LIE ALGEBRAS, LECTURE NOTES

P. SOSNA

Contents

These are notes for a lecture (14 weeks, $1,5 \times 90$ minutes per week) held at the University of Hamburg in the summer semester 2016. The goal was to introduce the necessary concepts in order to be able to outline the proof of the classification result for semisimple complex Lie algebras. In particular, the classification theorem for irreducible root systems is only stated but not proved and the construction of a (semi-)simple Lie algebra from a given root system is only outlined.

For the most part I closely followed [2], basically just sometimes rearranging the order in which topics are presented, explaining some of the proofs in more detail and omitting several topics.

1. Basic concepts

Definition. Let $\mathfrak{g}$ be a vector space over some field K endowed with an operation $\mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$ denoted by $(x, y) \longmapsto [x, y]$ and called the *bracket* or *commutator*. The pair $(\mathfrak{g}, [\,,\,])$ is called a *Lie algebra* over K if the following axioms are satisfied:

(L1) The bracket is bilinear.
(L2) $[x,x]=0$ for all $x\in\mathfrak{g}$.
(L3) $[x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0$ for all $x,y,z\in\mathfrak{g}$.

The third axiom is called the *Jacobi idenity.*

Note that

$$0=[x+y,x+y]=[x,x]+[x,y]+[y,x]+[y,y]=[x,y]+[y,x]\quad\forall x,y\in\mathfrak{g},$$

hence we have (L2') $[x,y]=-[y,x]$, that is, the bracket is anticommutative. If the characteristic of K is not 2, then (L2') implies (L2) ($[x,x]=-[x,x]$).

Definition. A *K-algebra* is a K-vector space A together with a K-bilinear map $m\colon A\times A\longrightarrow A$ called the *multiplication.* One usually writes $x\cdot y$ instead of $m(x,y)$. A K-algebra is called *associative* if $(x\cdot y)\cdot z=x\cdot(y\cdot z)$ for all $x,y,z\in A$. A K-algebra is called *unital* if there exists an element $1\in A$ such that $1\cdot x=x\cdot 1=x$ for all $x\in A$. A K-algebra is called *commutative* if $x\cdot y=y\cdot x$ for all $x,y\in A$.

We will frequently write xy instead of $x\cdot y$ for the multiplication of two elements in a K-algebra. If the underlying field is clear we will also frequently simply say algebra instead of K-algebra. The dimension of an algebra is its dimension as a K-vector space.

{e:Algebras}

Example 1.1. (1) The polynomial ring $K[x_1,\ldots,x_n]$ over a field is an associative commutative unital K-algebra.
(2) Let V be any K-vector space and $A=\operatorname{End}(V)$ be the space of K-linear endomorphisms of V. This is an associative unital K-algebra with respect to the composition of maps, the identity being the unit element.
(3) The space of $n\times n$-matrices over a field K is an associative unital K-algebra.

{e:gl}

Example 1.2. Let A be any associative algebra. Define a new operation $A\times A\longrightarrow A$ by sending (x,y) to $xy-yx=:[x,y]$. This operation clearly satisfies (L1) and (L2) and an easy computation, which is left to the reader (see Exercise 1 on Sheet 1), shows that (L3) also holds. Therefore, any associative algebra defines a Lie algebra.

As an explicit example, let $A=\operatorname{End}(V)$. The corresponding Lie algebra will be denoted by $\mathfrak{gl}(V)$. It is called the *general linear Lie algebra.*

Definition. Let A and B be two K-algebras.

(1) A K-algebra homomorphism from A to B is a K-linear map $\varphi\colon A\longrightarrow B$ which is compatible with the respective multiplications, that is, $\varphi(x)\varphi(y)=\varphi(xy)$ for all $x,y\in A$.
(2) A K-algebra homomorphism which is an isomorphism of K-vector spaces is called an *isomorphism of K-algebras.*
(3) Assume that A and B are unital algebras. An algebra homomorphism $\varphi\colon A\longrightarrow B$ is called *unital* if $\varphi(1_A)=1_B$.

{e:MapsAlg}

Example 1.3. The map $\varphi\colon K[x] \longrightarrow K$ defined by $p \longmapsto p(0)$ is a unital homomorphism of K-algebras. On the other hand, the map

$$\det\colon \mathrm{Mat}(n, K) \longrightarrow K$$

is compatible with multiplication, but is not linear, hence is not a homomorphism of K-algebras.

{r:Liealghom}

Remark 1.4. Since a Lie algebra is in particular an algebra, the above definitions apply. Hence, a Lie algebra homomorphism is a linear map $\varphi\colon \mathfrak{g} \longrightarrow \mathfrak{g}'$ between Lie algebras respecting the brackets, that is $\varphi([x, y]) = [\varphi(x), \varphi(y)]$.

{r:Algring}

Remark 1.5. An associative unital K-algebra A is simply a unital ring together with a unital ring homomorphism $K \longrightarrow A$ (which is automatically injective since K has no non-trivial ideals) whose image is in the center of A. Recall that the center of a ring is the space $\{x \in A \mid xy = yx \ \ \forall y \in A\}$.

For example, if $A = \mathrm{End}(V)$, then the required ring homomorphism $K \longrightarrow A$ is given by sending $\lambda \in K$ to $\lambda \cdot \mathrm{id}_V$.

Definition. Let A be an algebra. A *subalgebra* is a vector subspace B such that for all $x, y \in B$ the element xy is also in B. Note that B is again an algebra.

A subalgebra B of a unital algebra A is called *unital* if $1_A \in B$.

Any subalgebra of the Lie algebra $\mathfrak{gl}(V)$ is called a *linear Lie algebra*.

Example 1.6. Let $\mathfrak{g}$ be a Lie algebra and $0 \neq x \in \mathfrak{g}$. Then Kx is a Lie algebra with trivial bracket (that is, $[y, z] = 0$ for all $y, z \in Kx$), because of (L2).

For a different example, consider $A = K[x, y]$ and $B = K[x]$.

Note that the intersection of two subalgebras is again a subalgebra. However, this does not hold for the (span of the) union: Take $A = \mathrm{Mat}(3, \mathbb{C})$, $B_1 = \{\begin{pmatrix} 0 & \lambda & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mid \lambda \in \mathbb{C}\}$, $B_2 = \{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \lambda \\ 0 & 0 & 0 \end{pmatrix} \mid \lambda \in \mathbb{C}\}$.

{e:Derivations}

Example 1.7. Let A be any K-algebra. A *derivation* of A is a linear map $\delta\colon A \longrightarrow A$ which satisfies the *Leibniz rule* $\delta(xy) = \delta(x)y + x\delta(y)$. The space of all derivations of A is a vector space (this is left to the reader), denoted by $\mathrm{Der}(A)$. The reader can check that if δ, δ' are derivations, then $\delta\delta' - \delta'\delta$ is also a derivation (see Exercise 1 on Sheet 1). In other words, $\mathrm{Der}(A) \subset \mathfrak{gl}(A)$ is a Lie subalgebra.

In the following we will introduce some examples of linear Lie algebras. First note that if V is finite-dimensional, then we can identify it with K^n after choosing a basis. The algebra $\mathrm{End}(V)$ can then be identified with the algebra $\mathrm{Mat}(n, K)$. We will denote the corresponding Lie algebra by $\mathfrak{gl}(n, K)$.

{e:sl}

Example 1.8. Consider the space

$$\mathfrak{sl}(n, K) := \{A \in \mathfrak{gl}(n, K) \mid \operatorname{tr}(A) = 0\}.$$

Clearly, this is a vector subspace. For any $A, B \in \mathfrak{gl}(n, K)$ we have the equality

$$\operatorname{tr}([A, B]) = \operatorname{tr}(AB - BA) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0,$$

hence $\mathfrak{sl}(n, K)$ is a subalgebra of $\mathfrak{gl}(n, K)$, called the *special linear algebra*. Since the definition of the trace does not depend on the choice of a basis, it also makes sense to talk about the Lie algebra $\mathfrak{sl}(V)$ if V is finite-dimensional.

Note that the subspace of tracefree matrices is not a subalgebra of the associative algebra $\operatorname{Mat}(n, K)$ (e.g., take the product with itself of the diagonal matrix 2×2-matrix with entries 1 and -1).

Let us compute the dimension of $\mathfrak{sl}(n + 1, K)$. First of all note that the dimension of $\mathfrak{gl}(n + 1, K)$ is $(n + 1)^2$. Since $\mathfrak{sl}(n + 1, K)$ is a proper subalgebra of $\mathfrak{gl}(n + 1, K)$ its dimension is at most $(n + 1)^2 - 1$. Write $E_{i,j}$ for the matrix which has a 1 in the (i, j) position and 0 everywhere else and note that the matrices $E_{i,j}$ for $i \neq j$ and $E_{i,i} - E_{i+1,i+1}$ for $1 \leq i \leq n$ are all in $\mathfrak{sl}(n + 1, K)$ and are linearly independent. These are $(n+1)^2 - (n+1) + n = (n+1)^2 - 1$ matrices, hence $\dim_K \mathfrak{sl}(n+1, K) = (n+1)^2 - 1$.

{e:OrthLie}

Example 1.9. Let V be a K-vector space and $f\colon V \times V \longrightarrow W$ be a bilinear map into a vector space W. Consider the space

$$\mathfrak{o}(V, f) := \{x \in \mathfrak{gl}(V) \mid f(xv, w) + f(v, xw) = 0 \ \forall v, w \in V\}.$$

This is a Lie subalgebra of $\mathfrak{gl}(V)$, since:

$$\begin{aligned} f([x, y]v, w) &= f((xy - yx)v, w) = f(xyv, w) - f(yxv, w) \\ &= f(v, yxw) - f(v, xyw) = -f(v, [x, y]w), \end{aligned}$$

so if $x, y \in \mathfrak{o}(V, f)$, then also $[x, y] \in \mathfrak{o}(V, f)$.

Let us consider some specific examples, where W is always the base field.

(1) Let $V \simeq K^{2n}$ and let f be the bilinear map given by the matrix $\begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$. Note that this form is antisymmetric. In this case the Lie algebra $\mathfrak{o}(V, f)$ is denoted by $\mathfrak{sp}(2n, K)$ and called the *symplectic Lie algebra*.

(2) Let $V \simeq K^n$ and take f to be the identity matrix. The resulting Lie algebra is called the *orthogonal Lie algebra* and denoted by $\mathfrak{so}(n, K)$. It can be easily checked that it consists of the $n \times n$-matrices A with $A = -A^t$ (see Remark 1.14 below).

For computations it is sometimes more useful to have a different description. Namely, if $n = 2m$ is even, we let f be defined by the matrix $\begin{pmatrix} 0 & I_m \\ I_m & 0 \end{pmatrix}$, while for n odd, say $n = 2m + 1$, we take the matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_m \\ 0 & I_m & 0 \end{pmatrix}$.

{e:TriandM}

Example 1.10. The set $\mathfrak{t}(n, K)$ of all *upper triangular $n \times n$-matrices A* (that is, $A_{i,j} = 0$ for all $i > j$) is a Lie subalgebra of $\mathfrak{gl}(n, K)$. Similarly, the set $\mathfrak{n}(n, K)$ of all *strictly upper triangular matrices A* (that is, $A_{i,j} = 0$ for all $i \geq j$) is also a Lie subalgebra of $\mathfrak{gl}(n, K)$. The set $\mathfrak{d}(n, K)$ of all diagonal matrices is also easily seen to be a subalgebra of $\mathfrak{gl}(n, K)$. Compare Exercise 4 on Sheet 1.

Definition. A Lie algebra $\mathfrak{g}$ is called *abelian* if $[x, y] = 0$ for all $x, y \in \mathfrak{g}$.

{e:AbLieAlg}

Example 1.11. Any vector space can be viewed as an abelian Lie algebra. Also note that the diagonal matrices are an abelian subalgebra of $\mathfrak{gl}(n, K)$. If $\mathfrak{g}$ is any Lie algebra, then $Kx \subset \mathfrak{g}$ is an abelian algebra for any $x \neq 0$. In fact, there is, up to isomorphism, only over one-dimensional Lie algebra, which is abelian.

Note that most examples we have looked at so far are linear Lie algebras. In fact, every finite dimensional Lie algebra is isomorphic to some linear Lie algebra. This is the content of theorems due to Ado and Iwasawa.

It might also be useful to contemplate abstract Lie algebras for a moment. As noted above, every vector space V can be endowed with the trivial bracket and hence with the structure of a Lie algebra, which is then abelian. Now assume that $\mathfrak{g}$ is a finite-dimensional Lie algebra with basis $x_1, \ldots, x_n$, then the multiplication table of $\mathfrak{g}$ can be recovered from the *structure constants* a_{ij}^k which occur in the expressions

$$[x_i, x_j] = \sum_{k=1}^{n} a_{ij}^k x_k.$$

Due to the anticommutativity of the bracket, it is enough to consider the constants a_{ij}^k for $i < j$.

Conversely, one can define an abstract Lie algebra by specifying a set of structure constants. That is, one takes a finite-dimensional vector space, fixes a basis and a set of constants $\{a_{ij}^k\}$ satisfying the relations

$$a_{ii}^k = 0 = a_{ij}^k + a_{ji}^k$$

and

$$\sum_k (a_{ij}^k a_{kl}^m + a_{jl}^k a_{ki}^m + a_{li}^k a_{kj}^m) = 0.$$

The first equation stems from axiom (L2) while the second stems from (L3). This concludes the detour concerning abstract algebras.

Definition. Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h}$, $\mathfrak{i}$ be subalgebras of $\mathfrak{g}$. We will write $[\mathfrak{h}, \mathfrak{i}]$ for the subspace of $\mathfrak{g}$ spanned by commutators $[x, y]$ with $x \in \mathfrak{h}$ and $y \in \mathfrak{i}$.

A Lie algebra is called *simple* if it is non abelian, that is $[\mathfrak{g}, \mathfrak{g}] \neq 0$, and every non-zero homomorphisim into some other Lie algebra is injective.

A lot of concepts and notions which we will introduce are needed in order to prove the following classification result, whose proof we will outline in the course of the lecture:

{t:MainThm}

Theorem 1.12 (Killing, Cartan)**.** *Each simple finite-dimensional complex Lie algebra is isomorphic to one of the following* classical Lie algebras

$$\begin{aligned} \mathfrak{sl}(n+1, \mathbb{C}) &\qquad n \geq 1 \quad (A_n) \\ \mathfrak{so}(2n+1, \mathbb{C}) &\qquad n \geq 2 \quad (B_n) \\ \mathfrak{sp}(2n, \mathbb{C}) &\qquad n \geq 3 \quad (C_n) \\ \mathfrak{so}(2n, \mathbb{C}) &\qquad n \geq 4 \quad (D_n). \end{aligned}$$

or one of the following exceptional Lie algebras*:* e_6, e_7, e_8*,* f_4 *or* g_2*.*

In order to make a comment on the above theorem, we need the following

Definition. Let A and B be two algebras. The *direct sum algebra* is the vector space $A \oplus B$ endowed with the componentwise multiplication.

{r:ExcIsom}

Remark 1.13*.* Note that $\mathfrak{so}(2, \mathbb{C})$ is one-dimensional (a basis vector is, for example, the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$), hence abelian and in particular not simple. There are also the following isomorphisms:

$$\begin{aligned} &\mathfrak{so}(3, \mathbb{C}) \simeq \mathfrak{sp}(2, \mathbb{C}) \simeq \mathfrak{sl}(2, \mathbb{C}), \\ &\mathfrak{sp}(4, \mathbb{C}) \simeq \mathfrak{so}(5, \mathbb{C}), \\ &\mathfrak{so}(4, \mathbb{C}) \simeq \mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C}), \\ &\mathfrak{so}(6, \mathbb{C}) \simeq \mathfrak{sl}(4, \mathbb{C}). \end{aligned}$$

{r:DimSympLie}

Remark 1.14*.* While we will not prove the isomorphisms of Remark 1.13 (Exercise 5 on Sheet 1 deals with the first chain of isomorphisms), let us at least outline how to compute the dimensions of the Lie algebras occurring there.

So, let $f \colon K^n \times K^n \longrightarrow K$ be a bilinear form given by an $n \times n$-matrix F, that is, $f(x, y) = x^t F y$. A matrix $A \in \mathfrak{gl}(n, K)$ is in $\mathfrak{o}(K^n, f)$ if and only if $(Ax)^t F y + x^t F A y = 0$ for all $x, y \in K^n$. To put it differently,

$$A \in \mathfrak{o}(K^n, f) \Longleftrightarrow A^t F = -FA.$$

So, if, for example, $F = I_n$, then $A^t = -A$ is skew-symmetric as stated in Example 1.9(2). Note that the dimension of the space of skew-symmetric $n \times n$-matrices depends on the characteristic of the base field: If the characteristic is not equal to 2, then this space has dimension $\frac{n(n-1)}{2}$. But if $\mathrm{char}(K) = 2$, then the dimension is $\frac{n(n-1)}{2} + n$.

Now, let us consider the case $\mathfrak{sp}(2n,K)$, hence $F=\begin{pmatrix}0 & I_n\\ -I_n & 0\end{pmatrix}$. Writing a matrix $A\in\mathfrak{gl}(2n,K)$ in block form (that is, the blocks are $n\times n$-matrices), we get

$$\begin{aligned}
&A=\begin{pmatrix}A_1 & A_2\\ A_3 & A_4\end{pmatrix}\in\mathfrak{sp}(2n,K)\iff\\
&\iff\begin{pmatrix}A_1^t & A_3^t\\ A_2^t & A_4^t\end{pmatrix}\begin{pmatrix}0 & I_n\\ -I_n & 0\end{pmatrix}=-\begin{pmatrix}0 & I_n\\ -I_n & 0\end{pmatrix}\begin{pmatrix}A_1 & A_2\\ A_3 & A_4\end{pmatrix}\\
&\iff\begin{pmatrix}-A_3^t & A_1^t\\ -A_4^t & A_2^t\end{pmatrix}=\begin{pmatrix}-A_3 & -A_4\\ A_1 & A_2\end{pmatrix}\\
&\iff A_2=A_2^t,\ A_3=A_3^t,\ A_1=-A_4^t.
\end{aligned}$$

Therefore, we get

$$\dim_K\mathfrak{sp}(2n,K)=n^2+2\frac{n(n+1)}{2}=2n^2+n.$$

2. Connection to Lie groups

This section is only motivational.

Recall that a *topological space* is a pair (X,σ), where X is a set and σ is a subset of the power set of X satisfying the following conditions: a) $\emptyset\in\sigma$, b) $X\in\sigma$, c) any union of elements of σ is in σ and 4) any finite intersection of elements of σ is an element in σ. The elements of σ are called the *open sets* of the topology.

Given two topological spaces, their product is a topological space by defining the open subsets to be products of open subsets.

A map $f\colon X\longrightarrow Y$ between topological spaces is *continuous* if the preimage of any open subset if open. Equivalently, one can work with closed subsets, that is, the complements of the open sets. For instance, the projection $X\times Y\longrightarrow Y$ is continuous. A bijective continuous map whose inverse is also continuous is called a *homeomorphism*.

A topological space X is called *Hausdorff* if for any $x\neq y\in X$ there exist open subsets $U_x\ni x$ and $U_y\ni y$ such that $U_x\cap U_y=\emptyset$.

A *basis* for a topology is a collection B of open subsets such that every open subset can be written as a union of elements of B.

Definition. An m-dimensional C^k-manifold is a topological space M together with an open covering $M=\bigcup U_i$ and homeomorphisms

$$\varphi_i\colon U_i\simeq V_i$$

onto open subsets $V_i\subset\mathbb{R}^m$ such that a) M is Hausdorff, b) the topology of M admits a countable basis and c) the functions $\varphi_j\circ\varphi_i^{-1}\colon\varphi_i(U_i\cap U_j)\longrightarrow\varphi_j(U_i\cap U_j)$ are C^k-maps.

A *differentiable manifold* is a C^∞-manifold of some finite dimension.

While a C^k-manifold is defined without an embedding into some $\mathbb{R}^n$, the Whitney embedding theorem states that any n-dimensional differentiable manifold can in fact be embedded into $\mathbb{R}^{2n}$. However, note that defining manifolds without reference to a surrounding space allows for constructions such as taking quotients by (reasonable) group actions.

The datum $\{(U_i, \varphi_i)\}$ is called an *atlas* and each tuple (U_i, φ_i) a *chart*. One says that two atlases $\{U_i, \varphi_i\}$ and $\{V_j, \psi_j\}$ define the same manifold if the transition functions $\psi_j \circ \varphi_i^{-1}$ are differentiable for all tuples (i, j).

Let $U \subset M$ be an open subset. A function $f\colon U \longrightarrow \mathbb{R}$ is *differentiable* if

$$f \circ \varphi_i^{-1}\colon \varphi_i(U_i \cap U) \longrightarrow \mathbb{R}$$

is differentiable for any chart (U_i, φ_i). We will write $C_M(U)$ for the set of all differentiable functions $U \longrightarrow \mathbb{R}$. Note that if $f \in C_M(U)$ and $V \subset U$ is open, then we can restrict f to V and get an element of $C_M(V)$. Also note that if $M = \bigcup V_j$ is an open covering and we are given elements $f_j \in C_M(V_j)$ for all j which agree on the intersections, then they glue to a unique element $f \in C_M(M)$.

Given two differentiable manifolds it is then clear what a differentiable map between them is.

Definition. A *Lie group* is a differentiable manifold G together with a group structure such that the multiplication $G \times G \longrightarrow G$, $(x, y) \longmapsto xy$ and taking the inverse $G \longrightarrow G$, $x \longmapsto x^{-1}$ are differentiable.

Example 2.1. (1) Every finite dimensional real vector space is a Lie group with respect to addition.
(2) The space of all upper triangular $n \times n$-matrices with 1 on the diagonal is a Lie group with respect to matrix multiplication.
(3) Since any open subset of a manifold is a manifold, the space of invertible $n \times n$-matrices is a Lie group.

Recall that the tangent space T_xM to $x \in M$ is the vector space of equivalence classes of derivatives in 0 of curves $\gamma\colon (-1, 1) \longrightarrow M$ satisfying $\gamma(0) = x$. Of course, we need to pick a chart containing x to make sense of this definition, but the end result does not depend on this choice. One can also define T_xM using derivations. Of course, every open subset U of $\mathbb{R}^m$ is a manifold and we have already used that a basis of the tangent space to $x \in U$ is given by the partial derivatives.

There is the following result

{p:LieOnePara}

Proposition 2.2. *Let G be a Lie group. There is a bijection between the space of C^∞-group homomorphisms $\varphi\colon \mathbb{R} \longrightarrow G$, also called one-parameter subgroups, and T_eG, where e is the neutral element of G. The bijection is defined by sending φ to $\varphi'(0)$.* □

Now, if G is an arbitrary Lie group, let $A \in T_eG$ be a tangent vector. By Proposition 2.2 there is a unique one-parameter subgroup $\varphi_A\colon \mathbb{R} \longrightarrow G$ with $\varphi_A'(0) = A$. Define the *exponential map*

$$\exp\colon T_eG \longrightarrow G, \quad A \longmapsto \varphi_A(1).$$

It can be checked that $\varphi_{tA}(s) = \varphi_A(ts)$ for all $t, s \in \mathbb{R}$. It follows that

$$\exp(tA) = \varphi_{tA}(1) = \varphi_A(t).$$

Definition. Let G be a Lie group. The *Lie algebra* of G is by definition the tangent space T_eG, where e is the neutral element of G, endowed with the following bracket

$$\begin{aligned}[A,B] &:= \lim_{t\longrightarrow 0} \frac{1}{t^2}\{\varphi_A(t)^{-1}\varphi_B(t)^{-1}\varphi_A(t)\varphi_B(t) - e\} \\ &= \lim_{t\longrightarrow 0} \frac{1}{t^2}\{\exp(-tA)\exp(-tB)\exp(tA)\exp(tB) - e\}.\end{aligned}$$

Example 2.3. Recall that $G = \mathrm{GL}(n,\mathbb{R})$ is a Lie group. In this case the one-parameter subgroups are precisely the maps

$$\mathbb{R} \longrightarrow \mathrm{GL}(n,\mathbb{R}), \quad t \longmapsto \exp(tA) = 1 + tA + \frac{t^2A^2}{2!} + \dots$$

We can now compute

$$\begin{aligned}&\exp(-tA)\exp(-tB)\exp(tA)\exp(tB) \\ &= 1 + t(-A - B + A + B) + \\ &+ \frac{t^2}{2}(A^2 + B^2 + A^2 + B^2 + 2AB - 2A^2 - 2AB - 2BA - 2B^2 + 2AB) + \\ &+ O(t^3) \\ &= 1 + t^2(AB - BA) + O(t^3).\end{aligned}$$

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

3. Ideals

Definition. Let A be a K-algebra. A vector subspace $I \subset A$ is called a *left-sided ideal* of A if $AI \subset I$, a *right-sided ideal* if $IA \subset I$ and an *ideal* if it is a left- and right-sided ideal. Here, for example, $AI \subset I$ means that $ai \in I$ for all $a \in A$ and $i \in I$.

Note that if A is a commutative or a Lie algebra, then any left- or right-sided ideal is automatically an ideal. In the Lie algebra case this follows from axiom (L2'), anticommutativity. Also note that any ideal is a subalgebra, but will usually not contain the unit of A if it exists.

Remark 3.1. A vector subspace $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is a subalgebra if $[\mathfrak{h},\mathfrak{h}] \subset \mathfrak{h}$ and an ideal if $[\mathfrak{g},\mathfrak{h}] \subset \mathfrak{h}$.

{e:ExIdeals}

Example 3.2. Since $\mathrm{tr}(xy - yx) = 0$ for all $x, y \in \mathfrak{gl}(n,K)$, the subspace $\mathfrak{sl}(n,K)$ is an ideal in $\mathfrak{gl}(n,K)$.

It is easy to see that $\mathfrak{n}(n,K)$ is an ideal in $\mathfrak{t}(n,K)$, the space of upper triangular matrices. Note that it is not an ideal in $\mathfrak{gl}(n,K)$. Indeed, take $n = 2$ and compute

$$[\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}] = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

{r:ExIdeals}

Remark 3.3. Let A be an algebra.

(1) The subspaces 0 and A are ideals, called the *trivial ideals.*
(2) The intersection of ideals is an ideal.
(3) Any subset S of A defines an ideal, namely the smallest ideal containing S. It is equal to the intersection of all ideals containing S.
(4) If I and J are ideals in A, then $I + J := \{x + y \mid x \in I, y \in J\}$ is also an ideal.
(5) Let $\mathfrak{g}$ be a Lie algebra. The *center*

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

is an ideal (see Exercise 3 on Sheet 2). Note that $\mathfrak{g}$ is abelian if and only if $Z(\mathfrak{g}) = \mathfrak{g}$.
(6) Let $\mathfrak{g}$ be a Lie algebra. The space $[\mathfrak{g}, \mathfrak{g}]$ spanned by elements of the form $[x, y]$ for $x, y \in \mathfrak{g}$ is easily seen to be an ideal. It is called the *derived algebra* of $\mathfrak{g}$. Note that $\mathfrak{g}$ is abelian if and only $[\mathfrak{g}, \mathfrak{g}] = 0$.
(7) Let $\mathfrak{g}$ be a Lie algebra. If I and J are ideals in $\mathfrak{g}$, then $[I, J] := \{\sum_k [x_k, y_k] \mid x_k \in I, y_k \in J\}$ is an ideal (the sums are of course finite). Note that the derived algebra $[\mathfrak{g}, \mathfrak{g}]$ is a special case of this construction.

We have seen above that if $\mathfrak{g}$ is an abelian Lie algebra, then the derived algebra $[\mathfrak{g}, \mathfrak{g}]$ is trivial. On the other end, consider $\mathfrak{g} = \mathfrak{sl}(n, K)$ (if $\mathrm{char}(K) = 2$, then take $n \neq 2$). Recall from Example 1.8 that it has a basis consisting of the matrices $E_{i,j}$ for $i \neq j$ and $E_{i,i} - E_{i+1,i+1}$ for $1 \leq i \leq n$. One can use the (easily checked) formula (compare Exercise 1 on Sheet 1)

$$[E_{i,j}, E_{k,l}] = \delta_{jk} E_{i,l} - \delta_{li} E_{k,j} \quad (3.1) \quad \{\texttt{eq:Commut}\}$$

to show that $[\mathfrak{sl}(n, K), \mathfrak{sl}(n, K)] = \mathfrak{sl}(n, K)$ (compare Exercise 2 on Sheet 1) as follows.

Of course, inclusion $\subset$ is clear. But $\mathfrak{sl}(n, K) \subset [\mathfrak{sl}(n, K), \mathfrak{sl}(n, K)]$ as well. Indeed, for $i \neq j$ one easily checks that $[E_{i,i+1}, E_{i+1,i+1}] = E_{i,i} - E_{i+1,i+1}$, while $[E_{i,i} - E_{i+1,i+1}, E_{i,j}] = E_{i,j}$.

{p:OpIdeals}

Proposition 3.4. *Let A and B be algebras.*
(1) *If $\varphi\colon A \longrightarrow B$ is an algebra homomorphism, then $\ker(\varphi)$ is an ideal in A.*
(2) *If $I \subset A$ is an ideal, then there exists a unique algebra structure on the quotient vector space A/I such that the canonical projection $\pi\colon A \longrightarrow A/I$ is an algebra homomorphism.*
(3) *If $\varphi\colon A \longrightarrow B$ is an algebra homomorphism and $I \subset A$ is an ideal contained in the kernel of φ, then there exists a unique algebra homomorphism $\widetilde{\varphi}\colon A/I \longrightarrow B$ such that $\widetilde{\varphi} \circ \pi = \varphi$. In particular, $A/\ker(\varphi) \simeq \mathrm{im}(\varphi)$ (the reader should check that the image of a homomorphism is a subalgebra).*
(4) *If $\varphi\colon A \longrightarrow B$ is an algebra homomorphism and $J \subset B$ is an ideal, then $\varphi^{-1}(J)$ is an ideal in A.*
(5) *If $\varphi\colon A \longrightarrow B$ is a* surjective *algebra homomorphism and $I \subset A$ is an ideal, then $\varphi(I)$ is an ideal in B.*

Proof. Exercise (1 on Sheet 3). □

{r:IdealsCorr}

Remark 3.5. In the following we collect some more statements concerning ideals. The proofs are only outlined.

