LIE ALGEBRAS AND QUANTUM GROUPS

Jouko Mickelsson

CHAPTER 1 GENERAL STRUCTURE OF LIE ALGEBRAS

1.1. Lie algebras and homomorphisms

Let $\mathbb{F}$ be the field of real or complex numbers. A *Lie algebra* is a vector space $\mathbf{g}$ over $\mathbb{F}$ with a *Lie product* (or *commutator*) $[\cdot,\cdot]:\mathbf{g}\times\mathbf{g}\to\mathbf{g}$ such that

(1) $x\mapsto[x,y]$ is linear for any $y\in\mathbf{g}$,

(2) $[x,y]=-[y,x]$,

(3) $[x,[y,z]]+[y,[z,x]]+[z,[x,y]]=0$.

The last condition is called the *Jacobi identity* . From (1) and (2) it follows that also $y\mapsto[x,y]$ is linear for any $x\in\mathbf{g}$. In this chapter we shall consider only finite-dimensional Lie algebras. In any vector space $\mathbf{g}$ one can always define a *trivial Lie product* $[x,y]\equiv 0$. A Lie algebra with this commutator is *Abelian.* The space $\mathbf{gl}(n,\mathbb{R})$ of all real $n\times n$ matrices is naturally a Lie algebra with respect to the matrix commutator $[X,Y]=XY-YX$, and correspondingly the complex algebra $\mathbf{gl}(n,\mathbb{C})$.

Some other nontrivial examples follow:

Example 1.1.1. Let $\mathbf{o}(n)$ denote the space of all real antisymmetric $n\times n$ matrices. The commutator of a pair of matrices is defined by

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

(ordinary matrix multiplication in xy). Since $(xy)^t=y^tx^t$, where x^t denotes the transpose of the matrix x, the commutator of two antisymmetric matrices is again antisymmetric. The commutator clearly satisfies (1) and (2); (3) is checked by a simple computation. The dimension of the real vector space $\mathbf{o}(n)$ is $\frac{1}{2}n(n-1)$.

The matrix Lie algebras, like $\mathbf{o}(n)$ above, are closely related to groups of matrices. Let $O(n)$ denote the group of all orthogonal $n\times n$ matrices A, $A^tA=1$. Then the Lie algebra $\mathbf{o}(n)$ consists precisely of those matrices x for which $A(s)=\exp sx\in O(n)$ for all $s\in\mathbb{R}$. Namely, taking the derivative of $A(s)^tA(s)$ at $s=0$ one gets x^t+x. So $A(s)\in O(n)$ implies $x\in\mathbf{o}(n)$. On the other hand if $x\in\mathbf{o}(n)$ then $(\exp sx)^t=\exp sx^t=\exp(-sx)=(\exp sx)^{-1}$, so $A(s)\in O(n)$.

Example 1.1.2. The real vector space $\mathbf{u}(n)$ consisting of *anti-Hermitian* $n\times n$ matrices $x, x^*=-x$, where $x^*=\overline{x}^t$ and the bar means complex conjugation, is a Lie algebra with respect to the matrix commutator. Its dimension is n^2. Denoting by $U(n)$ the group of *unitary* matrices A, $A^*A = 1$, one can prove as in the case of orthogonal matrices that $\exp sx \in U(n)\, \forall s \in \mathbb{R}$ iff $x \in \mathbf{u}(n)$.

Example 1.1.3. The traceless anti-Hermitian $n \times n$ matrices form a Lie algebra to be denoted by $\mathbf{su}(n)$ and it corresponds to the group $SU(n) = \{A \in U(n) \mid \det A = 1\}$. The dimension of $\mathbf{su}(n)$ is $n^2 - 1$.

Example 1.1.4. Let J be the antisymmetric $2n \times 2n$ matrix

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

Since $\det J = (-1)^{n+1} \neq 0$ the form $\langle x, y\rangle = x^t J y$ is nondegenerate (the vectors x, y are written as column matrices). Define $\mathbf{sp}(2n, \mathbb{R})$ to consist of all real $2n \times 2n$ matrices x such that $x^t J + Jx = 0$. This is a Lie algebra and one can associate to $\mathbf{sp}(2n, \mathbb{R})$ the group $Sp(2n, \mathbb{R})$ consisting of real matrices A such that $A^t J A = J$, or equivalently such that A preserves the form $\langle u, v\rangle = u^t J v$, $\langle Au, Av\rangle = \langle u, v\rangle$ for all $u, v \in \mathbb{R}^{2n}$. $Sp(2n, \mathbb{R})$ is the *symplectic group* defined by J.

Exercise 1.1.5. Find a basis for $\mathbf{sp}(2n, \mathbb{R})$ and show that dim $\mathbf{sp}(2n, \mathbb{R}) = 2n^2 + n$.

One can analogously define the complex orthogonal Lie algebra $\mathbf{o}(n, \mathbb{C})$ and the complex symplectic Lie algebra $\mathbf{sp}(2n, \mathbb{C})$.

We have also the Lie algebra $\mathbf{sl}(n, \mathbb{C})$ of complex traceless $n \times n$ matrices and correspondingly the real Lie algebra $\mathbf{sl}(n, \mathbb{R})$.

Let $\{X_1, X_2, \dots, X_n\}$ be a vector space basis of a Lie algebra $\mathbf{g}$. We define the *structure constants* c_{ij}^k by

$$[X_i, X_j] = c_{ij}^k X_k$$

(sum over the repeated index k; we shall use the same summation convention also later). From the defining properties (1) and (2) follows that the commutator $[X, Y]$ for arbitrary $X, Y \in \mathbf{g}$ is determined by the structure constants. The Jacobi identity can be written as

$$c_{ij}^l c_{lk}^m + c_{jk}^l c_{li}^m + c_{ki}^l c_{lj}^m = 0$$

$\forall i, j, k, m$. By the antisymmetry of the Lie product we have $c_{ij}^k = -c_{ji}^k$.

Example 1.1.6. Let $\mathbf{g}$ be a two dimensional Lie algebra with a basis $\{X_1, X_2\}$. If $\mathbf{g}$ is not commutative we can define a nonzero element

$$e_1 = [X_1, X_2] = \alpha X_1 + \beta X_2.$$

Choose a pair of numbers γ, δ such that $\alpha\delta - \beta\gamma = 1$ and set

$$e_2 = \gamma X_1 + \delta X_2.$$

Then $[e_1, e_2] = e_1$. Thus we have found the general structure of a noncommutative two dimensional Lie algebra.

Let $\mathbf{g}$ and $\mathbf{g}'$ be Lie algebras. A linear map $\phi : \mathbf{g} \to \mathbf{g}'$ is a *homomorphism* if

$$\phi([x, y]) = [\phi(x), \phi(y)]$$

$\forall x, y \in \mathbf{g}$. An invertible homomorphism is an *isomorphism* . The inverse of an isomorphism is also an isomorphism. An isomorphism of $\mathbf{g}$ into itself is an *automorphism* of the Lie algebra $\mathbf{g}$.

A linear subspace $\mathbf{k} \subset \mathbf{g}$ is a *subalgebra* of $\mathbf{g}$ if $[x, y] \in \mathbf{k} \, \forall x, y \in \mathbf{k}$. A subalgebra is a Lie algebra in its own right.

Exercise 1.1.7. Let $\phi : \mathbf{g} \to \mathbf{g}'$ be a homomorphism. Show that the *kernel* $\ker\phi = \{x \in \mathbf{g} \mid \phi(x) = 0\} \subset \mathbf{g}$ and the *image* $\mathrm{im}\phi = \{\phi(x) \mid x \in \mathbf{g}\} \subset \mathbf{g}'$ are subalgebras.

A subspace $\mathbf{k} \subset \mathbf{g}$ is an *ideal* if $[x, y] \in \mathbf{k} \, \forall x \in \mathbf{g}$ and $y \in \mathbf{k}$. In particular, an ideal is always a subalgebra. If $\mathbf{k} \subset \mathbf{g}$ is an ideal then the quotient space $\mathbf{g}/\mathbf{k}$ is naturally a Lie algebra: The commutator of the cosets $x + \mathbf{k}$ and $y + \mathbf{k}$ is by definition the coset $[x, y] + \mathbf{k}$. If $x' + \mathbf{k} = x + \mathbf{k}$ and $y' + \mathbf{k} = y + \mathbf{k}$ (i.e., $x' - x \in \mathbf{k}$ and $y' - y \in \mathbf{k}$) then $[x', y'] = [x + (x' - x), y + (y' - y)] \equiv [x, y] \bmod \mathbf{k}$ by the ideal property of $\mathbf{k}$; thus $[x', y']$ represents the same element in $\mathbf{g}/\mathbf{k}$ as $[x, y]$ and so the commutator is well-defined in $\mathbf{g}/\mathbf{k}$.

Proposition 1.1.8. *Let $\phi : \mathbf{g} \to \mathbf{g}'$ be a homomorphism which is onto* (i.e., $\mathbf{g}' = \mathrm{im}\phi$). Then the Lie algebras $\mathbf{g}'$ and $\mathbf{g}/\ker\phi$ are isomorphic.

Proof. Define $\psi : \mathbf{g}/\mathrm{ker}\phi \to \mathbf{g}'$ by $\psi(x+\mathrm{ker}\phi) = \phi(x)$. Obviously ϕ is one-to-one and it is a homomorphism by $\psi([x+\mathrm{ker}\phi, y+\mathrm{ker}\phi]) = \psi([x, y]+\mathrm{ker}\phi) = \phi([x, y]) = [\psi(x+\mathrm{ker}\phi), \psi(y+\mathrm{ker}\phi)]$.

A linear map $\delta : \mathcal{A} \to \mathcal{A}$ in an algebra is a *derivation* if

$$\delta(a * b) = \delta(a) * b + a * \delta(b)$$

for all $a, b \in \mathcal{A}$.

Let $\mathrm{Der}(\mathcal{A})$ be the set of all derivations of $\mathcal{A}$. Then $\mathrm{Der}(\mathcal{A})$ is a Lie subalgebra of the Lie algebra of all endomorphisms of $\mathcal{A}$.

In the special case when $\mathcal{A} = \mathbf{g}$ is a Lie algebra we can define a derivation ad_X of $\mathbf{g}$ for any $X \in \mathbf{g}$ by

$$\mathrm{ad}_X : \mathbf{g} \to \mathbf{g},\ \mathrm{ad}_X(Y) = [X, Y].$$

This defines a homomorphism ad: $\mathbf{g} \to \mathrm{Der}(\mathbf{g})$; this is called *the adjoint representation* of $\mathbf{g}$. The derivations ad_X are called *inner derivations,* the rest are *outer derivations.*

Exercise 1.1.9 Let $\mathbf{g}$ be the three dimensional Lie algebra which as a vector space is $\mathbb{R}^3$, equipped with the commutator $[X, Y] = X \wedge Y$, the vector product in $\mathbb{R}^3$. Show that $\mathbf{g}$ is a Lie algebra and that it is isomorphic with the Lie algebra $\mathbf{o}(3)$.

Exercise 1.1.10 Let $x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ be a basis of the Lie algebra $\mathbf{sl}(2, \mathbb{C})$. Determine explicitely the adjoint representation, i.e., the matrices ad_x, ad_y, ad_h.

Exercise 1.1.11 Show that the Lie algebras $\mathbf{o}(3), \mathbf{su}(2)$, and $\mathbf{sp}(2)$ (the anti-hermitean part of $\mathbf{sp}(2, \mathbb{C})$) are isomorphic. Show that $\mathbf{o}(6)$ and $\mathbf{su}(4)$ are isomorphic.

Exercise 1.1.12 Find a two dimensional Lie algebra of 2×2 matrices which is isomorphic to the noncommutative two dimensional Lie algebra discussed earlier in this section.

1.2. Ideals in Lie algebras

A left (right) ideal in an algebra $\mathcal{A}$ is a linear subspace $I \subset \mathcal{A}$ such that $x * y \in I$ ($y * x \in I$) for all $x \in \mathcal{A}$ and $y \in I$. An (two sided) ideal is both left and right ideal.

If $\mathcal{A}$ is a Lie algebra, there is no difference between left and right ideals since $x * y = [x, y] = -[y, x]$.

The *center* of a Lie algebra $\mathbf{g}$ is the subspace $Z(\mathbf{g}) = \{x \in \mathbf{g} | [x, y] = 0 \forall y \in \mathbf{g}\}$. Clearly the center is an ideal. Another ideal is the subspace $[\mathbf{g}, \mathbf{g}]$ consisting of all linear combinations of commutators in the Lie algebra.

Lemma 1.2.1. *The vector space sum of two ideals in* $\mathbf{g}$ *is again an ideal in* $\mathbf{g}$. *The commutator* $[I, J]$ *of a pair of ideals is also an ideal.*

Proof. The first claim follows directly from the definition. The second is a simple consequence of the Jacobi identity.

A Lie algebra $\mathbf{g}$ is *simple* if its only ideals are the trivial ideals 0 and $\mathbf{g}$ itself and if $\mathbf{g}$ is not the commutative one dimensional Lie algebra. If $\mathbf{g}$ is simple then $\mathbf{g} = [\mathbf{g}, \mathbf{g}]$ and $Z(\mathbf{g}) = 0$.

The basic example. Let $\mathbf{g} = \mathbf{sl}(2, \mathbb{C})$. We choose a basis as in the exercise 1.1.10. Then

$$[h, x] = 2x, \ \ [h, y] = -2y, \ \ [x, y] = h.$$

Let $I \subset \mathbf{g}$ be a nonzero ideal. We choose $0 \neq z = ax + by + ch \in I$. Then

$$[x, z] = bh - 2cx \text{ and } [x, bh - 2cx] = -2bx.$$

Thus $bx \in I$ and $[y, [y, z]] = -2ay \in I$.

1) If $a \neq 0$ then $y \in I$ and so $[x, y] = h \in I$ and $-\frac{1}{2}[x, h] = x \in I$ and so $I = \mathbf{g}$. Likewise the case $b \neq 0$.

2) If $a = b = 0$ then $c \neq 0$ and $z = ch \in I$, so $h \in I$, $y = \frac{1}{2}[y, h] \in I$. and $x = -\frac{1}{2}[x, h] \in I$. It follows that $I = \mathbf{g}$.

Thus $\mathbf{sl}(2, \mathbb{C})$ is simple. Actually, the above proof holds for $\mathbf{sl}(2, \mathbb{F})$ when $\mathbb{F}$ is an arbitrary field of characteristic not equal to 2.

Theorem 1.2.2.

(1) *Let* $\phi : \mathbf{g} \to \mathbf{g}'$ *be a Lie algebra homomorphism and* $I \subset \mathbf{g}$ *an ideal such that* $I \subset ker\phi$. *Then there exists a unique homorphism* $\psi : \mathbf{g}/I \to \mathbf{g}'$ *such that* $\phi = \psi \circ \pi$, *where* $\pi : \mathbf{g} \to \mathbf{g}/I$ *is the canonical homomorphism.*

(2) *If* $I, J \subset \mathbf{g}$ *is a pair of ideals with* $I \subset J$ *then* J/I *is an ideal in* $\mathbf{g}/I$ *and* $(\mathbf{g}/I)/(J/I) \simeq \mathbf{g}/J$.

(3) *If* $I, J \subset \mathbf{g}$ *is any pair of ideals then* $(I + J)/J \simeq I/(I \cap J)$.

Proof.

(1) Define the map $\psi : \mathbf{g}/I \to \mathbf{g}'$ by $\psi(x + I) = \phi(x)$. It is easy to see that this is a homomorphism which satisfies the requirement. If ψ' is another such a homomorphism, then $(\psi' - \psi) \circ \pi = 0$ and so $\psi' - \psi = 0$ since π is onto.

(2) The first statement follows directly from definitions. For the second, define a map $f : (\mathbf{g}/I)/(J/I) \to \mathbf{g}/J$ by $f((x+I)+J/I) = x+J$. This map is the required isomorphism.

(3) Define $f : I/(I \cap J) \to (I + J)/J$ by $f(x + I \cap J) = x + J$ and check that this is an isomorphism.

A *representation* of a Lie algebra $\mathbf{g}$ in a vector space V is a Lie algebra homomorphism $\phi : \mathbf{g} \to \mathrm{End}(V)$. As an example, any Lie algebra has the natural adjoint representation in the vector space $V = \mathbf{g}$, $\mathrm{ad}_x(y) = [x, y]$.

A representation is *irreducible* if the representation space V does not have any invariant subspaces except of course 0 and V; a subspace $W \subset V$ is invariant if $\phi(x)v \in W$ for all $x \in \mathbf{g}$ and $v \in W$.

If $\mathbf{g}$ is a simple Lie algebra then the adjoint representation is necessarily irreducible. Conversely, if $\mathbf{g}$ is noncommutative and the adjoint representation is irreducible then $\mathbf{g}$ is simple.

If $\mathbf{g}$ is simple then $Z(\mathbf{g}) = 0$ and it follows that the kernel of the adjoint representation ad: $\mathbf{g} \to \mathrm{End}(\mathbf{g})$ is zero. Thus $\mathbf{g}$ is isomorphic to a subalgebra of $\mathrm{End}(\mathbf{g})$. Choosing a basis in $\mathbf{g}$ we see that any simple Lie algebra is isomorphic to a Lie algebra of matrices.

Let $\delta \in \mathrm{Der}(\mathbf{g})$, $\mathbf{g}$ any finite-dimensional Lie algebra. Since δ is a linear operator in a finite-dimensional vector space we may form the exponential

$$e^{\delta} = 1 + \delta + \frac{1}{2!}\delta^2 + \frac{1}{3!}\delta^3 + \dots$$

to define a linear operator $\exp(\delta) : \mathbf{g} \to \mathbf{g}$.

Proposition 1.2.3. *The map* $\exp(\delta)$ *is an automorphism of* $\mathbf{g}$.

Proof. First, $\exp(\delta)$ is a linear isomorphism since it has the inverse $\exp(-\delta)$. But

$$\begin{aligned}\exp(\delta)[x,y] &= \sum_n \frac{1}{n!}\delta^n[x,y] \\ &= \sum_n \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}[\delta^k(x),\delta^{n-k}(y)] \\ &= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty}[\frac{1}{k!}\delta^k(x),\frac{1}{i!}\delta^i(y)] = [e^\delta(x),e^\delta(y)]\end{aligned}$$

and so $\exp(\delta)$ is a Lie algebra homomorphism. Here $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$ are the binomial coefficients.

The automorphisms of the type $\exp(\delta)$ when $\delta = \mathrm{ad}_x$ are called *inner automorphisms.* They generate a group (upon multiplication), to be denoted by $\mathrm{Int}(\mathbf{g})$; this is a subgroup of the group $\mathrm{Aut}(\mathbf{g})$ of all automorphims of $\mathbf{g}$.

Proposition 1.2.4. *The group Int*$(\mathbf{g})$ *is a normal subgroup of Aut*$(\mathbf{g})$.

Proof. Let $\phi \in \mathrm{Aut}(\mathbf{g})$ and $x, y \in \mathbf{g}$. Then

$$\phi \circ \mathrm{ad}_x \circ \phi^{-1}(y) = \phi([x,\phi^{-1}(y)]) = [\phi(x),y] = \mathrm{ad}_{\phi(x)}(y)$$

and thus $\phi \circ \mathrm{ad}_x \circ \phi^{-1} = \mathrm{ad}_{\phi(x)}$ which proves the statement.

Exercise 1.2.5 Let $\mathbf{g}$ be a given subalgebra of $\mathrm{End}(V)$, where V is a finite-dimensional vector space. Show that $e^{\mathrm{ad}_x}(y) = e^x y e^{-x}$ for any $x, y \in \mathbf{g}$.

Exercise 1.2.6 Let $\mathbf{g} = \mathbf{sl}(2,\mathbb{C})$ and choose a basis $\{x,h,y\}$ as in the exercise 1.1.10. Determine the matrices $e^{\mathrm{ad}_x}, e^{\mathrm{ad}_h}$, and e^{ad_y} in this basis.

Exercise 1.2.7 Let $\mathbf{g} = \mathbf{o}(n)$. Let g be any orthogonal matrix. Show that the map $x \mapsto gxg^{-1}$ defines an automorphism of $\mathbf{g}$. Is this automorphism inner?

1.3 Solvable and nilpotent Lie algebras

Let $\mathbf{g}$ be a Lie algebra and define $\mathbf{g}^0 = \mathbf{g}$ and for any $k = 0, 1, 2, \ldots$ $\mathbf{g}^{k+1} = [\mathbf{g}^k, \mathbf{g}^k]$. Then $\mathbf{g}^{k+1}$ is an ideal in $\mathbf{g}^k$. The Lie algebra $\mathbf{g}$ is *solvable* if $\mathbf{g}^k = 0$ for some integer k.

Of course any commutative Lie algebra is solvable. A basic nontrivial example is:

Example 1.3.2 Let $\mathbf{g} = \mathbf{t}(n, \mathbb{F})$ be the space of upper triangular $n \times n$ matrices A over the field $\mathbb{F}$, $A_{ij} = 0$ for $i > j$. In this case $\mathbf{g}^1$ is contained in the set of upper triangular matrices with $A_{ii} = 0$ and in general $\mathbf{g}^k$ is contained in the space of matrices A with $A_{ij} = 0$ for $i > j - 2^{k-1}$. It follows that $\mathbf{t}(n, \mathbb{F})$ is solvable.

Theorem 1.3.3.

(1) *Any subalgebra of a solvable Lie algebra is solvable. The image of a solvable Lie algebra in a homomorphism is solvable.*
(2) *If $\mathbf{k}$ is an ideal in $\mathbf{g}$ and if both $\mathbf{k}$ and $\mathbf{g}/\mathbf{k}$ are solvable then $\mathbf{g}$ is solvable.*
(3) *A sum of two solvable ideals in a Lie algebra is also solvable.*

Proof. i) Clearly $\mathbf{k}^k \subset \mathbf{g}^k$ when $\mathbf{k} \subset \mathbf{g}$ is a subalgebra. This implies implies the solvability of $\mathbf{k}$. If $\phi : \mathbf{g} \to \mathbf{g}'$ is a homomorphism then $[\phi(\mathbf{g}), \phi(\mathbf{g})] = \phi([\mathbf{g}, \mathbf{g}])$ and in general $\phi(\mathbf{g}^k) = \phi(\mathbf{g})^k$ from which the sovability of $\phi(\mathbf{g})$ follows.

ii) For some m, n we have $\mathbf{k}^m = 0$ and $(\mathbf{g}/\mathbf{k})^n = 0$. If $\pi : \mathbf{g} \to \mathbf{g}/\mathbf{k}$ is the canonical homomorphism then $\pi(\mathbf{g}^n) = (\pi(\mathbf{g}))^n = (\mathbf{g}/\mathbf{k})^n = 0$. This implies $\mathbf{g}^n \subset \mathbf{k}$ and so $\mathbf{g}^{mn} \subset \mathbf{k}^m = 0$.

iii) Let $\mathbf{k}, \mathbf{k}'$ be a pair of ideals in $\mathbf{g}$. According to 1.2.2 we have $(\mathbf{k} + \mathbf{k}')/\mathbf{k} \simeq \mathbf{k}'/(\mathbf{k} \cap \mathbf{k}')$. The canonical projection $\pi : \mathbf{k} \to \mathbf{k}/(\mathbf{k} \cap \mathbf{k}')$ is a homomorphism and thus the image is solvable. By ii) we see that $\mathbf{k} + \mathbf{k}'$ is solvable.

