to demonstrate the method of investigation. For a systematic study see [23, Chap. 17] and [21, Chap. 10, §2]. For the classical groups of arbitrary dimensions one has the following topological results:

Remark 2.19 (More topological results about the classical groups).

1. *General linear group, $m \geq 1$:*

The complex group $GL(m,\mathbb{C})$ is connected due to Theorem 1.29 about the surjecitivity of the exponential map. Moreover,

$$\pi_1(GL(n,\mathbb{C}),*) = \mathbb{Z}.$$

The real group $GL(m,\mathbb{R})$ has two connected components. The component of the identity is not simply connected for $m \geq 2$.

2. *Series* $A_r, r \geq 1, m = r+1$:

For $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ the group $SL(m,\mathbb{K})$ is connected: For the proof one considers the homeomorphism

$$GL(m,\mathbb{K}) \xrightarrow{\simeq} SL(m,\mathbb{K}) \times K^*,\ X \mapsto (X',\ det\ X),$$

with X' obtained by diving the first column of X by $det\ X$. The projection onto the first factor

$$pr_1 : GL(m,\mathbb{K}) \to SL(m,\mathbb{K})$$

maps the connected $GL(m,\mathbb{K})$ onto $SL(m,\mathbb{K})$, which therefore is connected too.

The groups $SL(m,\mathbb{C})$ are simply connected. One has

$$\pi_1(SL(m,\mathbb{R}),*) = \begin{cases} \mathbb{Z} & m=2 \\ \mathbb{Z}/2 & m \geq 3 \end{cases}$$

3. *Series* $B_r, r \geq 2, m = 2r$, and *series* $D_r, r \geq 4, m = 2r+1$:

For $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ the group $SO(m,\mathbb{K})$ is connected. The groups $SO(m,\mathbb{R})$ are compact. One has

$$\pi_1(SO(m,\mathbb{R}),*) = \begin{cases} \mathbb{Z} & m=2 \\ \mathbb{Z}/2 & m \geq 3 \end{cases}$$

For $m=2$ the universal covering projection is

$$\mathbb{R} \to SO(2,\mathbb{R}),\ t \mapsto \begin{pmatrix} cos\ t & sin\ t \\ -sin\ t & cos\ t \end{pmatrix},$$

with

$$SO(2,\mathbb{R}) \simeq \mathbb{R}/\mathbb{Z}.$$

For $m \geq 3$ the universal covering of $SO(m,\mathbb{R})$ is $Spin(m,\mathbb{R})$, see [16, Chap. IV, Sect. V.4]. One has

$$Spin(3,\mathbb{R}) = SU(2),$$

cf. Example 2.24. Moreover one has the isomorphy of groups

$$Spin(4,\mathbb{R}) \simeq SU(2) \times SU(2).$$

4. *Series* $C_r : r \geq 3, m = 2r$:

For $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ the symplectic groups $Sp(m, \mathbb{K})$ are connected. The groups $Sp(m, \mathbb{C}))$ are simply connected, the real symplectic groups satisfy

$$\pi_1(Sp(m, \mathbb{R}), *) = \mathbb{Z}.$$

5. *Special unitary group, $m \geq 1$*:

   The group $SU(m)$ is compact, connected and simply connected.

No complex group from this list is compact, because each complex, connected and compact Lie group is Abelian, see [23, Prop. 15.3.7].

Several relations exists between the classical groups. These relations can be made explicit by differentiable group morphisms. Hence it is advantageous not to study each group in isolation. Instead, one should focus on the *relations* between different classical groups and study those properties, which the groups have in common. For two examples in low dimension see Example 2.24 and Proposition 2.27.

We now make some remarks about the relation between real and complex Lie algebras. Recall Notation 2.12 for restricting scalars from $\mathbb{C}$ to $\mathbb{R}$. Complexification is a method of scalar extension.

Definition 2.20 (Complexification of a real Lie algebra, real form of a complex Lie algebra).

1. Consider a real Lie algebra M. The *complexification* of M is the complex Lie algebra
$$M \otimes_{\mathbb{R}} \mathbb{C}$$
with Lie bracket
$$[m_1 \otimes z_1, m_2 \otimes z_2] := [m_1, m_2] \otimes (z_1 \cdot z_2),\ m_1, m_2 \in M,\ z_1, z_2 \in \mathbb{C}.$$
2. A *real form* of a complex Lie algebra L is a real subalgebra $M \subset L_{\mathbb{R}}$ such that the complex linear map from the complexification
$$M \otimes_{\mathbb{R}} \mathbb{C} \to L,\ m \otimes 1 \mapsto m,\ m \otimes i \mapsto i \cdot m,$$
is an isomorphism of complex Lie algebras.

Note that for a complex vector space V restricting and extending scalars are not inverse operations: The complex vector space

$$(V_{\mathbb{R}}) \otimes_{\mathbb{R}} \mathbb{C}$$

has complex dimension $2 \cdot dim_{\mathbb{C}} V$.

Remark 2.21 (Non-isomorphic real forms).

1. The complex Lie algebra $sl(2,\mathbb{C})$ has the non-isomorphic real forms

$$sl(2,\mathbb{R}) \text{ and } su(2).$$

The proof goes along the following steps:

i) *Real forms*: Apparently, $sl(2,\mathbb{R})$ is a real form of $sl(2,\mathbb{C})$.

Concerning $su(2)$, the $\mathbb{C}$-linear map

$$su(2) \otimes_{\mathbb{R}} \mathbb{C} \to sl(2,\mathbb{C}),\ A \otimes 1 \mapsto A,\ A \otimes i \mapsto iA,$$

has the inverse

$$sl(2,\mathbb{C}) \to su(2) \otimes_{\mathbb{R}} \mathbb{C},\ Z = X + iY \mapsto (-iX) \otimes i + iY \otimes 1,$$

with the Hermitian matrices

$$X := \frac{Z + Z^*}{2} \text{ (“real” part of Z) and } Y := \frac{Z - Z^*}{2i} \text{ (“imaginary” part of Z)},$$

and therefore

$$(-iX), iY \in su(2)$$

Hence both real Lie algebras $sl(2,\mathbb{R})$ and $su(2)$ are real forms of the complex Lie algebra $sl(2,\mathbb{C})$.

ii) *Existence of a* 2*-dimensional subalgebra of* $sl(2,\mathbb{R})$: The real Lie algebra $sl(2,\mathbb{R})$ contains the 2-dimensional subspace

$$span_{\mathbb{R}} < A := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} >,$$

which is a Lie subalgebra because

$$[A, B] = 2B.$$

iii) *Isomorphy* $su(2) \simeq so(3,\mathbb{R})$: Cf. [18, Example 3.27]. Using the traceless Hermitian *Pauli matrices*

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

we introduce the basis of the real Lie algebra $su(2)$ of skew-Hermitian traceless matrices

$$\left(E_1 := \frac{i}{2}\cdot\sigma_3,\ E_2 := \frac{i}{2}\cdot\sigma_1,\ E_3 := \frac{-i}{2}\cdot\sigma_2\right)$$

The basis elements have the commutators

$$[E_1, E_2] = E_3,$$

and all further non-zero commutators result from cyclic permutation, i.e. using the Levi-Civita symbol ε_{jkl}

$$[E_j, E_k] = \varepsilon_{jkl}\cdot E_l$$

To investigate the Lie algebra $so(3,\mathbb{R})$ we consider the infinitesimal generators of the standard 1-parameter subgroups of rotations around the coordinate axes of $\mathbb{R}^3$ by

$$X := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},\ Y := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix},\ Z := \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in so(3,\mathbb{R}).$$

One checks

$$[X, Y] = Z$$

with the standard behaviour for cyclic permutation. Then the $\mathbb{R}$-linear map

$$f : su(2) \to so(3,\mathbb{R}),$$

defined by

$$f(E_1) := X,\ f(E_2) := Y,\ f(E_3) := Z,$$

is an isomorphism of real Lie algebras.

iv) *No* 2*-dimensional subalgebra of* $so(3,\mathbb{R})$: The Lie algebra $so(3,\mathbb{R})$ is isomorphic to the Lie algebra *vector product*

$$Vect := (\mathbb{R}^3, \times).$$

For any two linear independent vectors $x_1, x_2 \in \mathbb{R}^3$ the product

$$x_1 \times x_2 \in \mathbb{R}^3$$

is orthogonal to the plane generated by x_1 and x_2. Hence

$$x_1 \times x_2 \notin span_{\mathbb{R}} < x_1, x_2 >,$$

which shows that

$$Vect \simeq so(3,\mathbb{R}) \simeq su(2)$$

does not contain a 2-dimensional Lie subalgebra.

As a consequence, the two real Lie algebras $sl(2,\mathbb{R})$ and $su(2)$ are not isomorphic.

2. *Compactness versus non-compactness*: The real Lie algebra $su(2)$ is the Lie algebra of the compact matrix group $SU(2)$. The real Lie algebra $sl(2,\mathbb{R})$ is the Lie algebra of the matrix group $SL(2,\mathbb{R})$ which is not compact.

3. *Arbitrary dimensions*: Also for general $n \in \mathbb{N}$ the real matrix group $su(n)$ is a real form of the complex Lie algebra $sl(n,\mathbb{C})$. And the matrix group $SU(n)$ is compact.

4. *Complex representations of real forms*: Consider a complex vector space V and a real Lie algebra M. Each $\mathbb{R}$-linear representation

$$\rho : M \to gl(V)$$

induces by complexification the $\mathbb{C}$-linear representation

$$\rho \otimes_{\mathbb{R}} \mathbb{C} : M \otimes_{\mathbb{R}} \mathbb{C} \to gl(V),\ m \otimes z \mapsto z \cdot \rho(m),$$

which renders commmutative the following diagram

$$\begin{tikzcd} M \arrow[r, "\rho"] \arrow[d] & gl(V) \\ M \otimes_{\mathbb{R}} \mathbb{C} \arrow[ur, dashed, "\rho \otimes_{\mathbb{R}} \mathbb{C}"'] & \end{tikzcd}$$

with the canonical map

$$M \to M \otimes_{\mathbb{R}} \mathbb{C},\ m \mapsto m \otimes z,$$

in the vertical direction. In particular, the representations of $su(n)$ and $sl(n,\mathbb{C})$ as Lie algebras of complex-linear endomorphisms of a complex vector space V correspond bijectively to each other.

Note: Using the notation 2.12 the map

$$\rho : M \to gl(V)$$

can be considered a morphism of real Lie algebras

$$M \to gl(V)_{\mathbb{R}}$$

But do not mix up $gl(V)_{\mathbb{R}}$ and $gl(V_{\mathbb{R}})$.

We saw in Remark 2.21 that the complex Lie algebra $sl(2,\mathbb{C})$ has the real form $su(2)$, which is the Lie algebra of the *compact* real matrix group $SU(2)$. A similar result holds for all complex Lie algebras from the A-D series, see [31, Chap. VI.10].

Proposition 2.22 (Compact real form). *Each complex Lie algebra from Proposition 2.15 has a* compact *real form, i.e. a real form which is the Lie algebra of a compact real matrix group:*

1. Series A_r, $m = r+1$*: The real Lie algebra $su(m)$ is a compact real form of $sl(m,\mathbb{C})$ because the matrix group $SU(m)$ is compact.*

2. Series B_r, $m = 2r+1$*: The real Lie algebra $so(m,\mathbb{R})$ is a compact real form of $so(m,\mathbb{C})$ because the matrix group $SO(m,\mathbb{R})$ is compact.*

3. Series C_r, $m = r$*: The real Lie algebra*

$$sp(m) := \{X \in gl(m,\mathbb{H}) : \ X^* + X = 0\},$$

with $\mathbb{H}$ the real division algebra of quaternions, is a compact real form of $sp(m,\mathbb{C})$ because the matrix group of unitary quaternions

$$SP(m) := \{X \in GL(m,\mathbb{H}) : \ g \cdot g^* = \mathbb{1}\}$$

is compact.

4. Series D_r, $m = 2r$*: The real Lie algebra $so(m,\mathbb{R})$ is a compact real form of $so(m,\mathbb{C})$ because the matrix group $SO(m,\mathbb{R})$ is compact.*

According to Example 1.30 the exponential map

$$exp : Lie\ G \to G$$

is not surjective in the general case. Compact real matrix groups G have the nice property that their exponential map

$$exp : Lie\ G \to G$$

is surjective, see [21, Chap. II, Prop. 6.10]. A typical example is the surjectivity of

$$exp : su(n) \to SU(n)$$

for the special unitary group, see Proposition 2.17.

A further application of Lie algebra theory for the investigation of topological properties of matrix groups is the polar decomposition. It can be used to restrict the investigation of topological properties of certain matrix groups to the study of a maximal compact subgroup.

Remark 2.23 (Polar decomposition and examples).

1. *Polar decomposition*: Consider a matrix group $G \subset GL(n,\mathbb{C})$ which is the zero set of polynomials in the $2 \cdot n^2$ real parts and imaginary parts of the entries of its matrices (algebraic matrix group), and which is invariant with respect to Hermitian conjugation. Then G has a *polar decomposition*: Define the subgroup

$$K := G \cap U(n)$$

and the vector space

$$P := \{X \in Lie\ G :\ X^* = X\} \text{ (Hermitian matrices).}$$

The polar decomposition theorem states, see [31, Prop. 1.143], [23, Prop. 4.3.3, Prop. 16.1.9]:

The product map

$$K \times P \to G,\ (k,X) \mapsto k \cdot exp\, X,$$

is a homeomorphism, hence

$$G = K \cdot exp(P)$$

Moreover, $K \subset G$ is a maximal compact subgroup with Lie algebra

$$Lie\ K = \{X \in Lie\ G :\ X + X^* = 0\} \text{ (Skew-Hermitian matrices),}$$

see [32, Lect. 2, Theor. on p. 25].

Using for the vector spaces of Hermitian matrices the notation

$$Herm(n,\mathbb{C}) := \{X \in M(n \times n,\mathbb{C}) :\ X = X^*\},\ Herm_0(n,\mathbb{C}) := \{X \in Herm(n,\mathbb{C}) : tr\, X = 0\}$$

and for the symmetric matrices the notation

$$Symm(n,\mathbb{R}) := \{X \in M(n \times n,\mathbb{R}) :\ X = X^\top\},\ Symm_0(n,\mathbb{R}) := \{X \in Herm(n,\mathbb{R}) : tr\, X = 0\}$$

the prototype of polar decompositions are the polar representations

$$GL(n,\mathbb{C}) = U(n)\cdot exp(Herm(n,\mathbb{C})),\ SL(n,\mathbb{C}) = SU(n)\cdot exp(Herm_0(n,\mathbb{C}))$$

and

$$GL(n,\mathbb{R}) = O(n,\mathbb{R})\cdot exp(Symm(n,\mathbb{R})),\ SL(n,\mathbb{R}) = SO(n,\mathbb{R})\cdot exp(Symm_0(n,\mathbb{R})),$$

see [18, Theor. 2.17, Prop. 2.19].

The vector spaces of matrices in these product decomposition are connected and simply connected. Therefore the connected components of G with polar decomposition

$$G = K\cdot exp(P)$$

correspond bijectively to the connected components of the compact group K, and the fundamental groups are isomorphic

$$\pi_1(G,*) \simeq \pi_1(K,*).$$

2. *The groups $O(p,q)$*: The matrix group

$$O(3,1) := \{g \in GL(4,\mathbb{R}) :\ g^\top \cdot I_{3,1}\cdot g = I_{3,1}\}$$

with the block matrix of type (3,1)

$$I_{3,1} := \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -1 \end{pmatrix}$$

is an algebraic matrix group. It is isomorphic to the Lorentz group which will be introduced in Definition 2.25 and further investigated in Remark 2.26 and Proposition 2.27.

The group $O(3,1)$ belongs to the class of real matrix groups

$$O(p,q) \subset GL(p+q,\mathbb{R}),\ 1 \le p,q.$$

They are the isometry groups of $\mathbb{R}^{p+q}$ provided with the bilinear form of signature (p,q), defined by the block matrix of type (p,q)

$$I_{p,q} := \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \in GL(p+q,\mathbb{R}),$$

namely

$$O(p,q) := \{g \in GL(p+q,\mathbb{R}) :\ g^\top \cdot I_{p,q}\cdot g = I_{p,q}\}$$

with Lie algebra

$$Lie\ O(p,q) = \{X \in sl(p+q,\mathbb{R}) : X^\top \cdot I_{p,q} + I_{p,q}\cdot X = 0\}$$

A maximal compact subgroup of $O(p,q)$ is

$$K_{O(p,q)} = O(p,q) \cap O(p+q,\mathbb{R}) \simeq O(p,\mathbb{R}) \times O(q,\mathbb{R})$$

The group has 4 connected components. Moreover

$$P_{O(p,q)} = \{X \in Lie\ O(p,q) :\ X = X^\top\}$$

The polar decomposition is the homeomorphic product map

$$K_{O(p,q)} \times P_{O(p,q)} \xrightarrow{\simeq} O(p,q),\ (k,X) \mapsto k \cdot exp\, X.$$

Accordingly $O(p,q)$ has 4 connected components.

3. *The groups* $SO(p,q)$: By definition

$$SO(p,q) := \{g \in O(p,q) :\ det\ g = 1\}$$

A maximal compact subgroup is

$$K_{SO(p,q)} := K_{O(p,q)} \cap SO(p,q) \simeq \{(g_1,g_2) \in O(p,\mathbb{R}) \times O(q,\mathbb{R}) :\ det\ g_1 \cdot det\ g_2 = 1\}$$

The group has 2 connected components corresponding to

$$(det\ g_1, det\ g_2) = (1,1) \text{ and } (det\ g_1, det\ g_2) = (-1,-1),$$

because each of the two groups $O(p,\mathbb{R})$ and $O(q,\mathbb{R})$ has two connected components, similar to $O(3,\mathbb{R})$ from Proposition 2.18.

The rest of the section presents two examples from covering theory. Each example deals with a matrix group and the corresponding simply connected matrix group which is the universal covering.

According to Remark 2.21 the two real Lie algebras

$$su(2) \simeq so(3,\mathbb{R})$$

are isomorphic. We now show that the corresponding connected matrix groups

$$SU(2) \text{ and } SO(3,\mathbb{R})$$

are not isomorphic: Example 2.24 constructs a 2-fold covering projection

$$\Phi : SU(2) \to SO(3,\mathbb{R}).$$

Due to Proposition 2.18 the group $SU(2)$ is simply connected. The existence of the 2-fold covering projection Φ implies that $SO(3,\mathbb{R})$ is not simply connected, but

$$\pi_1(SO(3,\mathbb{R},*) = \mathbb{Z}_2.$$

In order to obtain the morphism Φ we have to find out first: How does a unitary matrix $U \in SU(2)$ act on the 3-dimensional real space $\mathbb{R}^3$?

Example 2.24 ($SU(2)$ as two-fold covering of $SO(3,\mathbb{R})$).

1. *Vector space of Hermitian matrices*: Consider the real vector space of complex Hermitian traceless two-by-two matrices

$$Herm_0(2) := \{X \in M(2\times 2,\mathbb{C}) : X = X^*,\ tr\,X = 0\}.$$

The family of Pauli matrices $(\sigma_j)_{j=1,2,3}$, see Remark 2.21, is a basis of $Herm_0(2)$. We define the map

$$\alpha : \mathbb{R}^3 \to Herm_0(2), x = (x_1,x_2,x_3) \mapsto X := \sum_{j=1}^{3} x_j\cdot\sigma_j = \begin{pmatrix} x_3 & x_1 - i\cdot x_2 \\ x_1 + i\cdot x_2 & -x_3 \end{pmatrix}.$$

On $\mathbb{R}^3$ we consider the Euclidean quadratic form

$$q_E : \mathbb{R}^3 \to \mathbb{R}, x = (x_1,x_2,x_3) \mapsto \sum_{j=1}^{3} x_j^2.$$

Correspondingly, on the real vector space $Herm_0(2)$ we consider the quadratic form

$$q_H : Herm_0(2) \to \mathbb{R}, X \mapsto -det\,X,$$

i.e.

$$q_H(X) = a^2 + |b\,|^2\ for\ X = \begin{pmatrix} a & b \\ \overline{b} & -a \end{pmatrix}, a \in \mathbb{R}, b \in \mathbb{C}.$$

Then the map

$$\alpha : (\mathbb{R}^3, q_E) \xrightarrow{\sim} H := (Herm_0(2), q_H)$$

is an isometric isomorphism of Euclidean spaces , i.e. α is an isomorphism of real vector spaces satisfying

$$q_H(\alpha(x)) = q_E(x),\ x \in \mathbb{R}^3.$$

By means of the isometric isomorphism α we identify $O(3,\mathbb{R})$, the group of isometries of $(\mathbb{R}^3, q_E)$, with the group of isometries of H

$$O(H) := \{g \in GL(Herm_0(2)) : q_H(g(X)) = q_H(X)\ for\ all\ X \in Herm_0(2)\}.$$

Moreover, we identify the subgroup $SO(3,\mathbb{R}) \subset O(3,\mathbb{R})$ with the subgroup

$$SO(H) := \{g \in O(H) : det\ g = 1\} \subset O(H),$$

the connected component of the neutral element $e \in O(H)$.

In the following we do no longer deal with $SO(3,\mathbb{R})$ but with $SO(H)$.

2. *Definition of Φ*: We define the group morphism

$$\Phi : SU(2) \to SO(H), B \mapsto \Phi_B,$$

as the conjugation by B, setting

$$\Phi_B : Herm_0(2) \to Herm_0(2),\ X \mapsto B \cdot X \cdot B^{-1}.$$

Note $B \cdot X \cdot B^{-1} \in Herm_0(2)$, because $B^{-1} = B^*$ and $X^* = X$ imply

$$(B \cdot X \cdot B^{-1})^* = (B \cdot X \cdot B^*)^* = B \cdot X \cdot B^* = B \cdot X \cdot B^{-1}.$$

We have $\Phi_B \in O(H)$ because

$$-det(B \cdot X \cdot B^{-1}) = -det\ X.$$

Because $SU(2)$ is connected due to Proposition 2.18, we even have

$$\Phi_B \in SO(H).$$

Apparently Φ is a group morphism.

3. *Tangent map of Φ*: We use from the theory of Lie groups without proof the following general results:

- Lie groups, in particular matrix groups, are differentiable manifolds. The underlying vector space of the Lie algebra of a matrix group G is the tangent space $Lie\ G = T_eG$ at the neutral element $e \in G$.

- For each differentiable homomorphism of matrix groups

$$\Psi : G_1 \to G_2$$

the induced tangent map at the neutral element $e \in G_1$ is a morphism

$$\psi := Lie\ \Psi = T_e\Psi : Lie\ G_1 \to Lie\ G_2$$

of Lie algebras. It linearizes Ψ at the unit element $e \in G_1$.

Therefore we now determine the linearization of

$$\Phi : SU(2) \to SO(H),\ B \mapsto \Phi_B,$$

at the neutral element $e \in SU(2)$ as a Lie algebra morphism

$$\phi : su(2) \rightarrow so(H),\ A \mapsto \phi_A.$$

Recall

$$so(H) := Lie\ SO(H) \subset gl(Herm_0(2)).$$

For

$$A \in su(2) \text{ and } X \in Herm_0(2)$$

set

$$B = exp\ A \in SU(2)$$

Then

$$\Phi_B(X) = \Phi_{exp\ A}(X) = exp(A) \cdot X \cdot exp(-A) =$$

$$= (\mathbb{1} + A + O(A^2)) \cdot X \cdot (\mathbb{1} - A + O(A^2)) =$$

$$= X + A \cdot X - X \cdot A + O(A^2) = X + [A,X] + O(A^2).$$

As linearization with respect to $A \in su(2)$ we obtain

$$\phi_A : H \rightarrow H,\ X \mapsto \phi_A(X) = [A,X],$$

i.e.

$$\phi_A = ad\ A,$$

the linearisation ϕ of Φ at the neutral element $e \in SU(2)$ is the adjoint representation of

$$H \simeq so(3,\mathbb{R}) \simeq su(2).$$

Note: To compute the linearisation ϕ one can also fix $A \in su(2)$ and expand $\Phi_{exp\ tA}(X)$ with respect to powers of $t \in \mathbb{R}$.

To determine the value of

$$\phi : su(2) \rightarrow so(H)$$

on the basis elements $(i \cdot \sigma_j)_{j=1,2,3}$ of $su(2)$ we compute:

$$\phi_{i\sigma_1} : H \rightarrow H$$

$$\phi_{i\sigma_1}(\sigma_1) = 0, \varphi_{i\sigma_1}(\sigma_2) = i \cdot [\sigma_1, \sigma_2] = -2 \cdot \sigma_3$$

$$\phi_{i\sigma_1}(\sigma_3) = i \cdot [\sigma_1, \sigma_3] = -i \cdot [\sigma_3, \sigma_1] = 2 \cdot \sigma_2$$

With respect to the basis $(\sigma_j)_{j=1,2,3}$ of H we obtain

$$\phi_{i\sigma_1} = 2 \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \in so(H).$$

A similar evaluation of ϕ on the other two basis elements $i \cdot \sigma_2$ and $i \cdot \sigma_3$ gives

$$\phi_{i\sigma_2} = 2 \cdot \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \phi_{i\sigma_3} = 2 \cdot \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in so(H).$$

Hence the family $(\phi_{i\cdot\sigma_j})_{j=1,2,3}$ is linearly independent in $so(H)$. As a consequence,

$$\phi : su(2) \xrightarrow{\sim} so(H)$$

is an isomorphism because domain and range of ϕ are 3-dimensional Lie algebras.

4. *Surjectivity of Φ*: The morphism

$$\Phi : SU(2) \to SO(3,\mathbb{R})$$

is a local isomorphism at $\mathbb{1} \in SU(2)$ because its tangent map is bijective due to part 3. In particular, Φ is an open map. The image

$$\Phi(SU(2)) \subset SO(3,\mathbb{R})$$

is compact because $SU(2)$ is compact due to Proposition 2.18. As a consequence, the open and closed subset

$$\Phi(SU(2)) \subset SO(3,\mathbb{R})$$

equals the connected set $SO(3,\mathbb{R})$, i.e. the map Φ is surjective.

5. *Discrete kernel*: We are left with calculating the kernel of Φ. We claim:

$$ker\ \Phi = \{\pm\mathbb{1}\} \subset SU(2).$$

For the proof consider an arbitrary but fixed matrix

$$B = \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} \in SU(2)$$

with

$$\Phi_B = id_{Herm_0(2)},$$

i.e. for all $X \in Herm_0(2)$

$$B \cdot X \cdot B^{-1} = X$$

We have

$$B^{-1} = B^* = \begin{pmatrix} \overline{z} & -w \\ \overline{w} & z \end{pmatrix}.$$

Choosing for X successively the basis elements $\sigma_1, \sigma_2 \in Herm_0(2)$ we obtain

$$B \cdot \sigma_1 \cdot B^{-1} = \sigma_1 \text{ and } B \cdot \sigma_2 \cdot B^{-1} = \sigma_2.$$

Equating for both equations respectively on both sides the components shows after some calculation

$$z^2 \pm w^2 = 1$$

Hence

$$w = 0 \text{ and } z = \pm 1$$

i.e

$$B = \pm \mathbb{1}$$

6. *Universal covering space*: We use from Lie group theory without proof that a surjective morphism between Lie groups with discrete kernel is a covering projection. The group $SU(2)$ is simply connected according to Proposition 2.18. Hence the map

$$\Phi : SU(2) \to SO(3,\mathbb{R})$$

is the univeral covering projection of $SO(3,\mathbb{R})$, and $SU(2)$ is a double cover of $SO(3,\mathbb{R})$.

7. *Fundamental group of* $SO(3,\mathbb{R})$: As a consequence of part 6 we have

$$\pi_1(SO(3,\mathbb{R}),*) = \mathbb{Z}_2.$$

Nearly the same method as used in Example 2.24 allows to compute the universal covering space of the connected component of the neutral element of the Lorentz group. More specific, the covering projection Φ of $SO(3,\mathbb{R})$ from Example 2.24 extends to a covering projection Ψ of the *orthochronous Lorentz group*, see Proposition 2.27.

Figure 2.1 shows Minkowski space with the embedded light cone. Points of Minkowski space are named *events*. Referring to the event e, the origin, the closure of the interior of the *light cone*, the set with $q_M(x) \leq 0$, splits into the disjoint union of e as well as the future and the past cone of e. This splitting refers to the causal structure of the world: Events from the past may have influenced e, while e may influence events in its future. All other events, $q_M(x) > 0$, have no causal relation to e. They form the presence of the event e.

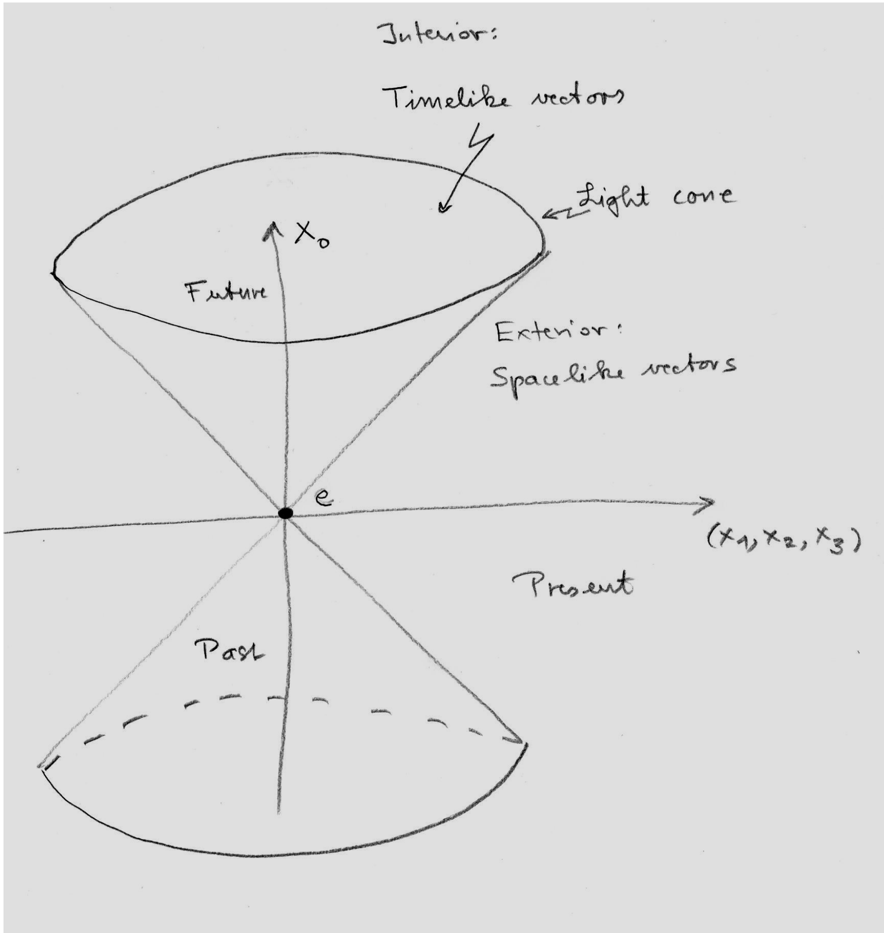

Fig. 2.1 The world model of flat Minkowski space

Definition 2.25 (Minkowski space and Lorentz group).

1. *Minkowski space* or *spacetime* is the pair

$$M := (\mathbb{R}^4, q_M)$$

with the quadratic form of signature (3,1)

$$q_M : \mathbb{R}^4 \to \mathbb{R}, q_M(x) := -x_0{}^2 + x_1{}^2 + x_2{}^2 + x_3{}^2,\ x = (x_0, ..., x_3)^\top.$$

We use the convention from Special Relativity concerning four-vectors $x \in M$ with coordinates $x = (x_0, x_1, x_2, x_3)$. The coordinate x_0 is interpreted as time, while x_1, x_2, x_3 are the usual space coordinates.

The Minkowski metric has the signature (3,1) with three positive eigenvalues $+1$ referring to the space coordinates and one negative eigenvalue -1 referring to the time coordinate. Using these conventions Euclidean space embeds into Minkowski space in a natural way:

$$\mathbb{R}^3 \hookrightarrow M,\ (x_1, x_2, x_3) \mapsto (0, x_1, x_2, x_3)$$

2. *Lorentz group*: The *Lorentz group* L is the real matrix group of isometries of Minkowski space

$$L := \{f \in GL(4,\mathbb{R}) : q_M(f(x)) = q_M(x)\ for\ all\ x \in \mathbb{R}^4\}.$$

Elements from L leave invariant the metric defined by the symmetric bilinear form

$$\eta := \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The Lorentz group L is isomorphic to the group $O(3,1)$ from Remark 2.23: While L employs the bilinear form η, the group $O(3,1)$ refers to the bilinear form

$$I_{3,1} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

with the same signature $(3,1)$.

Remark 2.26 (Lie algebra and orthochronous Lorentz group).

1. *Lie algebra*: The fact that all elements from the Lorentz group L have determinant ± 1, see Definition 2.25, implies for the Lie algebra of the Lorentz group

$$Lie\ L \simeq Lie\ O(3,1) = o(3,1) = so(3,1) = \{X \in sl(4,\mathbb{R}) :\ B^\top \cdot \eta + \eta \cdot B = 0\} =$$

$$= \left\{ \begin{pmatrix} 0 & b \\ b^\top & D \end{pmatrix} : b^\top \in \mathbb{R}^3,\ D \in o(3,\mathbb{R}) \right\}$$

A block matrix of type

$$\begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix} \in Lie\ L$$

is the infinitesimal generator of a 1-parameter group of rotations, while a block matrix of type

$$\begin{pmatrix} 0 & b \\ b^\top & 0 \end{pmatrix} \in Lie\ L$$

is the infinitesimal generator of a 1-parameter group of Lorentz boosts. The dimension is

$$dim_{\mathbb{R}}\ (Lie\ L) = dim_{\mathbb{R}}\ o(3,1) = 3 + dim_{\mathbb{R}}\ o(3,\mathbb{R}) = 6.$$

2. *The entry* b_{00}: By definition

$$B \in L \iff \eta = B^{\top} \cdot \eta \cdot B.$$

Hence for $B \in L$:

$$-1 = det\ \eta = det(B^{\top} \cdot \eta \cdot B) = det\ B^{\top} \cdot det\ \eta \cdot det\ B = -(det\ B)^2$$

which implies

$$det\ B = \pm 1.$$

Consider the timelike vector

$$e_0 := (1,0,0,0)^{\top}$$

Because each matrix

$$B = (b_{jk})_{0 \le j,k \le 3} \in L$$

acts as isometry on Minkowski space we obtain

$$-1 = q_M(e_0) = q_M(B \cdot e_0) = -b_{00}^2 + b_{10}^2 + b_{20}^2 + b_{30}^2.$$

Hence

$$b_{00}^2 = 1 + b_{10}^2 + b_{20}^2 + b_{30}^2 \ge 1.$$

3. *The orthochronous Lorentz group* $L_+^{\uparrow}$:

- According to the calculation in Definition 2.25, part 3 the group $SO(3,1)$ is the disjoint union of the two subsets

$$\left\{ (b_{jk})_{0 \le j,k \le 3} \in SO(3,1) : b_{00} \ge 1 \right\} \dot{\cup} \left\{ (b_{jk})_{0 \le j,k \le 3} \in SO(3,1) : b_{00} \le -1 \right\}$$

- and according to Remark 2.23, part 3 the group $SO(3,1)$ has two connected components.

Each of the two connected components is contained in only one of the two subsets. Hence the two subsets are the two connected components of $SO(3,1)$. As a consequence, the *orthochronous* Lorentz group

$$L_+^{\uparrow} := \{B = (b_{jk})_{0 \le j,k \le 3} \in L : det\ B = 1,\ b_{00} \ge 1\} \subset L$$

is the connected component of $SO(3,1)$ which contains the unit element of $SO(3,1)$. In particular, $L_+^{\uparrow}$ is a subgroup of L. The elements from $L_+^{\uparrow}$ keep the orientation of tetrads (German: 4-bein). They also keep the sign of the time component.

4. *The four connected components*: Besides $L_+^\uparrow \subset L$ the other three connected components of L are the neben-classes

$$L_-^\uparrow := \{B = (b_{jk})_{0 \le j,k \le 3} \in L : det\ B = -1,\ b_{00} \ge 1\}$$

$$L_+^\downarrow := \{B = (b_{jk})_{0 \le j,k \le 3} \in L : det\ B = 1,\ b_{00} \le -1\}$$

$$L_-^\downarrow := \{B = (b_{jk})_{0 \le j,k \le 3} \in L : det\ B = -1,\ b_{00} \le -1\}$$

The two connected components of $SO(3,1)$ correspond to the following two of the four connected components of L

$$L_+^\uparrow \text{ and } L_+^\downarrow \text{ (time reversal).}$$

Proposition 2.27 (Universal covering projection of the Lorentz group). *The proper orthochronous Lorentz group has the universal covering projection*

$$\Psi : SL(2,\mathbb{C}) \to L_+^\uparrow$$

with the group homomorphism Ψ a two-fold covering projection. The following diagram commutes

$$\begin{tikzcd} SU(2) \arrow[r, "\Phi"] \arrow[d] & SO(3,\mathbb{R}) \arrow[d] \\ SL(2,\mathbb{C}) \arrow[r, "\Psi"] & L_+^\uparrow \end{tikzcd}$$

with the canonical inclusions in the vertical direction and the map Φ from Example 2.24.

Proof. See also [42, Anhang L.8].

1. *Minkowski space as a vector space of matrices*: Let

$$H := (Herm(2), q_H)$$

denote the real vector space of Hermitian matrices

$$Herm(2) := \{X \in M(2 \times 2, \mathbb{C}) : X = X^*\}$$

equipped with the real quadratic form

$$q_H : Herm(2) \to \mathbb{R}, X \mapsto -det\ X,$$

i.e.

$$q_H(X) = -(a \cdot d) + |b|^2$$

for

$$X = \begin{pmatrix} a & b \\ \overline{b} & d \end{pmatrix}, a, d \in \mathbb{R}, b \in \mathbb{C}.$$

Set $\sigma_0 := \mathbb{1} \in Herm(2)$. Then the family $(\sigma_j)_{j=0,...,3}$ is a basis of the vector space $Herm(2)$. The map

$$\beta : M \to H, x = (x_0, ..., x_3) \mapsto X := \sum_{j=0}^{3} x_j \cdot \sigma_j = \begin{pmatrix} x_0 + x_3 & x_1 - i \cdot x_2 \\ x_1 + i \cdot x_2 & x_0 - x_3 \end{pmatrix}$$

is an isometric isomorphism, i.e. an isomorphism of vector spaces satisfying

$$q_H(\beta(x)) = q_M(x),\ x \in \mathbb{R}^4.$$

The isometry property is due to

$$q_H(\beta(x)) = -det\ \beta(x) = -x_0^2 + x_1^2 + x_2^2 + x_3^2 = q_M(x).$$

By means of the isometric isomorphism β we identify the group of isometries of H

$$O(H) := \{g \in GL(Herm(2)) : q_H(g(X)) = q_H(X)\ for\ all\ X \in Herm(2)\}$$

with $O(3,1)$ and denote by

$$L_+^{\uparrow}(H) \subset O(H)$$

the connected component of the neutral element $id_H \in O(H)$.

2. *Definition of* Ψ: The map

$$\Psi : SL(2,\mathbb{C}) \to O(H), B \mapsto \Psi_B,$$

defined by the conjugation

$$\Psi_B : H \to H, X \mapsto B \cdot X \cdot B^*,$$

is a well-defined morphism of real matrix groups. We have

$$\Psi(SL(2,\mathbb{C})) \subset L_+^{\uparrow}(H)$$

because $SL(2,\mathbb{C})$ is connected due to Proposition 2.18.

3. *Tangent map of* Ψ: The family $(A_j)_{j=1,...,6}$ with

$$A_j := \begin{cases} \sigma_j & \text{if } j = 1,2,3 \\ i \cdot \sigma_{j-3} & \text{if } j = 4,5,6 \end{cases}$$

is a basis of the real vector space $sl(2,\mathbb{C})_{\mathbb{R}}$. Note that

$$su(2) = span_{\mathbb{R}} < i \cdot \sigma_j : j = 1,2,3 >$$

is a real subalgebra of $sl(2,\mathbb{C})_{\mathbb{R}}$, but

$$span_{\mathbb{R}} < \sigma_j : j = 1,2,3 >$$

is not a subalgebra of $sl(2,\mathbb{C})_{\mathbb{R}}$. Denote by

$$o(H) := Lie\ O(H) \subset gl(Herm(2))$$

the Lie algebra of the matrix group $O(H)$ and by

$$\psi := Lie\ \Psi : sl(2,\mathbb{C})_{\mathbb{R}} \to o(H),\ A \mapsto \psi_A,$$

the tangent map of Ψ at $\mathbb{1} \in SL(2,\mathbb{C})$. One checks that ψ as the linearization of Ψ takes the values

$$\psi_A(X) = A \cdot X + X \cdot A^*,\ A \in sl(2,\mathbb{C})_{\mathbb{R}},\ X \in Herm(2).$$

One may also check that indeed ψ is a morphism of Lie algebras:

$$\psi_{[B_1,\ B_2]} = [\psi_{B_1}, \psi_{B_2}]$$

by using the commutator relation of the Pauli matrices and the Hermitian resp. skew-Hermitian properties

$$\sigma_j^* = \sigma_j,\ (i \cdot \sigma_j)^* = -i \cdot \sigma_j \text{ for } j = 1,2,3.$$

Explicit computation of the matrices representing

$$\psi_{A_j}, j = 1,...,6,$$

shows: The family $(\psi_{A_j})_{j=1,...,6}$ is linearly independent in the 6-dimensional vector space $o(H) \subset End(Herm(2))$.

4. *Surjectivity of* Ψ: The map

$$\Psi : SL(2,\mathbb{C}) \to L_+^{\uparrow}(H)$$

is open because its tangent map at $\mathbb{1} \in SL(2,\mathbb{C})$ is an isomorphism. The image

$$\Psi(SL(2,\mathbb{C})) \subset L_+^{\uparrow}(H)$$

is also closed because

$$L_+^{\uparrow}(H) = \bigcup_{g \in L_+^{\uparrow}(H)} g \cdot \Psi(SL(2,\mathbb{C}))$$

represents the complement

$$L_+^{\uparrow}(H) \setminus \Psi(SL(2,\mathbb{C}))$$

as a union of open subsets. Hence

$$\Psi(SL(2,\mathbb{C})) \subset L_+^{\uparrow}(H)$$

is also closed, and

$$\Psi : SL(2,\mathbb{C}) \to L_+^{\uparrow}(H)$$

is surjective, because $SL(2,\mathbb{C})$ is connected and $L_+^{\uparrow}(H)$ is the connected component of the unit element according to Remark 2.26.

5. *Discrete kernel*: The kernel is

$$ker\ \Psi = \{\pm\mathbb{1}\} \subset SL(2,\mathbb{C}) :$$

For the proof one evaluates for $B \in SL(2,\mathbb{C})$ the condition:

$$\Psi_B(X) = X$$

for all elements $X \in Herm(2)$ from the basis $(\sigma_j)_{j=0,...,3}$ of $Herm(2)$.

6. *Universal covering space*: Due to the previous parts the map

$$\Psi(SL(2,\mathbb{C})) \to L_+^{\uparrow}(H)$$

is a 2-fold covering projection. It is the universal covering projection because $SL(2,\mathbb{C})$ is simply connected according to Proposition 2.18.

7. *Fundamental group*: Hence the orthochronous Lorentz group has the fundamental group

$$\pi_1(L_+^{\uparrow}, *) = \mathbb{Z}_2,$$

the group of deck-tranformations of the universal covering projection.
□

Remark 2.28 (Lie algebras and simply connected matrix group).

1. There is a close relation between Lie algebras and simply connected matrix groups: Consider two matrix groups G and H. If G is connected and simply connected, then each morphism

$$\phi : Lie\ G \to Lie\ H$$

of Lie algebras lifts to a unique morphism

$$\Phi : G \to H$$

of matrix groups such that the following diagram commutes

$$\begin{tikzcd} G \arrow[r, dashed, "\Phi"] & H \\ Lie\ G \arrow[u, "exp"] \arrow[r, "\phi"] & Lie\ H \arrow[u, "exp"'] \end{tikzcd}$$

The result is not restricted to matrix groups, it holds for Lie groups in general: The categories of connected, simply-connected Lie groups over $\mathbb{R}$ respectively $\mathbb{C}$ and the category of Lie algebra over these fields are equivalent, see [41, Part II, Chap. V, §8 Theor. 2].

2. For a general, not necessarily simply connected matrix group G one applies the previous result to the universal covering projection

$$\pi : \tilde{G} \to G,$$

taking into account that $\tilde{G}$ and G have the same Lie algebra. Then one has to study the covering projection π in order to determine whether the morphism from the universal covering

$$\tilde{\Phi} : \tilde{G} \to H$$

projects down to a morphism

$$\Phi : G \to H$$

such that the diagram

$$\begin{tikzcd} \tilde{G} \arrow[r, "\tilde{\Phi}"] \arrow[d, "\pi"'] & H \\ G \arrow[ur, dashed, "\Phi"'] & \end{tikzcd}$$

commutes.

Note: The universal covering of a matrix group is a Lie group, but not necessarily a matrix group. A counter example is the universal covering of the matrix group $SL(2,\mathbb{R})$, see [18, Prop. 5.16].

Chapter 3
Nilpotent Lie algebras and solvable Lie algebras

If not stated otherwise, all Lie algebras and vector spaces in this chapter will be assumed finite dimensional over the base field $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$.

3.1 Engel's theorem for nilpotent Lie algebras

Recall Definition 1.11: An endomorphism $f \in End(V)$ with V a vector space is *nilpotent* iff an index $n \in \mathbb{N}$ exists with $f^n = 0$. Note that complex eigenvalues of a nilpotent endomorphism are *zero*.

Definition 3.1 (Ad-nilpotency). Consider a Lie algebra L. An element $x \in L$ is *ad-nilpotent* iff the induced endomorphism of L

$$ad\ x : L \to L,\ y \mapsto [x,y],$$

is nilpotent.

Nilpotency and ad-nilpotency refer to two different structures: Nilpotency iterates the *associative* product, while ad-nilpotency iterates the *Lie product*. If the Lie algebra results from a matrix algebra both concepts are related: A Lie algebra of nilpotent endomorphisms of a vector space acts "nilpotent" on itself by the adjoint representation.

Lemma 3.2 (Nilpotency implies ad-nilpotency). *Consider a vector space V, an embedded Lie algebra $L \subset gl(V)$ and an element $x \in L$. If the endomorphism*

$$x : V \to V$$

is nilpotent, then also the induced endomorphism

$$ad\ x : L \to L$$

is nilpotent.

The content of Lemma 3.2 can be paraphrased as: Nilpotency implies ad-nilpotency.

Proof. The endomorphism $x \in End(V)$ acts on $End(V)$ by left composition and right composition

$$l : End(V) \to End(V),\ y \mapsto x \circ y,$$

$$r : End(V) \to End(V),\ y \mapsto y \circ x.$$

Then

$$ad\ x = l - r$$

because for all $y \in End(V)$

$$(ad\ x)(y) = [x,y] = (l-r)(y).$$

Nilpotency of x implies that both actions are nilpotent. Both actions commute: For all $y \in End(V)$

$$[l,r](y) = (l \circ r - r \circ l)(y) = x \circ (y \circ x) - (x \circ y) \circ x = 0.$$

Proposition 1.12 implies the nilpotency of the difference

$$ad\ x = l - r$$

□

The adjoint representation respects the Jordan decomposition from Theorem 1.19.

Proposition 3.3 (Jordan decomposition of the adjoint representation). *Consider an n-dimensional complex vector space V, an endomorphism $f \in End(V)$ and its Jordan decomposition*

$$f = f_s + f_n.$$

Then the Jordan decomposition of

$$ad\ f \in L := gl(End\ V),$$

defined as

$$ad\ f : End(V) \to End(V),\ g \mapsto [f,g],$$

is

$$ad\ f = ad\ f_s + ad\ f_n \in L.$$

In particular, $ad\ f_s$ is semisimple, $ad\ f_n$ is nilpotent and

$$[ad\ f_s, ad\ f_n] = 0.$$

Proof. i) *Nilpotency of ad f_n*: According to Lemma 3.2 the endomorphism *ad* f_n is nilpotent.

ii) *Semisimplicity of ad f_s*: In order to show that *ad* f_s is semisimple we choose a base $(v_1, ..., v_n)$ of V consisting of eigenvectors of f_s with corresponding eigenvalues $\lambda_1, ..., \lambda_n$.

Let $(E_{ij})_{1 \le i,j \le n}$ denote the standard base of $End(V)$ relatively to $(v_1, ..., v_n)$, i.e.

$$E_{ij}(v_k) = \delta_{jk} v_i$$

mapping v_j to v_i and annihilating v_k for all $k \neq j$.

For $1 \le i, j, k \le n$:

$$((ad\ f_s)E_{ij})(v_k) = [f_s, E_{ij}](v_k) = f_s(E_{ij}(v_k)) - E_{ij}(f_s(v_k)) = (f_s - \lambda_k)(E_{ij}(v_k)) =$$

$$= (f_s - \lambda_k)(\delta_{jk} v_i) = (\lambda_i - \lambda_j)(\delta_{jk} v_i) = (\lambda_i - \lambda_j) \cdot E_{ij}(v_k)$$

Hence

$$(ad\ f_s)(E_{ij}) = (\lambda_i - \lambda_j) \cdot E_{ij}$$

and *ad* f_s acts with respect to the standard basis (E_{ij}) diagonally on the vector space $End(V)$ with eigenvalues

$$\lambda_i - \lambda_j,\ 1 \le i, j \le n.$$

Hence *ad* f_s is semisimple.

iii) *Commutator* $[ad\ f_s, ad\ f_n]$: Because

$$ad : gl(V) \to L$$

is a morphism of Lie algebras:

$$[ad\ f_s, ad\ f_n] = ad([f_s, f_n]) = 0.$$

□

Using the adjoint representation we carry over the concept of *semisimpleness* to elements of arbitrary complex Lie algebra. Definition 3.4 is analogous to Definition 3.1.

Definition 3.4 (Ad-semisimple element of a complex Lie algebra). Consider a complex Lie algebra L. An element $x \in L$ is *ad-semisimple* iff the induced endomorphism

$$ad\ x : L \to L,\ y \mapsto [x, y]$$

is semisimple.

It is a trivial observation that a nilpotent endomorphism of a non-zero vector space V has an eigenvector with eigenvalue zero. Theorem 3.5 strongly generalizes this fact: It proves the existence of a common eigenvector for a *whole* Lie algebra of nilpotent endomorphisms.

Theorem 3.5 (Annihilation of a common eigenvector).

Consider a vector space $V \neq \{0\}$ and an embedded Lie algebra $L \subset gl(V)$. If each endomorphism $x \in L$ is nilpotent, then all elements $x \in L$ annihilate a common eigenvector, i.e., a nonzero vector $v \in V$ exists with

$$x(v) = 0 \text{ for all } x \in L.$$

The proof of Theorem 3.5 is standard, cf. [24, Chap.3.3], [13, Theor. 9.9]. The idea of the induction step is to find an ideal $I \subset L$ of codimension 1 and then to split

$$L = I \oplus \mathbb{K} \cdot x_0 \text{ with } x_0 \in L \setminus I,$$

as the direct sum of two subalgebras.

Proof. The proof goes by induction on $dim\, L \in \mathbb{N}$, while the dimension of the finite-dimensional vector space V is left arbitrary.

Apparently the theorem is true for $dim\, L = 0$.

For the induction step assume $dim\, L > 1$ and assume that the theorem is true for all Lie algebras of less dimension.
i) *Existence of an ideal of codimension 1*: By assumption all $x \in L$ are nilpotent endomorphisms of V. Hence by Lemma 3.2 each endomorphism

$$ad\, x : L \to L,\ x \in L,$$

is nilpotent.

Define the set

$$\mathscr{A} := \{K \subsetneq L :\ K \text{ Lie subalgebra}\}$$

of all proper Lie subalgebras of L. We have $\mathscr{A} \neq \emptyset$, because the zero-dimensional vector subspace $K = \{0\} \subset L$ is a subalgebra of L.

Choose

$$I \in \mathscr{A}$$

as a subalgebra of L having maximal dimension with respect to all Lie algebras from $\mathscr{A}$. By definition

$$dim\, I < dim\, L$$

Because I is a Lie algebra, for all $x \in I$

$$(ad\ x)(I) \subset I$$

Hence the adjoint representation of L induces a representation of I on the vector space L/I: For $x \in I$ the endomorphism

$$\overline{ad\ x} : L/I \to L/I,\ y+I \mapsto [x,y]+I$$

is well-defined. For each $x \in I$ the nilpotency of $ad\ x \in L$ implies the nilpotency of $\overline{ad\ x}$. The induction hypothesis applies to the Lie algebra I and the vector space

$$V := L/I,$$

and provides a common eigenvector with eigenvalue zero

$$\overline{y} = y+I \in L/I$$

for all endomorphisms

$$\overline{ad\ x} : L/I \to L/I$$

Eigenvectors are non-zero vectors, hence

$$y \in L \setminus I \text{ with } [x,y] \in I \text{ for all } x \in I.$$

As a consequence

$$y \in N_L(I) \setminus I$$

i.e. the Lie subalgebra $I \subset L$ is properly contained in its normalizer. Due to the maximality of the dimension of I we have $N_L(I) \notin \mathscr{A}$, hence

$$N_L(I) = L,$$

i.e. $I \subset L$ is a proper ideal.

The ideal $I \subset L$ induces a canonical projection of Lie algebras

$$\pi : L \to L/I.$$

In case

$$dim\ (L/I) \geq 2$$

one chooses a 1-dimensional Lie subalgebra $K \subset L/I$. The inverse image

$$\pi^{-1}(K) \subset L$$

is a proper Lie subalgebra of L, properly containing I, a contradiction to the maximality of I. Hence

$$dim\ (L/I) = 1$$

ii) *Constructing a common eigenvector*: The construction from part i) implies the vector space decomposition

$$L = I + \mathbb{K} \cdot x_0$$

with an arbitrary element $x_0 \in L \setminus I$. By induction assumption the elements of the Lie algebra I annihilate a common non-zero eigenvector, i.e.

$$W := \{w \in V : x(w) = 0 \; for \; all \; x \in I\} \neq \{0\}.$$

In addition, the subspace $W \subset V$ is stable with respect to the endomorphism $x_0 \in End(V)$, i.e. for all $w \in W$ holds

$$x_0(w) \in W :$$

For all $x \in I$

$$x(x_0(w)) = x_0(x(w)) - [x_0, x](w) = 0.$$

Here the first summand vanishes because $w \in W$ and $x \in I$. The second summand vanishes because $I \subset L$ is an ideal, which implies

$$[x_0, x] \in I \text{ and } [x_0, x](w) = 0$$

The restriction

$$x_0 | W : W \to W$$

is a nilpotent endomorphism, and therefore has an eigenvector $\mathrm{v}_0 \in W \subset V$ with eigenvalue 0. From the decomposition

$$L = I + \mathbb{K} \cdot x_0$$

follows for all $x \in L$

$$x(\mathrm{v}_0) = 0.$$

This finishes the induction step. □

The fact, that any embedded Lie algebra $L \subset gl(V)$ of nilpotent endomorphisms annihilates a common eigenvector, allows the simultaneous triagonalization to strict upper triangular matrices for all endomorphisms of L. We state the result by using the concept of a flag from Definition 1.11.