(1) One checks that there is a one-to-one correspondence between surjective homomorphisms and ideals of an algebra. Indeed, to an ideal I one associates the map $A \twoheadrightarrow A/I$ and to a surjective map one associates its kernel.
(2) If I and J are ideals in an algebra A and $I \subset J$, then J/I is an ideal of A/I. Conversely, the preimage of an ideal in A/I is an ideal J of A containing I. This gives a one-to-one correspondence between ideals in A/I and ideals in A contaning I.
(3) If I and J are ideals in A such that $I \subset J$, then
$$(A/I)/(J/I) \simeq A/J.$$
Note that since $I \subset J$, there is a natural surjective homomorphism $A/I \twoheadrightarrow A/J$ by Proposition 3.4(3). One checks that its kernel if precisely J/I.
(4) If I and J are ideals in A, then there is a natural isomorphism between $I/I \cap J$ and $(I+J)/J$. Indeed, one considers the canonical map $I \longrightarrow I+J \twoheadrightarrow (I+J)/J$, checks that it is surjective with kernel $I \cap J$ and concludes by Proposition 3.4(3).

{c:Liesimple}

Corollary 3.6. *A Lie algebra* $\mathfrak{g}$ *is simple if and only if* $[\mathfrak{g}, \mathfrak{g}] \neq 0$ *and its only ideals are* 0 *and* $\mathfrak{g}$.

Proof. If the only ideals are 0 and $\mathfrak{g}$, then any non-trivial homomorphism $\varphi\colon \mathfrak{g} \longrightarrow \mathfrak{h}$ is injective by Proposition 3.4(1). Conversely, if any non-trivial homomorphism is injective, then there cannot be any non-trivial ideals since otherwise the morphism $\mathfrak{g} \twoheadrightarrow \mathfrak{g}/I$ would be a non-injective homomorphism by Proposition 3.4(2). □

{c:QuotAb}

Corollary 3.7. *If* $\mathfrak{g}$ *is any Lie algebra, then* $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ *is an abelian Lie algebra.*

Proof. This immediately follows from the definition of the multiplication on $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$. □

{r:LieSimple}

Remark 3.8. Note that if $\mathfrak{g}$ is simple, then in particular $[\mathfrak{g}, \mathfrak{g}] \neq 0$, so $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$, thus $\mathfrak{g}$ is not abelian. Therefore, $Z(\mathfrak{g}) = 0$, since the center is an ideal and it is not equal to $\mathfrak{g}$.

{e:sl2}

Example 3.9. Consider $\mathfrak{g} = \mathfrak{sl}(2, K)$ and assume that $\operatorname{char}(K) \neq 2$. In the following we will show that this is a simple Lie algebra. Take the following matrices as a basis for $\mathfrak{g}$:
$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$
One checks immediately that $[X, Y] = H$, $[H, X] = 2X$ and $[H, Y] = -2Y$.

Let I be an ideal in $\mathfrak{g}$ and let $aX + bY + cH$ be an arbitrary nonzero element in I. Then
$$[X, aX + bY + cH] = bH - 2cX, \quad \text{so} \quad [X, [X, aX + bY + cH]] = -2bX.$$
Therefore, if $b \neq 0$, then $X \in I$, hence $[X, Y] = H \in I$, and also $[H, Y] = -2Y \in I$, so $Y \in I$. Therefore, if $b \neq 0$, then $I = \mathfrak{g}$. Similarly, if $a \neq 0$, then computing

$[Y, [Y, [X, aX + bY + cH]]]$ shows that $I = \mathfrak{g}$. Finally, if $a = b = 0$, then $0 \neq cH \in I$, hence $H \in I$ and then, as before, $I = \mathfrak{g}$ follows. Therefore, $\mathfrak{g}$ is a simple Lie algebra as predicted by Theorem 1.12. Over a field of characteristic two this algebra is not simple, compare Exercise 2 on Sheet 1.

4. Solvable and nilpotent Lie algebras

Definition. Let $\mathfrak{g}$ be a Lie algebra. The *derived series* of $\mathfrak{g}$ is the following sequence of ideals

$$\mathfrak{g}^{(0)} = \mathfrak{g},\ \mathfrak{g}^{(1)} = [\mathfrak{g}, \mathfrak{g}],\ \mathfrak{g}^{(2)} = [\mathfrak{g}^{(1)}, \mathfrak{g}^{(1)}], \dots, \mathfrak{g}^{(i+1)} = [\mathfrak{g}^{(i)}, \mathfrak{g}^{(i)}], \dots$$

The algebra $\mathfrak{g}$ is called *solvable* if $\mathfrak{g}^{(n)} = 0$ for some n.

The fact that $\mathfrak{g}^{(i)}$ is indeed an ideal follows by a direct computation, compare Exercise 2 on Sheet 3.

{e:ExSolv}

Example 4.1. Any abelian algebra is solvable. Indeed, if $\mathfrak{g}$ is abelian, then all brackets are trivial, hence $\mathfrak{g}^{(1)} = 0$.

On the contrary, if $\mathfrak{g}$ is simple, then it cannot be solvable. Indeed, by definition $\mathfrak{g}^{(1)} \neq 0$ and, in fact $\mathfrak{g}^{(1)} = \mathfrak{g}$. It follows that $\mathfrak{g}^{(n)} = \mathfrak{g}$ for all n.

Note that $\mathfrak{g}^{(i)} \subset \mathfrak{g}^{(i-1)}$ is an ideal for all i.

{e:TriangSolv}

Example 4.2. Recall that $\mathfrak{g} = \mathfrak{t}(n, K)$ denotes the Lie algebra of upper triangular $n \times n$-matrices over K. Clearly, it has a basis given by the matrices $E_{i,j}$ for $i \leq j$. In particular, its dimension is $\frac{n(n+1)}{2}$. We will now show that this algebra is solvable.

Recall Equation (3.1): $[E_{i,j}, E_{k,l}] = \delta_{jk} E_{i,l} - \delta_{li} E_{k,j}$. It implies that

$$[E_{i,i}, E_{i,l}] = E_{i,l} \quad \forall i < l. \tag{4.1}$$

{eq:Matrices}

This, in turn, implies that the algebra of strictly upper triangular matrices is contained in the derived algebra, that is, $\mathfrak{n}(n, K) \subset [\mathfrak{g}, \mathfrak{g}]$. Now recall from Exercise 4 on Sheet 1 that as vector spaces

$$\mathfrak{g} = \mathfrak{n}(n, K) \oplus \mathfrak{d}(n, K),$$

where $\mathfrak{d}(n, K)$ is the algebra of diagonal matrices. Note that this algebra is abelian. Also note the following equality:

$$[\mathfrak{d}(n, K), \mathfrak{n}(n, K)] = \mathfrak{n}(n, K),$$

which follows from Equation (4.1). Therefore,

$$[\mathfrak{g}, \mathfrak{g})] \subset \mathfrak{n}(n, K),$$

so, summarising, $\mathfrak{g}^{(1)} = \mathfrak{n}(n, K)$.

We will from now on work inside $\mathfrak{n}(n, K)$. Define the level of $E_{i,j}$ to be $j - i$. Now, if $l \neq i$, then

$$[E_{i,j}, E_{k,l}] = E_{i,l} \quad \text{if } j = k \quad \text{and} \quad 0 \quad \text{otherwise}.$$

So if $i < j$ and $k < l$, then every $E_{i,l}$ is the commutator of two matrices whose levels add up to those of $E_{i,l}$. It follows that $\mathfrak{g}^{(2)}$ is spanned by those $E_{i,j}$ of level at least 2, $\mathfrak{g}^{(i)}$ by those of level at least 2^{i-1}. In particular, $\mathfrak{g}^{(i)} = 0$ whenever $2^{i-1} > n - 1$.

Let us now record some observations concerning solvability.

{p:PropSolv}

Proposition 4.3. *Let $\mathfrak{g}$ be a Lie algebra.*

(1) *If $\mathfrak{g}$ is solvable, then any subalgebra and any homomorphic image of $\mathfrak{g}$ is solvable.*
(2) *If I is a solvable ideal in $\mathfrak{g}$ such that $\mathfrak{g}/I$ is solvable, then $\mathfrak{g}$ is solvable.*
(3) *If I and J are solvable ideals of $\mathfrak{g}$, then $I + J$ is solvable as well.*

Proof. (1) If $\mathfrak{h} \subset \mathfrak{g}$ is a subalgebra, then by definition $\mathfrak{h}^{(i)} \subset \mathfrak{g}^{(i)}$. If $\varphi\colon \mathfrak{g} \twoheadrightarrow \mathfrak{g}'$ is a surjective homomorphism, then by induction,

$$\begin{aligned}\varphi(\mathfrak{g}^{(i)}) &= \varphi[\mathfrak{g}^{(i-1)}, \mathfrak{g}^{(i-1)}] = [\varphi(\mathfrak{g}^{(i-1)}), \varphi(\mathfrak{g}^{(i-1)})] \\ &= [\mathfrak{g}'^{(i-1)}, \mathfrak{g}'^{(i-1)}] = \mathfrak{g}'^{(i)}.\end{aligned}$$

(2) Assume that $(\mathfrak{g}/I)^{(n)} = 0$. Consider the canonical surjective homomorphism $\pi\colon \mathfrak{g} \twoheadrightarrow \mathfrak{g}/I$. Applying (1) to it, we get $\pi(\mathfrak{g}^{(n)}) = (\mathfrak{g}/I)^{(n)} = 0$. This means that $\mathfrak{g}^{(n)} \subset I = \ker(\pi)$. If $I^{(m)} = 0$, then clearly $(\mathfrak{g}^{(n)})^{(m)} = \mathfrak{g}^{(n+m)} \subset I^{(m)} = 0$.
(3) By Remark 3.5(4), $(I + J)/J \simeq I/(I \cap J)$. The right-hand side is a homomorphic image of the solvable ideal I, hence solvable by (1). But J is solvable by assumption, hence so is $I + J$ by (2).

□

Note that the sum of solvable ideals is a solvable ideal by Proposition 4.3(3). In particular, we can consider the sum of all solvable ideals in $\mathfrak{g}$. This is a solvable ideal. This leads to the following

Definition. Assume that $\dim_K \mathfrak{g} < \infty$. Then there is a unique maximal solvable ideal in $\mathfrak{g}$, called the *radical* of $\mathfrak{g}$ and denoted by $\operatorname{rad} \mathfrak{g}$.

Convention. From now on we will only consider finite dimensional Lie algebras. In particular, the radical will always exist.

Definition. A Lie algebra $\mathfrak{g}$ is called *semisimple* if $\operatorname{rad} \mathfrak{g} = 0$.

{e:ExSemisim}

Example 4.4. A simple algebra is semisimple, since it has no non-trivial ideals, hence $\operatorname{rad} \mathfrak{g}$ is either 0 or $\mathfrak{g}$ and it cannot be $\mathfrak{g}$, because $\mathfrak{g}$ is not solvable by Example 4.1.

The zero algebra is of course also semisimple.

For a less trivial example, consider the Lie algebra $\mathfrak{g}/\operatorname{rad} \mathfrak{g}$, which is semisimple. Indeed, if $\operatorname{rad}(\mathfrak{g}/\operatorname{rad} \mathfrak{g}) \neq 0$, then its preimage is a solvable ideal J (J is solvable, since $J/\operatorname{rad} \mathfrak{g} \simeq \operatorname{rad}(\mathfrak{g}/\operatorname{rad} \mathfrak{g})$, so we can use Proposition 4.3(2)) strictly contaning $\operatorname{rad} \mathfrak{g}$ which is a contradiction.

{r:SolvAlg}

Remark 4.5. Let $\mathfrak{g}$ be a Lie algebra. Then $\mathfrak{g}$ is solvable if and only if there exists a sequence of subalgebras

$$\mathfrak{g} \supset I_1 \supset I_2 \supset \ldots \supset I_m = 0$$

such that $I_{i+1} \subset I_i$ is an ideal for all possible i and I_i/I_{i+1} is an abelian Lie algebra for all possible i.

The proof is left to the reader, see Exercise 4 on Sheet 3.

{r:ExSeqSolv}

Remark 4.6. A *short exact sequence* of Lie algebras is a sequence of Lie algebra homomorphisms

$$\mathfrak{g}' \xrightarrow{\varphi} \mathfrak{g} \xrightarrow{\psi} \mathfrak{g}''$$

such that φ is injective, ψ is surjective and $\ker(\psi) = \mathrm{im}(\varphi)$.

In this case $\mathfrak{g}$ is solvable if and only if $\mathfrak{g}'$ and $\mathfrak{g}''$ are solvable. Indeed, by assumption, $\mathfrak{g}/\mathfrak{g}' \simeq \mathfrak{g}''$, so we can use Proposition 4.3(2) for the "if"-direction and 4.3(1) for the "only if"-direction.

Having had a look at solvability, we now move on to a different concept.

Definition. Let $\mathfrak{g}$ be a Lie algebra. The *descending central series* or the *lower central series* is the following sequence of ideals in $\mathfrak{g}$:

$$\mathfrak{g}^0 = \mathfrak{g},\ \mathfrak{g}^1 = [\mathfrak{g}, \mathfrak{g}],\ \mathfrak{g}^2 = [\mathfrak{g}, \mathfrak{g}^1], \dots, \mathfrak{g}^i = [\mathfrak{g}, \mathfrak{g}^{i-1}], \dots$$

A Lie algebra $\mathfrak{g}$ is called *nilpotent* if $\mathfrak{g}^n = 0$ for some n.

Again, the fact that the members of the lower central series are indeed ideals follows by a straightforward computation.

{r:NilpLie}

Remark 4.7.
(1) Any abelian Lie algebra is nilpotent.
(2) Any nilpotent algebra is solvable, since $\mathfrak{g}^{(i)} \subset \mathfrak{g}^i$ for all i. Also note that $\mathfrak{g}^{(1)} = \mathfrak{g}^1$.

{e:NilpLie}

Example 4.8. The algebra $\mathfrak{g} = \mathfrak{n}(n, K)$ is easily seen to be nilpotent. Namely, $\mathfrak{g}^1$ is spanned by the matrices $E_{i,j}$ of level ≥ 2, $\mathfrak{g}^2$ by those of level ≥ 3 and so on.

Recall that $\mathfrak{h} = \mathfrak{t}(n, K)$ is a solvable algebra, by Example 4.2. But $\mathfrak{h}$ is not nilpotent, since $\mathfrak{h}^2 = [\mathfrak{h}, \mathfrak{h}^1] = \mathfrak{h}^1 = \mathfrak{h}$ and, consequently, $\mathfrak{h}^n = \mathfrak{h}$ for all $n \geq 1$.

We now collect some properties of nilpotency.

{p:NilpProp}

Proposition 4.9. *Let $\mathfrak{g}$ be a Lie algebra.*
(1) *If $\mathfrak{g}$ is nilpotent, then any subalgebra and homomorphic image of $\mathfrak{g}$ is nilpotent.*
(2) *If $\mathfrak{g}/Z(\mathfrak{g})$ is nilpotent, then so is $\mathfrak{g}$.*
(3) *If $\mathfrak{g}$ is nilpotent and $0 \neq \mathfrak{g}$, then $Z(\mathfrak{g}) \neq 0$.*

Proof.
(1) The proof is the same as that of Proposition 4.3(1).
(2) Writing out what it means that $\mathfrak{g}/Z(\mathfrak{g})$ is nilpotent, shows that $\mathfrak{g}^n \subset Z(\mathfrak{g})$ for some n. Then $\mathfrak{g}^{n+1} = [\mathfrak{g}, \mathfrak{g}^n] \subset [\mathfrak{g}, Z(\mathfrak{g})] = 0$.
(3) Say $\mathfrak{g}^n = 0$ (by assumption $n \geq 1$), while $\mathfrak{g}^{n-1} \neq 0$, then $[\mathfrak{g}, \mathfrak{g}^{n-1}] = 0$, which means that $\mathfrak{g}^{n-1} \subset Z(\mathfrak{g})$.

□

{r:NilpExSeq}

Remark 4.10. Let $\mathfrak{g}' \to \mathfrak{g} \to \mathfrak{g}''$ be an exact sequence of Lie algebras. If $\mathfrak{g}'$ and $\mathfrak{g}''$ are nilpotent, then $\mathfrak{g}$ need not be nilpotent. For example, consider the two-dimensional complex Lie algebra $\mathfrak{g} = \mathrm{span}_{\mathbb{C}}(x, y)$ with the commutator $[x, y] = x$ (it is left to the reader to check that this is indeed a Lie algebra; compare Exercise 2 on Sheet 2). Clearly,

$\mathfrak{g}^{(1)} = \mathfrak{g}^1 = \mathrm{span}(x) =: \mathfrak{g}_1$, so $\mathfrak{g}^{(2)} = 0$, hence $\mathfrak{g}$ is solvable. But $\mathfrak{g}^i = \mathfrak{g}_1$ for all $i \geq 1$, hence $\mathfrak{g}$ is not nilpotent. Setting $\mathfrak{g}_2 = \mathrm{span}_{\mathbb{C}}(y)$ gives a short exact sequence

$$\mathfrak{g}_1 \longrightarrow \mathfrak{g} \longrightarrow \mathfrak{g}_2.$$

It remains to note that $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are abelian, hence in particular nilpotent.

We now come to a very important concept. Let $\mathfrak{g}$ be a Lie algebra and $x \in \mathfrak{g}$. Define the map $\mathrm{ad}_x \colon \mathfrak{g} \longrightarrow \mathfrak{g}$ by sending y to $\mathrm{ad}_x(y) := [x, y]$. Another typical notation is $\mathrm{ad}\, x$ instead of ad_x. Let us check that this gives a homomorphism of Lie algebras $\mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g})$:

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

In fact, $\mathrm{ad}_x \in \mathrm{Der}(\mathfrak{g})$ for all x. Indeed,

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

{r:CentAd}

Remark 4.11. Consider the homomorphism $\mathrm{ad} \colon \mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g})$. Its kernel is the set of elements $x \in \mathfrak{g}$ such that $[x, y] = 0$ for all $y \in \mathfrak{g}$. In other words, $\ker(\mathrm{ad}) = Z(\mathfrak{g})$. In particular, if $\mathfrak{g}$ is simple, then $Z(\mathfrak{g}) = 0$ by Remark 3.8, so ad is an injective map, hence a simple Lie algebra is linear.

Definition. Let $\mathfrak{g}$ be a Lie algebra. An element $x \in \mathfrak{g}$ is called *ad-nilpotent* if $\mathrm{ad}_x \colon \mathfrak{g} \longrightarrow \mathfrak{g}$ is a nilpotent endomorphism, that is, there exists an $n \geq 1$ such that $(\mathrm{ad}_x)^n = 0$.

{r:NilpAd}

Remark 4.12. Let $\mathfrak{g}$ be a nilpotent algebra. Then there exists some n such that

$$\mathrm{ad}\, x_1 \circ \ldots \circ \mathrm{ad}\, x_n(y) = 0 \quad \forall x_i, y \in \mathfrak{g}.$$

Also note that any element of a nilpotent Lie algebra is ad-nilpotent. Indeed, if $y \in \mathfrak{g}^i$, then $\mathrm{ad}_x(y) = [x, y] \in \mathfrak{g}^{i+1}$. By induction, $(\mathrm{ad}_x)^k(y) \in \mathfrak{g}^{i+k}$. Hence, if $\mathfrak{g}^n = 0$, then $(\mathrm{ad}_x)^n = 0$.

Our next goal is to prove the following result.

{t:Engel}

Theorem 4.13 (Engel)**.** *If all elements in a Lie algebra $\mathfrak{g}$ are ad-nilpotent, then $\mathfrak{g}$ is nilpotent.*

Before proving this result, let us establish a lemma which can be used to prove that $\mathfrak{n}(n, K)$ is nilpotent.

{l:nilp}

Lemma 4.14. *Let $x \in \mathfrak{gl}(V)$ be a nilpotent endomorphism. Then $\mathrm{ad}\, x$ is also nilpotent.*

Proof. Define $\lambda_x \colon \mathrm{End}(V) \longrightarrow \mathrm{End}(V)$ by sending y to xy (composition of endomorphisms). Furthermore, set $\rho_x \colon \mathrm{End}(V) \longrightarrow \mathrm{End}(V)$, $y \longmapsto yx$. Clearly, λ_x and ρ_x commute. If x is nilpotent, then so are λ_x and ρ_x. The binomial formula $(a+b)^n = \sum_{i=1}^{n} \binom{n}{i} a^i b^{n-i}$

holds for commuting elements of any ring, hence $\lambda_x - \rho_x$ is nilpotent. Now note that $\operatorname{ad} x = \lambda_x - \rho_x$, since $\operatorname{ad} x(y) = xy - yx = \lambda_x(y) - \rho_x(y)$. $\square$

Now, clearly any matrix x in $\mathfrak{n}(n, K)$ is nilpotent, so $\operatorname{ad} x$ is nilpotent by the lemma, hence $\mathfrak{n}(n, K)$ is nilpotent by Engel's Theorem.

Remark 4.15. A matrix can be ad-nilpotent in $\mathfrak{gl}(n, K)$ without being nilpotent. For instance, the identity matrix has this property. {r:NilpAdnilp}

For the next result we need the following

Definition. The *normaliser* of a subspace $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is defined to be

$$N_{\mathfrak{g}}(\mathfrak{h}) := \{x \in \mathfrak{g} \mid [x, \mathfrak{h}] \subset \mathfrak{h}\}.$$

Note that if $\mathfrak{h}$ is a subalgebra, then $N_{\mathfrak{g}}(\mathfrak{h})$ is a subalgebra of $\mathfrak{g}$. Indeed, if $x, y \in N_{\mathfrak{g}}(\mathfrak{h})$, then by definition $[x, h] \in \mathfrak{h}$ and $[y, h] \in \mathfrak{h}$ for all $h \in \mathfrak{h}$. Then $[[x, y], h] = -[h, [x, y]] = [x, [y, h]] + [y, [h, x]] \in \mathfrak{h}$.

Lemma 4.16. *Let $0 \neq V$ be a finite-dimensional K-vector space and let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$. If $\mathfrak{g}$ consists of nilpotent endomorphisms, then there exists a $0 \neq v \in V$ such that $\mathfrak{g}.v = 0$, that is, $x(v) = 0$ for all $x \in \mathfrak{g}$.* {l:Engel}

Proof. First of all recall that a nilpotent linear map has at least one eigenvector corresponding to its unique eigenvalue 0. This gives the induction hypothesis if we perform induction over $\dim(\mathfrak{g})$.

Let $\mathfrak{h} \subsetneq \mathfrak{g}$ be any subalgebra. Note that by Lemma 4.14 the algebra $\mathfrak{h}$ acts via the map ad as a Lie algebra of nilpotent linear transformations on the vector space $\mathfrak{g}$. But then it also acts on $\mathfrak{g}/\mathfrak{h}$. Since $\dim(\mathfrak{h}) < \dim(\mathfrak{g})$, by the induction hypothesis we get the existence of a vector $x + \mathfrak{h} \neq \mathfrak{h} \in \mathfrak{g}/\mathfrak{h}$ which is killed by the image of $\mathfrak{h}$ in $\mathfrak{gl}(\mathfrak{g}/\mathfrak{h})$. In other words, $[y, x] \in \mathfrak{h}$ for all $y \in \mathfrak{h}$, while by construction $x \notin \mathfrak{h}$. This means that $\mathfrak{h} \subsetneq N_{\mathfrak{g}}(\mathfrak{h})$.

Now take $\mathfrak{h}$ to be a maximal proper subalgebra of $\mathfrak{g}$. By the previous paragraph, $N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{g}$, hence $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. If $\dim(\mathfrak{g}/\mathfrak{h}) > 1$, then the inverse image of a one-dimensional subalgebra of $\mathfrak{g}/\mathfrak{h}$ (take Kw for any $w \in \mathfrak{g}/\mathfrak{h}$) would be a proper subalgebra properly contaning $\mathfrak{h}$, a contradiction. Therefore, $\dim(\mathfrak{h}) = \dim(\mathfrak{g}) - 1$. Thus, we can write $\mathfrak{g} = \mathfrak{h} + Ku$ for any $u \in \mathfrak{g} \setminus \mathfrak{h}$.

By induction, $W = \{v \in V \mid \mathfrak{h}.v = 0\}$ is not zero. If $x \in \mathfrak{g}$, $y \in \mathfrak{h}$ and $w \in W$, then $y(w) = 0$ and $[x, y](w) = 0$, because $\mathfrak{h}$ is an ideal. Therefore, $y(x(w)) = x(y(w)) - [x, y]w = 0$. Therefore, W is stable under $\mathfrak{g}$. Choose $u \in \mathfrak{g} \setminus \mathfrak{h}$ as above. The nilpotent endomorphism u acts on W and has an eigenvector $0 \neq v_0 \in W$ such that $u(v_0) = 0$. Therefore, $\mathfrak{g}.v_0 = 0$. $\square$

Proof of Engel's theorem. Assume that $\mathfrak{g} \neq 0$ (the case $\mathfrak{g} = 0$ is trivial) and consider the Lie algebra $\operatorname{ad} \mathfrak{g} \subset \mathfrak{gl}(\mathfrak{g})$. By assumption, Lemma 4.16 applies to $\operatorname{ad} \mathfrak{g}$. Therefore, there exists an element $0 \neq x \in \mathfrak{g}$ such that $\operatorname{ad} \mathfrak{g}.x = [\mathfrak{g}, x] = 0$. In other words, $Z(\mathfrak{g}) \neq 0$. Now note that $\mathfrak{g}/Z(\mathfrak{g})$ consists of ad-nilpotent elements and has cmaller dimension than $\mathfrak{g}$. By induction, $\mathfrak{g}/Z(\mathfrak{g})$ is nilpotent, hence $\mathfrak{g}$ is as well, by Proposition 4.9(2). $\square$

Definition. Let V be a finite dimensional vector space. A *flag* in V is a chain of subspaces

$$0 = V_0 \subset V_1 \subset \ldots \subset V_n = V, \quad \dim V_i = i.$$

If $x \in \operatorname{End}(V)$, then we say that x *stabilises* a flag if $x(V_i) \subset V_i$ for all i.

{c:Engel}

Corollary 4.17. *Let $0 \neq V$ be a finite-dimensional K-vector space and let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$. If $\mathfrak{g}$ consists of nilpotent endomorphisms, then there exists a flag (V_i) in V such that $x.V_i \subset V_{i-1}$ for all i and $x \in \mathfrak{g}$. In other words, there exists a basis of V relative to which the matrices of $\mathfrak{g}$ are all in $\mathfrak{n}(n, K)$.*

Proof. Let $v \in V$ be any nonzero element killed by $\mathfrak{g}$ which exists by Lemma 4.16. Set $V_1 = Kv$. Let $W = V/V_1$ and note that the action of $\mathfrak{g}$ on W is also by nilpotent endomorphisms. By induction on $\dim(V)$, W has a flag which is stabilised by $\mathfrak{g}$. Taking the inverse image of this flag gives the required flag in V. □

5. Representations of Lie algebras

Definition. Let $\mathfrak{g}$ be a Lie algebra. A *representation* of $\mathfrak{g}$ is a homomorphism $\rho\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$, where V is a vector space over K.

{e:ReprLie}

Example 5.1. (1) Probably the most important representation is the *adjoint representation*: It is the homomorphism

$$\operatorname{ad}\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g}), \quad x \longmapsto \operatorname{ad}_x .$$

Recall that we have already checked that this map is indeed a homomorphism. In fact, ad_x is a derivation for all $x \in \mathfrak{g}$.

(2) If $\mathfrak{g}$ and V are arbitrary, then the zero map $\mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ is a representation of $\mathfrak{g}$.

(3) The *trivial representation* of a Lie algebra $\mathfrak{g}$ is the base field K together with the zero map $\mathfrak{g} \longrightarrow \mathfrak{gl}(K)$.

Definition. Let $\mathfrak{g}$ be a Lie algebra. A $\mathfrak{g}$-*module* is a vector space V endowed with an operation $\mathfrak{g} \times V \longrightarrow V$ denoted by $(x, v) \longmapsto x.v = xv = x(v)$ satisfying the following conditions for all $x, y \in \mathfrak{g}$, $a, b \in K$ and $v, w \in V$:

(M1) $(ax + by).v = a(x.v) + b(y.v)$;
(M2) $x.(av + bw) = a(x.v) + b(x.w)$;
(M3) $[x, y].v = x.y.v - y.x.v$.

Note that (M1) and (M2) just say that the operation is bilinear.

{r:RepMod}

Remark 5.2. If V is a $\mathfrak{g}$-module, then $\rho\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$, $x \longmapsto x(-)$ is a representation. Indeed, this map is well-defined by (M2), linear by (M1) and a homomorphism of Lie algebras by (M3).

Conversely, if $\rho\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ is a representation, then setting $x.v := \rho(x)(v)$ defines a $\mathfrak{g}$-module structure on V. Since ρ maps to $\mathfrak{gl}(V)$, the map $x(-)$ is linear for all x, hence (M2) holds. Item (M1) holds since ρ is linear, while (M3) holds, since ρ is a homomorphism.

Thus, there is a one-to-one correspondence between $\mathfrak{g}$-modules and representations of $\mathfrak{g}$.

Definition. If $\rho\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ and $\rho'\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(W)$ are two representations, a *morphism of representations* (sometimes called *intertwiner*) is a linear map $\phi\colon V \longrightarrow W$ such that $\phi(\rho(x)(v)) = \rho'(x)(\phi(v))$ for all $v \in V$. If ϕ is an isomorphism of vector spaces, then we call it an *isomorphism* of representations.

A representation $\rho\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ is called *faithful* if ρ is injective.

A *homomorphism* of $\mathfrak{g}$-modules is a linear map $\phi\colon V \longrightarrow W$ such that $\phi(x.v) = x.\phi(v)$. This is, of course, just a reformulation of the above definition.

If V and W are representations of $\mathfrak{g}$, we will sometimes write $\mathrm{Hom}_{\mathfrak{g}}(V, W)$ for the space of morphisms between them. In particular, $\mathrm{End}_{\mathfrak{g}}(V)$ is then defined.

Definition. If V is a $\mathfrak{g}$-module and $U \subset V$ is a vector subspace, we call U a *submodule* if $x.v \in U$ for all $x \in \mathfrak{g}$ and $v \in U$.

A $\mathfrak{g}$-module $0 \neq V$ is called *irreducible* or *simple* if it has precisely two $\mathfrak{g}$-submodules, namely 0 and V.

If V and W are two $\mathfrak{g}$-modules, then the *direct sum* $V \oplus W$ is a $\mathfrak{g}$-module, with respect to the operation $x.(v, w) := (x.v, x.w)$.

A $\mathfrak{g}$-module V is called *completely reducible* or *semisimple* if it is a direct sum of irreducible $\mathfrak{g}$-modules.

We leave it to the reader to rewrite the notions introduced in the definition using the language of representations.