Let $\mathbf{g}$ be a finite-dimensional Lie algebra. Then by 1.3.3. iii) the sum of all its solvable ideals is solvable. It follows that it has a unique *maximal solvable ideal*. This ideal is called *the radical* of $\mathbf{g}$ and denoted by $\operatorname{rad}\mathbf{g}$. A Lie algebra $\mathbf{g}$ is *semisimple* if $\mathbf{g} \neq 0$ and $\operatorname{rad}\mathbf{g} = 0$. Any simple Lie algebra is semisimple since the only ideals in a simple Lie algebra $\mathbf{g}$ are 0 and $\mathbf{g}$ and $\mathbf{g}$ is not solvable.

Assume that $0 \neq \mathbf{g} \neq \operatorname{rad}\mathbf{g}$. Then $\mathbf{g}/\operatorname{rad}\mathbf{g}$ is semisimple: In the opposite case there would be a nonzero solvable ideal $\mathbf{k} = \mathbf{t}/\operatorname{rad}\mathbf{g}$ in $\mathbf{g}/\operatorname{rad}\mathbf{g}$, where $\mathbf{t} \subset \mathbf{g}$ is some ideal. But by 1.3.3 ii) $\mathbf{t}$ is solvable, which implies that there is a larger solvable ideal in $\mathbf{g}$ than $\mathrm{rad}\,\mathbf{g}$, a contradiction.

For any Lie algebra $\mathbf{g}$ we set $\mathbf{g}_0 = \mathbf{g}$ and $\mathbf{g}_{k+1} = [\mathbf{g}, \mathbf{g}_k]$ for $k = 0, 1, 2, \dots$. We get a descending set of ideals in $\mathbf{g}$. The Lie algebra $\mathbf{g}$ is *nilpotent* if $\mathbf{g}_n = 0$ for some n. Since $\mathbf{g}^k \subset \mathbf{g}_k$, any nilpotent Lie algebra is solvable. The basic example:

Example 1.3.4 Let $\mathbf{g} = \mathbf{n}(n, \mathbb{F})$ be the Lie algebra of upper triangular matrices A such that $A_{ij} = 0$ for all $i \geq j$. Then $\mathbf{g}_k$ consists of matrices A for which $A_{ij} = 0$ for $i \geq j - k$ and thus $\mathbf{g}$ is nilpotent.

Theorem 1.3.5.

(1) *Any subalgebra of a nilpotent Lie algebra is nilpotent. The image of a nilpotent Lie algebra is nilpotent.*
(2) *Let $Z(\mathbf{g})$ be the center of a Lie algebra $\mathbf{g}$. If $\mathbf{g}/Z(\mathbf{g})$ is nilpotent then $\mathbf{g}$ is nilpotent.*
(3) *The center of a nonzero nilpotent Lie algebra is nonzero,*

Proof. i) As in the proof of 1.3.3. i)

ii) Let $(\mathbf{g}/Z(\mathbf{g}))_n = 0$. Then $\mathbf{g}_n \subset Z(\mathbf{g})$ and therefore $\mathbf{g}_{n+1} = 0$.

iii) Let $\mathbf{g}_{n+1} = 0$ but $\mathbf{g}_n \neq 0$. Then $\mathbf{g}_n \subset Z(\mathbf{g})$ and thus $Z(\mathbf{g}) \neq 0$.

An element $x \in \mathbf{g}$ is called *ad-nilpotent* if $(ad_x)^n = 0$ for some n. Since for any $y \in \mathbf{g}$ we have $0 = (ad_x)^n(y) \in \mathbf{g}_{n+1}$ we observe that in a nilpotent Lie algebra any element is ad-nilpotent.

Theorem 1.3.6. *(Engel) If all elements in $\mathbf{g}$ are ad-nilpotent then $\mathbf{g}$ is nilpotent.*

To prove the theorem we need some preparations.

Lemma 1.3.7. *Let $x \in \mathbf{gl}(n, \mathbb{F})$ be nilpotent, $x^k = 0$ for some k. Then x is ad-nilpotent.*

Proof. We write $ad_x = \rho_x - \lambda_x$, where $\rho_x(y) = xy$ and $\lambda_x(y) = -yx$. From $x^k = 0$ follows $\rho_x^k = \lambda_x^k = 0$. But

$$(ad_x)^m = \sum_{i=0}^{m} \binom{m}{i} \rho_x^i (-\lambda_x)^{m-i}$$

which is equal to zero for $m \geq 2k$.

Theorem 1.3.8. *Let* $\mathbf{g}$ *be a Lie subalgebra of* $\mathbf{gl}(n,\mathbb{F})$ *for some* $n = 1, 2, 3, \ldots$. *If all elements of* $\mathbf{g}$ *are nilpotent as matrices then there exists* $0 \neq v \in \mathbb{F}^n$ *such that* $xv = 0$ *for all* $x \in \mathbf{g}$.

Proof. The statement is clearly true when $\dim \mathbf{g} = 1$. We perform an induction on $\dim \mathbf{g}$. Thus we assume that the statement holds for $\dim \mathbf{g} < n$ and prove it for the case $\dim \mathbf{g} = n$. Let $0 \neq \mathbf{k} \neq \mathbf{g}$ be a subalgebra of $\mathbf{g}$. Now $\mathbf{g}/\mathbf{k}$ is a vector space of dimension $m < n$ and we have a homomorphism $\phi : \mathbf{k} \to \mathbf{gl}(m,\mathbb{F})$ (after selecting a basis in $\mathbf{g}/\mathbf{k}$) by $\phi(x)(y + \mathbf{k}) = [x, y] + \mathbf{k}$. By Lemma 1.3.7 each $\phi(x)$ is nilpotent, as a linear transformation of $\mathbf{g}/\mathbf{k}$. By the induction assumption there is a nonzero vector $y + \mathbf{k}$ in $\mathbf{g}/\mathbf{k}$ such that $\phi(x)(y + \mathbf{k}) = 0$ for all $x \in \mathbf{k}$. This means that $[x, y] \in \mathbf{k}$ for all $x \in \mathbf{k}$. We define the *normalizer* of a subalgebra by

$$N(\mathbf{g},\mathbf{k}) = \{y \in \mathbf{g} | [x, y] \in \mathbf{k} \forall x \in \mathbf{k}\}.$$

We see that the vector y above belongs to $N(\mathbf{g},\mathbf{k})$. In particular, $N(\mathbf{g},\mathbf{k})$ is strictly larger than $\mathbf{k}$.

Let now $\mathbf{k} \subset \mathbf{g}$ be a maximal subalgebra; this means that if $\mathbf{k}'$ is a subalgebra of $\mathbf{g}$ containing $\mathbf{k}$ then either $\mathbf{k}' = \mathbf{k}$ or $\mathbf{k}' = \mathbf{g}$. It is easy to see that maximal subalgebras exist. In this case $\mathbf{k} \neq N(\mathbf{g},\mathbf{k})$ and so $N(\mathbf{g},\mathbf{k}) = \mathbf{g}$. From this follows that $\mathbf{k}$ is an ideal in $\mathbf{g}$.

Now $\dim \mathbf{g}/\mathbf{k} = 1$; otherwise, there would be a one dimensional subalgebra $\mathbf{s} \subset \mathbf{g}/\mathbf{k}$ which implies that there is a subalgebra $\mathbf{k}'$ such that $\mathbf{k}' \neq \mathbf{k}$ and $\mathbf{k}' \neq \mathbf{g}$. This is in contradiction with the maximality of $\mathbf{k}$.

Thus indeed $\dim \mathbf{g} = \dim \mathbf{k} + 1$. Choose $z \neq 0$ in the complement of $\mathbf{k}$ in $\mathbf{g}$. By the induction assumption,

$$W = \{v \in V | \mathbf{k}v = 0\} \neq 0.$$

Since $\mathbf{k}$ is an ideal of $\mathbf{g}$, $[x, z] \in \mathbf{k}$ for $x \in \mathbf{k}$ and thus $x(zw) = 0$ for $w \in W, x \in \mathbf{k}$ and so W is a z-invariant subspace. Since z is a nilpotent transformation in W there is an element $0 \neq v \in W$ such that $zw = 0$ which implies $\mathbf{g}v = 0$.

Proof of theorem 1.3.6. Let $\mathbf{g} \neq 0$ with each $x \in \mathbf{g}$ ad-nilpotent. We apply 1.3.8 to the algebra $\operatorname{ad} \mathbf{g} \subset \mathbf{gl}(\mathbf{g})$. There exists a vector $0 \neq x \in \mathbf{g}$ such that $[y, x] = 0$ for all $y \in \mathbf{g}$. Thus $Z(\mathbf{g}) \neq 0$.

We use induction in $\dim \mathbf{g} = n$. For $n = 1$ the claim is clearly true. Assume then that the claim is true for $\dim \mathbf{g} < n$. By the induction assumption $\mathbf{g}/Z(\mathbf{g})$ is nilpotent. By 1.3.5. ii), the Lie algebra $\mathbf{g}$ is nilpotent.

Theorem 1.3.9. *Let $\mathbf{g} \subset \mathbf{gl}(V)$ a subalgebra consisting of nilpotent endomorphisms, $V \neq 0$, $\dim V < \infty$. There exists a flag of subspaces, $0 = V_0 \subset V_1 \subset V_2 \subset \dots V_n = V$, such that $xV_i \subset V_{i-1}$ for each $x \in \mathbf{g}$. In other words, we can choose a basis of V such that in this basis the transformations x are upper triangular with zeros on the diagonal.*

Proof. Choose $v_1 \in V$ such that $\mathbf{g}v_1 = 0$. Set $V_1 = \mathbb{F}v_1$. Set $W_1 = V/V_1$. Then we have a homomorphism $\phi : \mathbf{g} \to \mathbf{gl}(W_1)$ by $\phi(x)(v + V_1) = xv + V_1$. All endomorphisms $\phi(x)$ are nilpotent and therefore we may choose $0 \neq v_2 + V_1 \in W_1$ such that $\phi(\mathbf{g})(v_2 + V_1) = 0$, so $\mathbf{g}v_2 \subset V_1$. Next we set $V_2 = V_1 + \mathbb{F}v_2, W_2 = V/V_1$, and continue as in the first step. The process stops at some point since V is finite-dimensional.

Corollary 1.3.10. *Let $\mathbf{g}$ be a nilpotent Lie algebra and $\mathbf{k} \subset \mathbf{g}$ a nonzero ideal. Then $\mathbf{k} \cap Z(\mathbf{g}) \neq 0$.*

Proof. Set $\phi(x)(y) = ad_x(y)$ for $x \in \mathbf{g}$ and $y \in \mathbf{k}$. By 1.3.8 there is a vector $0 \neq y \in \mathbf{k}$ such that $\phi(\mathbf{g})y = 0$, i.e., $[x, y] = 0$ for all $x \in \mathbf{g}$. Thus $y \in Z(\mathbf{g})$.

Exercise 1.3.11 Let $\operatorname{char} \mathbb{F} = 2$. Show that $\mathbf{sl}(2, \mathbb{F})$ is nilpotent.

Exercise 1.3.12 Let $\mathbf{k}, \mathbf{k}'$ be a pair of nilpotent ideals in a Lie algebra $\mathbf{g}$. Show that $\mathbf{k} + \mathbf{k}'$ is nilpotent. From this follows that the Lie algebra has a unique maximal nilpotent ideal, the so called *nilradical.* Determine the nilradical of a Lie algebra defined by the relations $[x, y] = z, [x, z] = y, [y, z] = 0$.

Exercise 1.3.13 Let $\mathbf{g}$ be a nonzero nilpotent Lie algebra. Show that it has an ideal of codimension $= 1$.

Exercise 1.3.14 Show that a Lie algebra $\mathbf{g}$ is solvable if and only if there is a sequence of ideals $\mathbf{g}_{(k)} \subset \mathbf{g}_{(k-1)}$ such that $\mathbf{g}_{(0)} = \mathbf{g}$ and $\mathbf{g}_{(n)} = 0$ for some n, and such that $\mathbf{g}_{(k-1)}/\mathbf{g}_{(k)}$ is commutative for each k.

Exercise 1.3.15 Let $\mathbf{k} \neq \mathbf{g}$ be a subalgebra of a nilpotent Lie algebra $\mathbf{g}$. Show that $\mathbf{k} \subset N(\mathbf{g}, \mathbf{k})$ is a proper subalgebra.

CHAPTER 2 SEMISIMPLE LIE ALGEBRAS

2.1 Lie's and Cartan's theorems

In this section $\mathbb{F}$ is an algebraically closed field of char$=0$ (typically, $\mathbb{F}=\mathbb{C}$.)

Theorem 2.1.1. *Let* $\mathbf{g}$ *be a solvable subalgebra of* $\mathbf{gl}(V)$*, where* V *is a finite-dimensional vector space over* $\mathbb{F}$*,* $V \neq 0$*. Then there exists* $0 \neq v \in V$ *and a linear map* $\lambda : \mathbf{g} \to \mathbb{F}$ *such that* $xv = \lambda(x)v$ *for all* $x \in \mathbf{g}$*.*

Proof. The case $\dim \mathbf{g} = 1$ is clear because any matrix over $\mathbb{F}$ has an eigenvector. We use induction $\dim \mathbf{g}$. So let $\dim \mathbf{g} = n > 1$ and assume that the claim is true for dimension less than n.

First we observe that there exists an ideal $\mathbf{k} \subset \mathbf{g}$ of codimension one. Since $\mathbf{g}$ is solvable, $[\mathbf{g},\mathbf{g}] \neq \mathbf{g}$ and we may choose a subspace $\mathbf{k} \subset \mathbf{g}$ of codimension one, containing $[\mathbf{g},\mathbf{g}]$. This subspace is an ideal since $[\mathbf{g},\mathbf{k}] \subset [\mathbf{g},\mathbf{g}] \subset \mathbf{k}$.

From the induction hypothesis follows that there is a vector $0 \neq v \in V$ and a linear map $\lambda : \mathbf{k} \to \mathbb{F}$ such that $xv = \lambda(x)v$ for all $x \in \mathbf{k}$. Let

$$W = \{w \in V | xw = \lambda(x)w \forall x \in \mathbf{k}\}.$$

We know already that $W \neq 0$.

Next we prove that $\mathbf{g}W \subset W$. Let $x \in \mathbf{g}$, $w \in W$ and $y \in \mathbf{k}$. Then

$$yxw = xyw - [x,y]w = \lambda(y)xw - \lambda([x,y])w.$$

We want to prove that $\lambda([x,y]) = 0$ for all $x \in \mathbf{g}, y \in \mathbf{k}$. Let n be the smallest integer for which $w, xw, x^2w, \ldots, x^nw$ are linearly dependent. Let W_i be the subspace spanned by the vectors $w, xw, \ldots, x^{i-1}w$. Then $\dim W_i = i$ for $i = 0, 1, \ldots, n$. Furthermore, W_n is invariant under the transformation x.

Subinduction. We prove by induction on i that $yx^iw \equiv \lambda(y)x^iw \bmod W_i$, for $y \in \mathbf{k}$. The case $i = 0$ is clear, so assume that the claim is true for integers less or equal to i. Now

$$yx^{i+1}w = yxx^iw = xyx^iw - [x,y]x^iw = x(\lambda(y)x^iw + w') - \lambda([x,y])x^iw - w'',$$

for some $w', w'' \in W_i$. Since $xW_i \subset W_{i+1}$ and $W_i \in W_{i+1}$ we get

$$yx^{i+1}w \equiv \lambda(y)x^{i+1}w \bmod W_{i+1}.$$

End of subinduction.

Thus the linear map $y : W_n \to W_n$ is represented in the basis $w, xw, \ldots, x^{n-1}w$ as a matrix with diagonal entries equal to $\lambda(y)$. Thus $\mathrm{tr}_{W_n}(y) = n\lambda(y)$. In particular, $0 = \mathrm{tr}_{W_n}([x, y]) = n\lambda([x, y])$ and thus $\lambda([x, y]) = 0$ for $y \in \mathbf{k}$ from which follows $\mathbf{g}W \subset W$.

We can write $\mathbf{g} = \mathbf{k} + \mathbb{F}z$ for some $0 \neq z \in \mathbf{g}$. Since $zW \subset W$ there is an eigenvector $0 \neq v_0$ of z in W, $zv_0 = \lambda(z)v_0$. Thus we may extend the map $\lambda : \mathbf{k} \to \mathbb{F}$ to a linear map $\lambda : \mathbf{g} \to \mathbb{F}$ such that $xv_0 = \lambda(x)v_0$ for all $x \in \mathbf{g}$.

Corollary 2.1.2. *(Lie's theorem) Let* $\mathbf{g}$ *be a solvable subalgebra of* $\mathbf{gl}(V)$. *Then we may choose a basis in* V *such that all elements of* $\mathbf{g}$ *are presented as upper triangular matrices.*

Proof. The claim is clearly true when the dimension n of V is $n = 1$. We use induction on n. So we assume that the claim is true when the dimension is less than n. By the previous theorem there is a nonzero vector $0 \neq v_1 \in V$ such that $xv_1 = \lambda(x)v_1$ for some linear functional λ on $\mathbf{g}$. Then we pass to the quotient space $V_1 = V/\mathbb{F}v_1$ and use the induction hypothesis to see that there is a basis $\{v_i + \mathbb{F}v_1\}$, with $i = 2, 3, \ldots n$ such that the $\mathbf{g}$ action is upper triangular in this basis. Then $\{v_i\}_{i=1}^n$ is a basis of V with the required property.

Let $\mathbf{g}$ be a Lie algebra and $\phi : \mathbf{g} \to \mathbf{gl}(V)$ a representation of $\mathbf{g}$ in a vector space. We set

$$V_\lambda = \{v \in V | (\phi(x) - \lambda(x))^n v = 0 \text{ for } x \in \mathbf{g} \text{ and some } n = n_x\},$$

where λ is a linear functional on $\mathbf{g}$.

The linear subspaces $V_\lambda \subset V$ are called the *weight subspaces* of ϕ, corresponding to the *weights* λ. The vectors $0 \neq v \in V_\lambda$ are called weight vectors.

Example 2.1.3 Let $\mathbf{g}$ be the Lie algebra with basis $\{a, b, h\}$ and commutation relations

$$[a, b] = -a, \ [h, a] = [h, b] = 0.$$

Let ϕ be the 2-dimensional representation of $\mathbf{g}$ defined by

$$\phi(a) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \phi(b) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \phi(h) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then the representation has two weights, $\lambda(h) = 1, \lambda(a) = \lambda(b) = 0$, and $\mu(h) = 1, \mu(a) = 0, \mu(b) = 1$. The weight vectors are the unit vectors $v_\lambda = e_2, v_\mu = e_1$.

According to the Jordan decomposition theorem in matrix algebra, in any finite-dimensional vector space V and for any $T \in \operatorname{End}(V)$ there is a decomposition

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_p} \text{ with } V_{\lambda_k} = \{v \in V | (T - \lambda_k)^n v \text{ for some } n\}.$$

This result generalizes to nilpotent Lie algebras.

Theorem 2.1.4. *Let $\phi : \mathbf{g} \to \mathbf{gl}(V)$ be a representation of a nilpotent Lie algebra in a finite-dimensional vector space. Then*

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_p},$$

where $\lambda_i : \mathbf{g} \to \mathbb{F}$ are the weights of the representation ϕ and V_{λ_i} are the corresponding weight spacces. Furthermore, $\phi(\mathbf{g})V_{\lambda_i} \subset V_{\lambda_i}$ for all i.

Proof. Induction on $n = \dim \mathbf{g}$. The case $n = 1$ is clear by the matrix algebra theorem mentioned above. So we asume that the induction hypothesis is true for dimension less than n. When $\dim \mathbf{g} = n$, we observe:

1) Since $\mathbf{g}$ is nilpotent, $\mathbf{g}_1 = [\mathbf{g}, \mathbf{g}] \neq \mathbf{g}$. Let $\mathbf{k} \subset \mathbf{g}$ be a subspace of codimension one, containing $\mathbf{g}_1$. Then $\mathbf{k}$ is a nilpotent subalgebra and we may write

$$V = V_{\beta_1} \oplus \cdots \oplus V_{\beta_r}$$

for some weights of $\mathbf{k}$ in V.

2) Let $0 \neq x$ be a vector in the complement of $\mathbf{k}$ in $\mathbf{g}$. Then $\mathbf{g} = \mathbf{k} \oplus \mathbb{F}x$. We shall show that $\phi(x)V_{\beta_i} \subset V_{\beta_i}$ for all i. Since $\mathbf{g}$ is nilpotent there exists an integer n_0 such that $(ad_y)^n x = 0$ for all $n \geq n_0$ and $y \in \mathbf{g}$. Let $v \in V_{\beta_i}$. Choose m_0 such that $(\phi(y) - \beta_i(y))^{m_0} v = 0$ for all $y \in \mathbf{k}$. By Lemma 2.1.5 below,

$$(\phi(y) - \beta_i(y))^{n_0+m_0} \phi(x) v = \sum_{j=0}^{n_0+m_0} \binom{n_0 + m_0}{j} \phi((ad_y)^j x)(\phi(y) - \beta_i(y))^{n_0+m_0-j} v = 0$$

for all $y \in \mathbf{k}$. It follows that $\phi(x)v \in V_{\beta_i}$ for $v \in V_{\beta_i}$.

3) By (2) we can write

$$V_{\beta_i} = V_{i,1} \oplus \cdots \oplus V_{i,n_i}$$

where

$$V_{i,j} = \{v \in V_{\beta_i} | (\phi(x) - \alpha_{i,j})^n v = 0 \text{ for some } n\}.$$

Repeating the argument in (2) with Lemma 2.1.5, we get $\phi(y)V_{i,j} \subset V_{i,j}$ for all $y \in \mathbf{k}$. Since $\mathbf{g} = \mathbf{k} + \mathbb{F}x$ we have $\phi(\mathbf{g})V_{i,j} \subset V_{i,j}$. Setting

$$\lambda_{i,j}(y + ax) = \beta_i(y) + a\alpha_{i,j}$$

for $y \in \mathbf{k}$ we observe that each $V_{i,j}$ is a weight subspace of the representation ϕ, corresponding to the weight $\lambda_{i,j}$.

Lemma 2.1.5. *Let ϕ be a representation of a Lie algebra $\mathbf{g}$ in a vector space V. Let $x, y \in \mathbf{g}, v \in V$, and $\alpha, \beta \in \mathbb{F}$. Then*

$$(\phi(y) - \alpha - \beta)^n \phi(x)v = \sum_{i=0}^{n} \binom{n}{i} \phi((ad_y - \beta)^i x)(\phi(y) - \alpha)^{n-i} v$$

for any $n \in \mathbb{N}$.

Proof. The case $n = 0$ is clear. We use induction on n, so we assume that the formula holds for exponents less or equl to n and we prove it for $n + 1$. Denote $x_i = (\mathrm{ad}_y - \beta)^i x$.

$$\begin{aligned}(\phi(y) - \alpha - \beta)^{n+1}\phi(x)v &= (\phi(y) - \alpha - \beta) \sum \binom{n}{i} \phi(x_i)(\phi(y) - \alpha)^{n-i} v \\ &= \sum \binom{n}{i} (\phi(y) - \alpha - \beta)\phi(x_i)(\phi(y) - \alpha)^{n-i} v.\end{aligned}$$

Now

$$\begin{aligned}(\phi(y) - \alpha - \beta)\phi(x_i) &= \phi(x_i)\phi(y) + [\phi(y), \phi(x_i)] - (\alpha + \beta)\phi(x_i) \\ &= \phi(x_i)(\phi(y) - \alpha) + \phi((\mathrm{ad}_y - \beta)x_i) = \phi(x_i)(\phi(y) - \alpha) + \phi(x_{i+1})\end{aligned}$$

so that

$$\begin{aligned}&\sum \binom{n}{i} (\phi(y) - \alpha - \beta)\phi(x_i)(\phi(y) - \alpha)^{n-i} v \\ &= \sum \binom{n}{i} \phi(x_i)(\phi(y) - \alpha)^{n-i+1} v + \sum \binom{n}{i} \phi(x_{i+1})(\phi(y) - \alpha)^{n-i} v \\ &= \sum_{i=0}^{n+1} \binom{n+1}{i} \phi(x_i)(\phi(y) - \alpha)^{n-i+1} v\end{aligned}$$

because of $\binom{n}{i} + \binom{n}{i-1} = \binom{n+1}{i}$.