Corollary 3.6 (Simultaneous strict triagonalization of nilpotent endomorphisms). *Consider an n-dimensional $\mathbb{K}$-vector space V and an embedded Lie algebra $L \subset gl(V)$. Then the following properties are equivalent:*

1. *Each endomorphism $x \in L$ is nilpotent.*
2. *A flag $(V_i)_{i=0,...,n}$ of subspaces of V exists such that all $x \in L$ satisfy for all $i = 1, ..., n$*

$$x(V_i) \subset V_{i-1}$$

3. *The Lie algebra L is isomorphic to a Lie subalgebra of the Lie algebra of strictly upper triangular matrices*

$$\mathfrak{n}(n,\mathbb{K}) = \left\{ \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix} \in gl(n,\mathbb{K}) \right\}.$$

Proof. 1) $\implies$ 2): The proof goes by induction on $n = dim\,V$. The implication is valid for $n = 0$.

Assume $n > 0$ and assume part 2) valid for all vector spaces of less dimension. Set

$$V_0 := \{0\}.$$

According to Theorem 3.5 all elements from L annihilate a common eigenvector $v_1 \in V$. Consider the quotient vector space

$$W := V/(\mathbb{K} \cdot v_1)$$

with the canonical projection of vector spaces

$$\pi : V \to W.$$

Each endomorphism $x \in L \subset gl(V)$ annihilates v_1, and therefore induces an endomorphism $\overline{x} : W \to W$ such that the following diagram commutes

$$\begin{array}{ccc} V & \xrightarrow{x} & V \\ \downarrow \pi & & \downarrow \pi \\ W & \xrightarrow{\overline{x}} & W \end{array}$$

The endomorphism $\overline{x} \in End(W)$ is nilpotent. The induction assumption applied to W with

$$dim\,W < dim\,V$$

provides a flag $(W_i)_{i=0,...,n-1}$ of subspaces of W with

$$\overline{x}(W_i) \subset W_{i-1} \ for\ all\ i = 1,...,n-1.$$

Now define

$$V_i := \pi^{-1}(W_{i-1}),\ i = 1,...,n.$$

Then

$$\overline{x}(W_i) \subset W_{i-1}$$

implies

$$x(V_i) \subset V_{i-1},\ i = 1,...,n.$$

2) $\implies$ 3) For the proof one constructs step by step a basis of V

$$(\mathrm{v}_1,...,\mathrm{v}_n)$$

which satisfies for all $i = 1, ..., n$

$$V_i = span_{\mathbb{K}} < \mathrm{v}_1, ..., \mathrm{v}_i >$$

3) $\implies$ 1) The proof is obvious. $\square$

Due to Lemma 3.2 the non-trivial implication of Corollary 3.6 states: Consider a vector space V. Then each embedded Lie algebra $L \subset gl(V)$ of ad-nilpotent endomorphisms is nilpotent.

A Lie algebra is *Abelian* when the commutator of any two elements vanishes. Nilpotent Lie algebras are a first step to generalize this property: A Lie algebra L is named *nilpotent* when a number $N \in \mathbb{N}$ exists such that all N-fold commutators of elements from L vanish. The concept of the *descending central series* of a Lie algebra formalizes this property.

Definition 3.7 (Descending central series and nilpotent Lie algebra). Consider a Lie algebra L.

1. The *descending central series* or *lower central series* of L is the sequence $(C^iL)_{i\in\mathbb{N}}$ of subsets $C^iL \subset L$, inductively defined as

$$C^0L := L \text{ and } C^{i+1}L := [L, C^iL],\ i \in \mathbb{N}.$$

2. The Lie algebra L is *nilpotent* iff an index $i_0 \in \mathbb{N}$ exists with

$$C^{i_0}L = 0.$$

The smallest index with this property is named the *length* of the descending central series.

3. An ideal $I \subset L$ is *nilpotent* if I is nilpotent when considered as a Lie algebra with Lie bracket the restricted Lie bracket of L.

By induction on $i \in \mathbb{N}$ one easily verifies

$$C^{i+1}L \subset C^iL.$$

As a consequence, all elements of the descending central series are ideals

$$C^iL \subset L,\ i \in \mathbb{N}.$$

Apparently, any Lie subalgebra and any quotient algebra of a nilpotent Lie algebra L is nilpotent too: The descending central series of a subalgebra respectively of a quotient of L arise from the descending central series of L by restriction respectively taking quotients. Concerning the reverse implication consider Lemma 3.8.

The Lie algebra $\mathfrak{n}(n,\mathbb{K})$ of strictly upper triangular matrices - and hence also all Lie subalgebras of $\mathfrak{n}(n,\mathbb{K})$ - are nilpotent. Due to Corollary 3.6 an embedded Lie algebra $L \subset gl(V)$ is nilpotent if each element $x \in L$ is a nilpotent endomorphism of V.

Lemma 3.8 (Central extension of a nilpotent Lie algebra). *Consider a Lie algebra L and an ideal $I \subset L$ which is contained in the center, i.e.*

$$I \subset Z(L).$$

If the quotient L/I is nilpotent then L, named an extension *of L/I by I, is nilpotent too.*

In Lemma 3.8 the extension is named *central* extension because $I \subset Z(L)$. In particular, I is a nilpotent Lie algebra.

Proof. The canonical morphism

$$\pi : L \to L' := L/I$$

of Lie algebras relates the descending central series of L and L' as

$$C^i L' = \pi(C^i L),\ i \in \mathbb{N}.$$

If $C^{i_0} L' = 0$ then $\pi(C^{i_0} L) = 0$, i.e.

$$C^{i_0} L \subset I \subset Z(L).$$

The center satisfies $[L, Z(L)] = 0$, hence

$$C^{i_0+1} L := [L, C^{i_0} L] = 0.$$

□

In Lemma 3.8 one must not drop the assumption $I \subset Z(L)$, i.e. that the ideal $I \subset L$ belongs to the center of L. It is not enough to assume that I is nilpotent:

Example 3.9 (Counter example against nilpotency of extensions). The descending central series of the Lie algebra of upper triangular matrices

$$L := \mathfrak{t}(2,\mathbb{K}) = \left\{ \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \in gl(2,\mathbb{K}) \right\}$$

starts with

$$C^0L := L, \text{ and } C^1L = [L,L] = \mathfrak{n}(2,\mathbb{K}) = \mathbb{K} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

the Lie algebra of strictly upper triangular matrices. The elements of the derived series of L are

$$C^2L = [L, C^1L] = [L, \mathfrak{n}(2,\mathbb{K})] = \mathfrak{n}(2,\mathbb{K}) = C^1L \neq 0.$$

As a consequence for all $i \in \mathbb{N}^*$

$$C^iL = C^1L \neq \{0\}.$$

Hence L is not nilpotent. But

$$I := \mathfrak{n}(2,\mathbb{K}) \subset L$$

is a nilpotent ideal, and the quotient

$$L/I \simeq \mathfrak{d}(2,\mathbb{K}),$$

the Lie algebra of diagonal matrices, is also nilpotent. We have

$$I \not\subset Z(L) = \mathbb{K} \cdot \mathbb{1}$$

We now state Engel's theorem. It is the main result about nilpotent Lie algebras. It characterizes the nilpotency of a Lie algebra by the ad-nilpotency of all its elements. Engel's theorem follows as a corollary from Theorem 3.5.

Theorem 3.10 (Engel's theorem for nilpotent Lie algebras). *A Lie algebra L is nilpotent if and only if every element $x \in L$ is ad-nilpotent.*

Proof. i) Assume that every endomorphism

$$ad\ x : L \to L,\ x \in L,$$

is nilpotent. According to Corollary 3.6 the embedded Lie algebra

$$ad\ L \subset gl(L)$$

is isomorphic to a subalgebra of

$$\mathfrak{n}(n,\mathbb{K}),\ n = dim\ L,$$

of strictly upper triangular matrices, hence $ad\ L$ is nilpotent. Due to Lemma 3.8 the isomorphy

$$L/Z(L) \simeq ad\ L$$

implies the nilpotency of L.

ii) Suppose that L is nilpotent. An index $i_0 \in \mathbb{N}$ exists with $C^{i_0}L = 0$. Hence in particular for all $x \in L$

$$ad^{i_0}(x) = 0,$$

i.e. $ad\ x$ is nilpotent.

□

From the categorical point of view the concept of a *short exact sequence* is a useful tool to handle injective or surjective morphisms and their cokernels respectively kernels.

Definition 3.11 (Exact sequence of Lie algebra morphisms).

1. A *chain complex* of Lie algebra morphisms is a sequence of Lie algebra morphisms
$$(L_i \xrightarrow{f_i} L_{i+1})_{i \in \mathbb{Z}}$$
such that for all $i \in \mathbb{Z}$ the composition $f_{i+1} \circ f_i = 0$, i.e.
$$im[L_{i-1} \xrightarrow{f_{i-1}} L_i] \subset ker[L_i \xrightarrow{f_i} L_{i+1}].$$
2. A chain complex of Lie algebra morphisms $(L_i \xrightarrow{f_i} L_{i+1})_{i \in \mathbb{Z}}$ is *exact* or an *exact sequence* of Lie algebra morphisms if for all $i \in \mathbb{Z}$
$$im[L_{i-1} \xrightarrow{f_{i-1}} L_i] = ker[L_i \xrightarrow{f_i} L_{i+1}].$$
3. A *short exact sequence* of Lie algebra morphisms is an exact sequence of the form
$$0 \to L_0 \xrightarrow{f_0} L_1 \xrightarrow{f_1} L_2 \to 0.$$

A short exact sequence

$$0 \to L_0 \xrightarrow{f_0} L_1 \xrightarrow{f_1} L_2 \to 0$$

expresses the following facts about the two morphisms:

- $f_0 : L_0 \to L_1$ is injective,
- $f_1 : L_1 \to L_2$ is surjective and
- $im\ f_0 = ker\ f_1$, in particular $L_2 \simeq L_1/L_0$.

Using the concept of exact sequences we restate from Lemma 3.8 and its prologue the relation between the nilpotency of a Lie algebra L, an ideal of L and a quotient of L as follows:

Proposition 3.12 (Nilpotency and short exact sequences). *Consider a short exact sequence of Lie algebra morphisms*

$$0 \to L_0 \xrightarrow{j} L_1 \xrightarrow{\pi} L_2 \to 0.$$

1. *If the Lie algebra L_1 is nilpotent, then also L_0 and L_2 are nilpotent.*
2. *If L_2 is nilpotent and $j(L_0) \subset Z(L_1)$, then also L_1 is nilpotent.*

For a Lie algebra L one obtains for any $i \in \mathbb{N}$ a short exact sequence

$$0 \to C^iL/C^{i+1}L \to L/C^{i+1}L \to L/C^iL \to 0$$

By definition of the descending central series, the Lie algebra on the left-hand side is contained in the center of the Lie algebra in the middle, i.e.

$$C^iL/C^{i+1}L \subset Z(L/C^{i+1}L).$$

For a nilpotent Lie algebra L, with $i_0 \in \mathbb{N}$ the length of the descending central series, the exact sequence

$$0 \to C^{i_0-1}L \to L \to L/C^{i_0-1}L \to 0$$

presents L as a central extension of the nilpotent Lie algebra $L/C^{i_0-1}L$.
Successively one obtains L as a finite sequence of central extensions of nilpotent Lie algebras: One starts with the nilpotent Lie algebra $L/C^{i_0-1}L$ on the right-hand side and

$$C^{i_0-1} \subset Z(L)$$

on the left-hand side. The next step presents $L/C^{i_0-1}L$ as a central extension of the nilpotent Lie algebra L/C^{i_0-2} ...
This sequence of central extensions of nilpotent Lie algebras explains the attribute *central* in the name of the descending central series $(C^iL)_{i\in\mathbb{N}}$.

For later use we prove two results about the centralizer of a nilpotent Lie algebra L and about the normalizer of a proper subalgebra of a nilpotent Lie algebra. Due to Corollary 3.13 each ideal of a nilpotent Lie algebra L contains non-zero elements from the center of L.

Corollary 3.13 (Center of nilpotent Lie algebras). *Consider a nilpotent Lie algebra $L \neq \{0\}$. For each non-zero ideal $I \subset L$ holds*

$$Z(L) \cap I \neq \{0\}.$$

In particular,

$$\{0\} \neq Z(L) \subset C_L(I).$$

Proof. According to Theorem 3.10 all endomorphisms

$$ad\ x : L \to L,\ x \in L,$$

are nilpotent. Because $I \subset L$ is an ideal each restriction

$$(ad\ x)|I : I \to I,\ x \in L,$$

is well-defined and also nilpotent. Theorem 3.5, applied to the embedded Lie algebra

$$ad\ L \subset gl(I),$$

provides a non-zero element $x_0 \in I$ with

$$[L, x_0] = 0,$$

i.e. $0 \neq x_0 \in Z(L) \cap I$. □

In a nilpotent Lie algebra L each proper subalgebra $M \subsetneq L$ embeds properly into its normalizer. The result explains a posteriori the construction made in the proof of Theorem 3.5.

Proposition 3.14 (Normalizer in nilpotent Lie algebras). *Any proper subalgebra $M \subsetneq L$ of a nilpotent Lie algebra L is properly contained in its normalizer, i.e.*

$$M \subsetneq N_L(M).$$

Proof. We consider the descending central series of L. Because $M \subsetneq L = C^0L$ we start with

$$M \subsetneq M + C^0L.$$

Due to $C^{i_0}L = 0$ for a suitable index $i_0 \in \mathbb{N}$ we end with

$$M = M + C^{i_0}L.$$

Let $i < i_0$ be the largest index with

$$M \subsetneq M + C^iL.$$

Then

$$M = M + C^{i+1}L.$$

Therefore

$$[M + C^iL, M] \subset [M, M] + [C^iL, M] \subset M + [C^iL, L] = M + C^{i+1}L = M,$$

which implies

$$M + C^iL \subset N_L(M).$$

We obtain

$$M \subsetneq M + C^i L \subset N_L(M).$$

□

The logical dependencies between the results of this section is clarified by the diagram from Figure 3.1. It shows the fundamental role of Theorem 3.5 about the existence of a common eigenvector annihilated by all elements of an embedded Lie algebra of nilpotent endomorphisms:

$$\begin{tikzcd}
Theorem\ 3.5 \arrow[d, Rightarrow] & \\
Corollary\ 3.6 \arrow[d, Rightarrow] & \\
Theorem\ 3.10 \arrow[d, Rightarrow] & Lemma\ 3.8 \arrow[l, Rightarrow] \\
Corollary\ 3.13 &
\end{tikzcd}$$

Fig. 3.1 Logical relations of the results in Section 3.1

3.2 Lie's theorem for solvable Lie algebras

Solvability generalizes nilpotency. Solvable Lie algebras relate to nilpotent Lie algebras like upper triangular matrices relate to strictly upper triangular matrices.

Definition 3.15 (Derived series and solvable Lie algebra). Consider a Lie algebra L.

1. The *derived series* of L is the sequence $(D^i L)_{i \in \mathbb{N}}$ inductively defined as

$$D^0 L := L \textit{ and } D^{i+1} L := [D^i L, D^i L],\ i \in \mathbb{N}.$$

2. The Lie algebra L is *solvable*(Deutsch: *aufloesbar*) iff an index $i_0 \in \mathbb{N}$ exists with $D^{i_0} L = 0$. The smallest index with this property is named the *length* of the derived series.

3. An ideal $I \subset L$ is *solvable* iff I is solvable when considered as Lie algebra.

By induction on $i \in \mathbb{N}$ one easily shows that each $D^iL, i \in \mathbb{N}$, is an ideal in L, and that the series $(D^iL)_{i\in\mathbb{N}}$ is descending. Note that

$$D^1L = [L,L]$$

is the derived or commutator algebra. Comparing the derived series with the lower central series one has $D^iL \subset C^iL$ for all $i \in \mathbb{N}$. Hence for Lie algebras:

$$Abelian \implies nilpotent \implies solvable.$$

Solvability behaves well with respect to short exact sequences.

Lemma 3.16 (Solvability and short exact sequences). *Consider an exact sequence of Lie algebra morphisms*

$$0 \to L_0 \to L_1 \xrightarrow{\pi} L_2 \to 0.$$

Then solvability of L_1 is equivalent to solvability of L_0 and L_2.

Proof. i) *L_0 and L_2 solvable $\implies$ L_1 solvable*: Assume

$$D^{i_0}L_2 = 0 \text{ and } D^{i_0}L_0 = 0.$$

Note that one may find a common index i_0 for both Lie algebras. For all $i \in \mathbb{N}$

$$D^iL_2 = D^i\pi(L_1) = \pi(D^iL_1).$$

Hence $D^{i_0}L_2 = 0$ implies $D^{i_0}L_1 \subset L_0$. Then

$$D^{2\cdot i_0}L_1 \subset D^{i_0}L_0 = 0.$$

ii) *L_1 solvable $\implies$ L_0 and L_2 solvable*: The proof uses the following relations between the derived series:

$$D^iL_0 \subset D^iL_1,\ \pi(D^iL_1) = D^iL_2.$$

□

Example 3.9 demonstrates that an analogous statement concerning the nilpotency of the Lie algebra in the middle is not valid.

Corollary 3.17 (Solvable ideals). *Consider a Lie algebra L. The sum $I+J$ of two solvable ideals $I,\ J \subset L$ is solvable.*

Proof. Lemma 3.16 applied to the exact sequence

$$0 \to I \cap J \to I \to I/(I \cap J) \to 0$$

shows the solvability of the quotient

$$I/(I \cap J)$$

The canonical isomorphy

$$I/(I \cap J) \simeq (I+J)/J$$

and Lemma 3.16 applied to the exact sequence

$$0 \to J \to I+J \to (I+J)/J \to 0$$

shows the solvability of

$$I+J.$$

□

Definition 3.18 (Radical of a Lie algebra). The unique solvable ideal of a Lie algebra L which is maximal with respect to all solvable ideals of L, is the *radical* of L, denoted *rad* L.

To verify that the concept is well-defined consider a maximal solvable ideal

$$I_{max} \subset L.$$

For an arbitrary solvable ideal $I \subset L$ also

$$I + I_{max}$$

is solvable by Corollary 3.17. The inclusion

$$I_{max} \subset I + I_{max}$$

implies by the maximality of I_{max} that

$$I_{max} = I + I_{max}.$$

Hence $I \subset I_{max}$. Therefore I_{max} is the uniquely determined, maximal solvable ideal of L.

For a *nilpotent* embedded Lie algebra Theorem 3.5 always deals with the same eigenvalue zero. In the more general context of a *solvable* embedded Lie algebra L the eigenvalues of the common eigenvector depend linearly on the elements of L.

The following Lemma 3.19 prepares the proof of Theorem 3.20 and therefore also of Lie's theorem about the existence of a common eigenvector for embedded solvable Lie algebras. The lemma makes the assumption that a common eigenvector exists. It investigates: How do the corresponding eigenvalues for the different endomorphism from an ideal depend on the endomorphisms, in particular which endomorphism annihilate the eigenvector?

Lemma 3.19 (Dynkin lemma). *Consider a $\mathbb{K}$-vector space V, an embedded Lie algebra $L \subset gl(V)$, and an ideal $I \subset L$. Assume the existence of a common eigenvector $v \in V$ for all endomorphisms*

$$x : V \to V,\ x \in I,$$

and consider the corresponding linear functional

$$\lambda : I \to \mathbb{K}$$

satisfiying for all $x \in I$

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

Then λ vanishes on commutators, i.e.

$$\lambda|[I,I] = 0,$$

and

$$U := \{v \in V :\ x.v = \lambda(x) \cdot v \text{ for all } x \in I\}$$

is an L-module.

Proof. Apparently the eigenvector equation defines a linear functional λ. Let $y \in L$ be an arbitrary but fixed element. We have to show: For all $x \in I$ holds

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

i) *Simultaneous triangularization on a stable subspace $W = W(y) \subset V$*: Let $n \in \mathbb{N}^*$ be a maximal exponent such that the family of iterates

$$\mathscr{B} = (\mathrm{v}, y(\mathrm{v}), y^2(\mathrm{v}), ..., y^{n-1}(\mathrm{v}))$$

is linearly independent. Denote by

$$W := span_{\mathbb{K}} < y^i(\mathrm{v}) :\ i = 0, ..., n-1 >$$

the n-dimensional subspace of V spanned by $\mathscr{B}$. By definition W is stable with respect to y, i.e. $y(W) \subset W$.

Claim: The family

$$(W_i := span < \mathrm{v}, y(\mathrm{v}), \dots, y^{i-1}(\mathrm{v}) >)_{i=0,\dots,n},\ \text{ with } W_0 = \{0\},$$

is a flag of W, stable with respect to all endomorphisms

$$x : V \to V,\ x \in I$$

We prove $x(W_i) \subset W_i$ by induction on $i = 0, ..., n$: The induction claim holds for $i = 0$ and $i = 1$. For the induction step $i \mapsto i+1$ we consider $x \in I$ and decompose

$$x(y^i(\mathrm{v})) = (x \circ y)(y^{i-1}(\mathrm{v})) = [x,y](y^{i-1}(\mathrm{v})) + (y \circ x)(y^{i-1}(\mathrm{v})).$$

The induction assumption applied to $y^{i-1}(\mathrm{v}) \in W_i$ proves: For the first summand

$$[x,y](y^{i-1}(\mathrm{v})) \in \{z(W_i) : z \in I\} \subset W_i \subset W_{i+1}$$

because $y^{i-1}(\mathrm{v}) \in W_i$ and $[x,y] \in I$, and for the second summand

$$(y \circ x)(y^{i-1}(\mathrm{v})) = y(x(y^{i-1}(\mathrm{v}))) \in y(W_i) \subset W_{i+1}.$$

Hence $x(y^i(\mathrm{v})) \in W_{i+1}$.

Therefore each restricted endomorphism

$$x|W : W \to W,\ x \in I,$$

is represented with respect to the basis $\mathscr{B}$ of W by an upper triangular matrix

$$A_x = \begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix} \in \mathfrak{t}(n, \mathbb{K}).$$

ii) *Diagonal elements of the triangularization*: We claim that for each $x \in I$ all eigenvalues of $x|W$ are equal to $\lambda(x)$, i.e.

$$A_x = \begin{pmatrix} \lambda(x) & & * \\ & \ddots & \\ 0 & & \lambda(x) \end{pmatrix}$$

To prove the claim we have to shows for each $i = 0, ..., n-1$: The vector $y^i(\mathrm{v}) \in W_{i+1}$ satisfies

$$x(y^i(\mathrm{v})) - \lambda(x) \cdot y^i(\mathrm{v}) \in W_i$$

The proof is by induction on $i = 0, ..., n-1$. The induction start for $i = 0$ is the eigenvalue equation

$$x(\mathrm{v}) = \lambda(x) \cdot \mathrm{v}.$$

For the induction step $i-1 \mapsto i$ consider the decomposition from part i)

$$x(y^i(\mathrm{v})) = [x,y](y^{i-1}(\mathrm{v})) + (y \circ x)(y^{i-1}(\mathrm{v})).$$

- For the first summand we showed in part i)

$$w_i := [x,y](y^{i-1}(\mathrm{v})) \in W_i$$

- Concerning the second summand we apply to the induction assumption

$$x(y^{i-1}(\mathrm{v})) - \lambda(x) \cdot y^{i-1}(\mathrm{v}) \in W_{i-1}$$

the endomorphism $y \in L$, and obtain

$$(y \circ x)(y^{i-1}(\mathrm{v}) - \lambda(x) \cdot y^{i}(\mathrm{v}) \in y(W_{i-1}) \subset W_i$$

Taken together, both steps imply the induction claim for the exponent i:

$$x(y^{i}(\mathrm{v})) - \lambda(x) \cdot y^{i}(\mathrm{v}) = w_i + (y \circ x)(y^{i-1}(\mathrm{v})) - \lambda(x) \cdot y^{i}(\mathrm{v}) \in W_i$$

iii) *Vanishing of the trace of a commutator*: Part ii) shows: All elements $x \in I$ act on the subspace $W \subset V$ as endomorphisms with

$$trace(x|W) = n \cdot \lambda(x).$$

For all $x \in I$ also $[y,x] \in I$ because $I \subset L$ is an ideal. The trace of a commutator of two endomorphisms vanishes. Hence for all $x \in I$

$$0 = tr([y|W, x|W]) = n \cdot \lambda([y,x])$$

which proves $\lambda([y,x]) = 0$. Because $y \in L$ is arbitrary the proof shows

$$\lambda|[I,I] = 0.$$

iv) *Eigenspace is L-module*: For arbitrary $u \in U$, $x \in I$ and $y \in L$ we have to show

$$x.(y.u) = \lambda(x) \cdot (y.u)$$

One has due to part iii)

$$x.(y.u) = [x,y].u + y.(x.u) = \lambda([x,y]) \cdot u + y.(\lambda(x) \cdot u) = \lambda(x) \cdot (y.u)$$

□

The present section deals with eigenvalues of certain endomorphisms. We need the fact that a polynomial with coefficients from $\mathbb{K}$ splits completely over $\mathbb{K}$ into a product of linear factors. Therefore we assume that the underlying field $\mathbb{K}$ is algebraically closed, i.e. we now consider complex Lie algebras.

Theorem 3.20 proves the existence of a common eigenvector for the endomorphisms of an embedded, complex solvable Lie algebra. It is an analogue of Theorem 3.5 for the nilpotent case.

Theorem 3.20 (Common eigenvector of a solvable embedded complexLie algebra).

Consider a complex vector space $V \neq \{0\}$. *Each solvable embedded Lie algebra* $L \subset gl(V)$ *has a common eigenvector: A non-zero vector* $v \in V$ *and a linear functional*

$$\lambda : L \to \mathbb{C}$$

exist such that

$$x(v) = \lambda(x) \cdot v \text{ for all } x \in L.$$

The proof of Theorem 3.20 is standard. It imitates the proof of Theorem 3.5, see [24, Chap.4.1], [13, Theor. 9.11].

Proof. The proof goes by induction on $dim\ L$. The claim trivially holds for $dim\ L = 0$. For the induction step we assume $dim\ L > 0$, hence $L \neq \{0\}$. We suppose that the claim is true for all solvable Lie algebras of smaller dimension.

i) *Construction of an ideal* $I \subset L$ *of codimension* $codim_L(I) = 1$: The derived series of L ends with $D^{i_0}L = 0$, hence it starts with

$$D^1 L \subsetneq D^0 L, \text{ i.e. } L/[L,L] \neq 0.$$

Let

$$\pi : L \to L/[L,L]$$

be the canonical projection of Lie algebras. The Lie algebra $L/[L,L]$ is Abelian. Therefore any arbitrary choosen vector subspace

$$D \subset L/[L,L]$$

of codimension 1 is even an ideal. Then

$$I := \pi^{-1}(D)$$

is an ideal of L with

$$codim_L(I) := dim\ L - dim\ I = 1.$$

The last formula about the codimension is a general statement from the theory of vector spaces:

$$dim\ I = dim\ \pi^{-1}(D) = dim\ D + dim\ [L,L]$$

implies

$$codim_L\ I = dim\ L - dim\ I = dim\ L - dim\ [L,L] - dim\ D =$$

$$= dim\ L/[L,L] - dim\ D = codim_{L/[L,L]} D = 1.$$

ii) *Subspace of eigenvector candidates*: The action of L on V restricts to an action of I on V. The induction assumption applied to I provides an element $0 \neq \mathrm{v} \in V$ and a linear functional $\lambda : I \to \mathbb{C}$ with

$$x(\mathrm{v}) = \lambda(x) \cdot \mathrm{v} \ for\ all\ x \in I.$$

We consider the non-zero subspace of V of common eigenvectors for all endomorphisms of I

$$W := \{\mathrm{v} \in V : x(\mathrm{v}) = \lambda(x) \cdot \mathrm{v} \ for\ all\ x \in I\}.$$

The subspace $W \subset V$ is stable under the action of L, i.e. for $\mathrm{v} \in W$ and $y \in L$ holds $y(\mathrm{v}) \in W$: For arbitrary $x \in I, \mathrm{v} \in W, y \in L$ we have $[x,y] \in I$ and

$$x(y(\mathrm{v})) = [x,y](\mathrm{v}) + y(x(\mathrm{v})) = \lambda([x,y]) \cdot \mathrm{v} + \lambda(x) \cdot y(\mathrm{v}).$$

Here

$$[x,y](\mathrm{v}) = \lambda([x,y]) \cdot \mathrm{v}$$

because $[x,y] \subset I$.

Lemma 3.19, applied to $I \subset gl(W)$, gives $\lambda([x,y]) = 0$. Hence

$$x(y(\mathrm{v})) = \lambda(x) \cdot y(\mathrm{v})$$

and $y(\mathrm{v}) \in W$.

iii) 1-*dimensional complement*: According to the choice of the ideal $I \subset L$ in part i) each element $x_0 \in L \setminus I$ provides the vector space decomposition

$$L = I \oplus \mathbb{C} \cdot x_0.$$

By part ii) the subspace W is stable under the action of the restricted endomorphism $x_0|W$. Because the field $\mathbb{C}$ is algebraic closed, the restricted endomorphism $x_0|W \in End(W)$ has an eigenvector $\mathrm{v}_0 \in W$ of x_0 with eigenvalue λ':

$$x_0(\mathrm{v}_0) = \lambda' \cdot \mathrm{v}_0.$$

Due to the definition of W in part ii) the vector $\mathrm{v}_0 \in W$ is also an eigenvector of all endomorphisms from

$$L = I \oplus \mathbb{C} \cdot x_0.$$

The linear functional $\lambda : I \to \mathbb{C}$ extends to a linear functional $\widetilde{\lambda} : L \to \mathbb{C}$ by defining $\widetilde{\lambda}(x_0) := \lambda'$. Then

$$x(\mathrm{v}_0) = \widetilde{\lambda}(x) \cdot \mathrm{v}_0 \ for\ all\ x \in L,$$

which ends the induction step and completes the whole proof.

□

A corollary of Theorem 3.20 is Lie's theorem about the simultaneous triangularization of a complex solvable matrix algebra.

Theorem 3.21 (Lie's theorem on complex solvable embedded Lie algebras).

Consider a n-dimensional complex vector space V. Each solvable embedded Lie algebra $L \subset gl(V)$ is isomorphic to a Lie subalgebra of the Lie algebra of upper triangular matrices

$$\mathfrak{t}(n,\mathbb{C}) = \left\{ \begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix} \in gl(n,\mathbb{C}) \right\},$$

more precisely: There exists an invertible matrix

$$S \in GL(V) \simeq GL(n,\mathbb{C})$$

such that the conjugation satisfies

$$L' := S \cdot L \cdot S^{-1} \subset \mathfrak{t}(n,\mathbb{C}).$$

Proof. Similar to the proof of Corollary 3.6 we construct by induction on $dim\ V$ a flag $(V_i)_{i=0,...,n}$ of V, each V_i stable with respect to all endomorphisms

$$x : V \to V,\ x \in L.$$

To start we set $V_0 := \{0\}$. Theorem 3.20 provides a common eigenvector $\mathrm{v}_1 \in V$, satisfying

$$x(\mathrm{v}_1) = \lambda(x) \cdot \mathrm{v}_1$$

for all endomorphisms $x \in L$. Set

$$V_1 := \mathbb{C} \cdot \mathrm{v}_1 \text{ and } \overline{V} := V/V_1$$

We consider for all $x \in L$ the induced endomorphisms

$$\overline{x} : \overline{V} \to \overline{V}$$

and apply the induction assumption to the vector space $\overline{V}$ with

$$dim_{\mathbb{C}}\ \overline{V} < dim_{\mathbb{C}}\ V :$$

There exists a basis $(\overline{\mathrm{v}}_j)_{j=2,...,n}$ of $\overline{V}$ with triagonalizes all endomorphims

$$\overline{x} : \overline{V} \to \overline{V}$$

For $j = 2,...,n$ we choose elements $\mathrm{v}_j \in V$ with

$$\overline{\mathrm{v}}_j = v_j + V_1$$

The family $(\mathrm{v}_j)_{j=1,...,n}$ is a basis of V: A given vector $\mathrm{v} \in V$ has the residue class

$$\pi(\mathrm{v}) = \sum_{j=2}^{n} (\alpha_j \cdot (\mathrm{v}_j + V_1)) = \sum_{j=2}^{n} (\alpha_j \cdot \mathrm{v}_j + V_1) = \left(\sum_{j=2}^{n} \alpha_j \cdot \mathrm{v}_j \right) + V_1$$

Hence

$$\mathrm{v} - \sum_{j=2}^{n} \alpha_j \cdot \mathrm{v}_j \in V_1$$

i.e.

$$\mathrm{v} = \left(\sum_{j=2}^{n} \alpha_j \cdot \mathrm{v}_j \right) + \alpha_1 \cdot \mathrm{v}_1$$

with complex coefficients

$$\alpha_j,\ j = 1, ..., n.$$

Hence

$$dim_{\mathbb{C}} < \mathrm{v}_1, ..., \mathrm{v}_n >= dim_{\mathbb{C}}\ V$$

We set for each $j = 2, ..., n$

$$V_j := span_{\mathbb{C}} < v_1, ..., v_j >$$

Then the family $(V_j)_{j=1,...,n}$ is a flag of V, and each V_j is stable with respect to all endomorphisms of L. $\square$

Corollary 3.22 (Solvability and nilpotency). *Consider a complex Lie algebra L. Then L is solvable iff its commutator algebra*

$$DL := [L, L] = D^1 L$$

is nilpotent.

Proof. The derived series of L relates to the lower central series of the Lie algebra DL as

$$D^i L \subset C^{i-1}(DL)$$

for all $i \geq 1$. The proof goes by induction on $i \in \mathbb{N}^*$. The inclusion holds for $i = 1$. Induction step $i \mapsto i+1$:

$$D^{i+1} L := [D^i L, D^i L] \subset [C^{i-1}(DL), C^{i-1}(DL)] \subset [DL, C^{i-1}(DL)] = C^i(DL).$$

i) Assume that DL is nilpotent. The vanishing of $C^{i_0}(DL)$ for an index $i_0 \in \mathbb{N}$ implies the vanishing of $D^{i_0+1}L$. Hence L is solvable.

ii) Assume that L is solvable and $dim\ L = n$. Applying Lemma 3.16 to the exact sequence of Lie algebras

$$0 \to Z(L) \to L \xrightarrow{ad} ad\ L \to 0$$

proves that $ad\ L$ is solvable. Lie's theorem, see Theorem 3.21, implies for the embedded Lie algebra $ad\ L$

$$ad\ L \subset \mathfrak{t}(n,\mathbb{C}).$$

If $x \in DL = [L,L]$ then

$$ad\ x \in [\mathfrak{t}(n,\mathbb{C}),\mathfrak{t}(n,\mathbb{C})] \subset \mathfrak{n}(n,\mathbb{C}).$$

By Engel's theorem, see Theorem 3.10, the Lie algebra DL is nilpotent. □

Note that Corollary 3.22 is also valid for an $\mathbb{R}$-Lie algebra L, because the complexification satisfies

$$D^iL \otimes_{\mathbb{R}} \mathbb{C} = D^i(L \otimes_{\mathbb{R}} \mathbb{C}),$$

and similarly for the descending central series.

3.3 Semidirect product of Lie algebras

The semidirect product of two Lie algebras is a Lie algebra structure on the Cartesian product of the vector spaces of the two Lie algebras. The most simple case of a semidirect product is the direct product. It is obtained by taking the Lie bracket in each component separately. But using certain derivations one can define a Lie bracket which mixes the Lie brackets of the components, see [4, §1.8]. The semidirect product is an important tool to construct new Lie algebras from given ones, and also for splitting Lie algebras into factors of smaller Lie algebras.

Definition 3.23 (Semidirect product). Consider two Lie algebras I and M with Lie brackets respectively $[-,-]_I$ and $[-,-]_M$, together with a morphism of Lie algebras to the Lie algebra of derivations

$$\theta : M \to Der(I).$$

The *semidirect product* of I and M with respect to θ is defined as

$$I \rtimes_\theta M := (L,[-,-])$$

with vector space $L := I \times M$ and bracket

$$[-,-] : L \times L \to L$$

$$[(i_1,m_1),(i_2,m_2)] := ([i_1,i_2]_I + \theta(m_1)(i_2) - \theta(m_2)(i_1),[m_1,m_2]_M)$$

for $i_1,i_2 \in I, m_1,m_2 \in M$.

According to Definition 3.23 the semidirect product $I \rtimes_\theta M$ differs from the direct product

$$I \times M,$$

which has the Lie bracket

$$[[i_1, i_2]_I, [m_1, m_2]_M],$$

by the additional summand of the first component

$$\theta(m_1)(i_2) - \theta(m_2)(i_1) \in I$$

The semidirect product captures how M acts on I via θ, and the direct product is the particular case $\theta = 0$.

Proposition 3.24 (Semidirect product). *Consider two Lie algebras I and M, and a morphism of Lie algebras*

$$\theta : M \to Der(I).$$

1. *The semidirect product*

$$I \rtimes_\theta M$$

is a Lie algebra.

2. *The semidirect product fits into the exact sequence of Lie algebra morphisms*

$$0 \to I \xrightarrow{j} I \rtimes_\theta M \xrightarrow{\pi} M \to 0$$

with $j(i) := (i, 0),\ \pi(i, m) := m$.

3. *The Lie algebra morphism π from part 2 has a section s, i.e. a morphism of Lie algebras exists*

$$M \xrightarrow{s} I \rtimes_\theta M$$

satisfying

$$\pi \circ s = id_M.$$

4. *We have*

$$I \simeq j(I) \subset I \rtimes_\theta M$$

an ideal, and

$$M \simeq s(M) \subset I \rtimes_\theta M$$

a subalgebra of $I \rtimes_\theta M$.

Proof. 1. *Lie bracket*: The Lie bracket is $\mathbb{K}$-bilinear and alternating. In order to verify for

$$x = (x_1, x_2), y = (y_1, y_2), z = (z_1, z_2) \in I \times M$$

the Jacobi identity in the form

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

we distinguish four cases:

- If $x,y,z\in I$ then the claim follows from the Jacobi identity of I.
- If $x,y\in I$ and $z\in M$ then in $I\rtimes_\theta M$

$$[x,y]=([x_1,y_1],0)\implies$$
$$[[x,y],z]=([[x_1,y_1],0]-\theta(z_2)([x_1,y_1]),0)=(-\theta(z_2)([x_1,y_1]),0),$$
$$[y,z]=[(y_1,0),(0,z_2)]=(-\theta(z_2)(y_1),0)\implies$$
$$[[y,z],x]=[-(\theta(z_2)(y_1),0),(x_1,0)]=([-\theta(z_2)(y_1),x_1],0),$$
$$[z,x]=[(0,z_2),(x_1,0)]=(\theta(z_2)(x_1),0)\implies$$
$$[[z,x],y]=[(\theta(z_2)(x_1),0),(y_1,0)]=([\theta(z_2)(x_1),y_1],0)$$

  The claim reduces to the claim that in I

$$-\theta(z_2)([x_1,y_1])+[-\theta(z_2)(y_1),x_1]+[\theta(z_2)(x_1),y_1]=0,$$

  i.e.

$$\theta(z_2)([x_1,y_1])=[\theta(z_2)(x_1),y_1]+[x_1,\theta(z_2)(y_1)]$$

  The latter equation holds because $\theta(z_2)$ is a derivation of I.

- If $x\in I$ and $y,z\in M$ then in $I\rtimes_\theta M$

$$[x,y]=(-\theta(y_2)(x_1),0);\ [[x,y],z]=(\theta(z_2)(\theta(y_2)(x_1)),0)$$
$$[y,z]=(0,[y_2,z_2]);\ [[y,z],x]=(\theta([y_2,z_2])(x_1),0)$$
$$[z,x]=(\theta(z_2)(x_1),0);\ [[z,x],y]=(-\theta(y_2)(\theta(z_2)(x_1)),0)$$

  The claim reduces to the claim in I

$$\theta(z_2)(\theta(y_2)(x_1))+\theta([y_2,z_2])(x_1)-\theta(y_2)(\theta(z_2)(x_1))=0,$$

  i.e.

$$\theta([y_2,z_2])(x_1)=\theta(y_2)(\theta(z_2)(x_1))-\theta(z_2)(\theta(y_2)(x_1))$$

  The latter equation holds because θ is a morphism of Lie algebras.

- If $x,y,z\in M$ then the claim follows from the Jacobi identity of M.

2. *Exact sequence*: The definition of $I\rtimes_\theta M$ shows that j and π are morphisms of Lie algebras. The exactness of the sequence is obvious: j is injective, π is surjective, and

$$im\ j=j(I)=ker\ \pi.$$

3. *Existence of a section*: The map

$$s : M \to I \rtimes_\theta M,\ m \mapsto (0,m),$$

is a morphism of Lie algebras, satisfying

$$\pi(s(m)) = \pi((m,0)) = m.$$

4. *Embedding I and M*: The kernel of a morphism of Lie algebras is an ideal:

$$ker\ \pi = im\ j = j(I) \simeq I.$$

The image of a morphism of Lie algebras is a subalgebra. Hence the injectivity of s implies that

$$s(M) \simeq M$$

is a subalgebra of $I \rtimes_\theta M$.

□

Note. The semidirect product $I \rtimes_\theta M$ reduces to the direct product of Lie algebras

$$id : I \rtimes_\theta M \simeq I \times M$$

if and only if $\theta = 0$.

Remark 3.25 (Product versus coproduct). See also [35].

1. The direct product $L_1 \times L_2$ of two $\mathbb{K}$-Lie algebras together with the two canonical projections

$$p_i : L_1 \times L_2 \to L_i,\ i = 1,2,$$

is the product in the category $\underline{Lie}_{\mathbb{K}}$ of $\mathbb{K}$-Lie algebras. The two embeddings

$$j_1 : L_1 \hookrightarrow L_1 \times L_2,\ x \mapsto (x,0) \text{ and } j_2 : L_2 \hookrightarrow L_1 \times L_2,\ x \mapsto (0,x)$$

satisfy

$$p_i \circ j_i = id_{L_i},\ i = 1,2.$$

2. The direct sum

$$(L_1 \oplus L_2, [-,-])$$

of two $\mathbb{K}$-Lie algebras L_1 and L_2 is the $\mathbb{K}$-vector space $L_1 \oplus L_2$ equipped with the Lie bracket

$$[X,Y] := \begin{cases} [X,Y]_{L_1} & \text{if } X,Y \in L_1 \\ [X,Y]_{L_2} & \text{if } X,Y \in L_2 \\ 0 & \text{if } X_1 \in L_1, X_2 \in L_2 \end{cases}$$

The direct sum $L_1 \oplus L_2$ together with the two injective morphisms of Lie algebras

$$j_i : L_i \hookrightarrow L_1 \oplus L_2,\ i = 1,2,$$

is not the coproduct of Lie algebras: Consider a non-Abelian Lie algebra L. Assume there exists a morphism

$$f : L \oplus L \to L$$

which renders commutative the following diagram

$$\begin{tikzcd} L \arrow[d, "j_1"'] \arrow[dr, "\mathrm{id}"] & \\ L \oplus L \arrow[r, dashed, "f"] & L \\ L \arrow[u, "j_2"] \arrow[ur, "\mathrm{id}"'] & \end{tikzcd}$$

Then for all $(z_1,z_2) \in L \oplus L$

$$f(z_1,z_2) = z_1 + z_2 \in L$$

which implies for

$$x = (x_1,x_2),\ y = (y_1,y_2) \in L \oplus L$$

on one hand,

$$f([x,y]) = f(([x_1,x_2],[y_1,y_2])) = [x_1,y_1] + [x_2,y_2].$$

On the other hand, f being a morphism satisfies

$$f([x,y]) = [f(x),f(y)] \text{ i.e. } f([x,y]) = [x_1+x_2,y_1+y_2] = [x_1,y_1]+[x_2,y_2]+[x_1,y_2]+[x_2,y_1],$$

a contradiction in the non-Abelian case.

3. Note that $\underline{Lie}_{\mathbb{K}}$ is not an Abelian category: For non-Abelian L the set of Lie algebra endomorphisms

$$Hom_{Lie_{\mathbb{K}}}(L,L)$$

is not even additively closed, because

$$id_L + id_L = 2 \cdot id_L \notin Hom_{Lie_{\mathbb{K}}}(L,L)$$

due to

$$(2 \cdot id_L)([x,y]) = 2 \cdot [x,y] \text{ while } [(2 \cdot id_L)(x),(2 \cdot id_L)(y)] = 4 \cdot [x,y]$$

A corner stone of quantum mechanics is the canonical commutation relation

$$[\mathrm{Q},\mathrm{P}] = \hbar \cdot \mathbb{1}$$

with Q the position operator, P the momentum operator, $\mathbb{1}$ the identity operator, and Planck's constant

$$\hbar = \frac{h}{2\pi} = 1.054 \times 10^{-27} erg \cdot sec.$$

858

Zur Quantenmechanik.

Von **M. Born** und **P. Jordan** in Göttingen.

(Eingegangen am 27. September 1925.)

Die kürzlich von Heisenberg gegebenen Ansätze werden (zunächst für Systeme von einem Freiheitsgrad) zu einer systematischen Theorie der Quantenmechanik entwickelt. Das mathematische Hilfsmittel ist die Matrizenrechnung. Nachdem diese kurz dargestellt ist, werden die mechanischen Bewegungsgleichungen aus einem Variationsprinzip abgeleitet und der Beweis geführt, daß auf Grund der Heisenbergschen Quantenbedingung der Energiesatz und die Bohrsche Frequenzbedingung aus den mechanischen Gleichungen folgen. Am Beispiel des anharmonischen Oszillators wird die Frage der Eindeutigkeit der Lösung und die Bedeutung der Phasen in den Partialschwingungen erörtert. Den Schluß bildet ein Versuch, die Gesetze des elektromagnetischen Feldes der neuen Theorie einzufügen.

...

festgelegt. Zusammenfassend erhalten wir unter Benutzung der durch (6) definierten Einheitsmatrix $\mathbf{1}$ die Gleichung

$$\boldsymbol{p}\boldsymbol{q} - \boldsymbol{q}\boldsymbol{p} = \frac{h}{2\pi i}\mathbf{1}, \tag{38}$$

die wir die „verschärfte Quantenbedingung" nennen und auf der alle weiteren Schlüsse beruhen.

Fig. 3.2 The birth of the strict quantum condition in [3]

This relation has been introduced by Born and Jordan and termed *strict quantum condition* (German: *verschärfte Quantenbedingung*), see Figure 3.2.

Proposition 3.26 (Heisenberg algebra of n-dimensional quantum mechanics). *The matrices from*

$$heis(n) := \left\{ \begin{pmatrix} 0 & p & c \\ 0 & 0_n & q \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{n}(n+2,\mathbb{R}) : p = (p_1,...,p_n) \in \mathbb{R}^n, q = (q_1,...,q_n)^\top \in \mathbb{R}^n, c \in \mathbb{R} \right\}$$

form a real $2n+1$*-dimensional Lie subalgebra of* $\mathfrak{n}(n+2,\mathbb{R})$, *named the* Heisenberg algebra *of n-dimensional quantum mechanics. In particular:*

$$heis(1) = \mathfrak{n}(3,\mathbb{R}).$$

A basis of $heis(n)$ *is the family*

$$(P_1,...,P_n,Q_1,...,Q_n,I)$$

with elements

- $P_j = E_{1,1+j}$*, the matrix with the only non-zero entry* $p_j = 1$*,*
- $Q_j = E_{1+j,n+2}$*, the matrix with the only non-zero entry* $q_j = 1$*, and*
- $I = E_{1,n+2}$*, the matrix with the only non-zero entry* $c = 1$

and the commutators

$$[P_j,Q_k] = \delta_{jk} \cdot I \text{ and } [P_j,I] = [Q_j,I] = [P_j,P_k] = [Q_j,Q_k] = 0.$$

A typical element of $heis(n)$ looks like

$$\begin{pmatrix} 0 & p_1 & ... & p_j & ... & p_n & c \\ 0 & 0 & ... & 0 & ... & 0 & q_1 \\ \\ 0 & 0 & ... & 0 & ... & 0 & q_j \\ \\ 0 & 0 & ... & 0 & ... & 0 & q_n \\ 0 & 0 & ... & 0 & ... & 0 & 0 \end{pmatrix} \in \mathfrak{n}(n+2,\mathbb{R})$$

with

$$p = (p_1,...,p_n),\ q = (q_1,...,q_n)^\top,$$

see also [49, Chap. 13.1]. For further information on a graduate level about $heis(n)$ see [29, Chap. 2].

Proof. The commutator relations follow from the commutator formula

$$[E_{i,j},E_{s,t}] = \delta_{js} \cdot E_{i,t} - \delta_{ti} \cdot E_{s,j},$$

for example

$$[P_j,Q_k] = [E_{1,1+j},E_{1+k,n+2}] = E_{1,1+j} \cdot E_{1+k,n+2} - E_{1+k,n+2} \cdot E_{1,1+j} = \delta_{jk} \cdot E_{1,n+2}$$

because $n+2 \neq 1$. Hence $heis(n) \subset \mathfrak{n}(n+2,\mathbb{R})$ is closed with respect to the Lie bracket. $\square$

The Heisenberg algebra is nilpotent, being a subalgebra of the nilpotent Lie algebra $\mathfrak{n}(n+2,\mathbb{R})$ of strictly upper triangular matrices.

The Heisenberg algebra captures the *kinematics* of an n-dimensional quantum mechanical problem. For a complete description which also covers the *dynamics* of the problem one needs a further operator: The Hamiltonian H of the problem and its relation to the kinematical operators.

How to extend the Heisenberg algebra to include also the Hamiltonian H?

The solution is to construct the *dynamical* Lie algebra of quantum mechanics as the semidirect product of the Heisenberg algebra and the 1-dimensional Lie algebra generated by the Hamiltonian. For the case of 1-dimensional quantum mechanics see also [23, Example 5.1.19].

Definition 3.27 (Dynamical Lie algebra of quantum mechanics). For $n \in \mathbb{N}$ consider the Heisenberg algebra $heis(n)$ of n-dimensional quantum mechanics with basis

$$(P_1,...,P_n,Q_1,...,Q_n,I),$$

the 1-dimensional Abelian Lie algebra $\mathbb{R}$ with basis (H), and the morphism of Lie algebras

$$\theta : \mathbb{R} \to Der(heis(n))$$

to the Lie algebra of derivations, defined as

$$\theta(H)(P_j) := Q_j,\ \theta(H)(Q_j) := -P_j,\ j = 1,...,n, \text{ and } \theta(H)(I) := 0.$$

The semidirect product

$$quant(n) := heis(n) \rtimes_\theta \mathbb{R}$$

is the *dynamical Lie algebra of n-dimensional quantum mechanics*. It fits into the short exact sequence

$$0 \to heis(n) \xrightarrow{j} quant(n) \xrightarrow{\pi} \mathbb{R} \to 0.$$

Lemma 3.28 (Dynamical Lie algebra of quantum mechanics). *The dynamical Lie algebra of n-dimensional quantum mechanics quant(n) is a well-defined real Lie algebra. Identifying heis(n) with the ideal*

$$j(heis(n)) \subset heis(n) \rtimes_\theta \mathbb{R}$$

and $\mathbb{R}$ *with the subalgebra*

$$s(\mathbb{R}) \subset heis(n) \rtimes_\theta \mathbb{R}$$

the distinguished commutators of $quant(n)$ *are*

$$[H,P_j] = Q_j,\ [H,Q_j] = -P_j,\ j = 1,...,n,\ [H,I] = 0.$$

Proof. i) *The endomorphism* $\theta(H)$ *is a derivation*: Using the shorthand $D := \theta(H)$:

- Left-hand side:
$$D([P_j,P_k]) = D(0) = 0$$
Right-hand side:
$$[D(P_j),P_k] + [P_j,D(P_k)] = [Q_j,P_k] + [P_j,Q_k] = -\delta_{jk} \cdot I + \delta_{jk} \cdot I = 0.$$
- Left-hand side:
$$D([Q_j,Q_k]) = D(0) = 0$$
Right-hand side:
$$[D(Q_j),Q_k] + [Q_j,D(Q_k)] = [-P_j,Q_k] + [Q_j,-P_k] = -\delta_{jk} \cdot I + \delta_{jk} \cdot I = 0.$$
- Left-hand side:
$$D([P_j,Q_k]) = D(\delta_{jk} \cdot I) = 0$$
Right-hand side:
$$[D(P_j),Q_k] + [P_j,D(Q_k)] = [Q_j,Q_k] + [P_j,-P_k] = 0.$$
- Left-hand side:
$$D([I,P_j]) = D(0) = 0$$
Right-hand side:
$$[D(I),P_j] + [I,D(P_j)] = 0 + [I,Q_j] = 0$$
- Left-hand side:
$$D([I,Q_j]) = D(0) = 0$$
Right-hand side:
$$[D(I),Q_j] + [I,D(Q_j)] = 0 - [I,P_j] = 0.$$

ii) *The map* θ *is a morphism of Lie algebras*: Because $\mathbb{R}$ is 1-dimensional and Abelian, the claim follows from $[D,D] = 0$.

iii) *Computing the distinguished commutators*: For $j = 1, ..., n$ the distinguished commutators follow from the definition of the Lie bracket of a semidirect product. In

$$quant(n) := heis(n) \rtimes_\theta \mathbb{R}$$

holds

$$[(0,H),(P_j,0)] = ([0,P_j] + D(P_j),[H,0]) = (Q_j,0)$$
$$[(0,H),(Q_j,0)] = ([0,Q_j] + D(Q_j),[H,0]) = (-P_j,0)$$
$$[(0,H),(I,0)] = (D(I),0) = (0,0),$$

i.e. under the identification of

$$heis(n) \text{ with } j(heis(n)) \subset quant(n) \text{ and of } \mathbb{R} \text{ with } s(\mathbb{R}) \subset quant(n)$$

holds

$$[H,P_j] = Q_j,\ [H,Q_j] = -P_j,\ [H,I] = 0.$$

□

Proposition 3.29 (Solvability of $quant(n)$**).** *The dynamical Lie algebra of quantum mechanics*

$$quant(n) := heis(n) \rtimes_\theta \mathbb{R}$$

is solvable, but not nilpotent. Its derived algebra is

$$D^1 quant(n) = heis(n).$$

Proof. The descending central series of

$$L := quant(n) = span_{\mathbb{R}} < P_j, Q_j, I, H : \ j = 1, .., n >$$

is

$$C^1 L := span_{\mathbb{R}} < P_j, Q_j, I : \ j = 1, .., n)$$

and

$$C^2 L = [L, C^1 L] = span_{\mathbb{R}} < P_j, Q_j, I : \ j = 1, .., n) = C^1 L$$

Hence for all $i \in \mathbb{N}$

$$C^i L \neq \{0\},$$

which shows that $quant(n)$ is not nilpotent. Corollary 3.22 and the subsequent note for the base field $\mathbb{R}$ imply that $quant(n)$ is solvable. □

Remark 3.30 (Heisenberg picture and Schroedinger picture).