{e:SubRepr}

Example 5.3. If $\phi\colon V \longrightarrow W$ is a morphism of representations, then $\ker(\phi)$ and $\mathrm{im}(\phi)$ are also representations.

An ideal of a Lie algebra $\mathfrak{g}$ is simply a $\mathfrak{g}$-submodule of the adjoint representation $\mathrm{ad}\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g})$.

{r:RemRepr}

Remark 5.4. Note that the Lie algebra $\mathfrak{g} = K$ is simple as a $\mathfrak{g}$-module but not simple as a Lie algebra, since it is abelian.

Also note that the adjoint representation of a simple Lie algebra is faithful. More generally, since $\ker(\mathrm{ad}) = Z(\mathfrak{g})$, the adjoint representation is faithful if and only if the center of $\mathfrak{g}$ is trivial. In particular, if $0 \neq \mathfrak{g}$ is abelian, then its adjoint representation is not faithful. As a last remark, recall that the center of a non-trivial nilpotent algebra is not trivial by Proposition 4.9(3), hence, for instance, the adjoint representation of $\mathfrak{n}(n, K)$ is not faithful.

We already have seen that we can define direct sums of representations. Here are some other operations.

Definition. Let V be a $\mathfrak{g}$-module and U be a submodule. Then there exists a $\mathfrak{g}$-module structure on V/U defined by $x.(v + U) := x.v + U$. This is called the *quotient* $\mathfrak{g}$*-module*. This is the unique module structure such that $\pi\colon V \longrightarrow V/U$ is a homomorphism of modules.

The *dual* $\mathfrak{g}$-*module* is the vector space V^* endowed with the following module structure: if $f \in V^*$, $v \in V$ and $x \in \mathfrak{g}$, then set $(x.f)(v) := -f(x.v)$.

Let us check that the definition of the dual $\mathfrak{g}$-module makes sense. Clearly, (M1) and (M2) are satisfied. Let us check (M3):

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

Recall that if V and W are vector spaces, their tensor product $V \otimes W$ is the quotient of the vector space Z with basis $V \times W$ by bilinear relations. The image of (v, w) in $V \otimes W$ is denoted by $v \otimes w$. Hence, $V \otimes W$ is generated by elements of the form $v \otimes w$ with $v \in V$ and $w \in W$. If V has a basis $(x_1, \dots, x_n)$ and W has a basis $(y_1, \dots, y_m)$, then $V \otimes W$ has as basis the vectors $v_i \otimes w_j$, $1 \le i \le n$, $1 \le j \le m$. In particular, if $\dim_K(V) = n$ and $\dim_K(W) = m$, then $\dim_K(V \otimes W) = mn$.

Definition. Let V and W be two $\mathfrak{g}$-modules. Their *tensor product* is the vector space $V \otimes W$ endowed with the operation $\mathfrak{g} \times (V \otimes W) \longrightarrow V \otimes W$, $x.(v \otimes w) \longmapsto x.v \otimes w + v \otimes x.w$.

Yet again, let us check that this is well-defined. Axioms (M1) and (M2) are easy and left to the reader. As for (M3):

$$\begin{aligned}[x,y].(v \otimes w) &= [x,y].v \otimes w + v \otimes [x,y].w\\ &= (x.y.v - y.x.v) \otimes w + v \otimes (x.y.w - y.x.w)\\ &= (x.y.v \otimes w + v \otimes x.y.w) - (y.x.v \otimes w + v \otimes y.x.w)\\ &= x.(y.v \otimes w + v \otimes y.w) - y.(x.v \otimes w + v \otimes x.w)\\ &= (x.y).(v \otimes w) - (y.x).(v \otimes w) = (xy - yx)(v \otimes w).\end{aligned}$$

Now recall that if $\dim_K(V) = n < \infty$, then $V^* \otimes V \longrightarrow \mathrm{End}(V)$, $f \otimes v \longmapsto [w \longmapsto f(w)v]$ is an isomorphism of vector spaces (compare below for a more general statement). One way to see this is by using a basis $x_1, \dots, x_n$ of V, the corresponding dual basis $x_1^*, \dots, x_n^*$ of V^* and check that the given map is then surjective, hence injective as well, since both spaces have the same dimension.

Note that if V is a $\mathfrak{g}$-module, the above isomorphism endows the space $\mathrm{End}(V)$ with a $\mathfrak{g}$-module structure. More generally, there is the following

Definition. Let V and W be two finite-dimensional $\mathfrak{g}$-modules. The space of linear maps $\mathrm{Hom}(V, W)$ has the structure of a $\mathfrak{g}$-module by setting $(x.f)(v) := x.f(v) - f(x.v)$.

In fact, linear algebra tells us that $V^* \otimes W$ and $\mathrm{Hom}(V, W)$ are isomorphic via the map which sends $f \otimes w \in V^* \otimes W$ to the linear map $v \longmapsto f(v)w$. One can check that the formula in the definition is precisely the one arising from the $\mathfrak{g}$-module structure on $V^* \otimes W$ defined above.

Lemma 5.5. *Let $\mathfrak{g}$ be a Lie algebra, let U, V, W be $\mathfrak{g}$-modules and let $\alpha\colon U \longrightarrow V$, $\beta\colon V \longrightarrow W$ and $\gamma\colon U \longrightarrow W$ be module homomorphisms. Then* {l:LemSchur}

(1) *If U is irreducible, then α is either 0 or injective.*
(2) *If W is irreducible, then β is either 0 or surjective.*
(3) *If U and W are irreducible, then $\gamma = 0$ or an isomorphism.*

Proof. (1) We know by Example 5.3 that $\ker(\alpha)$ is a submodule of U. Hence, either $\ker(\alpha) = 0$ or $\ker(\alpha) = U$.
(2) Use that $\mathrm{im}(\beta)$ is a submodule of W.
(3) Combine (1) and (2).

□

{r:DivAlg}

Remark 5.6. An associative unital algebra A over a field K is a *division algebra* if every element $0 \neq x$ has a multiplicative inverse. Thus, Lemma 5.5(3) in particular says that if U is simple, then $\mathrm{End}_{\mathfrak{g}}(U)$ is a division algebra. If U is finite dimensional and K is algebraically closed, then $\mathrm{End}_{\mathfrak{g}}(U)$ is already isomorphic to K. In other words, any intertwiner $\phi\colon U \longrightarrow U$ is a scalar multiple of the identity.

Later we will classify the simple finite dimensional representations of $\mathfrak{sl}(2, K)$ (if K is algebraically closed and of characteristic zero). This classification plays an important role in the general theory.

6. Jordan decomposition

Convention. Frow now on the base field K is assumed to be algebraically closed.

Let V be a finite dimensional K-vector space and $x \in \mathrm{End}(V)$ be an endomorphism. Recall from linear algebra that one can represent x in the Jordan normal form. Namely, there exists a basis of V with respect to which x consists of blocks of the following form

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ & & \ddots & \ddots & & \\ & & & & & \\ & & & & & 1 \\ 0 & \dots & \dots & \dots & \dots & \lambda \end{pmatrix}.$$

Note that the above matrix is the sum of a nilpotent matrix and a scalar multiple of the identity matrix. In particular, both commute.

Definition. Let V be a finite dimensional K-vector space and $x \in \mathrm{End}(V)$ be an endomorphism. We say that x is *semisimple* if and only if x is diagonalisable.

{r:DiagRoots}

Remark 6.1. Recall that x is diagonalisable if and only if the roots of its minimal polynomial over K are all distinct.

Note that if $W \subset V$ is a subspace and a semisimple x maps W into W, then the rectriction of x to W is semisimple.

One can extend the notion of semisimplicity of an endomorphism to arbitary fields as follows. Namely, if K is not necessarily algebraically closed, then write $\overline{K}$ for its algebraic closure. We then call $x \in \mathrm{End}(V)$ semisimple if the map

$$\overline{x}\colon \overline{K} \otimes_K V \longrightarrow \overline{K} \otimes_K V, \quad , \lambda \otimes v \longmapsto \lambda \otimes x(v)$$

is diagonalisable.

{p:Jordan}

Proposition 6.2. *Let V be a finite dimensional K-vector space and $x \in \mathrm{End}(V)$ be an endomorphism.*

(1) *There exist unique $x_s, x_n \in \mathrm{End}(V)$ satisfying the conditions $x_s + x_n = x$, $x_s x_n = x_n x_s$ and x_s is semisimple, x_n is nilpotent.*
(2) *There exist polynomials $p(T)$, $q(T)$ with $p(0) = q(0) = 0$ such that $x_s = p(x)$ and $x_n = q(x)$. In particular, x_s and x_n commute with any endomorphism y which satisfies $xy = yx$.*
(3) *If $U \subset W \subset V$ are subspace and x maps W into U, then x_s and x_n also map W into U.*

The decomposition $x = x_s + x_n$ is called the *Jordan-Chevalley decomposition* or the *Jordan decomposition* of x. The endomorphisms x_s and x_n are called the *semisimple part* and the *nilpotent part* of x, respectively.

Proof. Let $\chi(T) = \prod_{i=1}^{k}(T - a_i)^{m_i}$ be the characteristic polynomial of x. If we set $V_i = \ker(x - a_i \cdot \mathrm{id})^{m_i}$, then

$$V \simeq V_1 \oplus \dots V_k$$

and on V_i the endomorphism x has characteristic polynomial $(T - a_i)^{m_i}$.

Apply the Chinese Remainder Theorem to $K[T]$ to find a polynomial p with the following properties

$$p(T) \equiv a_i \pmod{(T - a_i)^{m_i}}$$
$$p(T) \equiv 0 \pmod{T}.$$

Define $q(T) = T - p(T)$. Since $p(T) \equiv 0$ modulo T, both q and p have no constant term.

Set $x_s = p(x)$ and $x_n = q(x)$. Then $x_s x_n = x_n x_s$, since both are polynomials in x and, for the same reason, if some endomorphism y commutes with x, it also commutes with x_s and x_n. Furthermore, x_s and x_n stabilise the spaces V_i (since x does). Since $p(T) \equiv a_i \pmod{(T - a_i)^{m_i}}$, the restriction of $x_s - a_i \cdot \mathrm{id}$ to V_i is zero for all i, hence x_s acts diagonally on V_i with single eigenvalue a_i. Since $x_n = x - x_s$, this shows that x_n is nilpotent.

If $x(W) \subset U$, then $x^i(W) \subset U$ for all $i \geq 1$, hence x_s and x_n map W into U, because they are defined by the polynomials p and q which do not have a constant term.

So, we have proved all of (1)-(3), except for the uniqueness statement in (1). Let $x = s + n$ be another decomposition as in (1). Then $x_s - s = x_n - n$. By (2), all these endomorphisms commute. The sum of two commuting nilpotent endomorphisms is again nilpotent (use the binomial formula). Furthermore, the sum of two commuting semisimple endomorphisms is again semisimple, because they can be diagonalised simultaneously. Therefore, $x_s - x = x_n - n$ is both a semisimple and a nilpotent endomorphism, which means that it is zero. This shows $x_s = x$ and $x_n = n$. □

{r:Hau}

Remark 6.3. In more classical terms, x_s is defined by requiring the generalised eigenspace $\mathrm{Hau}(x, \lambda) = \bigcup_{k \geq 0} \ker(x - \lambda)^k$ of x for the eigenvalue λ to be the eigenspace of x_s for λ.

Note that if $x \in \mathrm{End}(V)$, $y \in \mathrm{End}(W)$ and $f\colon V \longrightarrow W$ satisfies $fx = yf$, then $fx_s = y_s f$ and $fx_n = y_n f$ by (2), but this can also be seen since f respects the generalised eigenspaces.

Let us prove a first useful application of the Jordan decomposition.

{l:Jordan}

Lemma 6.4. *Let $x \in \mathrm{End}(V)$ with V finite dimensional and let $x = x_s + x_n$ be its Jordan decomposition. Then $\operatorname{ad} x = \operatorname{ad} x_s + \operatorname{ad} x_n$ is the Jordan decomposition of $\operatorname{ad} x$ in $\mathrm{End}(\mathfrak{gl}(V))$.*

Proof. We are considering the adjoint representation $\mathfrak{gl}(V) \longrightarrow \mathfrak{gl}(\mathfrak{gl}(V))$. If $y \in \mathfrak{gl}(V)$ is nilpotent, then $\operatorname{ad} y$ is also nilpotent by Lemma 4.14.

Now let y be semisimple. Choose a basis $(v_1, \ldots, v_n)$ of V with respect to which y is a diagonal matrix $(a_1, \ldots, a_n)$ (of course, the a_i are not necessarily distinct). Let $E_{i,j}$ be the basis of $\mathfrak{gl}(V)$ such that $E_{i,j}(v_k) = \delta_{jk}(v_i)$. Then

$$\begin{aligned} \operatorname{ad} y(E_{i,j})(v_j) &= [y, E_{i,j}](v_j) = y(E_{i,j}(v_j)) - E_{i,j}(y(v_j)) \\ &= y(v_i) - E_{i,j}(a_j v_j) = (a_i v_i - a_j v_i) = (a_i - a_j)E_{i,j}(v_j), \end{aligned}$$

hence

$$\operatorname{ad} y(E_{i,j}) = (a_i - a_j)E_{i,j} \text{ if } y \text{ is semisimple.} \tag{6.1}$$ {eq:sem}

. This means that with respect to $E_{i,j}$ the endomorphism $\operatorname{ad} y$ is diagonal.

So, if $x = x_s + x_n$, then $\operatorname{ad} x_s$ is semisimple and $\operatorname{ad} x_n$ is nilpotent. Since

$$[\operatorname{ad} x_s, \operatorname{ad} x_n] = \operatorname{ad}[x_s, x_n] = 0,$$

these two endomorphisms commute. The claim follows by Proposition 6.2(1). □

7. The theorems of Lie and Cartan

Convention. Frow now on the base field K is assumed to be of characteristic zero and algebraically closed, unless stated otherwise.

{t:Eigenv}

Theorem 7.1. *Let K be an algebraically closed field of characteristic zero. Let $\mathfrak{g}$ be a solvable subalgebra of $\mathfrak{gl}(V)$ for some finite dimensional vector space V. If $0 \neq V$, then V contains a common eigenvector for all the endomorphisms contained in $\mathfrak{g}$.*

Proof. We will use induction on $\dim(\mathfrak{g})$. Note that the case $\dim(\mathfrak{g}) = 0$ is trivial. The strategy is as follows: (1) find an ideal $\mathfrak{h}$ in $\mathfrak{g}$ of codimension one; (2) show by induction that a common eigenvector exists for $\mathfrak{h}$; (3) verify that the space of $\mathfrak{h}$-eigenvectors is stable under $\mathfrak{g}$ and (4) find in this space an eigenvector for a $u \in \mathfrak{g}$ satisfying $\mathfrak{g} = \mathfrak{h} + Ku$.

So, let $\mathfrak{g}$ be of dimension at least 1. Since $\mathfrak{g}$ is solvable and of positive dimension, $[\mathfrak{g},\mathfrak{g}] \subsetneq \mathfrak{g}$. Now recall that $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$ is an abelian Lie algebra by Corollary 3.7, hence every subspace of it is an ideal. Take a subspace of codimension one. Its preimage is an ideal by Proposition 3.4(4) and it is of course of codimension one in $\mathfrak{g}$. Thus, step (1) is done.

Assume that $\dim_K(\mathfrak{g}) = 1$. Then an eigenvector for a basis vector of $\mathfrak{g}$ shows the statement of the theorem.

So, assume that $\dim_K(\mathfrak{g}) \geq 2$ and let $\mathfrak{h} \subset \mathfrak{g}$ be an ideal of codimension one. By induction, there exists a common eigenvector $v \in V$ for $\mathfrak{h}$. This means that $x.v = \lambda(x).v$ for some function $\lambda\colon \mathfrak{h} \longrightarrow K$. Note that λ is linear by (M1). Define

$$W = \{w \in V \mid x.w = \lambda(x).w \ \forall x \in \mathfrak{h}\}.$$

By what we have seen above, $W \neq 0$.

Now, let $w \in W$ and $x \in \mathfrak{g}$. We want to prove that $x.w \in W$. By definition, this means that we need to check that for any $y \in \mathfrak{h}$ we have $yx.w = \lambda(y)x.w$. Compute

$$yx.w = xy.w - [x,y].w = \lambda(y)x.w - \lambda([x,y]).w, \tag{7.1}$$ {eq:L}

where we used that $[x,y] \in \mathfrak{h}$, since $\mathfrak{h}$ is an ideal and that $y \in \mathfrak{h}$. So it is enough to show that $\lambda([x,y]).w = 0$ to prove our claim.

For w, x as above, let $n > 0$ be the smallest integer for which $\{w, x.w, \ldots, x^n.w\}$ is a linearly dependent set. Set $W_0 = 0$ and $W_i = \operatorname{span}_K(w, x.w, \ldots, x^{i-1}.w)$. In particular, $\dim_K W_n = n$, $W_n = W_{n+i}$ for all $i \geq 0$ and x maps W_n into W_n. Furthermore, one can use Equation (7.1) and the fact that $\mathfrak{h}$ is an ideal to see that

$$(\diamond)\ y.W_i \subset W_i \ \forall y \in \mathfrak{h}.$$

We will next prove the following claim:

$$(*)\ yx^i.w - \lambda(y)x^i.w \in W_i.$$

We use induction on i, the case $i = 0$ being obvious. Now $yx^i.w = yxx^{i-1}.w = xyx^{i-1}.w - [x,y]x^{i-1}.w$. By induction,

$$yx^{i-1}.w - \lambda(y)x^{i-1}.w \in W_{i-1},$$

so $x(yx^{i-1}.w - \lambda(y)x^{i-1}.w) \in W_i$. Since $[x,y] \in \mathfrak{h}$, the element $[x,y]x^{i-1}.w$ is in W_{i-1} by $(\diamond)$. Therefore,

$$\begin{aligned} yx^i.w &= x(yx^{i-1}.w - \lambda(y)x^{i-1}.w + \lambda(y)x^{i-1}.w) - [x,y]x^{i-1}.w \\ &= \lambda(y)x^i.w + x(yx^{i-1}.w - \lambda(y)x^{i-1}.w) - [x,y]x^{i-1}.w, \end{aligned}$$

and we are done proving $(*)$ since the last two summands are in W_i.

Now, $(*)$ means that the action of $y \in \mathfrak{h}$ on W_n with respect to the ordered basis $w, x.w, \ldots, x^{n-1}.w$ is represented by an upper triangular matrix whose diagonal entries equal $\lambda(y)$. Hence, $\operatorname{tr}_{W_n}(y) = n\lambda(y)$ for any $y \in \mathfrak{h}$. In particular, this holds for $[x,y]$, where $y \in \mathfrak{h}$ is arbitrary and $x \in \mathfrak{g}$ is the fixed element from above. Now, both x and y map W_n to itself, hence $[x,y]$ acts on W_n as the commutator of two endomorphisms of W_n, meaning that $\mathrm{tr}_{W_n}([x,y]) = 0 = n\lambda([x,y])$. Since $\mathrm{char}(K) = 0$, this means that $\lambda([x,y]) = 0$. In other words, $\mathfrak{g}$ maps W into W.

To conclude the proof of the theorem, write $\mathfrak{g} = \mathfrak{h} + Ku$. Since K is algebraically closed, there exists an eigenvector $v_0 \in W$ of u. Then v_0 is the common eigenvector for $\mathfrak{g}$ we were looking for. □

We can now prove

{t:LiesThm}

Theorem 7.2 (Lie's theorem)**.** *Let V be an n-dimensional K-vector space and let $\mathfrak{g}$ be a solvable subalgebra of $\mathfrak{gl}(V)$. Then $\mathfrak{g}$ stabilises some flag in V. In other words, with respect to a suitable basis of V the matrices of $\mathfrak{g}$ are all upper triangular.*

Proof. Use Theorem 7.1 and induction on $\dim_K(V)$. □

{c:LiesCor}

Corollary 7.3. *Let $\mathfrak{g}$ be a solvable Lie algebra of dimension n. Then there exists a chain of ideals of $\mathfrak{g}$*

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

such that $\dim_K \mathfrak{g}_i = i$.

Proof. If $\mathfrak{g}$ is any solvable finite dimensional Lie algebra and $\phi\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ is a finite dimensional representation of $\mathfrak{g}$, then $\phi(\mathfrak{g})$ is a solvable algebra by Proposition 4.3(1), hence it stabilises a flag by Theorem 7.2. If we take ϕ to be the adjoint representation, then note that a flag in $V = \mathfrak{g}$ which is stabilised by $\mathfrak{g}$ is just a chain of ideals in $\mathfrak{g}$ by Example 5.3, and each is of codimension 1 in the next. □

Notation. If $\mathfrak{h} \subset \mathfrak{g}$ is a subalgebra and $x \in \mathfrak{h}$, then $\mathrm{ad}_{\mathfrak{h}}\, x$ is considered as an element in $\mathfrak{gl}(\mathfrak{h})$, while $\mathrm{ad}_{\mathfrak{g}}\, x$ is considered as an element in $\mathfrak{gl}(\mathfrak{g})$.

As an example, let x be a diagonal matrix. Consider the inclusion of Lie algebras $\mathfrak{d}(n,K) \subset \mathfrak{gl}(n,K)$. Then $\mathrm{ad}_{\mathfrak{d}(n,K)}(x) = 0$, since x commutes with any diagonal matrix, but $\mathrm{ad}_{\mathfrak{gl}(n,K)}(x)$ need not be zero.

{c:LiesCor2}

Corollary 7.4. *Let $\mathfrak{g}$ be solvable of dimension n. If $x \in [\mathfrak{g},\mathfrak{g}]$, then $\mathrm{ad}\, x$ is nilpotent. In particular, the algebra $[\mathfrak{g},\mathfrak{g}]$ is nilpotent.*

Proof. Find a chain of ideals as in Corollary 7.3. Fix a basis $(x_1,\ldots,x_n)$ of $\mathfrak{g}$ such that $\mathfrak{g}_i = \mathrm{span}_K(x_1,\ldots,x_i)$. With respect to this basis, the matrices of $\mathrm{ad}\,\mathfrak{g}$ are all upper triangular. Therefore, the matrices of $[\mathrm{ad}\,\mathfrak{g}, \mathrm{ad}\,\mathfrak{g}]$ are all in $\mathfrak{n}(n,K)$, as an easy computation shows. Therefore, $\mathrm{ad}\, x$ is nilpotent for $x \in [\mathfrak{g},\mathfrak{g}]$. Hence, $\mathrm{ad}_{[\mathfrak{g},\mathfrak{g}]}\, x$ is nilpotent, so $[\mathfrak{g},\mathfrak{g}]$ is nilpotent by Engel's Theorem 4.13. □

Having proved Lie's theorem and some of its consequences, we move on to Cartan's Criterion which is a criterion for solvability of a Lie algebra. We will require the following

{l:Cartan}

Lemma 7.5. *Let $A \subset B$ be subspaces of $\mathfrak{gl}(V)$, $\dim_K V = m < \infty$. Set $M = \{x \in \mathfrak{gl}(V) \mid \mathrm{ad}\, x(b) \in A\ \forall b \in B\}$. Suppose that $x \in M$ satisfies the equality $\mathrm{tr}(xy) = 0$ for all $y \in M$. Then x is nilpotent.*

Proof. Let $x = x_s + x_n$ be the Jordan decomposition of x. For simplicity, write $x_s = s$ and $x_n = n$. Fix a basis $(v_1, \dots, v_m)$ of V with respect to which s has matrix $\operatorname{diag}(a_1, \dots, a_m)$. Since K has characteristic zero, K is a $\mathbb{Q}$-vector space. Let E be the $\mathbb{Q}$-vector subspace of K spanned by the eigenvalues $a_1, \dots, a_m$. We have to show that $s = 0$ or, equivalently, that $E = 0$. By construction, E is finite dimensional over $\mathbb{Q}$, hence it suffices to show that its dual space E^* is zero. In other words, we will show that any $\mathbb{Q}$-linear function $f\colon E \longrightarrow \mathbb{Q}$ is zero.

So, let $f\colon E \longrightarrow \mathbb{Q}$ be a $\mathbb{Q}$-linear function. Define $y \in \mathfrak{gl}(V)$ to be given by the diagonal matrix $\operatorname{diag}(f(a_1), \dots, f(a_m))$. Recall that $\{E_{i,j}\}$ denotes the usual basis of $\mathfrak{gl}(V)$ and that by Equation (6.1)

$$\operatorname{ad} s(E_{i,j}) = (a_i - a_j)E_{i,j}, \quad \operatorname{ad} y(E_{i,j}) = (f(a_i) - f(a_j))E_{i,j}.$$

Note that if $a_i - a_j = a_k - a_l$, then $f(a_i) - f(a_j) = f(a_k) - f(a_l)$. Therefore, there exists (by Lagrange interpolation, compare Exercise 1 on Sheet 7) a polynomial $r \in K[T]$ without constant term satisfying $r(a_i - a_j) = f(a_i) - f(a_j)$. By construction, $\operatorname{ad} y = r(\operatorname{ad} s)$.

Recall that by Lemma 6.4, $\operatorname{ad} s$ is the semisimple part of $\operatorname{ad} x$. By Proposition 6.2(2), it can be written as a polynomial in $\operatorname{ad} x$ without constant term. Since $\operatorname{ad} y = r(\operatorname{ad} s)$, $\operatorname{ad} y$ is also a polynomial in $\operatorname{ad} x$ without constant term. By hypothesis, $\operatorname{ad} x$ maps B into A, so the same holds for $\operatorname{ad} y$. In other words, $y \in M$. By our assumption,

$$0 = \operatorname{tr}(xy) = \sum_{i=1}^{m} f(a_i)a_i.$$

The right-hand side is an element in E. Applying f to it, we get $\sum_{i=1}^{m} f(a_i)^2 = 0$. But $f(a_i) \in \mathbb{Q}$ for all i, hence $f(a_i) = 0$ for all i. Since the a_i span E, $f = 0$. Therefore, $E^* = 0$, so $E = 0$, hence $s = 0$ and $x = n$ is nilpotent as claimed. ☐

{l:TrThree}

Lemma 7.6. *Let V be a finite dimensional vector space over an arbitrary field. If $x, y, z \in \operatorname{End}(V)$, then*

$$\operatorname{tr}([x, y]z) = \operatorname{tr}(x[y, z]).$$

Proof. We have $[x, y]z = xyz - yxz$, $x[y, z] = xyz - xzy$. Now apply the well-known equality $\operatorname{tr}(ab) = \operatorname{tr}(ba)$ to $a = y$, $b = xz$. ☐

{l:DANilpSo}

Lemma 7.7. *Let $\mathfrak{g}$ be a finite dimensional Lie algebra. If $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent, then $\mathfrak{g}$ is solvable.*

Proof. A nilpotent algebra is in particular solvable, hence $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}^{(1)}$ is solvable. Then $\mathfrak{g}$ is solvable by definition. ☐

{t:Cartan}

Theorem 7.8. *[Cartan's Criterion] Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$, where V is fnite dimensional. The following statements are equivalent.*

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

Proof. "(2)⇒(1)" If $\mathfrak{g}$ is solvable, then, by Lie's Theorem 7.2, the matrices of $\mathfrak{g}$ are upper triangular with respect to a suitable basis of V. Then any element in $[\mathfrak{g},\mathfrak{g}]$ is a strict upper triangular matrix and (1) follows by an easy computation.

"(1)⇒(2)" By Lemma 7.7 it is enough to show that $[\mathfrak{g},\mathfrak{g}]$ is nilpotent. By Lemma 4.14 if x is nilpotent, then so is $\operatorname{ad} x$ and by Engel's Theorem 4.13 an algebra whose all element are ad-nilpotent is nilpotent. Therefore, it is enough to show that all $x \in [\mathfrak{g},\mathfrak{g}]$ are nilpotent.

We apply Lemma 7.5 as follows: V is as given, $A = [\mathfrak{g},\mathfrak{g}]$, $B = \mathfrak{g}$. Therefore,

$$M = \{x \in \mathfrak{gl}(V) \mid \operatorname{ad} x(\mathfrak{g}) \subset [\mathfrak{g},\mathfrak{g}]\}.$$

Of course, $\mathfrak{g} \subset M$. Our hypothesis tells us that $\operatorname{tr}(xy) = 0$ for all $x \in [\mathfrak{g},\mathfrak{g}]$ and all $y \in \mathfrak{g}$. To apply Lemma 7.5 we need the stronger statement that $\operatorname{tr}(xy) = 0$ for all $x \in [\mathfrak{g},\mathfrak{g}]$ and $y \in M$.

Now recall that $[\mathfrak{g},\mathfrak{g}]$ is generated by the commutators $[u,w]$. If $z \in M$, then, by Lemma 7.6, $\operatorname{tr}([u,w]z) = \operatorname{tr}(u[w,z]) = \operatorname{tr}([w,z]u)$. Now note that since $z \in M$, the element $[w,z]$ is in $[\mathfrak{g},\mathfrak{g}]$ by definition of M. Therefore, $\operatorname{tr}([w,z]u) = 0$, hence Lemma 7.6 applies and we are done. □

{c:Cartan}

Corollary 7.9. *Let $\mathfrak{g}$ be a finite dimensional Lie algebra such that $\operatorname{tr}(\operatorname{ad}_x \operatorname{ad}_y) = 0$ for all $x \in [\mathfrak{g},\mathfrak{g}]$, $y \in \mathfrak{g}$. Then $\mathfrak{g}$ is solvable.*

Proof. Consider the adjoint representation $\operatorname{ad}\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g})$. Cartan's Criterion shows that $\operatorname{ad}(\mathfrak{g}) \subset \mathfrak{gl}(\mathfrak{g})$ is solvable. Now note that $\ker(\operatorname{ad}) = Z(\mathfrak{g})$ is solvable (being abelian) and that $\operatorname{ad}(\mathfrak{g}) \simeq \mathfrak{g}/Z(\mathfrak{g})$. We conclude by Proposition 4.3(2). □

8. The Killing form and semisimplicity

Definition. Let $\mathfrak{g}$ be a Lie algebra. If $x, y \in \mathfrak{g}$, set

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

This is a symmetric bilinear form on $\mathfrak{g}$, called the *Killing form*.