The Killing form on a finite-dimensional Lie algebra $\mathbf{g}$ is the symmetric bilinear form defined as

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

In the next section we shall prove that a Lie algebra is semisimple if and only if its Killing form is nondegenerate.

Exercise 2.1.6 Let $\{x, y, h\}$ be the standard basis of $\mathbf{sl}(2, \mathbb{F})$. Compute the determinant of the Killing form in the standard basis. Compute the dual basis to this basis. (Two basis e_i, f_i are dual to each other if $(e_i, f_j) = \delta_{ij}$.)

Exercise 2.1.7 Let $\mathbf{g} = \mathbf{sl}(n, \mathbb{F}), char\, \mathbb{F} = 0$. Using Lie's theorem show that $\mathrm{rad}\, \mathbf{g} = Z(\mathbf{g})$ and that $\mathbf{g}$ is semisimple.

Exercise 2.1.8 We assume here that the field $\mathbb{F}$ has characteristics $p \neq 0$. Consider the the two dimensional Lie subalgebra in $\mathbf{gl}(p, \mathbb{F})$ spanned by the matrices x, y with $y = diag(0, 1, 2, \ldots, p-1)$ and

$$x = \begin{pmatrix} 0 & 1 & 0 \ldots 0 \\ 0 & 0 & 1 \ldots 0 \\ \ldots & & \\ 0 & 0 & 0 \ldots 1 \\ 1 & 0 & 0 \ldots 0 \end{pmatrix}.$$

Show that Lie's theorem fails, the matrices $x, , y$ do not have any common nonzero eigenvector.

2.2 Cartan subalgebras

A nilpotent subalgebra $\mathbf{h} \subset \mathbf{g}$ is called a *Cartan subalgebra* if the normalizer $N(\mathbf{g}, \mathbf{h})$ of $\mathbf{h}$ in $\mathbf{g}$ is equal to $\mathbf{h}$.

Example 2.2.1 Let $\mathbf{g} = \mathbf{sl}(n, \mathbb{F})$ and $\mathbf{h}$ the subalgebra of diagonal matrices in $\mathbf{g}$. Then $\mathbf{h}$ is a Cartan subalgebra of $\mathbf{g}$. Of course, since here $\mathbf{h}$ is commutative, it is nilpotent. The only thing to check is that if $[x, y]$ is diagonal for every diagonal matrix y then also x is diagonal; this is an easy exercise in matrix algebra.

Let $\mathbf{h} \subset \mathbf{gl}(V)$ be any subalgebra. We define

$$V_0(\mathbf{h}) = \{v \in V | x^n v = 0 \text{ for all } x \in \mathbf{h}, \text{ for some } n \in \mathbb{N}\}.$$

For a single element x we put $V_0(x) = V_0(\mathbb{F}x)$. By repeated use of the Jacobi identity,

$$(ad_x)^n[y, z] = \sum \binom{n}{i} [(ad_x)^i y, (ad_x)^{n-i} z].$$

From this follows that $\mathbf{g}_0(ad_x) \subset \mathbf{g}$ is a subalgebra.

The minumum of the dimension of $\mathbf{g}_0(ad_x)$ (when x goes through all elements in $\mathbf{g}$) is called *the rank* of the Lie algebra $\mathbf{g}$. If $x \in \mathbf{g}$ such that $\dim \mathbf{g}_0(ad_x) = \operatorname{rank} \mathbf{g}$ then x is a *regular element* in $\mathbf{g}$.

Lemma 2.2.2. *Let $\mathbf{k} \subset \mathbf{g}$ be a subalgebra and $z \in \mathbf{k}$ such that $dim\, \mathbf{g}_0(ad_z) = min_{x \in \mathbf{k}}\ dim\, \mathbf{g}_0(ad_x)$. If $\mathbf{k} \subset \mathbf{g}_0(ad_z)$ then $\mathbf{g}_0(ad_z) \subset \mathbf{g}_0(ad_x)$ for all $x \in \mathbf{k}$.*

Proof. Let $x \in \mathbf{k}$. Since $\mathbf{k} \subset \mathbf{g}_0(ad_z)$, we have a linear map

$$ad_{z+cx} : \mathbf{g}_0(ad_z) \to \mathbf{g}_0(ad_z)$$

for all $c \in \mathbb{F}$. We have then also the induced linear map

$$ad_{z+cx} : \mathbf{g}/\mathbf{g}_0(ad_z) \to \mathbf{g}/\mathbf{g}_0(ad_z).$$

It is a standard result in linear algebra that the characteristic polynomial $f_{\mathbf{g}}$ of the linear map ad_{z+cx} in $\mathbf{g}$ factorizes (think about determinants of block upper triangular matrices!) as $f_{\mathbf{g}} = f_{\mathbf{g}_0} \cdot f_{\mathbf{g}/\mathbf{g}_0}$ to the characteristic polynomials in $\mathbf{g}_0 = \mathbf{g}_0(ad_z)$ and in $\mathbf{g}/\mathbf{g}_0$. We can write

$$f_{\mathbf{g}_0}(a) = a^r + p_1(c)a^{r-1} + \dots p_r(c)$$

and

$$f_{\mathbf{g}/\mathbf{g}_0}(a) = a^{n-r} + q_1(c)a^{n-r-1} + \ldots q_{n-r}(c),$$

where $n = \dim \mathbf{g}, r = \dim\, \mathbf{g}_0$, and each p_i is a polynomial at most of degree r in the parameter c and each q_i is a polynomial of at most degree $n-r$.

If $\mathbf{g}_0(ad_z) = \mathbf{g}$ there is nothing to prove, so let us assume that $\mathbf{g}_0(ad_z)$ is a proper subalgebra. All eigenvectors of $ad_z = ad_{z+0\cdot x}$ corresponding to the eigenvalue =0 belong to the subspace $\mathbf{g}_0(ad_z)$ so that $\lambda = 0$ is not an eigenvalue of ad_z in $\mathbf{g}/\mathbf{g}_0$. It follows that $q_{n-r}(0) \neq 0$. It follows that we may choose parameter values $c_1, c_2, \ldots, c_{r+1}$ such that $q_{n-r}(c_i) \neq 0$ and $c_i \neq c_j$ for $i \neq j$. Then $ad_{z+c_i x}$ does not have eigenvalue 0 in the quotient $\mathbf{g}/\mathbf{g}_0$ which implies that

$$\mathbf{g}_0(ad_{z+c_i x}) \subset \mathbf{g}_0(ad_z).$$

By the assumption, $\mathbf{g}_0(ad_{z+c_i x}) = \mathbf{g}_0(ad_z)$. Thus the linear map $ad_{z+c_i x} : \mathbf{g}_0(ad_z) \to \mathbf{g}_0(ad_z)$ has zero as its only eigenvalue so that

$$f_{\mathbf{g}_0}(a) = a^r \text{ for each parameter value } c = c_1, \ldots c_{r+1}.$$

It follows that $p_j(c_i) = 0$ for each i. Since p_j is at most of degree r we must have $p_j \equiv 0$. Thus

$$\mathbf{g}_0(ad_z) \subset \mathbf{g}_0(ad_{z+cx})$$

for all c. In particular, setting $c = 1$ and replacing x by $x - z$ we have completed the proof.

Lemma 2.2.3. *Let* $\mathbf{k}$ *be a subalgebra in* $\mathbf{g}$ *and assume that* $\mathbf{g}_0(ad_x) = \mathbf{k}$ *for some* $x \in \mathbf{k}$. *Then* $N(\mathbf{g}, \mathbf{k}) = \mathbf{k}$.

Proof. If $z \in N(\mathbf{g}, \mathbf{k})$ then $[z, x] \in \mathbf{k}$ since $x \in \mathbf{k}$. But then $ad_x^n[z, x] = 0$ for some n and so $ad_x^{n+1}(z) = 0$ and $z \in \mathbf{g}_0(ad_x) = \mathbf{k}$. □

Remark If we choose $\mathbf{k} = \mathbf{g}_0(ad_x)$ for some x then we have $N(\mathbf{g}, \mathbf{g}_0(ad_x)) = \mathbf{g}_0(ad_x)$.

Theorem 2.2.4. *Let* $\mathbf{h}$ *be a subalgebra in a Lie algebra* $\mathbf{g}$. *Then* $\mathbf{h}$ *is a Cartan subalgebra if and only if there is a regular element* $x \in \mathbf{g}$ *such that* $\mathbf{h} = \mathbf{g}_0(ad_x)$.

Proof. 1) Let $x \in \mathbf{g}$ be regular and set $\mathbf{h} = \mathbf{g}_0(ad_x)$. By Lemma 2.2.3, $\mathbf{h} = N(\mathbf{g}, \mathbf{h})$. Since $x \in \mathbf{h}$, by Lemma 2.2.2 we have $\mathbf{h} = \mathbf{g}_0(ad_x) \subset \mathbf{g}_0(ad_y)$ for all $y \in \mathbf{h}$. Thus

$$(ad_y)^n z = 0 \text{ for some } n, \forall z \in \mathbf{h}.$$

This means that each $y \in \mathbf{h}$ is ad-nilpotent. Theorem 1.3.6 implies that $\mathbf{h}$ is nilpotent, so $\mathbf{h}$ is a Cartan subalgebra in $\mathbf{g}$.

2) Let $\mathbf{h}$ be a Cartan subalgebra of $\mathbf{g}$. Since $\mathbf{h}$ is nilpotent, we have $\mathbf{h} \subset \mathbf{g}_0(ad_x)$ for any $x \in \mathbf{h}$. Choose $z \in \mathbf{h}$ such that the dimension of $\mathbf{g}_0(ad_z)$ is the minimum of the dimensions of $\mathbf{g}_0(ad_x)$ for $x \in \mathbf{h}$. From Lemma 2.2.2 follows that $\mathbf{g}_0(ad_z) \subset \mathbf{g}_0(ad_x)$ for all $x \in \mathbf{h}$. *We claim that* $\mathbf{h} = \mathbf{g}_0(ad_z)$. If this is not the case, we have a representation

$$\phi : \mathbf{h} \to \mathbf{gl}(\mathbf{g}_0(ad_z)/\mathbf{h}), \;\; \phi(x)(y + \mathbf{h}) = [x, y] + \mathbf{h}$$

in a nonzero vector space. From $\mathbf{g}_0(ad_z) \subset \mathbf{g}_0(ad_x)$ (for all $x \in \mathbf{h}$) follows that each $\phi(x)$ is nilpotent as a linear transformation. From 1.3.8 follows that there exists a nonzero vector $y + \mathbf{h}$ such that

$$[x, y] \in \mathbf{h} \text{ for all } x \in \mathbf{h}.$$

This implies $y \in N(\mathbf{g}, \mathbf{h})$ and so $N(\mathbf{g}, \mathbf{h})$ is strictly larger than $\mathbf{h}$, a contradiction.

Let $\mathbf{h}$ be a Cartan subalgebra of $\mathbf{g}$. The weights of the representation ad of $\mathbf{h}$ in $\mathbf{g}$ are called the *roots* of the pair $(\mathbf{g}, \mathbf{h})$. By 2.1.4 we can write

$$\mathbf{g} = \mathbf{g}_0 \oplus_{\gamma \neq 0} \mathbf{g}_\gamma,$$

where $\mathbf{g}_\gamma$ is the *root subspace* corresponding to the root γ. By 2.2.4 the subspace $\mathbf{g}_0$ corresponding to the zero root is equal to $\mathbf{h}$.

Lemma 2.2.5. *Let* $\mathbf{h} \subset \mathbf{g}$ *be a Cartan subalgebra and* γ, γ' *a pair of roots. Then* $[\mathbf{g}_\gamma, \mathbf{g}_{\gamma'}] \subset \mathbf{g}_{\gamma+\gamma'}$. *In particular, if* $\gamma + \gamma'$ *is not a root then* $[\mathbf{g}_\gamma, \mathbf{g}_{\gamma'}] = 0$.

Proof. Let $x \in \mathbf{g}_\gamma, y \in \mathbf{g}_{\gamma'}, h \in \mathbf{h}$. Then

$$(ad_h - (\gamma + \gamma')(h))^n [x, y] = \sum \binom{n}{i} [(ad_h - \gamma(h))^i x, (ad_h - \gamma'(h))^{n-i} y] = 0$$

when n is large enough, Lemma 2.1.5. Thus $[x, y] \in \mathbf{g}_{\gamma+\gamma'}$. □

Lemma 2.2.6. *Let* $\mathbf{h} \subset \mathbf{g}$ *be a Cartan subalgebra and* $\phi : \mathbf{g} \to \mathbf{gl}(V)$ *a representation of* $\mathbf{g}$ *in* V. *Let* γ *be a root of* $(\mathbf{g}, \mathbf{h})$ *and* α *a weight of the restriction of* ϕ *to the subalgebra* $\mathbf{h}$. *Then* $\phi(x)v \in V_{\alpha+\gamma}$ *for all* $v \in V_\alpha, x \in \mathbf{g}_\gamma$. *In particular,* $\phi(x)v = 0$ *if* $\alpha + \gamma$ *is not a weight.*

Proof. Use Lemma 2.1.5.

Lemma 2.2.7. *Let* $\mathbf{g}$ *be a Lie algebra such that* $[\mathbf{g}, \mathbf{g}] = \mathbf{g}$. *Let* $\mathbf{h} \subset \mathbf{g}$ *be a Cartan subalgebra and* ϕ *a representation of* $\mathbf{g}$ *in a finite-dimensional vector space* V. *Assume that* $tr((\phi(x))^2) = 0$ *for all* $x \in \mathbf{h}$. *Then each* $\phi(x)$ *is nilpotent,* $x \in \mathbf{h}$.

Proof. We have $\mathbf{g} = \mathbf{h} \oplus_{\gamma \neq 0} \mathbf{g}_\gamma$. Now

$$\mathbf{g} = [\mathbf{g}, \mathbf{g}] = \sum_{\gamma,\gamma'} [\mathbf{g}_\gamma, \mathbf{g}_{\gamma'}] \subset \sum_{\gamma,\gamma'} \mathbf{g}_{\gamma+\gamma'}$$

so that $\mathbf{h} = \mathbf{g}_0 = \sum_\gamma [\mathbf{g}_\gamma, \mathbf{g}_{-\gamma}]$. Let α be any root and η a weight of the representation $\phi|_{\mathbf{h}}$. Set

$$V' = \oplus_{k \in \mathbb{Z}} V_{\eta+k\alpha}.$$

Since $\mathbf{g}_\alpha V_\gamma \subset V_{\gamma+\alpha}$, we observe that the subspace V' is invariant under the linear transformations $\phi(e_{\pm\alpha})$, where $0 \neq e_{\pm\alpha} \in \mathbf{g}_{\pm\alpha}$. We set $h = [e_\alpha, e_{-\alpha}] \in \mathbf{h}$ and $\psi(x) = \phi(x)|_{V'}$. Then

$$\mathrm{tr}\,(\psi(h)) = \mathrm{tr}\,(\psi([e_\alpha, e_{-\alpha}]) = \mathrm{tr}\,[\psi(e_\alpha), \psi(e_{-\alpha})] = 0.$$

When p is large enough,

$$(\phi(h) - \eta(h) - k\alpha(h))^p V_{\eta+k\alpha} = 0$$

and so the restriction of $\phi(h) - \eta(h) - k\alpha(h)$ to the subspace $V_{\eta+k\alpha}$ is nilpotent. The trace of a nilpotent matrix is zero, so that the trace of the restriction of $\phi(h)$ to $V_{\eta+k\alpha}$ is equal to $(\eta(h) + k\alpha(h)) \cdot \dim V_{\eta+k\alpha}$. It follows that

$$0 = \mathrm{tr}\,(\psi(h)) = \mathrm{tr}\,(\phi(h)|_{V'}) = \sum_k \mathrm{tr}\,(\phi(h)_{V_{\eta+k\alpha}}) = \sum_k (\eta(h) + k\alpha(h)) \dim V_{\eta+k\alpha}$$

so that

$$\eta(h) = -\alpha(h) \cdot \frac{\sum k \dim V_{\eta+k\alpha}}{\sum \dim V_{\eta+k\alpha}}$$

when $h \in [\mathbf{g}_\alpha, \mathbf{g}_{-\alpha}]$. Note that by the assumption $\mathbf{g} = [\mathbf{g}, \mathbf{g}]$, any $h \in \mathbf{h}$ is a linear combination of elements of this type for different α's. We have now

$$\eta(h) = r(\eta, \alpha)\alpha(h) \text{ for } h \in [\mathbf{g}_\alpha, \mathbf{g}_{-\alpha}]$$

where r is a rational number. By $x^2 - y^2 = (x+y)(x-y)$, also the operators $(\phi(h))^2 - (\eta(h))^2$ are nilpotent in the subspace V_η. It follows that the trace vanishes in this subspace, so

$$0 = \mathrm{tr}_{V_\eta} (\phi(h))^2 - \mathrm{tr}_{V_\eta} (\eta(h))^2.$$

Thus

$$0 = \mathrm{tr}\,(\phi(h)^2) = \sum_\eta \mathrm{tr}\,(\phi(h)^2)|_{V_\eta} = \sum_\eta (\eta(h))^2 \mathrm{dim}\, V_\eta$$

and so $\eta(h) = 0$ for all $h \in \mathbf{h}$. As a consequence $V = V_0$ and $(\phi(h) - \eta(h))^p = (\phi(h))^p = 0$ for p large enough.

Corollary 2.2.8. *Let $\mathbf{h} \subset \mathbf{g}$ be a Cartan subalgebra and ϕ a representation of $\mathbf{g}$ in a finite-dimensional vector space V. If $[\mathbf{g}, \mathbf{g}] = \mathbf{g}$ and $h \in [\mathbf{g}_\alpha, \mathbf{g}_{-\alpha}]$ then*

$$\eta(h) = r(\eta, \alpha) \cdot \alpha(h)$$

for some rational number r and for any weight η. Furthermore,

$$r(\eta, \alpha) = -\frac{\sum k\, dim\, V_{\eta + k\alpha}}{\sum dim\, V_{\eta + k\alpha}}.$$

Theorem 2.2.9. *(Cartan's criterium) A finite-dimensional nonzero Lie algebra is semisimple if and only if its Killing form is nondegenerate.*

Proof. 1) Assume that the Killing form of $\mathbf{g}$ is *degenerate,* Let

$$\mathbf{s} = \{x \in \mathbf{g} | (x, y) = 0 \ \forall y \in \mathbf{g}\} \neq 0.$$

Let $x \in \mathbf{s}$ and $y, z \in \mathbf{g}$. Then

$$([x, y], z) = \mathrm{tr}\,(\mathrm{ad}_{[x,y]} \cdot \mathrm{ad}_z) = \mathrm{tr}\,([\mathrm{ad}_x, \mathrm{ad}_y], \mathrm{ad}_z) = \mathrm{tr}\,(\mathrm{ad}_x \cdot [\mathrm{ad}_y, \mathrm{ad}_z]) = (x, [y, z]).$$

Therefore $[x, y] \in \mathbf{s}$ so that $\mathbf{s}$ is an ideal. *We claim that* $\mathbf{s}$ *is solvable.* If this is not the case, there is an integer k such that $\mathbf{s}^{k+1} = [\mathbf{s}^k, \mathbf{s}^k] \neq 0$. Let $\mathbf{k} = \mathbf{s}^k$ and $\mathbf{h}$

a Cartan subalgebra of $\mathbf{k}$. The linear map $x \mapsto \mathrm{ad}_x$ is a representation of $\mathbf{k}$ in $\mathbf{g}$. Since

$$\mathrm{tr}\,((\mathrm{ad}_x)^2) = (x, x) = 0 \text{ for } x \in \mathbf{h} \subset \mathbf{s},$$

it follows from Lemma 2.2.7 that ad_x is nilpotent for all $x \in \mathbf{h}$, so that $\mathbf{h} = \mathbf{k}$ (Theorem 2.2.4). In particular $\mathbf{k}$ is nilpotent, a contradiction. Thus $\mathbf{g}$ has a solvable ideal and it is not semisimple.

2) Assume now that $\mathbf{g}$ is not semisimple. Then it has a solvable nonzero ideal $\mathbf{s}$. Let n be the smallest integer for which $\mathbf{s}^{n+1} = 0$. If $x \in \mathbf{s}^n$ then $\mathrm{ad}_x(\mathbf{g}) \subset \mathbf{s}^n$ and $\mathrm{ad}_x(\mathbf{s}^n) = 0$. Thus $(\mathrm{ad}_x \cdot \mathrm{ad}_y)^2 = 0$ for all $y \in \mathbf{g}$ and so it has zero trace,

$$0 = \mathrm{tr}\,(\mathrm{ad}_x \cdot \mathrm{ad}_y) = (x, y)$$

and the Killing form is degenerate.

We repeat the simple but important observation in the proof above:

Corollary 2.2.10. *$(ad_x(y), z) = -(y, ad_x(z)$, so that the matrices ad_x are anti-symmetric with respect to the Killing form.*

Exercise 2.2.11 Let $\mathbf{h}$ be the set of diagonal matrices in $\mathbf{g} = \mathbf{sl}(n, \mathbb{C})$. Determine the roots of $(\mathbf{g}, \mathbf{h})$.

Exercise 2.2.12 Let $\mathbf{g}$ be nilpotent. Show that the Killing form of $\mathbf{g}$ vanishes identically.

Exercise 2.2.13 Compute the Killing form for the two dimensional Lie algebra in Example 1.1.6.

Exercise 2.2.14 Let $\mathbf{g}$ be a Lie algebra over a field $\mathbb{F}$ of characteristics $p \neq 0$. Show that $\mathbf{g}$ is semisimple if its Killing form is nondegenerate.

2.3 The system of roots

In this section $\mathbf{g}$ is a semisimple Lie algebra over an algebraically closed field $\mathbb{F}$ of characteristic zero and $\mathbf{h} \subset \mathbf{g}$ is a Cartan subalgebra.

We denote by Φ the set of (nonzero) roots of $(\mathbf{g}, \mathbf{h})$. For $\alpha \in \Phi$ we denote by $\mathbf{g}_\alpha$ the corresponding root subspace.

Lemma 2.3.1. *If α, β is a pair of roots such that $\alpha+\beta \neq 0$ then $(x,y)=0$ for $x \in \mathbf{g}_\alpha$ and $y \in \mathbf{g}_\beta$.*

Proof. By Lemma 2.2.6 $(\mathrm{ad}_x \cdot \mathrm{ad}_y)^n \mathbf{g}_\gamma \subset \mathbf{g}_{\gamma+n(\alpha+\beta)}$ for $n \in \mathbb{N}$. Since $\dim \mathbf{g} < \infty$ and $\alpha+\beta \neq 0$ we must have $(\mathrm{ad}_x \cdot \mathrm{ad}_y)^n \mathbf{g}_\gamma = 0$ for large n. The Lie algebra $\mathbf{g}$ is a sum of root subspaces, so $(\mathrm{ad}_x \cdot \mathrm{ad}_y)^n = 0$ for large n. The trace of a nilpotent matrix vanishes, which implies $(x,y)=0$.