1. Proposition 3.26 and Definition 3.27 introduce the Heisenberg Lie algebra $heis(n)$ respectively the Lie algebra $quant(n)$ of n-dimensional quantum mechanics as abstract mathematical objects.

The role of these Lie algebras for quantum mechanics results from representing their elements as selfadjoint operators on complex Hilbert spaces, which are adapted to the particular physical system. Only after applying a representation the Lie algebra gets a meaning for physics. In the most simple case of 1-dimensional quantum mechanics the representation ρ maps the elements of

$$heis(1) = span_{\mathbb{R}} < Q,\ P,\ I >$$

to selfadjoint operators, defined on functions f from a dense subspace of $L^2(\mathbb{R})$, see [27, Chap. 3.6]:

$$\rho(P)(f(x)) := x \cdot f(x) \text{ (multiplication)}$$

$$\rho(Q)(f) := -i \cdot \frac{df}{dx} \text{ (derivation)}$$

$$\rho(I) := -i \cdot id \text{ (scalar)}$$

Here the definition of $\rho(I)$ follows from the canonical commutator relation

$$[P,Q] = I$$

because it forces

$$\rho(I) = \rho([P,Q]) = [\rho(P), \rho(Q)] = -i \cdot id.$$

The canonical commutator relations have no non-zero representation on a *finite-dimensional* vector space V: Taking the trace of a matrix equation

$$\rho(I) = \rho([P,Q]) = [\rho(P), \rho(Q)] = -i \cdot id$$

shows

$$0 = trace[\rho(P), \rho(Q)] = trace(-i \cdot id) = -i \cdot trace\ id_V$$

and implies $V = \{0\}$.

2. Like classical mechanics also quantum mechanics distinguishes between states and observables to describe a physical system, cf. [37]. Pure states, the most simple states, are represented by the 1-dimensional subspaces of a complex Hilbert space $Hilb$ with a Hermitian product $< -,- >$, and the observables are selfadjoint operators on $Hilb$. The expectation value of measuring the observable Ω when the system has been prepared in state ϕ is

$$< \phi|\Omega|\phi > := < \phi, \Omega(\phi) > = < \Omega\phi, \phi > \in \mathbb{R}.$$

3. A pure state ϕ is given by a function $\psi = \psi(t)$, and the temporal development of the system, i.e. its dynamics, is governed by the Hamiltonian H of the system according to the *Schrödinger equation*, the ordinary differential equation

$$H(\psi(t)) = i \cdot \dot{\psi}(t), \text{ normalization: } \hbar = 1.$$

The Hamiltonian H is the observable of the total energy of the system. The *Schrödinger picture* considers the temporal development of the states and fixes the observables. For conservative systems the Hamiltonian has no explicit time dependency.

4. The *Heisenberg picture* takes the opposite approach: It fixes the states and considers the temporal development of the observables. This approach resembles the Hamiltonian approach of classical mechanics using canonical coordinates. In the Heisenberg picture, the temporal development of an observable Ω is governed by the commutator equation

$$\frac{1}{i}\,\dot{\Omega}(t) = [H, \Omega(t)] + \frac{\partial}{\partial t}\,\Omega(t).$$

Here the partial derivation $\dfrac{\partial}{\partial t}\,\Omega$ vanishes in case Ω does not explicitly depend on the time t. In particular, for a set of canonical observables $\mathrm{P}_j,\ \mathrm{Q}_j,\ j = 1, ..., n$:

$$\frac{1}{i}\,\dot{\mathrm{Q}}_j(t) = [H, \mathrm{Q}_j(t)] = -\mathrm{P}_j(t)$$

$$\frac{1}{i}\,\dot{\mathrm{P}}_j(t) = [H, \mathrm{P}_j(t)] = \mathrm{Q}_j(t).$$

A simple example of a 1-dimensional quantum system is the 1-dimensional oscillator. It has the Hamiltonian

$$H = \frac{1}{2m}\,\mathrm{P}^2 + \frac{m\omega^2}{2}\,\mathrm{Q}^2$$

with respect to the momentum observable P and the position observable Q. The real number ω denotes the oscillator frequency, the real number m denotes the oscillator mass.

The *mathematical formalism* of quantum mechanics is well understood: The theory of self-adjoint operators in a Hilbert space forms part of the domain of functional analysis. One needs the existence of the spectral representation.
But the *interpretation* of the physical content, even more the implications considered from the viewpoint of *philosophy of nature* are still debated.
Historically, the first elaborated interpretation of quantum mechanics was the Copenhagen interpretation, see [20]. Its main thesis: Observables of quantum mechanical systems get a specific value not until the moment of observation.

The Copenhagen interpretation is sharpened further in Rovelli's Relational Quantum Mechanics, see [39]: Observables get a specific value in the act of interaction of two physical systems, and this value has a relative meaning:

"Quantum mechanics is a theory about the physical description of physical system *relative* [emphasis JW] to other systems, and this is a complete description of the world."

Rovelli's formulation does not refer to any observation, in particular it does not refer to an observer. Concerning the Copenhagen interpretation Rovelli finishes his paper:

"Heisenberg's insistence on the fact that the lesson to be taken from the atomic experiments is that we should stop thinking of the *state of the system* has been obscured by the subsequent terse definition of the theory in terms of states given by Dirac. Here, I have taken Heisenberg's lesson to some extreme consequences."

Chapter 4
Killing form and semisimple Lie algebras

All vector spaces and Lie algebras are assumed finite-dimensional if not stated otherwise. The composition $f_2 \circ f_1$ of two endomorphisms will be denoted as product $f_2 \cdot f_1$ or also $f_2 f_1$.

4.1 The trace of endomorphisms

The present section introduces a powerful tool for the study of Lie algebras: Important properties of a Lie algebra L are encoded in a bilinear form, which derives from the trace of the endomorphisms of the adjoint representation of L.

Lemma 4.1 (Basic properties of the trace). *Consider a vector space V and endomorphisms $x, y, z \in End(V)$.*

1. *For nilpotent x holds*
$$tr\, x = 0.$$
2. *With respect to two endomorphisms the trace is symmetric, i.e.*
$$tr\ (xy) = tr\ (yx) \text{ or } tr\ [x,y] = 0.$$
3. *With respect to cyclic permutation the trace is invariant, i.e.*
$$tr\ (xyz) = tr\ (yzx)$$
4. *With respect to the commutator the trace is "associative"*
$$tr\ ([x,y]z) = tr\ (x[y,z]).$$

Proof. 1) All complex eigenvalues of a nilpotent endomorphism are zero.

2) With respect to matrix representations of the endomorphisms note

$$tr\ (xy) = \sum_{i,j} x_{ij} y_{ji} = \sum_{i,j} y_{ji} x_{ij} = tr\ (yx)$$

3) According to part 2) we have

$$tr(xyz) = tr(x(yz)) = tr((yz)x) = tr(yzx).$$

4) We have

$$[x,y]z = xyz - yxz \text{ and } x[y,z] = xyz - xzy.$$

The ordering yxz results from the ordering xzy by cyclic permutation. Hence part 3) implies

$$tr(yxz) = tr(xzy)$$

which proves the claim.

Definition 4.2 (Killing form). Let L be a $\mathbb{K}$-Lie algebra.

1. The *trace form* of a representation

$$\rho : L \to gl(V)$$

on a $\mathbb{K}$-vector space V is the symmetric bilinear map

$$\beta : L \times L \to \mathbb{K},\ \beta(x,y) := tr\ (\rho(x) \circ \rho(y)).$$

2. The *Killing form* of L

$$\kappa : L \times L \to \mathbb{K},\ \kappa(x,y) := tr\ (ad(x) \circ ad(y))$$

is the trace form of the adjoint representation $ad : L \to gl(L)$,

Theorem 4.3 (Cartan condition for the solvability of embedded Lie algebras). *For a $\mathbb{K}$-vector space V each Lie subalgebra $L \subset gl(V)$ of matrices with vanishing trace form, i.e. satisfying for all $x,\ y \in L$*

$$tr(x \circ y) = 0,$$

is solvable.

For the proof cf. [13, Theor. C.5].

Proof. A real Lie algebra is solvable if and only if its complexification is solvable. Therefore we may assume $\mathbb{K} = \mathbb{C}$ for the base field.

i) According to Corollary 3.22 it suffices to prove the nilpotency of the commutator algebra

$$DL = [L,L].$$

Therefore it is sufficient according to Corollary 3.6 to show, that each element

$$x \in DL \subset gl(V)$$

is a nilpotent endomorphism of V.

ii) Consider an arbitrary but fixed endomorphism $x \in [L,L]$. It's Jordan decomposition according to Theorem 1.19

$$x = x_s + x_n \in End\ V$$

provides a basis of V consisting of generalized eigenvectors of x. With respect to this basis x is an upper triangular matrix and its semisimple summand x_s is a diagonal matrix

$$x_s = diag(\lambda_1, ..., \lambda_n)$$

with each eigenvalue λ_j of x counted with its algebraic multiplicity. Define the complex conjugate diagonal matrix

$$\bar{x}_s = diag(\overline{\lambda}_1, ..., \overline{\lambda}_n)$$

Then the product

$$\bar{x}_s \circ x$$

is an upper triangular matrix with the values

$$|\lambda_j|^2,\ j = 1, ..., n,$$

on the main diagonal. Hence

$$tr(\bar{x}_s \circ x) = \sum_{j=1}^{n} |\lambda_j|^2 \geq 0$$

We have to show

$$tr(\bar{x}_s \circ x) = 0$$

By assumption $x \in [L,L]$ is a sum of commutators of the form $[y,z]$ with $y,\ z \in L$. Hence

$$tr(\bar{x}_s \circ x)$$

is a sum of terms of the form

$$tr(\bar{x}_s \circ [y,z]) = tr([\bar{x}_s, y] \circ z),$$

with the last equality due to Lemma 4.1.

iii) *Relating diagonal matrices A and $\overline{A}$*: According to the proof of Proposition 3.3 the matrices of $ad\ x_s$ and of $ad\ \overline{x}_s$ are diagonal with respect to the canonical basis $(E_{ij})_{1 \le i,j \le n}$ of $End(V)$:

$$ad\ x_s = diag\left((\lambda_i - \lambda_j)_{ij}\right) \text{ and } ad\ \overline{x}_s = diag\left((\overline{\lambda}_i - \overline{\lambda}_j)_{ij}\right)$$

We choose a polynomial $q \in \mathbb{C}[T]$ satisfying for all $1 \le i, j \le n$

$$q(\lambda_i - \lambda_j) = \overline{\lambda}_i - \overline{\lambda}_j$$

Applying the polynomial in the context of the two diagonal matrices shows

$$q(ad\ x_s) = ad\ \overline{x}_s$$

Because the Jordan decomposition of $ad\ x$ is obtained by applying ad to the Jordan decomposition of x, there exists a polynomial $s \in \mathbb{C}[T]$ with

$$ad\ x_s = s(ad\ x).$$

Hence

$$(ad\ x)(L) \subset L \implies (ad\ x_s)(L) \subset L \implies (ad\ \overline{x}_s)(L) \subset L$$

We obtain for each of the summands from part ii)

$$tr((ad\ \overline{x}_s)(y) \circ z) = 0$$

because by assumption

$$u := (ad\ \overline{x}_s)(y) \in L \implies tr(u \circ z) = 0$$

Collecting all summands gives

$$tr(\overline{x}_s \circ x) = 0$$

□

In Theorem 4.3 the converse implication does not hold: The Lie algebra $\mathfrak{t}(n, \mathbb{K})$ is solvable but does not fulfill the Cartan condition.

Corollary 4.4 (Cartan's characterization of solvability). *For a complex Lie algebra L are equivalent:*

- *The Lie algebra L is solvable.*

- *The Killing form κ of L satisfies*

$$\kappa(L, [L, L]) = 0.$$

Proof. • Assume L solvable. Applying Lemma 3.16 to the exact sequence

$$0 \to Z(L) \to L \xrightarrow{ad} ad(L) \to 0$$

shows that

$$ad(L) \subset gl(End\ L)$$

is solvable. Lie's theorem, Theorem 3.21, implies that the Lie algebra $ad(L)$ can be considered a matrix algebra of upper triangular matrices, and elements of

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

as strictly upper triangular matrices.

Set

$$DL := D^0 L = [L,L]$$

as a shorthand for the commutator algebra of L.

$$\kappa(L,DL) = \{tr(ad\ x \circ ad\ y) :\ x \in L,\ y \in DL\}$$

For $x \in L$ and $y \in DL$ the matrix product

$$ad\ x \circ ad\ y$$

is a strictly upper triangular matrix and therefore nilpotent. Then the Killing form annihilates the pair:

$$\kappa(x,y) := tr(ad\ x \circ ad\ y) = 0$$

Because $x \in L$ and $y \in [L,L]$ are arbitrary we obtain

$$\kappa(L,[L,L]) = 0$$

• Assume $\kappa(L,[L,L]) = 0$. Then in particular

$$\kappa([L,L],[L,L]) = \kappa(DL,DL) = 0.$$

Theorem 4.3 implies that the embedded Lie algebra $ad\ DL$ is solvable. The kernel

$$ker[ad : DL \to gl(DL)] = Z(DL)$$

is Abelian, hence solvable. Lemma 3.16 implies the solvability of DL. Eventually

$$DL = [L,L] \text{ solvable} \implies L \text{ solvable}$$

according to the definition of solvability via the derived series of L.

□

Note. Corollary 4.4 is valid also for a real Lie algebra L: For the necessary calculation with the Killing forms of L and $L \otimes_{\mathbb{R}} \mathbb{C}$ see [31, Proof of Prop. 1.46].

4.2 Fundamentals of semisimple Lie algebras

Definition 4.5 (Simple, semisimple and reductive Lie algebras). Consider a Lie algebra L.

1. L is *simple* iff $[L,L] \neq \{0\}$ and L has no ideal different than $\{0\}$ and L.
2. L is *semisimple* iff L has no Abelian ideal $I \neq \{0\}$.
3. L is *reductive* iff it splits as the direct sum
$$L = Z \oplus S$$
with an Abelian ideal $Z \subset L$ and a semisimple ideal $S \subset L$.

These concepts apply also to an ideal $I \subset L$ when the ideal is considered a Lie algebra.

Note: For each Lie algebra L the derived algebra $[L,L] \subset L$ is an ideal. Hence
$$[L,L] = L$$
for simple L. By definition the trivial Lie algebra $L = \{0\}$ is semisimple, but not simple. A semisimple Lie algebra $L \neq \{0\}$ is not Abelian. One has
$$simple \implies semisimple \implies reductive.$$

Remark 4.6 (Solvable Lie algebra, semisimple Lie algebra).

1. If a Lie algebra L has a solvable Ideal $I \neq \{0\}$ then L has also an Abelian ideal $\neq \{0\}$:

   Let $i \in \mathbb{N}$ be the largest index with $D^i I \neq \{0\}$ for the derived series of I. Then $D^i I \subset L$ is an Abelian ideal because
$$D^{i+1} I = [D^i I, D^i I] = \{0\}.$$
   The reverse implication is obvious: Each Abelian ideal of L is in particular a solvable ideal.

2. As a consequence for a Lie algebra L:

$$L \text{ semisimple} \iff rad(L) = \{0\}.$$

The equivalence indicates a complementarity between semisimplicity and solvability. The only Lie algebra which is both semisimple and also solvable is the zero Lie algebra $L = \{0\}$.

A simple application of the characterization of semisimpleness is Proposition 4.7.

Proposition 4.7 (Semisimple Lie algebras are embedded Lie algebras).

The adjoint representation

$$ad : L \to gl(L)$$

of a semisimple Lie algebra L is a faithful representation of L. In particular,

$$L \simeq ad(L) \subset Der(L) \subset gl(L)$$

represents L as a matrix algebra.

Proof. The kernel of the adjoint representation of L is the center of L, an Abelian ideal of L. Because L is semisimple, one concludes

$$Z(L) = \{0\}.$$

Remember that Proposition 2.8 demonstrates: The adjoint representation already maps to the Lie algebra of derivations of L. $\square$

Each Lie algebra becomes semisimple after dividing out its radical.

Proposition 4.8 (Semisimpleness after dividing out the radical). *For any Lie algebra L the quotient $L/rad(L)$ is semisimple.*

Proof. Consider the canonical projection of Lie algebras

$$\pi : L \to L/rad(L).$$

Each ideal

$$I \subset L/rad(L)$$

has the form

$$I = J/rad(L)$$

with the ideal

$$J := \pi^{-1}(I) \subset L$$

If I is an Abelian ideal, then $[I,I] = 0$, hence

$$[J,J] \subset rad(L)$$

Therefore the derived algebra

$$D^1 J = [J,J]$$

is solvable. The quotient $J/D^1 J$ is Abelian, hence solvable. Lemma 3.16 applies to the canonical exact sequence

$$0 \to D^1 J \to J \to J/D^1 J \to 0$$

and shows that J is a solvable ideal. Hence

$$J \subset rad(L)$$

and

$$I = \pi(J) = 0$$

□

According to Proposition 4.8 each Lie algebra L fits into an exact sequence of Lie algebras

$$0 \to I \xrightarrow{j} L \xrightarrow{\pi} M \to 0$$

with $I = rad(L)$, a solvable subalgebra, and $M = L/I$ a semisimple quotient. Levi's theorem provides a Lie algebra morphism

$$s : M \to L$$

which is a section against π. When identifying I with $j(I)$ the section s induces the Lie algebra morphism

$$\theta : M \to Der(I),\ \theta(m) := (ad\ s(m))|I,$$

to the derivations of I. One obtains L as the semidirect product

$$L \simeq I \rtimes_\theta M$$

For a proof of Levi's theorem see [4, Chap. I: §6.8 Theor. 5 and §1.8].

Lemma 4.9 (Radical). *Consider a Lie algebra L. Then for each ideal $L' \subset L$ the radical $rad\ L' \subset L$ is also an ideal.*

Proof. The quotient $L/rad\ L$ is semisimple due to Proposition 4.8. The inclusion

$$L' \subset L$$

implies the inclusion of Lie algebras

$$L'/(L' \cap rad\ L) \subset (L/rad\ L)$$

Due to Corollary 4.20 also

$$L'/(L' \cap rad\ L)$$

is semisimple. The vanishing of its radical implies

$$rad\ L' \subset (L' \cap rad\ L).$$

The right-hand side

$$L' \cap rad\ L$$

is a solvable ideal of L', hence contained in *rad* L'. We obtain:

$$rad\ L' = L' \cap\ rad\ L$$

is the intersections of two ideals in L, and therefore itself an ideal in L. $\square$

Definition 4.10 (Orthogonal space of a symmetric bilinear form).

Consider a $\mathbb{K}$-vector space V and a symmetric bilinear form

$$\beta : V \times V \to \mathbb{K}.$$

i) The *orthogonal space* with respect to β of a subspace $M \subset V$ is

$$M^{\perp} := \{x \in V : \beta(x, M) = 0\}.$$

The orthogonal space $V^{\perp}$ is named the *nullspace* of β or the *radical* of β.

ii) The form β is *non-degenerate* if its nullspace is trivial, i.e. $V^{\perp} = \{0\}$.

Remark 4.11 (Orthogonal space). Consider a $\mathbb{K}$-vector space V with a symmetric, non-degenerate bilinear form

$$\beta : V \times V \to \mathbb{K}.$$

Then β induces an isomorphism to the dual space

$$j : V \to V^*,\ j(\mathrm{v}) := \beta(\mathrm{v}, -).$$

For a subspace $M \subset V$ one obtains

$$M^{\perp} \simeq j(M^{\perp}) = \{\lambda \in V^* :\ \lambda | M = 0\} \simeq (V/M)^*$$

Lemma 4.12 (Orthogonal space of ideals). *Consider a vector space V, a Lie algebra $L \subset gl(V)$ and the trace form*

$$\beta : L \times L \to \mathbb{K},\ \beta(x,y) := tr(xy).$$

For an ideal $I \subset L$ the orthogonal space $I^{\perp} \subset L$ of β is also an ideal.

Proof. Consider $x \in I^{\perp}$. We have to show: For arbitrary $y \in L$ and all $u \in I$ holds

$$[y,x] \in I^{\perp},\ \text{i.e. } \beta([y,x],u) = 0$$

We have

$$-\beta([y,x],u) = \beta([x,y],u) = \beta(x,[y,u])$$

according to Lemma 4.1. Because $[y,u] \in I$ and $x \in I^{\perp}$

$$\beta(x,[y,u]) = 0.$$

Hence $[x,y] \in I^{\perp}$. □

The main step in characterizing semisimplicity of a Lie algebra by its Killing form is Proposition 4.13.

Proposition 4.13 (Non-degenerateness of the trace form of an embedded semisimple Lie algebra). *Consider a vector space V and a semisimple Lie algebra $L \subset gl(V)$. Then the trace form*

$$\beta : L \times L \to \mathbb{K},\ \beta(x,y) := tr(xy),$$

is non-degenerate.

Proof. The nullspace

$$S := L^{\perp} = \{x \in L : \ tr(xy) = 0 \ for\ all\ y \in L\}$$

is an ideal according to Lemma 4.12. We consider the Lie algebra S: By definition for all $x, y \in S$ holds

$$tr(x \circ y) = 0$$

Hence Cartan's trace condition, Theorem 4.3, shows that S is solvable. Hence

$$S \subset L$$

is a solvable ideal. Semisimpleness of L implies $S = \{0\}$. □

The main theorem of the present section is the following Cartan criterion for semisimplicity. It derives from Proposition 4.13. It is therefore a consequence of Cartan's trace criterion for solvability.

Theorem 4.14 (Cartan's characterization of semisimplicity). *For a Lie algebra L the following properties are equivalent:*

- *L is semisimple*

- *The Killing form κ of L is non-degenerate.*

Proof. i) Assume L semisimple. According to Proposition 4.7 the adjoint representations identifies L with the subalgebra

$$ad\ L \subset gl(L).$$

Therefore κ is non-degenerate according to Proposition 4.13.

ii) Assume κ non-degenerate, i.e. $L^{\perp} = 0$. Consider an Abelian ideal

$$I \subset L.$$

We claim $I \subset L^{\perp}$: For arbitrary, but fixed $x \in I$ and for all $y \in L$ the composition

$$ad(x) \circ ad(y) : L \to I \subset L$$

is a nilpotent endomorphism of L, because

$$L \xrightarrow{ad\ y} L \xrightarrow{ad\ x} I \xrightarrow{ad\ y} I \xrightarrow{ad\ x} [I,I] = \{0\}.$$

Because the trace of nilpotent endomorphisms vanishes we obtain

$$\kappa(x,y) = tr(ad(x) \circ ad(y)) = 0.$$

As a consequence $x \in L^{\perp}$. Because $x \in I$ was arbitrary, we obtain

$$I \subset L^{\perp}.$$

The inclusion implies $I = 0$ because κ is non-degenerate by assumption.
Therefore $\{0\}$ is the only Abelian ideal of L, and L is semisimple.
□

Lemma 4.15 (Killing form of an ideal). *Consider a $\mathbb{K}$-Lie algebra L. For any ideal $I \subset L$ the Killing form κ_I of I is the restriction of the Killing form κ of L to $I \times I$*

$$\kappa_I = \kappa|(I \times I) : I \times I \to \mathbb{K}.$$

Proof. Any base of the vector subspace $I \subset L$ extends to a base of L. Then for $x,\ y \in I$ the map

$$(ad\ x) \circ (ad\ y) : L \to I \subset L$$

has the matrix representation

$$\begin{pmatrix} A & B \\ 0 & 0 \end{pmatrix}.$$

And the matrix A represents the restriction

$$(ad\ x)|I \circ (ad\ y)|I : I \to I.$$

Hence

$$tr((ad\ x) \circ (ad\ y)) = tr\,A = tr((ad\ x)|I \circ (ad\ y)|I),$$

i.e.

$$\kappa_I(x,y) = tr((ad(x) \circ ad(y))|I) = tr(ad(x) \circ ad(y)) = \kappa(x,y).$$

□

Proposition 4.16 (Semisimple ideal as a direct summand). *If a Lie algebra L contains a semisimple ideal $I \subset L$, then L splits as*

$$L = I \oplus J$$

with an ideal $J \subset L$.

Proof. i) *Directness*: With respect to the Killing form κ of L consider the orthogonal space of I in L

$$I^\perp = \{y \in L : \ \kappa(I,y) = 0\}$$

Lemma 4.1 shows that $J := I^\perp$ is also an ideal in L. The intersection

$$A := (I \cap J) \subset L$$

is an ideal in L. Due to Lemma 4.15 the Killing form κ_I of I is the restriction of the Killing form κ to arguments from L. Hence

$$0 = \kappa(I,A) = \kappa_I(I,A),$$

and the semisimpleness of I implies $A = \{0\}$.

ii) *Dimension formula*: One has

$$\dim\ I^\perp \geq dim\,L - dim\,I$$

because the number of linear equations, which define the reduction from L to $I^\perp$, is equal to $dim\,I$. Hence

$$dim\,I + \dim\ I^\perp \geq dim\,L,$$

which implies

$$dim\ I + \dim\ I^{\perp} = dim\ L \text{ and } L = L \oplus J.$$

□

Theorem 4.17 (Criterion for reductiveness). *Consider a complex vector space V and an embedded subalgebra $L \subset gl(V)$. If V is an irreducible L-module, then L splits as*

$$L = L_0 \oplus J$$

- *with the semisimple subalgebra*

$$L_0 := \{x \in L :\ tr\ x = 0\}$$

- *and a central, scalar ideal*

$$J \subset \mathbb{C} \cdot \mathbb{1}_V \subset Z(L).$$

Proof. i) *Semisimpleness of* L_0: The kernel of the trace map

$$L_0 = ker\ [tr : End(V) \to \mathbb{C}]$$

is an ideal in L of codimension

$$codim_L\ L_0 \leq 1.$$

Lemma 4.9 implies

$$rad\ L_0 \subset L$$

is an ideal. The Lie algebra

$$M := rad\ L_0$$

is solvable. Due to Theorem 3.20 there exists a common eigenvector $\mathrm{v} \in V$ of all endomorphism of M. The eigenvector v defines a linear functional

$$\lambda : M \to \mathbb{C}$$

satisfying for all $X \in M$

$$X.\mathrm{v} = \lambda(X) \cdot \mathrm{v}$$

The vector space

$$W := \{w \in V :\ X.w = \lambda(w) \cdot w \text{ for all } w \in M\}$$

is not zero. Due to Dynkin's Lemma, Proposition 3.19, the vector space $W \subset V$ is even an L-module. The irreducibility of the L-module V implies

$$W = V.$$

Hence for each $x \in M$ the endomorphism

$$x : V \to V$$

is the scalar

$$x = \lambda(x) \cdot \mathbb{1}_V$$

with

$$tr\, x = (dim\, V) \cdot \lambda(X)$$

The inclusion

$$x \in rad\, L_0 = ker(tr\, X)$$

implies $\lambda(x) = 0$, and a posteriori $x = 0$. We obtain

$$M = rad\, L_0 = \{0\},$$

which shows the semisimplicity of L_0.

ii) *The central ideal* J: Proposition 4.16 implies the existence of an ideal $J \subset L$ such that

$$L = L_0 \oplus J$$

The estimates

$$codim_{L_1}\, L \leq 1 \text{ and } dim\, L + dim\, J \leq dim\, L_1$$

imply

$$dim\, J \leq 1.$$

Hence J is an Abelian ideal in L, and

$$L = L_0 \oplus J$$

The argument, which was applied above to the solvable Lie algebra M, also applies to the Abelian Lie algebra J. It shows that J is generated by a scalar, in particular J is a central ideal.

□

Corollary 4.18 (Semisimpleness of the classical Lie algebras from the *ABCD*-series). *Each classical $\mathbb{K}$-Lie algebra L from the ABCD-series within the parameter domain (r,m) from Proposition 2.15 is semisimple.*

Proof. 1. *Complex base field*: Denote by L a complex Lie algebra within the range of the corollary. Let V the complex vector space where the defining matrices of L act as endomorphisms. Then

$$L \subset gl(V)$$

is a Lie subalgebra. In order to apply Theorem 4.17 we show that the L-module V is irreducible.

- *A-series* $L = sl(r,\mathbb{C})$: Consider the canonical basis $(e_j)_{1 \le j \le m}$ of $\mathbb{C}^m \simeq V$. For each $j = 1,...,m$, exists a matrix $A \in L$ satisfying

$$A \cdot e_1 = e_j$$

Hence the L-module V is irreducible.

For the Lie algebras L from the remaining B,C,D-series we denote by $(E_{rs})_{1 \le r,s \le m}$ the canonical basis of the vector space $End\ V$. Due to their definition all matrices from $L \subset gl(B)$ are symmetric. For a given L-submodule $W \subset V$ we introduce the associative matrix algebra

$$T := \{f \in End\ V : \ f(W) \subset W\}.$$

It satisfies $L \subset T$, because W is an L-module.

- *B-series and D-series* $L = so(r,\mathbb{C})$: For pairwise distinct indices $1 \le i,j,k \le m$ the elements

$$E_{ij} - E_{ji} \text{ and } E_{jk} - E_{kj}$$

belong to L, hence also to T. Because T is an associative matrix algebra, also for the product

$$E_{ik} := (E_{ij} - E_{ji}) \cdot (E_{jk} - E_{kj}) \in T.$$

And for all indices $1 \le i \le m$

$$E_{ii} = E_{ik} \cdot E_{ki},\ k \neq i,$$

implies

$$E_{ii} \in T.$$

Hence

$$T = End\ V \text{ and } W = V.$$

Therefore V is an irreducible L-module.

- *C-series* $L = sp(r,\mathbb{C})$: For arbitrary indices $1 \le i,j \le r$ one starts with the block matrices

$$\begin{pmatrix} 0 & E_{ii} \\ 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 \\ E_{ij} + E_{ji} & \end{pmatrix} \in L.$$

Hence their product

$$\begin{pmatrix} 0 & E_{ii} \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ E_{ij} + E_{ji} & \end{pmatrix} = \begin{pmatrix} E_{ii} \cdot (E_{ij} + E_{ji}) & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} E_{ij} & 0 \\ 0 & 0 \end{pmatrix}$$

belongs to T. Analogously, one shows

$$\begin{pmatrix} 0 & 0 \\ 0 & E_{ij} \end{pmatrix} \in T.$$

Then for arbitrary $\alpha \in M(r \times r, \mathbb{C})$ the product in block form satisfies

$$\begin{pmatrix} \alpha & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & \mathbb{1} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix} \in T$$

and also

$$\begin{pmatrix} 0 & 0 \\ 0 & \alpha \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ \mathbb{1} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \alpha & 0 \end{pmatrix} \in T$$

Hence

$$T = End\ V \text{ and } W = V.$$

Therefore V is an irreducible L-module.

As a consequence of Theorem 4.17, each complex Lie algebra L of the A, B, C, D-series within the given range is reductive. Because all matrices of L are traceless, the scalar ideal $J \subset Z(L)$ vanishes, and L is semisimple.

2. *Real base field*: A real Lie algebra M is semisimple iff its complexification

$$M \otimes_{\mathbb{R}} \mathbb{C}$$

is semisimple. According to this general result part i) implies the semisimpleness of L.

□

Note: Theorem 7.10 will show that all complex Lie algebras from Corollary 4.18 are even simple. The proof of Corollary 4.18 does not generalize to type D_1. Indeed, the Lie algebra L of type D_1 is Abelian and not semisimple.

The first step on the way to split a semisimple Lie algebra as a direct sum of simple Lie algebras is Proposition 4.19.

The direct sum

$$L := L_1 \oplus L_2$$

of two Lie algebras L_1 and L_2 has as underlying vector space the direct sum of the vector spaces underlying L_1 and L_2, and the Lie bracket of L is by definition

$$[L_1, L_2] := \{0\}.$$

Hence both Lie algebras L_1 and L_2 become ideals in L.

Conversely, taking the direct sum

$$I \oplus J$$

of two ideals of a Lie algebra presupposes that the sum of the underlying vector spaces is direct. Then $I \cap J = \{0\}$, which implies for the Lie bracket

$$[I,J] \subset (I \cap J) = \{0\}.$$

As a consequence: If a Lie algebra L splits in the category of *vector spaces* as the direct sum

$$L = I \oplus J$$

with two ideals $I,\ J \subset L$, then L also splits in the category of *Lie algebras* as the direct sum of the Lie algebras I and J.

All of these considerations also apply to the direct sum of arbitrary many Lie algebras.

Proposition 4.19 (Splitting a semisimple Lie algebra with respect to an ideal). *Consider a semisimple Lie algebra L and an ideal $I \subset L$. Then:*

1. *The Lie algebra L splits as the direct sum of ideals*

$$L = I \oplus I^{\perp}$$

with $I^{\perp}$ the orthogonal space with respect to the Killing form of L.

2. *Both ideals $I,\ I^{\perp} \subset L$ are semisimple.*

Proof. 1) *Direct sum*: First we prove a dimension formula. According to Cartan's criterion for semisimplicity, Theorem 4.14, the Killing form κ of L is non-degenerate. Hence the induced map

$$j : L \to L^*,\ j(x) := \kappa(x, -))$$

is an isomorphism. Remark 4.11 implies

$$j(I^{\perp}) = (L/I)^*$$

Hence

$$dim\ I^{\perp} = dim(L/I)^* = dim(L/I) = dim\ L - dim\ I$$

or

$$dim\ I + dim\ I^{\perp} = dim\ L$$

Secondly, we show

$$I \cap I^{\perp} = \{0\} :$$

According to Lemma 4.12 also the orthogonal space $I^{\perp} \subset L$ is an ideal. Due to Proposition 4.7 we may identify L with *ad L*. Then Cartan's trace condition for solvability applies to the ideal

$$J := I \cap I^{\perp}$$

considered as a Lie algebra: The Killing form κ_J of J is the restriction to J of the Killing form κ of L, according to Lemma 4.15. We have

$$\kappa_J([J,J],\ J) = \kappa([J,J],\ J) = 0$$

because

$$[J,J] \subset J \subset I \text{ and } J \subset I^{\perp}.$$

The Cartan criterion for solvability, Corollary 4.4, implies that $J \subset L$ is solvable, and semisimpleness of L implies

$$J = \{0\}.$$

As a consequence, the dimension formula for the sum of vector spaces

$$dim\ (I + I^{\perp}) = dim\ I + dim\ I^{\perp} - dim(I \cap I^{\perp}) = dim\ I + dim\ I^{\perp} = dim\ L$$

implies the direct sum decomposition

$$L = I \oplus I^{\perp}.$$

2) *Semisimpleness*: Due to part 1) any Abelian ideal $J \subset I$ is also an Abelian ideal of L, because

$$[J, I^{\perp}] \subset [I, I^{\perp}] = \{0\}.$$

The semisimple Lie algebra L has no non-zero Abelian ideal, hence $J = 0$. Analogously for an Abelian ideal of $I^{\perp}$. $\square$

Corollary 4.20 (Semisimplenes in exact sequences). *Consider a short exact sequence of Lie algebra morphisms*

$$0 \to L_0 \xrightarrow{j} L_1 \xrightarrow{\pi} L_2 \to 0$$

Then are equivalent:

- *The Lie algebra L_1 is semisimple.*
- *Both Lie algebras L_0 and L_2 are semisimple.*

Proof. We identify L_0 with the ideal $j(L_0) \subset L_1$.

- Assume L_1 semisimple. Due to Proposition 4.19 the Lie algebra L_1 splits as

  $$L_1 = L_0 \oplus L_0^{\perp},$$

  and the ideals

$$L_0 \subset L_1 \text{ and } L_0^{\perp} \subset L_1$$

are semisimple. Hence

$$L_2 \simeq L_1/L_0 \simeq L_0^{\perp}$$

is semisimple.

- Assume L_0 and L_2 semisimple. Consider an Abelian ideal $J \subset L_1$. Then

$$\pi(J) \subset \pi(L_1) = L_2$$

is an ideal because $\pi : L_1 \to L_2$ is surjective. In addition, $\pi(J)$ is Abelian. The semisimpleness of L_2 implies $\pi(J) = 0$, i.e.

$$J \subset L_0$$

is an Abelian ideal. Semisimpleness of L_0 implies $J = 0$.

□

Theorem 4.21 (Splitting a semisimple Lie algebra into simple summands). *Consider a Lie algebra L.*

1. *The following properties are equivalent:*
   - *L is semisimple*
   - *L splits as a direct sum*

     $$L = \bigoplus_{\alpha \in A} I_\alpha, \ card\ A < \infty,$$

     of simple ideals $I_\alpha \subset L$.
2. *Assume L semisimple with a splitting according to part 1*

   $$L = \bigoplus_{\alpha \in A} I_\alpha, \ card\ A < \infty.$$

   Each ideal $I \subset L$ splits as

   $$I = \bigoplus_{\substack{\alpha \in A \\ \pi_\alpha(I) \neq \{0\}}} I_\alpha$$

   If I is simple then for exactly one $\alpha \in A$

   $$I = I_\alpha.$$

   Hence the splitting of L is uniquely determined up to the order of the summands.

3. A semisimple Lie algebra L equals its derived algebra, i.e.

$$L = D^1 L = [L, L].$$

Proof. Each simple Lie algebra L equals its derived algebra

$$L = [L, L]$$

because $[L, L] \subset L$ is an ideal, but L is not Abelian and has no other ideals than $\{0\}$ and L itself.

1. i) Assume L semisimple. In case $L = \{0\}$ take the empty sum with index set $A = \emptyset$. Otherwise $L \neq \{0\}$ and L is not Abelian.

   If L has no ideal different from L and different from $\{0\}$, then L is simple. Otherwise we choose an ideal I_1 of minimal dimension

   $$\{0\} \subsetneq I_1 \subsetneq L.$$

   Proposition 4.19 provides a direct sum representation

   $$L = I_1 \oplus I_1^{\perp}.$$

   The two ideals I_1 and $I_1^{\perp}$ have the following properties:

   - Any ideal of I_1 respectively of $I_1^{\perp}$ is also an ideal of L because of the direct sum representation.
   - The ideals I_1 and $I_1^{\perp}$ are semisimple due to Proposition 4.19.
   - The ideal I_1 is even simple: Due to the semisimpleness of L the ideal I_1 is not Abelian, and therefore

     $$[I_1, I_1] \neq \{0\}.$$

     And due to its minimality I_1 has no ideal different than $\{0\}$ and I_1.

   Continuing with $I_1^{\perp}$ the decomposition can be iterated until no summand in the direct sum representation

   $$L = I_1 \oplus I_2 \oplus \ldots \oplus I_n$$

   contains a proper ideal. The decomposition stops after finitely many steps because each step decreases the dimension of the ideals in question.

   ii) Assume a direct sum decomposition

   $$L = \bigoplus_{\alpha \in A} I_\alpha, \; card\ A < \infty.$$

   For the semisimpleness of L we have to show: The only Abelian ideal $I \subset L$ is $L = \{0\}$.

For each $\alpha \in A$ the canonical projection

$$\pi_\alpha : L \to I_\alpha$$

is a surjective morphism of Lie algebras. Hence it maps ideals to ideals. Therefore the image

$$\pi_\alpha(I) \subset I_\alpha$$

is an Abelian ideal of the simple Lie algebra I_α. Hence either

$$\pi_\alpha(I) = \{0\} \text{ or } \pi_\alpha(I) = I_\alpha.$$

The latter case is excluded because the simple Lie algebra I_α is not Abelian. As a consequence for each $\alpha \in A$ holds

$$\pi_\alpha(I) = \{0\},$$

which implies

$$I = \{0\}.$$

Hence L has no Abelian ideals different from $\{0\}$.

2. Claim: For each $\alpha \in A$ with $\pi_\alpha(I) \neq \{0\}$ holds

$$\pi_\alpha(I) = I_\alpha \subset I$$

For the proof consider the chain of equalities respectively inclusions

$$\pi_\alpha(I) = I_\alpha = [I_\alpha, I_\alpha] = [I_\alpha, \pi_\alpha(I)] = [I_\alpha, I] \subset I$$

Simpleness of I_α implies the first and the second equality. The third equality is implied by the first equality. Concerning the fourth equality note: The splitting

$$L = \bigoplus_{\beta \in A} I_\beta$$

implies for each element $x \in I$ the decomposition

$$x = \sum_{\beta \in A} x_\beta \text{ with } x_\beta := \pi_\beta(x) \in I_\beta,\ \beta \in A.$$

Because for $\beta \neq \alpha$ holds

$$[I_\alpha, I_\beta] = 0,$$

we have

$$[I_\alpha, x_\alpha] = [I_\alpha, x].$$

As a consequence

$$[I_\alpha, \pi_\alpha(I)] = [I_\alpha, I].$$

The final inclusion follows from the fact that $I \subset L$ is an ideal.

3. The splitting of L implies the direct sum decomposition

$$[L,L] = [\bigoplus_{\alpha \in A} I_\alpha, \bigoplus_{\beta \in A} I_\beta] = \sum_{\alpha,\beta \in A} [I_\alpha, I_\beta] = \sum_{\alpha \in A} [I_\alpha, I_\alpha] =$$

$$= \sum_{\alpha \in A} I_\alpha = \bigoplus_{\alpha \in A} I_\alpha = L$$

□

The logical dependencies between the results of the last two sections is clarified by the diagram from Figure 4.1. It shows the fundamental role of the Killing form as part of the Cartan criteria for solvability and semisimplicity.

$$\begin{tikzcd}
 & \mathit{Theorem}\ 4.3 \arrow[ddl, Rightarrow] \arrow[dr, Rightarrow] & \mathit{Lemma}\ 4.1 \arrow[l, Rightarrow] \\
 & & \mathit{Proposition}\ 4.13 \arrow[d, Rightarrow] \\
\mathit{Corollary}\ 4.4 & & \mathit{Theorem}\ 4.14 \arrow[d, Rightarrow] \\
 & & \mathit{Theorem}\ 4.21
\end{tikzcd}$$

Fig. 4.1 Logical relations of the results in Section 4.1 and 4.2

4.3 Weyl's theorem on complete reducibility

Alike to splitting a semisimple Lie algebra as a direct sum of simple Lie algebras Weyl's Theorem 4.30 splits an arbitrary finite-dimensional representation of a semisimple Lie algebra L as a direct sum of irreducible representations of L.

Consider a $\mathbb{K}$-Lie algebra L, a vector space V, and a representation of L

$$\rho : L \to gl(V).$$

Recall from Definition 2.4 that V is named an L-module with respect to ρ. As a shorthand one often uses for the module operation the notation from commutative algebra

$$L \times V \to V, (x, \mathrm{v}) \mapsto x.\mathrm{v} := \rho(x)(\mathrm{v}).$$

Definition 4.22 (Reducible and irreducible modules). Consider a Lie algebra L and an L-module V.

1. A *submodule* W of V is a subspace $W \subset V$ stable under the action of L, i.e.

$$L.W \subset W.$$

2. An L-module V is *irreducible* iff V has exactly two different L-submodules, namely V and $\{0\}$. Otherwise V is *reducible*. Notably, the zero-module is reducible.

3. A submodule W of V has a *complement* iff a submodule $W' \subset V$ exists with

$$V = W \oplus W'$$

as a direct sum of vector spaces.

4. An L-module V is *completely reducible* iff a decomposition exists

$$V = \bigoplus_{j=1}^{k} W_j$$

with irreducible L-modules $W_j, j = 1, ..., k$.

Applying the standard constructions from linear algebra to L-modules creates a series of new L-modules based on existing L-modules, cf. [24, Chap. 6.1].

Definition 4.23 (Induced representations). Consider a Lie algebra L. Two representations of L

$$\rho : L \to gl(V) \text{ and } \sigma : L \to gl(W)$$

with corresponding L-modules V and W induce further representations of L in a canonical manner:

1. *Direct sum* $\rho \oplus \sigma : L \to gl(V \oplus W)$ with

$$(\rho \oplus \sigma)(x)(\mathrm{v} + w) := \rho(x)(\mathrm{v}) + \sigma(x)(w),\ (x \in L,\ \mathrm{v} \in V,\ w \in W).$$

Corresponding module: *Direct sum* $V \oplus W$ with

$$x.(\mathrm{v} + w) := x.\mathrm{v} + x.w$$

2. *Dual representation* $\rho^* : L \to gl(V^*)$ with - note the minus sign -

$$(\rho^*(x)\lambda)(\mathrm{v}) := -\lambda(\rho(x)\mathrm{v}),\ (x \in L, \lambda \in V^*, \mathrm{v} \in V).$$

Corresponding module: *Dual module* V^* with

$$(x.\lambda)(\mathrm{v}) = -\lambda(x.\mathrm{v}).$$

3. *Tensor product* $\rho \otimes \sigma : L \to gl(V \otimes W)$ with

$$(\rho \otimes \sigma)(x)(\mathrm{v} \otimes w) := (\rho(x) \otimes id_W + id_V \otimes \sigma(x))(\mathrm{v} \otimes w) =$$

$$= \rho(x)\mathrm{v} \otimes w + v \otimes \sigma(x)w,\ (x \in L, \mathrm{v} \in V, w \in W).$$

Corresponding module: *Tensor product* $V \otimes W$ with

$$x.(\mathrm{v} \otimes w) = x.\mathrm{v} \otimes w + \mathrm{v} \otimes x.w$$

4. *Exterior product* $\rho \wedge \rho : L \to gl(\bigwedge^2 V)$ with

$$(\rho \wedge \rho)(x)(\mathrm{v}_1 \wedge \mathrm{v}_2) := \rho(x)\mathrm{v}_1 \wedge \mathrm{v}_2 + \mathrm{v}_1 \wedge \rho(x)\mathrm{v}_2,\ (x \in L, \mathrm{v}_1, \mathrm{v}_2 \in V).$$

Corresponding module: *Exterior product* $\bigwedge^2 V$ with

$$x.(u \wedge \mathrm{v}) = x.u \wedge \mathrm{v} + u \wedge x.\mathrm{v}.$$

5. *Symmetric product* $Sym^2(\rho) : L \to gl(Sym^2 V)$ with

$$(Sym^2(\rho)(x))(\mathrm{v}_1 \cdot \mathrm{v}_2) := (\rho(x)\mathrm{v}_1) \cdot \mathrm{v}_2 + \mathrm{v}_1 \cdot (\rho(x)\mathrm{v}_2),\ (x \in L, \mathrm{v}_1, \mathrm{v}_2 \in V).$$

Corresponding module: *Symmetric product* $Sym^2 V$ with

$$x.(u \cdot \mathrm{v}) = (x.u) \cdot \mathrm{v} + u \cdot (x.\mathrm{v}).$$

6. *Hom-representation* $Hom_{\mathbb{K}}(\rho, \sigma) := \tau : L \to gl(Hom_{\mathbb{K}}(V, W))$ with

$$(\tau(x)f)(\mathrm{v}) := \sigma(x)(f(\mathrm{v})) - f(\rho(x)(\mathrm{v})),\ (x \in L, f \in Hom_{\mathbb{K}}(V, W), \mathrm{v} \in V).$$

Corresponding module: *Vector space* $Hom_{\mathbb{K}}(V, W)$ *of* $\mathbb{K}$*-linear maps* with

$$(x.f)(\mathrm{v}) = x.f(\mathrm{v}) - f(x.\mathrm{v}).$$

The constructions from part 1) and from part 3) - 5) generalize to sums respectively products of more than two components. The notations emphasize the close relationship to similar constructions from commutative algebra for modules over a ring.

Remark 4.24 (Induced representations).

1. It remains to check that the constructions in Definition 4.23 actually yield L-modules, in particular that they are compatible with the Lie bracket. As an example we consider the case of the dual module V^* and verify that the induced map
$$\rho^* : L \to gl(V^*),$$
defined as
$$(\rho^*(x)\lambda)(\mathrm{v}) := -\lambda(\rho(x)\mathrm{v})$$
or
$$(x.\lambda)(\mathrm{v}) = -\lambda(x.\mathrm{v}),$$
preserves the Lie bracket. Claim: For all $x, y \in L, \lambda \in V^*$ holds
$$[x,y].\lambda = x.(y.\lambda) - y.(x.\lambda)$$
Evaluating both sides on an arbitrary vector $\mathrm{v} \in V$ shows:
   - *Left-hand side*:
   $$([x,y].\lambda)(\mathrm{v}) = -\lambda([x.y].\mathrm{v}) = -\lambda(x.(y.\mathrm{v})) + \lambda(y.(x.\mathrm{v}))$$
   - *Right-hand side*:
   $$(x.(y.\lambda) - y.(x.\lambda))(\mathrm{v}) = \lambda(y.(x.\mathrm{v})) - \lambda(x.(y.\mathrm{v}))$$
   Note in the last equation: The functional $x.(y.\lambda)$ means to apply x to the functional $y.\lambda$, hence
   $$(x.(y.\lambda))(\mathrm{v}) = -(y.\lambda)(x.\mathrm{v}) = \lambda(y.(x.\mathrm{v})),$$
   switching the order of x and y.

2. One checks that the canonical isomorphism in the category of $\mathbb{K}$- vector spaces
$$V^* \otimes W \xrightarrow{\simeq} Hom_{\mathbb{K}}(V,W), \lambda \otimes w \mapsto \lambda(-) \cdot w,$$
extends to an isomorphism in the category of L-modules.

3. For a $\mathbb{K}$-linear map $f \in Hom_{\mathbb{K}}(V,W)$:
$$L.f = 0 \iff f : V \to W \text{ is } L\text{-linear.}$$
The vector space of L-linear morphisms $V \to W$ is denoted
$$Hom_L(V,W)$$
Elements from $Hom_L(V,W)$ are sometimes named *intertwiner*, because they relate the L-modules V and W by an L-module morphism.

Theorem 4.25 (Lemma of Schur for irreducible representations). *Consider a $\mathbb{K}$-Lie algebra L.*

1. *Then each L-module morphism*

$$f : V \to W$$

between two irreducible L-modules is either zero or an isomorphism.

2. *Consider an irreducible representation with a complex vector space V*

$$\rho : L \to gl(V).$$

Then each endomorphism $f \in End(V)$, which commutes with all endomorphisms from $\rho(L)$, is a scalar multiple of the identity, i.e. if for all $x \in L$ holds

$$[f, \rho(x)] = 0,$$

then a complex number $\mu \in \mathbb{C}$ exists with

$$f = \mu \cdot id_V.$$

3. *Consider two morphisms*

$$f_1, f_2 : V \to W$$

between irreducible complex L-modules. If $f_2 \neq 0$ then

$$f_1 = \mu \cdot f_2$$

for a suitable $\mu \in \mathbb{C}$, i.e.

$$Hom_L(V,W) = \begin{cases} \mathbb{C} & V \simeq W \\ \{0\} & V \not\simeq W \end{cases}$$

Proof. 1. Because f is a morphism, its kernel

$$\ker f \subset V$$

is a submodule. Irreducibility of V implies

$$ker f = \{0\} \text{ or } ker f = V.$$

If $ker f = \{0\}$ then f is injective and $f(V) \subset W$ is a submodule with $f(V) \neq \{0\}$. Irreducibility of W implies $f(V) = W$. Therefore f is injective and surjective, hence an isomorphism. If ker $f \neq \{0\}$, then $ker\ f = V$, hence f is zero.

2. The endomorphism f has a complex eigenvalue μ. Its eigenspace $W \subset V$ is an L-submodule: Each eigenvector $w \in W$ satisfies

$$(f \circ \rho(x))(w) = (\rho(x) \circ f)(w) = \rho(x)(\mu \cdot w) = \mu \cdot \rho(x)(w).$$

Now $W \neq \{0\}$ and V irreducible imply $W = V$.

3. Because $f_2 \neq 0$ the morphism f_2 is an isomorphism according to part 1. Consider the L-module morphism
$$f := f_1 \circ f_2^{-1} : W \to W.$$
Then
$$f \in Hom_L(W,W),$$
i.e. f commutes with the action of L on W. Part 2 implies the existence of a suitable $\mu \in \mathbb{C}$ with
$$f = \mu \cdot id_W$$
which proves
$$f_1 = \mu \cdot f_2.$$

□

Definition 4.26 (Quadratic Casimir element of a representation). Consider a semisimple Lie algebra L and a faithful representation $\rho : L \to gl(V)$ on a vector space V. The trace form of ρ

$$\beta : L \times L \to \mathbb{K},\ \beta(x,y) := tr(\rho(x)\rho(y)),$$

is non-degenerate according to Proposition 4.13. For a base $(x_i)_{i=1,...,n}$ of L denote by $(y_j)_{j=1,...,n}$ the dual base with respect to the trace form β, i.e.

$$\beta(x_i, y_j) = \delta_{ij}.$$

The *quadratic Casimir element* of ρ is defined as the $\mathbb{K}$-linear endomorphism

$$c_\rho := \sum_{i=1}^{n} \rho(x_i)\rho(y_i) \in End(V).$$

Note: One checks that the quadratic Casimir element does not depend on the choice of the basis $(x_i)_{i=1,...,n}$. The Casimir element is an element of the associative algebra $End(V)$. It depends in a quadratic way on the elements of L.

Remark 4.27 (Reduction to faithful representations). If the representation ρ is not faithful then one considers the direct decomposition

$$L = ker\ \rho \oplus L'$$

with

$$L' := (ker\ \rho)^{\perp} \subset L.$$

The Lie algebra L' is semisimple according to Proposition 4.19. The restricted representation

$$\rho|L' : L' \to gl(V)$$

is faithful. The Casimir element of ρ is by definition the Casimir element of the restriction $\rho|L'$.