It is clear that κ is symmetric, and bilinearity follows from the fact that ad is a Lie algebra homomorphism and that the trace is bilinear. Also note that

$$\begin{aligned}\kappa([x,y],z) &= \operatorname{tr}(\operatorname{ad}[x,y]\operatorname{ad} z) = \operatorname{tr}([\operatorname{ad} x, \operatorname{ad} y]\operatorname{ad} z)\\ &= \operatorname{tr}(\operatorname{ad} x, [\operatorname{ad} y, \operatorname{ad} z]) = \operatorname{tr}(\operatorname{ad} x, \operatorname{ad}[y,z])\\ &= \kappa(x,[y,z]),\end{aligned}$$

where we used Lemma 7.6 for the third equality. We will call this property of κ *associativity.*

{l:Killing}

Lemma 8.1. *Let I be an ideal of $\mathfrak{g}$. If κ is the Killing form of $\mathfrak{g}$ and κ_I is the Killing form of I (recall that any ideal is in particular a Lie algebra), then $\kappa_I = \kappa_{|I\times I}$.*

Proof. We will use the following easy fact from linear algebra: If W is a subspace of a finite dimensional vector space V and $\phi \in \mathrm{End}(V)$ maps V into W, then $\mathrm{tr}(\phi) = \mathrm{tr}(\phi_{|W})$. Indeed, take a basis of W, extend it to a basis of V and consider the resulting matrix of ϕ.

Now, if $x, y \in I$, then $\mathrm{ad}\, x\, \mathrm{ad}\, y$ is an endomorphism of $\mathfrak{g}$ and it maps $\mathfrak{g}$ into I, since I is an ideal. Therefore, $\kappa(x, y) = \mathrm{tr}(\mathrm{ad}\, x\, \mathrm{ad}\, y)$ coincides with the trace of $(\mathrm{ad}\, x\, \mathrm{ad}\, y)_{|I} = \mathrm{ad}_I\, x\, \mathrm{ad}_I\, y$. □

Definition. Let V be a finite dimensional vector space and let

$$\beta \colon V \times V \longrightarrow K$$

be a symmetric bilinear form. We call β *nondegenerate* if its *radical* S is 0, where

$$S = \{x \in V \mid \beta(x, y) = 0\ \forall y \in V\}.$$

By bilinearity, S is a vector subspace of V. Note that β is nondegenerate if and only if the map $V \longrightarrow V^*$, $x \longmapsto \beta(x, -)$ is an isomorphism. A practical way of checking nondegeneracy is the following. Fix a basis $(x_1, \ldots, x_n)$ of V and consider the matrix A whose (i, j)-entry is $\beta(x_i, x_j)$. Then β is nondegenerate if and only if A is invertible if and only if $\det(A) \neq 0$.

In particular, this definition applies to a symmetric bilinear form on a Lie algebra. For the Killing form we have

{l:RadKill}

Lemma 8.2. *If $\mathfrak{g}$ is a Lie algebra and κ its Killing form, then the radical of κ is an ideal in $\mathfrak{g}$.*

Proof. If $x \in S$ and $y \in \mathfrak{g}$, then $\kappa([x, y], z) = \kappa(x, [y, z]) = 0$ for all $z \in \mathfrak{g}$ by associativity of κ. □

{e:sl2Killing}

Example 8.3. Recall that the Lie algebra $\mathfrak{sl}(2, K)$ has as basis the matrices

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We fix the ordered basis $\mathcal{B} = (X, H, Y)$. In Example 3.9 we already computed the following equalities:

$$\begin{aligned} [X, Y] &= \mathrm{ad}\, X(Y) = H, \\ [H, X] &= \mathrm{ad}\, H(X) = 2X, \\ [H, Y] &= \mathrm{ad}\, H(Y) = -2Y. \end{aligned}$$

From this we get that with respect to $\mathcal{B}$, $\mathrm{ad}\, H$ is a diagonal matrix $\mathrm{diag}(2, 0, -2)$. Furthermore,

$$\mathrm{ad}\, X = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$\operatorname{ad} Y = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

Using this, one checks that $\kappa(X, Y) = \operatorname{tr}(\operatorname{ad} X \operatorname{ad} Y) = 4 = \kappa(Y, X)$ and $\kappa(H, H) = 8$, while all the other pairings have trace zero. This means that κ is given by the following matrix

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

Its determinant is -128. In particular, if $\operatorname{char}(K) \neq 2$, then κ is non-degenerate.

Recall that a Lie algebra is called semisimple if $\operatorname{rad}(\mathfrak{g}) = 0$, where $\operatorname{rad}(\mathfrak{g})$ is the unique maximal solvable ideal. We have the following easy

{l:SemiAb}

Lemma 8.4. *A Lie algebra $\mathfrak{g}$ is semisimple if and only if it has no nonzero abelian ideals.*

Proof. If $0 \neq I$ is a nonzero abelian ideal, then $I \subset \operatorname{rad}(\mathfrak{g})$ (since any abelian ideal is solvable), so $\mathfrak{g}$ cannot be semisimple.

Conversely, if $\mathfrak{h} = \operatorname{rad}(\mathfrak{g}) \neq 0$, then its derived series is

$$\mathfrak{h} \supset \mathfrak{h}^{(1)} \supset \dots \mathfrak{h}^{(k-1)} \supset \mathfrak{h}^{(k)} = 0.$$

If we assume that k is minimal with the property that $\mathfrak{h}^{(k)} = 0$, then, by Remark 4.5, $0 \neq \mathfrak{h}^{(k-1)}$ is an abelian algebra. In particular, $\mathfrak{h}^{(k-1)}$ is an abelian ideal in $\mathfrak{g}$. $\square$

The following result relates semisimplicity and the Killing form.

{t:KillingSeSi}

Theorem 8.5. *A Lie algebra $\mathfrak{g}$ is semisimple if and only if its Killing form is nondegenerate.*

Proof. Suppose that $\mathfrak{g}$ is semisimple, hence $\operatorname{rad}(\mathfrak{g}) = 0$. Let S be the radical of κ. By definition, $\operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$ for all $x \in S$ and $y \in \mathfrak{g}$ (in particular, for $y \in [S, S]$). By Corollary 7.9, S is solvable. Therefore, $S \subset \operatorname{rad}(\mathfrak{g}) = 0$, so κ is nondegenerate.

Conversely, suppose that $S = 0$. By Lemma 8.4, it suffices to show that any abelian ideal I of $\mathfrak{g}$ is contained in S. So, let $x \in I$ and let $y \in \mathfrak{g}$. Then $\operatorname{ad} x \operatorname{ad} y$ maps $\mathfrak{g}$ into I. Therefore, $(\operatorname{ad} x \operatorname{ad} y)^2$ maps $\mathfrak{g}$ into $[I, I] = 0$. Thus, $\operatorname{ad} x \operatorname{ad} y$ is nilpotent, hence $0 = \operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = \kappa(x, y)$, so $x \in S$. $\square$

A little later we will show that a Lie algebra is semisimple of and only if it is a direct sum of simple algebras. But first we want to introduce yet another class of Lie algebras. Recall that the center of a Lie algebra is of course solvable, so $Z(\mathfrak{g}) \subset \operatorname{rad}(\mathfrak{g})$.

Definition. A Lie algebra is called *reductive* if $Z(\mathfrak{g}) = \operatorname{rad}(\mathfrak{g})$.

Actually, there is an equivalent description.

{l:reductive}

Lemma 8.6. *A Lie algebra* $\mathfrak{g}$ *is reductive if and only if every solvable ideal* I *is central, that is* $[\mathfrak{g}, I] = 0$.

Proof. If $\mathfrak{g}$ is reductive, then $\operatorname{rad}(\mathfrak{g})$ is central, hence every solvable ideal is as well.

Conversely, if every solvable ideal is central, this in particular holds for $\operatorname{rad}(\mathfrak{g})$. $\square$

{e:reductive}

Example 8.7. Every semisimple Lie algebra is reductive.

The converse does not hold: If $0 \neq \mathfrak{g}$ is an abelian Lie algebra, then it is reductive, but not semisimple.

{p:reductive}

Proposition 8.8. *Let* $\rho\colon \mathfrak{g} \subset \mathfrak{gl}(V)$ *be a faithful irreducible representation of a Lie algebra. Then* $\mathfrak{g}$ *is reductive and* $\dim_K Z(\mathfrak{g}) \leq 1$. *If, furthermore,* $\operatorname{tr}(\rho(x)) = 0$ *for all* $x \in \mathfrak{g}$, *then* $\mathfrak{g}$ *is semisimple.*

Proof. Let I be a solvable ideal in $\mathfrak{g}$. As in the proof of Theorem 7.1 one shows that there exists a basis of V such that for all $y \in I$ the matrix $\rho(y)$ is upper triangular with $\lambda(y)$ on the diagonal, where $\lambda\colon I \longrightarrow K$ is a linear function. But the proof there shows even more. Noamely, let $x \in [\mathfrak{g}, I] \subset I$. Then $\operatorname{tr}(\rho(x)) = 0$, since the trace of a commutator vanishes. Therefore, $\lambda_{[\mathfrak{g},I]} = 0$. Going back to the proof of Theorem 7.1, this shows that the subspace

$$W = \{w \in V \mid x.w = \lambda(x).w \; \forall x \in I\}$$

is stable under $\mathfrak{g}$, hence a subrepresentation of V, hence equal to V. Therefore, $\rho(y)$ is a diagonal matrix $\operatorname{diag}(\lambda(y), \ldots, \lambda(y))$ for all I. Since ρ is faithful, it follows that $\dim_K I \leq 1$. Furthermore, $[\mathfrak{g}, I] = 0$, so $\mathfrak{g}$ is reductive. If $\operatorname{tr}(\rho(x)) = 0$ for all $x \in \mathfrak{g}$, then, in particular, $\operatorname{tr}(\rho(y)) = 0$ for all $y \in I$, hences $I = 0$, so every solvable ideal in $\mathfrak{g}$ is trivial and, therefore, $\mathfrak{g}$ admits no nonzero abelian ideals, thus $\mathfrak{g}$ is semisimple by Lemma 8.4. $\square$

{c:GLred}

Corollary 8.9. *The Lie algebra* $\mathfrak{gl}(n, K)$ *is reductive. The Lie algebra* $\mathfrak{sl}(n, K)$ *is semisimple.* $\square$

Of course, Theorem 1.12 predicts that $\mathfrak{sl}(n, \mathbb{C})$ is even simple. Exercise 4 on Sheet 5 tells us that $\mathfrak{sl}(n, K)$ is simple if $\operatorname{char}(K) = 0$.

Definition. Leg $\mathfrak{g}$ be a Lie algebra. We say that $\mathfrak{g}$ is the *direct sum of ideals* $I_1, \ldots, I_t$ if $\mathfrak{g}$ is the direct sum of its subspaces $I_1, \ldots, I_t$.

In particular, $I_i \cap I_j = 0$ for $i \neq j$. Since $[I_i, I_j] \subset I_i \cap I_j$, this means that also $[I_i, I_j] = 0$ for $i \neq j$. Note that to say that $\mathfrak{g}$ is the direct sum of ideals just means that it is a direct sum of the corresponding vector spaces and the bracket of $\mathfrak{g}$ is defined componentwise.

{t:SimSemisim}

Theorem 8.10. *Let* $\mathfrak{g}$ *be a semisimple Lie algebra. Then there exist ideals* $\mathfrak{g}_1, \ldots, \mathfrak{g}_t$ *of* $\mathfrak{g}$ *which are simple as Lie algebras and such that* $\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_t$. *Every simple ideal* I *of* $\mathfrak{g}$ *coincides with one of the* $\mathfrak{g}_i$. *Moreover, the Killing form of* $\mathfrak{g}_i$ *is the restriction of* κ *to* $\mathfrak{g}_i \times \mathfrak{g}_i$.

Proof. Let I be an arbitrary ideal of $\mathfrak{g}$. Define

$$I^\perp := \{x \in \mathfrak{g} \mid \kappa(x, y) = 0\ \forall y \in I\}.$$

This is an ideal in $\mathfrak{g}$, because I is and because κ is associative. Note that $\kappa_{|I\cap I^\perp}$ is 0. In particular, applying Cartan's Criterion (Theorem 7.8) to the Lie algebra I, we see that $I \cap I^\perp$ is solvable. Since $\mathfrak{g}$ is semisimple, this implies that $I \cap I^\perp = 0$. Now, κ is nondegenerate, therefore $\dim I + \dim I^\perp = \dim \mathfrak{g}$, hence $\mathfrak{g} = I \oplus I^\perp$.

Note that any ideal J of I is automatically an ideal in $\mathfrak{g}$ (namely, $J \oplus 0$), because the bracket on $\mathfrak{g}$ is defined componentwise. The same argument holds for ideals of $I^\perp$. In particular, both I and $I^\perp$ are also semisimple. We can thus continue by induction on $\dim(\mathfrak{g})$ to arrive at the claimed decomposition of $\mathfrak{g}$.

Next we prove that these ideals are unique. If J is any simple ideal of $\mathfrak{g}$, then $[\mathfrak{g}, J]$ is also an ideal of J. Since $Z(\mathfrak{g}) = 0$, this means that $[\mathfrak{g}, J] = J$. On the other hand, $J = [\mathfrak{g}, J] = \oplus_{i=1}^t [\mathfrak{g}_i, J]$. Therefore, all summands bar one are zero. Say, $[\mathfrak{g}_i, J] = J$. Then $J \subset \mathfrak{g}_i$, hence $J = \mathfrak{g}_i$, since $\mathfrak{g}_i$ is simple.

The last assertion is just Lemma 8.1. □

{c:SimSemisim}

Corollary 8.11. *If $\mathfrak{g}$ is a semisimple Lie algebra, then $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ and all ideals and homomorphic images of $\mathfrak{g}$ are semisimple. Moreover, each ideal of $\mathfrak{g}$ is a sum of certain simple ideals of $\mathfrak{g}$.*

Proof. To check that $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ use that $[\mathfrak{g}_i, \mathfrak{g}_j] = 0$ for $i \neq j$ and $[\mathfrak{g}_i, \mathfrak{g}_i] = \mathfrak{g}_i$ since $\mathfrak{g}_i$ is simple, compare Remark 3.8. The other assertions are rather obvious. □

Recall that a derivation is an endomorphism of a vector space which satisfies the Leibniz rule.

{t:SemiSiDer}

Theorem 8.12. *If $\mathfrak{g}$ is semisimple, then* $\operatorname{ad} \mathfrak{g} = \operatorname{Der}(\mathfrak{g})$.

Proof. Recall that $Z(\mathfrak{g}) \subset \operatorname{rad}(\mathfrak{g}) = 0$, so $\operatorname{ad}\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g})$ is injective. We already remarked in Example 5.1(1) that $M = \operatorname{ad} \mathfrak{g} \subset \operatorname{Der}(\mathfrak{g}) = D$. A simpe computation shows the following formula (which is needed to show Exercise 2 on Sheet 4)

$$(\diamond) \quad [\delta, \operatorname{ad} x] = \operatorname{ad}(\delta x)\ \forall x \in \mathfrak{g}, \forall \delta \in D.$$

In other words, $[D, M] \subset M$. Since $\mathfrak{g} \simeq \operatorname{ad}(\mathfrak{g}) = M$, M is semisimple, hence the Killing form κ_M of M is nondegenerate. On the other hand, κ_M is the restriction of κ_D to $M \times M$ by Lemma 8.1. If $I = M^\perp$ is the orthogonal to M under κ_D, then the nondegeneracy of κ_M forces $I \cap M = 0$. Now I and M are ideals of D, so $[I, M] = 0$. If $\delta \in I = M^\perp$, then $\operatorname{ad}(\delta x) = [\delta, \operatorname{ad} x] = 0$, for all $x \in \mathfrak{g}$, since $\operatorname{ad} x \in M$. Since ad is injective, this means that $\delta x = 0$ for all $x \in \mathfrak{g}$, so $\delta = 0$. Hence, $I = 0$, so $M = \operatorname{ad}(\mathfrak{g}) = \mathfrak{g} = D = \operatorname{Der}(\mathfrak{g})$. □

9. Weyl's Theorem

Convention. In this section all representations will be finite dimensional. The conventions concerning the base field remain as does the convention that any Lie algebra we consider is finite dimensional.

In this section we will study representations of a *semisimple* Lie algebra. Weyl's theorem will tell us that any representaion is completely reducible in this case.

First we will need some preparations. Let V be a finite dimensional K-vector space and V^* its dual space. Fix a basis $(x_1, \dots, x_n)$ of V and consider the dual basis $(x_1^*, \dots, x_n^*)$ of V^* characterised by the property $x_i^*(x_j) = \delta_{ij}$. Recall that $\Phi\colon V^* \otimes V \simeq \mathrm{End}(V)$, where Φ is defined via $f \otimes v \longmapsto [w \longmapsto f(w)v]$. Note that

$$\sum_{i=1}^{n} x_i^* \otimes x_i = \Phi^{-1}(\mathrm{id}_V).$$

Indeed, $\Phi(x_i^* \otimes x_i)$ is the map which sends x_i to x_i and x_j to 0 if $j \neq i$. Therefore, $\Phi(\sum_{i=1}^{n} x_i^* \otimes x_i)$ sends any x_j to itself, that is, it is the identity map. In particular, the element $\sum_{i=1}^{n} x_i^* \otimes x_i$ does not depend on the choice of a basis of V.

Now, let $\mathfrak{g}$ be a semisimple Lie algebra and let $\phi\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ be a faithful representation of $\mathfrak{g}$, that is, ϕ is injective. Define a map

$$\beta\colon \mathfrak{g} \times \mathfrak{g} \longrightarrow K, \quad (x, y) \longmapsto \mathrm{tr}(\phi(x)\phi(y)).$$

Note that β is a symmetric bilinear form on $\mathfrak{g}$. Due to Lemma 7.6, the form β is associative. In particular, its radical S_β is an ideal of $\mathfrak{g}$. Consider $\phi(S_\beta) \simeq S_\beta$. By Cartan's Criterion (Theorem 7.8) and the definition of S_β, the algebra S_β is solvable. But $\mathfrak{g}$ is semisimple, hence the maximal solvable ideal $\mathrm{rad}(\mathfrak{g}) = 0$, so $S_\beta = 0$. In other words, β is nondegenerate. Note that the Killing form is just β if $\phi = \mathrm{ad}$.

Let $\mathfrak{g}$ be a semisimple Lie algebra and β be any nondegenerate symmetric associative bilinear form on $\mathfrak{g}$. If $(x_1, \dots, x_n)$ is a basis of $\mathfrak{g}$, there is a uniquely determined dual basis $(y_1, \dots, y_n)$ described by the property $\beta(x_i, y_j) = \delta_{ij}$.

If $x \in \mathfrak{g}$, we can of course write

$$[x, x_i] = \sum_{j=1}^{n} a_{ij} x_j$$

for some $a_{ij} \in K$ and similarly,

$$[x, y_i] = \sum_{j=1}^{n} b_{ij} y_j.$$

Now we compute

$$\begin{aligned} a_{ik} = \sum_{j=1}^{n} a_{ij}\beta(x_j, y_k) &= \beta([x, x_i], y_k) \\ &= \beta(-[x_i, x], y_k) = \beta(x_i, -[x, y_k]) \\ &= -\sum_{j=1}^{n} b_{kj}\beta(x_i, y_j) = -b_{ki}. \end{aligned}$$

If β is any nondegenerate symmetric associative bilinear form on $\mathfrak{g}$ and $\phi\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ is any representation of $\mathfrak{g}$, set

$$c_\phi(\beta) = \sum_{i=1}^{n} \phi(x_i)\phi(y_i) \in \operatorname{End}(V),$$

where x_i and y_i are elements of dual bases with respect to β.

Note that in $\operatorname{End}(V)$ we have

$$[x, yz] = xyz - yzx = xyz - yxz + yxz - yzx = [x, y]z + y[x, z].$$

Using this and the equation $a_{ik} = -b_{ki}$ established above, we now compute for any $x \in \mathfrak{g}$

$$\begin{aligned} [\phi(x), c_\phi(\beta)] &= [\phi(x), \sum_{i=1}^{n} \phi(x_i)\phi(y_i)] \\ &= \sum_{i=1}^{n} [\phi(x), \phi(x_i)\phi(y_i)] \\ &= \sum_{i=1}^{n} [\phi(x), \phi(x_i)]\phi(y_i) + \sum_{i=1}^{n} \phi(x_i)[\phi(x), \phi(y_i)] \\ &= \sum_{i=1}^{n} \phi([x, x_i])\phi(y_i) + \sum_{i=1}^{n} \phi(x_i)\phi([x, y_i]) \\ &= \sum_{i,j=1}^{n} a_{ij}\phi(x_j)\phi(y_i) + \sum_{i,j=1}^{n} b_{ij}\phi(x_i)\phi(y_j) = 0. \end{aligned}$$

In other words, we have proved

{p:Casicomm}

Proposition 9.1. *The map $c_\phi(\beta)$ commutes with $\phi(\mathfrak{g})$.* □

Definition. If β is any nondegenerate symmetric associative bilinear form on $\mathfrak{g}$, $\phi\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ is any representation of $\mathfrak{g}$ and x_i and y_i are elements of dual bases with respect to β, then $c_\phi(\beta) \in \operatorname{End}(V)$ is called the *Casimir element of* ϕ. If β is understood from the context, we will simply write c_ϕ.

{l:TrCasi}

Lemma 9.2. *We have* $\operatorname{tr}(c_\phi) = \dim \mathfrak{g}$.

Proof. Indeed, $\operatorname{tr}(c_\phi) = \operatorname{tr}(\sum_{i=1}^{n} \phi(x_i)\phi(y_i)) = \sum_{i=1}^{n} \beta(x_i, y_i) = \dim \mathfrak{g}$. □

{r:Casi}

Remark 9.3. Assume that ϕ is an irreducible faithful representation of $\mathfrak{g}$. By Remark 5.6, the only elements commuting with $\phi(\mathfrak{g})$ are the scalar multiples of the identity. Hence, $c_\phi = \lambda \cdot \mathrm{id}_V$, so $\lambda = \frac{\dim \mathfrak{g}}{\dim V}$.

{e:Casi}

Example 9.4. Let $\mathfrak{g} = \mathfrak{sl}(2, K)$, $V = K^2$, $\phi = \mathrm{id}\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$. Let (X, H, Y) be the basis of $\mathfrak{g}$ used before. With respect to the trace form, one quickly computes $\mathrm{tr}(XY) = 1$, $\mathrm{tr}(XX) = 0$, $\mathrm{tr}(XH) = 0$, $\mathrm{tr}(HH) = 2$, $\mathrm{tr}(HY) = 0$ and $\mathrm{tr}(YY) = 0$. It follows that the dual basis is of the form $(Y, \frac{1}{2}H, X)$, hence

$$c_\phi = XY + \frac{1}{2}H^2 + YX = \begin{pmatrix} \frac{3}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}.$$

Note that $\frac{3}{2} = \frac{\dim \mathfrak{g}}{\dim V}$.

{r:CasiFai}

Remark 9.5. Assume that ϕ is not necessarily faithful. Then its kernel $\ker(\phi)$ is a sum of certain simple ideals of $\mathfrak{g}$ by Corollary 8.11. If we denote by $\mathfrak{g}'$ the sum of the remaining simple ideals, then the restriction of ϕ to $\mathfrak{g}'$ is a faithful representation of $\mathfrak{g}'$ and one can apply the above construction to it.

The following result will be useful in the proof of Weyl's Theorem.

{l:SemiSiSL}

Lemma 9.6. *Let $\phi\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ be a representation of a semisimple Lie algebra $\mathfrak{g}$. Then $\phi(\mathfrak{g}) \subset \mathfrak{sl}(V)$. In particular, $\mathfrak{g}$ acts trivially on any one dimensional $\mathfrak{g}$-module.*

Proof. By Corollary 8.11, $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$. Now use that $\mathfrak{sl}(V) = [\mathfrak{gl}(V), \mathfrak{gl}(V)]$ (for this one can, for instance, use Equation 3.1) to see the first claim. For the second just note that $\mathfrak{sl}(K) = 0$. $\square$

We also need the following

Definition. Let V be a $\mathfrak{g}$-module. The set

$$V^{\mathfrak{g}} := \{v \in V \mid x.v = 0 \; \forall x \in \mathfrak{g}\}$$

is called the *invariant submodule* of V. Its elements are called *invariant vectors.*

The quotient space $V_{\mathfrak{g}} := V/\mathfrak{g}V$ is called the space of *coinvariants.*

For example, if V and W are two $\mathfrak{g}$-modules, then we have seen that $\mathrm{Hom}(V, W)$ is also a $\mathfrak{g}$-module. Recall that $\mathrm{Hom}_{\mathfrak{g}}(V, W)$ denotes the space of morphisms of representations. It follows directly from the definitions that $\mathrm{Hom}_{\mathfrak{g}}(V, W) = \mathrm{Hom}(V, W)^{\mathfrak{g}}$.

{l:Weyl}

Lemma 9.7. *Let (V, ρ) be a representation of a semisimple Lie algebra $\mathfrak{g}$. Then*

$$V = V^{\mathfrak{g}} \oplus \mathfrak{g}V,$$

where we abuse notation by writing $\mathfrak{g}V$ for $\rho(\mathfrak{g})V$.

Proof. We use induction on $\dim(V)$. Note that the case $\dim(V) = 1$ is precisely the content of Lemma 9.6, since it gives that $V = V^{\mathfrak{g}}$ in this case.

Now assume that $\dim(V) > 1$. If $V = V^{\mathfrak{g}}$, there is nothing to prove. So assume that $V \neq V^{\mathfrak{g}}$. By Corollary 8.11, $\rho(\mathfrak{g})$ is a semisimple subalgebra of $\mathfrak{gl}(V)$. We have the Casimir operator $c_\phi\colon V \longrightarrow V$, where we use the bilinear form $\beta(x, y) = \mathrm{tr}(xy)$ for $x, y \in \mathfrak{gl}(V)$.

Since K is algebraically closed, the vector space V decomposes into a direct sum of generalised eigenspaces V_λ of c_ϕ. We have seen in Proposition 9.1 that c_ϕ commutes with $\rho(\mathfrak{g})$, hence every V_λ is a subrepresentation of V. There are now two cases.

- c_ϕ has at least two eigenvalues. Then V is a direct sum of smaller dimensional subrepresentations V_λ, to which we can apply the induction hypothesis.
- c_ϕ has a single eigenvalue. We have seen in Lemma 9.2 that $\mathrm{tr}(c_\phi) = \dim \mathfrak{g} \neq 0$. Therefore, the unique eigenvalue of c_ϕ is not zero, hence $V = c_\phi V$, thus also $V = \mathfrak{g}V$.

□

We are now ready to prove the following

{t:Weyl}

Theorem 9.8 (Weyl's Theorem). *Any representation $\phi\colon \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$ of a semisimple algebra $\mathfrak{g}$ is completely reducible.*

Proof. Let $U \subset V$ be a subrepresentation. We have a morphism of representations

$$\mathrm{Hom}(V, U) \longrightarrow \mathrm{Hom}(U, U)$$

given by restriction. In particular, this morphism is surjective. By Lemma 9.7, we can restrict this surjective morphism to the invariant vectors and get a surjection

$$\mathrm{Hom}(V, U)^{\mathfrak{g}} \longrightarrow \mathrm{Hom}(U, U)^{\mathfrak{g}}.$$

In particular, we have an element $f \in \mathrm{Hom}(V, U)^{\mathfrak{g}}$ mapping to id_U.

We claim that $V = U \oplus \ker(f)$. Indeed, $\ker(f)$ is of course a subrepresentation of V. Furthermore, $U \cap \ker(f) = 0$, since if $v \in U \cap \ker(f)$, then $f(v) = 0 = v$. Finally, $V = U + \ker(f)$, since for any $v \in V$ we can write $v = f(v) + (v - f(v))$. Of course, $f(v) \in U$. But $f(v - f(v)) = f(v) - f(v) = 0$, hence $v - f(v) \in \ker(f)$ as claimed. Summarising, we have proved that $V = U \oplus \ker(f)$, hence we can continue decomposing the components U and $\ker(f)$ until we reach irreducible representations. □

10. Jordan decomposition of a semisimple Lie algebra

We begin this section with a quite general result.

{p:JordanDer}

Proposition 10.1. *Let A be a finite dimensional algebra and let $\delta \in \mathrm{Der}(A) \subset \mathrm{End}(A)$ be a derivation. Consider the Jordan decomposition $\delta = \sigma + \nu$ of δ in $\mathrm{End}(A)$ with σ semisimple and ν nilpotent. Then σ and ν are derivations.*

Proof. Since $\delta = \sigma + \nu$ and sums of derivations are derivations, it will be enough to show that $\sigma \in \mathrm{Der}(A)$. For any $\alpha \in K$, set

$$A_\alpha = \{x \in A \mid \exists k \geq 1 : (\delta - \alpha \cdot \mathrm{id})^k x = 0\},$$

that is, A_α is the generalised eigenspace of α. Therefore, A is isomorphic to the direct sum of those A_α for which α is an eigenvalue of δ or, equivalently, of σ. Since σ is semisimple, it acts on A_α via multiplication with α. Now the reader can check by induction on n the following formula:

$$(\delta - (\alpha+\beta)\cdot \mathrm{id})^n(xy) = \sum_{i=0}^{n} \binom{n}{i} (\delta - \alpha\cdot\mathrm{id})^{n-i}(x)(\delta-\beta\cdot\mathrm{id})^i(y).$$

It implies that $A_\alpha A_\beta \subset A_{\alpha+\beta}$. Therefore, if $x \in A_\alpha$ and $y \in A_\beta$, then $\sigma(xy) = (\alpha+\beta)xy$, since $xy \in A_{\alpha+\beta}$ (the latter space could of course be zero). On the other hand,

$$\sigma(x)y + x\sigma(y) = (\alpha+\beta)xy.$$

Since the decomposition into generalised eigenspaces is direct, it follows that σ is indeed a derivation. □

Now recall from Theorem 8.12 that for a semisimple Lie algebra $\mathfrak{g}$ we have $\mathrm{ad}(\mathfrak{g}) = \mathrm{Der}(\mathfrak{g})$. Furthermore, $\mathfrak{g} = \mathrm{ad}(\mathfrak{g})$, since $\ker(\mathrm{ad}) = Z(\mathfrak{g}) \subset \mathrm{rad}(\mathfrak{g}) = 0$. Hence,

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

If $x \in \mathfrak{g}$, then $\mathrm{ad}\, x$ is a derivation, hence by Proposition 10.1 its semisimple and nilpotent parts are again derivations. Therefore, any $x \in \mathfrak{g}$ determines unique elements $s, n \in \mathfrak{g}$ such that $\mathrm{ad}\, x = \mathrm{ad}\, s + \mathrm{ad}\, n$, where the latter is the usual Jordan decomposition of $\mathrm{ad}\, x \in \mathrm{End}(\mathfrak{g})$. Thus, $x = s+n$, $[s,n]=0$, $\mathrm{ad}\, s$ is semisimple and $\mathrm{ad}\, n$ is nilpotent. This is the *abstract Jordan decomposition* of a semisimple Lie algebra. This decomposition is unique.