Corollary 2.3.2. *If $\alpha \in \Phi$ then also $-\alpha \in \Phi$ and we can choose $e_{\pm\alpha} \in \mathbf{g}_{\pm\alpha}$ such that $(e_\alpha, e_{-\alpha}) = 1$.*

Proof. Choose any $0 \neq e_\alpha \in \mathbf{g}_\alpha$. Since the Killing form is nondegenerate, there exists $x \in \mathbf{g}$ such that $(e_\alpha, x) \neq 0$. Now all the root subspaces $\mathbf{g}_\beta$ with $\beta \neq -\alpha$ are orthogonal to e_α and $\mathbf{g}$ is the sum of root subspaces. Thus we can choose x to be an element of $\mathbf{g}_{-\alpha}$. After a normalization, we obtain the required element $e_{-\alpha} \in \mathbf{g}_{-\alpha}$.

Corollary 2.3.3. *The restriction of the Killing form to the Cartan subalgebra $\mathbf{h}$ is nondegenerate.*

Proof. Let $0 \neq h \in \mathbf{h}$. Choose any $x \in \mathbf{g}$ such that $(h,x) \neq 0$. Let x_0 be the projection of x to $\mathbf{g}_0 = \mathbf{h}$. Then $0 \neq (h,x) = (h,x_0)$ and so $(\cdot,\cdot)|_{\mathbf{h}}$ is nondegenerate.

Lemma 2.3.4. *Let ϕ be a representation of the nilpotent Lie algebra $\mathbf{h}$ in a finite-dimensional vector space V. Let $V_\alpha \subset V$ be a weight subspace. Then for any $x, x' \in \mathbf{h}$ the restriction of the linear map $\phi(x)\phi(x') - \alpha(x)\alpha(x')$ to the subspace $V_\alpha \subset V$ is nilpotent.*

Proof. Since the weight subspaces are $\mathbf{h}$-invariant, we may assume for simplicity that $V = V_\alpha$. Set $\psi(x) = \phi(x) - \alpha(x)$. Now ψ is a representation of $\mathbf{h}$ in V : By Theorem 2.1.1 there is a nonzero vector $v \in V$ such that $\phi(x)v = \alpha(x)v$ for all $x \in \mathbf{h}$. Then

$$\phi([x,y])v = [\phi(x), \phi(y)]v = 0$$

and so $\alpha([x,y]) = 0$ for all $x, y \in \mathbf{h}$. This implies

$$[\psi(x), \psi(y)] = [\phi(x), \phi(y)] = \phi([x,y]) = \psi([x,y])$$

and so ψ is indeed a representation of $\mathbf{h}$. By the definition of weight subspaces, the matrix $\psi(x)$ is nilpotent for each $x \in \mathbf{h}$. From Theorem 1.3.9 follows that in some basis all the matrices $\psi(x)$ are upper triangular, $\psi(x)_{ij} = 0$ for $i \geq j$. Then also the matrix

$$\psi(x)\psi(x') + \alpha(x)\psi(x') + \alpha(x')\psi(x) = \phi(x)\phi(x') - \alpha(x)\alpha(x')$$

is upper triangular and thus nilpotent.

Theorem 2.3.5. *If $x \in \mathbf{h}$ and $\alpha(x) = 0$ for all $\alpha \in \Phi$ then $x = 0$.*

Proof. If $x, x' \in \mathbf{h}$ then by the previous Lemma the restriction of $\mathrm{ad}_x \cdot \mathrm{ad}_{x'} - \alpha(x)\alpha(x')$ to the subspace $\mathbf{g}_\alpha \subset \mathbf{g}$ is nilpotent for $\alpha \in \Phi$. So its trace vanishes and

$$\mathrm{tr}_{\mathbf{g}_\alpha}(\mathrm{ad}_x \cdot \mathrm{ad}_{x'}) = \alpha(x)\alpha(x')\dim \mathbf{g}_\alpha.$$

Thus we obtain

$$(x, x') = \mathrm{tr}\,(\mathrm{ad}_x \cdot \mathrm{ad}_{x'}) = \sum_{\alpha \in \Phi} \alpha(x)\alpha(x')\dim \mathbf{g}_\alpha.$$

If now $\alpha(x) = 0$ for all α then $(x, x') = 0$ for all $x' \in \mathbf{h}$ and by 2.3.3 we get $x = 0$.

Theorem 2.3.6. *A Cartan subalgebra of a semisimple Lie algebra is commutative.*

Proof. From the proof of 2.3.4 we observe that $\alpha([x, y]) = 0$ for any $\alpha \in \Phi$ and $x, y \in \mathbf{h}$. From 2.3.5 follows then that $[x, y] = 0$. □

We denote by $\mathbf{h}^*$ the dual vector space of $\mathbf{h}$, i.e., the space of linear functionals $\lambda : \mathbf{h} \to \mathbb{F}$. Let $\{x_i\}_{i=1}^{\ell}$ be a basis of $\mathbf{h}$. We denote $\lambda_i = \lambda(x_i)$. Consider the following system of linear equations:

$$(\sum_i a_i x_i, x_j) = \sum_i a_i (x_i, x_j) = \lambda_j$$

for $j = 1, 2 \ldots, \ell$. Here λ_j's are given numbers and the a_i's the variables to be determined. Since the Killing form is nondegenerate in the subspace $\mathbf{h} \subset \mathbf{g}$ the determinant of the matrix (x_i, x_j) is nonzero. It follows that the linear system has a unique solution $a = (a_1, \ldots, a_\ell)$. Thus for any $\lambda = (\lambda_1, \ldots, \lambda_\ell) \in \mathbf{h}^*$ there is a unique $h_\lambda = \sum_i a_i x_i \in \mathbf{h}$ such that

$$\lambda(y) = (h_\lambda, y) \text{ for all } y \in \mathbf{h}.$$

This map gives a linear isomorphism $\mathbf{h}^* \to \mathbf{h}$, $\lambda \mapsto h_\lambda$.

Theorem 2.3.7. *Let α be a nonzero root of $(\mathbf{g}, \mathbf{h})$. Then dim $\mathbf{g}_\alpha = 1$ and $\mathbf{g}_{k\alpha} = 0$ for $k = 2, 3, \dots$.*

Proof. By Theorem 2.1.1 and Corollary 2.3.2 there is a common nonzero eigenvector $e_{-\alpha} \in \mathbf{g}_{-\alpha}$ for all linear maps ad_h for $h \in \mathbf{h}$, with $[h, x] = -\alpha(h)x$ for $x \in \mathbf{g}_{-\alpha}$. Also by Corollary 2.3.2 we can choose $e_\alpha \in \mathbf{g}_\alpha$ such that $(e_\alpha, e_{-\alpha}) = 1$. Define the subspace $V \subset \mathbf{g}$ by

$$V = \mathbb{F}e_{-\alpha} \oplus \mathbf{h} \oplus_{k=1,2,\dots} \mathbf{g}_{k\alpha}.$$

Set $h = [e_\alpha, e_{-\alpha}]$. Then for the restrictions to the subspace V we have

$$\mathrm{tr}_V\,(\mathrm{ad}_h) = \mathrm{tr}_V[\mathrm{ad}_{e_\alpha}, \mathrm{ad}_{e_{-\alpha}}] = 0.$$

Since $\mathrm{ad}_h - k\alpha(h)$ is nilpotent in the subspace $\mathbf{g}_{k\alpha}$, $\mathrm{tr}_{V_{k\alpha}}\,(\mathrm{ad}_h) = k\alpha(h)\mathrm{dim}\,\mathbf{g}_{k\alpha}$. Thus

$$\mathrm{tr}_V\,(\mathrm{ad}_h) = -\alpha(h) + \sum_k \alpha(h)\mathrm{dim}\,\mathbf{g}_{k\alpha} = \alpha(h)(-1 + \sum_k \mathrm{dim}\,\mathbf{g}_{k\alpha}).$$

If the theorem does not hold the expression in the brackets on the right would be positive so that $\alpha(h) = 0$. From Corollary 2.2.8 follows that $\beta(h) = 0$ for all roots β so that $h = 0$. But then

$$0 = (x, [e_\alpha, e_{-\alpha}] = ([e_{-\alpha}, x], e_\alpha) = \alpha(x)(e_{-\alpha}, e_\alpha) = \alpha(x)$$

for all $x \in \mathbf{h}$ and so $\alpha = 0$, a contradiction.

Corollary 2.3.8. *If $h \in \mathbf{h}$ and $\alpha \in \Phi$ then $[h, x] = \alpha(h)x$ for all $x \in \mathbf{g}_\alpha$.*

Proof. We know that $\dim \mathbf{g}_\alpha = 1$ and $\mathrm{ad}_h - \alpha(h)$ is nilpotent in this subspace, so $\mathrm{ad}_h - \alpha(h)$ is zero in $\mathbf{g}_\alpha$.

Corollary 2.3.9. *Let α be a nonzero root and $e_{\pm\alpha} \in \mathbf{g}_{\pm\alpha}$ such that $(e_\alpha, e_{-\alpha}) = 1$. Let $h = [e_\alpha, e_{-\alpha}]$. Then $h = h_\alpha$, that is, $(h, x) = \alpha(x)$ for all $x \in \mathbf{h}$. Furthermore, $h_{-\alpha} = -h_\alpha$ and $h_{\alpha+\beta} = h_\alpha + h_\beta$.*

Proof.

$$\alpha(x)(e_\alpha, e_{-\alpha}) = ([x, e_\alpha], e_{-\alpha}) = (x, [e_\alpha, e_{-\alpha}]) = (x, h)$$

so that $(h, x) = \alpha(x)$ for all $x \in \mathbf{h}$. From this equation follows at once that $h_{-\alpha} = -h_\alpha$ and $h_{\alpha+\beta} = h_\alpha + h_\beta$.

Corollary 2.3.10. *The vectors h_α for $\alpha \in \Phi$ span the vector space $\mathbf{h}$.*

Proof. Let $V \subset \mathbf{h}$ be the subspace spanned by all the h_α's. If $V \neq \mathbf{h}$ then there exists $h \in \mathbf{h}$ such that $(h, x) = 0$ for all $x \in V$. This means that $(h_\alpha, h) = 0$ for all $\alpha \in \Phi$ so that $\alpha(h) = 0$ for all α and therefore $h = 0$. □

We have earlier constructed an vector space isomorphism $\mathbf{h}^* \simeq \mathbf{h}$, $\lambda \mapsto h_\lambda$. From 2.3.10 follows that *the roots of* $(\mathbf{g}, \mathbf{h})$ *span the space* $\mathbf{h}^*$.

Theorem 2.3.11. *Let $\alpha, \beta \in \Phi$ be a pair of nonzero roots and $0 \neq e_\alpha \in \mathbf{g}_\alpha, 0 \neq e_\beta \in \mathbf{g}_\beta$. If $\alpha + \beta$ is a root then $0 \neq [e_\alpha, e_\beta] \in \mathbf{g}_{\alpha+\beta}$.*

Proof. By Lemma 2.2.6, $[e_\alpha, e_\beta] \in \mathbf{g}_{\alpha+\beta}$. For each root γ choose $e_\gamma \in \mathbf{g}_\gamma$ such that $(e_\gamma, e_{-\gamma}) = 1$. Then $h_\gamma = [e_\gamma, e_{-\gamma}]$ (Corollary 2.3.8). Set $P = \{k \in \mathbb{Z} | \alpha + k\beta \in \Phi\}$. Let k_+ be the largest number in P and k_- the smallest.

We claim that P is the interval $[k_-, k_+]$ of integers. In the opposite case there would be a smallest integer $k' \notin P$ with $k_- < k' < k_+$. We set

$$V = \bigoplus_{k_- \leq k < k'} \mathbf{g}_{\alpha+k\beta} \subset \mathbf{g}.$$

Then $\mathrm{ad}_{e_{\pm\beta}} V \subset V$ and

$$\mathrm{tr}_V \, \mathrm{ad}_{h_\beta} = \mathrm{tr}_V \, [\mathrm{ad}_{e_\beta}, \mathrm{ad}_{e_{-\beta}}] = 0.$$

On the other hand, by 2.3.7 and 2.3.8,

$$0 = \mathrm{tr}_V \, \mathrm{ad}_{h_\beta} = \sum_{k_- \leq k < k'} (\alpha(h_\beta) + k\beta(h_\beta))$$

which implies $\alpha(h_\beta)/\beta(h_\beta) = -\frac{1}{2}(k' + k_- - 1)$.

Note that $\beta(h_\beta)$ does not vanish by Lemma 2.3.13 below. Since $k' \notin P$ and $k' < k_+$ there exists a nonempty interval $[k'', k_+] \subset P$ with $k' < k''$. We choose k'' as small as possible. In the same way as above,

$$-\frac{\alpha(h_\beta)}{\beta(h_\beta)} = \frac{1}{2}(k'' + k_+) \neq \frac{1}{2}(k' + k_- - 1),$$

a contradiction. Thus $P = [k_-, k_+]$. *We claim that* $[e_\beta, \mathbf{g}_{\alpha+k\beta}] \neq 0$ and $[e_{-\beta}, \mathbf{g}_{\alpha+k\beta}] \neq 0$ for all $k_- < k < k_+$. In the opposite case there would be an $\mathrm{ad}_{e_{\pm\beta}}$ invariant subspace $V' = \bigoplus_{k_1 \leq k \leq k_2} \mathbf{g}_{\alpha+k\beta}$ where either $k_1 = k_-, k_2 < k_+$ or $k_1 > k_-, k_2 = k_+$. Again, as above we could reduce that $-\alpha(h_\beta)/\beta(h_\beta) = \frac{1}{2}(k_1 + k_2)$, a contradiction since $k_1 + k_2 \neq k_- + k_+$. □

Corollary 2.3.12. *Let α, β be a pair of roots. Then the set of those integers k for which $\alpha + k\beta \in \Phi$ is an interval $[k_-, k_+]$. Furthermore, $-\alpha(h_\beta)/\beta(h_\beta) = \frac{1}{2}(k_- + k_+)$.*

Lemma 2.3.13. *$\beta(h_\beta) \neq 0$ for each nonzero root β.*

Proof. Let $e_{\pm\beta} \in \mathbf{g}_{\pm\beta}$ such that $(e_\beta, e_{-\beta}) = 1$. If now $\beta(h_\beta) = 0$ then

$$[e_\beta, e_{-\beta}] = h_\beta \text{ and } [h_\beta, e_{\pm\beta}] = \pm\beta(h_\beta)e_{\pm\beta} = 0.$$

Then the subalgebra $\mathbf{s}$ spanned by $e_{\pm\beta}, h_\beta$ would be solvable and so also $\mathrm{ad}(\mathbf{s}) \subset \mathbf{gl}(\mathbf{g})$ is solvable. By the Corollary 2.1.2 we can choose a basis in $\mathbf{g}$ such that each ad_x for $x \in \mathbf{s}$ is represented by an upper triangular matrix. On the other hand, ad_{h_β} is diagonalizable (Cor. 2.3.8) so that $\mathrm{ad}_{h_\beta} = 0$ and $h_\beta \in Z(\mathbf{g})$. This implies that $h_\beta = 0$, a contradiction.

Corollary 2.3.14. *If α, β is a pair of nonzero roots then also $\alpha - <\alpha, \beta> \beta$ is a root, where $<\alpha, \beta> = 2(\alpha, \beta)/(\beta, \beta)$.*

Proof. Let $k_\pm$ be as in Cor. 2.3.12. Since $\alpha + 0 \cdot \beta \in \Phi$, we must have $k_- \leq 0 \leq k_+$. Then $- <\alpha, \beta> = k_- + k_+ \in [k_-, k_+]$ and we are done.

Exercise 2.3.15 Prove Schur's Lemma: If ϕ is an irreducible representation of a Lie algebra $\mathbf{g}$ in a finite-dimensional vector space V then the only matrices in V which commute with all $\phi(x)$ ($x \in \mathbf{g}$) are scalar multiples of the unit matrix. ($\mathbb{F}$ is algebraically closed.)

Exercise 2.3.16 Using Schurs lemma show that in any simple Lie algebra $\mathbf{g}$ any nondegenerate symmetric bilinear form κ which satisfies $\kappa([x, y], z) = -\kappa(y, [x, z])$ for all x, y, x is proportional to the Killing form.

Exercise 2.3.17 We prove later that $\mathbf{sl}(n, \mathbb{F})$ is simple. Show that the Killing form in this case can be written as

$$(x, y) = 2n\mathrm{tr}(xy)$$

with the ordinary trace in the algebra of $n \times n$ matrices in $\mathbf{sl}(n, \mathbb{F})$.

Exercise 2.3.18 Let $\mathbf{g} = D_{2\ell}$ be the Lie algebra of complex antisymmetric $2\ell \times 2\ell$ matrices. Let $\mathbf{h}$ be the subalgebra spanned by the matrices $h_i = e_{2i-1,2i} - e_{2i,2i-1}$ for $i = 1, 2, \ldots \ell$. Compute the roots and root subspaces for $(\mathbf{g}, \mathbf{h})$ and reduce from the results that $\mathbf{h}$ is a Cartan subalgebra of $\mathbf{g}$.

Since $\lambda \to h_\lambda$ is an isomorphism $\mathbf{h}^* \to \mathbf{h}$ we can define a nondegenerate bilinear form

$$(\lambda, \mu) = (h_\lambda, h_\mu) \text{ with } \lambda, \mu \in \mathbf{h}^*$$

in the dual vector space $\mathbf{h}^*$.

According to what we have defined before,

$$(\lambda, \mu) = \lambda(h_\mu) = \mu(h_\lambda).$$

As we have seen, the roots span the vector space $\mathbf{h}^*$. Thus we may select a set of roots $\alpha_1, \dots \alpha_\ell$ such that they give a basis in $\mathbf{h}^*$. Then any root β can be written uniquely as

$$\beta = \sum_{i=1}^{\ell} c_i \alpha_i \text{ with } c_i \in \mathbb{F}.$$

We claim that the coefficients c_i are *rational numbers.* Now

$$(\beta, \alpha_j) = \sum c_i(\alpha_i, \alpha_j)$$

and so $<\beta, \alpha_j>= \sum < \alpha_i, a_j > c_i$. This gives ℓ linear equations to determine ℓ values c_i. By Corollary 2.3.12 the coefficients in the linear system are integers and it follows that the solution is rational.

Set $E_{\mathbb{Q}}$ to be the linear span of the roots α_i with rational coefficients. Next

$$(\alpha, \alpha) = (h_\alpha, h_\alpha) = \operatorname{tr}(ad_{h_\alpha} \cdot ad_{h_\alpha}) = \sum_{\beta \in \Phi} (\beta(h_\alpha))^2 = \sum_\beta r(\beta, \alpha)^2 (\alpha(h_\alpha))^2 = r \cdot (\alpha, \alpha)^2,$$

where r is positive rational as a sum of squares of rational numbers; we have used the Corollary 2.2.8. It follows that $(\alpha, \alpha) = r^{-1}$ is a positive rational number. This implies then that $(\alpha, \beta) = \frac{1}{2} < \alpha, \beta > (\beta, \beta)$ is rational for all roots β. In case of an arbitrary rational linear combination λ of the roots α_i we can again write $(\lambda, \lambda) = \sum_\beta (\beta(h_\lambda))^2$ and since $h_\lambda = \sum_i c_i h_{\alpha_i}$ we see that also (λ, λ) is a sum of squares of rational numbers. Thus the bilinear form is *an inner product* in $E_{\mathbb{Q}}$.

Finally we define the extension $E = E_{\mathbb{Q}} \otimes_{\mathbb{Q}} \mathbb{R}$, a vector space over real numbers.

We gather some of the most important results above to a theorem:

Theorem 2.3.19. *Let* $\mathbf{h}$ *be a Cartan subalgebra of a semisimple Lie algebra* $\mathbf{g}$, Φ *the set of nonzero roots and* E *the real subspace of* $\mathbf{h}^*$ *spanned by the roots. Then*

1. *If* $\alpha \in \Phi$ *then* $-\alpha \in \Phi$ *but* $k\alpha \notin \Phi$ *for* $k \neq \pm 1$
2. *If* $\alpha, \beta \in \Phi$ *then* $\beta - < \beta, \alpha > \alpha \in \Phi$ *where* $< \alpha, \beta >= 2(\beta, \alpha)/(\alpha, \alpha)$
3. *If* $\alpha, \beta \in \Phi$ *then* $< \alpha, \beta > \in \mathbb{Z}$.

Theorem 2.3.20. *A Lie algebra* $\mathbf{g}$ *is semisimple if and only if it has simple ideals* $\mathbf{g}_i$ *such that* $\mathbf{g} = \mathbf{g}_1 \oplus \cdots \oplus \mathbf{g}_n$.

Proof. 1) Let $\mathbf{g}$ be semisimple. If $\mathbf{g}$ is simple, there is nothing to prove. Let us then assume that $\mathbf{g}$ has a nonzero ideal $\mathbf{g}' \neq \mathbf{g}$. Let

$$\mathbf{g}'' = \{x \in \mathbf{g} | (x, y) = 0 \ \forall\, y \in \mathbf{g}'\}.$$

Since the Killing form $(\cdot, \cdot)$ is nondegenarate, the dimension of $\mathbf{g}''$ is equal to $\dim \mathbf{g} - \dim \mathbf{g}'$. For $x \in \mathbf{g}'', y \in \mathbf{g}, z \in \mathbf{g}'$ we have

$$([y, x], z) = -(x, [y, z]) = 0$$

which implies that $\mathbf{g}''$ is an ideal. The intersection $\mathbf{g}' \cap \mathbf{g}''$ is by the proof of Theorem 2.2.9 solvable. But since $\mathbf{g}$ is semisimple, it has no nontrivial solvable ideals and so $\mathbf{g}' \cap \mathbf{g}'' = 0$. It follows that $\mathbf{g} = \mathbf{g}' \oplus \mathbf{g}''$.

We claim that $\mathbf{g}', \mathbf{g}''$ are semisimple. Otherwise, there would be a solvable nonzero ideal, say $\mathbf{s} \subset \mathbf{g}'$. But $[\mathbf{g}, \mathbf{s}] = [\mathbf{g}' \oplus \mathbf{g}'', \mathbf{s}] = [\mathbf{g}', \mathbf{s}] \subset \mathbf{s}$ and so $\mathbf{s}$ would be a solvable nonzero ideal in $\mathbf{g}$, a contradiction.

We can continue this process and split both $\mathbf{g}', \mathbf{g}''$ to semisimple ideals; the process stops at some point since the algebra is finite-dimensional.

2) Assume that $\mathbf{g} = \mathbf{g}_1 \oplus \cdots \oplus \mathbf{g}_n$ is a sum of simple ideals. If $x = \sum x_i$ and $y = \sum y_i$ are arbitrary elements in the sum, $x_i, y_i \in \mathbf{g}_i$, then

$$(x, y)_{\mathbf{g}} = \sum_i \operatorname{tr}(ad_{x_i} \cdot ad_{y_i}) = \sum_i (x_i, y_i)_{\mathbf{g}_i}.$$

If now $(x, y)_{\mathbf{g}} = 0$ for all y then each $x_i = 0$ and so $x = 0$ since the Killing forms in $\mathbf{g}_i$'s are nondegenerate (a simple Lie algebra is always semisimple). Thus $(\cdot, \cdot)_{\mathbf{g}}$ is nondegenerate and $\mathbf{g}$ is semisimple.

Corollary 2.3.21. *If* $\mathbf{g}$ *is semisimple then* $[\mathbf{g}, \mathbf{g}] = \mathbf{g}$.