Theorem 4.28 (Properties of the Casimir element). *For a semisimple Lie algebra L the quadratic Casimir element of a faithful representation ρ of L on a vector space V*

$$c_\rho \in End(V)$$

has the following properties:

- Commutation*: The Casimir element commutes with all elements of the representation*

  $$[c_\rho, \rho(L)] = 0,$$

  and the $\mathbb{K}$-linear endomorphism

  $$c_\rho : V \to V$$

  is even an L-module morphism.

- Trace*: $tr(c_\rho) = dim\ L$*

- Scalar*: For an irreducible representation ρ of L on a complex vector space V holds*

  $$c_\rho = \frac{dim\ L}{dim\ V} \cdot id_V.$$

Proof. The faithful representation ρ is an embedding

$$\rho : L \to gl(V)$$

Due to Proposition 4.13 the trace form β of ρ is non-degenerate. Set $n = dim\ L$. By definition

$$c_\rho = \sum_{i=1}^{n} \rho(x_i)\rho(y_i) \in End(V)$$

with a pair of bases $(x_i)_{i=1,...,n}$ and $(y_j)_{j=1,...,n}$ of L, which are dual with respect to β.

- *Commutation*: For $x \in L$ we show $[\rho(x), c_\rho] = 0$: Define the coefficients (a_{ij}) and (b_{jk}) according to

$$[x, x_i] = \sum_{k=1}^{n} a_{ik} \cdot x_k, \ [x, y_j] = \sum_{k=1}^{n} b_{jk} \cdot y_k.$$

Because the families $(x_j)_{1 \le j \le n}$ and $(y_k)_{1 \le k \le n}$ are dual bases with respect to β

$$a_{ij} = \beta(\sum_{k=1}^{n} a_{ik} \cdot x_k, y_j) = \beta([x, x_i], y_j) = -\beta([x_i, x], y_j) = -\beta(x_i, [x, y_j]) =$$

$$= -\beta(x_i, \sum_{k=1}^{n} b_{jk} \cdot y_k) = -b_{ji}.$$

Here we made use of the associativity of the trace form according to Lemma 4.1. To compute

$$[\rho(x), c_\rho] = \sum_{i=1}^{n} [\rho(x), \rho(x_i)\rho(y_i)]$$

we use the formula

$$[A, BC] = [A, B]C + B[A, C]$$

for endomorphisms $A, B, C \in End(V)$. The formula follows easily by expanding both sides.

Therefore each summand of the last sum decomposes as

$$[\rho(x), \rho(x_i)\rho(y_i)] = [\rho(x), \rho(x_i)]\rho(y_i) + \rho(x_i)[\rho(x), \rho(y_i)]$$

and therefore

$$[\rho(x), c_\rho] = \sum_{i=1}^{n} ([\rho(x), \rho(x_i)]\rho(y_i) + \rho(x_i)[\rho(x), \rho(y_i)]) =$$

$$= \sum_{i=1}^{n} (\rho([x, x_i])\rho(y_i) + \rho(x_i)\rho([x, y_i]) = \sum_{i,j=1}^{n} a_{ij} \cdot \rho(x_j)\rho(y_i) + \sum_{i,k=1}^{n} b_{ik} \cdot \rho(x_i)\rho(y_k) =$$

$$= \sum_{i,j=1}^{n} a_{ij} \cdot \rho(x_j)\rho(y_i) + \sum_{i,j=1}^{n} b_{ji} \cdot \rho(x_j)\rho(y_i) =$$

$$= \sum_{i,j=1}^{n} (a_{ij} + b_{ji}) \cdot (\rho(x_j)\rho(y_i)) = 0.$$

Here we have changed in the second sum the summation indices $(i, k) \mapsto (j, i)$.

The commutation

$$[c_\rho, \rho(L)] = 0$$

is equivalent to the fact that
$$c_\rho : V \to V$$
is L-linear.

- *Trace*: We have
$$tr(c_\rho) = \sum_{i=1}^{n} tr(\rho(x_i)\rho(y_i)) = \sum_{i=1}^{n} \beta(x_i, y_i) = n = dim\ L.$$

- *Scalar*: For an irreducible representation ρ we get with the first part of the proof and with the Lemma of Schur, Theorem 4.25, part 2
$$c_\rho = \mu \cdot id_V,$$
and with the second part of the present proof
$$tr(c_\rho) = \mu \cdot dim\ V = dim\ L.$$
Hence
$$\mu = \frac{dim\ L}{dim\ V}$$

□

In Theorem 4.28 the properties of the Casimir element of a representation are not a happy incidence. They follow from the fact that the Casimir operators have their origin in the center of the universal enveloping algebra of L, see [24, Chap. 22.1].

Lemma 4.29 will be used in the proof of Theorem 4.30.

Lemma 4.29 (Representations of semisimple Lie algebras are traceless). *Consider a semisimple Lie algebra L and a representation $\rho : L \to gl(V)$.*

- *Then $\rho(L) \subset sl(V)$, i.e. for all $x \in L$*
$$tr(\rho(x)) = 0.$$

- *In particular, each 1-dimensional representations of L is trivial, i.e. if $dim\ V = 1$ then $\rho = 0$.*

Proof. Because L is semisimple Theorem 4.21 implies
$$L = [L, L].$$
If $x = [u, \mathrm{v}] \in L$ then
$$tr(\rho(x)) = tr(\rho([u, \mathrm{v}])) = tr([\rho(u), \rho(\mathrm{v})]) = 0$$

according to Lemma 4.1. For 1-dimensional V, i.e. $V = \mathbb{K}$, holds for all $x \in L$

$$0 = tr(\rho(x)) = \rho(x).$$

□

Theorem 4.30 (Weyl's theorem on complete reducibility). *For a semisimple Lie algebra L each non-zero L-module is completely reducible. The isomorphism class of each irreducible direct summand as well as the multiplicity of each isomorphism class is uniquely determined.*

Proof. For the existence of the splitting we have to show: Each submodule of an L-module has a complement. For the proof we may assume $L \neq \{0\}$. We will consider all pairs (V,W) with an L-module V and a submodule $W \subset V$, and proceed along the following steps:

1. *Particular case* $codim_V W = 1$: Proof by induction on $n = dim\, W$.

   - Subcase 1a): W reducible. The proof is elementary and relies on a separate induction.

   - Subcase 1b): W irreducible. The proof relies on the Casimir element and Schur's Lemma.

2. *General case* $codim_V W$ *arbitrary*: The proof constructs a complement of W as the kernel of a certain section against the injection $W \hookrightarrow V$. The section is obtained by considering the L-module

$$Hom_{\mathbb{K}}(V,W)$$

   and constructing a pair $(\mathscr{V},\mathscr{W})$ to which the particular case of codimension $= 1$ applies.

All exact sequences in the following refer to the category of L-modules.

1. *Particular case* $codim_V W = 1$: Consider all pairs (V,W) with an L-module V and a submodule

$$W \subset V \text{ satisfying } codim_V\ W = 1.$$

   Due to Lemma 4.29 the 1-dimensional quotient V/W fits into an exact sequence of L-modules

$$0 \to W \to V \to V/W \to 0$$

   We construct a complement of W by induction on $dim\ W$ with the induction assumption: For all pairs of L-modules

$$(V_1,V_2) \text{ with } codim_{V_1} V_2 = 1 \text{ and } dim\, V_2 < dim\, W$$

exists a complement of V_2 in V_1.

The induction step employs one of two alternative subcases. Subcase 1a) uses the induction assumption twice, while subcase 1b) uses the Casimir element.

Subcase 1*a), W reducible*: Then a proper submodule

$$\{0\} \subsetneq W' \subsetneq W,$$

exists, in particular

$$dim\, W' < dim\, W.$$

Dividing out the proper submodule $W' \subset W$ induces the exact sequence

$$0 \to W/W' \to V/W' \to \frac{V/W'}{W/W'} \to 0$$

The submodule

$$W/W' \subset V/W'$$

has

$$codim_{V/W'}\, W/W' = codim_V\, W = 1$$

and satisfies

$$dim\,(W/W') < dim\, W.$$

- Hence the pair $(V/W', W/W')$ satisfies the induction assumption. We obtain a complement in the form

  $$\tilde{W}/W',\ W' \subsetneq \tilde{W} \subsetneq V,$$

  and a first splitting

  $$V/W' = W/W' \oplus \tilde{W}/W'.$$

  It induces the exact sequence

  $$0 \to W' \to \tilde{W} \to \tilde{W}/W' \to 0$$

  with the isomorphy of vector spaces

  $$\tilde{W}/W' \simeq \frac{V/W'}{W/W'} \simeq \mathbb{K}$$

  because the complement of W/W' in V/W' has codimension $= 1$.

  We recall also $dim\, W' < dim\, W$.

- Hence also the pair $(\tilde{W}, W')$ satisfies the induction assumption. We obtain a complement $X \subset \tilde{W}$ and a second splitting

$$\tilde{W} = W' \oplus X.$$

Claim: Combining the two splittings provides the final splitting

$$V = W \oplus X.$$

For the proof, on one hand

$$dim\, V = dim\, W + dim\, \tilde{W} - dim\, W' \text{ and } dim\, \tilde{W} = dim\, W' + dim\, X.$$

Hence

$$dim\, V = dim\, W + dim\, X.$$

On the other hand,

$$X \subset \tilde{W} \implies (W \cap X) \subset (W \cap \tilde{W}).$$

Due to the first splitting

$$\{0\} = (W/W') \cap (\tilde{W}/W') \implies (W \cap \tilde{W}) \subset W',$$

and therefore

$$(W \cap X) \subset (W \cap \tilde{W}) \subset W'.$$

As a consequence

$$W \cap X = (W \cap X) \cap W' = W \cap (X \cap W').$$

The second splitting implies

$$X \cap W' = \{0\}.$$

Hence

$$W \cap X = \{0\}.$$

Therefore

$$V = W \oplus X$$

which finishes the induction step for reducible W.

Subcase 1*b), W irreducible*: Assume that the representation

$$\rho : L \to gl(V)$$

defines the L-module structure of V. Due to Remark 4.27 we may assume ρ faithful. We consider the Casimir element of ρ

$$c_\rho := \sum_{j=1}^{dim\, L} \rho(x_j)\rho(y_j) \in End(V).$$

- Refering to V: The Casimir element

$$c_\rho : V \to V$$

is an L-module morphism due to Theorem 4.28, part 1. Therefore

$$X := ker\ c_\rho \subset V$$

is an L-submodule.

- Refering to W: Because the 1-dimensional L-module V/W is trivial, we have

$$\rho(L)(V) \subset W.$$

The definition of c_ρ implies that also

$$c_\rho(L)(V) \subset W.$$

One can find a basis of the vector space V such that the matrix of c_ρ has block form

$$c_\rho = \begin{pmatrix} A & * \\ 0 & 0 \end{pmatrix}$$

and the matrix A represents the restriction

$$c_\rho|W : W \to W.$$

Hence

$$tr\ (c_\rho|W) = tr\ c_\rho$$

Due to Theorem 4.28, part 2

$$tr\ c_\rho = dim\ L \neq 0,$$

which implies

$$tr(c_\rho|W) \neq 0.$$

Due to Theorem 4.25, part 1 the irreducibility of W implies that

$$c_\rho|W : W \to W$$

is an isomorphism. Then

$$dim\ V = dim(im\ c_\rho) + dim(ker\ c_\rho) = dim\ W + dim\ X$$

implies the splitting of L-modules

$$V = W \oplus X$$

and finishes the induction step for irreducible W.

2. *General case $codim_V W$ arbitrary*:

   Consider an arbitrary proper submodule
$$\{0\} \subsetneq W \subsetneq V.$$
   We want to construct a complement of W as the kernel of an L-module morphism. Therefore we claim the existence of a section against the canonical injection
$$j : W \hookrightarrow V,$$
   i.e. we claim the existence of an L-linear map
$$\tilde{f} : V \to W$$
   such that
$$\tilde{f} \circ j = id_W \text{ i.e. } \tilde{f}|W = id_W$$
   Then $\tilde{f}$ is surjective, and
$$X := ker\ \tilde{f}$$
   is a complement of W because $V/X \simeq W$.

   The idea is to translate the question on the existence of the section $\tilde{f}$ to a problem about L-modules in a context where case 1 applies. We consider the induced L-module
$$Hom_{\mathbb{K}}(V,W)$$
   to reduce the question on sections to a problem concerning pairs of L-modules of $\mathbb{K}$-linear homomorphism
$$(\mathscr{V},\mathscr{W})$$
   with $codim_{\mathscr{V}}\mathscr{W} = 1$. The latter problem can be solved by case 1.

   Note that elements of $Hom_{\mathbb{K}}(V,W)$ are morphisms in the category of vector spaces, not necessarily morphisms of L-modules.

   Consider the following submodules of the L-module $Hom_{\mathbb{K}}(V,W)$
$$\mathscr{V} := \{f \in Hom_{\mathbb{K}}(V,W) : f|W = \lambda \cdot id_W,\ \lambda \in \mathbb{K}\}$$
$$\mathscr{W} := \{f \in \mathscr{V} : f|W = 0\}.$$
   In order to prove that $\mathscr{V}$ is a L-module, consider $f \in Hom_{\mathbb{K}}(V,W)$ with $f|W = \lambda \cdot id_W$ and $x \in L, w \in W$:
$$(x.f)(w) = x.(f(w)) - f(x.w) = x.(\lambda \cdot w) - \lambda \cdot (x.w) =$$
$$= \lambda \cdot (x.w) - \lambda \cdot (x.w) = 0.$$
   Therefore even

$$L.\mathscr{V} \subset \mathscr{W} \subset \mathscr{V}$$

and both $\mathscr{V}$ and $\mathscr{W}$ are L-modules. By definition

$$codim_{\mathscr{V}} \mathscr{W} = 1$$

because $\mathscr{W}$ as a subspace of $\mathscr{V}$ is defined by the single linear equation $\lambda = 0$.

We now apply the result of case 1 to the pair $(\mathscr{V}, \mathscr{W})$: The submodule $\mathscr{W} \subset \mathscr{V}$ has a complement, i.e. there exists a $\mathbb{K}$-linear map

$$\tilde{f} \in \mathscr{V} \text{ with } \tilde{f}|W \neq 0,$$

which splits

$$\mathscr{V} = \mathscr{W} \oplus \mathbb{K} \cdot \tilde{f}.$$

By definition of $\mathscr{V}$ the restriction has the form

$$\tilde{f}|W = \mu \cdot id_W$$

with a non-zero scalar $\mu \in \mathbb{K}$ because

$$\tilde{f} \notin \mathscr{W},$$

and we may assume $\mu = 1$.

The L-module

$$\mathbb{K} \cdot \tilde{f}$$

is 1-dimensional, hence trivial according to Lemma 4.29. According to Remark 4.24 the equality $L.\tilde{f} = 0$ implies the L-linearity of $\tilde{f}$. Therefore

$$X := ker\ \tilde{f} \subset V$$

is an L-submodule. Due to our considerations at the beginning of case 2

$$V = W \oplus X.$$

3. *Uniqueness of the isomorphism classes*: Consider an irreducible submodule of $W \subset V$, and a splitting

$$V = \bigoplus_{j \in J} V_j$$

with irreducible L-modules $V_j,\ j \in J$. For each $j \in J$ the canonical projection

$$p_j : V \to V_j$$

is an L-module morphism. There exists at least one index $j \in J$ with

$$\{0\} \neq p_j(W) \subset V_j$$

Because both L-modules V_j as well as W are irreducible we have

$$p_j(W) = V_j \text{ and } ker\ (p_j|W) = \{0\}.$$

Hence

$$p_j|W : W \xrightarrow{\simeq} V_j$$

is an isomorphism of L-modules. Iterating the argument proves the claim.
□

Note. Theorem 4.30 is due to Weyl, who gave an analytic proof using the theory of maximal compact subgroups of a semisimple Lie group ("Unitarian trick"), see [48, Kap. I, §5, Satz 5] and also Serre's explanation in [40, Chap. VIII, no. 7]. The proof given above stays completely in the algebraic domain, see [41, Part 1, Chap. 6.3].

The following Proposition 4.31 is a consequence of Theorem 4.30. It refers to the Jordan decomposition for embedded semisimple Lie algebras.

Proposition 4.31 (Jordan decomposition for an embedded semisimple Lie algebra). *Consider a complex vector space V and an embedded semisimple Lie algebra $L \subset gl(V)$. If an element $x \in L$, considered as endomorphism of V*

$$x : V \to V,$$

has the Jordan decomposition

$$x = x_s + x_n \in End_{\mathbb{C}}(V),$$

then both summands belong to L, i.e. $x_s, x_n \in L$. In addition,

$$ad\ x = ad\ x_s + ad\ x_n$$

is the Jordan decomposition of

$$ad_x : L \to L.$$

Proof. The proof has two separate parts. The first part considers L as an embedded Lie algebra $L \subset gl(V)$. The second part considers the induced L-module $End_{\mathbb{C}}(V,V)$.

Part 1. *The embedded Lie algebra $L \subset gl(V)$*: We introduce the shorthand

$$E := End_{\mathbb{C}}(V).$$

Proposition 3.3 implies for the endomorphism

$$ad_E\ x : E \to E,\ f \mapsto [x, f] := x \circ f - f \circ x,$$

the Jordan decomposition

$$ad_E\ x = ad_E\ x_s + ad_E\ x_n$$

with endomorphisms

$$x_s,\ x_n \in E.$$

There exist polynomials $p_s(T),\ p_n(T) \in \mathbb{C}[T]$ without constant term such that

$$ad_E\ x_s = p_s(ad_E\ x),\ ad_E\ x_n = p_n(ad_E\ x).$$

Let

$$N := N_{gl(V)}L \subset gl(V)$$

be the normalizer of L. Then $x \in L \subset N$. Hence also

$$x_s,\ x_n \in N.$$

Part 2. *Semisimpleness of* L: In order to show $x_s,\ x_n \in L$ the result $x_s,\ x_n \in N$ is not sufficient, because in general

$$L \subsetneq N.$$

Therefore we will now employ the semisimpleness of L to construct a specific L-submodule

$$\tilde{L} \subset End_{\mathbb{C}}(V)$$

satisfying

$$L \subset \tilde{L} \subset N \text{ and } x_s,\ x_n \in \tilde{L},$$

and eventually show $\tilde{L} = L$.

i) *Construction of* $\tilde{L}$: For any L-submodule $W \subset V$ we consider the vector subspace of endomorphisms

$$L_W := \{y \in End_{\mathbb{C}}(V) : y(W) \subset W \text{ and } tr(y|W) = 0\} \subset End_{\mathbb{C}}(V).$$

E.g.

$$L_W = sl(V) \text{ if } W := V$$

and

$$L_W = gl(V) \text{ if } W := \{0\}$$

Because W is an L-module, also L_W is an L-submodule with respect to the induced L-module structure on $Hom_{\mathbb{C}}(V,V)$: For each endomorphism

$$(z : V \to V) \in L \subset End_{\mathbb{C}}(V)$$

and $y \in L_W,\ w \in W$ holds

$$(z.y)(w) = z(y(w)) - y(z(w)) \in W$$

and

$$tr((z.y)|W) = tr([z,y]|W) = 0.$$

We define

$$\tilde{L} := N \cap \bigcap_{W \subset V} L_W = \bigcap_{W \subset V} (N \cap L_W)$$

with the intersection taken for all L-submodules $W \subset V$. Because N and each L_W are L-modules of the L-module $End_{\mathbb{C}}(L)$, also $\tilde{L}$ is an L-module. It satisfies:

- $L \subset \tilde{L}$: Because L is semisimple, Lemma 4.29 implies $tr(y|W) = 0$ for all $y \in L$. As a consequence
$$L \subset \bigcap_{W \subset V} L_W \text{ and } L \subset \tilde{L}.$$
- $x_s, x_n \in \tilde{L}$: Because the vector subspace W is stable with respect to the endomorphism $x : V \to V$, the same is true for its Jordan components which depend on x in a polynomial way, i.e.
$$x_s(W) \subset W \text{ and } x_n(W) \subset W.$$
Again according to Lemma 4.29, the semisimpleness of L implies $tr(x|W) = 0$. Hence $x \in L_W$. With x_n also the restriction $x_n|W$ is nilpotent, and therefore
$$tr\ (x_n|W) = 0 \text{ and } x_n \in L_W.$$
As a consequence also
$$x_s = x - x_n \in L_W.$$
We obtain
$$x_s, x_n \in \tilde{L}$$
because $x_s, x_n \in N$ and $x_s, x_n \in L_W$ for all L-submodules $W \subset V$.

ii): *The equality* $L = \tilde{L}$: It remains to show $L = \tilde{L}$. Weyl's theorem on complete reducibility, Theorem 4.30, applies to the L-module $\tilde{L}$. Hence there exists a L-submodule $M \subset \tilde{L}$ with

$$\tilde{L} = L \oplus M.$$

We claim: $M = 0$. Because $\tilde{L} \subset N$, the normalizer of L, we have

$$[L, \tilde{L}] = [\tilde{L}, L] \subset [N, L] \subset L$$

which implies

$$[L, M] \subset (L \cap M)$$

because M is an L-module. Due to

$$L \cap M = \{0\}$$

the action of L on M is trivial.

In order to conclude $M = \{0\}$ we consider an arbitrary, but fixed endomorphism $y \in M$. The annihilation

$$[L, y] = 0$$

means that the endomorphism

$$y : V \to V$$

commutes with all endomorphism of L. Because V is a complex vector space, Schur's Lemma, Theorem 4.25, part 2 implies for each irreducible submodule $W \subset V$ the existence of a scalar $\mu \in \mathbb{C}$ with

$$y|W = \mu \cdot id_W.$$

On the other hand, $y \in L_W$ implies $tr(y|W) = 0$, hence

$$y|W = 0$$

for each irreducible L-sumbdule of V. And the splitting of V as a direct sum of irreducible L-modules shows $y = 0$. Because $y \in M$ can be choosen arbitrarily, we obtain $M = \{0\}$ and

$$L = \tilde{L}.$$

By construction $x_s,\ x_n \in \tilde{L} = L$.

Part 3. *Jordan decomposition*: The final claim about the Jordan decomposition of $ad\ x$ follows from the result

$$ad\ x = (ad_E\ x_s)|L + (ad_E\ x_n)|L$$

from part 1, and the result $x_s,\ x_n \in L$ from part 2. $\square$

We proved Theorem 4.31 for a complex embedded semisimple Lie algebra because our proof employs the strong form of Schur's Lemma for complex-linear module endomorphism. For a real Lie algebra L the complexification

$$L \otimes_{\mathbb{R}} \mathbb{C}$$

remains semisimple, because its Killing form is the complexification of the Killing form of L, see [4, Chap.I, §6, no. 3 Prop. 3]. Theorem 4.31 applies to the complex Lie algebra $L \otimes_{\mathbb{R}} \mathbb{C}$ and endomorphisms

$$x \in L \otimes_{\mathbb{R}} \mathbb{C} \subset gl(V \otimes_{\mathbb{R}} \mathbb{C}).$$

The previous Proposition 4.31 assures: For an embedded semisimple Lie algebra L of endomorphisms of a complex vector space V the Jordan decomposition of the endomorphisms of V resulting from $ad\ L$ induces a decomposition of elements from L. We know from Proposition 4.7 that each semisimple Lie algebra L embeds via its adjoint representation as a semisimple Lie algebra of endomorphisms. Hence a complex semisimple Lie algebra satisfies the assumption of Proposition 4.31. Definition 4.32 defines the abstract Jordan decomposition of L. Then Corollary 4.33 shows why the abstract Jordan decomposition is a useful concept.

Definition 4.32 (Abstract Jordan decomposition). Consider a complex semisimple Lie algebra L. Its adjoint representation

$$ad : L \to ad\ L \subset gl(L)$$

is an isomorphism and represents L as an embedded Lie algebra of endomorphisms. For each $x \in L$ the adjoint endomorphism

$$ad\ x : L \to L,\ y \mapsto [x, y]$$

has the Jordan decomposition

$$ad\ x = f_s + f_n \in End(L),$$

and Proposition 4.31 ensures

$$f_s \in ad(L) \text{ and } f_n \in ad(L).$$

One defines

$$s := ad^{-1}(f_s) \in L \text{ and } n := ad^{-1}(f_n) \in L.$$

Then the decomposition

$$x = s + n$$

with $s \in L$ ad-semisimple, $n \in L$ ad-nilpotent and $[s, n] = 0$ is named the *abstract Jordan decomposition* of $x \in L$.

Note. The abstract Jordan decomposition

$$x = s + n$$

is uniquely determined by the property, that the components $s,\ n \in L$ are respectively ad-semisimple and ad-nilpotent and satisfy $[s, n] = 0$. The uniqueness follows from the uniqueness of the Jordan decomposition of the endomorphism

$$ad\ x : L \to L$$

and from the isomorphism

$$ad : L \xrightarrow{\simeq} ad(L).$$

In the abstract case, neither $x \in L$ nor its components $s,\ n \in L$ from the abstract Jordan decomposition are endomorphisms of a vector space. But Corollary 4.33 will show: For any representation

$$\rho : L \to gl(V)$$

the abstract Jordan decomposition of $x \in L$ induces the Jordan decomposition of the endomorphism $\rho(x) \in End(V)$.

Corollary 4.33 (Jordan decomposition for representations of semisimple Lie algebras). *Consider a complex vector space V and a representation*

$$\rho : L \to gl(V)$$

of a complex semisimple Lie algebra L. If $x \in L$ has the abstract Jordan decomposition

$$x = s + n$$

then $\rho(x) \in End_{\mathbb{C}}(V)$ has the Jordan decomposition

$$\rho(x) = \rho(s) + \rho(n).$$

Proof. Due to Corollary 4.20 the Lie algebra

$$F := \rho(L)$$

is semisimple.

i) *Abstract Jordan decomposition of $\rho(x) \in F$*: We choose a basis $\mathscr{B} = (v_j)_{j=1,\dots,n}$ of L of eigenvectors of the semisimple endomorphism

$$ad_L(s) : L \to L.$$

Then the non-zero elements from

$$\rho(\mathscr{B}) := (\rho(v_j))_{j=1,\dots,n}$$

form a family of eigenvectors of $ad_F\rho(s)$, and spans F. Hence the endomorphism

$$ad_F\rho(s) \in End_{\mathbb{C}}(F)$$

is semisimple.

The nilpotency of $ad_L n \in End_{\mathbb{C}}(L)$ implies the nilpotency of the endomorphism

$$ad_F(\rho(n)) \in End_{\mathbb{C}}(F)$$

Both endomorphisms commute

$$[ad_F(\rho(s)),\ ad_F(\rho(n))] = ad_F(\rho([s,n])) = 0$$

Hence

$$ad_F(\rho(x)) = ad_F(\rho(s)) + ad_F(\rho(n))$$

is the Jordan decomposition of the endomorphism

$$ad_F(\rho(x)) \in End_{\mathbb{C}}(F)$$

As a consequence,

$$\rho(x) = \rho(s) + \rho(n)$$

is by definition the uniquely determined abstract Jordan decomposition of $\rho(x) \in F$ from Definition 4.32.

ii) *Jordan decomposition of* $\rho(x) \in End_{\mathbb{C}}(V)$: Now we consider the element $\rho(x) \in F$ as an endomorphism

$$\rho(x) : V \to V$$

of the vector space V. Due to Theorem 1.19 the endomorphism $\rho(x)$ has the Jordan decomposition

$$\rho(x) = f_s + f_n$$

with semisimple $f_s \in End_{\mathbb{C}}(V)$ and nilpotent $f_n \in End_{\mathbb{C}}(V)$. Proposition 4.31 applies to the semisimple embedded Lie algebra $F \subset gl(V)$ and shows

$$f_s, f_n \in F.$$

The proof of Proposition 4.31, part 1 shows: Semisimpleness of $f_s \in End_{\mathbb{C}}(V)$ implies that f_s is ad_F-semisimple, i.e.

$$ad_F(f_s) \in End_{\mathbb{C}}(F)$$

is semisimple. And nilpotency of f_n implies ad_F-nilpotency, see Lemma 3.2. Hence by Definition 4.32

$$\rho(x) = f_s + f_n$$

is also the abstract Jordan decomposition of $\rho(x) \in F$. From the uniqueness of the abstract Jordan decomposition in F derives

$$f_s = \rho(s) \text{ and } f_n = \rho(n).$$

□

Part II
Complex semisimple Lie algebras

Chapter 5
Root space decomposition

The base field in this chapter is $\mathbb{K} = \mathbb{C}$, the field of complex numbers. All Lie algebras are *complex* Lie algebras unless stated otherwise.

The present chapter starts to investigate the structure of complex semisimple Lie algebras. The subject is a classical topic of mathematics from the 20th century. An excellent overview of Chapter 5 and 6 and the outlook in Chapter 7 is given by Knapp as a survey to Chapter II of his book [31]. At this point we will only name the keywords: Maximal toral subalgebra, Cartan subalgebra, root space decomposition, root system, Cartan matrix, Weyl group, Coxeter graph, Dynkin diagram.

According to Theorem 4.21 each semisimple Lie algebras splits as the direct sum of simple Lie algebras. The simple complex Lie algebras are completely classified. They are the members of the *ABCD*-series from Proposition 2.15 together with five exceptional Lie algebras.

The most elementary member, and at the same time the prototype of the *ABCD*-series is the complex simple Lie algebra $sl(2,\mathbb{C})$ of type A_1. The canonical basis elements of $sl(2,\mathbb{C})$

$$h := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ x := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ y := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \in sl(2,\mathbb{C})$$

make up two classes: The elements h is *ad*-semisimple and gerates a maximal Abelian subalgebra. While the second class contains the two elements x and y. They are eigenvectors of h under the adjoint representation

$$(ad\ h)(x) = [h,x] = 2x,\ (ad\ h)(y) = [h,y] = -2y,$$

and their commutator is

$$[x,y] = h.$$

These two classes generalize for a semisimple Lie algebra L to the concept of a *maximal toral subalgebra* T of L, and the generators of the *root spaces* of L with respect to T.

5.1 Toral subalgebra

Consider a Lie algebra L. We recall from Definition 3.4: An element $x \in L$ is *ad-semisimple* iff the endomorphism

$$ad\ x : L \to L,\ y \mapsto [x,y],$$

is semisimple.

It is well-known from Linear Algebra that a set of commuting semisimple endomorphisms can be simultaneously diagonalized. Therefore, starting from one ad-semisimple element of L one tries to find all pairwise commuting ad-semisimple elements of L, which also commute with the given one.

- The first result about a semisimple Lie algebra L states: Any subalgebra $T \subset L$ with only ad-semisimple elements, named a *toral subalgebra*, is Abelian, see Proposition 5.2.

- The second result, Theorem 5.17, shows that no further element of L commutes with a maximal toral subalgebra $T \subset L$, i.e. the centralizer of T satisfies

$$C_L(T) = T.$$

Definition 5.1 (Toral subalgebra). Consider a Lie algebra L.

- A *toral subalgebra* of L is a subalgebra $T \subset L$ with all elements $x \in T$ ad-semisimple.

- A toral subalgebra $T \subset L$ is a *maximal toral subalgebra* iff T is not properly contained in any other toral subalgebra of L.

We show: The existence of non-zero toral subalgebras of a non-zero semisimple Lie algebra follows from Engel's theorem and the abstract Jordan decomposition. The result, together with Theorem 5.17, is fundamental for the root space decomposition from Definition 5.18.

Proposition 5.2 (Existence of non-zero toral subalgebras). *Consider a semisimple Lie algebra L.*

1. *If $L \neq \{0\}$ then exists a non-zero toral subalgebra in L, hence also a non-zero maximal toral subalgebra.*

2. *Each toral subalgebra of L is Abelian.*

Proof. 1. *Existence of a non-zero toral subalgebra*: If each element $x \in L$ were ad-nilpotent, then Engel's Theorem 3.10 would imply that L is nilpotent, a contradiction to the semisimpleness of L. Hence we can choose an element $x \in L$ with abstract Jordan decomposition

$$x = s + n \text{ and } ad\ s \neq 0.$$

The 1-dimensional subalgebra

$$\mathbb{C} \cdot s \subset L$$

is a toral subalgebra. Because L is finite-dimensional, there also exists a maximal toral subalgebra.

2. *Toral subalgebras are Abelian*: Consider a toral subalgebra $T \subset L$. For a pair of non-zero vectors $x, y \in T$ we have to show

$$[x,y] = 0$$

Because $T \subset L$ is a subalgebra, the toral subalgebra T is stable with respect to $ad\ x$ and with respect to $ad\ y$. Both are semisimple. According to Lemma 1.18 also their restrictions

$$ad_x := (ad\ x)|T \text{ and } ad_y := (ad\ y)|T$$

are semisimple.

Because T is spanned by eigenvectors of ad_x, we may assume $y \in T$ as an eigenvector of ad_x, i.e. for a suitable $\lambda \in \mathbb{C}$

$$ad_x(y) = [x,y] = \lambda \cdot y.$$

We develop $x \in T$ with respect to a basis $(y_j)_{j \in J}$ of T of eigenvectors of ad_y with eigenvalues $(\lambda_j)_{j \in J}$

$$x = \sum_{j \in J} \alpha_j \cdot y_j.$$

Then

$$-\lambda \cdot y = -[x,y] = [y,x] = ad_y(x) = \sum_{j \in J} (\lambda_j \cdot \alpha_j) \cdot y_j.$$

Assume $\lambda \neq 0$. Because $y \neq 0$, for at least one $j \in J$

$$\lambda_j \cdot \alpha_j \neq 0, \text{ notably } \lambda_j \neq 0.$$

Apparently, the vector $y \in T$ is an eigenvector of ad_y with eigenvalue 0. The representation

$$-\lambda \cdot y = \sum_{j \in J} (\lambda_j \cdot \alpha_j) \cdot y_j$$

shows, that y is a linear combination of eigenvectors belonging to eigenvalues $\lambda_j \neq 0$, a contradiction. Therefore $\lambda = 0$, which implies

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

□

Lemma 5.3 will be used later.

Lemma 5.3 (Centralizer of a maximal toral subalgebra).

Consider a pair (L, T) with L a semisimple Lie algebra and $T \subset L$ a maximal toral subalgebra.

i) The centralizer $C_L(T)$ contains with each element $x \in C_L(T)$ also the ad-semisimple and the ad-nilpotent part

$$s,\ n \in C_L(T)$$

from the abstract Jordan decomposition $x = s + n$.

ii) Each ad-semisimple element $x \in C_L(T)$ belongs to T.

Proof. Set $C := C_L(T)$.

i) *Abstract Jordan decomposition*: Consider $x \in C$ with abstract Jordan decomposition

$$x = s + n.$$

Then

$$ad\ x = ad\ s + ad\ n \in End(L)$$

is the Jordan decomposition of the endomorphism $ad\ x \in End(L)$. In particular

$$ad\ s = p_s(ad\ x) \text{ and } ad\ n = p_n(ad\ x)$$

with polynomials $p_s(Z), p_n(Z) \in \mathbb{C}[Z]$ satisfying $p_s(0) = p_n(0) = 0$.

As a consequence: For each $h \in T$ with $(ad\ x)(h) = 0$ also

$$(ad\ s)(h) = 0 \text{ and } (ad\ n)(h) = 0.$$

Hence $s,\ n \in C$.

ii) *Each ad-semisimple element $x \in C$ belongs to T*: Any ad-semisimple element $x \in C$ commutes with all elements from T. Therefore all elements from

$$span_{\mathbb{C}} < x, T >$$

are pairwise commuting and therefore ad-semisimple. The maximality of T implies

$$span_{\mathbb{C}} < x, T > = T$$

i.e. $x \in T$. $\square$

5.2 Structure and representations of $sl(2,\mathbb{C})$

The 3-dimensional complex Lie algebra $sl(2,\mathbb{C})$ is the prototype of complex semisimple Lie algebras. Its representation theory is of fundamental importance:

- The representation theory of $sl(2,\mathbb{C})$ is the means to clarify the structure of general complex semisimple Lie algebras. Proposition 5.4 presents the structure of $sl(2,\mathbb{C})$ in a form which generalizes to arbitrary complex semisimple Lie algebras, cf. Proposition 7.3 about the root space decomposition.

- The representation theory of $sl(2,\mathbb{C})$ is also paradigmatic for the representation theory of general semisimple Lie algebras. Proposition 5.7, Corollary 5.8 and Theorem 5.10 present the representation theory of $sl(2,\mathbb{C})$-modules in a form which generalizes to representations of arbitrary complex semisimple Lie algebras.

Proposition 5.4 (Structure of $sl(2,\mathbb{C})$). *The Lie algebra*

$$L := sl(2,\mathbb{C})$$

has the standard basis $\mathscr{B} := (h, x, y)$ *with matrices*

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \in sl(2,\mathbb{C}).$$

- *The non-zero commutators of the elements from $\mathscr{B}$ are*

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

- *With respect to $\mathscr{B}$ the matrices of the adjoint representation*

$$ad : L \to gl(L)$$

are

$$ad\ h = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \ ad\ x = \begin{pmatrix} 0 & 0 & 1 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ ad\ y = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}$$

- *The element $h \in L$ is ad-semisimple. The endomorphism*

  $$ad\ h : L \to L$$

  has the eigenvalues

  $$0,\ \alpha := 2,\ -\alpha$$

  with corresponding eigenspaces

  $$H := L^0 = \mathbb{C} \cdot h,\ L^\alpha = \mathbb{C} \cdot x,\ L^{-\alpha} = \mathbb{C} \cdot y.$$

  In particular

  $$L = H \oplus (L^\alpha \oplus L^{-\alpha})$$

  as a direct sum of complex vector spaces.

- *The Lie algebra L is simple.*

- *The Abelian subalgebra $H \subset L$ is a maximal toral subalgebra of L.*

- *With respect to the basis $\mathscr{B}$ the Killing form κ of L, see Definition 4.2, has the symmetric matrix with integer coefficients*

  $$4 \cdot \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \in M(3 \times 3, \mathbb{Z}).$$

  The Killing form is non-degenerate.

- *The restriction*

  $$\kappa_H := \kappa | (H \times H)$$

  of the Killing form is positive definite

  $$\kappa_H(h,h) = 8.$$

  The scalar product κ_H induces the isomorphism of the maximal toral subalgebra H and its dual space

  $$j_{\kappa_H} : H \xrightarrow{\sim} H^*,\ h \mapsto \kappa_H(h,-) = 8 \cdot h^*.$$

  The element $t_\alpha \in H$, which is defined as

  $$j_{\kappa_H}(t_\alpha) = \alpha \cdot h^*,$$

is
$$t_\alpha = \frac{h}{4} \text{ satisfying } h = \frac{2 \cdot t_\alpha}{\kappa_H(t_\alpha, t_\alpha)}.$$

Proof. Only the following claims need a separate proof:

i) *Simpleness*: Assume the existence of a proper ideal
$$I \subsetneq L.$$

The ideal I satisfies $[I,L] \subset I$.

- First, $h \notin I$: Otherwise $2x = [h,x] \in I$ and $-2y = [h,y] \in I$, which implies $I = L$, a contradiction.
- Secondly, $x \notin I$: Otherwise $h = [x,y] \in I$, contradicting the first part.
- Thirdly, $y \notin I$: Otherwise $-h = [y,x] \in I$, contradicting the first part.

Consider an element
$$z = \alpha \cdot x + \beta \cdot y + \gamma \cdot h \in I \text{ with } \alpha, \beta, \gamma \in \mathbb{C}.$$

The commutator relations imply
$$(ad\ x)(z) = \beta \cdot [x,y] + \gamma \cdot [x,h] = \beta \cdot h - 2 \cdot \gamma \cdot x$$

and
$$(ad\ x)^2(z) = -2 \cdot \beta \cdot x \in I$$

Similarly
$$(ad\ y)^2(z) = -2 \cdot \alpha \cdot y \in I$$

If $\alpha \neq 0$ or $\beta \neq 0$ then $x \in I$ or $y \in I$, a contradiction. Hence
$$\alpha = \beta = 0.$$

If $\gamma \neq 0$ then $h \in I$, a contradiction. Hence $z = 0$. As a consequence
$$I = \{0\} \text{ and } L \text{ is simple.}$$

ii) *Maximal toral subalgebra*: To prove that the toral subalgebra $H \subset L$ is a maximal toral subalgebra, we show
$$C_L(H) = H.$$

Assume
$$z = \alpha \cdot h + \beta \cdot x + \gamma \cdot y \in C_L(H);\ \alpha, \beta, \gamma \in \mathbb{C}.$$

Then
$$[\alpha \cdot h + \beta \cdot x + \gamma \cdot y, h] = -2\beta \cdot x + 2\gamma \cdot y = 0 \iff \beta = \gamma = 0,$$

hence $z \in H$. Proposition 5.2 ensures that toral subalgebras are Abelian, which proves that H is a maximal toral subalgebra. $\square$

The simpleness of $L := sl(2,\mathbb{C})$ also follows from the general theory: According to Corollary 4.18 the Lie algebra L is semisimple. Hence L splits due to Theorem 4.21 as a direct sum of simple Lie algebras. If the splitting of the 3-dimensional Lie algebra L comprises at least two simple summands, then at least one of them is 1-dimensional and therefore Abelian, which contradicts its simpleness.

Remark 5.5 (Basis of Pauli matrices). A different basis of $sl(2,\mathbb{C})$ is the family $(\sigma_j)_j = 1,2,3$ of the Pauli matrices, see Remark 2.21: We have

$$h = \sigma_3,\ x = \frac{\sigma_1 + i\sigma_2}{2},\ y = \frac{\sigma_1 - i\sigma_2}{2}$$

We now investigate the theory of finite-dimensional representations of

$$L := sl(2,\mathbb{C}).$$

The element $h \in L$ is ad-semisimple because $ad\ h \in End(L)$ is semisimple. Therefore $h \in L$ coincides with its semisimple component in the abstract Jordan decomposition of L. Corollary 4.33 shows the far reaching consequences: The element h acts as semisimple endomorphism on each L-module V. Hence each L-module decomposes as a direct sum of eigenspaces with respect to the action of h.

Definition 5.6 introduces some basic concepts from the representation theory of semisimple Lie algebras.

Definition 5.6 (Weight, weight space and primitive element). Consider an $sl(2,\mathbb{C})$-module V, not necessarily finite-dimensional, with respect to a representation

$$\rho : L \to gl(V).$$

1. For $\lambda \in \mathbb{C}$ set
$$V^\lambda = \{\mathrm{v} \in V : \rho(h)(\mathrm{v}) = \lambda \cdot \mathrm{v}\}.$$
If $V^\lambda \neq 0$ then $\lambda \in \mathbb{C}$ is a *weight* of V, the eigenspace V^λ of $\rho(h) \in End(L)$ is a *weight space* of V, and the non-zero elements of V^λ are named *weight vectors*.

2. A weight vector $e \in V^\lambda$ is named a *primitive element* of V with weight λ if
$$\rho(x)(e) = 0$$
with respect to the action of the element x from the standard basis $\mathscr{B}$ of $sl(2,\mathbb{C})$.

Note: Different than in Chapter 1 we denote here and in the following the eigenspace with eigenvalue λ by an *upper* index λ.

For any $sl(2,\mathbb{C})$-module V not only the action of $h \in sl(2,\mathbb{C})$, but also the action of the two other elements of the standard basis

$$\mathscr{B} = (h, x, y)$$

can be easily described. Hereby the role of a primitive element can be parafrased as "germ" of an $sl(2,\mathbb{C})$-module.

Proposition 5.7 (Action of the standard basis). *Consider an $sl(2,\mathbb{C})$-module V, not necessarily finite-dimensional. Assume the existence of a primitive element $e \in V$ with weight $\lambda \in \mathbb{C}$. Then the elements*

$$e_i := \frac{1}{i!} \cdot (y^i.e) \in V,\ i \geq 0,$$

satisfy for all $i \geq 0$:

1. Weight vector: $h.e_i = (\lambda - 2i) \cdot e_i$.
2. Lowering the weight: $y.e_i = (i+1) \cdot e_{i+1}$.
3. Raising the weight: $x.e_i = (\lambda - i + 1) \cdot e_{i-1},\ e_{-1} := 0$.
4. Linear dependency respectively independency:
   - *Either the family $(e_i)_{i \geq 0}$ is linearly independent*
   - *or the highest weight λ is a non-negative integer, the family*

     $$(e_i)_{i=0,\dots,\lambda}$$

     is linearly independent, and $e_i = 0$ for all $i > \lambda$.

Figure 5.1 illustrates the content of Proposition 5.7.

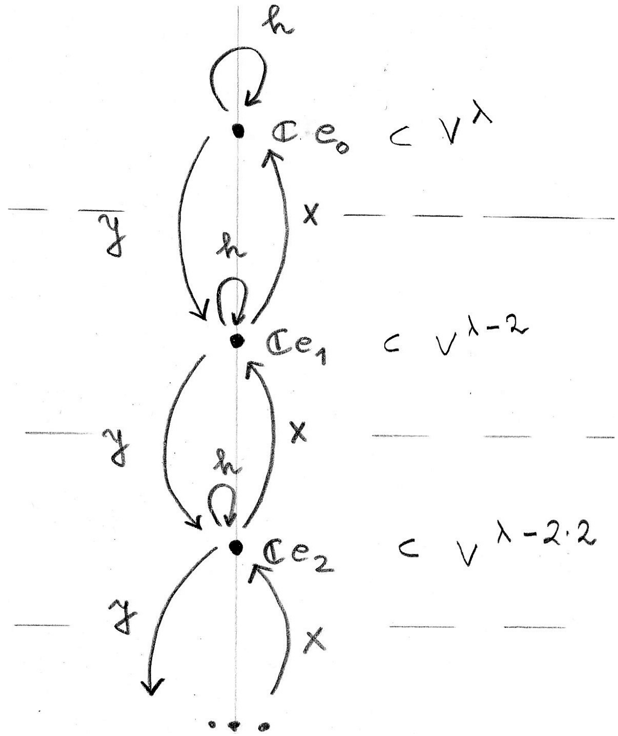

Fig. 5.1 Lowering/raising weights in irreducible $sl(2,\mathbb{C})$-modules

Proof. ad 2) By definition

$$y.e_i = \frac{1}{i!} \cdot (\ y.(y^i.e)) = \frac{1}{i!} \cdot (y^{i+1}.e) = \frac{(i+1)!}{i!} \cdot e_{i+1} = (i+1) \cdot e_{i+1}.$$

ad 1) By induction on $i \in \mathbb{N}$: $i = 0$ by definition.
Induction step $i \mapsto i+1$: Due to part 2)

$$(i+1) \cdot (h.e_{i+1}) = h.(y.e_i) = [h,y].e_i + y.(h.e_i) = -2y.e_i + y.((\lambda - 2i)e_i) =$$

$$= (\lambda - 2i - 2)\ (y.e_i) = (\lambda - 2i - 2) \cdot (i+1) \cdot e_{i+1},$$

hence

$$h.e_{i+1} = (\lambda - 2(i+1)) \cdot e_{i+1}.$$

ad 3) With the definition $e_{-1} := 0$ the formula

$$x.e_i = (\lambda - i + 1) \cdot e_{i-1}$$

follows by induction on $i \in \mathbb{N}$ by using the commutator formula

$$[x,y] = h \in sl(2,\mathbb{C})$$

together with the result from part 1) and part 2). The formula holds for $i = 0$ because e is a primitive element.
Induction step $i \mapsto i+1$:

$$(i+1)\cdot(x.e_{i+1}) = x.(y.e_i) = [x,y].e_i + y.(x.e_i) = h.e_i + y.((\lambda - i + 1)\cdot e_{i-1}) =$$

$$= (\lambda - 2i)\cdot e_i + i\cdot(\lambda - i + 1)\cdot e_i = (i+1)\cdot(\lambda - i)\cdot e_i$$

and after dividing by $(i+1)$

$$x.e_{i+1} = (\lambda - i)\cdot e_i$$

ad 4) The non-zero elements e_i are weight vectors with pairwise distinct weights, i.e. different eigenvalues. Hence they are linearly independent. The family $(e_i)_{i\geq 0}$ is linearly independent if $e_i \neq 0$ for all $i \in \mathbb{N}$.

Otherwise there exists a largest index $m \in \mathbb{Z}_+ = \{0,1,...\}$ with all elements $e_0,...,e_m$ non-zero. Then $e_i = 0$ for all $i > m$. We apply the formula from part 3) and obtain

$$0 = x.e_{m+1} = (\lambda - m)\cdot e_m$$

which implies

$$\lambda - m = 0 \text{ or } \lambda = m \in \mathbb{Z}_+.$$

□

Corollary 5.8 (Primitive element). *Consider a finite-dimensional $sl(2,\mathbb{C})$-module V.*

1. *If V is irreducible, then V has a primitive element.*
2. *Each primitive element $e \in V$ has a weight $\lambda \in \mathbb{Z}_+$. It generates an irreducible $sl(2,\mathbb{C})$-submodule*

$$V(\lambda) := span < e_i : \ i = 0,....,\lambda > \subset V$$

   using the notation from Proposition 5.7.

3. *A primitive element of an irreducible $sl(2,\mathbb{C})$-module is uniquely determined up to a scalar from $\mathbb{C}^*$.*

Proof. 1. *Existence of a primitive element*: Set

$$L := sl(2,\mathbb{C}).$$

We have $V \neq \{0\}$ because V is irreducible. Denote by

$$\rho : L \to gl(V)$$

the representation which defines the L-module structure on V. The kernel

$$ker\ \rho \subset L$$

is an ideal in the simple Lie algebra L. Depending on $ker\ \rho$ we distinguish two cases:

- If

  $$ker\ \rho = L$$

  then each non-zero element $e \in V$ is a primitive element and has weight $\lambda = 0 \in \mathbb{Z}_+$.

- Otherwise

  $$ker\ \rho = 0$$

  and

  $$\rho : L \to gl(V)$$

  is injective. The subalgebra

  $$B := span_{\mathbb{C}} < h, x > \subset L$$

  is solvable because

  $$[h,x] = 2x \text{ and therefore } D^2 B = 0$$

  The restriction

  $$\rho | B : B \to gl(V)$$

  is injective. According to Theorem 3.20, the forerunner of Lie's theorem, the subalgebra

  $$B \simeq \rho(B) \subset gl(V)$$

  has a common eigenvector $e \in V$. Consider the eigenvalues $\lambda_h,\ \lambda_x$ defined by

  $$h.e = \lambda_h \cdot e \text{ and } x.e = \lambda_x \cdot e.$$

  The commutator $[h,x] = 2x$ implies

  $$2\lambda_x \cdot e = 2x.e = [h,x].e = h.(x.e) - x.(h.e) = \lambda_h \lambda_x \cdot e - \lambda_x \lambda_h \cdot e = 0,$$

  hence $\lambda_x = 0$. Therefore $e \in V$ is a primitive element.

2. *The primitive element as "germ"*: Let $e \in V$ be a primitive element with weight $\lambda \in \mathbb{C}$. Because V is finite-dimensional, Proposition 5.7 implies:

$$\lambda \in \mathbb{Z}_+$$

and successive application of $y \in L$ lowers the weight down to the value $-\lambda$.

The vector space

$$V(\lambda) = span_{\mathbb{C}} < e_i : \ i = 0, ..., \lambda > \subset V$$

is a submodule of V. In $V(\lambda)$ each weight space with weight μ is 1-dimensional, generated by the element e_i with weight $\mu = \lambda - 2i$.

Any non-zero submodule

$$W' \subset V(\lambda)$$

contains at least one weight vector, because

$$\rho(h)|W' : W' \to W'$$

has an eigenvector, i.e. a weight vector of $V(\lambda)$ is contained in W'. Hence for at least one index $i = 0, ..., \lambda$ holds

$$e_i \in W'.$$

The formulas from Proposition 5.7 for raising and lowering the weight imply that W' contains also e and a posteriori the whole L-module $V(\lambda)$, cf. Figure 5.1. Therefore

$$W' = V(\lambda),$$

which proves the irreducibility of $V(\lambda)$.

3. *Primitive element*: Due to part 2) each weight space of an irreducible $sl(2,\mathbb{C}$-module is 1-dimensional.

   □

Remark 5.9 (Eigenspaces of the angular momentum vector).

1. *The real Lie algebra $so(3,\mathbb{R})$ of infinitesimal rotations*: The Lie algebra $so(3,\mathbb{R})$ is the Lie algebra of the rotation group $SO(3,\mathbb{R})$. The infinitesimal generators of the 1-parameter subgroups of rotations around the coordinate axes are the elements

$$J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \ J_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \ J_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in so(3,\mathbb{R}).$$

   Their commutator relations are

$$[J_x, J_y] = J_z$$

   and cyclic permutation, see Remark 2.21.

2. *The complex Lie algebra* $so(3,\mathbb{C})$ *of the angular momentum*: The 3 generators

$$J_1 := iJ_x,\ J_2 := iJ_y,\ J_3 := iJ_z \in so(3,\mathbb{C}),$$

of the complexification

$$so(3,\mathbb{C}) \simeq so(3,\mathbb{R}) \otimes_{\mathbb{R}} \mathbb{C}$$

satisfy the commutator relation

$$[J_1, J_2] = i \cdot J_3$$

and cyclic permutation. The complexification provides the map

$$sl(2,\mathbb{C}) \xrightarrow{\simeq} so(3,\mathbb{C}),\ h/2 \mapsto J_3,\ x \mapsto J_+,\ y \mapsto J_-,$$

with

$$J_+ := J_1 + i \cdot J_2 \text{ and } J_- := J_1 - i \cdot J_2,$$

which is an isomorphism of Lie algebras due to the commutator relations

$$[J_3, J_+] = J_+,\ [J_3, J_-] = -J_-, [J_+, J_-] = 2 \cdot J_3,$$

note the factor $\frac{1}{2} \cdot h$ within the definition of the isomorphism.

The generators

$$J_1, J_2, J_3 \in so(3,\mathbb{C})$$

are Hermitian matrices. Therefore they have real eigenvalues, and represent three quantum mechanical observables in the Hilbert space $\mathbb{C}^3$. The complex Lie algebra $so(3,\mathbb{C})$ is named the Lie algebra of the *angular momentum vector*, the vector of three observables

$$\vec{J} := (J_1, J_2, J_3).$$

3. *The square of the angular momentum vector*: Leaving the Lie algebra $so(3,\mathbb{C})$ we consider in the associative algebra $M(3 \times 3,\mathbb{C})$ the scalar product

$$J^2 := \vec{J}^{\,2} = J_1^2 + J_2^2 + J_3^2 \in M(3 \times 3,\mathbb{C}).$$

The matrix J^2 is also Hermitian, hence also J^2 represents a quantum mechanical observable in the Hilbert space $\mathbb{C}^3$. In addition, the matrix J^2 is positive semidefinite.