{t:AbstractJ}

Theorem 10.2. *Let $\mathfrak{g} \subset \mathfrak{gl}(V)$ be a semisimple Lie algebra. Then $\mathfrak{g}$ contains the semisimple and nilpotent parts in $\mathfrak{gl}(V)$ of all its elements. In particular, the abstract and usual Jordan decomposition of $\mathfrak{g}$ coincide.*

Proof. The second statement follows from the first because both the abstract and usual Jordan decompositions are unique.

Let $x \in \mathfrak{g}$ be arbitrary and consider its Jordan decomposition $x = x_s + x_n$ in $\mathfrak{gl}(V)$. Since $\mathrm{ad}\, x(\mathfrak{g}) \subset \mathfrak{g}$, it follows from Proposition 6.2(3) that $\mathrm{ad}\, x_s(\mathfrak{g}) \subset \mathfrak{g}$ and $\mathrm{ad}\, x_n(\mathfrak{g}) \subset \mathfrak{g}$. Here $\mathrm{ad} = \mathrm{ad}_{\mathfrak{gl}(V)}$. In other words,

$$x_s, x_n \in N_{\mathfrak{gl}(V)}(\mathfrak{g}) =: N,$$

where N is the normaliser of $\mathfrak{g}$ in $\mathfrak{gl}(V)$. We have seen that N is a subalgebra of $\mathfrak{gl}(V)$. Clearly, it contains $\mathfrak{g}$ as an ideal.

At this point we would like to have the equality $N = \mathfrak{g}$, but this does not hold in general. Indeed, $\mathfrak{g} \subset \mathfrak{sl}(V)$ and the scalar multiples of the identity are in N but not in $\mathfrak{g}$.

To circumvent this problem, we will show that x_s, x_n are contained in a smaller subalgebra of N and then show that this subalgebra is equal to $\mathfrak{g}$.

So, let W be any $\mathfrak{g}$-submodule of V. Define

$$\mathfrak{g}_W := \{y \in \mathfrak{gl}(V) \mid y(W) \subset W, \operatorname{tr}(y_{|W}) = 0\}.$$

As an example: $\mathfrak{g}_V = \mathfrak{sl}(V)$. In general, any $\mathfrak{g}_W$ is a Lie algebra.

Note that since W is a $\mathfrak{g}$-submodule and $\mathfrak{g} \subset \mathfrak{sl}(V)$, $\mathfrak{g}$ is contained in $\mathfrak{g}_W$ for any W.

We now consider $\mathfrak{g}'$, the intersection of N with all the algebras $\mathfrak{g}_W$. Of course, $\mathfrak{g}'$ is a subalgebra of N and $\mathfrak{g}$ is contained in $\mathfrak{g}'$. Furthermore, $\mathfrak{g}$ is an ideal in $\mathfrak{g}'$ (since it is one in N). In fact, more is true: if $x \in \mathfrak{g}$, then $x_s, x_n \in \mathfrak{g}_W$ for all W, hence $x_s, x_n \in \mathfrak{g}'$. Indeed, x_s, x_n map W into W as x does and x_n is tracefree, hence so is x_s.

Our goal is to show that $\mathfrak{g} = \mathfrak{g}'$. Now, being completely reducible is equivalent to the following statement: every submodule has a complement (Exercise 1 on Sheet 6). Since $\mathfrak{g}'$ is a finite dimensional $\mathfrak{g}$-module, $\mathfrak{g}$ is semisimple and $\mathfrak{g} \subset \mathfrak{g}'$, we can write $\mathfrak{g}' = \mathfrak{g} \oplus M$ for some $\mathfrak{g}$-module M. Since $\mathfrak{g}' \subset N$, we have $[\mathfrak{g}, \mathfrak{g}'] \subset \mathfrak{g}$. In other words, the action of $\mathfrak{g}$ on M is trivial.

Let $y \in M$ and let W be an irreducible $\mathfrak{g}$-submodule of V. Since $[y, \mathfrak{g}] = 0$, Remark 5.6 tells us that y acts on W as a scalar matrix. Since $y \in \mathfrak{g}_W$, we have $\operatorname{tr}(y_{|W}) = 0$. Hence, y acts trivially on W. By Weyl's theorem 9.8, V decomposes into a direct sum of irreducible submodules, so by the preceding argument $y \in \mathfrak{g}' \subset \mathfrak{gl}(V)$ acts trivially on V. Hence, it is trivial. We have thus proved that $M = 0$, hence $\mathfrak{g} = \mathfrak{g}'$. ☐

{c:Jord}

Corollary 10.3. *Let $\phi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of a semisimple Lie algebra $\mathfrak{g}$. If $x = s + n$ is the abstract Jordan decomposition of an element $x \in \mathfrak{g}$, then $\phi(x) = \phi(s) + \phi(n)$ is the usual Jordan decomposition of $\phi(x)$ in $\mathfrak{gl}(V)$.*

Proof. The image $\phi(\mathfrak{g}) \subset \mathfrak{gl}(V)$ is a semisimple Lie algebra by Corollary 8.11. Let $x \in \mathfrak{g}$. Consider the following diagram

$$\begin{tikzcd}
\mathfrak{g} \arrow[r] \arrow[d, "\operatorname{ad} x"] & \phi(\mathfrak{g}) \arrow[r] \arrow[d, "\operatorname{ad} \phi(x)"] & \mathfrak{gl}(V) \arrow[d, "\operatorname{ad} \phi(x)"] \\
\mathfrak{g} \arrow[r] & \phi(\mathfrak{g}) \arrow[r] & \mathfrak{gl}(V).
\end{tikzcd}$$

Here, by abuse of notation we denote the action of $\phi(x)$ on $\phi(g)$ and $\mathfrak{gl}(V)$ by the same symbol. It is obvious that the diagram is commutative. By Remark 6.3, this remains true if we take the semisimple or the nilpotent part of the respective endomorphisms. Let us consider the diagram with semisimple parts.

Recall that by definition $(\operatorname{ad} x)_s = \operatorname{ad} s$. By Lemma 6.4, the semisimple part of $\operatorname{ad} \phi(x)$ in $\mathfrak{gl}(V)$ is the ad of the semisimple part of $\phi(x)$. In symbols, $(\operatorname{ad} \phi(x))_s = \operatorname{ad}(\phi(x)_s)$. This leads to the following commutative diagram

$$\begin{tikzcd}
\mathfrak{g} \arrow[r] \arrow[d, "\operatorname{ad} s"] & \phi(\mathfrak{g}) \arrow[r] \arrow[d] & \mathfrak{gl}(V) \arrow[d, "\operatorname{ad}(\phi(x)_s)"] \\
\mathfrak{g} \arrow[r] & \phi(\mathfrak{g}) \arrow[r] & \mathfrak{gl}(V).
\end{tikzcd}$$

Therefore, $\operatorname{ad}\phi(s) = \operatorname{ad}(\phi(x)_s)$. Since $\operatorname{ad}\colon \phi(\mathfrak{g}) \longrightarrow \mathfrak{gl}(\phi(\mathfrak{g}))$ is injective, we get $\phi(s) = \phi(x)_s$. Similarly, $\phi(n) = \phi(x)_n$. □

11. Representations of $\mathfrak{sl}(2, K)$

Recall that we work over an algebraically closed field K of characteristic zero and assume all representations to be finite dimensional over K. Also recall that the Lie algebra $\mathfrak{g} = \mathfrak{sl}(2, K)$ has as basis the matrices

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Let V be an arbitrary $\mathfrak{g}$-module. Note that H is semisimple, hence Corollary 10.3 gives that H acts on V as a diagonal matrix. Therefore, we can decompose V as a direct sum of eigenspaces $V_\lambda = \{v \in V \mid H.v = \lambda v\}$ for some $\lambda \in K$.

Definition. If $\lambda \in K$ is an eigenvalue of H acting on V, that is, $V_\lambda \neq 0$, then we call λ a *weight* of H in V and we call V_λ a *weight space*.

{l:Weightsl2}

Lemma 11.1. *If $v \in V_\lambda$, then $X.v \in V_{\lambda+2}$ and $Y.v \in V_{\lambda-2}$.*

Proof. Recall that $[X, Y] = H$, $[H, X] = 2X$ and $[H, Y] = -2Y$. Therefore,

$$H.X.v = [H, X].v + X.H.v = 2X.v - \lambda X.v = (2 + \lambda)X.v.$$

The second claim is proved similarly. □

{r:WeightNilp}

Remark 11.2. Note that since $\dim_K(V) < \infty$ and we have a direct sum decomposition $V = \bigoplus_{\lambda \in K} V_\lambda$ (of course, "most" $V_\lambda = 0$), there exists a $V_\lambda \neq 0$ such that $V_{\lambda+2} = 0$. Any nonzero vector $v \in V_\lambda$ is called a *maximal vector* of weight λ.

Also note that the lemma implies that X and Y act as nilpotent endomorphisms of V. Of course, this is also implied by Corollary 10.3.

{l:ClassRepSL2}

Lemma 11.3. *Assume that V is an* irreducible *$\mathfrak{g}$-module. Choose a maximal vector $v \in V_\lambda$. Set $v_{-1} = 0$ and $v_i = \frac{1}{i!} Y^i.v_0$ for $i \geq 0$. Then*

1. *$H.v_i = (\lambda - 2i)v_i$;*
2. *$Y.v_i = (i + 1)v_{i+1}$;*
3. *$X.v_i = (\lambda - i + 1)v_{i-1}$ for all $i \geq 0$.*

Proof. Item (2) is trivial, while item (1) follows immediately from (a repeated application of) Lemma 11.1.

To prove (3), we use induction over i. Note that $v_{-1} = 0$ by definition, while $X.v_0 = 0$ by Lemma 11.1 and the fact that v_0 is a maximal vector, that is, $V_{\lambda+2} = 0$. Now compute

$$\begin{aligned}
iX.v_i &= iX.\frac{1}{i!}Y^i.v_0 = \frac{1}{(i-1)!}X.Y.Y^{-1}.v_0 \\
&= X.Y.v_{i-1} = [X,Y]v_{i-1} + Y.X.v_{i-1} \\
&= H.v_{i-1} + Y((\lambda - i + 2)v_{i-2}) \\
&= (\lambda - 2(i-1))v_{i-1} + (\lambda - i + 2)Y.v_{i-2} \\
&= (\lambda - 2(i-1))v_{i-1} + (\lambda - i + 2)(i-1)v_{i-1} \\
&= (\lambda - 2i + 2 - \lambda + i - 2 + (\lambda - i + 2)i)v_{i-1} \\
&= i(\lambda - i + 1)v_{i-1},
\end{aligned}$$

where we used the definition for the first and second equality, (1) and induction for the fifth and (2) for the seventh. □

Note that (1) implies that all the nonzero v_i are linearly independent, since they are eigenvectors associated with distinct eigenvalues. Since $\dim_K(V) < \infty$, there exists a smallest integer m such that $v_m \neq 0$, but $v_{m+1} = 0$. Obviously then also $v_{m+i} = 0$ for all $i > 0$. Combining these arguments shows that

$$0 \neq U = \operatorname{span}_K(v_0, \ldots, v_m) \subset V$$

is not only a sub vector space, but also a submodule. Since V is irreducible, $U = V$. Furthermore, with respect to the given ordered basis of V, the endomorphism V is represented by a diagonal matrix, X is represented by a strictly upper triangular matrix and Y by a strictly lower triangular matrix. In particular, as we already knew, X and Y are represented by nilpotent matrices.

Now consider the formula in (3) for $i = m+1$: $0 = X.v_{m+1} = (\lambda - m)v_m$. Since $v_m \neq 0$, we have $\lambda = m$. On other words, the weight of a maximal vector is a nonnegative integer, equal to $\dim_K(V) - 1$. We will call this integer the *highest weight* of V.

Moreover, each weight μ occurs with multiplicity one (that is, if $V_\mu \neq 0$, then $\dim_K(V_\mu) = 1$) by (1). Summarising, we have proved

{t:ReprSL2}

Theorem 11.4. *Let V be an irreducible $(m+1)$-dimensional representation of $\mathfrak{g} = \mathfrak{sl}(2, K)$. Then the following holds.*

1. *The representation V is a direct sum of weight spaces V_μ with respect to H, where $\mu = m, m-2, \ldots, -(m-2), -m$ and $\dim_K(V_\mu) = 1$ for all μ.*
2. *The representation V has up to nonzero scalar multiples a unique maximal vector, whose weight is m.*
3. *The action of $\mathfrak{g}$ on V is completely described by the formulas in Lemma 11.3. In particular, there exists at most one irreducible $\mathfrak{g}$-representation (up to isomorphism) of each possible dimension $m+1$, $m \geq 0$.* □

Corollary 11.5. *Let V be any finite dimensional $\mathfrak{sl}(2,K)$-module. The eigenvalues of H on V are all integers and each occurs along with its negative (an equal number of times). Moreover, in any decomposition of V into a direct sum of irreducible submodules, the number of summands is precisely* $\dim V_0 + \dim V_1$. {c:ReprSL2}

Proof. If $V = 0$, there is nothing to prove. Otherwise, we can write V as a direct sum of irreducible submodules by Weyl's Theorem 9.8. The latter are all described by Theorem 11.4 and this argument immediately gives the first assertion. For the second, observe that each irreducible $\mathfrak{g}$-module has a unique occurrence of either the weight 0 or the weight 1, but not both by Theorem 11.4(1). □

To conclude this section, we answer the question whether $\mathfrak{sl}(2,K)$ does indeed have an irreducible module of each possible highest weight. In fact, the formulas of Lemma 11.3 do define an irreducible representation on an $(m+1)$-dimensional vector space V with basis $(v_0,\ldots,v_m)$ (see Exercise 3 on Sheet 9). Such a representation is denoted by $V(m)$. To make this more explicit, one could take as $V(m)$ the vector space of degree m polynomials in two variables s,t, that is $V(m) = \operatorname{span}_K(s^i t^{m-i}) \subset K[s,t]$ for $i = 0,\ldots,m$. Clearly, this is an $(m+1)$-dimensional vector space. One can check that letting X act by $s\partial_t$, Y act by $t\partial_x$ and H act by $s\partial_s - t\partial_t$ gives $K[s,t]$ the structure of an $\mathfrak{sl}(2,K)$-representation and the spaces $V(m)$ are then irreducible representations (see Exercise 2 on Sheet 6).

12. Root space decomposition

Let $\mathfrak{g}$ be a semisimple finite dimensional Lie algebra over an algebraically closed field of characteristic zero. By Engel's Theorem 4.13 if all elements in a Lie algebra are ad-nilpotent, then the algebra is nilpotent. Since $\mathfrak{g}$ is not nilpotent, not all its elements are ad-nilpotent, so there exists an $x \in \mathfrak{g}$ whose semisimple part x_s in the abstract Jordan decomposition is nonzero. Therefore, $\mathfrak{g}$ possesses nonzero subalgebras consisting of semisimple elements.

Definition. Let $\mathfrak{g}$ be a semisimple Lie algebra. A subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is called *toral* if it consists of semisimple elements.

Lemma 12.1. *Any toral subalgebra of a semisimple algebra is abelian.* {l:toral}

Proof. Let T be a toral subalgebra and let $x \in T$. Since $\operatorname{ad} x$ is semisimple, it is diagonalisable by our assumption that $K = \overline{K}$. We need to show that $[x,y] = 0$ for all $y \in T$ or, in other words, that $\operatorname{ad}_T x = 0$. We will thus show that $\operatorname{ad} x$ has no nonzero eigenvalues. Of course, the assertion is trivial for $x = 0$, so let $x \neq 0$.

Suppose that $\operatorname{ad} x(y) = [x,y] = ay$ for some $0 \neq a \in K$ and some $y \in T$. Then $z := \operatorname{ad}_T y(x) = [y,x] = -ay$. It follows that z is an eigenvector of $\operatorname{ad}_T y$, since $\operatorname{ad}_T y(z) = [y,-ay] = 0$. On the other hand, since $\operatorname{ad}_T y$ is also diagonalisable, we can find a basis of T consisting of eigenvectors $(v_1,\ldots,v_n)$ of $\operatorname{ad}_T y$. Say, $\operatorname{ad}_T y(v_i) = \mu_i v_i$ and write

$x = \sum_{i=1}^n a_i v_i$. Then

$$z = \operatorname{ad}_T y(x) = \operatorname{ad}_T y(\sum_{i=1}^n a_i v_i) = [y, \sum_{i=1}^n a_i v_i] = \sum_{i=1}^n a_i \mu_i v_i$$

and thus

$$0 = \operatorname{ad}_T y(z) = \sum_{i=1}^n a_i \mu_i^2 v_i,$$

hence $a_i = 0$ for all indices i where $\mu_i \neq 0$. Hence, $x = \sum a_i v_i$, where $v_i \ker \operatorname{ad}_T y$, so $x \in \ker \operatorname{ad}_T y$, a contradiction. Therefore, $\operatorname{ad} x(y) = 0$ for all $y \in T$, so T is indeed abelian. □

Now fix a maximal toral subalgebra of $\mathfrak{g}$. We will denote such an algebra by $\mathfrak{h}$. By the lemma, $\operatorname{ad}_{\mathfrak{g}} \mathfrak{h}$ is a commuting family of diagonalisable endomorphisms of $\mathfrak{g}$. Therefore, $\operatorname{ad}_{\mathfrak{g}} \mathfrak{h}$ is simulteneously diagonalisable, that is, there exists a basis of common eigenvectors for all elements of $\operatorname{ad}_{\mathfrak{g}} \mathfrak{h}$. In other words, $\mathfrak{g}$ can be written as a direct sum of subspaces

$$\mathfrak{g}_\alpha := \{x \in \mathfrak{g} \mid \operatorname{ad} h(x) = [h, x] = \alpha(h)x \ \forall h \in \mathfrak{h}\},$$

where $\alpha \in \mathfrak{h}^*$. Indeed, if $x \in \mathfrak{g}$ is a common eigenvector for all $h \in \mathfrak{h}$, then $[h, x] = \lambda_h x$ for some $\lambda_h \in K$. Clearly, the map $h \longmapsto \lambda_h$ is linear, hence an element in $\mathfrak{h}^*$.

Note that by definition $\mathfrak{g}_0 = \{x \in \mathfrak{g} \mid [h, x] = 0 \ \forall h \in \mathfrak{h}\} = C_{\mathfrak{g}}(\mathfrak{h})$. By Lemma 12.1, $\mathfrak{h} \subset \mathfrak{g}_0$.

Definition. Let $\mathfrak{g}$ be a semisimple Lie algebra and $\mathfrak{h}$ a maximal toral subalgebra. The set of all $0 \neq \alpha \in \mathfrak{h}^*$ for which $\mathfrak{g}_\alpha \neq 0$ will be denoted by Φ; the (finitely many) elements of Φ are called the *roots* of $\mathfrak{g}$ relative to $\mathfrak{h}$. The decomposition

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$$

is called the *root space decomposition* or the *Cartan decomposition.*

{p:Cartandec}

Proposition 12.2. *Let $\mathfrak{g}$, $\mathfrak{h}$ and Φ be as above. Then*

(1) *For all $\alpha, \beta \in \mathfrak{h}^*$, we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$.*
(2) *If $x \in \mathfrak{g}_\alpha$ for $\alpha \neq 0$, then $\operatorname{ad} x$ is nilpotent.*
(3) *We have $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ for $\alpha, \beta \in \mathfrak{h}^*$ and $\alpha \neq -\beta$.*

Proof. Let $x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_\beta$ and $h \in \mathfrak{h}$. Then

$$\begin{aligned} \operatorname{ad} h[x, y] &= [h, [x, y]] = -[x, [y, h]] - [y, [h, x]] \\ &= [x, [h, y]] + [[h, x], y] = \beta(h)[x, y] + \alpha(h)[x, y] \\ &= (\alpha + \beta)(h)[x, y]. \end{aligned}$$

This proves (1). To see (2), let $x \in \mathfrak{g}_\alpha$ and write any $y \in \mathfrak{g}$ as $y_0 + \sum_{\beta \in \Phi} y_\beta$ with respect to the Cartan decomposition. By (1), $\operatorname{ad} x(y_\beta) \in \mathfrak{g}_{\alpha+\beta}$, $\operatorname{ad} x^2(y) \in \mathfrak{g}_{2\alpha+\beta}$ and so on. Hence, by the finiteness of the direct sum, $\operatorname{ad} x$ is nilpotent on $\bigoplus\limits_{\beta \in \Phi} \mathfrak{g}_\beta$. On the other hand, $\operatorname{ad} x(y_0) \in \mathfrak{g}_\alpha$ so by the same argument $\operatorname{ad} x$ is nilpotent on $\mathfrak{g}_0$. Therefore, $\operatorname{ad} x$ is indeed nilpotent.

Finally, to see (3), let $h \in \mathfrak{h}$ be an element such that $(\alpha+\beta)(h) \neq 0$ and use the associativity of the Killing form to compute, for $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$:

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

hence $(\alpha+\beta)(h)\kappa(x,y) = 0$, so $\kappa(x,y) = 0$. ☐

{c:Cartandec}

Corollary 12.3. *The restriction of the Killing form to $\mathfrak{g}_0 = C_{\mathfrak{g}}(\mathfrak{h})$ is nondegenerate.*

Proof. By Proposition 12.2(3), $\kappa(\mathfrak{g}_0, \mathfrak{g}_\alpha) = 0$ for all $\alpha \in \Phi$. If $x \in \mathfrak{g}_0$ is orthogonal to $\mathfrak{g}_0$, then $\kappa(x, \mathfrak{g}) = 0$. But κ is nondegenerate by Theorem 8.5, so $x = 0$. ☐

We have seen that $\mathfrak{h} \in \mathfrak{g}_0 = C_{\mathfrak{g}}(\mathfrak{h})$. We will now show equality. For this record the following obvious fact.

($\diamond$): If x, y are commuting elements in a ring such that y is nilpotent, then xy is nilpotent.

{p:ToralCentr}

Proposition 12.4. *Let $\mathfrak{h}$ be a maximal toral subalgebra of $\mathfrak{g}$. Then $\mathfrak{h} = C_{\mathfrak{g}}(\mathfrak{h})$.*

Proof. Write $C = C_{\mathfrak{g}}(\mathfrak{h})$.

Claim 1: The algebra C contains the semisimple and nilpotent parts of its elements.

Indeed, $x \in C$ if and only if $\operatorname{ad} x(\mathfrak{h}) = 0$. By Proposition 6.2(3), the same holds for the semisimple and nilpotent parts of $\operatorname{ad} x$, that is, $(\operatorname{ad} x)_s$ and $(\operatorname{ad}_x)_n$. By the discussion before Theorem 10.2, $(\operatorname{ad} x)_s = \operatorname{ad} x_s$ and similarly for $(\operatorname{ad}_x)_n$. Hence, $\operatorname{ad} x_s$ and $\operatorname{ad} x_n$ map $\mathfrak{h}$ to 0, so they are in C by definition of C.

Claim 2: All semisimple elements of C lie in $\mathfrak{h}$.

Let x be a semisimple element in C. Then $\mathfrak{h} + Kx$ is an abelian subalgebra of $\mathfrak{g}$ and it is toral, since the sum of commuting semisimple elements is again semisimple. Since $\mathfrak{h}$ is maximal, $\mathfrak{h} + Kx = \mathfrak{h}$, hence $x \in \mathfrak{h}$.

Claim 3: The restriction of the Killing form κ to $\mathfrak{h}$ is nondegenerate.

Let $h \in \mathfrak{h}$ be an element such that $\kappa(h, \mathfrak{h}) = 0$. If $x \in C$ is a nilpotent element (recall that this means that $\operatorname{ad} x$ is nilpotent), then xy is nilpotent for all $y \in \mathfrak{h}$ by ($\diamond$), because $[x, \mathfrak{h}] = 0$, so in particular $\operatorname{ad} x$ commutes with $\operatorname{ad} y$. Therefore, $\kappa(x,y) = \operatorname{tr}(\operatorname{ad} x \operatorname{ad} y) = 0$ for all $y \in \mathfrak{h}$. Now write any $c \in C$ as $c = s + n$, where s is semisimple and n is nilpotent. By Claim 2, $s \in \mathfrak{h}$, so $\kappa(h, s) = 0$. By Claim 1, $n \in C$ and n is nilpotent, so by the previous argument, $\kappa(h, n) = 0$. In summary, $\kappa(h, c) = 0$ for all $c \in C$. But $\mathfrak{h} \subset C$ and κ is nondegenerate on C by Corollary 12.3, which forces $h = 0$.

Claim 4: The algebra C is nilpotent.

Let $x \in C$. If x is semisimple, then $x \in \mathfrak{h}$ by Claim 2, so $\operatorname{ad}_C x$ is zero and thus nilpotent. If $x \in C$ is nilpotent, then $\operatorname{ad}_C x$ is of course also nilpotent. Thus, let $x \in C$ be arbitrary and write $x = x_s + x_n$. By Claim 1, x_s and x_n are in C, hence they are nilpotent, being a sum of nilpotent elements. We conclude by Engel's Theorem 4.13.

Claim 5: $\mathfrak{h} \cap [C, C] = 0$.

By definition, $[\mathfrak{h}, C] = 0$. By the associativity of κ, we have $\kappa(\mathfrak{h}, [C, C]) = \kappa(C, [\mathfrak{h}, C]) = 0$. So if $z \in \mathfrak{h} \cap [C, C]$, then $\kappa(h, z) = 0$ for all $h \in \mathfrak{h}$, forcing $z = 0$ by Claim 3.

Claim 6: The algebra C is abelian.

Assume the converse. Then $[C, C] \neq 0$. By Claim 4, C is nilpotent. We now claim that $Z(C) \cap [C, C] \neq 0$. This follows from the following general statement.

$(*)$: If D is a nilpotent Lie algebra and $0 \neq I$ is an ideal in D, then $I \cap Z(D) \neq 0$.

Indeed, D acts on I via the adjoint representation, so by Lemma 4.16 there exists a vector $0 \neq v \in I$ such that $D.v = 0$. In other words, $v \in Z(D)$.

So, applying $(*)$ to $I = [C, C]$ gives that $Z(C) \cap [C, C] \neq 0$. Let $z \in Z(C) \cap [C, C]$ be a nonzero element. Since $z \in [C, C]$, by Claim 5 $z \notin \mathfrak{h}$, so by Claim 2, z is not semisimple. Thus its nilpotent part n is nonzero and lies in C by Claim 1. Since $z \in Z(C)$, $n \in Z(C)$ by Proposition 6.2(2). Then $\kappa(n, C) = 0$ by $(\diamond)$, contradicting Corollary 12.3.

Claim 7: $\mathfrak{h} = C$.

Otherwise, C contains a nonzero nilpotent element x by a combination of Claims 1 and 2. By Claim 6 and $(\diamond)$, $\kappa(x, y) = \mathrm{tr}(\mathrm{ad}\, x\, \mathrm{ad}\, y) = 0$ for all $y \in C$, contradicting Corollary 12.3. □

{c:Kappah}

Corollary 12.5. *The restriction of κ to $\mathfrak{h}$ is nondegenerate.* □

Since κ is nondegenerate when restricted to $\mathfrak{h}$, it identifies $\mathfrak{h}$ and $\mathfrak{h}^*$, that is, the map

$$\mathfrak{h} \longrightarrow \mathfrak{h}^*, \quad h \longmapsto \kappa(-, h)$$

is an isomorphism. In other words, for all $\phi \in \mathfrak{h}^*$ there exists a unique $t_\phi \in \mathfrak{h}$ such that

$$\phi(h) = \kappa(t_\phi, h) \quad \forall h \in \mathfrak{h}.$$

{p:InteProp}

Proposition 12.6. *Let $\mathfrak{g}$, $\mathfrak{h}$, Φ be as above.*

(1) *Φ spans $\mathfrak{h}^*$.*
(2) *If $\alpha \in \Phi$, then $-\alpha \in \Phi$.*
(3) *If $\alpha \in \Phi$, $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_{-\alpha}$, then $[x, y] = \kappa(x, y) t_\alpha$.*
(4) *If $\alpha \in \Phi$, then $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ is one dimensional, with basis t_α.*
(5) *$\alpha(t_\alpha) = \kappa(t_\alpha, t_\alpha) \neq 0$ for any $\alpha \in \Phi$.*
(6) *If $\alpha \in \Phi$ and $0 \neq x_\alpha \in \mathfrak{g}_\alpha$, then there exists an element $y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $x_\alpha, y_\alpha, h_\alpha := [x_\alpha, y_\alpha]$ span a three dimensional subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}(2, K)$ via the map $x_\alpha \longmapsto X$, $y_\alpha \longmapsto Y$ and $h_\alpha \longmapsto H$.*
(7) *$h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$, $h_\alpha = -h_{-\alpha}$.*

Proof. (1) If Φ does not span $\mathfrak{h}^*$, then, using the isomorphism $\mathfrak{h} \simeq \mathfrak{h}^*$ given by κ, there exists an element $0 \neq h \in \mathfrak{h}$ such that $\alpha(h) = 0$ for all $\alpha \in \Phi$. In other words, $[h, \mathfrak{g}_\alpha] = 0$ for all $\alpha \in \Phi$. But $[h, \mathfrak{h}] = 0$ since $\mathfrak{h}$ is abelian, so $[h, \mathfrak{g}] = 0$, that is, $h \in Z(\mathfrak{g}) = 0$, contradiction.

(2) If $\alpha \in \Phi$, hen $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ for all $\beta \neq -\alpha \in \mathfrak{h}^*$ by Proposition 12.2(3). If $-\alpha \notin \Phi$, then $\mathfrak{g}_{-\alpha} = 0$, so $\kappa(\mathfrak{g}_\alpha, g_\beta) = 0$ for all $\beta \in \mathfrak{h}^*$, hence $\kappa(\mathfrak{g}_\alpha, \mathfrak{g}) = 0$, contradiction to the nondegeneracy of κ.

(3) Let $\alpha \in \Phi$, $x \in g_\alpha$, $y \in \mathfrak{g}_{-\alpha}$. Let $h \in \mathfrak{h}$ be arbitrary. Since κ is associative, we have

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

so $\kappa(h,[x,y]-\kappa(x,y)t_\alpha) = 0$ for all h. Since $[x,y]$ and $\kappa(x,y)t_\alpha$ are both in $\mathfrak{h}$, the nondegeneracy of κ on $\mathfrak{h}$ implies that $[x,y]-\kappa(x,y)t_\alpha = 0$.