Proof. Now $\mathbf{g} = \mathbf{g}_1 \oplus \cdots \oplus \mathbf{g}_n$ where $\mathbf{g}_i$'s are simple ideals. Then

$$[\mathbf{g}, \mathbf{g}] = [\mathbf{g}_1, \mathbf{g}_1] \oplus \cdots \oplus [\mathbf{g}_n, \mathbf{g}_n] = \mathbf{g}_1 \oplus \cdots \oplus \mathbf{g}_n = \mathbf{g}.$$

Exercise 2.3.22 Fix a Cartan subalgebra $\mathbf{h}$ in C_ℓ and let Φ be the root system of the Lie algebra C_ℓ. Determine the vectors $h_\alpha \in \mathbf{h}$ for $\alpha \in \Phi$.

Exercise 2.3.23 Let $\mathbf{g}$ be a 3-dimensional complex semisimple Lie algebra. Show that $\mathbf{g}$ is isomorphic to $\mathbf{sl}(2, \mathbb{C})$.

Exercise 2.3.24 Show that there is no semisimple Lie algebra of dimension four.

Exercise 2.3.25 Let $\mathbf{h}$ be the standard Cartan subalgebra of $A_\ell = \mathbf{sl}(\ell + 1, \mathbb{C})$ consisting of diagonal matrices in A_ℓ and Φ the set of roots. Determine the numbers $< \alpha, \beta >$ for $\alpha, \beta \in \Phi$.

CHAPTER 3 ROOT SYSTEMS

3.1 Reflections

In this Chapter E is a real finite-dimensional vector space with a positive definite inner product $(\cdot,\cdot)$.

A *reflection* of E is a linear map $\sigma : E \to E$ such that $\sigma(x) = -x$ for some nonzero vector x and $\sigma(y) = y$ when y belongs to the orthogonal complement $P_\sigma \subset E$ of x. The subspace P_σ is called the plane of reflection of σ. We set $\sigma_x = \sigma$ in this construction.

Explicitely, we van write $\sigma_\alpha(\beta) = \beta - 2\frac{(\beta,\alpha)}{(\alpha,\alpha)} \cdot \alpha = \beta - <\beta,\alpha> \alpha$ for $\alpha, \beta \in E$, $\alpha \neq 0$.

Theorem 3.1.1. *Let $\Phi \subset E$ be a finite subset which spans E and σ a linear automorphism of E. We assume*

(1) *$\sigma(\Phi) \subset \Phi$ and $\sigma_\beta(\Phi) \subset \Phi$ for all $\beta \in \Phi$,*
(2) *there exists a linear subspace $P \subset E$ of codimension one (a hyperplane) such that $\sigma(x) = x$ for all $x \in P$,*
(3) *there is a vector $\alpha \in \Phi$ such that $\sigma(\alpha) = -\alpha$.*

Then $P = P_\alpha$ and $\sigma = \sigma_\alpha$.

Proof. Set $\tau = \sigma\sigma_\alpha$. Then $\tau(\alpha) = \alpha$. Let P_α be the fixed point set of σ_α and P the fixed point set of σ. Then $\sigma(\beta + a\alpha) = \beta - a\alpha$ for all $\beta \in P$ and so the induced linear map $\tau : E/\mathbb{R}\alpha \to E/\mathbb{R}\alpha$ is the identity. But also $\tau(\alpha) = \alpha$ so that $\tau : \mathbb{R}\alpha \to \mathbb{R}\alpha$ is the identity. It follows that the characteristic polynomial of τ is $(\lambda - 1)^\ell$ with $\ell = \dim E$.

Let $\beta \in \Phi$. Since σ permutes the elements in the finite set Φ, we must have $\tau^m(\beta) = \beta$ for some $m = m_\beta \geq 1$. Let n be the product of the m_β's. Then $\tau^n(\beta) = \beta$ for all $\beta \in \Phi$ and $\tau^n = 1$ since Φ spans E. Therefore the *minimal polynomial* (the minimal polynomial of a matrix A is the polynomial p of smallest degree such that $p(A) = 0$) of τ divides $\lambda^n - 1$. On the other hand, the minimal polynomial divides the characteristic polynomial $(\lambda - 1)^\ell$ so that the minimal polynomial is $\lambda - 1$ and $\tau = 1$. This implies $\sigma = \sigma_\alpha$ and $P = P_\alpha$.

3.2 Axioms and basic properties of root systems

We say that a finite subset Φ of a real Euclidean vector space E is *a system of roots* if

(1) Φ spans E, $0 \notin \Phi$,

(2) if $\alpha \in \Phi$ then $k\alpha \in \Phi$ if and only if $k = \pm 1$,

(3) for any $\alpha \in \Phi$ also $\sigma_\alpha(\Phi) \subset \Phi$,

(4) for any $\alpha, \beta \in \Phi$ the real number $< \beta, \alpha >$ is an integer.

We denote by W the group generated by the reflections σ_α, $\alpha \in \Phi$. Since Φ is finite, as a subgroup of permutations of a finite set the *Weyl group* W is a finite group.

Theorem 3.2.1. *If σ is a linear automorphism of E such that $\sigma(\Phi) \subset \Phi$ then $\sigma\sigma_\alpha\sigma^{-1} = \sigma_{\sigma(\alpha)}$ for all $\alpha \in \Phi$ and*

$$< \beta, \alpha > = < \sigma(\beta), \sigma(\alpha) > \text{ for all } \alpha, \beta \in \Phi.$$

Proof. Let $\tau = \sigma\sigma_\alpha\sigma^{-1}$. Then $\tau(\Phi) \subset \Phi$ and $\tau(\sigma(\alpha)) = \sigma\sigma_\alpha(\alpha) = -\sigma(\alpha)$. When $\beta \in \sigma(P_\alpha)$ then $\tau(\beta) = \sigma\sigma_\alpha\sigma^{-1}(\beta) = \sigma\sigma^{-1}(\beta) = \beta$. We have used

$$\sigma_\alpha(\sigma^{-1}(\beta)) = \sigma^{-1}(\beta), \text{ since } \sigma^{-1}(\beta) \in P_\alpha.$$

From the previous theorem follows that $\tau = \sigma_{\sigma(\alpha)}$. For an arbitrary pair $\alpha, \beta \in \Phi$ we get

$$\sigma\sigma_\alpha\sigma^{-1}(\sigma(\beta)) = \sigma\sigma_\alpha(\beta) = \sigma(\beta - < \beta, \alpha > \alpha) = \sigma(\beta) - < \beta, \alpha > \sigma(\alpha) = \sigma_{\sigma(\alpha)}(\sigma(\beta)).$$

On the other hand,

$$\sigma\sigma_\alpha\sigma^{-1}(\sigma(\beta)) = \sigma_{\sigma(\alpha)}(\sigma(\beta)) = \sigma(\beta) - < \sigma(\beta), \sigma(\alpha) > \sigma(\alpha).$$

Comparing the right-hand-sides of these two equations we obtain $< \beta, \alpha > =$ $< \sigma(\beta), \sigma(\alpha) >$.

We say that two root systems (E, Φ) and (E', Φ') are *isomorphic* if there is a linear isomorphim $\psi : E \to E'$ such that $\psi(\Phi) = \Phi'$ and $< \psi(\beta), \psi(\alpha) > = < \beta, \alpha >$

for all $\alpha, \beta \in \Phi$. If W, W' are the corresponding Weyl groups it is easy to see that the map $\sigma \mapsto \psi \circ \sigma \circ \psi^{-1} = f(\sigma)$ gives an isomorphism of the Weyl groups: Namely, for any reflection $\sigma_\alpha \in W$ we have

$$f(\sigma_\alpha)(\beta) = \psi \circ \sigma_\alpha \circ \psi^{-1}(\beta) = \psi(\psi^{-1}(\beta) - <\psi^{-1}(\beta), \alpha> \alpha)$$
$$= \beta - <\psi^{-1}(\beta), \alpha> \psi(\alpha) = \beta - <\beta, \psi(\alpha)> \psi(\alpha).$$

It follows that $f(\sigma_\alpha) = \sigma_{\psi(\alpha)}$.

In the case $\ell = \dim E = 1$ there is only one root system, called A_1. This consists of a pair $\alpha, -\alpha$ of vectors in the real line. (The length of the vector α turns out to be irrelevant.)

When $\ell = 2$ there are several alternatives. These are denoted by $A_1 \times A_1, A_2, B_2$, and G_2 and they are described on the enclosed sheet, Appendix A:

[scale=0.80]lieqA.pdf

Let $\alpha, \beta \in \Phi$ be a pair of roots. The angle θ between α, β is defined by

$$\cos\,\theta = \frac{(\alpha, \beta)}{||\alpha|| \cdot ||\beta||}.$$

Since $<\beta, \alpha>= 2(\beta, \alpha)/(\alpha, \alpha) = 2\cos\,\theta \cdot ||\beta||/||\alpha||$ we get

$$<\beta, \alpha><\alpha, \beta>= 4\cos^2\theta.$$

According to the root system axioms $4\cos^2\theta$ is a nonnegative integer. Since $\cos^2\,\theta \leq 1$ the only options are $4\cos^2\,\theta = 0, 1, 2, 3, 4$. That is, $\theta = 0, \pi/6, \pi/4, \pi/3, \pi/2$, when $0 \leq \theta \leq \pi/2$, and $\theta = 5\pi/6, 3\pi/4, 2\pi/3, \pi$ when $\pi/2 < \theta \leq \pi$. Since we are only interested in $\cos\,\theta$, we may restrict $0 \leq \theta \leq \pi$.

Assuming that $||\beta|| \geq ||\alpha||$ and $\beta \neq \pm\alpha$ the various possibilities are listed in the table on the enclosed sheet.

Theorem 3.2.2. *Let $\alpha, \beta \in \Phi$ and $\alpha \neq \pm\beta$. If $(\alpha, \beta) > 0$ then $\alpha - \beta \in \Phi$. If $(\alpha, \beta) < 0$ then $\alpha + \beta \in \Phi$.*

Proof. Let first $(\alpha, \beta) > 0$. Then $<\alpha, \beta>> 0$ and $<\beta, \alpha>> 0$. According to the table on the enclosed sheet either $<\alpha, \beta>= 1$ or $<\beta, \alpha>= 1$. In the former case $\sigma_\beta(\alpha) = \alpha - \beta \in \Phi$ (root axioms). In the latter case $\sigma_\alpha(\beta) = \beta - \alpha \in \Phi$ so that $\alpha - \beta = -(\beta - \alpha) \in \Phi$. The case $(\alpha, \beta) < 0$ is treated similarly by replacing β by $-\beta$.

Theorem 3.2.3. *Let $\alpha, \beta \in \Phi$ and let S be the set of roots of the type $\beta + k\alpha$ for some $k \in \mathbb{Z}$. Then S is of the form $S = \{\beta + k\alpha | q \leq k \leq r\}$ for some $q \leq r \in \mathbb{Z}$.*

Proof. Antithesis: There are integers $q \leq p < s \leq r$ such that $\beta + p\alpha, \beta + s\alpha \in \Phi$ but $\beta+(s-1)\alpha, \beta+(p+1)\alpha \notin \Phi$. According to the previous theorem $(\beta+p\alpha, \alpha) \geq 0$ and $(\beta + s\alpha, \alpha) \leq 0$. Thus $(p - s)(\alpha, \alpha) \geq 0$, a contradiction since $(\alpha, \alpha) > 0$ and $p - s < 0$.

Compare this result with Corollary 2.3.12!

Theorem 3.2.4. *The reflection σ_α reverts the chain of roots $\alpha + k\beta$ (with $k \in \mathbb{Z}$). The root chain has at most four elements.*

Proof. Since $\sigma_\alpha(\beta + k\alpha) = \beta + k\alpha - < \beta + k\alpha, \alpha > \alpha = \beta + k'\alpha$, the reflection σ_α maps the chain onto itself. Here $k' = k- < \beta, \alpha > -k < \alpha, \alpha >=$
$k(1- < \alpha, \alpha >)- < \beta, \alpha >= -k- < \beta, \alpha >$. When k increases, k' decreases and because σ_α is a bijection we must have $\sigma_\alpha(\beta + r\alpha) = \beta + q\alpha$, using the notation in the previous theorem. Then

$$q = -r- < \beta, \alpha > \text{ so } q + r = - < \beta, \alpha > .$$

Since $< \beta + k\alpha, \alpha >=< \beta, \alpha > +2k$, the second statement follows from the fact that $| < \gamma, \alpha > | \leq 3$ for any root γ.

Exercise 3.2.5 Let Φ be a root system. Set $\alpha^\vee = 2\alpha/(\alpha, \alpha)$. Show that the set of all $\alpha^\vee$'s form a root system $\Phi^\vee$. Draw $\Phi^\vee$ when $\Phi = A_1, A_2, B_2, G_2$.

Exercise 3.2.6 Determine the root chains $\beta + k\alpha$ when $\Phi = G_2$.

Exercise 3.2.7 Show the the Weyl group of A_2 is isomorphic with the group of permutations S_3 of three objects.

Exercise 3.2.8 The automorphism group $\mathrm{Aut}\,\Phi$ of a root system (E, Φ) consists of all linear isomorphisms $\phi : E \to E$ with $\phi(\Phi) = \Phi$ and $< \phi(\alpha), \phi(\beta) >=< \alpha, \beta >$. Show that the Weyl group is a normal subgroup of $\mathrm{Aut}\,\Phi$.

3.3 Simple roots

A subset Δ in a system of roots $\Phi \subset E$ is a *system of simple roots* if

(1) Δ is a basis in the vector space E,

(2) and all roots in Φ can be expressed as $\sum k_\alpha \alpha$, where all coefficients k_α are either nonnegative integers or all of them are nonpositive integers.

If $\ell = \dim E$ then Δ has exactly ℓ elements. The *height* $h(\alpha)$ of a root γ is defined as the sum of the coefficients k_α in the expansion. We split $\Phi = \Phi^+ \cup \Phi^-$ as the union of *positive and negative roots* according to whether the coefficients are nonnegative or nonpositive.

We also define a partial order in Φ by declaring that $\alpha > \beta$ if $\alpha - \beta$ is a positive root.

Lemma 3.3.1. *For any pair of different simple roots* $(\alpha, \beta) \leq 0$ *and* $\alpha - \beta \notin \Phi$.

Proof. The second condition follows immediately from the definition of a set of simple roots Δ. If $(\alpha, \beta) > 0$ then by 3.2.2 $\alpha - \beta$ is a root, a contradiction.

For any $\gamma \in \Phi$ we denote $\Phi^+(\gamma) = \{\alpha \in \Phi | (\gamma, \alpha) > 0\}$.

We say that a vector $\gamma \in E$ is *regular* if it does not belong to any of the $\ell - 1$ dimensional hyperplanes P_α for $\alpha \in \Phi$; otherwise γ is *singular.*

For a regular vector γ we have clearly $\Phi = \Phi^+(\gamma) \cup \Phi^-(\gamma)$, where $\Phi^-(\gamma) = -\Phi^+(\gamma)$. We say that a root $\alpha \in \Phi^+(\gamma)$ is *decomposable* if $\alpha = \alpha_1 + \alpha_2$ for some $\alpha_1, \alpha_2 \in \Phi^+(\gamma)$, otherwise it is *indecomposable.*

Theorem 3.3.2. *Let* $\gamma \in E$ *be regular and* $\Delta(\gamma)$ *the set of indecomposable elements in* $\Phi^+(\gamma)$. *Then* $\Delta(\gamma)$ *is a set of simple roots. Any set of simple roots is of this form.*

Proof. (1) We claim that each $\alpha \in \Phi^+(\gamma)$ is a linear combination of elements in $\Delta(\gamma)$ with nonnegative coefficients. *Antithesis:* There exits $\alpha \in \Phi^+(\gamma)$ which cannot be expressed as such a linear combination. We choose α among those elements such that (α, γ) is minimal. Since $\alpha \notin \Delta(\gamma)$ we have $\alpha = \alpha_1 + \alpha_2$ for some $\alpha_1, \alpha_2 \in \Phi^+(\gamma)$. Then $(\gamma, \alpha) = (\gamma, \alpha_1) + (\gamma, \alpha_2)$ and so $(\gamma, \alpha_1) < (\gamma, \alpha)$ and $(\gamma, \alpha_2) < (\gamma, \alpha)$ and by the minimality property of α it follows that both α_1, α_2 are linear combinations of elements in $\Delta(\gamma)$ with nonnegative coefficients. Thus also α is a linear combination in $\Delta(\gamma)$ with nonnegative coefficients.

(2) We prove that for any $\alpha, \beta \in \Delta(\gamma)$ either $\alpha = \beta$ or $(\alpha, \beta) \leq 0$. Otherwise, we would have $\alpha - \beta \in \Phi$ (Theorem 3.2.2). If now $\alpha - \beta \in \Phi^+(\gamma)$ then $\alpha = \beta + (\alpha - \beta)$

is decomposable, a contradiction. But in the case $\alpha - \beta \in \Phi^-(\gamma)$ we have $\beta = \alpha + (\beta - \alpha)$, a contradiction since β was assumed to be indecomposable.

(3) We claim that the set $\Delta(\gamma)$ is linearly independent. Let $\sum_{\alpha\in\Delta(\gamma)} a_\alpha \alpha = 0$ for some $a_\alpha \in \mathbb{R}$. We can write

$$\theta = \sum a_\alpha \alpha = \sum_{\Delta_1} b_\alpha \alpha + \sum_{\Delta_2} c_\alpha \alpha = \theta_+ + \theta_- = 0$$

where $b_\alpha > 0$ and $c_\alpha < 0$ and Δ_i are disjoint subsets of $\Delta(\gamma)$. Then by (2) we have

$$(\theta_+, \theta_+) = -\sum_{\alpha,\beta} b_\alpha c_\beta (\alpha, \beta) \leq 0.$$

It follows that $\theta_+ = 0$. Now $0 = (\gamma, \theta_+) = \sum_{\Delta_1} b_\alpha(\gamma, \alpha)$ so that $b_\alpha = 0$ for all $\alpha \in \Delta_1$. In the same way $0 = (\gamma, \theta_+) = -\sum_{\Delta_2} c_\alpha(\gamma, \alpha)$ so $c_\alpha = 0$ for all $\alpha \in \Delta_2$ because $(\gamma, \alpha) > 0$ for all $\alpha \in \Delta(\gamma)$. Thus all coefficients a_α vanish.

(4) $\Delta(\gamma)$ is a system of simple roots: The second axiom follows from (1) above. Since Φ spans E the first axiom follows from (3) and (1).

(5) Let $\Delta \subset \Phi$ be a system of simple roots. We prove that $\Delta = \Delta(\gamma)$ for some regular γ. Set $\Delta = \{\alpha_1, \ldots, \alpha_\ell\}$. Consider the system of linear equations

$$\sum_i (\alpha_i, \alpha_j) x_i = a_j \text{ with } j = 1, \ldots, \ell; a_i, x_i \in \mathbb{R}.$$

The a_i's are given real numbers and x_i's are to be determined. The corresponding homogeneous system ($a_i = 0$) has only the trivial solution $x_i = 0$ since the system Δ is a basis of E. Thus the inhomogeneous system has a unique solution x for any vector a. For example, we may choose $a_i = 1$ for all i we have a unique solution x and we denote $\gamma = \sum_i x_i \alpha_i$. Then $(\gamma, \alpha) > 0$ for all $\alpha \in \Delta$. By the second axiom of simple roots we must have $(\gamma, \alpha) \neq 0$ for all $\alpha \in \Phi$ and γ is regular. Furthermore, $(\gamma, \alpha) > 0$ for all $\alpha \in \Phi^+$ and $(\gamma, \alpha) < 0$ for all $\alpha \in \Phi^-$ and therefore $\Phi^+ \subset \Phi^+(\gamma)$ and $\Phi^- \subset \Phi^-(\gamma)$. Consequently, $\Phi^\pm = \Phi^\pm(\gamma)$.

Let $\beta_1, \beta_2 \in \Phi^+$. Then the height $h(\beta_1 + \beta_2) = h(\beta_1) + h(\beta_2) \geq 2$. For $\beta \in \Delta$ the height $= 1$ by definition. On the other hand $h(\beta) = 1$ for all $\beta \in \Delta$ and β is indecomposable (with respect to γ). Thus $\Delta \subset \Delta(\gamma)$. Since both form a basis of E we have finally $\Delta = \Delta(\gamma)$.

The connected components of the open set $E \setminus \underset{\alpha\in\Phi}{U} P_\alpha$ are called *Weyl chambers.* There are finitely many Weyl chambers since the set of roots is finite. For any regular $\gamma \in E$ we denote $T(\gamma)$ the Weyl chamber containing the vector γ.

If $\gamma' \in T(\gamma)$ then both γ, γ' are on the same side of each hyperplane P_α and therefore $\Phi^+(\gamma) = \Phi^+(\gamma')$ and $\Delta(\gamma) = \Delta(\gamma')$. It follows from theorem 3.3.2 that there is a 1-1 correspondence between the set of Weyl chambers and the systems of simple roots, $T(\gamma) \mapsto \Delta(\gamma)$. Given a system of simple roots $\Delta = \Delta(\gamma)$ we call $T(\gamma)$ the *fundamental Weyl chamber,* denoted by $T(\Delta)$. Then $T(\Delta) = \{\gamma \in E | (\gamma, \alpha) > 0 \, \forall \alpha \in \Delta\}$.

Theorem 3.3.3. *Let $\gamma \in E$ be regular and $\sigma \in W$. Then $\sigma(T(\gamma)) = T(\sigma(\gamma))$.*

Proof. When $\gamma' \in T(\gamma)$ then γ' is on the same side of each hyperplane P_α as γ. Now (γ, α) and (γ', α) have the same signs is equivalent to the statement that $(\sigma(\gamma), \sigma(\alpha))$ and $(\sigma(\gamma'), \sigma(\alpha))$ have same signs, by theorem 3.2.1. Since σ permutes the roots, the last staement is equivalent to saying that $(\sigma(\gamma), \alpha)$ and $(\sigma(\gamma'), \alpha)$ have same signs for all $\alpha \in \Phi$. This means that $\sigma(\gamma), \sigma(\gamma')$ are on the same side of each hyperplane P_α and so belong to the same Weyl chamber. This implies $\sigma(\gamma') \in T(\sigma(\gamma)$ for all $\gamma' \in T(\gamma)$ and thus $\sigma(T(\gamma)) \subset T(\sigma(\gamma))$. Likewise, $\sigma^{-1}(T(\sigma(\gamma))) \subset T(\gamma)$ and so $T(\sigma(\gamma)) \subset \sigma(T(\gamma))$. Combining these two inclusions we obtain the claim.

Remark We have used the fact that each element of the Weyl group is a linear isometry in E. This follows from the fact that elements of W are products of reflections and from 3.3.1.

Lemma 3.3.4. *Let $\Delta \subset \Phi$ be a system of simple roots and α a positive root not included in Δ. Then there is a simple root β such that $\alpha - \beta$ is a positive root.*

Proof. If $(\alpha, \beta) \leq 0$ for all $\beta \in \Delta$ then by the proof of (3) in Theorem 3.3.2 the set $\{\Delta, \alpha\}$ would be linearly independent, which is absurd since $\Delta \subset E$ is a basis.

Thus $(\alpha, \beta) > 0$ at least for one simple root β. Then $\alpha - \beta \in \Phi$, by Theorem 3.2.2. But since any positive root is a linear combination of simple roots with nonnegative coefficients it follows that the coefficient of β in the linear combination $\alpha = \sum_{\gamma \in \Delta} k_\gamma \gamma$ must be at least one. Then $\alpha - \beta$ is also a linear combination with nonnegative coefficients, thus a positive root.