One checks the equations

$$[J_1^2, J_3] = J_1J_1J_3 - J_3J_1J_1 = J_1[J_1, J_3] + J_1J_3J_1 - J_3J_1J_1 =$$

$$= J_1[J_1, J_3] + [J_1, J_3]J_1 = -i \cdot J_1J_2 - i \cdot J_2J_1$$

and similarly
$$[J_2^2, J_3] = i \cdot J_2 J_1 + i \cdot J_1 J_2$$
Hence
$$[J_1^2 + J_2^2, J_3] = 0 \text{ and } [J^2, J_3] = [J_3^2, J_3] = 0.$$
Hence
$$[J^2, J_3] = 0,$$
and the matrices J^2 and J_3 can be diagonalized simultaneously with real eigenvalues. By symmetry also
$$[J^2, J_1] = [J^2, J_2] = 0.$$

4. *Complex representations*: The Lie algebras
$$su(2) \simeq so(3,\mathbb{R}) \text{ and } sl(2,\mathbb{R})$$
have the same complexification
$$sl(2,\mathbb{C}) \simeq su(2) \otimes_{\mathbb{R}} \mathbb{C} = so(3,\mathbb{R}) \otimes_{\mathbb{R}} \mathbb{C} = so(3,\mathbb{C})$$
of type $A_1 = B_1$. Hence the Lie algebras
$$su(2),\ so(3,\mathbb{R}),\ sl(2,\mathbb{R}),\ sl(2,\mathbb{C}),\ so(3,\mathbb{C})$$
have the same *complex* represesentations. These representations correspond bijectively to the complex representations of the simply connected, real matrix group $SU(2)$, see Figure 5.2.

$$\begin{tikzcd} SU(2) \arrow[r, "\rho"] & GL(V) \\ su(2) \arrow[u, "exp"] \arrow[r, "Lie\ \rho"] & gl(V) \arrow[u, "exp"'] \end{tikzcd}$$

Fig. 5.2 Representations of $su(2)$ and $SU(2)$

The group $SU(2)$ is the universal covering of the rotation group $SO(3,\mathbb{R})$, see Example 2.24. The covering projection
$$p : SU(2) \to SO(3,\mathbb{R})$$

is a 2-fold covering with kernel

$$ker\ p = \{\pm \mathbb{1}\}$$

The covering projection induces a bijection between the set of complex irreducible representations $\overline{\rho}$ of $SO(3,\mathbb{R})$ and the set of those complex irreducible representations ρ of $SU(2)$ with

$$\rho(-\mathbb{1}) = id_V$$

according to the commutative diagram from Figure 5.3

$$\begin{tikzcd} SU(2) \arrow[r, "\rho"] \arrow[d, "p"'] & GL(V) \\ SO(3,\mathbb{R}) \arrow[ur, dashed, "\overline{\rho}"] & \end{tikzcd}$$

Fig. 5.3 Representations of $SU(2)$ and $SO(3,\mathbb{R})$

5. *Ladder operators of* $so(3,\mathbb{C})$: The matrices

$$J_{\pm} = J_1 \pm i \cdot J_2 \in so(3,\mathbb{C})$$

are Hermitian conjugate to each other

$$(J_{\pm})^* = J_{\mp}$$

We now prove that $J_{\pm}$ are the "ladder" operators of $so(3,\mathbb{C})$-modules. For the proof we use the isomorphy from part 3

$$sl(2,\mathbb{C}) \xrightarrow{\simeq} so(3,\mathbb{C}),\ h/2 \mapsto J_3,\ x \mapsto J_+,\ y \mapsto J_-,$$

and carry over the result of Proposition 5.7 from $sl(2,\mathbb{C})$-modules to $so(3,\mathbb{C})$-modules:

Consider a finite-dimensional, irreducible $so(3,\mathbb{C})$-module V. Denote by $e \in V$ a primitive element when considering V as $sl(2,\mathbb{C})$-module under the isomorphism

$$sl(2,\mathbb{C}) \simeq so(3,\mathbb{C})$$

from part 2. It satisfies the eigenvalue equation

$$J_3.e = m \cdot e,\ m \in \frac{1}{2}\,\mathbb{Z}_+ \textit{ (half-integer)} \text{ and } J_+.e = 0.$$

The action of J_+ raises the eigenvalue of J_3 by 1, while the action of J_- lowers the eigenvalue by 1: For each eigenvector $\mathrm{v} \in V$ of the action of J_3 with

$$J_3.\mathrm{v} = k \cdot \mathrm{v},\ k \in \mathbb{R},$$

holds

$$J_3.(J_+.\mathrm{v}) = ([J_3, J_+] + J_+ J_3).\mathrm{v} = (J_+ + k \cdot J_+).\mathrm{v} = (1+k) \cdot (J_+.\mathrm{v})$$

and

$$J_3.(J_-.\mathrm{v}) = ([J_3, J_-] + J_- J_3).\mathrm{v} = (-J_- + k \cdot J_-).\mathrm{v} = (k-1) \cdot (J_-.\mathrm{v}).$$

The primitive element $e \in V$ is also an eigenvector of the action of J^2: The equation

$$J_\pm J_\mp = J_1^2 + J_2^2 \pm J_3 = J^2 - J_3^2 \pm J_3$$

implies

$$J^2 = J_+ J_- + J_3^2 - J_3$$

and

$$J^2 = J_- J_+ + J_3^2 + J_3$$

From the last equation and

$$J_+.e = 0$$

follows

$$J^2.e = J_3^2.e + J_3.e = m^2 \cdot e + m \cdot e = m(m+1) \cdot e.$$

Raising or lowering by $J_\pm$ the eigenvalue of J_3 does not change the eigenvalue of J^2: If

$$J^2.\mathrm{v} = \lambda \cdot \mathrm{v},\ \lambda \in \mathbb{R},$$

then due to $[J^2, J_\pm] = 0$ also

$$J^2.(J_\pm.\mathrm{v}) = ([J^2, J_\pm] + (J_\pm.J^2).\mathrm{v} = J_\pm.(J^2.\mathrm{v}) = \lambda \cdot (J_\pm.\mathrm{v})$$

Different from the basis of V, which derives according to Proposition 5.7 from the basis with elements

$$e_i := \frac{1}{i!} \cdot (y^i.e), i \geq 0,$$

textbooks from physics like [2, Eq. 5.74] consider the basis of V

$$(\mathrm{v}_k)_{-m \leq k \leq m}$$

successively defined as

$$\mathrm{v}_{k-1} := \frac{1}{\sqrt{(m+k)(m-k+1)}} \cdot (J_-.\mathrm{v}_k),\ -m+1 \leq k \leq m,\ \mathrm{v}_m := e.$$

These vectors satisfy, see [2, Eq. 5.71]:

$$J^2.\mathrm{v}_k = m(m+1)\cdot \mathrm{v}_k \text{ (\textit{Total angular momentum})}$$

and

$$J_3.\mathrm{v}_k = k\cdot \mathrm{v}_k,\ -m \le k \le m \text{ (\textit{Angular momentum around the z-axis})}.$$

Hence applying the ladder operators $J_\pm$ moves through the eigenspace of J^2 along the different eigenvectors of J_3.

The construction of the $sl(2,\mathbb{C})$-modules

$$V(\lambda),\ \lambda \in \mathbb{Z}_+,$$

from Proposition 5.7 generates all irreducible $sl(2,\mathbb{C})$-modules.

Theorem 5.10 (Classification of all finite-dimensional irreducible $sl(2,\mathbb{C})$-modules).

1. *For each $\lambda \in \mathbb{Z}_+$ exists a finite-dimensional, irreducible $sl(2,\mathbb{C})$-module with a primitive element of weight λ.*
   Each finite-dimensional, irreducible $sl(2,\mathbb{C})$-module with a primitive element of weight $\lambda \in \mathbb{Z}_+$ is isomorphic to the $sl(2,\mathbb{C})$-module $V(\lambda)$ from Corollary 5.8.

   The $sl(2,\mathbb{C})$-module $V(\lambda)$ splits in the category of vector spaces as the direct sum of 1-dimensional weight spaces

$$V(\lambda) = \bigoplus_{i=0}^{\lambda} V^{\lambda-2i}.$$

   with integer weights.

2. *The map to the isomorphy classes of finite-dimensional, irreducible $sl(2,\mathbb{C})$-modules*

$$\mathbb{Z}_+ \to \{[V] : V \text{ \textit{finite-dimensional, irreducible} } sl(2,\mathbb{C})\text{\textit{-module}}\},\ \lambda \mapsto [V(\lambda)],$$

   is bijective.

Proof. Set $L := sl(2,\mathbb{C})$.

1. *Existence*: Choose a vector space V of dimension $\lambda+1$ with basis $(e_0,...,e_\lambda)$. Define

$$L \times V \to V$$

   by linear etension of the formulas from Proposition 5.7:

- $h.e_i := (\lambda - 2i) \cdot e_i$
- $y.e_i := (i+1) \cdot e_{i+1},\ e_{\lambda+1} := 0$
- $x.e_i := (\lambda - i + 1) \cdot e_{i-1},\ e_{-1} := 0$

One checks, that in accordance with the commutator relations of $sl(2,\mathbb{C})$

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

these definitions satisfy the equations

- $h.(x.e_i) - x.(h.e_i) = 2x.e_i$
- $h.(y.e_i) - y.(h.e_i) = -2y.e_i$
- $x.(y.e_i) - y.(x.e_i) = h.e_i$

Hence V becomes the irreducible L-module $V(\lambda)$ from Corollary 5.8 with primitive element $e_0 \in V$.

2. *Classification*: According to part 1 the map from the theorem is well-defined and surjective. Concerning its injectivity: If $[V(\lambda_1)] = [V(\lambda_2)]$ then

$$V(\lambda_1) \simeq V(\lambda_2),$$

in particular

$$1 + \lambda_1 = dim\, V(\lambda_1) = dim\, V(\lambda_2) = 1 + \lambda_2,$$

hence

$$\lambda_1 = \lambda_2.$$

□

Combining Theorem 5.10 with the results displayed in Figure 5.2 and 5.3 shows: The complex finite-dimensional irreducible representations of $SO(3,\mathbb{R})$ correspond bijectively to the $sl(2,\mathbb{C})$-modules $V(\lambda)$ with highest weight $\lambda \in 2 \cdot \mathbb{Z}_+$.

Corollary 5.11 (Classification of all finite-dimensional $sl(2,\mathbb{C})$-modules). *Each finite-dimensional $sl(2,\mathbb{C})$-module V is isomorphic to a finite direct sum of irreducible modules of the type $V(\lambda)$ from Corollary 5.8*

$$V = \bigoplus_{\lambda \in \mathbb{Z}_+} n_\lambda \cdot V(\lambda),\ n_\lambda \in \mathbb{N},$$

with multiplicity $n_\lambda \neq 0$ for at most finitely many λ.

Proof. The proof follows from Weyl's theorem on complete reducibility, see Theorem 4.30, and Theorem 5.10. □

Example 5.12 (Explicit realization of the finite-dimensional irreducible $sl(2,\mathbb{C})$-modules by homogeneous polynomials).

Set $L := sl(2,\mathbb{C})$.

1. The irreducible L-module of highest weight $\lambda = 0$ is the 1-dimensional vector space $\mathbb{C}$ with the *trivial representation*
$$\rho : sl(2,\mathbb{C}) \to \{0\} \subset gl(\mathbb{C}).$$
Its weight space decomposition is
$$V(0) = V^0 \simeq \mathbb{C}.$$
Any non-zero element $e \in \mathbb{C}$ is a primitive element.

2. The irreducible L-module of highest weight $\lambda = 1$ is the 2-dimensional vector space $\mathbb{C}^2$ with the *tautological representation*
$$\rho : L = sl(2,\mathbb{C}) \hookrightarrow gl(\mathbb{C}^2).$$
Its weight space decomposition is
$$V(1) = V^1 \oplus V^{-1}$$
with
$$V^1 = \mathbb{C}\cdot\begin{pmatrix}1\\0\end{pmatrix},\ V^{-1} = \mathbb{C}\cdot\begin{pmatrix}0\\1\end{pmatrix}$$
and primitive element
$$e = \begin{pmatrix}1\\0\end{pmatrix},$$
because
$$x.e = \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \in \mathbb{C}^2$$

3. The irreducible L-module of highest weight $\lambda = 2$ is the 3-dimensional vector space L itself, considered as L-module with respect to the *adjoint representation*
$$ad : L \to gl(L).$$
Proposition 5.4 shows the weight space decomposition
$$V(2) = L = L^2 \oplus L^0 \oplus L^{-2}$$

with primitive element $e = x \in \mathscr{B}$.

4. In general, the irreducible L-module $V(\lambda)$ of highest weight $\lambda = n \in \mathbb{Z}_+$ is isomophic to the complex vector space of complex *homogeneous polynomials* in two variables
$$P(u,\mathrm{v}) \in \mathbb{C}[u,\mathrm{v}]$$
of degree $= n$.

The vector space $\mathbb{C}[u,\mathrm{v}]$ of polynomials in two variables u and v has a basis of monomials $(u^\mu \cdot \mathrm{v}^\nu)_{\mu,\nu\in\mathbb{N}}$. A homogeneous polynomial of degree $n \in \mathbb{N}$ is an element
$$P(u,\mathrm{v}) = \sum_{\mu=0}^{n} a_\mu \cdot u^\mu \cdot \mathrm{v}^{n-\mu} \in \mathbb{C}[u,\mathrm{v}],\ a_\mu \in \mathbb{C}.$$
Denote by
$$Pol^n \subset \mathbb{C}[u,\mathrm{v}]$$
the subspace of homogeneoups polynomials of degree n. One has
$$dim\ Pol^n = n+1$$
because the family of monomials $(u^{n-i} \cdot \mathrm{v}^i)_{i=0,\dots,n}$ is a base of Pol^n.

When identifying the canonical basis of $\mathbb{C}^2$ with the two variables
$$u = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \mathrm{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
then the tautological representation of L acts on
$$Pol^1 \simeq \mathbb{C} \cdot u \oplus \mathbb{C} \cdot \mathrm{v}$$
by definition as the matrix product
$$z. \begin{pmatrix} a \\ b \end{pmatrix} = z \cdot \begin{pmatrix} a \\ b \end{pmatrix},\ z \in L.$$
Here the dot on the line on the left-hand side denotes the action of the element $z \in L$, while the dot above the line on the right-hand side denotes the product of the matrix $z \in sl(2,\mathbb{C})$ with a vector from $\mathbb{C}^2$. We obtain
$$h.u = u,\ h.\mathrm{v} = -\mathrm{v};\ x.u = 0,\ x.\mathrm{v} = u;\ y.u = \mathrm{v},\ y.\mathrm{v} = 0.$$
More general, for each $z \in L$ we consider the linear differential operator
$$D_z : Pol^n \to Pol^n,\ P \mapsto D_z\, P,$$
defined as

$$(D_z\, P)(u,\mathrm{v}) := (z.u) \cdot \frac{\partial P(u,\mathrm{v})}{\partial u} + (z.\mathrm{v}) \cdot \frac{\partial P(u,\mathrm{v})}{\partial \mathrm{v}}$$

Notably for $u^{n-i} \cdot \mathrm{v}^i \in Pol^n$

$$h.(u^{n-i} \cdot \mathrm{v}^i) := D_h(u^{n-i} \cdot \mathrm{v}^i) = u \cdot (n-i) \cdot u^{n-i-1} \cdot \mathrm{v}^i - i \cdot \mathrm{v} \cdot u^{n-i} \cdot \mathrm{v}^{i-1} = (n-2i) \cdot u^{n-i} \cdot \mathrm{v}^i$$

$$y.(u^{n-i} \cdot \mathrm{v}^i) := D_y(u^{n-i} \cdot \mathrm{v}^i) = v \cdot (n-i) \cdot u^{n-i-1} \cdot \mathrm{v}^i = (n-i) \cdot u^{n-i-1} \cdot \mathrm{v}^{i+1}$$

$$x.(u^{n-i} \cdot \mathrm{v}^i) := D_x(u^{n-i} \cdot \mathrm{v}^i) = u \cdot u^{n-i} \cdot i \cdot \mathrm{v}^{i-1} = i \cdot u^{n-i+1} \cdot \mathrm{v}^{i-1}.$$

As a consequence, the map

$$L \times Pol^n \to Pol^n, (z,P) \mapsto D_z\, P,$$

defines an L-module structure on Pol^n: We set

$$e := u^n \in Pol^n$$

and define for $i = 0,...,n$

$$e_i := \frac{1}{i!} \cdot (y^i.e)$$

One checks by induction on $i = 0,...,n$

$$e_i = \frac{1}{i!} \cdot n \cdot (n-1) \cdot ... \cdot (n-i+1) \cdot u^{n-i} \cdot v^i = \binom{n}{i} \cdot u^{n-i} \cdot v^i.$$

We obtain

$$h.e_i = (n-2i) \cdot e_i$$

$$y.e_i = y.\left(\binom{n}{i} \cdot u^{n-i} \cdot \mathrm{v}^i\right) = \binom{n}{i} \cdot (n-i) \cdot u^{n-i-1} \cdot \mathrm{v}^{i+1} =$$

$$= (i+1) \cdot \binom{n}{i+1} \cdot u^{n-i-1} \cdot \mathrm{v}^{i+1} = (i+1) \cdot e_{i+1}$$

$$x.e_i = x.\left(\binom{n}{i} \cdot u^{n-i} \cdot \mathrm{v}^i\right) = i \cdot \binom{n}{i} \cdot u^{n-i+1} \cdot \mathrm{v}^{i-1} =$$

$$(n-i+1) \cdot \binom{n}{i-1} \cdot u^{n-i+1} \cdot \mathrm{v}^{i-1} = (n-i+1) \cdot e_{i-1}.$$

Due to Theorem 5.10 proof of part 1, these formulas prove $Pol^n \simeq V(n)$ with primitive element

$$e := u^n \in Pol^n.$$

5. One checks that the isomorphy of vector spaces

$$Pol^\lambda = Sym^\lambda(\mathbb{C} \cdot u \oplus \mathbb{C} \cdot \mathrm{v}) \simeq Sym^\lambda\ Pol^1$$

induces an isomorphy of $sl(2,\mathbb{C})$-modules between Pol^{λ} and the symmetric power Pol^{1} with exponent λ of the tautological $sl(2,\mathbb{C})$-module.

The example is the particular case where the symmetric power of an irreducible module stays irreducible. In general, tensor products of irreducible representations are reducible, and to determine the splitting behaviour of tensor products can be a tedious task.

5.3 Root space decomposition and Cartan subalgebra

In this section the Lie algebra L denotes a non-zero, semisimple complex Lie algebra if not stated otherwise. We generalize the splitting

$$sl(2,\mathbb{C}) = L^{0} \oplus \left(L^{\alpha} \oplus L^{-\alpha}\right),\ \alpha = 2,$$

from Proposition 5.4.

According to Proposition 5.2 there exists a maximal toral subalgebra $T \subset L$, and each toral subalgebra of L is Abelian. Hence all endomorphisms

$$ad\ h : L \to L,\ h \in T,$$

are simultaneously diagonizable, and the whole Lie algebra L splits as a direct sum of common eigenspaces of T, see Definition 5.13. The decomposition is named the *root space decomposition* of L with respect to T.

Definition 5.13 is a companion to Definition 5.6: The weights of the adjoint representation are named *roots*.

Definition 5.13 (Root space decomposition with L^0). Consider a pair (L,T) with a semisimple Lie algebra L and a maximal toral subalgebra $T \subset L$.

1. For a complex linear functional

$$\alpha : T \to \mathbb{C}$$

set

$$L^{\alpha} := \{x \in L : [h,x] = \alpha(h) \cdot x \ for\ all\ h \in T\}.$$

If $L^{\alpha} \neq \{0\}$ and $\alpha \neq 0$ then α is a *root*, L^{α} the common eigenspace with respect to α of all endomorphism $ad\ h,\ h \in T$, is the *root space* of α, and each non-zero vector $v \in L^{\alpha}$ is a *root vector* of (L,T). The *set of all roots* of (L,T) is denoted Φ.

2. The vector space decomposition

$$L = L^0 \oplus \left(\bigoplus_{\alpha \in \Phi} L^\alpha \right)$$

is the *root space decomposition* of L *with* L^0 with respect to T.

Because L is finite-dimensional there exist only finitely many roots. The zero eigenspace

$$L^0 := \{x \in L : \ [h,x] = 0 \text{ for all } h \in T\}$$

plays a distinguished role. The next task is to show

$$L^0 = T$$

and to improve Definition 5.13. By definition

$$L^0 = C_L(T) := \{h \in L : \ [h,T] = 0\},$$

the centralizer of T. Proposition 5.2 shows that T is Abelian, hence

$$T \subset C_L(T).$$

Therefore, the main task is to prove the opposite inclusion

$$C_L(T) \subset T,$$

i.e. elements which commute with the maximal toral algebra T already belong to T. This property will be proved in Theorem 5.17. The main steps of the proof are:

- The centralizer $C_L(T)$ contains with each element also its ad-semisimple and its ad-nilpotent summand, see Lemma 5.3, part i). Therefore one can consider both types of elements separately.
- The case of ad-semisimple elements is easy, because T is maximal with respect to ad-semisimple elements, Lemma 5.3, part ii).
- Because the subalgebra $C_L(T)$ is Abelian all its ad-nilpotent elements belong to the null space of the Killing form restricted to $C_L(T)$.
- The Killing form is nondegenerate on $C_L(T)$. Hence the nullspace of $\kappa|C_L(T)$ reduces to $\{0\}$.

We recall from Lemma 4.1 that the Killing form of L is "associative"

$$\kappa([x,y],z) = \kappa(x,[y,z]), x,y,z \in L.$$

Lemma 5.14 and Proposition 5.16 prepare the proof of Theorem 5.17.

Lemma 5.14 (Orthogonality of root spaces). *Consider a pair (L,T) with a semisimple Lie algebra L and a maximal toral subalgebra $T \subset L$. Then for each pair of functionals $\alpha, \beta \in T^*$ holds:*

•
$$[L^\alpha, L^\beta] \subset L^{\alpha+\beta}.$$
In particular, each element element
$$x \in L^\alpha,\ \alpha \neq 0,$$
is ad-nilpotent.

• *If $\alpha + \beta \neq 0$ then*
$$\kappa(L^\alpha, L^\beta) = 0.$$

Proof. i) Assume $x \in L^\alpha,\ y \in L^\beta,\ h \in T$: The Jacobi identity implies

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

hence $[x,y] \in L^{\alpha+\beta}$. Because L has only finitely many roots there exists an exponent $N \in \mathbb{N}$ with

$$(ad\ x)^n = 0 \text{ for all } x \in L^\alpha,\ n \geq N.$$

ii) By assumption there exists an element $h \in T$ with

$$(\alpha+\beta)(h) \neq 0.$$

For arbitrary elements $x \in L^\alpha,\ y \in L^\beta$ holds due to Lemma 4.1:

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

Hence
$$[h,x] = \alpha(h) \cdot x \text{ and } [h,y] = \beta(h) \cdot y$$
imply
$$\alpha(h) \cdot \kappa(x,y) = -\beta(h) \cdot \kappa(x,y),$$
and
$$(\alpha+\beta)(h) \cdot \kappa(x,y) = 0.$$
As a consequence
$$\kappa(x,y) = 0.$$

□

Corollary 5.15 (Negative of a root). *Consider a pair (L,T) with a semisimple Lie algebra L and a maximal toral subalgebra $T \subset L$. For each root $\alpha \in \Phi$ also the negative $-\alpha \in T^*$ is a root, i.e. $-\alpha \in \Phi$.*

Proof. To show that $-\alpha \in \Phi$, assume on the contrary that $-\alpha \in T^*$ is not a root. Then for all roots $\beta \in \Phi$

$$\beta \neq -\alpha \text{ i.e. } \alpha + \beta \neq 0$$

Lemma 5.14, part ii) implies for all $\beta \in \Phi$ and also for $\beta = 0$

$$\kappa(L^\alpha, L^\beta) = 0.$$

Therefore the root space decomposition with L^0

$$L = L^0 \oplus \left(\bigoplus_{\beta \in \Phi} L^\beta \right)$$

implies

$$\kappa(L^\alpha, L) = 0.$$

Theorem 4.14 on the non-degeneratedness of the Killling form implies $L^\alpha = 0$, a contradiction to α being a root. □

The proof of Theorem 5.17 relies on the fact that the Killing form stays non-degenerate when restricted to a maximal toral subalgebra. Proposition 5.16 proves this result in two steps - a posteriori Theorem 5.17 clarifies that both steps coincide.

Proposition 5.16 (Restriction of the Killing form to a maximal toral subalgebra).

Consider a pair (L,T) with L a semisimple Lie algebra with Killing form

$$\kappa : L \times L \to \mathbb{C},$$

and $T \subset L$ a maximal toral subalgebra.

- *The restriction of κ to the centralizer $C_L(T)$ of T*

  $$\kappa|(C_L(T) \times C_L(T)) : C_L(T) \times C_L(T) \to \mathbb{C},$$

  is non-degenerate.

- *The restriction of κ to T*

  $$\kappa|(T \times T) : T \times T \to \mathbb{C}$$

  is non-degenerate.

Proof. Set $C := C_L(T)$, the centralizer of T in L, and recall $C = L^0$.

- *Non-degenerateness of* $\kappa|(C \times C)$: Consider an arbitrary

  $$h \in C = L^0 \text{ with } \kappa(h,C) = 0, \text{ i.e. } \kappa(h,L^0) = 0.$$

  Lemma 5.14 implies for each root $\alpha \in \Phi$

  $$\kappa(h,L^\alpha) = 0.$$

  Hence the root-space decomposition of L with L^0 implies

  $$\kappa(h,L) = 0.$$

  Because κ is non-degenerate according to the Cartan criterion from Theorem 4.14, the last equation implies $h = 0$.

- *Non-degenerateness of* $\kappa|(T \times T)$: Consider an element $h \in T \subset C$ with

  $$\kappa(h,T) = 0.$$

  Due to part i): In order to show $h = 0$, it is sufficient to show

  $$\kappa(h,C) = 0.$$

  For an arbitrary element $x \in C$ consider its abstract Jordan decomposition within the semisimple Lie algebra L

  $$x = s + n$$

  Lemma 5.3, part ii) implies $n \in C$ and $s \in T$, in particular by assumption

  $$\kappa(h,s) = 0.$$

  Due to $[h,C] = 0$ by definition of C we conclude

  $$ad[h,C] = [ad\ h, ad(C)] = 0$$

  Therefore the ad-nilpotency of n implies the nilpotency of the composition

  $$ad\ h \circ ad\ n$$

  Lemma 4.1 concludes

  $$\kappa(h,n) = 0$$

  Hence

  $$\kappa(h,x) = \kappa(h,s) + \kappa(h,n) = 0.$$

  Because $x \in C$ is arbitrary, we obtain

  $$\kappa(h,C) = 0$$

Part i) concludes $h = 0$.
□

Now we are prepared to complete our task by proving Theorem 5.17.

Theorem 5.17 (A maximal toral subalgebra equals its centralizer). *Consider a semisimple complex Lie algebra L and a maximal toral subalgebra $T \subset L$. Then*

$$T = C_L(T).$$

Proof. According to Proposition 5.2 the toral subalgebra T is Abelian. Therefore $T \subset C_L(T)$. It remains to prove the opposite inclusion

$$C_L(T) \subset T.$$

For the proof set $C := C_L(T)$.

Due to Lemma 5.3, part ii) for each $x \in C$ with abstract Jordan decomposition

$$x = s + n$$

the semisimple summand s belongs to T. Hence it remains to show: Any ad-nilpotent element $n \in C$ belongs to T. The proof relies on the non-degenerateness of the restricted Killing κ_T form.

i) *C is nilpotent*: According to Engel's theorem, see Theorem 3.10, it suffices to show that for each element $x \in C$ the endomorphism

$$ad\ x \in End(C)$$

is nilpotent. For the proof consider the abstract Jordan decomposition in L

$$x = s + n.$$

On one hand, Lemma 5.3, part ii) implies $s \in T$. Hence $[s, C] = 0$ by definition. On the other hand, the endomorphism $ad\ n \in End(L)$ is nilpotent, hence a posteriori also the restriction $ad_C\ n \in End(C)$. As a consequence,

$$ad\ x = ad\ n$$

is nilpotent.

ii) $T \cap [C, C] = \{0\}$: By definition of the centralizer $[C, T] = \{0\}$. Lemma 4.1 implies

$$0 = \kappa([T, C], C) = \kappa(T, [C, C]).$$

Due to Proposition 5.16, part ii) the restriction $\kappa|(T \times T)$ is non-degenerate. Hence

$$T \cap [C,C] = \{0\}.$$

iii) *C is Abelian*: We argue by indirect proof. Assume on the contrary

$$[C,C] \neq 0$$

Part i) implies that the Lie algebra C is nilpotent. Corollary 3.13 applies to the ideal

$$\{0\} \neq I := [C,C] \subset C$$

and provides an element

$$0 \neq x \in Z(C) \cap [C,C].$$

The element x is not ad-semisimple, because ad-semisimple elements from C belong to T according to Lemma 5.3, part ii) and

$$T \cap [C,C] = \{0\}$$

according to part ii). Therefore $n \neq 0$ in the abstract Jordan decomposition

$$x = s + n$$

and $n \in C$ according to Lemma 5.3, part i). Moreover

$$x \in Z(C) \implies n \in Z(C)$$

according to Theorem 1.19. The nilpotency of $ad\ n$ and the property $[n,y] = 0$ for all $y \in C$ imply the nilpotency of

$$(ad\ n) \circ (ad\ y).$$

As a consequence $\kappa(n,y) = 0$ according to Lemma 4.1. We obtain

$$\kappa(n,C) = 0$$

Proposition 5.16 implies $n = 0$, a contradiction.

iv) $C \subset T$: We argue by indirect proof. Assume the existence of an element $x \in C \setminus T$, and consider the abstract Jordan decomposition

$$x = s + n.$$

First $n \neq 0$, because otherwise $x = s$ is semisimple and Lemma 5.3, part ii) implies $x \in T$, a contradiction. Secondly, Lemma 5.3, part i) implies $n \in C$. Thirdly, in order to show

$$\kappa(n,C) = 0$$

we consider an arbitrary element $y \in C$. The endomorphism $ad(n)$ is nilpotent. Moreover $[n,y] = 0$ because C is Abelian by part iii). Hence the composition

$$ad(n) \circ ad(y)$$

is nilpotent and

$$\kappa(n,y) = tr(ad(n) \circ ad(y)) = 0.$$

Proposition 5.16 part i) applies and shows $n = 0$, a contradiction.

□

Consider a pair (L,T) with a complex semisimple Lie algebra L and a maximal toral subalgebra $T \in L$. Theorem 5.17 allows to replace in the root space decomposition of L with L^0 from Definition 5.13 the eigenspace $L^0 = C_L(T)$ by T.

Due to Corollary 5.15 the roots of L appear in pairs $(\alpha, -\alpha)$. Chapter 6 and 7 will explain how to choose from each pair one root such that the chosen roots can be considered a set Φ^+ of positive roots.

Definition 5.18 (Root space decomposition or Cartan decomposition). Consider a pair (L,T) with a semisimple Lie algebra L and a maximal toral subalgebra $T \subset L$. Denote by Φ the root set of (L,T). The splitting of L as the direct sum of eigenspaces of T

$$L = T \oplus \left(\bigoplus_{\alpha \in \Phi} L^\alpha \right) = T \oplus \left(\bigoplus_{\alpha \in \Phi^+} \left(L^\alpha \oplus L^{-\alpha} \right) \right)$$

is named the *root space decomposition* or *Cartan decomposition* of L.

Definition 5.19 (Cartan subalgebra). Consider a Lie algebra L. A *Cartan subalgebra H* of L is a nilpotent subalgebra $H \subset L$ equal to its normalizer, i.e.

$$H = N_L(H)$$

Lemma 5.20 (Cartan subalgebras of a semisimple Lie algebra). *For a semisimple Lie algebra L each maximal toral subalgebra $T \subset L$ is a Cartan subalgebra of L.*

Proof. i) *T is nilpotent*: According to Proposition 5.2 any toral subalgebra $T \subset L$ is Abelian, in particular nilpotent.

ii) *T satisfies the normalizer condition*: Consider an arbitrary element $x \in N_L(T)$. We have to show $x \in T$. According to Theorem 5.17 it is sufficient for $x \in T$ to show

$$x \in C_L(T)$$

i.e. to show for all $h \in T$

$$[h,x]=0$$

For the proof apply the root space decomposition of L. It represents x uniquely as

$$x = x_T + \sum_{\alpha\in\Phi} x_\alpha,\ x_T \in T,\ x_\alpha \in L^\alpha.$$

The assumption $x \in N_L(T)$ implies

$$[h,x] = [h,x_T] + \sum_{\alpha\in\Phi} \alpha(h)\cdot x_\alpha = \sum_{\alpha\in\Phi} \alpha(h)\cdot x_\alpha \in T$$

Hence

$$[h,x] \in T \cap \bigoplus_{\alpha\in\Phi} L^\alpha = \{0\}.$$

The opposite inclusion $T \subset N_L(T)$ holds obviously because T is a subalgebra. As a consequence

$$N_L(T) = T.$$

□

Remark 5.21 (Cartan subalgebras of a semisimpleLie algebra).

1. For a semisimple Lie algebra L the two concepts *Cartan subalgebra* and *maximal toral subalgebra* are even equivalent, see [24, Chapter 15.3].

2. A Cartan subalgebra $H \subset L$ of a semisimple Lie algebra L is not uniquely determined. But each two Cartan subalgebras - and a posteriori each two maximal toral subalgebras $T \subset L$ - are conjugate under the group of inner automorphisms of L. By definition an inner automorphism of L is a map

   $$L \to L,\ z \mapsto (exp\ ad_x)(z),\ x \in L \text{ ad-nilpotent.}$$

   For a proof see [24, Chap. 16.4, Corollary] One defines the rank of the semisimple Lie algebra L as

   $$rank\ L := dim_{\mathbb{C}}\ H.$$

Chapter 6
Root systems from an axiomatic point of view

Our point of departure is the root space decomposition of a complex semisimple Lie algebra

$$L = T \oplus \left(\bigoplus_{\alpha \in \Phi^+} (L^\alpha \oplus L^{-\alpha}) \right),$$

see Definition 5.18.

We separate the concept of roots from the concept of a Lie algebra as its origin, and study the properties of Φ in the context of abstract root systems. Here we follow Serre's guide [40]. Different from many other authors Serre introduces a root system in an axiomatic way by its set of reflections. The existence of an invariant scalar product on the vector space generated by the elements of a root system is then a consequence and not a prerequisite. Serre develops the properties of a root system by focusing on the real vector space V spanned by the roots.

The base field in the present chapter is $\mathbb{R}$, all vector spaces are *real* and finite-dimensional.

6.1 Root system

The ambient space of an abstract root system is a real vector space V. One may conceive of V as the real vector space spanned by the root set of a semisimple Lie algebra L, i.e. as the real vector space spanned by certain non-zero, linear functionals on a maximal toral subalgebra of L.

Definition 6.1 (Symmetry). Consider a vector space V. A *symmetry* of V with vector $\alpha \in V, \alpha \neq 0$, is a $\mathbb{R}$-linear automorphism

$$\sigma : V \to V$$

with the following two properties

1. $\sigma(\alpha) = -\alpha$

2. The fixed space
$$H_\sigma := \{x \in V : \sigma(x) = x\}$$
of elements fixed by σ is a hyperplane in V, i.e. $codim_V H_\sigma = 1$.

Lemma 6.2 (Symmetry). *Consider a vector space V and a non-zero element $\alpha \in V$.*

1. *Each symmetry*
$$\sigma : V \to V$$
with vector α induces the splitting
$$V = \mathbb{R} \cdot \alpha \oplus H_\sigma$$
In particular
$$\sigma^2 = id_V.$$
If
$$x = \mu_1(x) \cdot \alpha + v(x) \in V \text{ with } \mu_1 \in V^*,\ v(x) \in H_\sigma,$$
then
$$\sigma(x) = x - 2\mu_1(x) \cdot \alpha$$

2. *Each symmetry σ with vector α induces the linear functional*
$$\alpha^* := 2\mu_1 : V \to \mathbb{R},$$
which satisfies
$$\alpha^*(\alpha) = 2 \text{ and } ker\ \alpha^* = H_\sigma.$$

3. *Conversely, for each non-zero linear functional*
$$\alpha^* : V \to \mathbb{R} \text{ with } \alpha^*(\alpha) = 2,$$
the map
$$\sigma : V \to V,\ x \mapsto x - \alpha^*(x) \cdot \alpha,$$
is a symmetry with vector α, induced functional α^*, and fixed space $H_\sigma = ker\ \alpha^*$.

In the following we define a root system Φ. It has the decisive property that for all $\alpha \in \Phi$ the linear functional $\alpha^* \in V^*$ takes on integer values on all elements $\beta \in \Phi$.

Definition 6.3 (Root system, rank and Cartan integers). A root system R in a vector space V is a pair
$$R = (V, \Phi)$$
with a real vector space V and a subset $\Phi \subset V$ with the following properties:

- (R1) *Finite and spanning*: The set Φ is finite, $0 \notin \Phi$, and $span_{\mathbb{R}}\Phi = V$.

- (R2) *Invariance under distinguished symmetries*: For each $\alpha \in \Phi$ there is a symmetry
$$\sigma : V \to V$$
with vector α, which leaves Φ invariant, i.e. $\sigma(\Phi) \subset \Phi$.

- (R3) *Cartan integers*: For all $\alpha \in \Phi$ each symmetry σ with vector α as in axiom $(R2)$ satisfies for all $\beta \in \Phi$
$$\sigma(\beta) = \beta - < \beta, \alpha > \cdot \alpha$$
with integer values $< \beta, \alpha > \in \mathbb{Z}$. The integers
$$< \beta, \alpha > \in \mathbb{Z}, \alpha, \beta \in \Phi,$$
are named the *Cartan integers* of Φ.

- (R4) *Reducedness*: For each $\alpha \in \Phi$ the only roots proportional to α are α itself and $-\alpha$, i.e.
$$(\mathbb{R} \cdot \alpha) \cap \Phi = \{\alpha, -\alpha\}.$$

The dimension of V is the *rank* of the root system, the elements of Φ are named the *roots* of the root system.

In the literature condition $(R4)$ is considered an additional requirement for a *reduced* root system. If the base field is not algebraically closed one has to distinguish between reduced and non-reduced root systems. These lecture notes consider only reduced root systems. Therefore we omit the attribute *reduced*.

Note that the function
$$< -, - >: V \times V \to \mathbb{R}$$
which defines the Cartan integers $< \beta, \alpha >$ is linear in the first argument β but not necessarily in the second argument α. For any $\alpha \in \Phi$:
$$< \alpha, \alpha >= 2.$$

Lemma 6.4 (Uniqueness of the symmetry). *The symmetry σ with vector α from Definition 6.3 part (R2) is uniquely determined by α.*

Proof. Assume two symmetries $\sigma_i,\ i=1,2$, of V with vector α satisfying

$$\sigma_i(\Phi) \subset \Phi,\ i=1,2.$$

The symmetries satisfy

$$\sigma_i(\Phi) = \Phi,\ i=1,2,$$

because Φ is a finite set, and σ_i is an automorphism. Consider the automorphism of V

$$u := \sigma_2 \circ \sigma_1 : V \to V.$$

It satisfies

$$u(\alpha) = \alpha \text{ and } u(\Phi) = \Phi.$$

Due to the formula from Lemma 6.2, part 1 for $i=1,2$ the symmetries

$$\sigma_i : V \to V,$$

and a posteriori also u, induce on the quotient $V/\mathbb{R}\alpha$ the identity. There exists a linear functional

$$f : V \to \mathbb{R} \text{ with } f(\alpha) = 0$$

such that for all $x \in V$

$$u(x) = x + f(x) \cdot \alpha$$

Iteration shows for arbitrary $n \in \mathbb{N}$ and all $x \in V$

$$u^n(x) = x + n \cdot f(x) \cdot \alpha$$

The restriction $u|\Phi$ is a permutation of Φ. Hence $u|\Phi$ has finite order, i.e. an exponent $n_0 \in \mathbb{N}$ exists with

$$(u|\Phi)^{n_0} = id|\Phi.$$

Hence

$$u^{n_0} = id$$

because Φ spans V. As a consequence

$$n_0 \cdot f = 0$$

which implies $f = 0$ and therefore $u = id_V$. We obtain

$$\sigma_1 \circ \sigma_2 = id_V$$

Because

$$\sigma_i = \sigma_i^{-1}$$

one obtains also

$$\sigma_1 \circ \sigma_2^{-1} = id_V \text{ or } \sigma_1 = \sigma_2$$

□

Notation 6.5 (Root σ_α). *For a root system Φ the unique symmetry σ from Lemma 6.4 with vector α will be denoted σ_α.*

Definition 6.6 (Weyl group). The *Weyl group* $\mathscr{W}$ of a root system (V,Φ) of V is the subgroup of $GL(V)$ generated by all symmetries

$$\sigma_\alpha : V \to V,\ \alpha \in \Phi.$$

The symmetries $\sigma_\alpha,\ \alpha \in \Phi$, are named the *Weyl reflections* of the root system.

The Weyl group $\mathscr{W}$ permutes the elements of the finite set Φ, hence $\mathscr{W}$ is a finite group. As a consequence, one may average an arbitrary scalar product over the elements of the Weyl group, obtaining a scalar product which is invariant with respect to $\mathscr{W}$.

Lemma 6.7 (Invariant scalar product). *Let Φ be a root system of a vector space V. Then a scalar product $(-,-)$ exists on V which is invariant under the Weyl group $\mathscr{W}$ of Φ. With respect to any invariant scalar product the Cartan integers satisfy*

$$<\beta,\alpha>= 2\cdot\frac{(\beta,\alpha)}{(\alpha,\alpha)},\ \alpha,\beta \in \Phi.$$

Proof. Take an arbitrary scalar product B on V and define the bilinear form $(-,-)$ as

$$(x,y) := \sum_{w\in\mathscr{W}} B(w(x),w(y)),\ x,y \in V,$$

as the average over the Weyl group, which is a finite group. Apparently, $(-,-)$ is positive definite, hence a scalar product. By construction, the scalar product is $\mathscr{W}$-invariant. For two roots $\alpha,\beta \in \Phi$ the corresponding Cartan integer $<\beta,\alpha>$ is defined by the equation

$$\sigma_\alpha(\beta) = \beta- <\beta,\alpha> \alpha.$$

Because

$$\sigma_\alpha(\alpha) = -\alpha,\ \text{and } (\sigma_\alpha)^2 = id$$

and due to the invariance of the scalar product:

$$-(\alpha,\beta) = (\sigma_\alpha(\alpha),\beta) = ((\sigma_\alpha)^2(\alpha),\sigma_\alpha(\beta)) =$$

$$= (\alpha,\sigma_\alpha(\beta)) = (\alpha,\beta- <\beta,\alpha>\ \alpha) =$$

$$= (\alpha,\beta)- <\beta,\alpha> (\alpha,\alpha),$$

which proves

$$< \beta, \alpha >= 2 \cdot \frac{(\alpha, \beta)}{(\alpha, \alpha)} = 2 \cdot \frac{(\beta, \alpha)}{(\alpha, \alpha)}$$

□

Corollary 6.8 (Fixed space of a symmetry). *Consider a root system Φ of a vector space V. The fixed space of the symmetry σ_α of a root $\alpha \in \Phi$ is the orthogonal space of the root*

$$H_{\sigma_\alpha} = \alpha^\perp$$

Hence σ_α is the reflection at the orthogonal hyperplane $\alpha^\perp$.

Proof. A given element $\mathrm{v} \in \alpha^\perp$ has a representation

$$\mathrm{v} = \sum_i t_i \cdot \alpha_i, \ \alpha_i \in \Phi.$$

Then

$$\sigma_\alpha(\mathrm{v}) = \sum_i t_i \cdot \sigma_\alpha(\alpha_i) = \sum_i t_i \cdot (\alpha_i - < \alpha_i, \alpha > \ \cdot \alpha) =$$

$$= \sum_i t_i \cdot \alpha_i - \sum_i t_i \cdot 2 \cdot \frac{(\alpha_i, \alpha)}{(\alpha, \alpha)} \cdot \alpha = \mathrm{v} - 2 \cdot \frac{(\mathrm{v}, \alpha)}{(\alpha, \alpha)} \cdot \alpha = \mathrm{v}$$

Hence

$$\alpha^\perp \subset H_{\sigma_\alpha}$$

Both subspaces of V have codimension $= 1$. Therefore

$$H_{\sigma_\alpha} = \alpha^\perp$$

□

Now, after constructing an Euclidean space $(V, (-,-))$ of the root system Φ, we can define the length of a root and the angle between two roots. In the following we always provide a root system Φ with a fixed $\mathscr{W}$-invariant Euclidean structure on its vector space V.

Definition 6.9 (Length of roots and angle between roots). Consider a root system Φ and the corresponding Euclidean vector space $(V, (-,-))$.

1. The *length* of a root $\alpha \in \Phi$ is defined as

$$\|\alpha\| := \sqrt{(\alpha, \alpha)}.$$

2. The *angle* included between two roots $\alpha, \beta \in \Phi$

$$\theta := \sphericalangle(\alpha, \beta) \text{ with } 0 \le \theta \le \pi$$

is defined according to

$$\cos(\theta) := \frac{(\alpha,\beta)}{\|\alpha\| \cdot \|\beta\|}$$

Lemma 6.10 (Possible angles and length ratio of two roots). *Consider a root system Φ and two non proportional roots $\alpha, \beta \in \Phi$. Table 6.1 displays the only possible angles $\sphericalangle(\alpha,\beta)$ included by α and β, and the ratios of the length of α and β if $\|\beta\| \geq \|\alpha\|$.*

No.	$<\alpha,\beta>$	$<\beta,\alpha>$	$\sphericalangle(\alpha,\beta)$	$\|\beta\|^2/\|\alpha\|^2$	$\Delta = \{\alpha,\beta\}$
1	0	0	$\frac{\pi}{2}$	undet.	$A_1 \times A_1$
2	1	1	$\frac{\pi}{3}$	1	
3	-1	-1	$\frac{2\pi}{3}$	1	A_2
4	1	2	$\frac{\pi}{4}$	2	
5	-1	-2	$\frac{3\pi}{4}$	2	B_2
6	1	3	$\frac{\pi}{6}$	3	
7	-1	-3	$\frac{5\pi}{6}$	3	G_2

Table 6.1 Angles and length of roots

The last column of the table indicates the type of the root system of rank $= 2$ with base $\Delta = \{\alpha, \beta\}$. The concept will be introduced in Definition 6.13, see also Theorem 6.31.

Proof. We employ the Cartan integers $<\alpha,\beta>, <\beta,\alpha> \in \mathbb{Z}$.
Both Cartan integers have equal sign due to Lemma 6.7. If α and β are not proportional, then the angle $\theta := \sphericalangle(\alpha,\beta)$ is different from zero and different from π, hence

$$|cos\ \theta| < 1$$

Therefore

$$4 > 4 \cdot \cos^2\ \theta = 2 \cdot \frac{(\alpha,\beta)}{\|\beta\|^2} \cdot 2 \cdot \frac{(\beta,\alpha)}{\|\alpha\|^2} = <\alpha,\beta> \cdot <\beta,\alpha> \geq 0.$$

If $\|\beta\| \geq \|\alpha\|$ then

$$|<\alpha,\beta>| \leq |<\beta,\alpha>|,$$

because

$$<\alpha,\beta> \neq 0 \implies \frac{\|\beta\|^2}{\|\alpha\|^2} = \frac{<\beta,\alpha>}{<\alpha,\beta>} = \frac{|<\beta,\alpha>|}{|<\alpha,\beta>|}$$

Hence only the combinations from the table are possible.

□

The formulas from the proof of Lemma 6.10 indicate : For two non-proportional roots $\alpha, \beta \in \Phi$ the product of their Cartan integers determines the included angle $\theta = \sphericalangle(\alpha, \beta)$, while the quotient of the Cartan integers determines their length ratio $\|\beta\|/\|\alpha\|$ with the only exception of the orthogonal case $\theta = \pi/2$.

Example 6.11 (Root systems of rank ≤ 3).

- Rank $= 1$: The only root system is $\Phi = \{\pm\alpha\}$ with $V = \mathbb{R}$.
- Rank $= 2$: Figure 6.1 displays all root systems of rank $= 2$, see also Theorem 6.27 and Proposition 6.29.
- Rank $= 3$: See the figures in [18, Chap. 8.9] as one example.

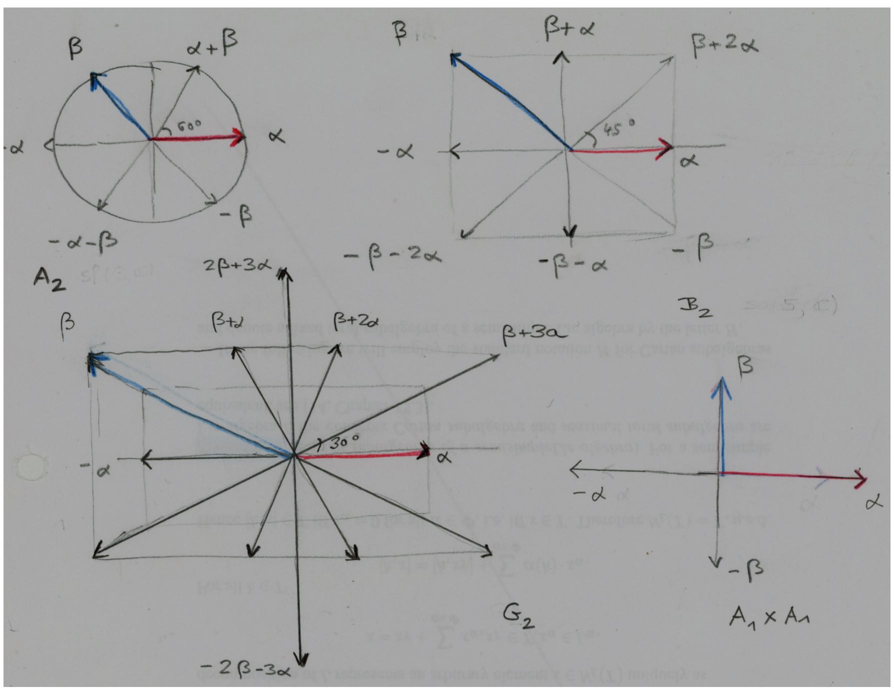

Fig. 6.1 Root systems of rank $= 2$

Lemma 6.12 (Roots with acute angle). *If two non-proportional roots $\alpha, \beta \in \Phi$ include an acute angle, i.e. $(\alpha, \beta) > 0$, then also $\alpha - \beta \in \Phi$.*

Proof. According to Table 6.1 the assumption $(\alpha, \beta) > 0$ implies, depending on the ratio
$$\|\beta\|^2/\|\alpha\|^2,$$
$$<\beta, \alpha>= 1 \text{ or } <\alpha, \beta>= 1.$$
In the first case $\sigma_\alpha(\beta) = \beta - \alpha$, in the second case $\sigma_\beta(\alpha) = \alpha - \beta$. In both cases $\alpha - \beta \in \Phi$ because the Weyl reflections
$$\sigma_\alpha,\ \sigma_\beta \in \mathscr{W}$$
leave Φ invariant and the negative of a root is also a root. □

Definition 6.13 (Base of a root system, positive and negative roots).

Consider a root system Φ in a vector space V.

i) A set
$$\Delta = \{\alpha_1, ..., \alpha_r\}$$
of roots $\alpha_i \in \Phi, i = 1, ..., r$, is a *base* of Φ and the elements of Δ are named *simple* roots iff

- The family $(\alpha_i)_{i=1,...,r}$ is a basis of V,
- and each root $\beta \in \Phi$ has a representation with integer coefficients
$$\beta = \sum_{i=1}^{r} k_i \cdot \alpha_i,\ k_i \in \mathbb{Z},$$
with either all $k_i \geq 0$ or all $k_i \leq 0$.

ii) With respect to a base Δ a root β is

- *positive*, $\beta \succ 0$, iff all $k_i \geq 0$
- *negative*, $\beta \prec 0$, iff all $k_i \leq 0$.

The subset $\Phi^+ \subset \Phi$ is defined as the set of all *positive* roots and the subset $\Phi^- \subset \Phi$ as the set of all *negative* roots.

Theorem 6.14 (Existence of a base). *Every root system Φ in a vector space V has a base Δ.*

Proof. The construction of a candidate for Δ is straightforward. But the proof that the candidate is indeed a base, will take several steps.

i) *Construction of* Δ: Because Φ is finite, a linear functional $t \in V^*$ exists with $t(\alpha) \neq 0$ for all $\alpha \in \Phi$. Set

$$\Phi_t^+ := \{\alpha \in \Phi : t(\alpha) > 0\}$$

and call $\alpha \in \Phi_t^+$ *decomposable* in Φ_t^+ iff

$$\alpha = \alpha_1 + \alpha_2 \textit{ with } \alpha_1, \alpha_2 \in \Phi_t^+$$

and otherwise *indecomposable* in Φ_t^+. We claim that the set

$$\Delta := \{\alpha \subset \Phi_t^+ : \alpha \text{ indecomposable in } \Phi_t^+\}$$

is a base of Φ.

ii) *Representation of elements from* Φ_t^+: We claim that each $\beta \in \Phi_t^+$ has the form

$$\beta = \sum_{\alpha \in \Delta} k_\alpha \cdot \alpha \text{ with all } k_\alpha \in \mathbb{Z}_+.$$

Otherwise consider the non-empty subset $C \subset \Phi_t^+$ of all elements which lack such a representation, and choose an element $\beta \in C$ with $t(\beta) > 0$ minimal. By construction $\beta \in \Phi_t^+$ cannot be indecomposable in Φ_t^+. Hence $\beta \in \Phi_t^+$ is decomposable in Φ_t^+, i.e.

$$\beta = \beta_1 + \beta_2$$

with $\beta_1, \beta_2 \in \Phi_t^+$ and $\beta_1 \in C$ or $\beta_2 \in C$. We get

$$t(\beta) = t(\beta_1) + t(\beta_2)$$

which implies

$$0 < t(\beta_1) < t(\beta) \text{ and } 0 < t(\beta_2) < t(\beta),$$

a contradiction to the minimality of $t(\beta)$ within $t(C)$.

iii) *Angle between elements from* Δ: We claim that two different roots $\alpha \neq \beta$ from Δ are either orthogonal or include an obtuse angle, i.e. $(\alpha, \beta) \leq 0$. Otherwise $(\alpha, \beta) > 0$ and we obtain from Lemma 6.12 the root

$$\gamma := \alpha - \beta \in \Phi.$$

As a consequence

$$\alpha = \beta + \gamma,$$

and $\gamma \notin \Phi_t^+$ because α is indecomposable in Φ_t^+. Hence $-\gamma \in \Phi_t^+$ which implies

$$\beta = \alpha + (-\gamma)$$

decomposable in Φ_t^+, a contradiction which proves the claim.

iv) *Linear independency*: We claim that each finite subset $A \subset V$ with all different elements $\alpha, \beta \in A$ satisfying

$$t(\alpha) > 0 \text{ and } (\alpha, \beta) \le 0$$

is linearly independent. For the proof assume the existence of a representation

$$0 = \sum_{\alpha \in A} n_\alpha \cdot \alpha$$

with coefficients $n_\alpha \in \mathbb{R}$ for all $\alpha \in A$. Separating summands with positive coefficients from those with negative coefficients gives an equation

$$\sum_{\beta \in A_1} k_\beta \cdot \beta = \sum_{\gamma \in A_2} k_\gamma \cdot \gamma =: \mathrm{v} \in V$$

with disjoint subsets $A_1, A_2 \subset A$ and all $k_\beta,\ k_\gamma \ge 0$. Then

$$(\mathrm{v}, \mathrm{v}) = \sum_{\beta \in A_1, \gamma \in A_2} k_\beta \cdot k_\gamma \cdot (\beta, \gamma) \le 0.$$

Hence $\mathrm{v} = 0$. Now

$$0 = t(\mathrm{v}) = \sum_{\beta \in A_1} k_\beta \cdot t(\beta)$$

with $t(\beta) > 0$ for all $\beta \in A_1$ implies $k_\beta = 0$ for all $\beta \in A_1$. Similarly $k_\gamma = 0$ for all $\gamma \in A_2$. Hence

$$n_\alpha = 0$$

for all $\alpha \in A$. Hence the family of elements from A is linearly independent.