(4) We have seen in (3) that t_α spans $[\mathfrak{g}_\alpha,\mathfrak{g}_{-\alpha}]$, assuming of course $[\mathfrak{g}_\alpha,\mathfrak{g}_{-\alpha}] \neq 0$. Let $0 \neq x \in \mathfrak{g}_\alpha$. If $\kappa(x,\mathfrak{g}_{-\alpha}) = 0$, then, as in the proof of (2), one immediately sees that $\kappa(x,\mathfrak{g}) = 0$, hence $x = 0$, contradiction. Therefore, there exists an element $0 \neq y \in \mathfrak{g}_{-\alpha}$ such that $\kappa(x,y) \neq 0$.

(5) Suppose $\alpha(t_\alpha) = 0$. Then $\alpha(t_\alpha)x = [t_\alpha,x] = 0 = [t_\alpha,y]$ for all $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_{-\alpha}$ (recall that $t_\alpha \in \mathfrak{g}_0$). As in (4) one can find x,y such that $\kappa(x,y) \neq 0$. Scaling one of them, we can wlog assume that $\kappa(x,y) = 1$. Then $[x,y] = t_\alpha$ by (3). It follows that $S = \mathrm{span}_K(x,y,t_\alpha)$ is a three dimensional solvable subalgebra of $\mathfrak{g}$, because $S^{(1)} = \mathrm{span}_K(t_\alpha)$, hence $S^{(2)} = 0$. One checks that $S \simeq \mathrm{ad}_\mathfrak{g} S \subset \mathfrak{gl}(\mathfrak{g})$. By Corollary 7.4, $\mathrm{ad}_\mathfrak{g} x$ is nilpotent for any $x \in [S,S]$. In particular, $\mathrm{ad}_\mathfrak{g} t_\alpha$ is nilpotent. But t_α is also semisimple (being in $\mathfrak{h}$), hence $\mathrm{ad}_\mathfrak{g} t_\alpha = 0$. Therefore, $t_\alpha \in Z(\mathfrak{g}) = 0$, contradiction.

(6) Let $0 \neq x \in \mathfrak{g}_\alpha$ and find $y \in \mathfrak{g}_{-\alpha}$ such that $\kappa(x_\alpha,y_\alpha) = \frac{2}{\kappa(t_\alpha,t_\alpha)}$ which is possible by (4) and (5). Set $h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha,t_\alpha)}$ and note that then $[x_\alpha,y_\alpha] = h_\alpha$ by (3). Also note in passing that $\alpha(h_\alpha) = 2$ by (5).

We now have

$$[h_\alpha,x_\alpha] = \frac{2}{\alpha(t_\alpha)}[t_\alpha,x_\alpha] = \frac{2\alpha(t_\alpha)}{\alpha(t_\alpha)}x_\alpha = 2x_\alpha,$$

where we used that $t_\alpha \in \mathfrak{h}$, $x_\alpha \in \mathfrak{g}_\alpha$ for the second equality. Similarly one shows that $[h_\alpha,y_\alpha] = -2y_\alpha$. Therefore, x_α, y_α and h_α span a three dimensional subalgebra of $\mathfrak{g}$ which has the same multiplication table as $\mathfrak{sl}(2,K)$.

(7) The first assertion was the definition of h_α.

We defined t_α by the equation $\kappa(t_\alpha,h) = \alpha(h)$ for all $h \in \mathfrak{h}$. This immediately shows that $t_\alpha = -t_{-\alpha}$ and hence $h_\alpha = -h_{-\alpha}$.

□ {r:Coroot}

Remark 12.7. We have seen in the course of the proof that for every root $\alpha \in \Phi \subset \mathfrak{h}^*$ there exists an element $h_\alpha \in [\mathfrak{g}_\alpha,\mathfrak{g}_{-\alpha}] \subset \mathfrak{h}$ with the property that $\alpha(h_\alpha) = 2$. Such an element is frequently called the *coroot* or *dual root* and denoted by $\alpha^\vee$.

We now prove the following

{l:RootTrace}

Lemma 12.8. *Let $\mathfrak{g}$, $\mathfrak{h}$, Φ be as above. If $\alpha \in \Phi$, then* $\dim_K \mathfrak{g}_\alpha = 1$ *and* $\mathbb{Z}\alpha \cap \Phi = \{\alpha,-\alpha\}$.

Proof. Pick $x_\alpha \in \mathfrak{g}_\alpha$ and $y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $[x_\alpha,y_\alpha] = h_\alpha$. Consider the space Q spanned over K by y_α, h_α and all the spaces $\mathfrak{g}_{t\alpha}$ for $t \geq 1$ ($t \in \mathbb{N}$). Note that $\mathrm{ad}\, x_\alpha(Q) \subset$

Q, by Proposition 12.2(1). Similarly, $\operatorname{ad} h_\alpha$ and $\operatorname{ad} y_\alpha$ leave Q invariant, since $\operatorname{ad} h_\alpha$ acts diagonally on all components involved and for $\operatorname{ad} y_\alpha$ we again invoke Proposition 12.2(1) (and note that it acts trivially on Ky_α). Now $\operatorname{ad} h_\alpha = \operatorname{ad}[x_\alpha, y_\alpha] = [\operatorname{ad} x_\alpha, \operatorname{ad} y_\alpha]$. Therefore, $\operatorname{tr}(\operatorname{ad} h_\alpha) = 0$. On the other hand, h_α acts on Q diagonally, namely via -2 on Ky_α, via 0 on Kh_α and on $\mathfrak{g}_{t\alpha}$ it acts as a diagonal matrix with entry $2t$. Therefore, setting $d_\beta = \dim_K \mathfrak{g}_\beta$, we have

$$0 = \operatorname{tr} \operatorname{ad} h_\alpha = -2 + 2d_\alpha + 4d_{2\alpha} + 6d_{3\alpha} + \dots$$

Here, the sum is of course finite. It follows that $d_\alpha = 1$ and $d_{t\alpha} = 0$ for all $t \geq 2$. $\square$

{p:RootTrace}

Proposition 12.9. *Let $\mathfrak{g}$, $\mathfrak{h}$, Φ be as above. If $\alpha \in \Phi$, then $K\alpha \cap \Phi = \{\alpha, -\alpha\}$.*

Proof. Consider the subspace M of $\mathfrak{g}$ spanned by $\mathfrak{h}$ along with the spaces $\mathfrak{g}_{c\alpha}$ for $c \in K^*$. By Proposition 12.2(1) this is an S_α-submodule of $\mathfrak{g}$, where $S_\alpha \simeq \mathfrak{sl}(2, K)$ is the subalgebra of $\mathfrak{g}$ constructed in Proposition 12.6(6). Note that by Lemma 12.8, S_α is the span of h_α, $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$.

Now, since M is a representation of S_α, the eigenvalues of h_α on M are integers by Corollary 11.5. By construction, h_α acts on $\mathfrak{g}_{c\alpha}$ with eigenvalue $2c$, so $2c \in \mathbb{Z}$ for all $c \in K^*$ or, in other words, $c \in \frac{1}{2}\mathbb{Z}$.

By Weyl's theorem, $M = \oplus_{n \geq 0} V(n)^{\mu(n)}$, where $V(n)$ is the irreducible S_α-representation of dimension $n+1$. Note that $M_0 = \mathfrak{h}$. By Corollary 11.5, $\sum_n \mu(n) = \dim M_0 + \dim M_1$. It follows that $\sum_{2n \geq 0} \mu(2n) = \dim \mathfrak{h}$. Set $\dim \mathfrak{h} = l$. Note that $\ker(\alpha) \subset \mathfrak{h}$ is $(\dim \mathfrak{h} - 1)$-dimensional (it is the complement to h_α), so $V(0)^{l-1} \subset \mathfrak{h} \subset M$. On the other hand, $V(2) \simeq S_\alpha \subset M$ as well. Write M_{ev} for the part of M belonging to the even n. It follows that $M_{\mathrm{ev}} = \mathfrak{h} \oplus S_\alpha$. In particular, the only even weights occurring in M are 0, 2 and -2. Note that this implies that if α is a root, then 2α is not a root. Therefore, $\frac{1}{2}\alpha$ is not a root (since α is). Hence, 1 is not a weight of h_α in M (because $\alpha(h_\alpha) = 2$ and $\frac{\alpha}{2}$ is not a root, there is no element $\beta \in \Phi$ such that $\beta(h_\alpha) = 1$), implying that $M = M_{\mathrm{ev}} = \mathfrak{h} \oplus S_\alpha$. The latter shows that $K\alpha \cap \Phi = \{\pm\alpha\}$. $\square$

Now let us understand how the algebra S_α acts on the root spaces $\mathfrak{g}_\beta$ if $\beta \neq \pm\alpha$. Set $N = \oplus_{i \in \mathbb{Z}} \mathfrak{g}_{\beta + i\alpha}$. By Lemma 12.8, the space $\mathfrak{g}_{\beta+i\alpha}$ is one dimensional if $\beta + i\alpha$ is a root. Also note that $\beta + i\alpha \neq 0$ by choice of β and Proposition 12.9. Similar arguments as in the case of M show that N is an S_α-submodule of $\mathfrak{g}$. If $\beta + i\alpha \in \Phi$, then the action of h_α on the one dimensional space $\mathfrak{g}_{\beta+i\alpha}$ is given by $\beta(h_\alpha) + i\alpha(h_\alpha) = \beta(h_\alpha) + 2i \in \mathbb{Z}$ by Remark 12.7. Hence, the eigenvalues of h_α on N are all of this form. In particular, either 0 or 1 can occur as a weight. By Corollary 11.5 this implies that the number of summands in a decomposition of N into irreducible submodules is 1, which is the dimension of the eigenspace to eigenvalue 0 or 1. In other words, N is irreducible. These arguments eventually lead to

{p:InteRoots}

Proposition 12.10. *Let $\mathfrak{g}$, $\mathfrak{h}$, Φ be as above. Then*

(1) *If $\alpha \in \Phi$, then $\dim \mathfrak{g}_\alpha = 1$. In particular, the algebra S_α constructed in Proposition 12.6(6) is the direct sum of $\mathfrak{g}_\alpha$, $\mathfrak{g}_{-\alpha}$ and $H_\alpha = [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$. Furthermore,*

for a given $0 \neq x_alpha \in \mathfrak{g}_\alpha$ *there exists a unique* $0 \neq y_alpha \in \mathfrak{g}_{-\alpha}$ *such that* $[x_\alpha, y_\alpha] = h_\alpha$.

(2) *If* $\alpha \in \Phi$*, then* $K\alpha \cap \Phi = \{\pm\alpha\}$.

(3) *If* $\alpha, \beta \in \Phi$*, then* $\beta(h_\alpha) \in \mathbb{Z}$ *and* $\beta - \beta(h_\alpha)\alpha \in \Phi$.

(4) *If* $\alpha, \beta, \alpha+\beta \in \Phi$*, then* $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$.

(5) *Let* $\alpha, \beta \in \Phi$*,* $\beta \neq \pm\alpha$*. Let* r *and* q *be the largest integers for which* $\beta - r\alpha$ *and* $\beta + q\alpha$ *are roots. Then* $\beta + i\alpha \in \Phi$ *for all* $-r \leq i \leq q$ *and* $\beta(h_\alpha) = r - q$.

(6) $\mathfrak{g}$ *is generated by the root spaces* $\mathfrak{g}_\alpha$ *as a Lie algebra.*

Proof. Statements (1) and (2) are Lemma 12.8 and Proposition 12.9, respectively.

To see (4) note that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$ by Proposition 12.2(1). Now $\mathfrak{g}_\alpha \subset S_\alpha$ acts on N, hence on $\mathfrak{g}_\beta \subset N$ and N is irreducible. By Lemma 11.3(3), the bracket $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta]$ is not trivial. Since all three spaces occurring in (4) are one dimensional, this shows the claim.

To prove (5), just note that if $\beta - r\alpha$ is a root, then $\mathfrak{g}_{\beta - r\alpha} \neq 0$ and h_α acts on it with weight $\beta(h_\alpha) - 2r$. Similarly, if $\beta + q\alpha$ is a root, then $\mathfrak{g}_{\beta+q\alpha} \neq 0$ and h_α acts on it with weight $\beta(h_\alpha) + 2q$. Since N is irreducible, the weights of h_α on it form an arithmetic progression with difference 2, so the roots $\beta + i\alpha$ form a string $\beta - r\alpha, \dots, \beta + q\alpha$. Also note that, since the lowest and the highest weight of the action of h_α on N have the same absolute value, we have

$$\beta(h_\alpha) - 2r = -\beta(h_\alpha) - 2q,$$

so $\beta(h_\alpha) = r - q$, showing (5).

To see (6) recall that Φ spans $\mathfrak{h}^*$, so the set $\{h_\alpha \mid \alpha \in \Phi\}$ spans $\mathfrak{h}$, because sending α to h_α defines an isomorphism $\mathfrak{h} \simeq \mathfrak{h}^*$. By (1), any h_α can be generated by the root spaces $\mathfrak{g}_\alpha$, hence (6) holds.

It remains to show (3). We have seen above that, for any $i \in \mathbb{Z}$ such that $\beta + i\alpha \in \Phi$, we have $\beta(h_\alpha) + i\alpha(h_\alpha) = \beta(h_\alpha) + 2i \in \mathbb{Z}$, so $\beta(h_\alpha) \in \mathbb{Z}$. Now, the second claim of (3) is obvious if $\beta = \pm\alpha$, so assume that $\beta \neq \pm\alpha$. The eigenvalue of h_α on $\mathfrak{g}_\beta$ is of course $\beta(h_\alpha)$ and the eigenvalue of h_α on $0 \neq \mathfrak{g}_{-\beta}$ is $-\beta(h_\alpha) = \beta(h_\alpha) - 2\beta(h_\alpha) = (\beta - \beta(h_\alpha)\alpha)(h_\alpha)$. This equation means that there is an eigenvector of h_α with eigenvalue $-\beta(h_\alpha)$, that is, that the space $\mathfrak{g}_{\beta - \beta(h_\alpha)\alpha}$ is not trivial. ☐

Let us rewrite some of the statements we have proved. Recall that if $\alpha \in \Phi$, then $h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$, so for any $\beta \in \Phi$ we have

$$\beta(h_\alpha) = \frac{2\beta(t_\alpha)}{\kappa(t_\alpha, t_\alpha)}.$$

The Killing form is nondegenerate on $\mathfrak{h}$ by Corollary 12.5, hence $\mathfrak{h}^* \simeq \mathfrak{h}$, where $\phi = \kappa(t_\phi, -) \in \mathfrak{h}^*$ is sent to $t_\phi \in \mathfrak{h}$. Therefore, we can define, for any $\gamma, \gamma' \in \mathfrak{h}^*$:

(12.1) {eq:Bilform} $$(\gamma, \gamma') := \kappa(t_\gamma, t_{\gamma'}).$$

Thus the above equation reads

(12.2) {eq:R1} $$\beta(h_\alpha) = \frac{2\beta(t_\alpha)}{\kappa(t_\alpha, t_\alpha)} = \frac{2}{\kappa(t_\alpha, t_\alpha)}\kappa(t_\beta, t_\alpha) = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z}.$$

This element is indeed in $\mathbb{Z}$ by Proposition 12.10(3).

Rewriting the second statement of Proposition 12.10(3) gives

$$\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha \in \Phi. \tag{12.3}$$ {eq:R2}

Lemma 12.11. *Let $\mathfrak{g}$, $\mathfrak{h}$ and Φ be as above. The elements h_α for $\alpha \in \Phi$ span $\mathfrak{h}$.*

Proof. let h be in the orthogonal complement with respect to κ to the span of the h_α in $\mathfrak{h}$. Let $\alpha \in \Phi$ and $x_\alpha \in \mathfrak{g}_\alpha$, $y_\alpha \in \mathfrak{g}_{-\alpha}$ be as above. Then

$$\begin{aligned} 0 = \kappa(h, h_\alpha) &= \kappa(h, [x_\alpha, y_\alpha]) \\ &= \kappa([h, x_\alpha], y_\alpha) = \alpha(h)\kappa(x_\alpha, y_\alpha). \end{aligned}$$

Since $\kappa(x_\alpha, y_\alpha) \neq 0$ by the proof of Proposition 12.6(6), we conclude that $\alpha(h) = 0$ for all $\alpha \in \Phi$. In other words, $[h, \mathfrak{g}_\alpha] = 0$ for all $\alpha \in \Phi$. Since in any case, $[h, \mathfrak{h}] = 0$, we get $[h, \mathfrak{g}] = 0$, hence $h = 0$. □

Proposition 12.12. *Consider the vector space $E^* := \mathfrak{h}_{K'} := \operatorname{span}_{K'}(h_\alpha)_{\alpha\in\Phi}$, where $K' = \mathbb{Q}$ or $\mathbb{R}$. Then for any $t_1, t_2 \in \mathfrak{h}_{K'}$, the number $\kappa(t_1, t_2)$ is in K'. Furthermore, κ is positive definite on $\mathfrak{h}_{K'}$. In other words, $\mathfrak{h}_{K'}$ is a rational/real vector space endowed with a scalar product.*

The same statements hold for $E = \operatorname{span}_{K'}(\alpha)_{\alpha\in\Phi}$.

Proof. Let $h, h' \in \mathfrak{h}$. Then (compare Exercise 2 on Sheet 10)

$$\kappa(h, h') = \operatorname{tr}(\operatorname{ad} h \operatorname{ad} h') = \sum_{\alpha\in\Phi} \alpha(h)\alpha(h'),$$

since the elements h and h' act as multiplication with $\alpha(h)$ and $\alpha(h')$ on $\mathfrak{g}_\alpha$, respectively.

The second claim is clear by Equation (12.1). □

Let us summarise the facts we proved so far.

{t:Roots}

Theorem 12.13. *Let $\mathfrak{g}$ be a semisimple Lie algebra, $\mathfrak{h}$ be a maximal toral subalgebra, $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Phi} \mathfrak{g}_\alpha$ be the Cartan decomposition of $\mathfrak{g}$ with respect to $\mathfrak{h}$ and E be the real vector space spanned by the elements of Φ. Then the following statements hold.*

1. *Φ spans E and $0 \notin \Phi$.*
2. *If $\alpha \in \Phi$, then $-\alpha \in \Phi$, but no other scalar multiple of α is an element in Φ.*
3. *If $\alpha, \beta \in \Phi$, then $\frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$.*
4. *If $\alpha, \beta \in \Phi$, then $\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha \in \Phi$.*

Proof. (1) is Proposition 12.6(1) and the definition of Φ, (2) is Proposition 12.9, (3) is Equation (12.2) and (4) is Equation (12.3). □

This is the first step on the way to proving Theorem 1.12. Namely, we have set up a correspondence which sends a pair $(\mathfrak{g}, \mathfrak{h})$ to a pair (E, Φ). We thus need to classify the pairs (E, Φ), check that the correspondence is actually 1-1 and does not depend on the choice of $\mathfrak{h}$.

13. Root systems

Definition. Let V be a finite dimensional Euclidean vector space with scalar product $(\,,\,)$. A *reflection* in V is an invertible linear transformation whose fixed point locus is some hyperplane.

If $0 \neq v \in V$ is arbitrary, then the corresponding reflection is

$$s_v \colon V \longrightarrow V, \quad w \longmapsto w - \frac{2(w,v)}{(v,v)} v.$$

Note that $s_v(v) = -v$ and $s_v(w) = w$ when $w \in P_v = \{z \in V \mid (z,v) = 0\}$. Clearly, a reflection is an isometry and is of order 2 in $\mathrm{GL}(V)$.

We will frequently write $\langle w, v\rangle$ for $\frac{2(w,v)}{(v,v)}$.

{l:refl}

Lemma 13.1. *Let Φ be a finite set which spans V. Assume $0 \notin \Phi$ and suppose that for all $\alpha \in \Phi$ the reflection s_α maps Φ to itself. If $\sigma \in \mathrm{GL}(V)$ leaves Φ invariant, fixes pointwise a hyperplane P of V and sends some nonzero $\alpha \in \Phi$ to $-\alpha$, then $\sigma = s_\alpha$ and $P = P_\alpha$.*

Proof. Consider the element $\tau := \sigma s_\alpha \in \mathrm{GL}(V)$. Clearly, $\tau(\Phi) = \Phi$, and $\tau(\alpha) = \alpha$, so τ acts as identity on $\mathbb{R}\alpha$. On the other hand, both σ and s_α act as identity on $V/\mathbb{R}\alpha$, hence τ does as well. We conclude that all the eigenvalues of τ are equal to 1, hence its minimal polynomial divides $(T-1)^{\dim(V)}$.

Now, Φ is finite by assumption, so for any $\beta \in \Phi$ there exists a positive integer k_β such that $\tau^{k_\beta}(\beta) = \beta$, because the set $\{\beta, \tau(\beta), \tau^2(\beta), \ldots\}$ is a subset of Φ. Letting $k = \prod_\beta k_\beta$ we see that $\tau^k(\beta) = \beta$ for all $\beta \in \Phi$. Since Φ spans V by assumption, this means that $\tau^k = \mathrm{id}_V$, hence the minimal polynomial of τ divides $T^k - 1$. Therefore, the minimal polynomial of τ divides the greatest common divisor (g.c.d.) of $T^k - 1$ and $(T-1)^{\dim(V)}$. Since the g.c.d. is $T-1$, the minimal polynomial of τ is $T-1$. In other words, $\tau = \mathrm{id}$. $\square$

Definition. Let V be a finite dimensional Euclidean vector space. A subset Φ of V is called a *root system* if the following conditions are satisfied.

(R1) The set Φ is finite, spans E and $0 \notin \Phi$.
(R2) If $\alpha \in \Phi$, then $-\alpha \in \Phi$, but no other scalar multiple of α is an element in Φ.
(R3) If $\alpha, \beta \in \Phi$, then $\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha = s_\alpha(\beta) \in \Phi$.
(R4) If $\alpha, \beta \in \Phi$, then $\frac{2(\beta,\alpha)}{(\alpha,\alpha)} = \langle \beta, \alpha\rangle \in \mathbb{Z}$.

The elements of Φ are called *roots*. The *rank* of Φ is the dimension of V.

In other words, we have seen that the Cartan decomposition provides us with a root system. In the following we will investigate properties of root systems.

Definition. Let Φ be a root system in V. The *Weyl group* of Φ is the subgroup W of $\mathrm{GL}(V)$ generated by the reflections s_α for $\alpha \in \Phi$.

Note that (R3) implies that W permutes the set Φ, hence can be considered as a subgroup of the symmetric group $S_{|\Phi|}$. In particular, W is a finite group.

{l:Weylgroup}

Lemma 13.2. *Let Φ be a root system in V with Weyl group W. If $\tau \in \mathrm{GL}(V)$ leaves Φ invariant, then $\tau s_\alpha \tau^{-1} = s_{\tau(\alpha)}$ for all $\alpha \in \Phi$ and $\langle \beta, \alpha\rangle = \langle \tau(\beta), \tau(\alpha)\rangle$ for all $\alpha, \beta \in \Phi$.*

Proof. Since $s_\alpha(\beta) \in \Phi$, we have $\tau s_\alpha \tau^{-1}(\tau(\beta)) = \tau s_\alpha(\beta) \in \Phi$. Computing the latter explicitly gives

$$\tau s_\alpha(\beta) = \tau(\beta - \langle \beta, \alpha\rangle \alpha) = \tau(\beta) - \langle \beta, \alpha\rangle \tau(\alpha).$$

Since τ is an automorphism of Φ, $\tau(\beta)$ runs over Φ when β runs over Φ. This means that $\tau s_\alpha \tau^{-1}$ leaves Φ invariant. Note that for any $v \in P_\alpha$ we have

$$\tau s_\alpha \tau^{-1}(\tau(v)) = \tau(v)$$

and, furthermore, $\tau s_\alpha \tau^{-1}(\tau(\alpha)) = -\tau(\alpha)$. By Lemma 13.1 ($\tau$ leaves $\tau(P_\alpha)$ invariant), $\tau s_\alpha \tau^{-1} = s_{\tau(\alpha)}$. Now compute

$$\tau(\beta) - \langle \beta, \alpha\rangle \tau(\alpha) = \tau s_\alpha \tau^{-1}(\tau(\beta)) = s_{\tau(\alpha)}(\tau(\beta)) = \tau(\beta) - \langle \tau(\beta), \tau(\alpha)\rangle \tau(\alpha).$$

Therefore, $\langle \beta, \alpha\rangle = \langle \tau(\beta), \tau(\alpha)\rangle$ for all $\alpha, \beta \in \Phi$. □

Definition. Let Φ be a root system in V and Φ' be a root system in V'. The pairs (V, Φ) and (V', Φ') are called *isomorphic* if there exists an isomorphism of vector spaces $f \colon V \longrightarrow V'$ mapping Φ into Φ' such that $\langle f(\beta), f(\alpha)\rangle = \langle \beta, \alpha\rangle$ for all $\alpha, \beta \in \Phi$.

Note that we do not assume f to be an isometry. Also note that

$$s_{f(\alpha)}(f(\beta)) = f(\beta) - \langle f(\beta), f(\alpha)\rangle f(\alpha) = f(\beta) - \langle \beta, \alpha\rangle f(\alpha) = f(s_\alpha(\beta)).$$

In particular, an isomorphism of root systems induces an isomorphism of Weyl groups $W \longrightarrow W'$ given by $\sigma \longmapsto f \circ \sigma \circ f^{-1}$.

Note that by Lemma 13.2 an automorphism of a root system Φ is the same thing as an automorphism of V which leaves Φ invariant.

Example 13.3. If $V \simeq \mathbb{R}$, then the only root system in V is $\Phi = \{1, -1\}$. This root system is denoted by A_1. Its Weyl group is isomorphic to S_2.

Before giving more interesting examples, let us examine the restrictions the axioms impose on a pair of roots. So, let $\alpha, \beta \in \Phi$ be two roots. Recall that the angle θ between two vectors v and w in a Euclidean vector space is defined by

$$\|v\|\|w\| \cos(\theta) = (v, w).$$

Now,

$$\langle \beta, \alpha\rangle = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} = \frac{2(\beta, \alpha)}{\|\alpha\|^2} = 2\frac{\|\beta\|}{\|\alpha\|} \cos(\theta).$$

This implies (exchange α and β in the previous formula) that

$$4\cos^2(\theta) = \langle \beta, \alpha\rangle \langle \alpha, \beta\rangle.$$

Since $0 \leq \cos^2(\theta) \leq 1$, this means that

$$0 \leq \langle \beta, \alpha\rangle \langle \alpha, \beta\rangle \leq 4.$$

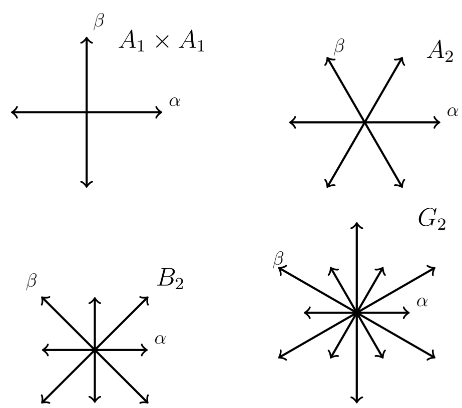
TABLE 1. {tab:1}

$\langle\alpha,\beta\rangle$	$\langle\beta,\alpha\rangle$	$\cos(\theta)$	θ	$\frac{\|\beta\|^2}{\|\alpha\|^2}$
0	0	0	$\frac{\pi}{2}$	??
1	1	$\frac{1}{2}$	$\frac{\pi}{3}$	1
-1	-1	$-\frac{1}{2}$	$\frac{2\pi}{3}$	1
1	2	$\frac{\sqrt{2}}{2}$	$\frac{\pi}{4}$	2
-1	-2	$-\frac{\sqrt{2}}{2}$	$\frac{3\pi}{4}$	2
1	3	$\frac{\sqrt{3}}{2}$	$\frac{\pi}{6}$	3
1	-3	$-\frac{\sqrt{3}}{2}$	$\frac{5\pi}{6}$	3

Now note that $\langle\beta,\alpha\rangle$ and $\langle\alpha,\beta\rangle$ are either both nonpositive or both nonnegative. So, without loss of generality, let us assume that $\langle\alpha,\beta\rangle \in \{0,1,-1\}$. This leads to the following possibilities

{ex:RootsR2}

Example 13.4. Let us describe some root systems in $V = \mathbb{R}^2$. Note that any root system has at least four roots, since it has to span V and for any root its negative is also a root. Using the above table, we get the following root systems.



Here, the $A_1 \times A_1$ corresponds to the first line in the table, A_2 corresponds to the third, B_2 corresponds to the fifth and G_2 corresponds to the seventh line.

Note that, for example, the Weyl group of $A_1 \times A_1$ is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

The following result will be useful later on.

{l:roots1}

Lemma 13.5. *Let Φ be a root system in V and let α, β be nonproportional roots. If $(\alpha,\beta) > 0$, then $\alpha - \beta$ is a root. If $(\alpha,\beta) < 0$, then $\alpha + \beta$ is a root.*

Proof. Note that (α,β) is positive if and only if $\langle\alpha,\beta\rangle$ is. According to the table above, this means that either $\langle\alpha,\beta\rangle = 1$ or $\langle\beta,\alpha\rangle = 1$. If $\langle\alpha,\beta\rangle = 1$, then $s_\beta(\alpha) = \alpha - \beta \in \Phi$ by (R3). If $\langle\beta,\alpha\rangle = 1$, then $s_\alpha(\beta) = \beta - \alpha \in \Phi$, again by (R3), and then $s_{\beta-\alpha}(\beta - \alpha) = \alpha - \beta \in \Phi$.

The second claim follows from the first by applying it to $-\beta$ instead of β. $\square$

Let us put this lemma to use. Consider nonproportional roots α and β and the α-*string through* β which by definition are all roots of the form $\beta + i\alpha$ for $i \in \mathbb{Z}$. Let $r, q \in \mathbb{Z}_{>0}$ be the largest integers for which $\beta - r\alpha \in \Phi$ and $\beta + q\alpha \in \Phi$. If there exists an i with $-r < i > q$ such that $\beta + i\alpha \notin \Phi$, we can find $p < s$ such that $\beta + p\alpha \in \Phi$, $\beta + (p+1)\alpha \notin \Phi$, $\beta + s\alpha \in \Phi$, $\beta + (s-1)\alpha \notin \Phi$ (indeed, if, for instance, i is the only element with the above property, we can simply take $p = i - 1$ and $s = i + 1$, otherwise call the smallest i with this property i_- and the biggest i_+ and set $p = i_- - 1$ and $s = i_+ + 1$).