Corollary 3.3.5. *Any positive root β can be written as a sum $\beta = \alpha_1 + \alpha_2 + \dots \alpha_n$, where each α_i is a simple root and each partial sum $\alpha_1 + \cdots + \alpha_i$ is a root.*

3.4 The Weyl group

Lemma 3.4.1. *Let α be a simple root. Then σ_α permutes the roots in $\Phi^+ \setminus \{\alpha\}$.*

Proof. Let $\beta \neq \alpha$ be a positive root. We can write $\beta = \sum_{\gamma \in \Delta} k_\gamma \gamma$ with nonnegative integers k_γ. Now $k_\gamma > 0$ for some $\gamma \neq \alpha$. But then

$$\sigma_\alpha(\beta) = \sum_\gamma k_\gamma(\gamma - <\gamma, \alpha> \alpha)$$

and so $\sigma_\alpha(\beta) = \sum k'_\gamma \gamma$ with $k'_\gamma > 0$ for some $\gamma \neq \alpha$ and $k'_\gamma \geq 0$ for all $\gamma \in \Delta$. This implies $\sigma_\alpha(\beta) \in \Phi^+$. Furthermore, $\sigma_\alpha(-\alpha) = \alpha$ and so $\sigma_\alpha(\beta) \neq \alpha$.

Corollary 3.4.2. *Let $\delta = \frac{1}{2}\sum_{\beta \in \Phi^+} \beta$. Then $\sigma_\alpha(\delta) = \delta - \alpha$ for each simple root α.*

Lemma 3.4.3. *Let $\alpha_1, \alpha_2 \dots, \alpha_k$ be some set of simple roots. Denote $\sigma_i = \sigma_{\alpha_i}$. If $\sigma_1\sigma_2 \dots \sigma_{k-1}(\alpha_k)$ is negative then $\sigma_1 \dots \sigma_k = \sigma_1 \dots \sigma_{s-1}\sigma_{s+1} \dots \sigma_{k-1}$ for some $1 \leq s < k$.*

Proof. Denote $\beta_j = \sigma_{j+1} \dots \sigma_{k-1}(\alpha_k)$ for $0 \leq j \leq k-2$ and set $\beta_{k-1} = \alpha_k$. Then $\beta_0 < 0$ and $\beta_{k-1} > 0$. Let s be the smallest number for which $\beta_s > 0$. Then $\sigma_s(\beta_s) = \sigma_{\alpha_s}(\beta_s) < 0$, so that by Lemma 3.4.1 $\alpha_s = \beta_s$. By Theorem 3.2.1 we have $\sigma\sigma_\alpha\sigma^{-1} = \sigma_{\sigma(\alpha)}$ so that

$$\sigma_s = \sigma_{\alpha_s} = (\sigma_{s+1} \dots \sigma_{k-1})\sigma_k(\sigma_{s+1} \dots \sigma_{k-1})^{-1}$$

from which follows $\sigma_1 \dots \sigma_k = \sigma_1 \dots \sigma_{s-1}\sigma_{s+1} \dots \sigma_{k-1}$, using $\sigma_\alpha = \sigma_\alpha^{-1}$.

Corollary 3.4.4. *Let $\sigma = \sigma_1 \dots \sigma_k$ (with $\sigma_i = \sigma_{\alpha_i}$, $\alpha_i \in \Delta$) be a shortest decomposition of σ to a product of simple reflections. Then $\sigma(\alpha_k) < 0$.*

Proof. If $\sigma(\alpha_k) > 0$ then $\sigma_1 \dots \sigma_{k-1}(\alpha_k) < 0$. This is in contradiction with the minimality of the decomposition, Lemma 3.4.3.

Theorem 3.4.5. *Let Δ be a system of simple roots.*

(1) *For any regular vector γ there is $\sigma \in W$ such that $\sigma(\gamma) \in T(\Delta)$*
(2) *If Δ' is another system of simple roots then $\sigma(\Delta') = \Delta$ for some $\sigma \in W$*
(3) *For any root α there is $\sigma \in W$ such that $\sigma(\alpha) \in \Delta$*
(4) *The simple reflections $\sigma_\alpha, \alpha \in \Delta$, generate the group W*
(5) *If $\sigma \in W$ is such that $\sigma(\Delta) = \Delta$ then $\sigma = 1$.*

Proof. Denote by W' the subgroup of W generated by the simple reflections. We prove first that (1)-(3) hold for the subgroup W'.

(1) Denote again by δ half the sum of positive roots. Choose first $\sigma \in W'$ such that $(\sigma(\gamma), \delta)$ obtains its maximum value. When $\alpha \in \Delta$ then $\sigma_\alpha \sigma \in W'$ so that

$$(\sigma(\gamma), \delta) \geq (\sigma_\alpha \sigma(\gamma), \delta) = (\sigma(\gamma), \sigma_\alpha(\delta)) = (\sigma(\gamma), \delta - \alpha) = (\sigma(\gamma), \delta) - (\sigma(\gamma), \alpha)$$

and thus $(\sigma(\gamma), \alpha) \geq 0$ for all $\alpha \in \Delta$. On the other hand, $(\sigma(\gamma), \alpha) = (\gamma, \sigma^{-1}(\alpha) \neq 0$ by the regularity of γ. Thus $(\sigma(\gamma), \alpha) > 0$ for all $\alpha \in \Delta$ and $\sigma(\gamma) \in T(\Delta)$.

(2) Let $\Delta' = \Delta(\gamma)$. By (1) above, there exists $\sigma \in W'$ such that $\sigma(\gamma) \in T(\Delta)$. But then $\sigma(\Delta') = \Delta(\sigma(\gamma)) = \Delta$.

(3) By (2) it is sufficient to show that any root belongs to *some* system of simple roots. So let α be a root. Since $P_\beta \neq P_\alpha$ for all roots $\beta \neq \pm\alpha$ we may choose $\gamma \in P_\alpha$ such that $\gamma \notin P_\beta$ for roots $\beta \neq \pm\alpha$. Set $\gamma' = \gamma + \epsilon\alpha$ where

$$\epsilon = \frac{1}{2} \cdot \min_{\beta \neq \pm\alpha} \frac{|(\gamma, \beta)|}{|(\alpha, \beta)| + (\alpha, \alpha)}.$$

Then $|(\gamma', \beta)| > (\gamma', \alpha) = \epsilon(\alpha, \alpha) > 0$ for all roots $\beta \neq \pm\alpha$. Now γ' is regular and $\alpha \in \Phi^+(\gamma')$. Because of $(\gamma', \beta) > (\gamma', \alpha)$ for $\beta \in \Phi^+(\gamma')$ the root α is indecomposable in $\Phi^+(\gamma')$ and thus $\alpha \in \Delta(\gamma')$ by Theorem 3.3.2.

(4) It is enough to show that each reflection σ_α with $\alpha \in \Phi$ is a product of simple reflections. Choose $\sigma \in W'$ such that $\beta = \sigma(\alpha) \in \Delta$. Then

$$\sigma_\beta = \sigma_{\sigma(\alpha)} = \sigma \sigma_\alpha \sigma^{-1} \text{ so } \sigma_\alpha = \sigma^{-1} \sigma_\beta \sigma \in W'.$$

(5) Let $\sigma(\Delta) = \Delta$ for $\sigma \in W$. If now $\sigma \neq 1$ then we can write $\sigma = \sigma_1 \dots \sigma_k$ with $\sigma_i = \sigma_{\alpha_i}$ and $\alpha_i \in \Delta$ and $k \geq 1$. We choose k minimal. From Cor. 3.4.4 follows that $\sigma(\alpha_k) < 0$ which is absurd since $\sigma(\alpha_k) \in \Delta$.

Exercise 3.4.6 Let Φ be a system of roots in $E = \mathbb{R}^2$. Show that it is isomorphic to one of the systems $A_1 \times A_1, A_2, B_2$ or G_2.

Exercise 3.4.7 Determine a system of simple roots for each of the cases in Exercise 3.4.6.

Exrcise 3.4.8 Let $\Delta \subset \Phi$ be a system of simple roots. Let $\alpha \neq \beta$ be a pair of simple roots and let Φ' be the subsystem consisting of roots in Φ which are integral linear combinations of α and β. Show that Φ' is a 2-dimensional root system.

Exercise 3.4.9 Let $\Delta \subset \Phi$ be a system of simple roots. Show that there is a unique $\sigma \in W$ such that $\sigma(\Phi^+) = \Phi^-$. Hint: The set $-\Delta$ is another system of simple roots. Use Theorem 3.4.5.

Exercise 3.4.10 Show by direct inspection of the root system that Cor. 3.3.5 holds for the root system G_2.

3.5 Classification of root systems

A root system (E, Φ) is *irreducible* if it is not possible to write $\Phi = \Phi_1 \cup \Phi_2$ as a union of two (nonempty) root systems with $\Phi_1 \perp \Phi_2$. It is clear that any root system is a direct sum of irreducible systems, so it is sufficient to classify the irreducible systems.

We shall skip most of the proofs in this section; they can be found in Section 11.4 in J. Humphrey's book *Introduction to Lie Algebras and Representation Theory.*

The first fact which we list without proof is that *in any irreducible root system there are at most two different root lengths;* the roots are either *long or short roots.* If all roots have the same length we call them long roots.

Let $\Delta = \{\alpha_1, \dots, \alpha_\ell\}$ be a system of simple roots. Denote

$$M_{ij} =< \alpha_i, \alpha_j >= 2 \cdot \frac{(\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)}.$$

The numbers M_{ij} form a $\ell \times \ell$ integral matrix, called the *Cartan matrix* of the root system. In the 2-dimensional cases we have the matrices

$$A_1 \times A_1 \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}; A_2 \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}; B_2 \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}; G_2 \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}.$$

When Δ' is another basis then $\sigma(\Delta) = \Delta'$ for some $\sigma \in W$. The brackets $< \alpha, \beta >$ are invariant under the Weyl group. It follows that the Cartan matrix does not depend on the choice of Δ, modulo reordering of the basis.

Theorem 3.5.1. *Let (E, Φ) and (E', Φ') be a pair of root systems with $\Delta \subset \Phi$ and $\Delta' \subset \Phi'$ systems of simple roots. If the Cartan matrices M and M' are equal (with some choice of ordering of basis) then the root systems are isomorphic.*

Proof. Set $\Delta = \{\alpha_1, \dots, \alpha_\ell\}$ and $\Delta' = \{a'_1, \dots, \alpha'_\ell\}$. We can define a linear isomorphism $\phi : E \to E'$ by $\phi(\alpha_i) = a'_i$ since the simple roots form a basis. Then for any

$\alpha, \beta \in \Delta$,

$$\sigma_{\phi(\alpha)}(\phi(\beta)) = \phi(\beta) - < \phi(\beta), \phi(\alpha) > \phi(\alpha)$$
$$= \phi(\beta) - < \beta, \alpha > \phi(\alpha) = \phi(\beta - < \beta, \alpha > \alpha) = \phi(\sigma_\alpha(\beta)).$$

The second equality follows from the asssumption that the Cartan matrices are equal. Since Δ is a basis, we obtain $\sigma_{\phi(\alpha)} \circ \phi = \phi \circ \sigma_\alpha$, that is, $\phi \circ \sigma_\alpha \circ \phi^{-1} = \sigma_{\phi(\alpha)}$ for all $\alpha \in \Delta$. Since the simple reflections generate the Weyl group, we reduce that the map $\sigma \to \phi \circ \sigma \circ \phi^{-1}$ from W to W' is *an isomorphims of Weyl groups.*

Let next $\beta \in \Phi$ and choose $\sigma \in W$ such that $\sigma(\beta) \in \Delta$, Theorem 3.4.5 (3). Then

$$\phi(\beta) = (\phi \circ \sigma^{-1} \circ \phi^{-1})\phi(\sigma(\beta)) \in \Phi'$$

and so $\phi(\Phi) \subset \Phi'$. In the same way one shows that $\phi^{-1}(\Phi') \subset \Phi$ and thus $\phi(\Phi) = \Phi'$. If γ is another element of Φ then, by the linearity of $< \cdot, \cdot >$ in the first argument and by the equality of Cartan matrices,

$$< \gamma, \beta > =< \sigma(\gamma), \sigma(\beta) >=< \phi \circ \sigma(\gamma), \phi \circ \sigma(\beta) >$$
$$=< (\phi \circ \sigma^{-1} \circ \phi^{-1})(\phi \circ \sigma(\gamma)), (\phi \circ \sigma^{-1} \circ \phi^{-1})(\phi \circ \sigma(\beta)) >=< \phi(\gamma), \phi(\beta) > .$$

We have used the fact that the Weyl groups W, W' preserve the brackets. We have shown that ϕ is an isomorphism of the root systems.

We have seen that if $\alpha \neq \beta$ is a pair of positive roots then $< \alpha, \beta >< \beta, \alpha >$ is one of the integers $0, 1, 2, 3$. We determine *the Coxeter graph* of the root system Φ from its Cartan matrix. The graph consists of ℓ nodes corresponding to the number of simple roots and lines connecting the nodes. The number of lines connecting the nodes α_i, α_j (for $i \neq j$) is equal to $< \alpha_i, \alpha_j >< \alpha_j, \alpha_i >$.

In the case when all simple roots have equal lengths the *Dynkin diagram* is equal to the Coxeter graph. In the case when a pair α_i, α_j of simple roots have unequal lengths we set an arrow to point towards the shorter root. On the enclosed sheet B we list all the Dynkin diagrams of simple Lie algebras.

[scale=0.80]lieqB.pdf

The Dynkin diagram determines completely the Cartan matrix and therefore also the root system of a semisimple Lie algebra. In the case when the simple root lengths are equal, we have $< \alpha_i, \alpha_j >= -(< \alpha_i, \alpha_j >< \alpha_j, \alpha_i >)^{1/2}$, for

$i \neq j$. This gives all the matrix elements of the Cartan matrix. Suppose then that $(\alpha_i, \alpha_i) \neq (\alpha_j, \alpha_j)$ but we know that α_i is shorter, for example. Then from the table of of root lengths and angles we see that $< \alpha_i, \alpha_j >< \alpha_j, \alpha_i >$ is either 2 or 3. In the former case $< \alpha_i, \alpha_j >= -1$ and $< \alpha_j, \alpha_i >= -2$. In the latter case $< \alpha_i, \alpha_j >= -1$ and $< \alpha_j, \alpha_i >= -3$.

For example, from the Dynkin diagram of F_4 we can read its Cartan matrix

$$F_4 : \qquad \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -2 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

A root system Φ is *irreducible* when its Dynkin diagram is connected. Let $\Delta = \Delta_1 \cup \Delta_2 \cdots \cup \Delta_t$ be a decomposition of the simple roots corresponding to the connected components of the Dynkin diagram. Then $\Delta_i \perp \Delta_j$ for $i \neq j$ and let E_i be the subspace of E spanned by the roots Δ_i, $E = E_1 \oplus \cdots \oplus E_t$. Denote Φ_i the subset of roots which are linear combinations of the roots Δ_i.

Now the Weyl group W maps Φ_i onto itself: To see this it is sufficient to show that $\sigma_\alpha(\Phi_i) \subset \Phi_i$ for any simple root α. If $\alpha \notin \Delta_i$ then $\sigma_\alpha(\beta) = \beta - < \beta, \alpha > \alpha = \beta$ for any $\beta \in \Phi_i$. But if $\alpha \in \Delta_i$ then $\sigma_\alpha(\beta) = \beta - < \beta, \alpha > \alpha \in \Phi_i$ by the definition of Φ_i.

If $\beta \in \Phi$ is an arbitrary root we may choose $\sigma \in W$ such that $\sigma(\beta) \in \Delta$. But then $\sigma(\beta)$ belongs to some Δ_i and by the observation above $\beta \in \Phi_i$. Thus we have

$$\Phi = \Phi_1 \cup \Phi_2 \cdots \cup \Phi_t.$$

We have proven:

Theorem 3.5.2. *Any root system $\Phi \subset E$ is a union of irreducible root systems $\Phi_i \subset E_i$ with $E = E_1 \oplus \cdots \oplus E_t$, as an orthogonal direct sum.*

Now we list all irreducible root systems in Theorems 3.5.3 - 3.5.11. We denote the standard basis vectors in $\mathbb{R}^\ell$ by $\epsilon_1, \ldots, \epsilon_\ell$.

Theorem 3.5.3. *Let E be the subspace of the euclidean space $\mathbb{R}^{\ell+1}$ with $\ell \geq 1$ consisting of vectors α such that $(\alpha, \sum \epsilon_i) = 0$. Let L be the integral lattice in E and set $\Phi = \{\alpha \in L | (\alpha, \alpha) = 2\}$. Then (E, Φ) is an irreducible root system and its Dynkin diagram is the Dynkin diagram of the Lie algebra A_ℓ.*

Proof. Clearly

$$\Phi = \{\epsilon_i - \epsilon_j | i \neq j\}.$$

Let Δ consist of the vectors $\alpha_i = \epsilon_i - \epsilon_{i+1}$ with $i = 1, 2, \ldots, \ell$. These vectors form a basis of E. Furthermore, each element in Φ is an integral linear combination of vectors in Δ with only nonnegative or only nonpositive coefficients, so it satisfies the requirements of a system of simple roots; we also observe that clearly the first two axioms of a root system are satisfied. Next $< \beta, \alpha >= 2(\beta, \alpha)/(\alpha, \alpha) = (\beta, \alpha) \in \{0, \pm 1, 2\}$ so that also the fourth axiom holds.

By a direct computation (Exercise!) we observe that
$\sigma_\alpha(\beta) = \beta - < \beta, \alpha > \alpha$ belongs to Φ for any $\alpha, \beta \in \Phi$ and so indeed Φ is a root system. Since $< \alpha_i, \alpha_{i+1} >= (\alpha_i, \alpha_{i+1}) = -1$ but $< \alpha_i, \alpha_j >= 0$ for $j \neq i \pm 1$ we see that the Dynkin diagram is really the diagram A_ℓ listed in the appendix B; one can then check by direct computation that the root system corresponding to the Cartan subalgebra of diagonal matrices in $\mathbf{sl}(\ell + 1, \mathbb{F})$, with the choice of simple roots corresponding to the root vectors $e_{i,i+1} \in \mathbf{sl}(\ell + 1, \mathbb{F})$, leads to the system (E, Φ, Δ).

Theorem 3.5.4. *Let $E = \mathbb{R}^\ell$ with $\ell \geq 2$ and Φ the set of vectors α in its integral lattice L such that $(\alpha, \alpha) = 1$ or $(\alpha, \alpha) = 2$. Then (E, Φ) is an irreducible system of roots with a Dynkin diagram corresponding to the Lie algebra B_ℓ.*

Proof. Now $\Phi = \{\pm\epsilon_i | 1 \leq i \leq \ell\} \cup \{\pm(\epsilon_i \pm \epsilon_j) | i \neq j\}$. The subset Δ of vectors $\alpha_i = \epsilon_i - \epsilon_{i+1}, i \leq \ell - 1$, and $\alpha_\ell = \epsilon_\ell$ is linearly independent and the number of vectors is equal to the dimension of E, thus it is a basis of E. Furthermore,

$$\begin{aligned}
\pm\epsilon_i &= \pm(\alpha_i + \ldots \alpha_\ell) \\
\pm(\epsilon_i - \epsilon_j) &= \pm(\alpha_i + \cdots + \alpha_j) \text{ for } i < j \\
\pm(\epsilon_i + \epsilon_j) &= \pm(\alpha_i + \cdots + \alpha_{j-1} + 2\alpha_j + 2\alpha_{j+1} + \ldots 2\alpha_\ell) \text{ for } i < j.
\end{aligned}$$

So Δ has the properties of a system of simple roots. When $i, j \leq \ell - 1$ the length of the roots α_i, α_j is equal to $\sqrt{2}$ and $< \alpha_i, \alpha_j >= 0$ for $i \neq j \pm 1, i \neq j$. For $j = i + 1$ we have $< \alpha_i, \alpha_{i+1} >< \alpha_{i+1}, \alpha_i >= 1$. The length of α_ℓ is 1 and $< \alpha_{\ell-1}, \alpha_\ell >< \alpha_\ell, \alpha_{\ell-1} >= 2$. It follows that the Dynkin diagram is the diagram B_ℓ in the appendix. This Dynkin diagram can be reduced from the results of last week's exercises; see the computations for the Lie algebra of antisymmetric $(2\ell+1)\times(2\ell+1)$ antisymmetric matrices.

Theorem 3.5.5. *Let $E=\mathbb{R}^\ell$ with $\ell \geq 3$ and $\Phi = \{\pm 2\epsilon_i | 1 \leq i \leq \ell\} \cup \{\pm(\epsilon_i \pm \epsilon_j) | i \neq j\}$. Then (E,Φ) is an irreducible root system corresponding to the Dynkin diagram C_ℓ.*

Remark We could have defined also C_2 but then $C_2 = B_2$.

Theorem 3.5.6. *Let $E=\mathbb{R}^\ell$ for $\ell \geq 4$ and define Φ as the set of vectors α in the integral lattice with $(\alpha,\alpha)=2$. Then $\Phi\{\pm(\epsilon_i \pm \epsilon_j)|i \neq j\}$ and it is an irreducible root system with Dynkin diagram D_ℓ corresponding to the Lie algebra of antisymmetric $2\ell \times 2\ell$ matrices.*

Proof. This is actually a subalgebra of B_ℓ, by leaving out the short roots $\pm\epsilon_i$. The simple roots are $\alpha_i = \epsilon_i - \epsilon_{i+1}$ for $i = 1,2,\ldots,\ell-1$ and $\alpha_\ell = \epsilon_{\ell-1}+\epsilon_\ell$.

Then we have the root systems of *exceptional simple Lie algebras.* It is left as an exercise to the reader to check that the axioms of root systems are satisfied.

Theorem 3.5.7. (G_2) *The following is an irreducible two dimensional root system: Let $\{\epsilon_1,\epsilon_2,\epsilon_3\}$ be the standard basis of $\mathbb{R}^3$ and let E be the plane orthogonal to $\epsilon_1+\epsilon_2+\epsilon_3$. A basis of E is given by $\{\epsilon_1-\epsilon_2, -2\epsilon_1+\epsilon_2+\epsilon_3\} = \Delta$. This is a system of simple roots for G_2. The positive roots are $\Phi^+ = \{\epsilon_1-\epsilon_2, -\epsilon_1+\epsilon_3, -\epsilon_2+\epsilon_3, -2\epsilon_1+\epsilon_2+\epsilon_3, \epsilon_1-2\epsilon_2+\epsilon_3, -\epsilon_1-\epsilon_2+2\epsilon_3\}$.*

Theorem 3.5.8. (F_4) *Let $E=\mathbb{R}^4$ and $\Delta = \{\epsilon_2-\epsilon_3, \epsilon_3-\epsilon_4, \epsilon_4, \frac{1}{2}(\epsilon_1-\epsilon_2-\epsilon_3-\epsilon_4)\}$. The root system of F_4 consists of all integral* linear combinations α of elements in Δ such that $\|\alpha\|^2 = 1$ or $\|\alpha\|^2=2$. Then $\Phi = \{\pm\epsilon_i\}_{i=1}^4 \cup \{\pm(\epsilon_i \pm \epsilon_j) \mid i \neq j\} \cup \{\pm\frac{1}{2}(\epsilon_1 \pm \epsilon_2 \pm \epsilon_3 \pm \epsilon_4) \mid \text{all signs}\}$. Thus the number of elements in Φ is 48.

Exercise What is the system of positive roots for F_4?