The sequence of all steps i) until iv) proves the claim of the theorem.
□

The proof of Theorem 6.14 constructs a base Δ by starting from a certain functional $t \in V^*$. Conversely Lemma 6.15 shows that any base can be obtained in this way: Each base of a root system Φ is the set of indecomposable elements in Φ_t^+ for a suitable linear functional $t \in V^*$.

Lemma 6.15 (Base and a determining linear functional). *In a vector space V consider a root system Φ with a base Δ. Denote by*

$$\Phi = \Phi^+ \dot{\cup} \Phi^-$$

the induced decomposition of Φ. Then

- *A functional $t \in V^*$ exists with $t(\alpha) \ne 0$ for all $\alpha \in \Phi$ such that*

$$\Phi^+ = \Phi_t^+ := \{\alpha \in \Phi : t(\alpha) > 0\} \text{ and } \Phi^- = \Phi_t^- := \{\alpha \in \Phi :\ t(\alpha) < 0\}$$

- *For each functional $t \in V^*$ with $\Phi^+ \subset \Phi_t^+$ holds*

$$\Delta = \{\alpha \in \Phi_t^+ : \alpha \textit{ indecomposable in } \Phi_t^+\}.$$

Proof. The base Δ defines the decomposition

$$\Phi = \Phi^+ \dot{\cup} \Phi^-.$$

First, because the family of elements from the base Δ are a basis of V there exists a linear functional $t \in V^*$ with $t(\alpha) > 0$ for all $\alpha \in \Delta$. Then

$$\Phi^+ = \Phi_t^+ \text{ and } \Phi^- = \Phi_t^-$$

Secondly, for each functional $t \in V^*$ with

$$\Delta \subset \Phi_t^+$$

follows

$$\Phi^+ \subset \Phi_t^+ \text{ and } \Phi^- \subset \Phi_t^-.$$

And the decomposition

$$\Phi^+ \dot{\cup} \Phi^- = \Phi = \Phi_t^+ \dot{\cup} \Phi_t^-$$

implies

$$\Phi^+ = \Phi_t^+ \text{ and } \Phi^- = \Phi_t^-.$$

As a consequence, the indecomposable elements from both sets Φ^+ and Φ_t^+ are equal, i.e. $\Delta = \Delta'$. $\square$

Corollary 6.16 (Two distinct roots of a base are orthogonal or include an obtuse angle). *Consider a base Δ of a root system Φ in a vector space V. Then any two different roots*

$$\alpha \neq \beta \in \Delta$$

are orthogonal or include an obtuse angle, i.e. $(\alpha, \beta) \leq 0$.

Proof. According to Lemma 6.15 a suitable functional $t \in V^*$ exists with

$$\Delta = \{\alpha \in \Phi_t^+ : \ \alpha \text{ indecomposable in } \Phi_t^+\}.$$

Part iii) in the proof of Theorem 6.14 shows $(\alpha, \beta) \leq 0$. $\square$

6.2 Action of the Weyl group

Our aim in the present chapter is the classification of all possible root systems. This result derives from two facts.

First, the Cartan numbers of a root system Φ of a real vector space V are integers. This fact restricts the angle and the relative length of two roots, see Lemma 6.10. The second fact results from investigating different group actions of the Weyl group, in particular the transitive action on the set of bases of a given root system. Besides the action of the Weyl group $\mathscr{W}$ on Φ

$$\mathscr{W} \times \Phi \to \Phi, (w, \alpha) \mapsto w(\alpha),$$

we study the action of $\mathscr{W}$

- on the linear functionals induced by $\alpha \in \Phi$ as

$$< -, \alpha >: V \to \mathbb{R},\ \mathrm{v} \mapsto < \mathrm{v}, \alpha > := 2 \cdot \frac{(\mathrm{v}, \alpha)}{(\alpha, \alpha)}$$

- on the Weyl reflections

$$\sigma_\alpha :\ V \to V,\ \alpha \in \Phi,$$

- and on the bases Δ of Φ.

We will show: A root system is characterized by the matrix of its Cartan integers, the *Cartan matrix*. Only finitely many types of Cartan matrices exist.

Definition 6.17 (Group action). A group G *acts* on a set X if a map exists

$$G \times X \to X,\ (g, x) \mapsto g.x$$

with the following properties:

$$e.x = x \text{ for all } x \in X,\ e \in G \text{ the neutral element,}$$

and

$$(g_1 \cdot g_2).x = g_1.(g_2.x) \text{ for all } g_1, g_2 \in G \text{ and all } x \in X.$$

The group acts *transitive* if for all $x \in X$ the induced map

$$G \to X,\ g \mapsto g.x,$$

is surjective.

Lemma 6.18 (Action of the Weyl group on Cartan integers and on symmetries). *Consider a root system Φ in a vector space V with Weyl group $\mathscr{W}$.*

1. *Set*
$$A := \{< -, \alpha > \in V^* : \alpha \in \Phi\}.$$
The map
$$\mathscr{W} \times A \to A,\ (w, < -, \alpha >) \mapsto < -, w(\alpha) >$$
is a group action.

2. *The conjugation*
$$\mathscr{W} \times \mathscr{W} \to \mathscr{W},\ (w_1, w_2) \mapsto w_1 \circ w_2 \circ w_1^{-1},$$
is a group action.

3. *Denote by*
$$C := \{\Delta \subset \Phi : \Delta \text{ a base of } \Phi\}$$
the set of all bases of Φ*. The map*
$$\mathscr{W} \times C \to C,\ (w, \Delta) \mapsto w(\Delta) := \{w(\alpha) : \alpha \in \Delta\},$$
is a group action.

Remark 6.19 (Action on the dual space). Consider a root system Φ with Weyl group $\mathscr{W}$.

1. For all $\alpha \in \Phi$ and $w \in \mathscr{W}$ holds
$$< -, w(\alpha) > = < -, \alpha > \circ w^{-1}$$
For the proof it is sufficient to consider a Weyl reflection $w = \sigma_\beta,\ \beta \in \Phi$. The proof employs the invariance of the scalar product $(-,-)$ with respect to σ_β, and the property
$$\sigma_\beta^2 = id, \text{ i. e. } \sigma_\beta = \sigma_\beta^{-1}.$$
Then
$$< -, \alpha > \circ \sigma_\beta^{-1} = < -, \alpha > \circ \sigma_\beta = < \sigma_\beta(-), \alpha > = 2 \cdot \frac{(\sigma_\beta(-), \alpha)}{(\alpha, \alpha)} =$$
$$= 2 \cdot \frac{(\sigma_\beta^2(-), \sigma_\beta(\alpha))}{(\alpha, \alpha)} = 2 \cdot \frac{(-, \sigma_\beta(\alpha))}{(\alpha, \alpha)} = 2 \cdot \frac{(-, \sigma_\beta(\alpha))}{(\sigma_\beta(\alpha), \sigma_\beta(\alpha))} =$$
$$= < -, \sigma_\beta(\alpha) >$$

2. Accordingly one extends the action of $\mathscr{W}$ to an action on the dual space V^*:
$$\mathscr{W} \times V^* \to V^*, (w, t) \mapsto w(t) := t \circ w^{-1}.$$

Proposition 6.20 (Properties of the Weyl group). *Consider a root system Φ of a vector space V and denote by $\mathscr{W}$ its Weyl group. Denote by Δ a fixed base of Φ. Then*

1. Embedding Δ into half-spaces*: For each functional $t \in V^*$ an element $w \in \mathscr{W}$ exists such that all $\alpha \in \Delta$ satisfy*
$$w(\alpha) \in \Phi_t^+$$
2. Transitive action on bases*: The action of $\mathscr{W}$ on the set of bases of Φ is transitive, i.e. for any base Δ' of Φ an element $w \in \mathscr{W}$ exists with*
$$w(\Delta) = \Delta'.$$
3. Any root extends to a base*: For any root $\alpha \in \Phi$ an element $w \in \mathscr{W}$ exists with*
$$w(\alpha) \in \Delta, \text{ i.e. } \Phi = \mathscr{W}(\Delta).$$
4. Generators of $\mathscr{W}$*: The Weyl group $\mathscr{W}$ is generated by the Weyl reflections σ_α of the roots $\alpha \in \Delta$.*

Proof. Denote by $\mathscr{W}_\Delta \subset \mathscr{W}$ the subgroup generated by the Weyl reflections σ_α of the roots $\alpha \in \Delta$. We first show that the first three claims of the Proposition can be satisfied with Weyl reflections from $\mathscr{W}_\Delta$. The final step of the proof will show
$$\mathscr{W} = \mathscr{W}_\Delta.$$

i) Each Weyl reflection $\sigma_\alpha \in \mathscr{W}$ with $\alpha \in \Delta$ leaves the set $\Phi^+ \setminus \{\alpha\}$ invariant: Assume that Δ comprises at least two roots. Consider an element $\beta \in \Phi^+ \setminus \{\alpha\}$. It has a representation
$$\beta = \sum_{\gamma \in \Delta} k_\gamma \cdot \gamma,\ k_\gamma \geq 0, \text{ for all } \gamma \in \Delta.$$

Due to axiom (R4) from Definition 6.3 the root β is not proportional to α, because $-\alpha \notin \Phi^+$. Hence $k_\gamma > 0$ for at least one $\gamma \in \Delta \setminus \{\alpha\}$. We get
$$\sigma_\alpha(\beta) = \beta - <\beta, \alpha> \alpha =$$
$$= \left(\sum_{\gamma \in \Delta} k_\gamma \cdot \gamma \right) - <\beta, \alpha> \alpha = (k_\alpha - <\beta, \alpha>) \cdot \alpha + \sum_{\gamma \in \Delta \setminus \{\alpha\}} k_\gamma \cdot \gamma.$$

Hence also $\sigma_\alpha(\beta)$ has at least one coefficient $k_\gamma > 0$. Because Δ is a base, all coefficients are non-negative and
$$\sigma_\alpha(\beta) \in \Phi^+ \setminus \{\alpha\}.$$

ii) We introduce the distinguished vector $\rho \in V$, which is defined as half the sum of all positive roots

$$\rho := (1/2) \cdot \sum_{\beta \in \Phi^+} \beta.$$

According to part i) each Weyl reflection $\sigma_\alpha \in \mathscr{W}$ with $\alpha \in \Delta$ permutes all positive roots different from α and $\sigma_\alpha(\alpha) = -\alpha$. Hence

$$\sigma_\alpha(\rho) = \sigma_\alpha((\rho - \alpha/2) + \alpha/2) = (\rho - \alpha/2) - \alpha/2 = \rho - \alpha.$$

We now show that the first three claims can be satisfied already by taking Weyl reflections from $\mathscr{W}_\Delta$.

iii) Claim: Part 1 of the Proposition can be achieved with an element $w \in \mathscr{W}_\Delta$: For a given functional $t \in V^*$ we choose an element $w \in \mathscr{W}_\Delta$ with $t(w(\rho))$ maximal with respect to all elements from $\mathscr{W}_\Delta$. For all $\alpha \in \Delta$ holds according to part ii):

$$\alpha = \rho - \sigma_\alpha(\rho)$$

$$w(\alpha) = w(\rho) - w(\sigma_\alpha(\rho))$$

$$t(w(\alpha)) = t(w(\rho)) - t(w(\sigma_\alpha(\rho))).$$

Because also $w \circ \sigma_\alpha \in \mathscr{W}_\Delta$ we have due to the choice of $w \in \mathscr{W}_\Delta$

$$t(w(\rho)) \geq t(w(\sigma_\alpha(\rho)))$$

or for all $\alpha \in \Delta$

$$t(w(\alpha)) \geq 0, \text{ i.e. } w(\alpha) \in \Phi_t^+$$

iv) Claim: Part 2 of the proposition can be achieved with an element $w \in \mathscr{W}_\Delta$: First, Lemma 6.15 applies to Δ' and provides a functional $t' \in V^*$

- satisfying

$$t'(\alpha) \neq 0$$

  for all $\alpha \in \Phi$

- and

$$t'(\alpha') > 0$$

  for all $\alpha' \in \Delta'$.

Secondly, the already proved Part 1 of the proposition applies to t' and provides an element $w \in \mathscr{W}_\Delta$ such that all $\alpha \in \Delta$ satisfy

$$t'(w(\alpha)) \geq 0$$

Then the functional

$$t := t' \circ w \in V^*$$

satisfies for all $\alpha \in \Delta$

$$t(\alpha) \geq 0.$$

Here the case $t(\alpha) = 0$ is excluded, because otherwise the root $w(\alpha) \in \Phi$ would satisfy

$$t'(w(\alpha)) = 0,$$

which is excluded because t' does not vanish on any root from Φ.

Applying now Lemma 6.15 to the pairs (Δ, t) and (Δ', t') characterizes Δ as the set of indecomposable elements of Φ_t^+ and Δ' as the set of indecomposable elements of $\Phi_{t'}^+$. Therefore

$$w(\Delta) = \{w(\alpha) \in \Phi : \alpha \in \Phi_t^+ \text{ indecomposable in } \Phi_t^+\} =$$

$$= \{w(\alpha) \in \Phi : \ t(\alpha) = t'(w(\alpha)) > 0, \ w(\alpha) \text{ indecomposable in } \Phi_{t'}^+\} =$$

$$= \{\beta \in \Phi_{t'}^+ : \ \beta \text{ indecomposable in } \Phi_{t'}^+\} = \Delta'$$

v) Claim: Part 3 of the Proposition can be achieved with an element $w \in \mathscr{W}_\Delta$: For fixed $\alpha \in \Phi$ we find a functional $t_0 \in V^*$ with

$$t_0(\alpha) = 0 \text{ but } t_0(\beta) \neq 0$$

for all roots β not proportional to α. Because Φ is a finite set, the minimum

$$\min\{|t_0(\beta)| : \ \beta \in \Phi \text{ not proportional to } \alpha\}$$

is positive. Hence a small perturbation of t_0 provides a linear functional $t \in V^*$ and an $\varepsilon > 0$ such that for all roots β not proportional to α

$$|t(\beta)| > \varepsilon$$

but

$$t(\alpha) = \varepsilon.$$

Denote by Δ_t the base of Φ induced by t according to the proof of Theorem 6.14. By part iv) an element $w \in \mathscr{W}_\Delta$ exists with

$$w(\Delta_t) = \Delta.$$

The root $\alpha \in \Phi$ is contained in Φ_t^+. It is indecomposable because $t(\beta) > \varepsilon$ for all roots $\beta \in \Phi_t^+$ which are non-proportional to α. Therefore $\alpha \in \Delta_t$, which implies $w(\alpha) \in \Delta$.

vi) Claim $\mathscr{W}_\Delta = \mathscr{W}$: Because the Weyl group is generated by all symmetries $\sigma_\alpha, \ \alpha \in \Phi$, it suffices to show that

$$\sigma_\alpha \in \mathscr{W}_\Delta$$

for all $\alpha \in \Phi$. We choose an arbitrary root $\alpha \in \Phi$. According to part v) an element $w \in \mathscr{W}_\Delta$ exists with

$$\beta := w(\alpha) \in \Delta.$$

First, we show

$$w^{-1} \circ \sigma_{w(\alpha)} = \sigma_\alpha \circ w^{-1} :$$

We apply both sides to a root $\gamma \in \Phi$. Left-hand side:

$$(w^{-1} \circ \sigma_{w(\alpha)})(\gamma) = w^{-1}(\gamma - < \gamma, w(\alpha) > \cdot w(\alpha)) =$$

$$= w^{-1}(\gamma) - < \gamma, w(\alpha) > \cdot w^{-1}(w(\alpha)) =$$
$$= w^{-1}(\gamma) - < \gamma, w(\alpha) > \cdot \alpha$$

Right-hand side:

$$(\sigma_\alpha \circ w^{-1})(\gamma) = \sigma_\alpha(w^{-1}(\gamma)) = w^{-1}(\gamma) - < w^{-1}(\gamma), \alpha > \cdot \alpha$$

It remains to show:

$$< \gamma, w(\alpha) > = < w^{-1}(\gamma), \alpha >,$$

i.e the equality of the two linear functionals

$$< -, w(\alpha) > = < w^{-1}(-), \alpha > .$$

According to Remark 6.19 for the left-hand side holds

$$< -, w(\alpha) > = < -, \alpha > \circ w^{-1} = < w^{-1}(-), \alpha >,$$

which proves the claim.

As a consequence

$$w^{-1} \circ \sigma_\beta = w^{-1} \circ \sigma_{w(\alpha)} = \sigma_\alpha \circ w^{-1}$$

i.e

$$\sigma_\alpha = w^{-1} \circ \sigma_\beta \circ w \in \mathscr{W}_\Delta.$$

□

Definition 6.21 (Cartan matrix).

Consider a root system Φ in a vector space V and a base $\Delta = \{\alpha_1, ..., \alpha_r\}$ of Φ. The *Cartan matrix* of Δ is the matrix of the Cartan integers of the roots from Δ

$$Cartan(\Delta) := \left(< \alpha_i, \alpha_j >_{1 \le i,j \le r}\right) \in M(r \times r, \mathbb{Z}).$$

Here the index i denotes the row and the index j denotes the column.

Note that the Cartan matrix is not necessarily symmetric. All diagonal elements of the Cartan matrix have the value

$$< \alpha_i, \alpha_i >= 2,$$

which follows from

$$-\alpha = \sigma_\alpha(\alpha) = \alpha - < \alpha, \alpha > \cdot \alpha$$

for each root $\alpha \in \Phi$.

For $i \neq j$ only values

$$< \alpha_i, \alpha_j > \leq 0$$

are possible according to Corollary 6.16. Moreover, these values are restricted to the set

$$\{0, -1, -2, -3\}$$

according to Lemma 6.10 and Corollary 6.16.

The Cartan matrix is defined with reference to a base Δ and with reference to a numbering of its elements. Lemma 6.22 shows: Each two bases of Φ have the same Cartan matrices.

Lemma 6.22 (Independence of the Cartan matrix from the choosen base).

Consider a root system Φ of rank

$$r = rank\ \Phi.$$

Any two bases Δ, Δ' of Φ have the same Cartan matrix up to a renumbering of the elements of the bases. More specifically:

An element $w \in \mathscr{W}$ of the Weyl group of Φ exists with $w(\Delta) = \Delta'$ and

$$Cartan(\Delta) = Cartan(\Delta').$$

Proof. According to Proposition 6.20 an element $w \in \mathscr{W}$ exists with $w(\Delta) = \Delta'$ after renumbering, i.e.

$$\Delta = \{\alpha_1, ..., \alpha_r\} \implies \Delta' = \{w(\alpha_1), ..., w(\alpha_r)\}.$$

According to Remark 6.19

$$< -, w(\alpha_j) >=< -, \alpha_j > \circ w^{-1}$$

which implies

$$< w(\alpha_i), w(\alpha_j) >=< w^{-1}(w(\alpha_i)), \alpha_j >=< \alpha_i, \alpha_j >$$

□

As a consequence of Lemma 6.22 one speaks of the Cartan matrix of a *root system*, independently from the choice of a base.

The Cartan matrix encodes the full information of the root system Φ, notably the dimension of its ambient space V. Theorem 6.23 shows that each bijective map between two bases of root systems with the same Cartan matrix extends to an isomorphism of the ambient vector spaces of the root systems.

Theorem 6.23 (The Cartan matrix characterizes the root system).

Consider a root system Φ in a vector space V and a base

$$\Delta = \{\alpha_1, ..., \alpha_r\}$$

of Φ. Let V' be a second vector space and

$$\Delta' = \{\alpha'_1, ..., \alpha'_r\}$$

a base of a root system Φ' in V'. If a bijective map

$$f : \Delta \to \Delta'$$

exists with

$$Cartan(\Delta) = Cartan(\Delta'),$$

then a unique isomorphism of vector spaces

$$F : V \to V'$$

exists with

$$F(\Phi) = \Phi' \text{ and } F|\Delta = f.$$

Proof. i) *Construction of F*: Because the elements from Δ form a basis of the vector space V we may define

$$F : V \to V'$$

as the uniquely determined linear extension of f. And because Δ and Δ' have the same cardinality the linear map F is an isomorphism.

ii) *Conjugation of the Wely groups*: We show that the Weyl groups $\mathscr{W}$ and $\mathscr{W}'$ are conjugate via F, i.e.

$$\mathscr{W}' = F \circ \mathscr{W} \circ F^{-1} \text{ i.e. } \mathscr{W}' \circ F = F \circ \mathscr{W} :$$

For the proof it is sufficient two consider two roots $\alpha, \beta \in \Delta$. Then

$$(\sigma_{f(\alpha)} \circ F)(\beta) = \sigma_{f(\alpha)}(f(\beta)) = f(\beta) - < f(\beta), f(\alpha) > f(\alpha)$$

and

$$(F \circ \sigma_\alpha)(\beta) = F(\sigma_\alpha(\beta)) = F(\beta - <\beta, \alpha, > \alpha) = f(\beta) - <\beta, \alpha, > f(\alpha)$$

Hence for every root $\alpha \in \Delta$

$$\sigma_{f(\alpha)} \circ F = F \circ \sigma_\alpha.$$

Moreover, the conjugation is compatible with taking the product of Weyl reflections. The Weyl groups $\mathscr{W}$ and $\mathscr{W}'$ are generated by the Weyl reflections of the elements from respectively Δ and Δ', see Proposition 6.20. Hence

$$\mathscr{W}' \circ F = F \circ \mathscr{W}.$$

iii) *Mapping* Φ: According to part ii) in combination with Proposition 6.20, part 3

$$\Phi' = \mathscr{W}(\Delta') = \mathscr{W}'(F(\Delta)) = F(\mathscr{W}(\Delta)) = F(\Phi)$$

□

6.3 Coxeter graph and Dynkin diagram

Due to Theorem 6.23 a root system Φ in a vector space is completely determined by its Cartan matrix. Hence the classification of root systems reduces to the classification of possible Cartan matrices. The Cartan integers satisfy a set of restrictions. The present section shows how the language of Cartan matrices translates to a data structure from Discrete Mathematics. The data structure is an undirected graph with multiple edges, called the *Coxeter graph* of the root system.

Therefore, we first define the Coxeter graph of Φ, and then classify all possible graphs from a class of graphs, which covers all Coxeter graphs. The classification is achieved by calculating within Euclidean vector spaces. Theorem 6.27 gives the final classification. A minor shortcoming of the Coxeter graph of a root system is the fact that it contains no information about the relative length of the roots. Therefore one upgrades the Coxeter graph by orientating the edges by a pointer from a long root to a short root. The result is the Dynkin diagram of Φ.

Definition 6.24 (Coxeter graph of a root system). Consider a root system Φ of a vector space V and a base Δ of Φ. The *Coxeter graph* of Φ is the undirected graph

$$Coxeter(\Phi) = (N, E)$$

with

- *vertex* set $N := \Delta$

- and *edge* set E: Each pair of distinct roots

$$\alpha, \beta \in \Delta,\ \alpha \neq \beta,$$

is joined by exactly

$$< \alpha, \beta > \cdot < \beta, \alpha >$$

undirected edges.

Note: The Coxeter graph of Φ does not depend on the choice of the base Δ of Φ, see Theorem 6.23.
Recall from Lemma 6.10 and Corollary 6.16: The angle θ between two roots $\alpha \neq \beta \in \Delta$ is determined by the product of their Cartan integers as

$$< \alpha, \beta >< \beta, \alpha >= 4 \cdot cos^2\ \theta \in \{0, 1, 2, 3\},\ \pi/2 \leq \theta < \pi.$$

If $(-,-)$ denotes a scalar product on V which is invariant under Φ, then

$$< \alpha, \beta >< \beta, \alpha >= 4 \cdot \frac{(\alpha, \beta)^2}{\|\alpha\|^2 \cdot \|\beta\|^2}.$$

In order to classify all Coxeter graphs, we introduce the concept of an *admissible* graph. Each Coxeter graphs defines an admissible graph. The next step, Theorem 6.27, classifies all connected admissible graphs. A second step must show that all connected admissible graphs are Coxeter graphs. We will show a partial result for the second step in Chapter 7 and give a reference for the remaining part.

Definition 6.25 (Admissible graph). Consider an Euclidean space $(V, (-,-))$ and an undirected graph (N, E) with a finite vertex set $N \subset V$ and with edge set E. Assume that the vertex set

$$N = \{\mathrm{v}_1, ..., \mathrm{v}_r\} \subset V$$

is a linearly independent family $(\mathrm{v}_i)_{i=1,...,r}$ of unit vectors satisfying for $1 \leq i \neq j \leq r$

$$(\mathrm{v}_i, \mathrm{v}_j) \leq 0,$$

and that each pair of distinct vertices $\mathrm{v}_i\ \mathrm{v}_j \in V$ is joined by

$$4 \cdot (\mathrm{v}_i, \mathrm{v}_j)^2 \in \mathbb{N}$$

edges from E. Then the graph (N, E) is named *admissible* if each pair of distinct nodes is joined by at most 3 edges, otherwise the graph is named *non-admissible*.

Lemma 6.26 (Coxeter graph and admissible graph). *Each Coxeter graph of a root system Φ defines an admissible graph after normalizing the lenght of the roots of a base Δ of Φ.*

Proof. After choosing a scalar product $(-,-)$ on V, which is invariant under the Weyl group of Φ, one defines for each $\alpha \in \Delta$ the unit vector

$$\alpha_u := \frac{\alpha}{\|\alpha\|} \in (V, (-,-)).$$

Then the following graph (N,E) is admissible: Vertex set

$$N := \{\alpha_u : \ \alpha \in \Delta\}$$

and edge set E: Each pair of distinct vertices $\alpha_u,\ \beta_u \in N$ is joined by exactly

$$< \alpha, \beta >< \beta, \alpha >$$

edges. The claim

$$4(\alpha_u, \beta_u)^2 \in \{0,1,2,3\}$$

follows immediately from the formula

$$4(\alpha_u, \beta_u)^2 = 4 \cdot \frac{(\alpha, \beta)^2}{\|\alpha\|^2 \cdot \|\beta\|^2} =< \alpha, \beta >< \beta, \alpha > \in \{0,1,2,3\}$$

as noted above. □

Theorem 6.27 (Classification of connected admissible graphs). *Each connected admissible graph belongs to exactly one of the classes from Figure 6.2 - up to a numbering of the vertices:*

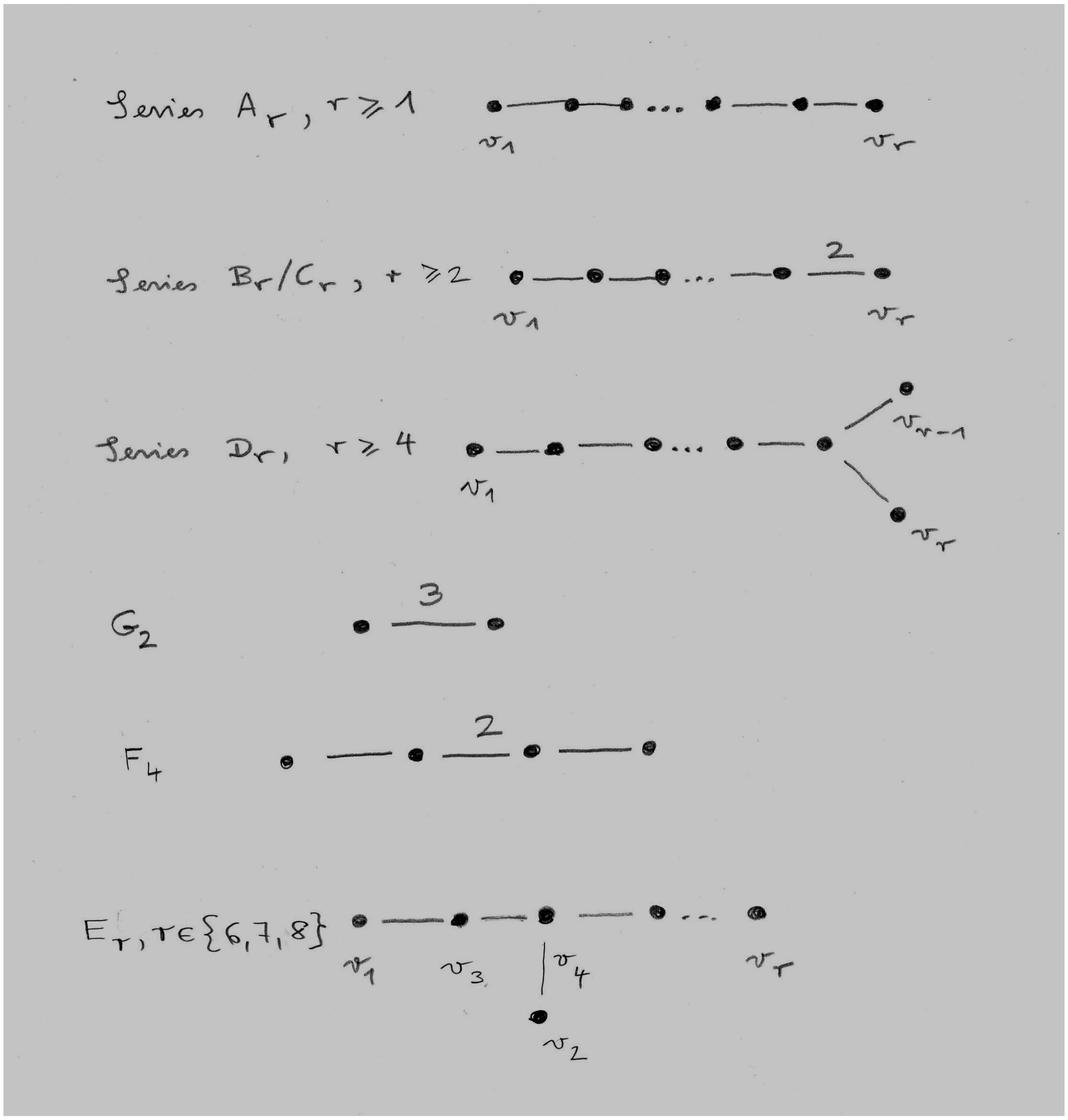

Fig. 6.2 Connected admissible graphs

Figure 6.2 means:

- *The integer at a link between two vertices $v_i \neq v_j$ is the number of edges between the two vertices; a link without a number indicates a single edge.*
- *Graphs from series $A_r, r \geq 1$: Each pair of subsequent roots include the angle $\frac{2\pi}{3}$.*

- *Graphs from series $B_r/C_r, r \geq 2$: The first $r-2$ pairs of subsequent roots include the angle $\frac{2\pi}{3}$, the last two roots include the angle $\frac{3\pi}{4}$.*

- *Graphs from series $D_r, r \geq 4$: The first $r-3$ subsequent pairs of roots include the angle $\frac{2\pi}{3}$, root α_{r-2} includes with each of the two roots α_{r-1} and α_r the angle $\frac{2\pi}{3}$.*

- *Exceptional graph G_2: $\Delta = (\alpha_1, \alpha_2)$. The two roots include the angle $\frac{5\pi}{6}$.*

- *Exceptional graph F_4: $\Delta = (\alpha_1, ..., \alpha_4)$. The pairs (α_1, α_2) and (α_3, α_4) include the angle $\frac{2\pi}{3}$, the pair (α_2, α_3) includes the angle $\frac{3\pi}{4}$.*

- *Exceptional graphs $E_r, r \in \{6, 7, 8\} : \Delta = (\alpha_1, ..., \alpha_r)$. All subsequent pairs of the chain include the angle $\frac{2\pi}{3}$. Also the distinguished root α_2 and the root α_4 include the angle $\frac{2\pi}{3}$.*

Proof. The proof is taken from [24, Chap. 11.4].

1. *Removing vertices and incident edges*: For an admissible graph each subgraph, which is obtained by removing a subset of vertices and their incident edges, is admissible.

2. *Number of vertices and edges*: An admissible graph has less edges, counted *without* multiplicity, than vertices, i.e. $|E| < |N|$:

   Assume $N = \{\mathrm{v}_1, ..., \mathrm{v}_r\} \subset V$. Consider the element

$$\mathrm{v} := \sum_{i=1}^{r} \mathrm{v}_i \in V.$$

   Then $\mathrm{v} \neq 0$ because the family $(\mathrm{v}_i)_{i=1,...,r}$ is linearly independent. We obtain

$$0 < (\mathrm{v}, \mathrm{v}) = \sum_{i=1}^{r} (\mathrm{v}_i, \mathrm{v}_i) + 2 \cdot \sum_{1 \leq i < j \leq r} (\mathrm{v}_i, \mathrm{v}_j).$$

   If $(\mathrm{v}_i, \mathrm{v}_j) \neq 0$ then

$$4 \cdot (\mathrm{v}_i, \mathrm{v}_j)^2 \in \{1, 2, 3\} \text{ and } (\mathrm{v}_i, \mathrm{v}_j) < 0,$$

   which implies

$$2 \cdot (\mathrm{v}_i, \mathrm{v}_j) \leq -1.$$

   Therefore

$$0 < r + 2 \cdot \sum_{1 \leq i < j \leq r} (\mathrm{v}_i, \mathrm{v}_j) \leq r + 2 \cdot |E| \cdot (-1/2) = r - |E|$$

and

$$|E| < |N|.$$

3. *Cycle-free*: A cycle $C = (N_C, E_C)$ were an admissible graph according to part 1, but would violate part 2 because

$$|E_C| = |N_C|.$$

4. *Bounded fan*: For each vertex of an admissible graph the number of incident edges, counted *with* multiplicity, is at most $= 3$: Denote by

$$Inc(\mathrm{v}) := \{e \in E : e \text{ incident with v}\}$$

the set of edges incident to a vertex v. We have to show

$$|Inc(\mathrm{v}| \leq 3.$$

Denote by $\{w_1, ..., w_k\}$ the set of vertices adjacent to v, see Figure 6.3.

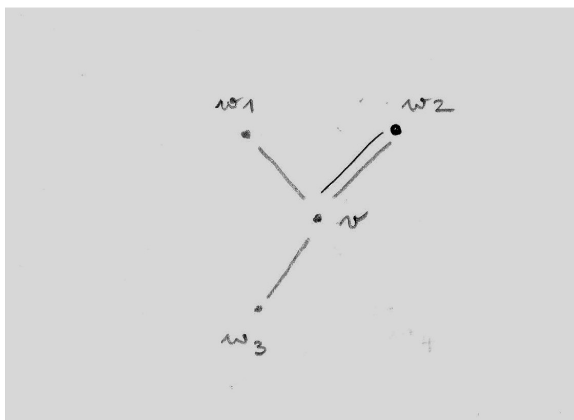

Fig. 6.3 Vertex v with incident edges

Because an admissible graph is cycle free due to part 3, two different vertices

$$w_i \neq w_j,\ 1 \leq i \neq j \leq k$$

are not adjacent, and therefore

$$(w_i, w_j) = 0$$

i.e. the family

$$\mathscr{B} := (w_i)_{i=1,...,k}$$

is an orthogonal family in $(V, (-,-))$. Because the family

$$(v, w_1, ..., w_k)$$

is linearly independent as a subfamily of the linearly independent family of all vertices in N, the family $\mathscr{B}$ extends to an orthomormal family

$$\tilde{\mathscr{B}} = (w_i)_{i=0,...,k}$$

by adding a unit-vector

$$w_0 \in \ span < v, w_1, ..., w_k >$$

It satisfies

$$(w_0, v) \neq 0$$

because otherwise $w_0 = 0$. The orthogonal decomposition with respect to $\tilde{\mathscr{B}}$

$$v = \sum_{i=0}^{k} (v, w_i) \cdot w_i.$$

implies

$$1 = (v, v) = \sum_{i=0}^{k} (v, w_i)^2$$

and therefore

$$\sum_{i=1}^{k} (v, w_i)^2 < 1.$$

As a consequence

$$\sum_{i=1}^{k} 4 \cdot (v, w_i)^2 < 4$$

which excludes more than 3 edges incident with v.

5. *Triple edge*: The only connected admissible graph with a triple edge is the graph of type G_2 from Figure 6.2. Apparently the graph is admissible. The fact that type G_2 is the only connected admissible graph with a triple edge follows from part 4.

6. *Blowing down simple paths* : Blowing down a simple path, i.e. a path without multiple edges, results in a new admissible graph (N', E'). Figure 6.4 shows a graph (N, E) which cannot be admissible because its blow-down (N', E') is non-admissible:

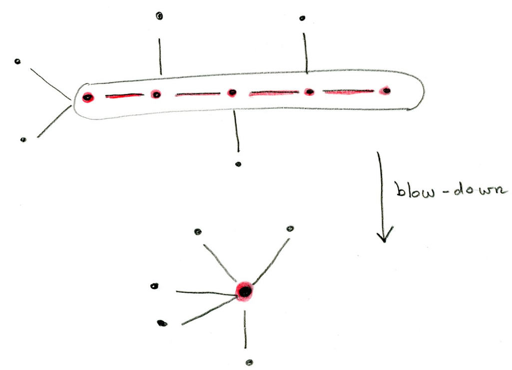

Fig. 6.4 Blowing down a simple path with non-admissible resulting graph

Denote by $C = \{w_1, ..., w_n\} \subset N$ the vertices of the path. By assumption for $i = 1, ..., n-1$

$$4 \cdot (w_i, w_{i+1})^2 = 1, \ i.e. \ 2 \cdot (w_i, w_{i+1}) = -1.$$

The graph (N', E'), resulting from blowing down the original path, has

- vertex set
$$N' = (N \setminus C) \cup \{w_0\}, \ w_0 := \sum_{i=1}^{n} w_i \in V,$$
- and edge set E' obtained from E by removing all edges of the path C and replacing each edge of E, which is incident with a vertex of C, by a corresponding edge incident with w_0.

We show that the graph (N', E') is admissible: Linear independence of the vertex set N' is obvious. We compute

$$(w_0, w_0) = \sum_{1 \le i,j \le n}^{n} (w_i, w_j) = \sum_{i=1}^{n} (w_i, w_i) + 2 \cdot \sum_{1 \le i < j \le n} (w_i, w_j) =$$

$$n + 2 \cdot \sum_{i=1}^{n-1} (w_i, w_{i+1}) = n + 2 \cdot (n-1) \cdot (-1/2) = n - (n-1) = 1,$$

which shows that also the vector w_0 is a unit vector. In (N, E) any vertex w from $N \setminus C$ is adjacent to at most one vertex from the path, because an admissible graph is cycle free according to part 3. Hence

- either $(w, w_0) = 0$
- or exactly one index $i = 1, ..., n$ exists with $0 \neq (w, w_i)$.

In either case holds
$$4 \cdot (w, w_0)^2 \in \{0, 1, 2, 3\}.$$

7. *Prohibited subgraphs*: A connected admissible graph does not contain any subgraph from Figure 6.5:

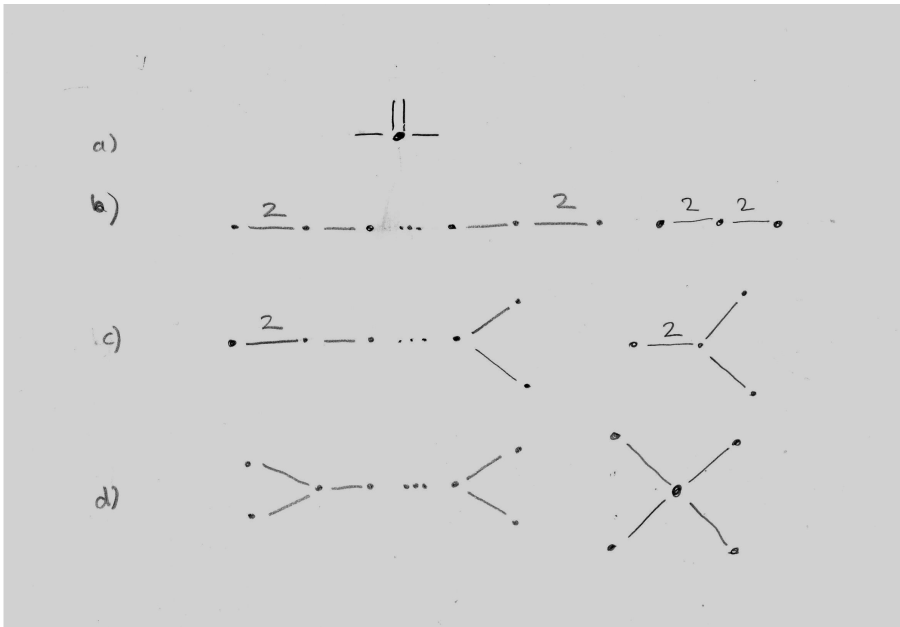

Fig. 6.5 Types of prohibited subgraphs and their non-admissible blow-down in cases b)-d)

Here a number 2 on a line connecting two vertices means that the two vertices are joined by 2 edges. The prohibited subgraphs from Figure 6.5 contain at least one of the following constellations of nodes respectively edges:

a) A vertex with more than 3 incident edges is prohibited according to part 4.

b) Two pairs of adjacent vertices connected by a multiple edge.

c) One pair of adjacent vertices connected by a multiple edge, and another vertex with 3 incident edges.

d) Two distinct vertices, both have at least 3 incident edges.

In any of the subgraphs b)-d) it would be possible to blow down a path to a vertex with at least 4 incident edges, which contradicts part 4 and part 6.

8. *The types of admissible connected graphs*: Each connected admissible graph belongs to one of the types from Figure 6.6:

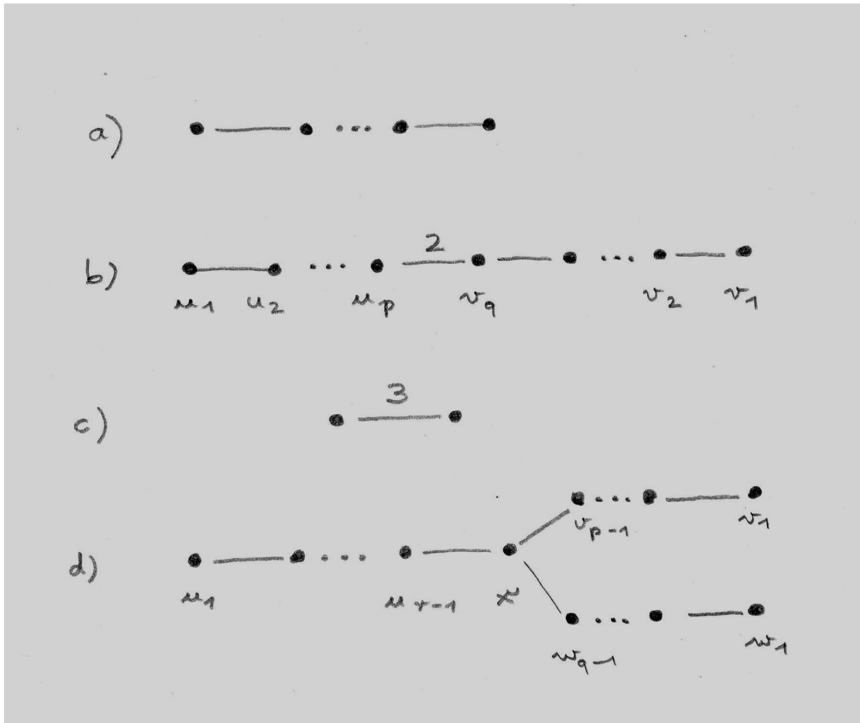

Fig. 6.6 Admissible connected graphs

a) No multiple edges, no vertex with 3 incident edges: All connected admissible graphs from Figure 6.6 type a) belong to type $A_r,\ r \geq 1$, from Figure 6.2.

b) A single pair of vertices with a double edge, no vertex with 3 incident edges: See part 9.

c) A single pair of vertices with a triple edge, no further vertices. The only connected admissible graph from Figure 6.6 type c) is the exceptional graph G_2 from Figure 6.2. For the proof see part 5.

d) No multiple edges, a single vertex with 3 incident single edges: See part 10.

These graphs are admissible. The fact, that there are no other types of admissible graphs, result from excluding the prohibited subgraphs from part 7.

9. *Admissible graphs from series B_r, C_r and exceptional graph F_4*: All connected admissible graphs from Figure 6.6 type b) belong to series B_r or are the exceptional graph F_4 in Figure 6.2.

Consider the two vectors from V

$$u := \sum_{i=1}^{p} i \cdot u_i \text{ and } \mathrm{v} := \sum_{i=1}^{q} i \cdot \mathrm{v}_i.$$

They are linearly independent. Using

$$2 \cdot (u_i, u_{i+1}) = -1$$

we compute

$$(u,u) = \sum_{1 \le i,j \le p} i \cdot j \cdot (u_i, u_j) = \sum_{i=1}^{p} i^2 \cdot (u_i, u_i) + 2 \cdot \sum_{i=1}^{p-1} i(i+1) \cdot (u_i, u_{i+1}) =$$

$$= \sum_{i=1}^{p} i^2 + 2 \cdot \sum_{i=1}^{p-1} (-1/2) \cdot i(i+1) = \sum_{i=1}^{p} i^2 - \sum_{i=1}^{p-1} i^2 - \sum_{i=1}^{p-1} i =$$
$$= p^2 - (1/2)p(p-1) = (p/2)(p+1).$$

Analogously

$$(\mathrm{v},\mathrm{v}) = (q/2)(q+1).$$

Because

$$4 \cdot (u_p, \mathrm{v}_q)^2 = 2$$

we obtain

$$(u,\mathrm{v})^2 = (p \cdot u_p, q \cdot \mathrm{v}_q)^2 = p^2 \cdot q^2 \cdot (u_p, \mathrm{v}_q)^2 = (1/2) \cdot p^2 \cdot q^2.$$

Employing the Cauchy-Schwarz inequality

$$(u,\mathrm{v})^2 < (u,u) \cdot (\mathrm{v},\mathrm{v})$$

with u and v linearly independent gives

$$(1/2) \cdot p^2 q^2 < (p/2)(p+1) \cdot (q/2)(q+1).$$

Multiplying both sides by 2 implies

$$p^2 \cdot q^2 < (1/2)p(p+1)q(q+1) = (1/2)pq(p+1)(q+1)$$
$$pq < (1/2)(p+1)(q+1) = (1/2)pq + (1/2)(p+q+1)$$
$$pq < (p+q+1)$$
$$(p-1)(q-1) - 2 < 0$$

and eventually

$$(p-1)(q-1) < 2.$$

This restriction allows only the possibilities

$$(p,q) = (1, \geq 2), (p,q) = (\geq 2, 1), (p,q) = (2,2), (p,q) = (1,1).$$

The first two give the same Coxeter graph. It has type $B_r = C_r, r \geq 3$. The third possibility is the exceptional type F_4. The fourth possibility is type B_2.

10. *Admissible graphs from series D_r and exceptional graphs E_r, $r = 6,7,8$* : All admissible graphs from Figure 6.6 type d) belong to series D_r, $r \geq 4$, in Figure 6.2 or are the exceptional graphs types $E_r, r \in \{6,7,8\}$ in Figure 6.2.

Similar to the proof of part 9 we define the vectors from V

$$u := \sum_{i=1}^{r-1} i \cdot u_i, \mathrm{v} := \sum_{i=1}^{p-1} i \cdot v_i \text{ and } w := \sum_{i=1}^{q-1} i \cdot w_i.$$

Note

$$r, p, q \geq 2.$$

The three vectors (u, v, w) are pairwise orthogonal and the four vectors (x, u, v, w) are linearly independent. Denote by

$$\theta_1 := \sphericalangle(x,u),\ \theta_2 := \sphericalangle(x,\mathrm{v}),\ \theta_3 := \sphericalangle(x,w)$$

the angles between the vector x and each of the other three vectors. Similarly to the calculation in part 4 we obtain

$$1 > \sum_{i=1}^{3} cos^2\ \theta_i.$$

Similarly to the calculation in part 9 we have

$$(u,u) = (r/2)(r-1),\ (\mathrm{v},\mathrm{v}) = (p/2)(p-1),\ (w,w) = (q/2)(q-1).$$

Using in addition

$$4 \cdot (x, u_{r-1})^2 = 1$$

we obtain

$$cos^2\ \theta_1 = \frac{(x,u)^2}{\|x\|^2 \cdot \|u\|^2} = \frac{(r-1)^2 \cdot (x, u_{r-1}))^2}{\|u\|^2} = \frac{(r-1)^2 \cdot 2 \cdot (1/4)}{r(r-1)} = (1/2)(1-(1/r))$$

and analogously for $cos^2\ \theta_2$ and $cos^2\ \theta_3$. Hence

$$1 > \sum_{i=1}^{3} cos^2\ \theta_i = (1/2)[(1-(1/r)) + (1-(1/p)) + (1-(1/q))]$$

or

$$1 < (1/r) + (1/p) + (1/q).$$

W.l.o.g we may assume

$$q \le p \le r$$

Hence

$$1 < (1/r) + (1/p) + (1/q) \le 3/q$$

which implies

$$3 > q \ge 2, \text{ i.e. } q = 2.$$

We obtain

$$1 < (1/r) + (1/p) + (1/2), \text{ i.e. } 1/2 < (1/r) + (1/p),$$

Because $p \le r$ we obtain

$$1/2 < 2/p \text{ and } 2 \le p < 4.$$

In case $p = 3$ we have $r < 6$. In case $p = 2$ the parameter r may have any value ≥ 2.

Summing up: When $r \ge p \ge q$ then the only possibilities for (r, p, q) are

$$(5,3,2), (4,3,2), (3,3,2), (\ge 2,2,2).$$

These possibilities refer to the exceptional graphs E_8, E_7, E_6 or to the graphs from series $D_r, r \ge 4$.
□

Definition 6.28 (Irreducible root system). Consider a root system Φ of a vector space V and denote by $(-,-)$ a scalar product on V invariant with respect to the Weyl group $\mathscr{W}$ of Φ.

1. The root system Φ is *reducible* iff it splits into two non-empty, orthogonal subsets. iff a decomposition

$$\Phi = \Phi_1 \dot{\cup} \Phi_2, \ \Phi_1 \neq \emptyset, \ \Phi_2 \neq \emptyset,$$

exists with

$$(\Phi_1, \Phi_2) = 0.$$

Otherwise Φ is *irreducible*.

2. Analogously defined are the terms *reducible* and *irreducible* for a base Δ of Φ.

Proposition 6.29 (Irreducibility of a root system and connectedness of its Coxeter graph). *Consider a root system Φ of a vector space V and a base Δ of Φ.*

1. Φ is irreducible if and only if Δ is irreducible.

2. Δ is irreducible if and only if the Coxeter graph of Φ is connected.

Proof. 1. i) Suppose Φ reducible with decomposition

$$\Phi = \Phi_1 \dot{\cup} \Phi_2;\ \Phi_1 \neq \emptyset,\ \Phi_2 \neq \emptyset.$$

Define $\Delta_i := \Phi_i \cap \Delta, i = 1,2$. Then

$$\Delta = \Delta_1 \dot{\cup} \Delta_2$$

and $(\Delta_1, \Delta_2) = 0$.

Assume $\Delta_1 = \emptyset$. Then $\Delta = \Delta_2 \subset \Phi_2$ which implies

$$(\Phi_1, \Delta_2) \subset (\Phi_1, \Phi_2) = 0.$$

Because

$$V = span_{\mathbb{R}}\ \Delta = span_{\mathbb{R}}\ \Delta_2$$

we even get

$$(\Phi_1, V) = (\Phi_1, \Delta_2) = 0.$$

Therefore

$$\Phi_1 = 0,$$

which is excluded. As a consequence:

$$\Delta_1 \neq \emptyset \text{ and similarly } \Delta_2 \neq \emptyset.$$

The decomposition

$$\Delta = \Delta_1 \dot{\cup} \Delta_2$$

proves the reducibility of Δ.

ii) For the opposite direction suppose Δ reducible with decomposition

$$\Delta = \Delta_1 \dot{\cup} \Delta_2.$$

Denote by $\mathscr{W}$ the Weyl group of Φ. Define

$$\Phi_i := \mathscr{W}(\Delta_i),\ i = 1,2.$$

According to Lemma 6.20 any root $\beta \in \Phi$ has the form $\beta = w(\alpha)$ for suitable $w \in \mathscr{W}$ and $\alpha \in \Delta$. Therefore

$$\Phi = \Phi_1\ \cup\ \Phi_2.$$

The Weyl group is generated by the symmetries $\sigma_\alpha, \alpha \in \Delta$. Explicit calculation shows for each root $\alpha \in span\ \Delta_1$:

- If $\alpha_1 \in \Delta_1$ then also $\sigma_{\alpha_1}(\alpha) \in span\ \Delta_1$.

- If $\alpha_2 \in \Delta_2$ then $\sigma_{\alpha_2}(\alpha) = \alpha$.