Since $\alpha + (\beta + p\alpha)$ is not a root, by Lemma 13.5 we have the inequality $(\alpha, \beta + p\alpha) \geq 0$. Similarly we get the inequality $(\alpha, \beta + s\alpha) \leq 0$. A short computation then gives $p \geq s$, a contradiction. Therefore, the α-string through β is unbroken, that is, an element i as above never exists. Any such string is invariant under s_α, which after all only adds and subtracts multiples of α. A moment's thought shows that s_α reverses the string (the string corresponds to the sequence $-r, -r+1, \ldots, q-1, q$; write $z = \langle \beta, \alpha \rangle$, then s_α maps q to $-(q+x)$, $-r$ to $(r-x)$ etc., now use that α gives a bijection), hence

$$\beta - r\alpha = s_\alpha(\beta + q\alpha) = \beta - 2\frac{(\beta + q\alpha, \alpha)}{(\alpha, \alpha)}\alpha = \beta - \langle \beta, \alpha \rangle \alpha - q\alpha$$

thus $r - q = \langle \beta, \alpha \rangle$. Therefore, all strings are of length at most 4.

Definition. Let Φ be a root system in an l-dimensional Euclidean space V. A subset Δ of Φ is called a *base* if

(B1) Δ is a basis of V.

(B2) each root β can be (by (B1) in a unique way) written as $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$ with $k_\alpha \in \mathbb{Z}$ and all $k_\alpha \geq 0$ or all $k_\alpha \leq 0$.

The elements of Δ are called *simple roots*.

The *height* of a root β with respect to a base Δ is defined to be the integer $\mathrm{ht}(\beta) = \sum_{\alpha \in \Delta} k_\alpha$.

A root β is called *positive* if all the k_α are nonnegative and *negative* if all the k_α are nonpositive. Write $\beta > 0$ in the former case and $\beta < 0$ in the latter.

As an example consider the root system $A_1 \times A_1$. Taking e_1 and e_2 as Δ, we see that these are the only positive roots and their negatives are the only negative roots. If we consider the root system A_2 and take α and β from the picture, then their sum is also a positive root. The other three roots are all negative.

The collection of all positive roots with respect to a base Δ is denoted by $\Phi^+(\Delta)$ or simply Φ^+. For the negative roots we will write $\Phi^-(\Delta) = \Phi^-$. Note that if α and β are in Φ^+ and $\alpha + \beta$ is a root, then $\alpha + \beta \in \Phi^+$. Also note that $\Phi^- = -\Phi^+$.

Lemma 13.6. *If Δ is a base of Φ, then $(\alpha, \beta) \leq 0$ whenever $\alpha \neq \beta \in \Delta$, and $\alpha - \beta$ is not in Φ.* {l:AngleNotroot

Proof. Assume to the contrary that $(\alpha, \beta) > 0$. By assumption, $\alpha \neq \beta$ and, of course, α cannot be equal to $-\beta$. By Lemma 13.5, $\alpha - \beta$ is a root. But this contradicts (B2). $\square$

{l:Samesidehype

Lemma 13.7. *Let $0 \neq v \in V$. If $w_1, \dots, w_k$ are vectors in V with the property that $(w_i, v) > 0$ for all i and $(w_i, w_j) \leq 0$ for all $i \neq j$, then the vectors $w_1, \dots, w_k$ are linearly independent.*

Proof. Let $\sum_{i=1}^{k} a_i w_i = 0$ and set $I^+ = \{i \mid a_i > 0\}$, $I^- = \{i \mid a_i < 0\}$. Then $n = \sum_{i \in I^+} a_i w_i = \sum_{i \in I^-} (-a_i) w_i$. Note that

$$(n, n) = \sum_{i \in I^+, j \in I^-} a_i(-a_j)(w_i, w_j) \leq 0,$$

hence $n = 0$. Then

$$0 = (n, v) = \sum_{i \in I^+} a_i (w_i, v),$$

a contradiction, since the right hand side is a sum of positive numbers. Therefore, $I^+ = \emptyset$ and similarly one shows that $I^- = \emptyset$. □

Our next order of business is proving that every root system admits a base. First we need some notation. If $v \in V$, then write

$$\Phi^+(v) = \{\alpha \in \Phi \mid (\alpha, v) > 0\}.$$

It is clear how to define $\Phi^-(v)$.

Definition. Let Φ be a root system of rank l in V.

A vector $v \in V$ is called *regular* if $v \in V \setminus \bigcup_{\alpha \in \Phi} P_\alpha$.

A root $\alpha \in \Phi^+(v)$ is called *decomposable* if $\alpha = \beta_1 + \beta_2$ with $\beta_i \in \Phi^+(v)$. Otherwise, call α *indecomposable*.

Note that if v is regular, then $\Phi = \Phi^+(v) \cup \Phi^-(v)$, since otherwise there would exist an $\alpha \in \Phi$ such that $(\alpha, v) = 0$.

{t:BaseRootSys}

Theorem 13.8. *Every root system Φ has a base.*

In fact, we will prove the following more precise

{t:BaseRootSys1

Theorem 13.9. *If $v \in V$ is a regular vector, then the set $\Delta(v)$ of all indecomposable roots in $\Phi^+(v)$ is a base of Φ. Furthermore, every base of Φ is obtained in this manner.*

Proof. Claim 1: Each root in $\Phi^+(v)$ is a nonnegative $\mathbb{Z}$-linear combination of elements in $\Delta(v)$.

Assume on the contrary that there exists an α in $\Phi^+(v)$ which cannot be written in the stated way. Choose an α such that (α, v) is minimal (possible since $\Phi^+(v) \subset \Phi$ is finite). In particular, $\alpha \notin \Delta(v)$, hence $\alpha = \beta_1 + \beta_2$ with $\beta_i \in \Phi^+(v)$. Since

$$(\alpha, v) = (\beta_1, v) + (\beta_2, v),$$

both summands on the left are positive and (α, v) was chosen minimal, we conclude that both β_1 and β_2 satisfy the claim, hence so does α.

Claim 2: If $\alpha, \beta \in \Delta(v)$, then $(\alpha, \beta) \leq 0$, unless $\alpha = \beta$.

Assume $(\alpha, \beta) > 0$. Then of course, $\alpha \neq \pm\beta$, so by Lemma 13.5, $\alpha - \beta$ is a root. In particular, either $\alpha - \beta \in \Phi^+(v)$ or $\beta - \alpha \in \Phi^+(v)$. In the first case, $\alpha = \beta + (\alpha - \beta)$ which means that α is decomposable. In the second case, $\beta = \alpha + (\beta - \alpha)$ is decomposable. In either case, we get a contradiction.

Claim 3: The set $\Delta(v)$ is linearly independent.

This follows from Claim 2 and Lemma 13.7.

Claim 4: The set $\Delta(v)$ is a base of Φ.

Recall that $\Phi = \Phi^+(v) \cup \Phi^-(v)$. Claim 1 shows that (B2) is satisfied. It is also clear that $\Delta(v)$ spans V, hence by Claim 3, $\Delta(v)$ is indeed a base of Φ.

Claim 5: Each base Δ of Φ is of the form $\Delta(v)$ for some regular v.

Let Δ be given. Choose $v \in V$ so that $(v, \alpha) > 0$ for all $\alpha \in \Delta$. Note that this is possible since the intersection of the finitely many positive open half-spaces associated with any basis of V is non-empty. By (B2), $(v, \alpha) \neq 0$ for all $\alpha \in \Phi$. In other words, v is regular. Furthermore, $\Phi^+ \subset \Phi^+(v)$ (any positive $\mathbb{Z}$-linear combination of elements of Δ has positive scalar product with v) and $\Phi^- \subset \Phi^-(v)$, hence equality holds in both cases (e.g., if $\beta \in \Phi^+(v)$, then $(\beta, v) > 0$, so β must be a positive combination of Δ-elements).

Now let $\alpha \in \Delta$. By the above arguments $\alpha \in \Phi^+(v)$. If α were decomposable, then we would be able to write it as a linear combination of at least two elements in Δ. Hence, α is indecomposable. In other words, $\Delta \subset \Delta(v)$. But both sets have cardinality l, so they are equal. $\square$

Of course, in dimension 1 and 2 it is easy to construct a basis just by looking at the pictures. The reader is invited to construct all possible bases for $l = 2$.

In order to continue, we need a notion from topology. A topological space X is called *connected* if it cannot be written as a disjoint union of open subsets. There is a partial order on connected subsets of X given by inclusion. The maximal elements with respect to this ordering are called *connected components*. These are disjoint. Indeed, if U_1 and U_2 are connected components and they were to intersect, then their union would be a strictly larger connected subset, contradiction. For example, if X is connected, then there is only one connected component, namely X itself. If $X = \{(x, y) \in \mathbb{R}^2 \mid x \neq 0\}$ with the subspace topology, then $X = \{(x, y) \in \mathbb{R}^2 \mid x > 0\} \cup \{(x, y) \in \mathbb{R}^2 \mid x < 0\}$ is the decomposition into connected components.

Definition. Let V be a finite dimensional Euclidean vector space, Φ a root system in V. The connected components of the space

$$V \setminus \bigcup_{\alpha \in \Phi} P_\alpha$$

are called the *Weyl chambers* of V.

In particular, every regular vector v belongs to exactly one Weyl chamber (since the Weyl chambers are disjoint), denoted by $C(v)$. If $C(v) = C(w)$, then v and w lie one the same side of P_α for all $\alpha \in \Phi$ (that is, $\forall \alpha \in \Phi$ either $(v, \alpha) > 0$ and $(w, \beta) > 0$ or

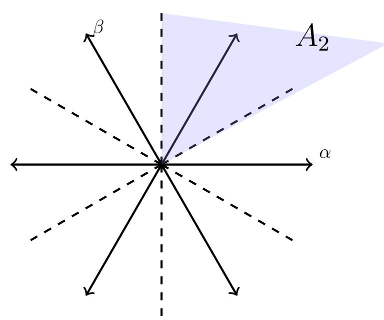
$(v, \alpha) < 0$ and $(w, \beta) < 0$). Therefore, $\Phi^+(v) = \Phi^+(w)$ and $\Delta(v) = \Delta(w)$. We thus have proved

Proposition 13.10. *There exists a one-to-one correspondence between bases of a root system and the Weyl chambers.* □ {p:ChambersBase

Definition. If $\Delta = \Delta(v)$, we thus can write $C(\Delta)$ for $C(v)$. This is called the *fundamental Weyl chamber relative to* Δ.

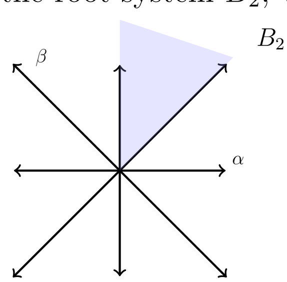
Note that the fundamental chamber consists of all $v \in V$ which satisfy the inequalities $(v, \alpha) > 0$ for all $\alpha \in \Delta$. In other words, $C(v)$ is the intersection of the open half-spaces $\{z \mid (z, \alpha) > 0\}$ for α running over Δ. Clearly, this set is open, being a finite intersection of open sets. It is also *convex*, that is, for $w, w' \in C(v)$, the line segment $(1-t)w + tw'$ is contained in $C(v)$ for all $t \in [0, 1]$.

Example 13.11. Let us consider the situation in the example of A_2. We pick $\{\alpha, \beta\}$ as a base.



In the above picture, the dotted lines are the hyperplanes and the shaded area is the fundamental chamber, which is bound by P_α and P_β.

Example 13.12. Now consider the root system B_2, with base $\{\alpha, \beta\}$.



Again, the shaded region is the fundamental chamber with respect to the base we picked. The hyperplanes here coincide with the directions of the roots, so they are omitted in the picture.

Note that the Weyl group W sends one Weyl chamber onto another. More precisely, $\sigma(C(v)) = C(\sigma(v))$ if $\sigma \in W$ and v is regular. Indeed, $v \in V \setminus \bigcup_{\alpha \in \Phi} P_\alpha$ iff $(v, \alpha) \neq 0$ for all $\alpha \in \Phi$ iff $(\sigma(v), \sigma(\alpha)) \neq 0$ for all $\alpha \in \Phi$. Now just note that $\sigma(\alpha)$ runs over Φ when α runs over Φ.

Furthermore, W permutes bases, since $\sigma \in W$ just sends Δ to $\sigma(\Delta)$ and the latter is clearly again a base. Note that $\sigma(\Delta(v)) = \Delta(\sigma(v))$. Indeed, $\Delta(v)$ consists of the indecomposable roots which pair positively with v and $\Delta(\sigma(v))$ consists of the indecomposable vectors which pair positively with $\sigma(v)$. Now use $(\sigma(v), \sigma(\alpha)) = (v, \alpha)$.

Ultimately our goal is to classify root systems in some reasonable way. For this we will in particular need some helpful results. In the following Δ is a fixed base of a root system Φ in V.

{l:RootsA}

Lemma 13.13. *If α is a positive root, which is not simple, then $\alpha - \beta$ is a positive root for some $\beta \in \Delta$.*

Proof. If $(\alpha, \beta) \leq 0$ for all $\beta \in \Delta$, then by Lemma 13.7 (which applies because $\Delta = \Delta(v)$ for some v and α is a positive root) the set $\Delta \cup \{\alpha\}$ would be linearly independent, which is absurd. Hence, $(\alpha, \beta_0) > 0$ for some $\beta_0 \in \Delta$. By Lemma 13.5 (note that $\alpha \neq \beta_0$ since α is assumed to be non simple), $\alpha - \beta_0$ is a root.

Write $\alpha = \sum_{\gamma\in\Delta} k_\gamma \gamma$, where $k_\gamma \geq 0$ and $k_\gamma > 0$ for some $\gamma \neq \beta_0$ (again, because α is not proportional to β_0). Subtracting β from α yields a linear combination of simple roots, where at least one coefficient is positive, hence they all have to be by (B2). □

{c:RootsA}

Corollary 13.14. *Each $\beta \in \Phi^+$ can be written in the form $\alpha_1 + \ldots + \alpha_k$ with $\alpha_i \in \Delta$ (not necessarily distinct, e.g. $2\alpha + \beta$ is allowed) such that every partial sum is also a root.*

Proof. Use the lemma and induction over the height of β. □

{l:RootsB}

Lemma 13.15. *Let α be a simple root. Then s_α permutes the positive roots other than α.*

Proof. Let $\beta \in \Phi^+(\Delta) \setminus \{\alpha\}$ and write $\beta = \sum_{\gamma\in\Delta} k_\gamma \gamma$ with $k_\gamma \in \mathbb{Z}_{\geq 0}$ for all γ. By assumption, $\beta \neq \pm\alpha$. Therefore, $k_{\gamma_0} \neq 0$ for some $\gamma_0 \neq \alpha$. Now

$$s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha,$$

so the coefficient of γ_0 in $s_\alpha(\beta)$ is still $k_{\gamma_0} > 0$. Hence, by (B2), all the coefficients are positive. Of course, $s_\alpha(\beta) \neq \pm\alpha$. □

{c:RootsB}

Corollary 13.16. *Set $\delta = \frac{1}{2}(\sum_{\beta>0} \beta) = \frac{1}{2}(\alpha + \sum_{\beta>0,\beta\neq\alpha} \beta)$. Then $s_\alpha(\delta) = \delta - \alpha$ for all $\alpha \in \Delta$.*

Proof. Follows from $s_\alpha(\delta) = \frac{1}{2}(-\alpha + \sum_{\beta>0,\beta\neq\alpha} \beta)$. □

{l:RootsC}

Lemma 13.17. *Let $\alpha_1, \ldots, \alpha_t \in \Delta$ (not necessarily distinct). For simplicity write s_i for s_{α_i}. Assume that $s_1 \ldots s_{t-1}(\alpha_t)$ is a negative root. Then for some index $1 \leq p < t$, we have $s_1 \ldots s_t = s_1 \ldots s_{p-1} s_{p+1} \ldots s_{t-1}$.*

Proof. Define $\beta_i = s_{i+1} \ldots s_{t-1}(\alpha_t)$ for $0 \leq i \leq t-2$ and $\beta_{t-1} = \alpha_t$. Note that $s_k(\beta_k) = \beta_{k-1}$ for all possible k.

Now β_0 is negative by assumption, while $\beta_{t-1} = \alpha_t$ is positive. Therefore, there exists a smallest index p such that $\beta_p > 0$. Then $s_p(\beta_p) = \beta_{p-1} < 0$. Since β_p is positive and s_p permutes the positive roots other than α_p by Lemma 13.15, we conclude that $\beta_p = \alpha_p$.

Now set $\tau = s_{p+1} \ldots s_{t-1}$, $\alpha = \alpha_t$. Applying Lemma 13.2 gives $s_{\tau(\alpha)} = \tau s_\alpha \tau^{-1}$. In other words, $s_{\tau(\alpha)} = s_{\beta_p} = s_{\alpha_p} = s_p = \tau s_\alpha \tau^{-1} = s_{p+1} \ldots s_{t-1} s_t s_{t-1} \ldots s_{p+1}$.

Multiplying with τs_t from the right and then with $s_1 \ldots s_{p-1}$ from the left gives the lemma. □

{c:RootsC}

Corollary 13.18. *If $\tau = s_1 \ldots s_t$ is an expression for $\tau \in W$ in terms of reflections corresponding to simple roots with t as small as possible, then $\tau(\alpha_t) < 0$.*

Proof. If $\tau(\alpha_t) > 0$, then $s_1 \ldots s_{t-1}(\alpha_t) = \tau(-\alpha_t) < 0$, hence Lemma 13.17 would provide a shorter expression for τ. □

{t:Weylgroup}

Theorem 13.19. *Let Δ be a base of a root system Φ.*

(1) *If $v \in V$ is a regular vector, then there exists $\sigma \in W$ such that $(\sigma(v), \alpha) > 0$ for all $\alpha \in \Delta$. In other words, $\sigma(v)$ lies in the fundamental chamber, which means that W operates transitively (in general, a group G operates transitively on a space X if for any elements $x, y \in X$ there exists an element $g \in G$ such that $g \cdot x = y$) on Weyl chambers.*
(2) *If Δ' is another base of Φ, then $\sigma(\Delta') = \Delta$ for some $\sigma \in W$. In other words, W acts transitively on bases.*
(3) *If $\alpha \in \Phi$, there exists $\sigma \in W$ such that $\sigma(\alpha) \in \Delta$.*
(4) *The Weyl group W is generated by s_α for $\alpha \in \Delta$.*
(5) *If $\sigma(\Delta) = \Delta$, then $\sigma = \mathrm{id}$. In other words, W acts simply transitively on bases.*

Proof. Denote by W' the subgroup of W generated by s_α for $\alpha \in \Delta$. We will first prove that (1)-(3) holds for W'.

(1) Write $\delta = \frac{1}{2} \sum_{\alpha > 0} \alpha$ and choose $\sigma \in W'$ for which $(\sigma(v), \delta)$ is maximal. For any simple root α the composition $s_\alpha \sigma$ is an element of W', hence
$$(\sigma(v), \delta) \geq (s_\alpha \sigma(v), \delta) = (\sigma(v), s_\alpha(\delta)) = (\sigma(v), \delta) - (\sigma(v), \alpha),$$
where we used Corollary 13.16 for the last equality. This implies that $(\sigma(v), \alpha) \geq 0$ for all $\alpha \in \Delta$. Since v is a regular vector, this inequality is actually strict. Indeed, if $0 = (\sigma(v), \alpha) = (v, \sigma^{-1}(\alpha))$, then this is a contradiction, since $\sigma^{-1}(\alpha) \in \Phi$ and v is regular. Therefore, $\sigma(v)$ lies in the fundamental chamber with respect to Δ.
(2) Proposition 13.10 established a 1-1-correspondence between Weyl chambers and bases, hence (1) implies (2).
(3) Using (2), it is enough to show that α is contained in some base of Φ. If $\beta \neq \pm\alpha$, then $P_\beta \neq P_\alpha$. Therefore, there exists $\gamma \in P_\alpha$, $\gamma \notin P_\beta$ $(\beta \neq \pm\alpha)$ (note that the union of the proper intersections $P_\alpha \cap P_\beta$ cannot be P_α). Choose v' such that $(v', \alpha) = \epsilon > 0$ while $|(v', \beta)| > \epsilon$ for all $\beta \neq \pm\alpha$. The first inequality shows that $\alpha \in \Phi^+(v')$, while the second shows that $\alpha \in \Delta(v')$ (some of the β's are

in $\Delta(v')$ and any linear combination of these will, by construction, have larger scalar product with v' than α).

(4) It is enough to show that for any $\alpha \in \Phi$ the reflection s_α is in W'. Use (3) to find $\tau \in W'$ such that $\tau(\alpha) \in \Delta$. Then $s_{\tau(\alpha)} = \tau s_\alpha \tau^{-1}$, hence $s_\alpha = \tau^{-1} s_{\tau(\alpha)} \tau \in W'$. Thus, we have shown that $W' = W$.

(5) Let $\sigma(\Delta) = \Delta$. If $\sigma \neq \mathrm{id}$, then by (4) we can write σ as a product of reflections in simple roots and we can choose a minimal such expression. But this contradicts Corollary 13.18.

□

So, in particular, the theorem says that the Weyl group W of Φ is generated by the reflections in simple roots with respect to any base. Furthermore, W acts simply transitively (so the action is not only transitive, but also free, that is, all stabilisers are trivial) on the Weyl chambers.

Definition. Let Δ be a base of a root system Φ and let τ be an element of the Weyl group of Φ. Write $\tau = s_{\alpha_1} \dots s_{\alpha_t}$ as a minimal product of reflections with $\alpha_i \in \Delta$. Such an expression will be called *reduced* and its length $t =: l(\tau)$ will be called the *length* of τ with respect to Δ. We set $l(\mathrm{id}) = 0$.

{p:length}

Proposition 13.20. *Let Δ be a base of a root system Φ and let τ be an element of the Weyl group of Φ. Write $n(\tau)$ for the number of positive roots α for which $\tau(\alpha) < 0$. Then $l(\tau) = n(\tau)$.*

Proof. We use induction over the length. If $l(\tau) = 0$, then $\tau = \mathrm{id}$, so $n(\tau) = 0$ as well.

Write $\tau = s_{\alpha_1} \dots s_{\alpha_t}$ and assume this expression is reduced. Set $\alpha = \alpha_t$. By Corollary 13.18, $\tau(\alpha) < 0$. By Lemma 13.15, $n(s_\alpha) = 1$, so $n(\tau s_\alpha) = n(\tau) - 1$ (s_α permutes the positive roots and sends α to its negative). On the other hand, $l(\tau s_\alpha) = l(\tau) - 1$ by choice of α. By the induction hypothesis, $l(\tau s_\alpha) = n(\tau s_\alpha)$, which proves the proposition. □

Before we prove the last result in this section, recall that the closure $\overline{A}$ of a subset A of a topological space X is the smallest closed subset contanining A. Of course, $A = \overline{A}$ if A is closed. Otherwise the closure is simply the intersection of all closed subsets containing A. For example, the closure of the open interval $(0,1)$ in $\mathbb{R}$ is the closed interval $[0,1]$.

{p:ClosureFund}

Proposition 13.21. *Let Δ be a base of a root system Φ and let λ, μ be two elements in the closure of the fundamental chamber $C(\Delta)$. If $\tau\lambda = \mu$ for some $\tau \in W$, then τ is a product of simple reflections which fix λ. In particular, $\lambda = \mu$.*

Proof. Yet again we use induction over the length of τ, the case $l(\tau) = 0$ being obvious. Assume $l(\tau) > 0$. By Proposition 13.20, $n(\tau) > 0$, so τ sends some positive root to a negative root. In particular, τ cannot send all simple roots to positive roots. Assume that $\tau\alpha_0 < 0$, $\alpha_0 \in \Delta$. Since $\mu \in C(\Delta)$ we have $0 \geq (\mu, \tau\alpha_0)$. The latter term is equal to $(\tau^{-1}\mu, \alpha_0) = (\lambda, \alpha_0)$. The right hand side is ≥ 0, since $\lambda \in C(\Delta)$. Combining everything, we see that $(\lambda, \alpha_0) = 0$, that is, $\lambda \in P_{\alpha_0}$. In other words, $s_{\alpha_0}\lambda = \lambda$, so $\tau s_{\alpha_0}\lambda = \mu$. By the same argument as in the proof of Proposition 13.20, $l(\tau s_{\alpha_0}) = l(\tau) - 1$, hence we are done by induction. $\square$

14. Classification of root systems

Throughout this section we fix an n-dimensional Euclidean vector space V and a root system Φ in V. As before, Δ denotes a base of Φ and W its Weyl group.

Definition. A root system Φ is called *irreducible* if it cannot be partitioned into a union of two proper subsets Φ_1, Φ_2 such that each root in Φ_1 is orthogonal to all roots in Φ_2 and vice versa.

Note that $A_1 \times A_1$ is not irreducible, while A_2, B_2 and G_2 are.

Definition. Let Δ be a base of Φ. We call Δ *irreducible* if it cannot be partitioned into a union of two proper subsets Δ_1, Δ_2 such that each root in Δ_1 is orthogonal to all roots in Δ_2 and vice versa.

{p:IrredBase}

Proposition 14.1. *A root system is irreducible if and only if all its bases are irreducible.*

Proof. Assume that $\Phi = \Phi_1 \cup \Phi_2$ is reducible. Let Δ be any base of Φ. If $\Delta \subset \Phi_1$, then $(\alpha, \beta) = 0$ for all $\alpha \in \Delta$ and for all $\beta \in \Phi_2$. Since Δ spans V, this implies that $(V, \Phi_2) = 0$, a contradiction. Similarly one shows that $\Delta \subsetneq \Phi_2$. Therefore, we can write $\Delta = \Delta_1 \cup \Delta_2$, where $\Delta_i = \Delta \cap \Phi_i$. Thus, Δ is not irreducible.

Now assume that Φ is irreducible and $\Delta = \Delta_1 \cup \Delta_2$ with $(\Delta_1, \Delta_2) = 0$. By Theorem 13.19(3), for any $\alpha \in \Phi$ there exists a $\tau \in W$ such that $\tau(\alpha) \in \Delta$. Writing Φ_i for the set of roots mapping to Δ_i under W, we get $\Phi = \Phi_1 \cup \Phi_2$ (we do not yet know that the union is disjoint).

An easy computation shows that if $(\alpha, \beta) = 0$, then $s_\alpha s_\beta = s_\beta s_\alpha$. By Theorem 13.19(4), W is generated by the simple reflections and looking at the formula of a reflection shows that any root in Φ_i is obtained from one in Δ_i by adding and subtracting elements of Δ_i. In other words, $\Phi_i \subset \operatorname{span}(\Delta_i) =: V_i \subset V$. In particular, $(\Phi_1, \Phi_2) = 0$, hence the union is disjoint. Therefore, $\Phi_1 = \emptyset$ or $\Phi_2 = \emptyset$, whence $\Delta_1 = \emptyset$ or $\Delta_2 = \emptyset$. $\square$

When $l = 2$, this gives an easy way of checking whether a root system is irreducible or not.

Recall that a root is called positive if all its coefficients are nonnegative. This actually defines a partial order on Φ: Say that $\alpha < \beta$ if $\beta - \alpha$ is a sum of positive roots or if $\alpha = \beta$.

{l:IrrRootA}

Lemma 14.2. *Let Φ be an irreducible root system. There exists a unique maximal root β, that is, for all $\alpha \neq \beta$ we have* $\operatorname{ht}(\alpha) < \operatorname{ht}(\beta)$ *and $(\beta, \alpha) \geq 0$ for all $\alpha \in \Delta$. If $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$, then $k_\alpha > 0$ for all α.*

Proof. Let $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$ be a maximal element with respect to the ordering $<$. Of course, β is positive. Let $\Delta_1 = \{\alpha \in \Delta \mid k_\alpha > 0\}$ and $\Delta_2 = \{\alpha \in \Delta \mid k_\alpha = 0\}$. Clearly,

$\Delta = \Delta_1 \cup \Delta_2$ is a partition. Assume that $\Delta_2 \neq \emptyset$. Then

$$(\alpha_0, \beta) = (\alpha_0, \sum_\alpha k_\alpha \alpha) \leq 0 \quad \forall \alpha_0 \in \Delta_2$$

by Lemma 13.6. Now Φ is irreducible, so $(\alpha_0, \alpha') < 0$ for some $\alpha' \in \Delta_1$. Therefore, $(\alpha_0, \beta) < 0$. By Lemma 13.5, this implies that $\alpha_0 + \beta$ is a root. This is a contradiction to the maximality of β, hence $\Delta_2 = \emptyset$ and $k_\alpha > 0$ for all α.

Note that we have also showed that $(\alpha, \beta) \geq 0$ for all $\alpha \in \Delta$. Of course, since Δ spans V, at least one scalar product is positive.

Now assume that β' is another maximal root. We already know that it is written as a sum with strictly positive coefficients and $(\alpha, \beta') \geq 0$ for all $\alpha \in \Delta$ (and at least one scalar product is strictly positive). The previous argument shows the existence of an $\alpha_0 \in \Delta$ with $(\alpha_0, \beta) > 0$. Therefore, $(\beta, \beta') > 0$ and $\beta - \beta'$ is a root by Lemma 13.5 unless $\beta = \beta'$. But if $\beta - \beta'$ is a root, then either $\beta < \beta'$ or $\beta' < \beta$, a contradiction. $\square$

To formulate the next result, recall that if a group G acts on a topological space X, then the orbit of an element $x \in X$ under G is the set $\{gx \mid g \in G\}$.

{l:IrrRootB}

Lemma 14.3. *Let Φ be an irreducible root system. Then any non-empty W-invariant subset of V spans V or, in other words, every non-trivial W-invariant subspace is V. In particular, the W-orbit of a root spans V.*

Proof. Let U be a subspace of V invariant under the Weyl group. Consider $U' = \{v \in V \mid (v, U) = 0\}$. This subspace is also invariant under W, since W acts by isometries. Therefore, we can write $V = U \oplus U'$. If $\alpha \in \Phi$, then $s_\alpha(U) = U$, so either $\alpha \in U$ or $U \subset P_\alpha$ by Exercise 1 on Sheet 12. If $\alpha \notin U$, then $U \subset P_\alpha$, hence $\alpha \in U'$ (otherwise write $\alpha = u + u'$, then $(u, u) + (u', u') = (\alpha, \alpha) = (\alpha, u + u') = (\alpha, u') = (u', u')$, hence $(u, u) = 0$, thus $u = 0$). This argument applies to every root, so we can partition Φ into orthogonal subsets, hence one of them is empty. Since Φ spans V, we conclude that $U = V$. $\square$

{l:IrrRootC}

Lemma 14.4. *Let Φ be an irreducible root system. Then at most two root lengths (in the sence of the norm on V) occur in Φ, and all roots of a given length are conjugate under W (that is, we can map any two roots of a given length to each other with an element of W).*

Proof. Let α, β be arbitrary roots. By Lemma 14.3, the orbit of α under W spans V, hence there exists a $\tau \in W$ such that $(\tau(\alpha), \beta) \neq 0$. Looking at the last column of Table 1, we see that the possible ratios of squared lengths of $\tau(\alpha)$ and β are $1, 2, 3, \frac{1}{2}$ and $\frac{1}{3}$. If there were a root with a third length, then the ratio $\frac{3}{2}$ would also appear, contradiction.