Theorem 3.5.9. (E_8) *Let $E=\mathbb{R}^8$ and $\Delta = \{\frac{1}{2}(\epsilon_1+\epsilon_8) - \frac{1}{2}(\epsilon_2+\ldots+\epsilon_7), \epsilon_1+\epsilon_2, \epsilon_2-\epsilon_1, \epsilon_3-\epsilon_2, \epsilon_4-\epsilon_3, \epsilon_5-\epsilon_4, \epsilon_6-\epsilon_5, \epsilon_7-\epsilon_6\}$. The root system $\Phi(E_8)$ consists of all integral linear combinations α of elements in Δ such that $\|\alpha\|^2=2$. Then*

$$\Phi = \{\pm(\epsilon_i \pm \epsilon_j) \mid i \neq j\} \cup \{\frac{1}{2}\sum_{i=1}^{8}(-1)^{\epsilon(i)}\epsilon_i \mid \epsilon(i) = 0,1;\ \sum \epsilon(i) \in 2\mathbb{Z}\}.$$

There are 240 elements in Φ.

Theorem 3.5.10. (E_7) *Δ and Φ are defined here in a similar way as in the case of E_8 except that the last vector $\epsilon_7 - \epsilon_6$ in Δ is left out. There are 126 roots.*

Theorem 3.5.11. (E_6) *Same as above, but now the two last vectors $\epsilon_6 - \epsilon_5$ and $\epsilon_7 - \epsilon_6$ in Δ are dropped. The number of roots is 72.*

Exercise 3.5.12 Let $\mathbf{g} = \mathbf{sl}(2, \mathbb{C})$ and $\mathbf{h}, \mathbf{h}'$ a pair of Cartan subalgebras of $\mathbf{g}$. Construct an automorphism $\phi : \mathbf{g} \to \mathbf{g}$ such that $\phi(\mathbf{h}) = \mathbf{h}'$. Hint: Any Cartan subalgebra of $\mathbf{g}$ is one dimensional. Show from the definition of a Cartan subalgebra that if $w = ax + by + ch$ is a basis of $\mathbf{h}$ (here x, y, h are the vectors in the standard basis) then $ab \neq -c^2$.

Exercise 3.5.13 A *Borel subalgebra* of a semisimple Lie algebra $\mathbf{g}$ is a maximal solvable subalgebra in $\mathbf{g}$. Let $\mathbf{h} \subset \mathbf{g}$ be a Cartan subalgebra and Φ the system of roots. Show that

$$\mathbf{b} = \mathbf{h} \bigoplus_{\alpha \in \Phi^+} \mathbf{g}_\alpha$$

is a Borel subalgebra.

Exercise 3.5.14 Show that the map $\alpha \to -\alpha$ is an isomorphism of the root system Φ of a semisimple Lie algebra.

3.6 Existence and uniqueness theorems

In the previous section we have listed all irreducible root systems. On the other hand, by inspection of the root systems of the simple Lie algebras $A_\ell, B_\ell, C_\ell, D_\ell, G_2$, F_4, E_6, E_7, E_8 one obeserves that these Lie algebras correspond exactly to the given root systems.

There are still several unanswered questions: Do these Lie algebras exhaust the list of all simple Lie algebras? What about general semisimple Lie algebras? What happens to the root system when we choose a different Cartan subalgebra? Is the correspondence between (isomorphism classes of) semisimple Lie algebras and (isomorphism classes of) root systems 1-1?

In this section we shall state the theorems answering these questions, but mostly without proofs. For proofs the reader should consult the book by J. Humphreys.

Theorem 3.6.1. *Let $\sigma : \mathbf{g} \to \mathbf{g}$ be an automorphism of the semisimple Lie algebra $\mathbf{g}$ with $\mathbf{h}' = \sigma(\mathbf{h})$, where $\mathbf{h}, \mathbf{h}'$ is a pair of Cartan subalgebras. Then the root systems Φ, Φ' determined by the Cartan subalgebras $\mathbf{h}, \mathbf{h}'$ are isomorphic.*

Proof. Define $\phi : \Phi \to \Phi'$ by $\phi(\alpha)(h) = \alpha(\sigma^{-1}(h))$ for $\alpha \in \Phi$ and $h \in \mathbf{h}'$. Choose $0 \neq e_\alpha \in \mathbf{g}_\alpha$. Then

$$\begin{aligned}[h, \sigma(e_\alpha)] &= \sigma([\sigma^{-1}(h), e_\alpha]) = \sigma(\alpha(\sigma^{-1}(h))e_\alpha) \\ &= \alpha(\sigma^{-1}(h))\sigma(e_\alpha) = \phi(\alpha)(h) \cdot \sigma(e_\alpha)\end{aligned}$$

for all $h \in \mathbf{h}'$ so that $\phi(\alpha) \in \Phi'$. We can extend by linearity $\phi : E \to E'$ where E, E' are the real vector spaces where the root systems are sitting. We show that $\phi \circ \sigma_\alpha \circ \phi^{-1} = \sigma_{\phi(\alpha)}$ for $\alpha \in \Phi$, $\sigma_\alpha \in W$:

$$\begin{aligned}\phi\sigma_\alpha\phi^{-1}(\phi(\beta)) &= \phi\sigma_\alpha(\beta) = \phi(\beta - <\beta, \alpha> \alpha) \\ &= \phi(\beta) - <\beta, \alpha> \phi(\alpha) = \phi(\beta) \text{ for } \beta \perp \alpha.\end{aligned}$$

This implies $\phi\sigma_\alpha\phi^{-1}(\gamma) = \gamma$ for all $\gamma \in \phi(P_\alpha)$. Furthermore, $\phi\sigma_\alpha\phi^{-1}(\phi(\alpha)) = \phi\sigma_\alpha(\alpha) = -\phi(\alpha)$ so that $\phi\sigma_\alpha\phi^{-1} = \sigma_{\phi(\alpha)}$. For an arbitrary pair $\alpha, \beta \in \Phi$,

$$\begin{aligned}\sigma_{\phi(\alpha)}(\phi(\beta)) &= \phi(\beta) - <\phi(\beta), \phi(\alpha)> \phi(\alpha) = \phi\sigma_\alpha\phi^{-1}(\phi(\beta)) \\ &= \phi\sigma_\alpha(\beta) = \phi(\beta) - <\beta, \alpha> \phi(\alpha)\end{aligned}$$

so that $<\phi(\beta), \phi(\alpha)> = <\beta, \alpha>$ and so ϕ is an isomorphism of the root systems.

Theorem 3.6.2. *Let* $\mathbf{h}, \mathbf{h}'$ *be a pair of Cartan subalgebras in a semisimple Lie algebra* $\mathbf{g}$. *Then there exists an automorphism* ϕ *of* $\mathbf{g}$ *such that* $\mathbf{h'} = \phi(\mathbf{h})$.

Proof. See J. Humphreys, Sections 16.1 - 16.5

Corollary 3.6.3. *The root system of a semisimple Lie algebra does not depend in an essential way (i.e. modulo isomorphism) on the choice of a Cartan subalgebra.*

Theorem 3.6.4. *The root system* Φ *of a simple Lie algebra* $\mathbf{g}$ *is irreducible.*

Proof. Assume the contrary: $\Phi = \Phi_1 \cup \Phi_2$ where Φ_i are nonempty orthogonal subsystems. Let $\mathbf{k}$ be the subalgebra of $\mathbf{g}$ generated by the root subspaces $\mathbf{g}_\alpha$ for $\alpha \in \Phi_1$. If now $\alpha \in \Phi_1$ and $\beta \in \Phi_2$ then $(\alpha + \beta, \alpha) \neq 0$ and $(\alpha + \beta, \beta) \neq 0$ and therefore $\alpha + \beta \notin \Phi_1$ and $\alpha + \beta \notin \Phi_2$ so that $\alpha + \beta$ is not a root. By Lemma 2.2.5, $[\mathbf{g}_\alpha, \mathbf{g}_\beta] = 0$. Since $\mathbf{g}$ is a direct sum of the root subspaces $\mathbf{g}_\gamma$ and of
$\mathbf{h} = \oplus_\gamma [\mathbf{g}_\gamma, \mathbf{g}_{-\gamma}]$, we reduce that $\mathbf{k}$ is an ideal in $\mathbf{g}$. But this is a contradiction, since there are no nontrivial ideals.

Lemma 3.6.5. *Let* $\mathbf{g} = \mathbf{g}_1 \oplus \cdots \oplus \mathbf{g}_t$ *be a semisimple Lie algebra, where the* $\mathbf{g}_i$*'s are its simple ideals and let* $\mathbf{h}$ *be a Cartan subalgebra of* $\mathbf{g}$. *Then* $\mathbf{h}_i = \mathbf{h} \cap \mathbf{g}_i$ *is a Cartan subalgebra of* $\mathbf{g}_i$ *for each* i.

Proof. First each $\mathbf{h}_i$ is commutative since $\mathbf{h}$ is commutative. We have to show that $N(\mathbf{g}_i, \mathbf{h}_i) = \mathbf{h}_i$. If this is not the case, then for some i_0 the ideal $\mathbf{h}'_{i_0} = N(\mathbf{g}_{i_0}, \mathbf{h}_{i_0})$ would be strictly larger than $\mathbf{h}_{i_0}$. But since $[\mathbf{g}_i, \mathbf{g}_j] = 0$ for $i \neq j$,

$$\mathbf{h} \underset{\neq}{\subset} \mathbf{h}_1 + \ldots \mathbf{h}'_{i_0} + \ldots \mathbf{h}_t \subset N(\mathbf{g}, \mathbf{h})$$

and so $N(\mathbf{g}, \mathbf{h}) \neq \mathbf{h}$, which is a contradiction, since $\mathbf{h}$ is a Cartan subalgebra of $\mathbf{g}$.

Theorem 3.6.6. *Let* $\mathbf{g} = \mathbf{g}_1 + \cdots + \mathbf{g}_t$ *be a semisimple Lie algebra composed of the simple ideals* $\mathbf{g}_i$. *Then the root system* Φ *of* $\mathbf{g}$ *decomposes to a union* $\Phi = \Phi_1 \cup \ldots \Phi_t$ *of mutually orthogonal irreducible subsystems, where* Φ_i *is a root system of* $\mathbf{g}_i$.

Proof. By the results above, it suffices to show that 1) each root of $\mathbf{g}$ belongs to some Φ_i, and 2) the subsystems are mutually orthogonal.

Now let $\alpha \in \Phi$. Since $[\mathbf{g}_i, \mathbf{g}_j] = 0$ for $i \neq j$ the ad_h eigenvectors must lie in the subspaces $\mathbf{g}_i$. So we have $\mathbf{g}_\alpha \subset \mathbf{g}_i$ for some i. But then $\mathbf{g}_\alpha$ is a root subspace of $(\mathbf{g}_i, \mathbf{h}_i)$ since $\mathbf{h}_i = \mathbf{g}_i \cap \mathbf{h}$. Thus $\alpha \in \Phi_i$.

Let next $\alpha \in \Phi_i$ and $\beta \in \Phi_j$ with $i \neq j$. Then $\mathbf{h}_\beta \in \mathbf{g}_j$ and $e_\alpha \in \mathbf{g}_i$, so

$$0 = [h_\beta, e_\alpha] = \alpha(h_\beta)e_\alpha \text{ and } \alpha(h_\beta) = 0$$

which implies $(\alpha, \beta) = \alpha(h_\beta) = 0$.

Lemma 3.6.7. *Let Φ be an irreducible root system and $\Delta \subset \Phi$ a system of simple roots. Then there is a unique maximal root $\beta \in \Phi$ with respect to the partial order defined by the set of positive roots Φ^+. If α is any root then the height $h(\alpha) < h(\beta)$ and $(\beta, \gamma) \geq 0$ for any simple root γ. All the coefficients in the decomposition $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$ are strictly positive.*

Proof. Let β be any maximal root (it exists since Φ is a finite set). Clearly $\beta \in \Phi^+$. Define $\Delta_1 \subset \Delta$ as the set of simple roots α such that $k_\alpha = 0$ in the expansion $\beta = \sum k_\alpha \alpha$ and $\Delta_2 = \Delta \setminus \Delta_1$. If $\Delta_1 \neq \emptyset$ then $(\beta, \alpha) \leq 0$ for all $\alpha \in \Delta_1$ by Lemma 3.3.1.

Since Φ is irreducible there are roots $\alpha_1 \in \Delta_1$ and $\alpha_2 \in \Delta_2$ such that $(\alpha_1, \alpha_2) \neq 0$ and thus also $(\beta, \alpha_1) < 0$. By Theorem 3.2.2 $\alpha_1 + \beta$ is a root. But since $\alpha_1 > 0$ we have $\alpha_1 + \beta > \beta$, contradiction. It follows that $\Delta_1 = \emptyset$ and $k_\alpha > 0$ for all $\alpha \in \Delta$.

In the same way we see that $(\beta, \alpha) \geq 0$ for all $\alpha \in \Delta$ (otherwise $\alpha + \beta \in \Phi$ and $\alpha + \beta > \beta$). Let then β' a another maximal root with $\beta' = \sum k'_\alpha \alpha$ as sum over simple roots. Again $(\beta', \alpha) \geq 0$ for all $\alpha \in \Delta$ and since $\beta' \neq 0$ we must have $(\beta', \alpha) > 0$ at least for one simple root α. Then

$$(\beta', \beta) = \sum k_\alpha (\beta', \alpha) > 0$$

and so $\beta - \beta' \in \Phi$ or $\beta = \beta'$. The former is absurd since then $\beta > \beta'$ or $\beta' > \beta$ by Theorem 3.2.2. So we must have $\beta' = \beta$.

Exercise 3.6.8 Let $\mathbf{g}$ be semisimple, $\mathbf{h} \subset \mathbf{g}$ a Cartan subalgebra, and $\Delta \subset \Phi$ a set of simple roots for $(\mathbf{g}, \mathbf{h})$. Choose $0 \neq x_\alpha \in \mathbf{g}_\alpha$ and $0 \neq y_\alpha \in \mathbf{g}_{-\alpha}$ for all $\alpha \in \Delta$. Show that the vectors x_α, y_α generate the Lie algebra $\mathbf{g}$, that is, any element in $\mathbf{g}$ is obtained by taking linear combinations of multiple commutators of these elements. Hint: Use repeatedly Theorem 2.3.11.

Theorem 3.6.9. *If the root systems Φ, Φ' of a pair of semisimple Lie algebras $\mathbf{g}, \mathbf{g}'$ are isomorphic then $\mathbf{g}$ is isomorphic with $\mathbf{g}'$.*

Proof. By assumption, there is a linear map $\phi : E \to E'$ such that $\phi(\Phi) = \Phi'$ and ϕ preserves the brackets $< \cdot, \cdot >$. Let $\Delta \subset \Phi$ be a system of simple roots. Then $\Delta' = \phi(\Delta)$ is a system of simple roots in Φ'. We fix $x_\alpha \in \mathbf{g}_\alpha$ and $y_\alpha \in \mathbf{g}_{-\alpha}$ for all $\alpha \in \Delta$ such that $h_\alpha = [x_\alpha, y_\alpha]$, Cor. 2.3.9. We fix likewise the elements $x'_{\alpha'}$ and $y'_{\alpha'}$ in $\mathbf{g}'$; we have denoted $\alpha' = \phi(\alpha)$.

We assume first that $\mathbf{g}, \mathbf{g}'$ are simple Lie algebras. We denote by $\mathbf{k}$ the subalgebra of $\mathbf{g} \oplus \mathbf{g}'$ generated by the vectors $\hat{x}_\alpha = x_\alpha \oplus x'_{\alpha'}$ and $\hat{y}_\alpha = y_\alpha \oplus y'_{\alpha'}$.

(1) We claim that $\mathbf{k} \neq \mathbf{g} \oplus \mathbf{g}'$. Since $\mathbf{g}, \mathbf{g}'$ were assumed to be simple, the root systems are irreducible. Let β, β' the maximal roots in Φ, Φ'. The map ϕ preserves the partial ordering since if $\gamma \in \Phi$ is a positive root then $\gamma = \sum_{\alpha \in \Delta} k_\alpha \alpha$ with nonnegative coefficients and so $\phi(\gamma)$ is a linear combination of the simple roots $\phi(\alpha) \in \Delta'$ with nonnegative coefficients and $\phi(\gamma)$ is positive. It follows that $\beta' = \phi(\beta)$.

Let $0 \neq x \in \mathbf{g}_\beta, 0 \neq x' \in \mathbf{g}'_{\beta'}$. Set $\hat{x} = x \oplus x'$. Let $V \subset \mathbf{g} \oplus \mathbf{g}'$ be the subspace generated from the vector $\hat{x}$ by repeated adjoint action by the elements in $\mathbf{k}$. Now $[x_\gamma, x] = 0 = [x'_{\gamma'}, x']$ for any positive roots γ, γ' by the maximality of the roots β, β'. It follows that it suffices to take commutators of $\hat{x}$ with the elements $\hat{y}_\alpha$ in order to generate the whole space V. This means that any vector in V is a linear combination of vectors

$$\mathrm{ad}_{\hat{y}_{\alpha_1}} \dots \mathrm{ad}_{\hat{y}_{\alpha_n}}(\hat{x}).$$

By the inspection of the weights of these vectors we conclude that the intersection $V \cap (\mathbf{g}_\beta \oplus \mathbf{g}'_{\beta'})$ consists only of the vector $\hat{x}$. On the other hand, $\dim(\mathbf{g}_\beta \oplus \mathbf{g}'_{\beta'}) = 2$ so that $V \neq \mathbf{g} \oplus \mathbf{g}'$.

We wanted to prove that $\mathbf{k} \neq \mathbf{g} \oplus \mathbf{g}'$. If this is not true then V is a nonzero ideal in $\mathbf{g} \oplus \mathbf{g}'$ and thus either $V = \mathbf{g}$ or $V = \mathbf{g}'$ which is absurd since $\hat{x} \in V$ but $\hat{x}$ is not an element of $\mathbf{g}$ or of $\mathbf{g}'$.

(2) Let $\pi : \mathbf{k} \to \mathbf{g}$ and $\pi' : \mathbf{k} \to \mathbf{g}'$ be the projections. Clearly both π, π' are homomorphisms of Lie algebras. By the exercise 3.6.8 these maps are surjective. We claim that they are also injective. If for example π' is not injective then there is an element $\hat{z} = z \oplus 0 \in \mathbf{k}$ with $z \neq 0$. Let $I \subset \mathbf{g}$ be the ideal generated by z, that is, the space of linear combinations of vectors obtained by taking multiple commutators of z with the vectors x_α, y_α. But $\mathbf{g}$ is simple, so this ideal must be

equal to $\mathbf{g}$. By the definition of $\mathbf{k}$ we have then $\mathbf{g} \subset \mathbf{k}$. This implies that the vector $0 \oplus x'$ is in $\mathbf{k}$. Again, the ideal in $\mathbf{g}'$ generated by x' must be all of $\mathbf{g}'$ and so $\mathbf{g}' \subset \mathbf{k}$. Now $\mathbf{g} \oplus \mathbf{g}' \subset \mathbf{k}$, which is in contradiction what we have shown in (1). It follows that the maps $\pi : \mathbf{k} \to \mathbf{g}$ and $\pi' : \mathbf{k} \to \mathbf{g}'$ are both isomorphisms and therefore the algebras $\mathbf{g}, \mathbf{g}'$ are isomorphic.

Note that in this isomorphism x_α is mapped to x'_α, via the element $\hat{x}_\alpha \in \mathbf{k}$, for each simple root $\alpha \in \Delta$. Likewise for the elements y_α and therefore also for h_α's.

(3) Consider finally the general case when $\mathbf{g}, \mathbf{g}'$ are not necessarily simple. If $\Phi = \Phi_1 \cup \dots \Phi_t$ is a decomposition of Φ to mutually orthogonal irreducible subsystems then $\Phi' = \Phi'_1 \dots \Phi'_t$ is a similar decomposition for Φ' with $\Phi'_i = \phi(\Phi_i)$, since ϕ is an isomorphism of root systems. Now Φ'_i is isomorphic to Φ_i. Denoting by $\mathbf{g}_i$ the subalgebra of $\mathbf{g}$ corresponding to the subsystem Φ_i and by $\mathbf{g}'_i \subset \mathbf{g}'$ the subalgebra corresponding to Φ'_i, we have by the previous results that $\mathbf{g}_i$ is isomorphic with $\mathbf{g}'_i$. The subalgebras $\mathbf{g}_i$ are simple ideals and

$$\mathbf{g} = \mathbf{g}_1 \oplus \cdots \oplus \mathbf{g}_t \text{ and } \mathbf{g}' = \mathbf{g}'_1 \oplus \cdots \oplus \mathbf{g}'_t.$$

It follows that $\mathbf{g}$ is isomorphic with $\mathbf{g}'$.

Exercise 3.6.10 Determine the maximal roots in each of the cases A_ℓ, B_ℓ, C_ℓ and D_ℓ.

Exercise 3.6.11 In a similar way as in the Exercise 3.5.13 we define a Borel subalgebra $\mathbf{b}'$ as

$$\mathbf{b}' = \mathbf{h} \bigoplus_{\alpha \in \Phi^-} \mathbf{g}_\alpha.$$

Show that the Borel subalgebras $\mathbf{b}$ and $\mathbf{b}'$ are isomorphic. Hint: Use Exercise 3.5.14.

Exercise 3.6.12 Let $\{\alpha_1, \dots, \alpha_\ell\}$ be a system of simple roots for a semisimple Lie algebra $\mathbf{g}$. Let $x_i \in \mathbf{g}_{\alpha_i}$ and $y_i \in \mathbf{g}_{-\alpha_i}$. Show that

$$(\mathrm{ad}_{x_i})^{-<\alpha_j,\alpha_i>+1}(x_j) = 0 = (\mathrm{ad}_{y_i})^{-<\alpha_j,\alpha_i>+1}(y_j)$$

when $i \neq j$. Hint: Use the Theorem on lengths of root chains.

Exercise 3.6.13 We know that the map $\alpha \to -\alpha$ is an automorphism of a root system Φ of a semisimple Lie algebra $\mathbf{g}$. Describe explicitly, in terms of basis in root subspaces, the corresponding automorphism of the Lie algebra $\mathbf{g}$.

CHAPTER 4: REPRESENTATION THEORY

4.1 The universal enveloping algebra

In this section we define an associative algebra $U(\mathbf{g})$ for any Lie algebra $\mathbf{g}$ which will be an important tool for constructing representations of $\mathbf{g}$.

First, for any set S we define *a free associative algebra* $\mathbb{F}(S)$ over a field $\mathbb{F}$, generated by S. As a vector space, $\mathbb{F}(S)$ is the space of formal linear combinations of *words* $a_1 a_2 \dots a_n$ where the a_i's are any (not necessarily different) elements in the set S. This means simply that $\mathbb{F}(S)$ is an infinite-dimensional vector space over $\mathbb{F}$ with a basis labelled by the words $a_1 \dots a_n$.

Next we define a product ab of words $a = a_1 \dots a_n$ and $b = b_1 \dots b_m$ by writing the words after each other,

$$ab = a_1 \dots a_n b_1 \dots b_m.$$

We extend this product by linearity to a pair of arbitrary vectors in $\mathbb{F}(S)$.

It is clear that the product is associative, by definition the standard distributive laws hold, so that indeed $\mathbb{F}(S)$ becomes an associative algebra over the field $\mathbb{F}$.

An empty word (no letters) is denoted by 1. This becomes the neutral element for multiplication, $1a = a1$ for all $a \in \mathbb{F}(S)$.