  As a consequence $\mathscr{W}(\Delta_1) \subset span\ \Delta_1$ and similarly $\mathscr{W}(\Delta_2) \subset span\ \Delta_2$. The orthogonality $(\Delta_1, \Delta_2) = 0$ implies the orthogonality

$$(\Phi_1, \Phi_2) = 0, \text{ notably } \Phi_1 \cap \Phi_2 = \emptyset.$$

  Because $\Delta_i \neq \emptyset$ and $id \in \mathscr{W}$ also $\Phi_i \neq \emptyset,\ i = 1, 2$. Therefore Φ is reducible with decomposition

$$\Phi = \Phi_1 \dot{\cup} \Phi_2.$$

2. The claim is obvious: Two roots of Δ are not joined by an edge of the Coxeter graph if and only if the roots are orthogonal. □

The Coxeter graph, which employs the *product* of Cartan integers as its weights, does not encode the relative length of two roots, i.e. their length ratio. The length ratio derives from the *quotient* of the Cartan integers

$$\left(\frac{\|\beta\|}{\|\alpha\|}\right)^2 = \frac{<\beta, \alpha>}{<\alpha, \beta>} \text{ if } <\alpha, \beta> \neq 0.$$

Knowing the product of the Cartan integers is not sufficient to reconstruct the Cartan matrix. Hence the Coxeter graph does not encode the full information about the root system. We will see that a base of the root systems belonging to the types B_r, C_r, G_2, F_4 is made up by roots with different lengths. As a consequence, after complementing the Coxeter graph by the information about the length ratio the series B_r and C_r of root systems will differ for $r \geq 3$, see Theorem 6.31.

The *Dynkin diagram* of a root system complements the Coxeter graph by the information about the length ratio of the elements from a base. This information can be encoded by an orientation of the edges pointing from the long root to the short root in case two non-orthogonal roots have different length.

But in any case, the (absolute) *length* of a root is not defined because an invariant scalar product of a root system is not uniquely determined.

Definition 6.30 (Dynkin diagram of a root system). Consider a root system Φ and its Coxeter graph

$$Coxeter(\Phi) = (N, E)$$

The *Dynkin diagram* of Φ is the directed graph

$$Dynkin(\Phi) := (N, E_D)$$

with

- the same vertex set: $N = \Delta$.
- and the same edges from E, but some edges provided with an orientation: Each edge between two vertices, which represent roots of different lenght, carries an arrow pointing from the vertex of the long root to the vertex of the short root.

Note: The Coxeter graph and the Dynkin diagram of a root system have the same set of vertices and edges. In the Dynkin diagram an edge with a orientation indicates that the incident roots have different length.

Theorem 6.31 (Classification of connected Dynkin diagrams).
Consider an irreducible root system Φ. Then its Dynkin diagram belongs to exactly one the following types, see Figure 6.7:

- *Series $A_r, r \geq 1$*
- *Series $B_r, r \geq 2$*
- *Series $C_r, r \geq 3$*
- *Series $D_r, r \geq 4$*
- *Exceptional type G_2*
- *Exceptional type F_4*
- *Exceptional types $E_r,\ r \in \{6,7,8\}$.*

In Theorem 6.31 the two types B_r and C_r are distinguished only by the length ratio of their roots. Moreover, if one dismisses the restriction of the rank r then one has the following repetitions of low rank:

$$C_1 = B_1 = A_1 \text{ and } C_2 = B_2 \text{ and } D_3 = A_3.$$

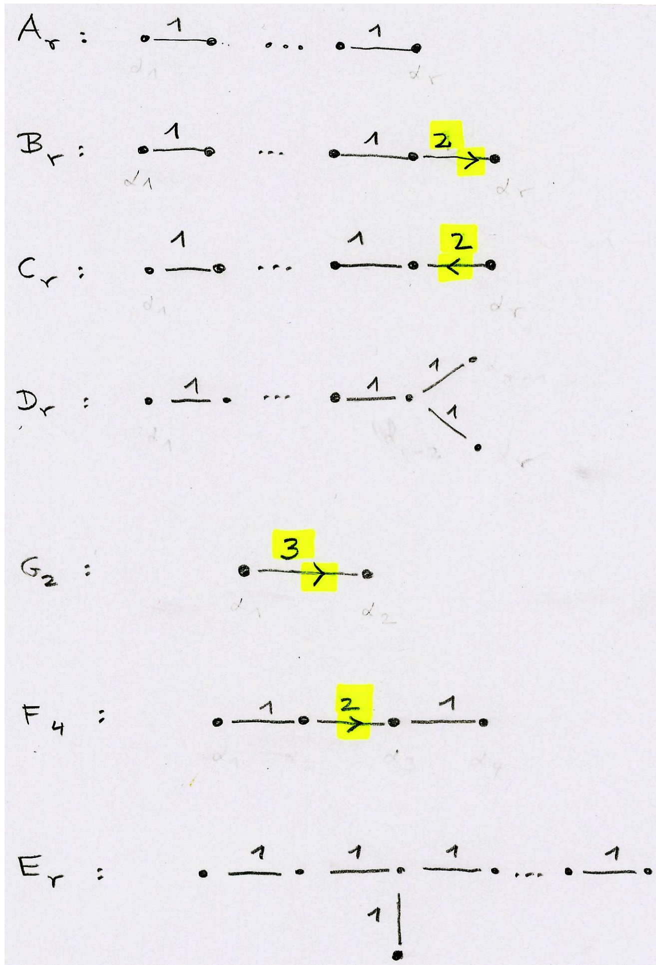

Fig. 6.7 The Dynkin diagrams of irreducible root systems

Proof. The statement follows from the classification of Coxeter graphs according to Theorem 6.27 and the restriction of the length ratio of two roots from a base according to Lemma 6.10: An edge from the Coxeter graph links two roots with length ratio $\neq 1$ if and only if the edge has multiplicity $m \in \{2,3\}$. Therefore the Dynkin diagrams distinguish between the two series B_r and C_r if $r \geq 3$. $\square$

The Dynkin diagram of a root system contains the full information of the Cartan matrix of the root system. Conversely from a Dynkin diagram one can construct a corresponding root system, cf. [24, Chap. 12.1].

Chapter 7
Explicit calculation of the root system

The objective of the present chapter is to classify complex semisimple Lie algebras L by their Dynkin diagram, more precisely the Dynkin diagram of a root system of L. The result is one of the highlights of Lie algebra theory. It completely encodes the structure of these Lie algebras by a certain finite graph, a data structure from discrete mathematics.

In the present chapter L denotes a complex semisimple Lie algebra, $H \subset L$ a maximal toral subalgebra and Φ the root set of L with respect to H, if not stated otherwise. The corresponding root set decomposition of (L, H) is

$$L = H \oplus \left(\bigoplus_{\alpha \in \Phi} L^{\alpha} \right).$$

According to the Cartan criterion the semisimpleness of L is equivalent to the non-degenerateness of the Killing form κ of L. Due to Proposition 5.16 also the restriction of the Killing form κ of L to H

$$\kappa|(H \times H) : H \times H \to \mathbb{C}$$

is non-degenerate.

7.1 Root systems of complex semisimple Lie algebras

In the present section Theorem 7.5 will show that the pair

$$R := (V := span_{\mathbb{R}} \Phi, \Phi)$$

satisfies the axioms of a root system from Definition 6.3.

Roots of L are linear functionals on H. Therefore we translate properties of H to properties of the dual space H^* and vice versa. The transfer is achieved by the restriction of the Killing form

$$\kappa|(H \times H) : H \times H \to \mathbb{C},$$

which is non-degenerate according to Proposition 5.16. We obtain an isomorphism

$$j : H \xrightarrow{\simeq} H^*$$

For each $\lambda \in H^*$ we denote by

$$t_\lambda := j^{-1}(\lambda) \in H$$

the inverse image. Then

$$\lambda = j(t_\lambda) = \kappa(t_\lambda, -).$$

In particular, each root $\alpha \in \Phi \subset H^*$ defines an elements $t_\alpha \in H$ with

$$\alpha = \kappa(t_\alpha, -) : H \to \mathbb{C}$$

relating roots from Φ to well-determined elements from the maximal toral subalgebra H.

Next we transfer the restriction $\kappa|(H \times H)$ of the Killing form to a non-degenerate bilinear form on the dual space H^*.

Definition 7.1 (Non-degenerate bilinear form on H^*). For the pair (L,H) the non-degenerate form $\kappa|(H \times H)$ on H induces a symmetric bilinear form on the dual space H^*

$$\kappa^* : H^* \times H^* \to \mathbb{C}, \kappa^*(\lambda, \mu) := \kappa(t_\lambda, t_\mu).$$

which is non-degenerate.

Combining the definition of $\kappa^*(\lambda, \mu)$ with the definition of t_μ and t_λ results in the following formula

$$\kappa^*(\lambda, \mu) = \kappa(t_\lambda, t_\mu) = \lambda(t_\mu) = \mu(t_\lambda).$$

Recall from Corollary 4.18 the semisimpleness of all complex Lie algebras L of the A,B,C,D-series within the range of Proposition 2.15. To motivate the separate steps in calculating the root system of L we first consider Example 7.2.

Example 7.2 (The root system of $sl(3,\mathbb{C})$).

Set $L := sl(3,\mathbb{C})$ and recall the standard basis $(E_{ij})_{1\le i,j\le 3}$ of $M(3\times 3,\mathbb{C})$.

i) *Maximal toral subalgebra*: The subalgebra

$$H := span\ < h_1 := E_{11} - E_{22}, h_2 := E_{22} - E_{33} > \subset L$$

is a maximal toral subalgebra. It has dimension $dim\ H = 2$.

ii) *Rootspace decomposition*: The rootspace decomposition of (L,H) is

$$L = H \oplus \left(\bigoplus_{\alpha\in\{\alpha_1,\alpha_2,\alpha_3\}} (L^{\alpha} \oplus L^{-\alpha}) \right).$$

All root spaces are 1-dimensional. The root set

$$\Phi = \{\pm\alpha_1, \pm\alpha_2, \pm\alpha_3\}$$

satisfies

- $L^{\alpha_1} = span < x_1 := E_{12} >, L^{-\alpha_1} = span\ < y_1 := E_{21} >$

$$\alpha_1 : H \to \mathbb{C},\ \alpha_1(h_1) = 2,\ \alpha_1(h_2) = -1$$

- $L^{\alpha_2} = span < x_2 := E_{23} >, L^{-\alpha_2} = span\ < y_2 := E_{32} >$

$$\alpha_2 : H \to \mathbb{C},\ \alpha_2(h_1) = -1,\ \alpha_2(h_2) = 2$$

- $L^{\alpha_3} = span < x_3 := E_{13} >, L^{-\alpha_3} = span\ < y_3 := E_{31} >$

$$\alpha_3 : H \to \mathbb{C},\ \alpha_3(h_1) = 1,\ \alpha_3(h_2) = 1$$

Therefore

$$\alpha_3 = \alpha_1 + \alpha_2$$

and

$$\Delta := \{\alpha_1, \alpha_2\}$$

is a base of Φ.

iii) *Killing form*: To compute the Killing form κ and its restriction

$$\kappa|(H\times H) : H\times H \to \mathbb{C}$$

one can use the formula

$$\kappa(z_1,z_2) = 2n\cdot tr(z_1\circ z_2), z_1,z_2 \in sl(n,\mathbb{C}),$$

with $n = 3$, see [24, Chapter 6, Ex. 7]. E.g.,

$$tr(h_1 \circ h_2) = tr((E_{11} - E_{22}) \circ (E_{22} - E_{33})) = tr(-E_{22} \circ E_{22}) = -tr(E_{22}) = -1.$$

We obtain

$$\big(\kappa(h_i, h_j)_{1 \le i,j \le 2}\big) = 6 \cdot \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$

a positive-definite matrix.

iv) *Lenght of roots*: With respect to the isomorphy

$$j : H \xrightarrow{\sim} H^*,\ t_\lambda \mapsto \lambda,$$

we get

$$t_{\alpha_1} = (1/6) \cdot h_1,\ t_{\alpha_2} = (1/6) \cdot h_2.$$

The family (α_1, α_2) is linearly independent and a basis of H^*. According to Definition 7.1, the Kiling form κ induces on H^* the bilinear form

$$\kappa^* : H^* \times H^* \to \mathbb{C}.$$

The induced bilinear from has with respect to the basis (α_1, α_2) of H^* the matrix

$$(\kappa^*(\alpha_i, \alpha_j)_{1 \le i,j \le 2}) = (\alpha_i(t_{\alpha_j})_{1 \le i,j \le 2}) = \Big(\big(\kappa(t_{\alpha_i}, t_{\alpha_j}\big)_{1 \le i,j \le 2}\Big) = (1/6) \cdot \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

In particular, all roots $\alpha_i, i = 1,2,3$, have the same lenght:

$$\kappa^*(\alpha_i, \alpha_i) = 2 \cdot (1/6) = 1/3.$$

The real vector space

$$V := span_{\mathbb{R}} < \alpha_1, \alpha_2 > .$$

satsifies

$$dim_{\mathbb{R}}\, V = dim_{\mathbb{C}}\, H^* = 2.$$

The restriction of the bilinear form κ^* to V is a real, positive-definite form $(-,-)$ on V, hence a scalar product. We now prove by explicit computation: The root set Φ of (L, H) defines a root system (V, Φ) in the sense of Definition 6.3.

v) *Weyl reflections*: Consider the Euclidean space $(V, (-,-))$. For $i = 1,2,3$ define the symmetries with vector $\alpha_i \in V$

$$\sigma_i : V \to V,\ x \mapsto x - 6 \cdot (x, \alpha_i) \cdot \alpha_i.$$

Here the factor 6 has been choosen in order to get

$$\sigma_i(\alpha_i) = -\alpha_i.$$

One checks

$$\sigma_1(\alpha_2) = \alpha_3, \sigma_1(\alpha_3) = \sigma_1(\alpha_1) + \sigma_1(\alpha_2) = -\alpha_1 + \alpha_2 + \alpha_1 = \alpha_2$$

$$\sigma_2(\alpha_1) = \alpha_3, \sigma_2(\alpha_3) = \alpha_1$$

$$\sigma_3(\alpha_1) = \alpha_1 - 6 \cdot (\alpha_1, \alpha_3) \cdot \alpha_3 = \alpha_1 - (\alpha_1 + \alpha_2) = -\alpha_2,$$

$$\sigma_3(\alpha_2) = \alpha_2 - (\alpha_1 + \alpha_2) = -\alpha_1.$$

Hence the restriction

$$\sigma_i | \Phi : \Phi \to \Phi$$

is well-defined and permutes the elements of Φ. Therefore

$$\sigma_i = \sigma_{\alpha_i},\ i = 1, 2, 3,$$

Moreover

$$\sigma_{-\alpha_i} = \sigma_{\alpha_i},\ i = 1, 2, 3,$$

because according to Corollary 6.8 these symmetries are the reflections on the hyperplanes

$$(-\alpha_i)^\perp = \alpha_i^\perp.$$

Each symmetry leaves the scalar product invariant, because for $\alpha \in \Phi,\ x, y \in V$,

$$(\sigma_\alpha(x), \sigma_\alpha(y)) = (x - 6(x, \alpha) \cdot \alpha, y - 6(y, \alpha) \cdot \alpha) =$$

$$(x, y) - 6(y, \alpha)(x, \alpha) - 6(x, \alpha)(\alpha, y) + 36(x, \alpha)(y, \alpha)(\alpha, \alpha) = (x, y)$$

using $(\alpha, \alpha) = 1/3$.

vi) *Cartan integers*: From the symmetries of part v) one reads off the Cartan numbers

$$< \alpha_i, \alpha_i >= 2,\ i = 1, 2, \text{ and } < \alpha_1, \alpha_2 >=< \alpha_2, \alpha_1 >= -1$$

which are integers indeed.

vii) *Reducedness*: Apparently for each $\alpha \in \Phi$ the only roots proportional to α are $\pm\alpha$.

viii) *Cartan matrix, Coxeter graph, Dynkin diagram*: The Cartan matrix with respect to

$$\Delta := \{\alpha_1, \alpha_2\}$$

and the given numbering is

$$Cartan(\Delta) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

According to Lemma 6.10 both roots from Δ have the same lenght and include the angle $(2/3) \cdot \pi$. Therefore Coxeter graph and Dynkin diagram of Φ contain the same information. According to the classification from Theorem 6.31 the root system of $sl(3,\mathbb{C})$ has type A_2. The scalar product $(-,-)$ induced from the Killing form is invariant with respect to the Weyl group

$$\mathscr{W} = span\ < \sigma_{\alpha_1}, \sigma_{\alpha_2} > .$$

The following two propositions collect the main properties of the root set Φ of a semisimple complex Lie algebra L. They generalize the result from Proposition 5.4 about the structure of $sl(2,\mathbb{C})$.

Proposition 7.3 considers the complex linear structure of L and its canonical subalgebras $sl(2,\mathbb{C}) \subset L$. While Proposition 7.4 considers the integrality properties of the root set Φ with respect to the bilinear form induced by the Killing form, see Definition 7.1. These properties assure that Φ is a root system in the sense of Definition 6.3. They allow to apply the classification of respectively Coxeter graphs from Theorem 6.27 and Dynkin diagrams from Theorem 6.31. They show which Dynkin diagrams result from the roots systems of the classical Lie algebras of the ABCD-series in Proposition 7.6 - 7.9.

Recall from Corollary 5.15 that for any root $\alpha \in \Phi$ of L also the negative $-\alpha$ is a root of L.

Proposition 7.3 prepares the proof of Theorem 7.5 by investigating in detail the root spaces $L^\alpha,\ \alpha \in \Phi$, of L.

Proposition 7.3 (Complex semisimple Lie algebras as $sl(2,\mathbb{C})$-modules). *Consider a pair (L,H) with L a complex semisimple Lie algebra and $H \subset L$ a maximal toral subalgebra. Denote by $(-,-)$ the non-degenerate bilinear form on H^* from Definition 7.1.*

Then the root set Φ and the root spaces L^α from the rootspace decomposition

$$L = H \oplus \left(\bigoplus_{\alpha \in \Phi} L^\alpha \right)$$

have the following properties:

1. Spanning*: The root set spans H^**

$$span_{\mathbb{C}} \Phi = H^*$$

2. Duality of root spaces*: For each $\alpha \in \Phi$ the vector spaces L^α and $L^{-\alpha}$ are dual with respect to the Killing form, i.e. the bilinear map*

$$L^\alpha \times L^{-\alpha} \to \mathbb{C},\ (x,y) \mapsto \kappa(x,y),$$

is non-degenerate. For each $x \in L^\alpha,\ y \in L^{-\alpha}$ *holds*

$$[x,y] = \kappa(x,y) \cdot t_\alpha,$$

in particular

$$[L^\alpha, L^{-\alpha}] = \mathbb{C} \cdot t_\alpha$$

3. Subalgebras $S_\alpha \simeq sl(2,\mathbb{C})$*: For each* $\alpha \in \Phi$ *holds*

$$(\alpha,\alpha) \neq 0$$

and there exists a unique element

$$h_\alpha \in [L^\alpha, L^{-\alpha}] \text{ with } \alpha(h_\alpha) = 2,$$

in particular

$$h_\alpha = \frac{2}{(\alpha,\alpha)} \cdot t_\alpha$$

For each non-zero $x_\alpha \in L^\alpha$ *exists an element* $y_\alpha \in L^{-\alpha}$ *satisfying*

$$[x_\alpha, y_\alpha] = h_\alpha$$

The morphism of Lie algebras

$$sl(2,\mathbb{C}) \to S_\alpha := span_{\mathbb{C}} < h_\alpha,\ x_\alpha,\ y_\alpha > \subset L,$$

defined on the standard basis of $sl(2,\mathbb{C})$ *as*

$$h := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mapsto h_\alpha,\ \ x := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \mapsto x_\alpha,\ \ y := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \mapsto y_\alpha,$$

is an isomorphism. Thereby L becomes an $sl(2,\mathbb{C})$*-module.*

Note: In Proposition 7.3, part 3 the element $y_\alpha \in L^{-\alpha}$ is uniquely determined, see Proposition 7.4.

Proof. 1. *Spanning*: For an indirect proof assume

$$span_{\mathbb{C}} \Phi \subsetneqq H^*$$

is a proper subspace. A non-zero linear functional

$$h \in (H^*)^* \simeq H$$

exists which satisfies for all $\alpha \in \Phi$

$$h \neq 0 \text{ and } \alpha(h) = 0.$$

Then for all $\alpha \in \Phi$

$$[h, L^\alpha] = 0,$$

and also

$$[h, H] = 0$$

because the maximal toral subalgebra subalgebra H is Abelian due to Proposition 5.2. The root space decomposition of L implies $[h, L] = 0$, i.e.

$$h \in Z(L).$$

Because the center of the semisimple Lie algebra L is trivial, one obtains

$$h = 0$$

This contradiction proves $span_{\mathbb{C}} \Phi = H^*$.

In order to prove the remaining claims we recall from Lemma 5.14: For two linear functionals $\lambda, \mu \in H^*$ with

$$\lambda + \mu \neq 0$$

the orthogonality of the eigenspaces, i.e.

$$\kappa(L^\lambda, L^\mu) = 0.$$

As a consequence, for two roots α, β with $\beta \neq -\alpha$, i.e.

$$\alpha + \beta \neq 0,$$

holds

$$\kappa(L^\alpha, L^\beta) = 0,$$

and due to $H = L^0$ also

$$\kappa(L^\alpha, H) = 0.$$

2. *Duality of root spaces*:

   - If
     $$x \in L^\alpha, \ \alpha \in \Phi,$$
     satisfies $\kappa(x, L^{-\alpha}) = 0$, then the root space decomposition from Definition 5.13 and Lemma 5.14 imply $\kappa(x, L) = 0$. The non-degenerateness of κ implies $x = 0$. Hence the bilinear map
     $$L^\alpha \times L^{-\alpha} \to \mathbb{C}, \ (x, y) \mapsto \kappa(x, y)$$
     is non-degenerate.

   - We have
     $$[L^\alpha, L^{-\alpha}] \subset L^0 = H$$

according to Theorem 5.17. The Killing form is associative according to Lemma 4.1. Hence for all $h \in H$ and arbitrary, but fixed $(x,y) \in L^\alpha \times L^{-\alpha}$

$$\kappa(h,[x,y]) = \kappa([h,x],y) = \kappa(\alpha(h)x,y) = \alpha(h)\kappa(x,y) =$$

$$= \kappa(t_\alpha,h)\cdot\kappa(x,y) = \kappa(h,t_\alpha)\cdot\kappa(x,y) = \kappa(h,\kappa(x,y)\cdot t_\alpha),$$

here we used the definition of t_α with

$$\alpha = \kappa(t_\alpha,-).$$

Non-degenerateness of the restriction $\kappa|(H\times H)$ implies

$$[x,y] = \kappa(x,y)\cdot t_\alpha.$$

The duality between L^α and $L^{-\alpha}$ provides for each non-zero $x_\alpha \in L^\alpha$ an element $y_\alpha \in L^{-\alpha}$, such that

$$\kappa(x_\alpha,y_\alpha) \neq 0,$$

which proves in particular

$$[L^\alpha,L^{-\alpha}] = \mathbb{C}\cdot t_\alpha$$

3. *Subalgebra* $S_\alpha \simeq sl(2,\mathbb{C})$:

- We first claim

$$\alpha(t_\alpha) \neq 0:$$

For the proof choose an arbitrary, but fixed non-zero $x_\alpha \in L^\alpha$. Because L^α and $L^{-\alpha}$ are dual according to part 2 one can find an element

$$y \in L^{-\alpha} \text{ with } \kappa(x_\alpha,y) \neq 0,$$

w.l.o.g.

$$\kappa(x_\alpha,y) = 1 \text{ and } [x_\alpha,y] = \kappa(x_\alpha,y)\cdot t_\alpha = t_\alpha$$

For an indirect proof of the claim assume on the contrary

$$\alpha(t_\alpha) = 0.$$

Consider the subalgebra

$$S := < x_\alpha,\ y,\ t_\alpha > \subset L.$$

Because

$$t_\alpha \in H,\ x_\alpha \in L^\alpha \text{ and } y \in L^{-\alpha},$$

the commutators satisfy

$$[t_\alpha, x_\alpha] = \alpha(t_\alpha)x_\alpha = 0,\ [t_\alpha, y] = -\alpha(t_\alpha)y = 0,\ [x_\alpha, y] = t_\alpha.$$

Apparently the Lie algebra S is nilpotent, in particular solvable. The semisimplicity of L implies that the adjoint representation

$$ad : L \to gl(L)$$

embeds L into the matrix Lie algebra $gl(L)$. According to Lie's theorem, see Theorem 3.21, with respect to a suitable basis of L the solvable subalgebra S embeds into the subalgebra of upper triangular matrices. Thereby the element

$$ad(t_\alpha) = ad\ [x_\alpha, y] = [ad\ x_\alpha, ad\ y],$$

the commutator of two endomorphisms, becomes a strict upper triangular matrix. Therefore $ad(t_\alpha)$ is a nilpotent endomorphism, i.e. $t_\alpha \in L$ is ad-nilpotent. Because H is a maximal toral subalgebra, the element $t_\alpha \in H$ is also ad-semisimple, hence $t_\alpha = 0$. This contradiction proves the claim

$$0 \neq \alpha(t_\alpha) = (\alpha, \alpha).$$

- Because

  $$[L^\alpha, L^{-\alpha}] = \mathbb{C}\ \cdot t_\alpha \text{ and } \alpha(t_\alpha) \neq 0$$

  there exists an element $h_\alpha \in [L^\alpha, L^{-\alpha}]$ satisfying

  $$\alpha(h_\alpha) = 2$$

  From

  $$\frac{\alpha(h_\alpha)}{2} = \frac{\alpha(t_\alpha)}{(\alpha, \alpha)}$$

  follows

  $$h_\alpha = \frac{2}{(\alpha, \alpha)} \cdot t_\alpha$$

  Multiplying $y \in L^{-\alpha}$ by a suitable constant provides an element $y_\alpha \in L^{-\alpha}$ with

  $$[x_\alpha, y_\alpha] = h_\alpha$$

  We obtain

  $$[h_\alpha, x_\alpha] = \alpha(h_\alpha) \cdot x_\alpha = 2 \cdot x_\alpha$$
  $$[h, y_\alpha] = -\alpha(h_\alpha) \cdot y_\alpha = -2 \cdot y_\alpha$$
  $$[x_\alpha, y_\alpha] = h_\alpha.$$

  Therefore the subalgebra

  $$S_\alpha := span_{\mathbb{C}}\ < h_\alpha,\ x_\alpha,\ y_\alpha >$$

is isomorphic to $sl(2,\mathbb{C})$.

□

Because L is a $sl(2,\mathbb{C})$-module with respect to each subalgebra S_α the results from section 5.2 about the structure of irreducible $sl(2,\mathbb{C})$-module apply. They imply a series of integrality and rationality properties of L, see Proposition 7.4. For two roots $\alpha,\beta \in \Phi$ we will often employ the formula

$$\beta(h_\alpha) = \frac{2\cdot\beta(t_\alpha)}{(\alpha,\alpha)} = \frac{2\cdot\kappa(t_\beta,t_\alpha)}{(\alpha,\alpha)} = \frac{2\cdot(\beta,\alpha)}{(\alpha,\alpha)}.$$

It derives from the relation between h_α and t_α from Proposition 7.3 and from the defining relation

$$\kappa_H(t_\beta,-) = \beta$$

The numbers $\beta(h_\alpha)$ will turn out as the Cartan integers $<\beta,\alpha>$ of the root system. Notably they are integers.

Proposition 7.4 (Integrality and rationality properties of the root set). *Consider a pair (L,H) with L a complex semisimple Lie algebra and $H \subset L$ a maximal toral subalgebra. Denote by Φ the roots of (L,H), Then the roots and the rootspaces from the rootspace decomposition*

$$L = H \oplus \left(\bigoplus_{\alpha\in\Phi} L^\alpha \right)$$

have the following properties:

1. Root spaces are 1-dimensional*: For each root $\alpha \in \Phi$ also $-\alpha \in \Phi$ and*

$$dim\ L^\alpha = dim\ L^{-\alpha} = 1.$$

As a consequence

$$|\Phi| = dim\ L - dim\ H$$

2. Integrality*: For each pair $\alpha,\beta \in \Phi$ and the element h_α from Proposition 7.3 holds*

$$\beta(h_\alpha) \in \mathbb{Z} \text{ and } \beta - \beta(h_\alpha)\cdot\alpha \in \Phi.$$

3. Rationality and scalar product*: Consider a basis $\mathscr{B} = (\alpha_1,...,\alpha_r)$ of the complex vector space H^* made up from roots $\alpha_i \in \Phi, i = 1,...,r$. Then any root $\beta \in \Phi$ is a* rational *combination of elements from $\mathscr{B}$, i.e.*

$$\Phi \subset V_{\mathbb{Q}} := span_{\mathbb{Q}}\{\alpha_1,...,\alpha_r\}$$

and

$$dim_{\mathbb{Q}}\, V_{\mathbb{Q}} = dim_{\mathbb{C}}\, H.$$

The bilinear form $(-,-)$ from Definition 7.1 restricts from H^* to a rational form $(-,-)_{\mathbb{Q}}$ on $V_{\mathbb{Q}} \subset H^*$, and

$$(-,-)_{\mathbb{Q}} : V_{\mathbb{Q}} \times V_{\mathbb{Q}} \to \mathbb{Q}$$

is positive definite, i.e. a scalar product.

4. Symmetries*: Extending scalars from $\mathbb{Q}$ to $\mathbb{R}$ shows*

$$dim_{\mathbb{R}}\, \Phi = dim_{\mathbb{C}}\, H^*$$

and extends the scalar product from $V_{\mathbb{Q}}$ to a scalar product $(-,-)$ on V.

For each root $\alpha \in \Phi$ the map

$$\sigma_\alpha : V \to V,\ v \mapsto \sigma_\alpha(v) := v - \frac{2\cdot(v,\alpha)}{(\alpha,\alpha)}\cdot\alpha,$$

is a symmetry of V with vector α and Cartan integers

$$<\beta,\alpha>:= \frac{2\cdot(\beta,\alpha)}{(\alpha,\alpha)} = \beta(h_\alpha) \in \mathbb{Z},\ \beta \in \Phi.$$

It satisfies

$$\sigma_\alpha(\Phi) \subset \Phi$$

In addition: Each symmetry σ_α leaves invariant the scalar product $(-,-)$ on V.

5. Proportional roots*: For each $\alpha \in \Phi$ the only roots proportional to α are $\pm\alpha$.*

Proof. 1. *Root spaces are* 1*-dimensional*: According to Proposition 7.3, part 2 the root spaces L^α and $L^{-\alpha}$ are dual with respect to the Killing form κ, i.e. the bilinear form

$$\kappa|(L^\alpha \times L^{-\alpha}) : L^\alpha \times L^{-\alpha} \to \mathbb{C},\ (x,y) \mapsto \kappa(x,y),$$

is non-degenerate. In order to show

$$dim\, L^\alpha = 1$$

we assume on the contrary

$$dim\, L^\alpha = dim\, L^{-\alpha} > 1.$$

Consider an element $x_\alpha \in L^\alpha,\ x_\alpha \neq 0$, and the correponding subalgebra

$$S_\alpha \simeq sl(2,\mathbb{C})$$

from Proposition 7.3, part 3. According to Proposition 7.3, part 1 an element $h_\alpha \in S_\alpha$ exists with
$$\alpha(h_\alpha) = 2.$$
The restriction
$$\kappa(x_\alpha, -)|L^{-\alpha} : L^{-\alpha} \to \mathbb{C}$$
is non-zero. The linear functional has a non-trivial kernel, i.e. an element
$$e \in L^{-\alpha},\ y \neq 0,$$
exists with
$$\kappa(x_\alpha, e) = 0$$
The latter formula implies $[x_\alpha, e] = 0$, see Proposition 7.3, part 2. As a consequence $e \in L$ is a primitive element of an irreducible S_α-submodule of L with weight
$$(-\alpha)(h_\alpha) = -2 < 0.$$
The latter property contradicts the fact that all primitive elements of an irreducible $sl(2,\mathbb{C})$-module have a non-negative weight, see Proposition 5.7.

2. *Integrality*: Choose a non-zero element $y \in L^\beta$. The element y can be considered in two different roles. On one hand, being a root vector of L the element $y \in L^\beta$ satisfies
$$[h_\alpha, y] = \beta(h_\alpha) \cdot y.$$
On the other hand considered as an element of the S_α-module L according to Proposition 7.3, part 3, the element $y \in L$ is a weight vector with weight
$$\beta(h_\alpha) \in \mathbb{Z}.$$
Applying $y_\alpha \in S_\alpha$ reduces the weight in K by subtracting the number 2 and reduces the corresponding root of L by subtracting the linear functional α. Similarly, applying $x_\alpha \in S_\alpha$ adds respectively the number 2 and the linear functional α.
$$K := \bigoplus_{j \in \mathbb{Z}} L^{\beta + j\cdot\alpha},$$
is a S_α-submodule of L. The action of S_α on K moves $y \in K$ through weight spaces of K: For each $j \in \mathbb{Z}$

- either
$$L^{\beta + j\cdot\alpha} = 0$$
- or
$$dim\, L^{\beta + j\cdot\alpha} = 1 \ \text{ and } (\beta + j\cdot\alpha \in \Phi \text{ or } \beta + j\cdot\alpha = 0).$$

After finitely many applications of x_α to $y \in L$ a primitive element of K is obtained and Proposition 5.7 applies, see Figure 7.1.

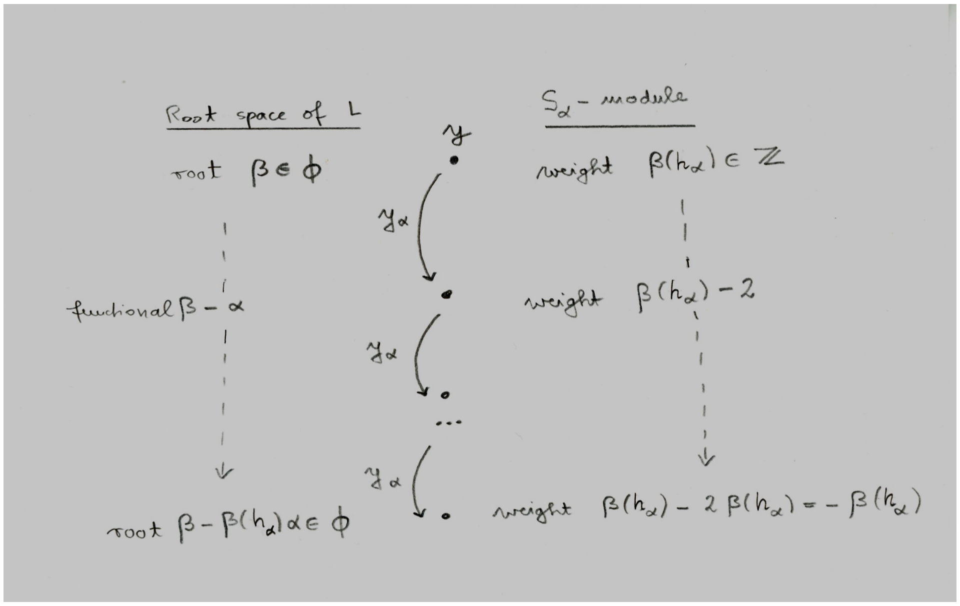

Fig. 7.1 The double role of $y \in L^\beta$ (Case $\beta(h_\alpha) \geq 0$)

In particular

$$\beta(h_\alpha) \text{ weight} \implies -\beta(h_\alpha) = \beta(h_\alpha) - 2 \cdot \beta(h_\alpha) \text{ weight}$$

which implies

$$\beta - \beta(h_\alpha) \cdot \alpha \in \Phi$$

Note: All weight spaces of K displayed on the right-hand side are non-zero. Hence also the corresponding eigenspaces of the linear functionals on the left-hand side are non-zero. In particular

$$\beta - \beta(h_\alpha) \cdot \alpha \in \Phi$$

is a root.

3. *Rationality and scalar product*: First, due to Proposition 7.3, part 1 the complex vector space H^* has a basis

$$\mathscr{B} = (\alpha_j)_{1 \leq j \leq r}$$

of roots.

The subsequent proof goes along the following steps:

- Claim: $\Phi \subset V_{\mathbb{Q}}$. We show for the proof that for each root $\beta \in \Phi$ the uniquely determined coefficients $c_i \in \mathbb{C},\ i = 1, ..., r$, in the representation

$$\beta = \sum_{i=1}^{r} c_i \cdot \alpha_i$$

are even *rational*, i.e. $c_i \in \mathbb{Q}$ for $i = 1, ..., r$.

Multiplying the above representation of β successively for $j = 1, ..., r$ by

$$\frac{2}{(\alpha_j, \alpha_j)}$$

and applying the bilinear form $(-, \alpha_j)$ to the resulting equation gives a system of linear equations

$$b = A \cdot c$$

for the vector of indeterminates $c := (c_1, ..., c_r)^\top$. The left-hand side is the vector

$$b = \left(\frac{2 \cdot (\beta, \alpha_j}{(\alpha_j, \alpha_j)} \right)^\top$$

and the coefficient matrix is

$$A = \left(a_{ji} := \frac{2 \cdot (\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)} \right)$$

The system is defined over the ring $\mathbb{Z}$ because for two roots $\gamma, \delta \in \Phi$

$$2 \cdot \frac{(\gamma, \delta)}{(\delta, \delta)} = \gamma(h_\delta)$$

and

$$\gamma(h_\delta) \in \mathbb{Z}$$

according to part 2. The coefficient matrix

$$A \in M(r \times r, \mathbb{Z})$$

is invertible as an element from $GL(r, \mathbb{Q})$. It originates by multiplying for $j = 1, ..., r$ the row with index j of the matrix

$$(\alpha_i, \alpha_j)_{1 \leq i, j \leq r} \in GL(r, \mathbb{C})$$

by the non-zero scalar $2/(\alpha_j, \alpha_j)$. And the latter matrix defines the bilinear non-degenerate form $(-, -)$ on H^*. As a consequence, the unique solution of the linear system of equations is already defined over the base field $\mathbb{Q}$, i.e.

$$c_j \in \mathbb{Q} \text{ for all } j = 1, ..., r.$$

Hence $\Phi \subset V_{\mathbb{Q}}$. As a consequence

$$dim_{\mathbb{Q}}\ span_{\mathbb{Q}}\ \Phi = dim_{\mathbb{Q}}\ V_{\mathbb{Q}} = dim_{\mathbb{C}}\ H^*,$$

and

$$V := span_{\mathbb{R}}\ \Phi \implies dim_{\mathbb{R}}\ V = dim_{\mathbb{C}}\ H^*.$$

- We claim that the restricted bilinear form

$$(-,-)_{\mathbb{Q}} := (-,-)|V_{\mathbb{Q}}$$

is a scalar product, i.e. that it is defined over the field $\mathbb{Q}$ and is positive definite:

– *Positive-definiteness*: For two linear functionals $\lambda, \mu \in V_{\mathbb{Q}}$ we compute

$$(\lambda, \mu) = \kappa(t_\lambda, t_\mu) = tr(ad\ t_\lambda \circ ad\ t_\mu)$$

for the endomorphism

$$ad\ t_\lambda \circ ad\ t_\mu : L \to L$$

In order to evaluate the trace we employ the defining property of a root space: If $z \in L^\gamma$ then

$$(ad\ t_\mu)(z) = \gamma(t_\mu) \cdot z$$

and

$$(ad\ t_\lambda \circ ad\ t_\mu)(z) = \gamma(t_\lambda) \cdot \gamma(t_\mu) \cdot z.$$

Using the Cartan decomposition

$$L = H \oplus \left(\bigoplus_{\gamma \in \Phi} L^\gamma \right)$$

and observing $[H,H] = 0$ and $dim\ L^\gamma = 1$ due to part 1, we obtain

$$(\lambda, \mu) = tr(ad\ t_\lambda \circ ad\ t_\mu) = \sum_{\gamma \in \Phi} \gamma(t_\lambda) \cdot \gamma(t_\mu)$$

as sum of the eigenvalues of

$$ad\ t_\lambda \circ ad\ t_\mu : L \to L.$$

In particular

$$(\lambda, \lambda) = \sum_{\gamma \in \Phi} \gamma(t_\lambda)^2 \ \geq 0.$$

The vanishing

$$(\lambda, \lambda) = 0$$

implies: For all $\gamma \in \Phi$ holds

$$0 = \gamma(t_\lambda) = \kappa^*(\lambda, \gamma),$$

and therefore $\lambda = 0$ because

$$span_{\mathbb{C}}\ \Phi = H^*$$

and

$$\kappa^* : H^* \times H^* \to \mathbb{C}$$

is non-degenerate.

– *Defined over* $\mathbb{Q}$: For each root $\alpha \in \Phi$ the relation between the two elements t_α and h_α from H due to Proposition 7.3, part 3 and the integrality $\gamma(h_\alpha) \in \mathbb{Z}$ from part 2 show

$$2 \cdot \frac{(\gamma, \alpha)}{(\alpha, \alpha)} = 2 \cdot \frac{\gamma(t_\alpha)}{(\alpha, \alpha)} = \gamma(h_\alpha) \in \mathbb{Z}$$

Hence

$$4 \cdot \gamma(t_\alpha)^2 = (\alpha, \alpha)^2 \cdot \gamma(h_\alpha)^2.$$

We obtain

$$(\alpha, \alpha) = \sum_{\gamma \in \Phi} \gamma(t_\alpha)^2 = (1/4) \cdot (\alpha, \alpha)^2 \cdot \sum_{\gamma \in \Phi} \gamma(h_\alpha)^2.$$

Dividing both sides by $(\alpha, \alpha) > 0$ shows

$$1 = (1/4) \cdot (\alpha, \alpha) \cdot \sum_{\gamma \in \Phi} \gamma(h_\alpha)^2$$

and

$$(\alpha, \alpha) = \frac{4}{\sum_{\gamma \in \Phi} \gamma(h_\alpha)^2} \in \mathbb{Q}_+.$$

As a consequence for each $\beta \in \Phi$,

$$(\beta, \alpha) = (1/2)(\alpha, \alpha) \cdot \beta(h_\alpha) \in \mathbb{Q}.$$

The general elements

$$\lambda,\ \mu \in V_{\mathbb{Q}}$$

have the representations

$$\lambda = \sum_{i=1}^{r} c_i \cdot \alpha_i \text{ and } \mu = \sum_{j=1}^{r} d_j \cdot \alpha_j \text{ with } c_i,\ d_j \in \mathbb{Q},\ i,\ j = 1, ..., r.$$

Therefore

$$(\lambda, \mu) = \sum_{i,j=1}^{r} c_i \cdot d_j \cdot (\alpha_i, \alpha_j) \in \mathbb{Q}$$

Hence the restriction

$$(-,-)|V_{\mathbb{Q}}$$

is defined over $\mathbb{Q}$.

4. *Symmetries*: The map

$$\sigma_\alpha : V \to V$$

is a symmetry of V with vector α. Due to part 2) its Cartan number is an integer

$$<\beta,\alpha>= \frac{2\cdot(\beta,\alpha)}{(\alpha,\alpha)} = \beta(h_\alpha) \in \mathbb{Z}.$$

The inclusion

$$\sigma_\alpha(\Phi) \subset \Phi$$

has been proven in part 2). In order to prove the invariance of the scalar product with respect to the Weyl group it suffices to consider three roots $\alpha, \beta, \gamma \in \Phi$:

$$(\sigma_\alpha(\beta), \sigma_\alpha(\gamma)) = (\beta - \beta(h_\alpha)\alpha, \gamma - \gamma(h_\alpha)\cdot\alpha) =$$

$$= (\beta,\gamma) - \beta(h_\alpha)(\alpha,\gamma) - \gamma(h_\alpha)(\beta,\alpha) + \beta(h_\alpha)\gamma(h_\alpha)(\alpha,\alpha)$$

Using

$$(\alpha,\gamma) = (1/2)(\alpha,\alpha)\gamma(h_\alpha) \text{ and } (\alpha,\beta) = (1/2)(\alpha,\alpha)\beta(h_\alpha)$$

we confirm

$$(\sigma_\alpha(\beta), \sigma_\alpha(\gamma)) = (\beta,\gamma).$$

5. *Proportional roots*: Assume the existence of a root $\gamma \in \Phi$ such that also

$$t\cdot\gamma \in \Phi \text{ for a suitable } t \in \mathbb{R}\setminus\{\pm 1\}$$

Then exists a root $\beta \in \Phi$ and $t \in \mathbb{R}$ with $0 < t < 1$ such that also

$$\alpha := t\cdot\beta \in \Phi$$

We calculate

$$\sigma_\beta(\alpha) = \alpha - \alpha(h_\beta)\cdot\beta = \alpha - t\cdot\beta(h_\beta)\cdot\beta$$

For the Cartan integer

$$<\alpha,\beta>= t\cdot\beta(h_\beta) = 2t$$

follows

$$2t \in \mathbb{Z}$$

which - due to $0 < t < 1$ - implies

$$t = 1/2 \text{ and } \alpha,\ 2\alpha \in \Phi.$$

But $3\alpha \notin \Phi$: Otherwise repeating the argument with the two proportional roots 2α and 3α would imply

$$2\alpha = \frac{1}{2} \cdot 3\alpha,$$

a contradiction.

With respect to the $sl(2,\mathbb{C})$-module structure of L induced from the action of

$$S_\alpha = < x_\alpha, y_\alpha, h_\alpha >,$$

see Proposition 7.3, part 3), we have for each root vector $z \in L^{2\alpha}$:

- $h_\alpha.z = 2 \cdot 2 \cdot z$, because $z \in L^{2\alpha}$.
- $x_\alpha.z \in L^{3\alpha} = \{0\}$ implies $x_\alpha.z = 0$
- From $h_\alpha = [x_\alpha, y_\alpha]$ results

  $$h_\alpha.z = x_\alpha.(y_\alpha.z) - y_\alpha.(x_\alpha.z) = x_\alpha.(y_\alpha.z)$$

  Then
  $$y_\alpha.z \in L^\alpha = \mathbb{C} \cdot x_\alpha \implies x_\alpha.(y_\alpha.z) \in \mathbb{C} \cdot [x_\alpha, x_\alpha] = 0,$$
  which implies
  $$h_\alpha.z = 4z = \{0\},$$
  a contradiction to $z \neq 0$.

□

We now combine the result of Proposition 7.3 and 7.4, and construct the root system of a semisimple complex Lie algebra L.

Theorem 7.5 (Root system of a complex semisimple Lie algebra). *Let L be a complex semisimple Lie algebra.*

1. *After choosing a maximal toral subalgebra $H \subset L$ the root space decomposition from Definition 5.18*

   $$L = H \oplus \left(\bigoplus_{\alpha \in \Phi} L^\alpha \right)$$

   determines the set Φ of roots, and

   $$R := (V := span_{\mathbb{R}} \Phi, \Phi)$$

   is a root system according to Definition 6.3.

2. *From the Killing form κ of L derives a scalar product $(-,-)$ on V*

$$(\lambda,\mu) := \kappa(t_\lambda, t_\mu),\ \lambda,\mu \in V,$$

which is defined over $\mathbb{Q}$, i.e. satisfying for $\alpha,\beta \in \Phi$

$$(\alpha,\beta) \in \mathbb{Q}.$$

The scalar product is invariant with respect to the Weyl group $\mathscr{W}$ of R.

Note: Due to Remark 5.21 each two maximal toral subalgebras of L are conjugate under an automorphism of L.

Proof. The pair $R = (V, \Phi)$ has the following properties:

- *(R1) Finite and spanning*: The root set Φ is finite because L has finite dimension, $0 \notin \Phi$, and Φ spans V by definition.

- *(R2) Invariance under distinguished symmetries*: We choose a basis

$$\mathscr{B} := (\beta_j)_{j=1,\dots,r},\ r := dim_{\mathbb{C}}\ H$$

of V formed by elements from Φ. For each $\alpha \in \Phi$ we define a symmetry

$$\sigma_\alpha : V \to V$$

as follows: For $\beta \in \mathscr{B}$ set

$$\sigma_\alpha(\beta) := \beta - \beta(h_\alpha)\cdot\alpha,$$

and extend the definition by linearity. Here

$$h_\alpha \in [L^\alpha, L^{-\alpha}] \subset H$$

denotes the uniquely determined element from Proposition 7.3, part 3 with

$$\alpha(h_\alpha) = 2.$$

Then for all roots $\beta \in \Phi$

$$\sigma_\alpha(\beta) = \beta - \beta(h_\alpha)\cdot\alpha.$$

and

$$\sigma_\alpha(\alpha) = -\alpha.$$

Proposition 7.4, part 4 and part 2 imply that σ_α is a symmetry with vector $\alpha \in V$ and

$$\sigma_\alpha(\Phi) \subset \Phi.$$

- *(R3) Cartan integers*: By construction, for each pair of roots $\alpha,\ \beta \in \Phi$ holds

$$<\beta,\alpha> = \beta(h_\alpha).$$

The integrality result from Proposition 7.4, part 2 shows

$$<\beta,\alpha> \in \mathbb{Z}.$$

- *(R4) Reducedness*: The reducedness of R holds due to Proposition 7.4, part 5.

The bilinear form $(-,-)$ was introduced in Definition 7.1. Its properties follow from Proposition 7.4, part 3 and part 4. □

7.2 Root systems of the A,B,C,D-series in explicit form

We will show in the present section that the Dynkin diagrams of type A_r, B_r, C_r, D_r from Theorem 6.31 are the Dynkin diagrams of the root systems of the complex Lie algebras belonging to the classical groups of the correponding types, see Proposition 2.15. We show the simpleness of these Lie algebras as a consequence of their Cartan decomposition. We follow [21, Chap. III, §8, Chap. X, §3] and [18, Chap. 7.7].

The Lie algebras of the classical groups are subalgebras of the Lie algebra $sl(n,\mathbb{C})$. We introduce the following notation for the elements of the canonical basis of the vector space $M(n\times n,\mathbb{C})$:

$$E_{ij} \in M(n\times n,\mathbb{C})$$

is the matrix with entry = 1 at place (i,j) and entry = 0 for all other places. Our matrix computations are based on the formulas

$$E_{ij}\cdot E_{kl} = \delta_{jk}\cdot E_{il}, 1 \le i,j,k,l \le n.$$

The family

$$(E_{ii})_{1\le i\le n}$$

is a basis of the subspace of diagonal matrices $\mathfrak{d}(n,\mathbb{C})$. Denote the elements of the dual base by

$$\hat{\varepsilon}_i := (E_{ii})^* \in \mathfrak{d}(n,\mathbb{C})^*, i = 1,...,n.$$

Due to Corollary 4.18 the Lie algebras

$$sl(n,\mathbb{C}),\ so(n,\mathbb{C}),\ sp(2n,\mathbb{C})$$

of the classical complex matrix groups from the A,B,C,D-series are semisimple.

Proposition 7.6 (Type A_r). *The Lie algebra*

$$L := sl(r+1,\mathbb{C}), r \geq 1,$$

has the following characteristics:

1. *dim* $L = (r+1)^2 - 1$.

2. *The subalgebra*
$$H := \mathfrak{d}(r+1,\mathbb{C}) \cap L$$
is a maximal toral subalgebra with dim $H = r$.

3. *The family* $(h_i)_{i=1,...,r}$ *with*
$$h_i := E_{ii} - E_{i+1,i+1}$$
is a basis of H.

4. *Define the functionals*
$$\varepsilon_i := (\hat{\varepsilon}_i|H) \in H^*,\ i = 1,...,r+1.$$
Then the root set Φ *of* L *has the elements*
$$\varepsilon_i - \varepsilon_j,\ 1 \leq i \neq j \leq r+1.$$
The corresponding root spaces are 1*-dimensional, generated by the elements*
$$E_{ij},\ 1 \leq i \neq j \leq r+1.$$
A base of Φ *is the set* $\Delta := \{\alpha_i : 1 \leq i \leq r\}$ *with*
$$\alpha_i := \varepsilon_i - \varepsilon_{i+1}.$$
The positive roots are the elements of $\Phi^+ = \{\varepsilon_i - \varepsilon_j : i < j\}$. *They have the representation*
$$\varepsilon_i - \varepsilon_j = \sum_{k=i}^{j-1} \alpha_k \in \Phi^+.$$

5. *For each positive root* $\alpha := \varepsilon_i - \varepsilon_j \in \Phi^+$ *the subalgebra*
$$S_\alpha \simeq sl(2,\mathbb{C})$$
is generated by the three elements
$$h_\alpha := E_{ii} - E_{jj},\ x_\alpha := E_{ij},\ y_\alpha := E_{ji}.$$

6. *The Cartan matrix of* Φ *is*

$$Cartan(\Delta) = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & \dots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix} \in M((r+1)\times(r+1),\mathbb{Z}).$$

All roots $\alpha \in \Delta$ have the same length. The only pairs of roots from Δ, which are not orthogonal, are

$$(\alpha_i, \alpha_{i+1}),\ i = 1, ..., r-1.$$

Each pair includes the angle $(2/3)\pi$. In particular the Dynkin diagram of the root system Φ from Figure 7.2 has type A_r from Theorem 6.31.

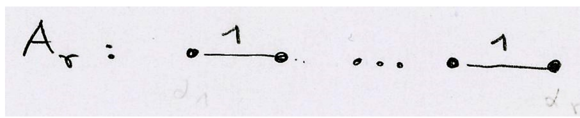

Fig. 7.2 Dynkin diagram of the root system of type A_r

Proof. 4) For $i \neq j$ elements $h \in H$ act on E_{ij} according to

$$[h, E_{ij}] = h \cdot E_{ij} - E_{ij} \cdot h = \varepsilon_i(h) \cdot E_{ij} - \varepsilon_j(h) \cdot E_{ij} = (\varepsilon_i(h) - \varepsilon_j(h)) \cdot E_{ij}$$

Due to the formula from Proposition 7.4, part 1 there are no further roots.

5) According to part 4) the commutators are

$$[h_\alpha, x_\alpha] = [h_\alpha, E_{ij}] = (\varepsilon_i(h_\alpha) - \varepsilon_j(h_\alpha))E_{ij} = 2 \cdot E_{ij} = 2 \cdot x_\alpha$$

$$[h_\alpha, y_\alpha] = [h_\alpha, E_{ji}] = (\varepsilon_j(h_\alpha) - \varepsilon_i(h_\alpha))E_{ji} = -2 \cdot E_{ji} = -2 \cdot y_\alpha$$

$$[x_\alpha, y_\alpha] = [E_{ij}, E_{ji}] = E_{ii} - E_{jj} = h_\alpha.$$

6) The Cartan matrix has the entries

$$< \beta, \alpha >= \beta(h_\alpha), \alpha, \beta \in \Delta.$$

We have to consider

- $\beta_k = \varepsilon_k - \varepsilon_{k+1}, 1 \leq k \leq r$ and
- $\alpha_j = \varepsilon_j - \varepsilon_{j+1}, 1 \leq j \leq r$, with corresponding elements $h_j = E_{jj} - E_{j+1,j+1} \in H$.