To prove the second claim, let α, β have the same length, so $\frac{\|\beta\|^2}{\|\alpha\|^2} = 1$. As above, we can assume, after possibly replacing one by a W-conjugate, that these roots are not orthogonal. If they are the same, we are done. Otherwise, Table 1 shows that $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle = \pm 1$. If necessary we can replace β by $s_\beta(\beta) = -\beta$ to achieve $\langle \alpha, \beta \rangle = 1$.

It follows that

$$s_\alpha s_\beta s_\alpha(\beta) = s_\alpha s_\beta(\beta - \alpha) = s_\alpha(-\beta - \alpha + \beta) = \alpha.$$

□

Definition. If Φ is an irreducible root system, then one calls the roots with the smaller length *short* and the ones with the bigger length *long*. If all roots have the same length, we call them all long.

{l:IrrRootD}

Lemma 14.5. *Let Φ be an irreducible root system with base Δ and with two distinct root lengths. Then the maximal root β whose existence is ensured by Lemma 14.2 is long.*

Proof. Let α be an arbitrary root. We only need to show that $(\beta, \beta) \geq (\alpha, \alpha)$. By Theorem 13.19 we can map α to an element of Δ and by Proposition 13.10 we can then assume that α lies in the closure of the fundamental Weyl chamber $C(\Delta)$. By Lemma 14.2, $\beta - \alpha > 0$, so $(\gamma, \beta - \alpha) \geq 0$ for all $\gamma \in \overline{C(\Delta)}$ (recall that $C(\Delta)$ is the set of all vectors having positive scalar product with all simple roots; now use that $\beta - \alpha$ is a positive combination of the latter). Now by Lemma 14.2 again, $\beta \in \overline{C(\Delta)}$, so setting $\gamma = \beta$ gives $(\beta, \beta) \geq (\beta, \alpha)$. Setting $\gamma = \alpha$ gives $(\alpha, \beta) \geq (\alpha, \alpha)$, and hence the claim. □

Definition. Let Φ be a root system with base Δ and fix an ordering $(\alpha_1, \dots, \alpha_n)$ of the simple roots. The matrix $(\langle \alpha_i, \alpha_j \rangle)_{i,j}$ is called the *Cartan matrix* of Φ. Its entries are called the *Cartan integers*.

{r:CartanInte}

Remark 14.6. Recall that if $\alpha, \beta \in \Delta$ and $\alpha \neq \beta$, then, by Lemma 13.6, $(\alpha, \beta) \leq 0$. In particular, the non-diagonal entries of a Cartan matrix are smaller or equal to 0, since $\langle \beta, \alpha \rangle = \frac{2(\beta,\alpha)}{(\alpha,\alpha)}$.

{e:CartanEx}

Example 14.7. Recall the root systems from Example 13.4. For the following matrices use Table 1.

For $A_1 \times A_1$ we get the matrix $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ (the base elements are orthogonal and of length 1), for A_2 the matrix $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, for B_2 the matrix $\begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}$ and for G_2 the matrix $\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}$.

{p:CartanMatr}

Proposition 14.8. *Let $\Phi' \subset V'$ be a second root system with base $\Delta' = (\alpha'_1, \dots, \alpha'_l)$. If $\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle$ for all i, j, then the bijection $\alpha_i \longrightarrow \alpha'_i$ extends to an isomorphism $f \colon V \longrightarrow V'$ mapping Φ onto Φ' and satisfying $\langle f(\alpha), f(\beta) \rangle = \langle \alpha, \beta \rangle$ for all $\alpha, \beta \in \Phi$. In particular, the Cartan matrix of Φ determines Φ up to isomorphism.*

Proof. Since Δ and Δ' are bases of the respective vector spaces, there exists a unique isomorphism $f \colon V \longrightarrow V'$ mapping $\alpha_i \longrightarrow \alpha'_i$. If α and β are in Δ, then

$$\begin{aligned} s_{f(\alpha)}(f(\beta)) &= s_{\alpha'}(\beta') = \beta' - \langle \beta', \alpha' \rangle \alpha' \\ &= f(\beta) - \langle \beta, \alpha \rangle f(\alpha) = f(\beta - \langle \beta, \alpha \rangle \alpha) \\ &= f(s_\alpha(\beta)). \end{aligned}$$

Hence, for every $\alpha \in \Delta$ we have the equality $s_{f(\alpha)} \circ f = f \circ s_\alpha \in \mathrm{Hom}(V, V')$.

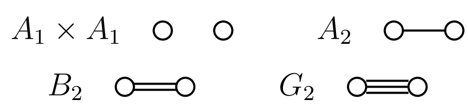
Recall that by Theorem 13.19(4), W is generated by the reflections s_α for $\alpha \in \Delta$ and W' is generated by the reflections $s_{\alpha'}$ for $\alpha' \in \Delta'$. The map $W \longrightarrow W'$, $\tau \longmapsto f \circ \tau \circ f^{-1}$ is an isomorphism, since any τ is a product of simple reflections and the above argument shows that this map sends s_α to $s_{f(\alpha)}$.

Now let $\beta \in \Phi$ be an arbitrary root. There exists $\tau \in W$ such that $\beta = \tau(\alpha)$ for some $\alpha \in \Delta$ (Theorem 13.19(3)). Therefore, $f(\beta) = (f \circ \tau \circ f^{-1})(f(\beta)) \in \Phi'$. Hence, f maps Φ into Φ'. Lastly, we have already seen that f preserves the Cartan integers. $\square$

Definition. The *Coxeter graph* of a root system Φ with respect to an ordered base Δ is a graph having n vertices the ith joined to the jth (for $i \neq j$) by $\langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle$ edges.

If more than one root length occurs and a double or triple edge occurs in the Coxeter graph, we add an arrow pointing to the shorter of the two roots. The resulting figure is called a *Dynkin diagram.*

Example 14.9. These are the Coxeter graphs for the root systems in Example 13.4.



The Dynkin diagrams of $A_1 \times A_1$ and A_2 are the same as the Coxeter graphs. For B_2 one adds an arrow pointing from left to right (this is a little confusing, since in the picture the shorter root comes first, but this can be solved by putting $\alpha = e_1 - e_2$ and $\beta = e_2$), while for G_2 one adds an arrow pointing from right to left.

{p:RootCoxe}

Proposition 14.10. *Any root system Φ in V decomposes, in a unique way, as the union of irreducible root systems in subspaces U_i of V such that $V = U_1 \oplus \ldots \oplus U_t$ is an orthogonal direct sum.*

Proof. Note that Φ is irreducible if and only if its Coxeter graph is connected, that is, any vertex can be connected to any other vertex via a sequence of edges.

In general, the Coxeter graph of Φ will have some connected components. Write Δ_i for the subset of a base Δ of Φ which corresponds to the ith connected component. This gives a partition $\Delta = \Delta_1 \cup \ldots \cup \Delta_t$. Set $U_i = \mathrm{span}(\Delta_i)$. Clearly, $V = U_1 \oplus \ldots \oplus U_t$ is an orthogonal direct sum.

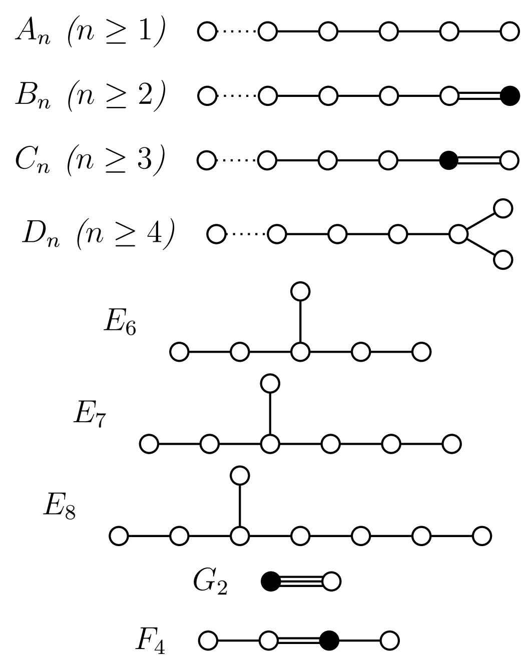
Define Φ_i to be $\mathrm{span}_{\mathbb{Z}}(\Delta_i) \subset U_i$. Of course, this is a root system in U_i. Also note that the Weyl group of Φ_i is the restriction to U_i of the subgroup of the Weyl group generated by the reflections s_α for $\alpha \in \Delta_i$.

If $\beta \notin \Delta_i$, then s_β acts trivially on U_i, compare the proof of Lemma 14.3. Therefore, each U_i is invariant under the Weyl group of Φ. By the proof of Lemma 14.3 again, if β is any root, then either $\beta \in U_i$ or $U_i \subset P_\beta$. In the former case, $\beta \in \Phi_i$, while the latter cannot occur for all i, since P_β has codimension one. Therefore, any root is contained in some Φ_i, hence $\Phi = \Phi_1 \cup \ldots \cup \Phi_t$. $\square$

We now come to the classification result for irreducible root systems. We will omit the proof and focus on the consequences instead.

{t:CarKil}

Theorem 14.11 (Cartan, Killing). *If Φ is an irreducible root system of rank n, then its Dynkin diagram is one of the following:*



In the above diagrams the black edges are the shorter roots, and the numbering of the edges starts on the left. □

Of course, in order to determine the Cartan matrices, one needs to number the edges. For all diagrams except for E_6, E_7 and E_8 we have already given the description. There are different conventions for the remaining three cases, the main question being whether the "upper" vertex should represent the second or the last root.

In the following we will construct the root systems A_n-D_n explicitly. In the following, $V = \mathbb{R}^m$ with the usual scalar product. The standard basis vectors will be denoted by $e_1, \ldots, e_m$. Their span over $\mathbb{Z}$ is of course $\mathbb{Z}^m$. Note that $A = (\mathbb{Z}^m \setminus \{0\}) \cap B_r(0)$ is a finite set for any $1 \leq r < \infty$. In the following we will consider subsets of A which contain vectors of one or two lengths. Obviously, any such subset will satisfy (R1). The choice of lengths will make it obvious that (R2) is also satisfied. The other two axioms will also be easy to check from the explicit description of the given subset.

{e:Al}

Example 14.12. Let $V = \mathbb{R}^{n+1}$ $(n \geq 1)$ and $E = \{v \in V \mid v_1 + \ldots + v_{n+1} = 0\}$. Define Φ to be the set of all vectors α in $\mathbb{Z}^{n+1} \cap E$ satisfying $(\alpha, \alpha) = 2$. It is easy to check that $\Phi = \{e_i - e_j \mid i \neq j\}$ and that Φ is a root system. Consider the vectors $\alpha_i = e_i - e_{i+1}$ for $1 \leq i \leq n$. Is is clear that the set $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ is linearly independent, hence a basis of E. Furthermore, $e_i - e_j = \alpha_i + \ldots + \alpha_{j-1}$ for all $i < j$. These last two statements show that Δ is a base of Φ. Note that $(\alpha_i, \alpha_j) = 0$ if $|i - j| \geq 2$, while $(\alpha_i, \alpha_{i+1}) = -1$. It follows that the Dynkin diagram of Φ is precisely A_n and the Cartan matrix is $2 \cdot \mathrm{Id}_n + \sum_{i=1}^{n-1}(-E_{i,i+1}) + \sum_{i=2}^{n}(-E_{i,i-1})$. Exercise 1 on Sheet 10 tells us that the root system of $\mathfrak{sl}(n+1, \mathbb{C})$ is A_n.

One can check that the reflection s_{α_i} sends e_i to e_{i+1} and vice versa and leaves everything else invariant. Therefore, this reflection corresponds to the transposition $(i, i+1) \in S_{n+1}$ and eventually one shows that the Weyl group of A_n is S_{n+1}.

{e:Bl}

Example 14.13. Let $E = \mathbb{R}^n$, $n \geq 2$ and $\Phi = \{\alpha \in \mathbb{Z}^n \mid (\alpha, \alpha) = 1 \text{ or } 2\}$. One easily checks that Φ consits of the vectors $\pm e_i$ of length 1 and $\pm(e_i \pm e_j)$ for $i \neq j$ (squared length 2). Define $\Delta = \{e_1 - e_2, \ldots, e_{n-1} - e_n, e_n\}$. Clearly, these are independent. Note that $e_i = (e_i - e_{i+1}) + \ldots + (e_{n-1} - e_n) + e_n$. As before, one checks that all the roots $e_i - e_j$ can be expressed by elements in Δ as a linear combination with nonnegative coefficients. Similarly, $e_i + e_j = (e_i - e_{i+1}) + \ldots + (e_{j-1} - e_j) + 2(e_j - e_{j+1}) + \ldots + 2(e_{n-1} - e_n) + 2e_n$. Note that $(e_{n-1} - e_n, e_n) = -1$, so $\langle e_{n-1} - e_n, e_n\rangle = -2$ and $\langle e_n, e_{n-1} - e_n\rangle = -1$. This means that the resulting Dynkin diagram is B_n and the Cartan matrix is $2 \cdot \mathrm{Id}_n + \sum_{i=1}^{n-2}(-E_{i,i+1}) + (-2)E_{n-1,n} + \sum_{i=2}^{n}(-E_{i,i-1})$. The Weyl group turns out to be the semidirect product of S_n and $(\mathbb{Z}/2\mathbb{Z})^n$, where the first factor corresponds to permutations (as in the A_n case) and the second to sign changes of the set $\{e_1, \ldots, e_n\}$.

{e:Cl}

Example 14.14. Let $n \geq 3$, $E = \mathbb{R}^n$, $\Phi = \{\pm 2e_i\} \cup \{\pm(e_i \pm e_j) \mid i \neq j\}$. A base of Φ is $\Delta = \{e_1 - e_2, \ldots, e_{n-1} - e_n, 2e_n\}$. Note that the last element of the base is long. The resulting Dynkin diagram is C_n and the Cartan matrix is $2 \cdot \mathrm{Id}_n + \sum_{i=1}^{n-1}(-E_{i,i+1}) + (-2)E_{n,n-1} + \sum_{i=2}^{n-1}(-E_{i,i-1})$. One can check that the Weyl group is isomorphic to that of B_n. Exercises 2 and 3 on the bonus sheet show that $\mathfrak{sp}(2n, \mathbb{C})$ has C_n as its root system.

{e:Dl}

Example 14.15. Let $n \geq 4$, $E = \mathbb{R}^n$, $\Phi = \{\alpha \in \mathbb{Z}^n \mid (\alpha, \alpha) = 2\} = \{\pm(e_i \pm e_j) \mid i \neq j\}$. For a base take the n independent vectors $e_1 - e_2, \ldots, e_{n-1} - e_n, e_{n-1} + e_n$. Note that, for example, $e_1 + e_2 = e_1 - e_2 + 2(e_2 - e_3) + \ldots + 2(e_{n-2} - e_{n-1}) + (e_{n-1} - e_n) + (e_{n-1} + e_n)$, so this is indeed a base. The resulting Dynkin diagram is of course D_n (note that the last two base elements are orthogonal to each other). The resulting Cartan matrix is $2 \cdot \mathrm{Id}_n + \sum_{i=1}^{n-2}(-E_{i,i+1}) + (-E_{n-2,n}) + \sum_{i=2}^{n-1}(-E_{i,i-1}) + (-E_{n,n-2})$. The Weyl group is the group of permutations and sign changes involving only even numbers of signs of the set $\{e_1, \ldots, e_n\}$ so it is isomorphic to the semidirect product of $(\mathbb{Z}/2\mathbb{Z})^{n-1}$ and S_n.

15. Sketch of the proof of Theorem 1.12

First we investigate the root system of a simple Lie algebra.

{p:irr-root}

Proposition 15.1. *Let $\mathfrak{g}$ be a simple Lie algebra, $\mathfrak{h}$ a maximal toral subalgebra and Φ the corresponding root system. Then Φ is an irreducible root system.*

Proof. Suppose $\Phi = \Phi_1 \cup \Phi_2$ such that $(\alpha, \beta) = 0$ for all $\alpha \in \Phi_1$ and all $\beta \in \Phi_2$. For two such elements we have $(\alpha + \beta, \alpha) \neq 0$ (so $\alpha + \beta \notin \Phi_2$) and $(\alpha + \beta, \beta) \neq 0$ (so $\alpha + \beta \notin \Phi_1$). Thus, $\alpha + \beta \notin \Phi$ and $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta} = 0$. Setting $T = \oplus_{\alpha \in \in \Phi_1}\mathfrak{g}_\alpha$, we thus see that $\oplus_{\beta \in \Phi_2}\mathfrak{g}_\beta \subset C_{\mathfrak{g}}(T)$. In particular, $T \neq \mathfrak{g}$, since $Z(\mathfrak{g}) = 0$. Of course, $\mathfrak{g}_\alpha \subset N_{\mathfrak{g}}(T)$ for all

$\alpha \in \Phi_1$, but by the above argument $\mathfrak{g}_\beta \subset N_\mathfrak{g}(T)$ for all $\beta \in \Phi_2$ as well. Since $\mathfrak{h} \subset N_\mathfrak{g}(T)$ in any case, we conclude that $N_\mathfrak{g}(T) = \mathfrak{g}$. In other words, T is a proper ideal of $\mathfrak{g}$, a contradiction. $\square$

{c:semisimplede

Corollary 15.2. *Let $\mathfrak{g}$ be a semisimple Lie algebra with maximal toral subalgebra $\mathfrak{h}$ and root system Φ. If $\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_t$ is the decomposition of $\mathfrak{g}$ into simple ideals, then $\mathfrak{h}_i = \mathfrak{h} \cap \mathfrak{g}_i$ is a maximal toral subalgebra of $\mathfrak{g}_i$ for all i, and the corresponding root system Φ_i can be regarded canonically as a subsystem of Φ in such a way that $\Phi = \Phi_1 \cup \ldots \cup \Phi_t$.*

Proof. Of course, $\mathfrak{h}_i$ is toral in $\mathfrak{g}_i$ and $\mathfrak{h} = \mathfrak{h}_1 \oplus \ldots \oplus \mathfrak{h}_t$ (compare Exercise 1 on Sheet 9). In fact, $\mathfrak{h}_i$ is maximal toral in $\mathfrak{g}_i$. Indeed, if $\mathfrak{h}_i \subsetneq \widetilde{\mathfrak{h}}_i$ with $\widetilde{\mathfrak{h}}_i$ toral in $\mathfrak{g}_i$, then $\widetilde{\mathfrak{h}}_i$ is toral in $\mathfrak{g}$ and it centralises $\mathfrak{h}_j$ for any $j \neq i$. Thus, $\mathfrak{h}_j$ for $j \neq i$ and $\widetilde{\mathfrak{h}}_i$ generate a toral subalgebra of $\mathfrak{g}$ bigger than $\mathfrak{h}$, a contradiction.

Now consider the root system Φ_i of $(\mathfrak{g}_i, \mathfrak{h}_i)$. If $\alpha \in \Phi_i$, then α can be considered as an element in $\mathfrak{h}^*$ by setting $\alpha(\mathfrak{h}_j) = 0$ for all $j \neq i$. Clearly, $\alpha \in \Phi$ and $\mathfrak{g}_\alpha \subset \mathfrak{g}_i$.

Conversely, if $\alpha \in \Phi$, then there exists an index i such that $[\mathfrak{h}_i, \mathfrak{g}_\alpha] \neq 0$ (otherwise α would be trivial on $\mathfrak{h}$). Thus, $\mathfrak{g}_\alpha \subset \mathfrak{g}_i$, and $\alpha_{|\mathfrak{h}_i}$ is a root of $\mathfrak{g}_i$ relative to $\mathfrak{h}_i$. It follows that $\Phi = \Phi_1 \cup \ldots \cup \Phi_t$. $\square$

The corollary reduces the study of semisimple Lie algebras by their root systems to the study of simple Lie algebras and their irreducible root systems.

Next, we exhibit a smaller set of generators of a semisimple Lie algebra.

{p:simplerootsg

Proposition 15.3. *Let $\mathfrak{g}$ be a semisimple Lie algebra, $\mathfrak{h}$, Φ be as usual. Fix a base Δ of Φ. Then $\mathfrak{g}$ is generated (as a Lie algebra) by the spaces $\mathfrak{g}_\alpha$, $\mathfrak{g}_{-\alpha}$ for $\alpha \in \Delta$.*

Proof. Recall that $\mathfrak{g}$ is generated by the $\mathfrak{g}_\gamma$ for $\gamma \in \Phi$. Let β be an arbitrary positive root. Then β is a sum of elements in Δ and each partial sum is a root. We know by Proposition 12.10(4) that if $\gamma_1, \gamma_2, \gamma_1 + \gamma_2 \in \Phi$, then $[\mathfrak{g}_{\gamma_1}, \mathfrak{g}_{\gamma_2}] = \mathfrak{g}_{\gamma_1+\gamma_2}$. By induction, $\mathfrak{g}_\beta$ is contained in the subalgebra generated by $\mathfrak{g}_\alpha$ for $\alpha \in \Delta$. Similarly, if β is negative, it is contained in the subalgebra generated by $\mathfrak{g}_{-\alpha}$ for $\alpha \in \Delta$. The claim follows. $\square$

The next step of the proof of Theorem 1.12 is the following result whose proof we have to omit.

{t:Cartanconj}

Theorem 15.4. *For any two maximal toral subalgebras $\mathfrak{h}_1$, $\mathfrak{h}_2$ in a semisimple Lie algebra $\mathfrak{g}$ there exists an automorphism ψ of $\mathfrak{g}$ such that $\psi(\mathfrak{h}_1) = \mathfrak{h}_2$.* $\square$

Next, we need to understand how to construct a Lie algebra from a given root system. For this we need the following construction.

Definition. Let V be a vector space. Set $T^0(V) = K$, $T^1(V) = V$,..., $T^m(V) = V \otimes \ldots \otimes V$ (m times), ...

Define an associative product on $T(V) = \bigoplus_{i \geq 0} T^i(V)$ by the rule

$$(v_1 \otimes \ldots \otimes v_k)(w_1 \otimes \ldots \otimes w_l) = v_1 \otimes \ldots \otimes v_k \otimes w_1 \otimes \ldots \otimes w_l,$$

on generators of $T^k(V)$ and $T^l(V)$, respectively, and extend this linearly. This is the *tensor algebra*. Note that there is a canonical inclusion $i \colon V \longrightarrow T(V)$.

An important fact about the tensor algebra is the following statement. If A is an associative unital algebra over K and $\phi\colon V \longrightarrow A$ is a K-linear map, then there exists a unique unit-preserving algebra homomorphism $\psi\colon T(V) \longrightarrow A$ sich that $\psi \circ i = \phi$ ($\psi(v_1 \otimes \ldots \otimes v_k) := \phi(v_1)\ldots\phi(v_k)$).

Definition. The *universal enveloping algebra* of a Lie algebra $\mathfrak{g}$ is a pair (U, j) where $U = U(\mathfrak{g})$ is an associative unital K-algebra and $j\colon \mathfrak{g} \longrightarrow U$ is a linear map satisfying the following property

$$(*)\, j([x,y]) = j(x)j(y) - j(y)j(x).$$

Furthermore, for any associative unital K-algebra A and any linear map $\phi\colon \mathfrak{g} \longrightarrow A$ satisfying $(*)$, there exists a unique algebra homomorphism $\psi\colon U \longrightarrow A$ such that $\psi \circ j = \phi$.

The uniqueness of U is proved in the usual manner. For the existence, let J be the ideal in $T(\mathfrak{g})$ generated by all elements of the form $x \otimes y - y \otimes x - [x,y]$ for $x, y \in \mathfrak{g}$. Define $U(\mathfrak{g}) = T(\mathfrak{g})/J$, let $\pi\colon T(\mathfrak{g}) \longrightarrow U(\mathfrak{g})$ be the canonical map and set $j = \pi \circ i$. Then j clearly satisfies $(*)$. If $\phi\colon \mathfrak{g} \longrightarrow A$ is given, we get an algebra homomorphism $\psi'\colon T(\mathfrak{g}) \longrightarrow A$ such that $\psi' \circ i = \phi$. Since ϕ satisfies $(*)$, J is in the kernel of ψ', hence we get a map ψ satisfying $\psi \circ \pi = \psi'$. Therefore, $\psi \circ j = \psi \circ \pi \circ i = \psi' \circ i = \phi$.

Definition. Let $\mathfrak{g}$ be a Lie algebra generated (as a Lie algebra) by a set X. We say that $\mathfrak{g}$ is *free on* X if, given any map ϕ from X into a Lie algebra $\mathfrak{g}'$, there exists a unique Lie algebra homomorphism $\psi\colon \mathfrak{g} \longrightarrow \mathfrak{g}'$ extending ϕ.

Yet again, uniqueness of such an algebra is obvious. To construct it, take V to be the vector space with basis X and form $T(V)$. This is an associative algebra, hence admits a Lie algebra structure via the commutator. Let $\mathfrak{g}$ be the Lie subalgebra of $T(V)$ generated by X. Given $\phi\colon X \longrightarrow \mathfrak{g}'$, extend ϕ to a map $V \longrightarrow \mathfrak{g}' \subset U(\mathfrak{g}')$. By the universal property of $T(V)$ we then get an algebra homomorphism $T(V) \longrightarrow \mathfrak{g}' \subset U(\mathfrak{g}')$. Restricting this map to $\mathfrak{g}$ gives the required ψ.

A variant of the above construction is as follows. If $\mathfrak{g}$ is free on X and R is the ideal of $\mathfrak{g}$ generated by a finite number of elements $f_1, \ldots, f_k$, then we call $\mathfrak{g}/R$ the Lie algebra with generators x_i and relations $f_j = 0$, where x_i are the images of the elements of X in $\mathfrak{g}/R$.

The next result is the penultimate ingredient we need.

{t:Serre}

Theorem 15.5. *Let Φ be a root system with base $\Delta = \{\alpha_1, \ldots, \alpha_l\}$. Let $\mathfrak{g}$ be the Lie algebra generated by $3l$ elements $\{x_i, y_i, h_i \mid 1 \leq i \leq l\}$ subject to the following relations:*

(S1) $[h_i, h_j] = 0$ *for all* $1 \leq i, j \leq l$.
(S2) $[x_i, y_i] = h_i$, $[x_i, y_j] = 0$ *if* $i \neq j$.
(S3) $[h_i, x_j] = \langle \alpha_j, \alpha_i \rangle x_j$, $[h_i, y_j] = \langle -\alpha_j, \alpha_i \rangle y_j$.
(S4) $(\operatorname{ad} x_i)^{-\langle \alpha_j, \alpha_i \rangle + 1}(x_j) = 0$ *(if* $i \neq j$*).*
(S5) $(\operatorname{ad} y_i)^{-\langle \alpha_j, \alpha_i \rangle + 1}(y_j) = 0$ *(if* $i \neq j$*).*

Then $\mathfrak{g}$ is a finite-dimensional semisimple Lie algebra with maximal toral subalgebra spanned by the h_i and root system Φ. □

The final ingredient is given by the following

{t:Isom}

Theorem 15.6. *Let $\mathfrak{g}$, $\mathfrak{g}'$ be semisimple Lie algebras over K with maximal toral subalgebras $\mathfrak{h}$ and $\mathfrak{h}'$ and root systems Φ, Φ'. Suppose there is an isomorphism $\varphi\colon \Phi \longrightarrow \Phi'$ inducing an isomorphism $\pi\colon \mathfrak{h} \longrightarrow \mathfrak{h}'$. Fix a base Δ of Φ, so $\Delta' = \{\varphi(\alpha) \mid \alpha \in \Delta\}$ is a base of Φ'. For each $\alpha \in \Delta$ and $\varphi(\alpha) =: \alpha' \in \Delta'$ choose nonzero vectors $x_\alpha \in \mathfrak{g}_\alpha$ and $x'_{\alpha'} \in \mathfrak{g}'_{\alpha'}$. These vectors define an isomorphism $\pi_\alpha\colon \mathfrak{g}_\alpha \longrightarrow \mathfrak{g}'_{\alpha'}$. Then there exists a unique isomorphism $\widetilde{\pi}\colon \mathfrak{g} \longrightarrow \mathfrak{g}'$ such that $\widetilde{\pi}_{|\mathfrak{h}} = \pi$ and $\widetilde{\pi}_{|\mathfrak{g}_\alpha} = \pi_\alpha$.* □

Proof of Theorem 1.12. Theorems 15.4, 15.5 and 15.6 show that the map from simple Lie algebras to irreducible root systems is one-to-one.

Let $\mathfrak{g}$ be a finite-dimensional simple complex Lie algebra. Its root system is irreducible and these were classified in Theorem 14.11. One first checks that any root system appearing in Theorem 14.11 can actually be constructed (we have done it for A_n-D_n). Then, one needs to show that all the classical algebras appearing in Theorem 1.12 are actually semisimple. This follows rather quickly from Proposition 8.8 for the algebras $\mathfrak{so}(2n, \mathbb{C})$ and $\mathfrak{so}(2n+1, \mathbb{C})$ (use Remark 1.14). The semisimplicity of the symplectic algebra is proved in Exercise 1 on the bonus sheet. The simplicity of $\mathfrak{sl}(n, \mathbb{C})$ was established in Exercise 4, Sheet 5.

For the classical algebras, it is straightforward to check that their root systems are precisely of type A_n, B_n, C_n or D_n. For the remaining cases one starts from a root system and constructs the associated simple Lie algebra as in Theorem 15.5. □

References

[1] W. Fulton and J. Harris, *Representation theory. A first course*, Springer-Verlag, New York, 1991.
[2] J. Humphreys, *Introduction to Lie algebras and representation theory*, Springer-Verlag, New York-Berlin, 1978.
[3] H. Samelson, *Notes on Lie algebras*, Springer-Verlag, New York, 1990.
[4] C. Schweigert, *Lie algebras, lecture notes (in German)*, available at http://www.math.uni-hamburg.de/home/schweigert/skripten.html.
[5] W. Soergel, *Lie algebras lecture notes (in German)*, available at http://home.mathematik.uni-freiburg.de/soergel/#Skripten.

Pawel Sosna, Fachbereich Mathematik der Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany
E-mail address: pawel.sosna@math.uni-hamburg.de