We compute

$$\beta_k(h_{\alpha_j}) = (\varepsilon_k - \varepsilon_{k+1})(E_{jj} - E_{j+1,j+1}) =$$

$$= \delta_{kj} - \delta_{k,j+1} - \delta_{k+1,j} + \delta_{k+1,j+1} = 2\delta_{kj} - \delta_{k,j+1} - \delta_{k+1,j} =$$

$$= \begin{cases} 2, & \text{if } k = j \\ -1, & \text{if } |k-j| = 1 \\ 0, & \text{if } |k-j| \geq 2 \end{cases}$$

The angle between each pair (α, β) of distinct, non-orthogonal roots in Δ is $(2/3)\pi$, and the symmetry of the Cartan matrix implies

$$1 = \frac{< \beta, \alpha, >}{< \alpha, \beta >} = \frac{\|\beta\|^2}{\|\alpha\|^2}$$

Hence all simple roots have the same length. The root system Φ has a connected Dynkin diagram, it has type A_r from Theorem 6.31. $\square$

In dealing with the Lie algebra $sp(2r, \mathbb{C})$ we note that $X \in sp(2r, \mathbb{C})$ iff

$$\sigma \cdot X^\top \cdot \sigma = X$$

with

$$\sigma := \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \text{ and } \sigma^{-1} = -\sigma.$$

This condition is equivalent to

$$X = \begin{pmatrix} A & B \\ C & -A^\top \end{pmatrix}$$

with symmetric matrices $B = B^\top$ and $C = C^\top$. The following proposition will refer to this type of decomposition.

Proposition 7.7 (Type C_r). *The Lie algebra*

$$L := sp(2r, \mathbb{C}), \ r \geq 3,$$

has the following characteristics:

1. *dim* $L = r(2r+1)$.
2. *The subalgebra*
$$H := \left\{ \begin{pmatrix} D & 0 \\ 0 & -D \end{pmatrix} \in L : D \in \mathfrak{d}(r, \mathbb{C}) \right\}$$
is a maximal toral subalgebra with dim $H = r$.
3. *The family*
$$(h_i := E_{ii} - E_{r+i,r+i})_{1 \leq i \leq r}$$
is a basis of H.

4. *Define the functionals*

$$\varepsilon_i := \hat{\varepsilon}_i | H \in H^*,\ i = 1,...,2r.$$

Then the root spaces and their corresponding roots are

$$\begin{cases} E_{ij} - E_{r+j,r+i}, 1 \le i \ne j \le r & \varepsilon_i - \varepsilon_j \\ E_{i,r+j} + E_{j,r+i}, 1 \le i \le j \le r & \varepsilon_i + \varepsilon_j \\ E_{r+i,j} + E_{r+j,i}, 1 \le i \le j \le r & -\varepsilon_i - \varepsilon_j \end{cases}$$

Grouped in a different way the root set Φ comprises the elements

$$\begin{cases} \varepsilon_i - \varepsilon_j : 1 \le i \ne j \le r & \textit{Type a} \\ \pm(\varepsilon_i + \varepsilon_j) : 1 \le i \ne j \le r & \textit{Type b} \\ \pm 2 \cdot \varepsilon_i : 1 \le i \le r & \textit{Type c} \end{cases}$$

A base of Φ is the set

$$\Delta := \{\alpha_1, ..., \alpha_r\}$$

with - note the different form of the last root α_r -

- $\alpha_j := \varepsilon_j - \varepsilon_{j+1}, 1 \le j \le r-1,$
- $\alpha_r = 2 \cdot \varepsilon_r$

The set Φ^+ of positive roots comprises the elements

$$\begin{cases} \varepsilon_i - \varepsilon_j : 1 \le i < j \le r & \textit{Type a} \\ \varepsilon_i + \varepsilon_j : 1 \le i < j \le r & \textit{Type b} \\ 2 \cdot \varepsilon_i : 1 \le i \le r & \textit{Type c} \end{cases}$$

5. *For each positive root $\alpha \in \Phi^+$ the subalgebra*

$$S_\alpha = < h_\alpha, x_\alpha, y_\alpha > \simeq sl(2,\mathbb{C})$$

has the generators:

- Type a*: If $\alpha := \varepsilon_i - \varepsilon_j \in \Phi^+, 1 \le i < j \le r$, then*

  $h_\alpha := (E_{ii} - E_{r+i,r+i}) - (E_{jj} - E_{r+j,r+j}),\ x_\alpha := E_{ij} - E_{r+j,r+i},\ y_\alpha := E_{ji} - E_{r+i,r+j}.$

- Type b*: If $\alpha := \varepsilon_i + \varepsilon_j \in \Phi^+, 1 \le i < j \le r$, then*

  $h_\alpha := (E_{ii} - E_{r+i,r+i}) + (E_{jj} - E_{r+j,r+j}),\ x_\alpha := E_{i,r+j} + E_{j,r+i},\ y_\alpha := E_{r+i,j} + E_{r+j,i}.$

- Type c*: If $\alpha := 2\varepsilon_i \in \Phi, 1 \le i \le r$, then*

$$h_\alpha := E_{ii} - E_{r+i,r+i},\ x_\alpha := E_{i,r+i},\ y_\alpha := E_{r+i,i}.$$

6. The Cartan matrix of the root system Φ referring to the basis Δ is

$$Cartan(\Delta) = \begin{pmatrix} 2 & -1 & 0 & \dots & & & 0 \\ -1 & 2 & -1 & \dots & & & 0 \\ & \ddots & \ddots & \ddots & & & \\ 0 & \dots & -1 & 2 & -1 & 0 \\ 0 & \dots & 0 & -1 & 2 & -1 \\ 0 & \dots & & 0 & -2 & 2 \end{pmatrix} \in M(r \times r, \mathbb{Z}),$$

entry $<\alpha_i, \alpha_j>$ at position $(row, column) = (i, j)$. Note the distinguished entry -2 in the last row: The Cartan matrix is not symmetric.
All roots $\alpha_j \in \Delta, 1 \le j \le r-1$ have equal length, they are the short roots. The root α_r is the long root:

$$\sqrt{2} = \frac{\|\alpha_r\|}{\|\alpha_j\|}, 1 \le j \le r-1.$$

The only pairs of simple roots with are not orthogonal are

$$(\alpha_i, \alpha_{i+1}), i = 1, ..., r-1.$$

For $i = 1, ..., r-2$ these pairs include the angle $\frac{2\pi}{3}$, while the pair (α_{r-1}, α_r) inludes the angle $\frac{3\pi}{4}$. In particular the Dynkin diagram of the root system Φ from Figure 7.3 has type C_r according to Theorem 6.31.

Fig. 7.3 Dynkin diagram of the root system of type C_r

Proof. 1) Using the representation

$$X = \begin{pmatrix} A & B \\ C & -A^{\top} \end{pmatrix} \in sp(2r, \mathbb{C})$$

with symmetric matrices $B = B^{\top}$ and $C = C^{\top}$ we obtain from the number of free parameters for A, B, C

$$dim\ L = r^2 + 2 \cdot \left(\frac{r^2 - r}{2} + r \right) = 2r^2 + r = r(2r+1).$$

4) Note for all $h \in H$

$$\varepsilon_i(h) = -\varepsilon_{r+i}(h).$$

For $h \in H$ the commutators are:

$$\begin{aligned}
[h, E_{ij} - E_{r+j,r+i}] &= h \cdot (E_{ij} - E_{r+j,r+i}) - (E_{ij} - E_{r+j,r+i}) \cdot h = \\
&= \varepsilon_i(h)E_{ij} - \varepsilon_{r+j}(h)E_{r+j,r+i} - \varepsilon_j(h)E_{ij} + \varepsilon_{r+i}(h)E_{r+j,r+i} = \\
&= (\varepsilon_i(h) - \varepsilon_j(h))E_{ij} - (\varepsilon_{r+j}(h) - \varepsilon_{r+i}(h))E_{r+j,r+i} = \\
&= (\varepsilon_i(h) - \varepsilon_j(h))E_{ij} + (\varepsilon_j(h) - \varepsilon_i(h))E_{r+j,r+i} = \\
&(\varepsilon_i(h) - \varepsilon_j(h))(E_{ij} - E_{r+j,r+i})
\end{aligned}$$

$$[h, E_{i,r+j} + E_{j,r+i}] = \varepsilon_i(h)E_{i,r+j} + \varepsilon_j(h)E_{j,r+i} - \varepsilon_{r+j}(h)E_{i,r+j} - \varepsilon_{r+i}(h)E_{j,r+i} =$$

$$\begin{aligned}
&(\varepsilon_i(h) - \varepsilon_{r+j}(h))E_{i,r+j} + (\varepsilon_j(h) - \varepsilon_{r+i}(h))E_{j,r+i} = \\
&= (\varepsilon_i(h) + \varepsilon_j(h))E_{i,r+j} + (\varepsilon_j(h) + \varepsilon_i(h))E_{j,r+i} = \\
&(\varepsilon_i(h) + \varepsilon_j(h))(E_{i,r+j} + E_{j,r+i})
\end{aligned}$$

$$[h, E_{r+i,j} + E_{r+j,i}] = \varepsilon_{r+i}(h)E_{r+i,j} + \varepsilon_{r+j}(h)E_{r+j,i} - \varepsilon_j(h)E_{r+i,j} - \varepsilon_i(h)E_{r+j,i} =$$

$$\begin{aligned}
&= (\varepsilon_{r+i}(h) - \varepsilon_j(h))E_{r+i,j} + (\varepsilon_{r+j}(h) - \varepsilon_i(h))E_{r+j,i} = \\
&= (-\varepsilon_i(h) - \varepsilon_j(h))E_{r+i,j} + (-\varepsilon_j(h) - \varepsilon_i(h))E_{r+j,i} = \\
&= (-\varepsilon_i(h) - \varepsilon_j(h))(E_{r+i,j} + E_{r+j,i})
\end{aligned}$$

The positive roots have the base representation

$$\varepsilon_i - \varepsilon_j = \sum_{k=i}^{j-1} \alpha_k, 1 \le i < j \le r.$$

$$\varepsilon_i + \varepsilon_j = 2 \cdot \varepsilon_r + \sum_{k=i}^{j-1} \alpha_k + 2 \cdot \sum_{k=j}^{r-1} \alpha_k, 1 \le i < j \le r.$$

$$2 \cdot \varepsilon_i = \alpha_r + \sum_{k=i}^{r-1} 2 \cdot \alpha_k, 1 \le i \le r-1.$$

Part 6) The Cartan matrix has the entries $\beta(h_\alpha), \alpha, \beta \in \Delta$. We have to consider the roots

- $\beta_k = \varepsilon_k - \varepsilon_{k+1}, 1 \le k \le r-1$, and
- $\beta_r = 2\varepsilon_r$

and the roots

- $\alpha_j = \varepsilon_j - \varepsilon_{j+1}, 1 \le j \le r-1$, with elements $h_{\alpha_j} = h_j - h_{j+1}$ and
- $\alpha_r = 2\varepsilon_r$ with element $h_{\alpha_r} = h_r$.

Accordingly, we calculate the cases:

- For $1 \le j,k \le r-1$:

$$\beta_k(h_{\alpha_j}) = (\varepsilon_k - \varepsilon_{k+1})(E_{jj} - E_{j+1,j+1}) = \delta_{k,j} - \delta_{k,j+1} - \delta_{k+1,j} + \delta_{k+1,j+1} =$$
$$= 2 \cdot \delta_{k,j} - \delta_{k,j+1} - \delta_{k,j-1}$$

- For $1 \le j \le r-1, k = r$:

$$\beta_r(h_{\alpha_j}) = 2 \cdot \varepsilon_r(E_{jj} - E_{j+1,j+1}) = 2(\delta_{rj} - \delta_{r,j+1}) = -2\delta_{r,j+1}$$

- For $1 \le k \le r-1, j = r$:

$$\beta_k(h_{\alpha_r}) = (\varepsilon_k - \varepsilon_{k+1})(E_{rr}) = \delta_{kr} - \delta_{k+1,r} = -\delta_{k+1,r}$$

- For $j = k = r$:

$$\beta_r(h_{\alpha_r}) = 2 \cdot \varepsilon_r(E_{rr}) = 2.$$

If $\alpha \ne \beta$ and $< \alpha, \beta >, < \beta, \alpha > \ne 0$ then

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

Hence

$$2 = \frac{-2}{-1} = \frac{\alpha_r(h_{\alpha_{r-1}})}{\alpha_{r-1}(h_{\alpha_r})} = \frac{\|\alpha_r\|^2}{\|\alpha_{r-1}\|^2}.$$

□

In dealing with the Lie algebra

$$so(2r, \mathbb{C})$$

it is useful to consider a matrix $M \in so(2r, \mathbb{C})$ as a scheme having 2×2-matrices as entries: We introduce the non-Abelian $\mathbb{C}$-algebra

$$R := M(2 \times 2, \mathbb{C})$$

as ring of coefficients and consider matrices

$$M = ((a_{jk})_{1 \le j,k \le r}) \in M(r \times r, R) \simeq M(2r \times 2r, \mathbb{C})$$

as matrices with entries from R

$$a_{jk} \in R, 1 \leq j,k \leq r.$$

For each $1 \leq j,k \leq r$ we introduce the matrix

$$\tilde{E}_{jk} \in M(r \times r, R)$$

with only one nonzero entry, namely $\mathbb{1} \in R$, at place (j,k). We distinguish the following elements from R, which can be expressed by using the Pauli matrices, see Remark 2.21, extended by

$$\sigma_0 := \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} :$$

$$h := \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$s := \frac{1}{2}\begin{pmatrix} i & -1 \\ -1 & -i \end{pmatrix} = (1/2) \cdot (i\sigma_3 - \sigma_1),\ t := \frac{1}{2}\begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}) = (1/2) \cdot (i\sigma_3 + \sigma_1)$$

$$u := \frac{1}{2}\begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} = (1/2) \cdot i(\sigma_0 + \sigma_2)\ \mathrm{v} := \frac{1}{2}\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} = (1/2) \cdot i(\sigma_0 - \sigma_2)$$

satisfying

- $s = s^\top,\ h \cdot s = s, s \cdot h = -s$
- $t = t^\top,\ h \cdot t = -t, t \cdot h = t$
- $u^\top = v,\ h \cdot u = u \cdot h = u$
- $\mathrm{v}^\top = u,\ h \cdot \mathrm{v} = \mathrm{v} \cdot h = -\mathrm{v}$
- $[s,t] = -h$
- $u^2 - \mathrm{v}\mathrm{v}^2 = -h$

E.g. the matrix $h \cdot \tilde{E}_{jj} \in M(r \times r, R)$ is the block matrix with the single block $h \in R$ at the diagonal place with index (j,j). The block is

$$h = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \in M(2 \times 2, \mathbb{C})$$

which implies

$$h \cdot \tilde{E}_{jj} = i \cdot (E_{2j+1,2j} - E_{2j,2j+1}) \in M(2r \times 2r, \mathbb{C}).$$

Proposition 7.8 (Type D_r). *The Lie algebra*

$$L := so(2r,\mathbb{C}),\ r \geq 4,$$

has the following characteristics:

1. *dim* $L = r(2r-1)$.

2. *The subalgebra*

$$H := span_{\mathbb{C}} < h\cdot\tilde{E}_{jj} : 1 \leq j \leq r > \subset (\mathfrak{d}(r,R) \cap L)$$

   is a maximal toral subalgebra with dim $H = r$.

3. *For any pair* $1 \leq j < k \leq r$ *each of the four elements*

$$s\cdot(\tilde{E}_{jk} - \tilde{E}_{kj}),\ t\cdot(\tilde{E}_{jk} - \tilde{E}_{kj}),\ u\cdot\tilde{E}_{jk} - v\cdot\tilde{E}_{kj},\ v\cdot\tilde{E}_{jk} - u\cdot\tilde{E}_{kj}$$

   generates a 1*-dimensional root space, belonging to the respective root*

$$\alpha = \begin{cases} \varepsilon_j + \varepsilon_k \\ -\varepsilon_j - \varepsilon_k \\ \varepsilon_j - \varepsilon_k \\ -\varepsilon_j + \varepsilon_k \end{cases}$$

   Here

$$\varepsilon_j : H \to \mathbb{C}$$

   are the $\mathbb{C}$*-linear functionals which are dual to the family*

$$h\cdot\tilde{E}_{jj},\ 1 \leq j \leq r,$$

   i.e.

$$\varepsilon_i(h\cdot\tilde{E}_{jj}) = \delta_{ij}.$$

4. *The root set* Φ *of* L *has the elements*

$$-\varepsilon_j - \varepsilon_k,\ \varepsilon_j + \varepsilon_k,\ -\varepsilon_j + \varepsilon_k,\ \varepsilon_j - \varepsilon_k, 1 \leq j < k \leq r,$$

   for short

$$\Phi = \{\pm\varepsilon_j \pm \varepsilon_k :\ 1 \leq j < k \leq r\}\ (each\ combination\ of\ signs).$$

   A base of Φ *is the set*

$$\Delta := \{\alpha_1, ..., \alpha_r\}$$

   with - note the different form of the last root α_r *-*

   - $\alpha_j := \varepsilon_j - \varepsilon_{j+1}, 1 \leq j \leq r-1,$

- $\alpha_r = \varepsilon_{r-1} + \varepsilon_r$

The set of positive roots is $\Phi^+ = \{\varepsilon_j \pm \varepsilon_k : 1 \leq j < k \leq r\}$.

5. *For each positive root* $\alpha := \varepsilon_j + \varepsilon_k \in \Phi^+, 1 \leq j < k \leq r$, *the subalgebra*

$$S_\alpha \simeq sl(2,\mathbb{C})$$

is generated by the three elements

$$h_\alpha := h \cdot (\tilde{E}_{jj} + \tilde{E}_{kk}),\ x_\alpha := s \cdot (\tilde{E}_{jk} - \tilde{E}_{kj}),\ y_\alpha := t \cdot (\tilde{E}_{jk} - \tilde{E}_{kj}).$$

For each positive root $\alpha := \varepsilon_j - \varepsilon_k \in \Phi^+, 1 \leq j < k \leq r$, *the subalgebra*

$$S_\alpha \simeq sl(2,\mathbb{C})$$

is generated by the three elements

$$h_\alpha := h \cdot (\tilde{E}_{jj} - \tilde{E}_{kk}),\ x_\alpha := u \cdot \tilde{E}_{jk} - v \cdot \tilde{E}_{kj},\ y_\alpha := v \cdot \tilde{E}_{jk} - u \cdot \tilde{E}_{kj}.$$

6. *The Cartan matrix of the root system* Φ *referring to the basis* Δ *is*

$$Cartan(\Delta) = \begin{pmatrix} 2 & -1 & 0 & \dots & & 0 \\ -1 & 2 & -1 & \dots & & 0 \\ & \ddots & \ddots & \ddots & & \\ 0 & \dots & -1 & 2 & -1 & -1 \\ 0 & \dots & 0 & -1 & 2 & 0 \\ 0 & & & -1 & 0 & 2 \end{pmatrix} \in M(r \times r, \mathbb{Z}).$$

Note the distinguished Cartan integers

$$< \alpha_{r-2}, \alpha_r > = < \alpha_r, \alpha_{r-2} >$$

All roots $\alpha \in \Delta$ *have the same length. The only pairs of simple roots with are not orthogonal are*

$$(\alpha_i, \alpha_{i+1}), i = 1, ..., r-2, \text{ and } (\alpha_{r-2}, \alpha_r).$$

These pairs include the angle $(2/3)\pi$. *The Dynkin diagram of the root system* Φ *from Figure 7.4 has type* D_r *from Theorem 6.31.*

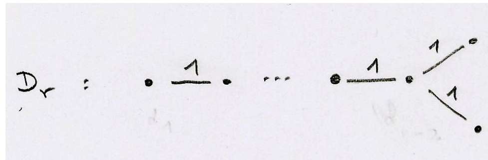

Fig. 7.4 The Dynkin diagram of the root system of type D_r

Proof. 3) The general element of H has the form

$$z=\sum_{\nu=1}^{r} a_\nu\cdot(h\cdot\tilde{E}_{\nu\nu}),\ a_\nu\in\mathbb{C}.$$

The element z acts on $s\cdot(\tilde{E}_{jk}-\tilde{E}_{kj})$ according to

$$[z,s\cdot(\tilde{E}_{jk}-\tilde{E}_{kj})]=\varepsilon_j(z)\cdot hs\cdot\tilde{E}_{jk}-\varepsilon_k(z)\cdot hs\cdot\tilde{E}_{kj}-\varepsilon_k(z)\cdot sh\cdot\tilde{E}_{jk}+\varepsilon_j(z)\cdot sh\cdot\tilde{E}_{kj}=$$

$$\varepsilon_j(z)\cdot s\cdot\tilde{E}_{jk}-\varepsilon_k(z)\cdot s\cdot\tilde{E}_{kj}+\varepsilon_k(z)\cdot s\cdot\tilde{E}_{jk}-\varepsilon_j(z)\cdot s\cdot\tilde{E}_{kj}=$$
$$=(\varepsilon_j(z)+\varepsilon_k(z))\cdot s\cdot(\tilde{E}_{jk}-\tilde{E}_{kj})$$

The element z acts on $t\cdot(\tilde{E}_{jk}-\tilde{E}_{kj})$ according to

$$[z,t\cdot(\tilde{E}_{jk}-\tilde{E}_{kj})]=\varepsilon_j(z)\cdot ht\cdot\tilde{E}_{jk}-\varepsilon_k(z)\cdot ht\cdot\tilde{E}_{kj}-\varepsilon_k(z)\cdot th\cdot\tilde{E}_{jk}+\varepsilon_j(z)\cdot th\cdot\tilde{E}_{kj}=$$

$$-\varepsilon_j(z)\cdot t\cdot\tilde{E}_{jk}+\varepsilon_k(z)\cdot t\cdot\tilde{E}_{kj}-\varepsilon_k(z)\cdot t\cdot\tilde{E}_{jk}+\varepsilon_j(z)\cdot t\cdot\tilde{E}_{kj}=$$
$$=(-\varepsilon_j(z)-\varepsilon_k(z))\cdot t\cdot(\tilde{E}_{jk}-\tilde{E}_{kj})$$

The element z acts on $u\cdot\tilde{E}_{jk}-\mathrm{v}\cdot\tilde{E}_{kj}$ according to

$$[z,u\cdot\tilde{E}_{jk}-\mathrm{v}\cdot\tilde{E}_{kj}]=\varepsilon_j(z)\cdot hu\cdot\tilde{E}_{jk}-\varepsilon_k(z)\cdot h\mathrm{v}\cdot\tilde{E}_{kj}-\varepsilon_k(z)\cdot uh\cdot\tilde{E}_{jk}+\varepsilon_j(z)\cdot \mathrm{v}h\cdot\tilde{E}_{kj}=$$

$$\varepsilon_j(z)\cdot u\cdot\tilde{E}_{jk}+\varepsilon_k(z)\cdot\mathrm{v}\cdot\tilde{E}_{kj}-\varepsilon_k(z)\cdot u\cdot\tilde{E}_{jk}-\varepsilon_j(z)\cdot\mathrm{v}\cdot\tilde{E}_{kj}=$$
$$=(\varepsilon_j(z)-\varepsilon_k(z))\cdot(u\cdot\tilde{E}_{jk}-\mathrm{v}\cdot\tilde{E}_{kj})$$

The element z acts on $\mathrm{v}\cdot\tilde{E}_{jk}-u\cdot\tilde{E}_{kj}$ according to

$$[z,\mathrm{v}\cdot\tilde{E}_{jk}-u\cdot\tilde{E}_{kj}]=\varepsilon_j(z)\cdot h\mathrm{v}\cdot\tilde{E}_{jk}-\varepsilon_k(z)\cdot hu\cdot\tilde{E}_{kj}-\varepsilon_k(z)\cdot \mathrm{v}h\cdot\tilde{E}_{jk}+\varepsilon_j(z)\cdot uh\cdot\tilde{E}_{kj}=$$

$$-\varepsilon_j(z)\cdot\mathrm{v}\cdot\tilde{E}_{jk}-\varepsilon_k(z)\cdot u\cdot\tilde{E}_{kj}+\varepsilon_k(z)\cdot\mathrm{v}\cdot\tilde{E}_{jk}+\varepsilon_j(z)\cdot u\cdot\tilde{E}_{kj}=$$
$$=(-\varepsilon_j(z)+\varepsilon_k(z))\cdot(\mathrm{v}\cdot\tilde{E}_{jk}-u\cdot\tilde{E}_{kj}).$$

We have

$$|\Phi|=4\cdot\frac{r(r-1)}{2}=2r\cdot(r-1)=r\cdot(2r-1)-r=dim\,L-dim\,H$$

in accordance with the formula from Proposition 7.4, part 1.

4) The positive roots have the base representation

$$\varepsilon_i - \varepsilon_j = \sum_{k=i}^{j-1} \alpha_k, 1 \le i < j \le r,$$

$$\varepsilon_i + \varepsilon_j = \sum_{k=i}^{r-2} \alpha_k + \sum_{k=j}^{r} \alpha_k, 1 \le i < j \le r.$$

5) For $\alpha := \varepsilon_j + \varepsilon_k \in \Phi^+, 1 \le j < k \le r$, the commutators are

$$[h_\alpha, x_\alpha] = [h \cdot (\tilde{E}_{jj} + \tilde{E}_{kk}), s \cdot (\tilde{E}_{jk} - \tilde{E}_{kj})] =$$

$$= hs \cdot (\tilde{E}_{jj} + \tilde{E}_{kk})(\tilde{E}_{jk} - \tilde{E}_{kj}) - sh \cdot (\tilde{E}_{jk} - \tilde{E}_{kj})(\tilde{E}_{jj} + \tilde{E}_{kk}) =$$

$$= s \cdot (\tilde{E}_{jk} - \tilde{E}_{kj}) + s \cdot (\tilde{E}_{jk} - \tilde{E}_{kj}) = 2 \cdot x_\alpha$$

$$[h_\alpha, y_\alpha] = [h \cdot (\tilde{E}_{jj} + \tilde{E}_{kk}), t \cdot (\tilde{E}_{jk} - \tilde{E}_{kj})] = -t \cdot (\tilde{E}_{jk} - \tilde{E}_{kj}) - t \cdot (\tilde{E}_{jk} - \tilde{E}_{kj}) = -2 \cdot y_\alpha$$

$$[x_\alpha, y_\alpha] = [s \cdot (\tilde{E}_{jk} - \tilde{E}_{kj}), t(\tilde{E}_{jk} - \tilde{E}_{kj})] = [s,t] \cdot (\tilde{E}_{jk} - \tilde{E}_{kj})^2 =$$

$$= (-h) \cdot (-\tilde{E}_{jj} - \tilde{E}_{kk}) = h_\alpha$$

For $\alpha := \varepsilon_j - \varepsilon_k \in \Phi^+, 1 \le j < k \le r$, the commutators are

$$[h_\alpha, x_\alpha] = [h \cdot (\tilde{E}_{jj} - \tilde{E}_{kk},\ u \cdot \tilde{E}_{jk} - \mathrm{v} \cdot \tilde{E}_{kj}] = hu \cdot \tilde{E}_{jk} + h\mathrm{v} \cdot \tilde{E}_{kj} + uh \cdot \tilde{E}_{jk} + \mathrm{v}h \cdot \tilde{E}_{kj} =$$

$$u \cdot \tilde{E}_{jk} - \mathrm{v} \cdot \tilde{E}_{kj} + u \cdot \tilde{E}_{jk} - \mathrm{v} \cdot \tilde{E}_{kj} = 2u \cdot \tilde{E}_{jk} - 2\mathrm{v} \cdot \tilde{E}_{kj} = 2 \cdot x_\alpha$$

$$[h_\alpha, y_\alpha] = [h \cdot (\tilde{E}_{jj} - \tilde{E}_{kk}), \mathrm{v} \cdot \tilde{E}_{jk} - u \cdot \tilde{E}_{kj}] = h\mathrm{v} \cdot \tilde{E}_{jk} + hu \cdot \tilde{E}_{kj} + \mathrm{v}h \cdot \tilde{E}_{jk} + uh \cdot \tilde{E}_{kj} =$$

$$-\mathrm{v} \cdot \tilde{E}_{jk} + u \cdot \tilde{E}_{kj} - \mathrm{v} \cdot \tilde{E}_{jk} + u \cdot \tilde{E}_{kj} = 2u \cdot \tilde{E}_{kj} - 2\mathrm{v} \cdot \tilde{E}_{jk} = -2 \cdot y_\alpha$$

$$[x_\alpha, y_\alpha] = [u \cdot \tilde{E}_{jk} - \mathrm{v} \cdot \tilde{E}_{kj}, \mathrm{v} \cdot \tilde{E}_{jk} - u \cdot \tilde{E}_{kj}] =$$

$$= -u^2 \tilde{E}_{jj} - \mathrm{v}^2 \tilde{E}_{kk} + \mathrm{v}^2 \tilde{E}_{jj} + u^2 \tilde{E}_{kk} =$$

$$= (\mathrm{v}^2 - u^2) \cdot \tilde{E}_{jj} + (u^2 - \mathrm{v}^2) \cdot \tilde{E}_{kk} = h \cdot (\tilde{E}_{jj} - \tilde{E}_{kk}) = h_\alpha$$

6) The Cartan matrix has the entries

$$< \beta, \alpha >= \beta(h_\alpha), \alpha, \beta \in \Delta.$$

We have to consider

- $\alpha_j = \varepsilon_j - \varepsilon_{j+1}, 1 \le j \le r-1$ with elements $h_{\alpha_j} = h \cdot (\tilde{E}_{jj} - \tilde{E}_{j+1,j+1})$ and
- $\alpha_r = \varepsilon_{r-1} + \varepsilon_r$ with element $h_{\alpha_r} = h \cdot (\tilde{E}_{r-1,r-1} + \tilde{E}_{r,r})$

and

- $\beta_k = \varepsilon_k - \varepsilon_{k+1}, 1 \le k \le r-1$ and

- $\beta_r = \varepsilon_{r-1} + \varepsilon_r$.

If $1 \leq j,k \leq r-1$ then

$$(\varepsilon_k - \varepsilon_{k+1})(h\cdot(\tilde{E}_{jj} - \tilde{E}_{j+1,j+1})) = \delta_{kj} - \delta_{k,j+1} - \delta_{k+1,j} + \delta_{k+1,j+1} = 2\delta_{kj} - \delta_{k,j+1} - \delta_{k+1,j} =$$

$$= \begin{cases} 2, & \text{if } j = k \\ -1, & \text{if } |j-k| = 1 \\ 0, & \text{if } |j-k| \geq 2 \end{cases}$$

If $k = r$ and $1 \leq j \leq r-1$ then

$$(\varepsilon_{r-1} + \varepsilon_r)(h\cdot(E_{jj} - E_{j+1,j+1})) = \delta_{r-1,j} - \delta_{r-1,j+1} + \delta_{r,j} - \delta_{r,j+1} = -\delta_{r,j+2}$$

$$= \begin{cases} -1, & \text{if } j = r-2 \\ 0, & \text{if } j \neq r-2 \end{cases}$$

If $1 \leq k \leq r-1$ and $j = r$ then

$$(\varepsilon_k - \varepsilon_{k+1})(h\cdot(E_{r-1,r-1} + E_{r,r})) = \delta_{k,r-1} + \delta_{k,r} - \delta_{k+1,r-1} - \delta_{k+1,r} = -\delta_{k,r-2}$$

$$= \begin{cases} -1, & \text{if } k = r-2 \\ 0, & \text{if } k \neq r-2 \end{cases}$$

If $k = r = 2$ then

$$(\varepsilon_{r-1} + \varepsilon_r)(h\cdot(\tilde{E}_{r-1,r-1} + \tilde{E}_{r,r})) = \delta_{r-1,r-1} + \delta_{r,r} = 2.$$

If $\alpha \neq \beta$ and $< \alpha, \beta >, < \beta, \alpha > \neq 0$ then

$$\frac{< \alpha, \beta >}{< \beta, \alpha >} = \frac{\|\beta\|^2}{\|\alpha\|^2} = 1.$$

We read off

$$< \alpha_{r-2}, \alpha_{r-1} >< \alpha_{r-1}, \alpha_{r-2} > = < \alpha_{r-2}, \alpha_r >< \alpha_r, \alpha_{r-2} > =$$

$$= (-1)\cdot(-1) = +1$$

and

$$< \alpha_{r-1}, \alpha_r >< \alpha_r, \alpha_{r-1} > = 0 \cdot 0 = 0$$

As a consequence, there is a single edge between the vertices of the pair

$$\alpha_{r-2} \text{ and } \alpha_{r-1}$$

and between the vertices of the pair

$$\alpha_{r-2} \text{ and } \alpha_r,$$

but no edge between the vertices of the pair

$$\alpha_{r-1} \text{ and } \alpha_r.$$

□

In order to employ the result of Proposition 7.8 for the investigation of the Lie algebra

$$so(2r+1,\mathbb{C})$$

we consider the matrices from $M(r \times r, R)$, used in Proposition 7.8, as matrices from $M(2r \times 2r, \mathbb{C})$, and embedd them via the canonical embedding

$$M(2r \times 2r, \mathbb{C}) \to M((2r+1) \times (2r+1), \mathbb{C})$$

as block matrices

$$A \mapsto \hat{A} := \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \in M((2r+1) \times (2r+1), \mathbb{C}).$$

Proposition 7.9 (Type B_r). *The Lie algebra*

$$L := so(2r+1,\mathbb{C}), r \geq 2,$$

has the following characteristics:

1. *dim* $L = r(2r+1)$.

2. *The subalgebra*

$$H := span_{\mathbb{C}} < h \cdot \hat{E}_{jj} : 1 \leq j \leq r > \subset (\mathfrak{d}(2r+1,\mathbb{C}) \cap L)$$

   is a maximal toral subalgebra with dim $H = r$.

3. *For* $1 \leq j \leq r$ *denote by*

$$\varepsilon_j := (h \cdot \hat{E}_{jj})^* \subset H^*$$

   the dual functionals. For each pair $1 \leq j < k \leq r$ *each of the four elements*

$$s \cdot (\hat{E}_{jk} - \hat{E}_{kj}),\ t \cdot (\hat{E}_{jk} - \hat{E}_{kj}),\ u \cdot \hat{E}_{jk} - v \cdot \hat{E}_{kj},\ v \cdot \hat{E}_{jk} - u \cdot \hat{E}_{kj}$$

   generates the 1*-dimensional root space belonging to the respective root*

$$\alpha = \begin{cases} \varepsilon_j + \varepsilon_k \\ -\varepsilon_j - \varepsilon_k \\ \varepsilon_j - \varepsilon_k \\ -\varepsilon_j + \varepsilon_k \end{cases}$$

In addition, for each index $j = 1, ..., r$

- *a* 1*-dimensional root space is generated by the matrix*

$$X_j \in so(2r+1, \mathbb{C})$$

with exactly four non-zero entries: The vector

$$B_1 := \begin{pmatrix} 1 \\ -i \end{pmatrix} \in M(2 \times 1, \mathbb{C})$$

at places $(2j-1, 2r+1)$ and $(2j, 2r+1)$ and the vector $-B_1^\top$ at places $(2r+1, 2j-1)$ and $(2r+1, 2j)$.

- *and a* 1*-dimensional root space is generated by the matrix*

$$Y_j \in so(2r+1, \mathbb{C})$$

with exactly four non-zero entries: The vector

$$B_2 := \begin{pmatrix} 1 \\ i \end{pmatrix} \in M(2 \times 1, \mathbb{C})$$

at places $(2j-1, 2r+1)$ and $(2j, 2r+1)$ and the vector $-B_2^\top$ at places $(2r+1, 2j-1)$ and $(2r+1, 2j)$.

For $j = 1, ..., r$ the respective roots belonging to X_j and Y_j are

$$\alpha = \pm \varepsilon_j.$$

4. *The root system Φ of L is*

$$\Phi = \{\pm\varepsilon_k \pm \varepsilon_n : 1 \le k < n \le r\} \cup \{\pm\varepsilon_j : j = 1, ..., r\}.$$

A base of ϕ is the set

$$\Delta = \{\alpha_1, ..., \alpha_r\}$$

with - note the different form of the last root -

- $\alpha_j := \varepsilon_j - \varepsilon_{j+1}, 1 \le j \le r-1,$
- $\alpha_r := \varepsilon_r.$

The set of positive roots is

$$\Phi^+ = \{\varepsilon_i \pm \varepsilon_j : 1 \le i < j \le r\} \cup \{\varepsilon_j : 1 \le j \le r\}.$$

5. *For a positive root* $\alpha \in \Phi^+$ *the subalgebra*

$$S_\alpha =< h_\alpha,\ x_\alpha,\ y_\alpha > \simeq sl(2,\mathbb{C})$$

has the generators:

- *If* $\alpha = \varepsilon_j + \varepsilon_k,\ 1 \le j < k \le r$, *then*

$$h_\alpha := h \cdot (\hat{E}_{jj} + \hat{E}_{kk}),\ x_\alpha := s \cdot (\hat{E}_{jk} - \hat{E}_{kj}),\ y_\alpha := t \cdot (\hat{E}_{jk} - \hat{E}_{kj}).$$

- *If* $\alpha = \varepsilon_j - \varepsilon_k, 1 \le j < k \le r$, *then*

$$h_\alpha := h \cdot (\hat{E}_{jj} - \hat{E}_{kk}),\ x_\alpha := u \cdot \hat{E}_{jk} - v \cdot \hat{E}_{kj},\ y_\alpha := v \cdot \hat{E}_{jk} - u \cdot \hat{E}_{kj}.$$

- *If* $\alpha = \varepsilon_j, 1 \le j \le r$, *then*

$$h_\alpha := 2h \cdot \hat{E}_{jj},\ x_\alpha = X_j,\ y_\alpha = Y_j.$$

6. *The Cartan matrix of the root system* Φ *referring to the base* Δ *is*

$$Cartan(\Delta) = \begin{pmatrix} 2 & -1 & 0 & \dots & & & 0 \\ -1 & 2 & -1 & \dots & & & 0 \\ & \ddots & \ddots & \ddots & & & 0 \\ & & & \dots & & & 0 \\ & & & \dots & & & 0 \\ 0 & \dots & & 0 & -1 & 2 & -2 \\ 0 & & & & 0 & -1 & 2 \end{pmatrix} \in M(r \times r, \mathbb{Z}).$$

Note the distinguished entries in the last row and the last column.
The roots $\alpha_j, j = 1, ..., r-1$, *have equal length; they are the long roots. The root* α_r *is the single short root. The length ratio is*

$$\frac{\|\alpha_j\|}{\|\alpha_r\|} = \sqrt{2}, j = 1, ..., r-1.$$

The non-orthogonal pairs of roots from Δ *are*

$$(\alpha_j, \alpha_{j+1}), j = 1, ..., r-1.$$

Each pair $(\alpha_j, \alpha_{j+1}), j = 1, ..., r-2$, *encloses the angle* $\dfrac{2\pi}{3}$, *while the last pair* (α_{r-1}, α_r) *encloses the angle* $\dfrac{3\pi}{4}$. *The Dynkin diagram of* Φ *from Figure 7.5 has type* B_r *from Theorem 6.31.*

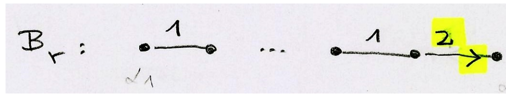

Fig. 7.5 The Dynkin diagram of the root system of type B_r

Proof. Most part of the proof follows from the corresponding statements in Proposition 7.8. In addition:

3) The general element of H has the form

$$z = \sum_{\nu=1}^{r} a_\nu \cdot h \cdot \hat{E}_{\nu\nu}, a_\nu \in \mathbb{C}.$$

In addition to the action of H on the root space elements from Proposition 7.8 one has the action on the additional elements X_j and Y_j: For $1 \le j \le r$ the element $z \in H$ acts according to

$$[z, X_j] = -\varepsilon_j(z) \cdot X_j, \ [z, Y_j] = \varepsilon_j(z) \cdot Y_j.$$

We have

$$|\Phi| = 4 \cdot \frac{r(r-1)}{2} + 2r = 2r^2 = r \cdot (2r+1) - r = dim\, L - dim\, H$$

in accordance with the formula from Proposition 7.4, part 1.4) The positive roots

have the base representation:

$$\varepsilon_j = \sum_{k=j}^{r} \alpha_k, 1 \le j \le r,$$

$$\varepsilon_i - \varepsilon_j = \sum_{k=i}^{j-1} \alpha_k, 1 \le i < j \le r.$$

5) Note:

$$B_1 \cdot B_2^\top - B_2 \cdot {B_1}^\top = 2 \cdot h \in M(2 \times 2, \mathbb{C}).$$

6) The computation of the Cartan matrix is similar to the computation in the proof of Proposition 7.8.

□

Theorem 7.10 (The classical complex Lie algebras are simple). *The Lie algebras of the complex classical matrix groups of the types*

$$A_r, r \ge 1; B_r, r \ge 2; C_r, r \ge 3; D_r, r \ge 4,$$

are simple.

Proof. We know from Corollary 4.18 that each classical Lie algebra L is semisimple. According to Theorem 4.21 the semisimple Lie algebra L splits into a direct sum of simple Lie algebras:

$$L = \bigoplus_{j=1}^{m} L_j$$

with simple Lie algebras $L_j,\ j = 1, ..., m$. The direct sum of maximal toral subalgebras

$$T_j \subset L_j,\ j = 1, ..., m$$

is a maximal toral subalgebra of L. For each pair $i \neq j$ and each pair of roots α_i of L_i and $\alpha_j \in L_j$ the corresponding Cartan integers vanish

$$< \alpha_i, \alpha_j >= 0.$$

Hence the Coxeter graph and a posteriori the Dynkin diagram has m connected components, which implies $m = 1$. $\square$

Remark 7.11 (Real simple Lie algebras). Also the Lie algebras of the real classical groups belonging to types

$$A_r,\ r \geq 1;\ B_r,\ r \geq 2;\ C_r,\ r \geq 3;\ D_r,\ r \geq 4;$$

are simple.

7.3 Review and outlook

Remark 7.12 (Classifying complex semisimple Lie algebras by Dynkin diagrams). Denote by

$$\mathscr{L},\ \mathscr{R},\ \mathscr{D}$$

respectively the set of isomorphism classes of complex semisimple Lie algebras, the set of roots systems, and the set of Dynkin diagrams. Then exist maps

$$F_{\mathscr{L}} : \mathscr{L} \to \mathscr{R} \text{ and } F_{\mathscr{R}} : \mathscr{R} \to \mathscr{D}$$

and

$$F := F_{\mathscr{R}} \circ F_{\mathscr{L}} : \mathscr{L} \to \mathscr{D},$$

defined as follows and satisfying the following properties:

1. Theorem 7.5 constructs the map $F_{\mathscr{L}}$.

2. The map $F_{\mathscr{L}}$ does not depend on the choice of a maximal toral subalgebra. For a proof see the reference in Remark 5.21.

3. The map $F_{\mathscr{L}}$ is bijective. The proof follows from a theorem of Serre, see [24, Sect. 18.4, Theor.]. For a simple Lie algebra L the root system $F_{\mathscr{L}}(L)$ is irreducible, see [24, Chap. 14.1, Prop.].

4. Definition 6.24 and 6.30 define the map $F_{\mathscr{R}}$.

5. Theorem 6.23 implies: The map $F_{\mathscr{R}}$ is injective.

6. The map $F_{\mathscr{R}}$ is surjective. For a proof see [24, Sect. 12.1, Theor.].

7. Propositions 7.6 - 7.8 show: At least the Dynkin diagrams of type A, B, C, D from Theorem 6.31 are contained in the image of F.

Part 1 - 6 imply: The map

$$F : \mathscr{L} \to \mathscr{D}$$

is bijective, i.e. the isomorphism classes of complex semisimple Lie algebras correspond bijectively to the Dynkin diagrams from Theorem 6.31.

For further topics in a more general context see Figure 0.1.

List of results

Chapter 1. Matrix functions

Operator norm (Def. 1.2)

Power series of matrices (Lem. 1.5)

Eigenspace and generalized eigenspace (Def. 1.9)

Jordan decomposition (Theor. 1.19)

Cayleigh-Hamilton theorem (Theor. 1.20)

Exponential of matrices (Def. 1.21)

Exponential of commuting matrices (Theor. 1.23)

Logarithm and exponential as locally inverse maps (Prop. 1.27)

Surjectivity of the exponential map of $M(n \times n, \mathbb{C})$ (Theor. 1.29)

Chapter 2. Fundamentals of Lie algebra theory

Lie algebra (Def. 2.2)

Specific matrix Lie algebras (Def. 2.3)

Representation of a Lie algebra (Def. 2.4)

1-parameter subgroup and its infinitesimal generator (Def. 2.9)

Lie algebra of a matrix group (Def. 2.14)

Lie algebras of the classical groups (Prop. 2.15)

Topology of some classical groups (Prop. 2.18)

Universal covering of $SO(3,\mathbb{R})$ and of the Lorentz group (Ex. 2.24, Prop. 2.27)

Chapter 3. Nilpotent Lie algebras and solvable Lie algebras

Annihilation of a common eigenvector for nilpotent matrix algebras (Theor. 3.5)

Engel's theorem about nilpotent Lie algebras (Theor. 3.10)

Exact sequence of Lie algebra morphisms (Def. 3.11)

Radical of a Lie algebra (Def. 3.18)

Existence of a common eigenvector for solvable matrix algebras (Theor. 3.20)

Lie's theorem about complex solvable matric algebras (Theor. 3.21)

Chapter 4. Killing form and semisimple Lie algebras

Killing form (Def. 4.2)

Cartan's characterization of solvability referring to the Killing form (Cor. 4.4)

Semisimple Lie algebra (Def. 4.5)

Cartan's characterization of semisimpleness referring to the Killing form (Theor. 4.14)

Semisimpleness of the classical Lie algebras from the *ABCD*-series (Cor. 4.18)

Semsimple Lie algebras split as direct sums of simple Lie algebras (Theor. 4.21)

Lemma of Schur (Theor. 4.25)

Quadratic Casimir element of a representation (Def. 4.26)

Weyl's theorem on complete reducibility (Theor. 4.30)

Abstract Jordan decomposition in semisimple Lie algebras (Def. 4.32)

Chapter 5. Root space decomposition

Existence of non-zero toral subalgebras (Prop. 5.2)

Structure of the simple Lie algebra $sl(2,\mathbb{C})$ (Prop. 5.4)

Classification of finite-dimensional irreducible $sl(2,\mathbb{C})$-modules (Theor. 5.10)

Maximal toral subalgebras are equal to their centralizer (Theor. 5.17)

Root space decomposition alias Cartan decomposition (Def. 5.18)

Cartan subalgebra (Def. 5.19)

Chapter 6. Root systems

Root system and Cartan integers (Def. 6.3)

Possible angles and length ratio of two roots (Lem. 6.10)

Base of a root system (Theor. 6.14)

Transitive action of the Weyl group (Prop. 6.20)

Cartan matrix of a root system (Theor. 6.23)

Classification of connected Coxeter graphs (Theor. 6.27)

Classification of connected Dynkin diagrams (Theor. 6.31)

Chapter 7. The root systems of complex semisimple Lie algebras

Semisimple Lie algebras as $sl(2,\mathbb{C})$-modules (Prop. 7.3)

Root system of a complex semisimple Lie algebra (Theor. 7.5)

Root systems and Dynkin diagrams from the A,B,C,D-series (Prop. 7.6 - 7.9)

FINIS

References

The main references for these notes are

- Lie algebra: Hall [18], Humphreys [24] and Serre [40]
- Lie group: Hall [18], Serre [41].

In addition,

- references with focus on mathematics:
  [4], [5], [13], [21], [22], [23], [24], [46]
- references with focus on physics:
  [3], [15], [17], [20], [36], [37]
- references with focus on both mathematics and physics:
  [2], [16], [17], [27], [42]

1. Barut, A. O.; Schneider, C. K. E.; Wilson, Raj: Quantum Theory of infinite component fields. Journal of Physics. Vol. 20, 2244 (1979)
2. Böhm, Manfred: Lie-Gruppen und Lie-Algebren in der Physik. Eine Einführung in die mathematischen Grundlagen. Springer, Berlin (2011)
3. Born, Max; Jordan, Pascual: Zur Quantenmechanik. Zeitschrift für Physik **34**, 858-888 (1925)
4. Bourbaki, Nicolas: Éléments de mathématique. Groupes et Algèbres de Lie. Chapitre I. Diffusion C.C.L.S., Paris (without year)
5. Bourbaki, Nicolas: Éléments de mathématique. Groupes et Algèbres de Lie. Chapitre VII, VIII. Algèbres de Lie. Diffusion C.C.L.S., Paris (without year)
6. Bröcker, Theodor; tom Dieck, Tammo: Representations of Compact Lie groups. Springer, New York (1985)
7. Brion, Michel: Lectures on the geometry of flag varieties. Notes de l'école d'été "Schubert" varieties (Varsovie 2003). http://www-fourier.ujf-grenoble.fr/ mbrion/notes.html Call 10.2.2023
8. Bryant, Robert: Answer to "Beautiful descriptions of exceptional groups" (2012), https://mathoverflow.net/questions/99736/beautiful-descriptions-of-exceptional-groups, call 10.11.2022
9. Dobrev, Vladimir K.: Elementary representations and interwining operators for SU(2,2).I Journal of Mathematical Physics. Vol. 26, 235 (1985)
10. Dugundji, James: Topology. Allyn and Bacon, Boston (1966)
11. Duistermaat, Johannes J.; Kolk, Johan A.C.: Lie Groups. Springer (2013)

12. Fischer, Gerd: Lineare Algebra. Eine Einführung für Studienanfänger. Springer Spektrum, 16. Aufl. (2008)
13. Fulton, William; Harris, Joe: Representation Theory. A First Course. Springer, New York, 3rd printing (1996)
14. Gasiorowicz, Stephen: Elementary Particle Physics. John Wiley & Sons, New York (1966)
15. Georgi, Howard: Lie Algebras in Particle Physics. Westview, 2nd ed. (1999)
16. Gilmore, Robert: Lie Groups, Lie Algebras, and some of their Applications. Dover Publications, Mineola (2005)
17. Hall, Brian: Quantum Theory for Mathematicians. Springer, New York (2013)
18. Hall, Brian: Lie Groups, Lie Algebras, and Representations. An Elementary Introduction. Springer, Heidelberg, 2^{ed} (2015)
19. Hatcher, Allen: Algebraic Topology. Cambridge University Press, Cambridge (2002). Also https://www.math.cornell.edu/ hatcher/AT/AT.pdf
20. Heisenberg, Werner: Die Kopenhagener Deutung der Quantentheorie. In ders.: Physik und Philosophie. Ullstein, Frankfurt/Main (1959)
Also https://www.marxists.org/reference/subject/philosophy/works/ge/heisenb3.htm Download 6.6.2018
21. Helgason, Sigurdur: Differential Geometry, Lie Groups and Symmetric Spaces. Academic Press, New York (1978)
22. Hilgert, Joachim, Neeb, Karl-Hermann: Lie Gruppen und Lie Algebren. Braunschweig (1991)
23. Hilgert, Joachim, Neeb, Karl-Hermann: Structure and Geometry of Lie Groups. New York (2012)
24. Humphreys, James E.: Introduction to Lie Algebras and Representation Theory. Springer, New York (1972)
25. Iachello, Francesco: Lie Algebras and Application. Lecture Notes in Physics 708, Springer, Berlin (2006)
26. Jacobson, Nathan: Exceptional Lie Algebras. Marcel Dekker, New York (1971)
27. Jauch, Josef: Foundations of Quantum Mechanics. Addison-Wesley, Reading, MA., (1968)
28. Kac, Victor, in *Introduction to Lie Algebras*, http://math.mit.edu/classes/18.745/index.html Cited 17 September 2018
29. Kirillov, Aleksandr Aleksandrovich: Lecture on the Orbit Method. Graduate Studies in Mathematics Vol. 64. American Mathematical Society (2004)
30. Knapp, Anthony: Lie Groups, Lie Algebras, and Cohomology. Princeton University Press, Princeton (1988)
31. Knapp, Anthony: Lie Groups Beyond an Introduction. Birkhäuser, Boston 2nd ed. (2005)
32. Knapp, Anthony; Trapa, Peter E.: Representations of Semisimple Lie Groups. American Mathematical Society (2000)
33. Knapp, Anthony; Speh, Birgit: Irreducible Unitary Representations of $SU(2,2)$. Journal of Functional Analysis 45, 41-73 (1982)
34. Lipkin, Harry J.: Lie Groups for Pedestrians (Deutsch: Anwendung von Lieschen Gruppen in der Physik). North-Holland Publishing. Amsterdam. 2nd ed. (1966)
35. Leytem, Alain: Some unexpected facts about Lie Algebras. https://orbilu.uni.lu/bitstream/10993/23377/1/1Lie%20Algebras.pdf Call 20.11.2022
36. Messiah, Albert: Quantum Mechanics. Volume 1. North-Holland Publishing Company, Amsterdam (1970)
37. Roman, Paul: Advanced Quantum Theory. An outline of the fundamental ideas. Addison-Wesley, Reading Mass. (1965)
38. Rim, Donsub: An elementary proof that symplectic matrices have determinant one. `https://arxiv.org/abs/1505.04240}` Download 9.5.2018
39. Rovelli, Carlo: Relational Quantum Mechanics. International Journal of Theoretical Physics, Vol 35, No. 8, (1996)
40. Serre, Jean-Pierre: Complex Semisimple Lie Algebras. Reprint 1987 edition, Springer, Berlin (2001)
41. Serre, Jean-Pierre: Lie Algebras and Lie Groups. 1964 Lectures given at Harvard University. 2nd edition, Springer, Berlin (2006)

42. Schottenloher, Martin: Geometrie und Symmetrie in der Physik. Leitmotiv der Mathematischen Physik. Vieweg, Braunschweig (1995)
43. Speh, Birgit: Degenerate Series Representations of the Universal Covering Group of $SU(2,2)$. Journal of Functional Analysis 33, 95-118 (1979)
44. Stöcker, Ralph; Zieschang, Heiner: Algebraische Topologie. Eine Einführung. Teubner, Stuttgart (1988)
45. Spanier, Edward: Algebraic Topology. Tata McGraw-Hill, New Delhi (Repr. 1976)
46. Varadarajan, Veeravalli Seshadri: Lie Groups, Lie Algebras, and their Representations. Springer, New York (1984)
47. Wehler, Joachim: Lie Groups. (2017) http://www.mathematik.uni-muenchen.de/%7ewehler/LieAlgebrasLieGroupsScript.pdf
48. Weyl, Hermann: Theorie der Darstellung kontinuierlicher halb-einfacher Gruppen durch lineare Transformationen. I. Math Z 23, 271–309 (1925)
49. Woit, Peter: Quantum Theory, Groups and Representations. An Introduction. Springer (2017)