Remark When the set S consists of a single element x then the algebra is simply the commutative polynomial algebra in one variable x : the words are $x^n = xx \dots x$ (n times) and a general element in the algebra is $\sum_{i=0}^n \alpha_i x^i$ with $\alpha_i \in \mathbb{F}$. In general however $\mathbb{F}(S)$ is noncommutative, xy and yx are different words.

We can also define the *commutative free associative algebra* generated by the set S by declaring that the order of the letters does not matter. For a finite set $S = \{x_1, \dots x_n\}$ this is the polynomial algebra in n variables x_i. The general element in this algebra is a linear combination of the basic monomials $x_1^{k_1} \dots x_n^{k_n}$.

Let next V be a vector space over $\mathbb{F}$. We define a new associative algebra $\mathbb{F}[V]$. This algebra is defined as $\mathbb{F}(V)$ but now we identify *a formal linear combination* $\alpha \cdot a + \beta \cdot b$ of one letter words $a, b \in V$ as the one letter word c, where $c = \alpha \cdot a + \beta \cdot b$ is the *linear combination in the vector space* V. Likewise, the products $w = a_1 \dots a_{i-1}(\alpha \cdot a + \beta \cdot b)a_{i+1} \dots a_n$ will be identified as vectors $w = \alpha \cdot a_1 \dots a_{i-1} a a_{i+1} \dots a_n + \beta \cdot a_1 \dots a_{i-1} b a_{i+1} \dots a_n$.

If $x_1 \dots x_p$ is a basis of V then a general element in $\mathbb{F}[V]$ is a linear combination of the words in the alphabet $x_1, \dots, x_p$. This means that actually the algebra $\mathbb{F}[V]$ is isomorphic to the free associative algebra $\mathbb{F}(S)$ where $S = \{x_1, \dots, x_p\}$.

When $V = \mathbf{g}$ is a Lie algebra we can perform a further reduction of the algebra $\mathbb{F}[\mathbf{g}]$. We want that the structure of the *universal enveloping algebra* $U(\mathbf{g})$ reflects the commutator structure of $\mathbf{g}$. Let $I \subset \mathbb{F}[\mathbf{g}]$ be the smallest two sided ideal containing all the elements $xy - yx - [x, y]$ for $x, y \in \mathbf{g}$. Note that these are linear combinations of words of length 1 and 2. If $e_1, \dots, e_n$ are basis vectors in $\mathbf{g}$,

$$[e_i, e_j] = \sum_k c_{ij}^k e_k,$$

then we can write

$$xy - yx - [x, y] = \sum x_i y_j (e_i e_j - e_j e_i) - \sum x_i y_j c_{ij}^k e_k$$

as elements in $\mathbb{F}[\mathbf{g}]$. Note that in the free algebra $\mathbb{F}[\mathbf{g}]$ the elements $e_i e_j - e_j e_i$ are completely independent of the Lie algebra commutators $[e_i, e_j] = c_{ij}^k e_k$.

By definition, the ideal I consists of all linear combinations of elements $u(xy - yx - [x, y])v$, where u, v are arbitrary elements in the algebra $\mathbb{F}[\mathbf{g}]$.

If $\mathcal{A}$ is any associative algebra and $I \subset \mathcal{A}$ is a two sided ideal then one can construct a new associative algebra $\mathcal{A}/I$, which, as a vector space, is just the quotient of two vector spaces. The product in $\mathcal{A}/I$ is defined through representatives of equivalence classes,

$$(u + I)(v + I) \equiv uv + I.$$

Exercise 4.1.1 Show that the product is well-defined (it does not depend on the choice of representatives of the classes) and defines an associative algebra.

In the case of $I \subset \mathbb{F}[\mathbf{g}]$ above, we define $U(\mathbf{g}) = \mathbb{F}[\mathbf{g}]/I$. This is the universal enveloping algebra of $\mathbf{g}$.

The universality properties refers to the following property:

Theorem 4.1.2. *If $\psi : \mathbf{g} \to \mathcal{A}$ is a homomorphism to an associative algebra $\mathcal{A}$, that is, ψ is linear map with the property $\psi([x, y]) = \psi(x)\psi(y) - \psi(y)\psi(x)$, then there exists a unique homomorphims $\phi : U(\mathbf{g}) \to \mathcal{A}$ of associative algebras such that*

$\psi = \phi \circ j$ where $j : \mathbf{g} \to U(\mathbf{g})$ is the canonical map which sends $x \in \mathbf{g}$ to the one letter word x in $U(\mathbf{g})$. The universal enveloping algebra is uniquely defined (up to isomorphism) by this property.

Proof. First the uniqueness. Let U' be another algebra with the above property, with $j' : \mathbf{g} \to U'$. Then there is a homomorphism $\phi : U(\mathbf{g}) \to U'$ such that $j' = \phi \circ j$. On the other hand, we have a homomorphism $\phi' : U' \to U(\mathbf{g})$ by the universality of U', such that $j = \phi' \circ j'$. Combining, we get a homomorphism $\theta = \phi \circ \phi' : U(\mathbf{g}) \to U(\mathbf{g})$ such that $j' = \theta \circ j'$. But also the identity map $U(\mathbf{g}) \to U(\mathbf{g})$ has this property. By the uniqueness of θ we must have $\theta = id$ and thus $\phi : U(\mathbf{g}) \to U'$ is an isomorphism.

Let then $\psi : \mathbf{g} \to \mathcal{A}$ be a homomorphism. We can define $\phi : U(\mathbf{g}) \to \mathcal{A}$ by setting $\phi(x) = \psi(x)$ for any one letter word x and $\psi(1) = 1$. The one letter words generate the whole algebra $U(\mathbf{g})$ and therefore ϕ extends by linearity to the whole algebra $U(\mathbf{g})$. It clearly has the required property, including uniqueness.

Corollary 4.1.3. *Let $\psi : \mathbf{g} \to End V$ be a representation of the Lie algebra $\mathbf{g}$ in a vector space V. Then there exists a unique representation $\phi : U(\mathbf{g}) \to End V$ such that $\psi = \phi \circ j$.*

Conversely, a representation of $U(\mathbf{g})$ gives by restriction to the one letter words a representation of the Lie algebra $\mathbf{g}$. Thus there is a one-to-one correspondence between representations of $\mathbf{g}$ and its universal enveloping algebra $U(\mathbf{g})$.

Theorem 4.1.4. *(Poincare-Birkhoff-Witt) Let $x_1, x_2, \dots, x_n$ be a basis of the Lie algebra $\mathbf{g}$. Then the ordered words $x_{i_1} x_{i_2} \dots x_{i_p}$ form a basis of $U(\mathbf{g})$, where $i_1 \leq i_1 \leq i_2 \cdots \leq i_p$. (We refer for a proof to J.Humphreys, Section 17.3.)*

Exercise 4.1.5 Prove the PBW theorem when $\mathbf{g}$ is a commutative Lie algebra.

Exercise 4.1.6 Let ψ be the representation of a finite-dimensional Lie algebra $\mathbf{g}$ in the vector space $U(\mathbf{g})$ defined by $\psi(x)u = xu - ux$. Although $U(\mathbf{g})$ itself is infinite-dimensional, show that any element $u \in U(\mathbf{g})$ lies in some finite-dimensional subspace $V \subset U(\mathbf{g})$ which is invariant under the representation ψ.

Exercise 4.1.7 Let $\mathbf{g} = \mathbf{sl}(2, \mathbb{C})$, $\lambda \in \mathbb{C}$ and let I_λ be the smallest left ideal in $U(\mathbf{g})$ containing the elements x and $h - \lambda \cdot 1$. Here x, y, h are the standard basis vectors of $\mathbf{g}$. Show that a basis of the vector space $U(\mathbf{g})/I_\lambda$ is given by the monomials y^n with $n = 0, 1, 2, \ldots$.

4.2 Representations of $\mathbf{sl}(2,\mathbb{F})$

We denote again the vectors in the standard basis of $\mathbf{g} = \mathbf{sl}(2,\mathbb{F})$ as $x = e_{12}, y = e_{21}, h = e_{11} - e_{22}$. The field $\mathbb{F}$ is of characteristic zero and algebraically closed. The standard Cartan subalgebra $\mathbf{h}$ is spanned by the vector h. We have the commutation relations

$$[h,x] = 2x \qquad [h,y] = -2y \qquad [x,y] = h.$$

First some new terminology. We have defined a representation of a Lie algebra $\mathbf{g}$ as a homomorphism $\rho : \mathbf{g} \to \operatorname{End} V$, where V is a vector space. Thus the action of an element $x \in \mathbf{g}$ to a vector $v \in V$ is written as $\rho(x)v$. We often drop the symbol ρ and write simply $\rho(x)v = xv$. That is, we have a multiplication $\mathbf{g} \times V \to V, (x,v) \mapsto xv$. The multiplication satisfies, besides being linear in both arguments, $[x,y]v = x(yv) - y(xv)$. In general, a vector space V together with this kind of multiplication $\mathbf{g} \times V \to V$ is called *a* $\mathbf{g}$ *module.* Thus a representation of $\mathbf{g}$ defines a $\mathbf{g}$ module and vice versa.

By the Corollary 4.1.3 any $\mathbf{g}$ module defines in a natural way a $U(\mathbf{g})$ module and any $U(\mathbf{g})$ module gives a $\mathbf{g}$ module by restriction to $\mathbf{g} \subset U(\mathbf{g})$.

A $\mathbf{g}$ module V is irreducible if there are no nontrivial $\mathbf{g}$ invariant subspaces $W \subset V$.

Assume next that V is an irreducible finite-dimensional nonzero $\mathbf{g}$ module. If $0 \neq v \in V$ then $U(\mathbf{g})v \subset V$ is clearly a $\mathbf{g}$ invariant subspace and therefore $U(\mathbf{g})v = V$.

Since V is finite-dimensional, the element h has at least one eigenvector in V. For the same reason there must be an eigenvector v_0 with maximal real part of the eigenvalue λ. Since

$$h(xv_0) = x(hv_0) + [h,x]v_0 = (\lambda + 2)v_0$$

we must have $xv_0 = 0$ by the maximality of the eigenvalue λ. By the Poincare-Birkhoff-Witt theorem a basis in $U(\mathbf{g})$ is given by the elements $y^p h^q x^r$ with $p,q,r = 0,1,2,\ldots$. But since $xv_0 = 0$ and $hv_0 = \lambda v_0$ we observe that

$$U(\mathbf{g})v_0 = \{\sum_p a_p y^p v_0 | a_p \in \mathbb{F}\}.$$

But V was assumed to be irreducible so we conclude that V is spanned by the vectors $y^p v_0$.

We denote $v_i = \frac{1}{i!} y^i v_0$.

Lemma 4.2.1.

(1) $hv_i = (\lambda - 2i)v_i$

(2) $yv_i = (i+1)v_{i+1}$

(3) $xv_i = (\lambda - i + 1)v_{i-1}$.

Proof. (1) The case $i = 0$ is clear. Induction on i :

$$\begin{aligned} hv_{i+1} &= (i+1)^{-1}h(yv_i) = (i+1)^{-1}(yhv_i + [h,y]v_i) \\ &= (i+1)^{-1}((\lambda - 2i)yv_i - 2yv_i) = (\lambda - 2(i+1))v_{i+1}. \end{aligned}$$

(2) This follows directly from the definition of v_i.

(3) The case $i = 0$ is clear. Induction on i :

$$\begin{aligned} xv_{i+1} &= (i+1)^{-1}xyv_i = (i+1)^{-1}(y(\lambda - i + 1)v_{i-1} + hv_i) \\ &= (i+1)^{-1}(i(\lambda - i + 1)v_i + (\lambda - 2i)v_i) = (\lambda - (i+1) + 1)v_i. \end{aligned}$$

Since by 4.2.1 (1) the vectors v_i for different values of i are linearly independent provided that they are not equal to zero, we must have $v_i = 0$ (by the finite-dimensionality of V) for $i > m$ for some m; choose this integer m to be the smallest possible. Then $v_i \neq 0$ for $i = 0, 1, \ldots, m$. It follws that the set $\{v_0, v_1, \ldots, v_m\}$ is a basis in V. Now we have, by Lemma 4.2.1,

$$0 = xv_{m+1} = (\lambda - m)v_m$$

Since $v_m \neq 0$ it follows that $\lambda - m = 0$. Thus the *the maximal eigenvalue* λ *of* h *is a nonnegative integer.*

Since the Cartan subalgebra is here one-dimensional, the weight subspaces V_μ of V are simply the eigenspaces of h. We have seen that the eigenvalues of μ of h are give as $\mu = \lambda - 2i$ with $i = 0, 1, 2, \ldots, m$ so that $\mu = -m, -m+2, \ldots, m$.

Theorem 4.2.2. *Let* V *be an irreducible nonzero* $\mathbf{sl}(2, \mathbb{F})$ *module. Then*

(1) *There is a unique (up to a multiplicative constant) maximal vector* v_0 *with the highest eigenvalue* $\lambda = 0, 1, 2 \ldots$ *of* h

(2) *V is a direct sum of one-dimensional weight subspaces V_μ with $\mu = -\lambda, -\lambda + 2, \dots, \lambda$*

(3) *There is a basis with $v_i \in V_{\lambda - 2i}$ such that the action of the elements x, y, h is given as in Lemma 4.2.1.*

Proof. First we observe that the maximal weight $\mu = \lambda$ determines the $\mathbf{sl}(2, \mathbb{F})$ module V up to an isomorphism. Given two different irreducible $\mathbf{g}$ modules V, V' with highest weight λ we can construct the isomorphism as the linear map $\phi : V \to V'$ with the property $\phi(v_i) = v_i'$, where the basis $\{v_i'\} \subset V'$ is chosen in a similar way as $\{v_i\} \subset V$.

The existence of the modules follows from a direct construction: Define $V = \mathbb{F}^{\lambda+1}$ an denote the basis vectors in the standard basis by $v_0, v_1, \dots v_\lambda$. Define the action of x, y, h using Lemma 4.2.1 and check by direct computation that the commutation relations of $\mathbf{g}$ hold.

4.3 The theorem of Weyl

Let $\mathbf{g}$ be any semisimple Lie algebra and choose a basis $x_1, \dots, x_n$ in $\mathbf{g}$. Let $\beta : \mathbf{g} \times \mathbf{g} \to \mathbb{F}$ be any symmetric nondegenerate bilinear form such that

$$\beta([x, y], z) = -\beta(y, [x, z]) \text{ for all } x, y, z \in \mathbf{g}.$$

We know that at least the Killing form satisfies this condition, and that if $\mathbf{g}$ is simple then any such a bilinear form is proportional to the Killing form.

Since β is nondegenerate, the determinant of the matrix $\beta_{ij} = \beta(x_i, x_j)$ is nonzero and the system of linear equations

$$\beta(y_j, x_i) = \delta_{ij} \text{ with } i = 1, 2, \dots, n$$

has a unique solution y_j for each index j. That is, the basis $x_1, \dots, x_n$ has a unique *dual basis* $y_1, \dots, y_n$.

Exercise 4.3.1 Let $\phi : \mathbf{g} \to \operatorname{End} V$ be a faithful representation of the semisimple Lie algebra $\mathbf{g}$ in a vector space V (that is, ϕ is injective). Then the symmetric bilinear form

$$\beta(x, y) = \operatorname{tr}(\phi(x)\phi(y))$$

is nondegenerate. Prove this!

In the situation of Exercise 4.3.1 we define an element $c_\phi \in \operatorname{End} V$ by

$$c_\phi = \sum_i \phi(x_i)\phi(y_i).$$

This endomorphism is not zero:

$$\operatorname{tr} c_\phi = \sum_i \operatorname{tr}\left(\phi(x_i)\phi(y_i)\right) = \sum_i (x_i, y_i) = \dim \mathbf{g} > 0.$$

Theorem 4.3.2. *c_ϕ commutes with every $\phi(x)$ and thus c_ϕ is equal to $\lambda \cdot \mathbf{1}$ in an irreducible representation (Schur's lemma), where $\lambda = \dim \mathbf{g} / \dim V$.*

Proof. Let $x \in \mathbf{g}$. We can write

$$[x, x_i] = \sum_j a_{ij} x_j \text{ and } [x, y_i] = \sum_j b_{ij} y_j.$$

We have

$$a_{ik} = \beta([x, x_i], y_k) = -\beta(x_i, [x, x_k]) = -\beta(x_i, \sum_j b_{kj} y_j) = -b_{ki}.$$

Using this and the identity $[A, BC] = [A, B]C + B[A, C]$ for matrices we get

$$\begin{aligned}
[\phi(x), c_\phi] &= \sum_i [\phi(x), \phi(x_i)\phi(y_i)] = \sum_i [\phi(x), \phi(x_i)]\phi(x_i) + \sum_i \phi(x_i)[\phi(x), \phi(y_i)] \\
&= \sum_i \phi([x, x_i])\phi(y_i) + \sum_i \phi(x_i)\phi([x, y_i]) \\
&= \sum_{ij} a_{ij}\phi(x_i)\phi(y_j) + \sum_{ij} b_{ij}\phi(x_i)\phi(y_j) = 0.
\end{aligned}$$

The endomorphism c_ϕ is called the *Casimir element* of the representation.

We can also define the (universal) Casimir element as a vector in the universal enveloping algebra $U(\mathbf{g})$ by setting $c = \sum_i x_i y_i$ where the dual basis $\{y_i\}$ is defined with respect to the Killing form, (y_i, x_i). One can then repeat the computation above and show that c commutes with very $x \in \mathbf{g}$ and therefore c commutes with every element in the enveloping algebra $U(\mathbf{g})$.

Any $\mathbf{g}$ module V defines the dual $\mathbf{g}$ module V^* module: As a vector space V^* is the space of linear functions $f : V \to \mathbb{C}$. The action of $x \in \mathbf{g}$ in V^* is given by

$$(x \cdot f)(v) = -f(x \cdot v).$$

This is really a $\mathbf{g}$ action:

$$([x,y]\cdot f)(v) = -f([x,y]v) = -f(x(yv)-y(xv))$$
$$= (x\cdot f)(yv) - (y\cdot f)(xv) = (-y(xf))(vv) + (x(yf))(v)$$

so that $[x,y]f = x(yf) - y(xf)$.

For a submodule $W \subset V$ the quotient module V/W is defined as usual: The action of $x \in \mathbf{g}$ on a vector $[v] = v + W$ is defined as $x[v] = [xv]$.

Lemma 4.3.3. *Let $\phi : \mathbf{g} \to End V$ be a representation of a semisimple Lie algebra $\mathbf{g}$. Then each endomorphism $\phi(x)$ is traceless.*

Proof. Since $\mathbf{g} = [\mathbf{g},\mathbf{g}]$, any $x \in \mathbf{g}$ is a linear combination of elmenents of the type $[y,z]$. But $\phi([y,z]) = \phi(y)\phi(z) - \phi(z)\phi(y)$ and has therefore vanishing trace.

A representation (or a $\mathbf{g}$ module) is *completely reducible* if it is a direct sum of irreducible representations (modules).

Theorem 4.3.4. *Any finite-dimensional representation of a semisimple Lie algebra is completely reducible.*

Proof. Let $\mathbf{g}$ be semisimple and $\phi : \mathbf{g} \to \mathrm{End}\, V$ a finite-dimensional representation.

(1) We assume first that there is a submodule $W \subset V$ of codimension $= 1$. We prove by induction on the dimension $p = \dim\, W$ that there is a complementary one-dimensional invariant subspace $X \subset V$. The case $p = 0$ is clear. Induction $p \mapsto p+1:$

(1a) If W is reducible then choose an invariant submodule $W' \subset W$ with $W' \neq 0, W$. Then $W/W' \subset V/W'$ is a submodule with $\dim(W/W') = \dim(V/W') - 1$. We may apply the induction hypothesis to W/W' to reduce that there is a complementary invariant submodule $W''/W' \subset V/W'$ of dimension one, $V/W' = W''/W' \oplus W/W'$. In the same way there is an invariant submodule $X \subset W''$ of dimension one such that $W'' = X \oplus W'$. Now $W \cap W'' \subset W'$ and therefore $W \cap X = 0$. Since $\dim\, X + \dim\, W = \dim\, V$ we get $V = W \oplus X$.

(1b) Let W be irreducible. Let c_ϕ be the Casimir element of the representation ϕ. Since $W \subset V$ is an invariant subspace we may view c_ϕ as a linear map $V/W \to V/W$. Since $\dim(V/W) = 1$ is a linear map in this space equal to its trace and by Lemma 4.3.3 $\phi(x) = 0$ in the quotient space V/W. But

$$\mathrm{tr}_V(c_\phi) = \mathrm{tr}_W(c_\phi) + \mathrm{tr}_{V/W}(c_\phi) = \mathrm{tr}_W(c_\phi)$$

and so $\mathrm{tr}_W(c_\phi) \neq 0$, by 4.3.2. Since W is irreducible we must have $c_\phi|_W = \lambda \times \mathbf{1}$ for some $\lambda \in \mathbb{F}$. Since the trace is nonzero, we have $\lambda \neq 0$ and thus $W \cap \ker c_\phi = 0$.

Since c_ϕ vanishes in V/W we have $c_\phi(V) \subset W$ and so $\ker c_\phi \neq 0$ in V. By a dimension argument we obtain

$$V = \ker c_\phi \oplus W.$$

$\ker c_\phi$ is a submodule of V since the Casimir element commutes with the representation. This completes the induction in the case $\operatorname{codim} W = 1$.

(2) The general case. Let $W \subset V$ be any nontrivial submodule. Let $T = \mathrm{Hom}(V, W)$ be the vector space of linear maps $V \to W$. This is a $\mathbf{g}$ module by setting $(xf)(v) = x(f(v)) - f(xv)$ for $x \in \mathbf{g}, v \in V$. Let $T' \subset T$ be the subspace consisting of linear maps f which are constant in the subspace $W \subset V$. It is clear that $T' \subset T$ is a submodule. Let $T'' \subset T'$ be the submodule consisting of functions f which vanish in W. Let $f : V \to W$ be any linear function such that $f(w) = w$ for all $w \in W$. Then $T' = T'' \oplus \mathbb{F} \cdot f$. On the other hand, xf is also such a linear function for any $x \in \mathbf{g}$. But since the complement of T'' in T' is one-dimensional, we may us step (1) and reduce that we may fix f such that it spans an invariant submodule S.

Since $\dim S = 1$ we have $x \cdot f = 0$ for all $x \in \mathbf{g}$, that is,

$$0 = (xf)(v) - f(xv).$$

This means that $f : V \to W$ is a homomorphism of $\mathbf{g}$ modules. The kernel $\ker f \subset V$ is a submodule and its intersection with W is zero ($f(w) = w$ for all $w \in W$). From this follows that

$$V = W \oplus \ker f.$$

This concludes the proof.

Exercise 4.3. 5 The elements of the Weyl group W determine automorphisms of the root system Φ. These are called the *inner automorphisms.* Show that in the case of A_2 the group of automorphisms is strictly larger than the group $\mathrm{Int}(\Phi)$ of inner automorphisms. Hint: Study the automorphism $\alpha \mapsto -\alpha$.

Exercise 4.3.6 Construct *all automorphims* of the root system A_2 (and thus of the Lie algebra A_2).

4.4 Some group theory and tensor analysis of representations

This section is a digression to group theory. We shall explain some constructions of representations of classical Lie groups without proofs. Because of the relation between Lie algebras and Lie groups explained in the beginning of Chapter I, any representation of a Lie group defines a representation of the corrresponding Lie algebra; the connection is given by the exponential map of matrices.

Tensor analysis provides some very simple constructions of representations. It is somewhat harder to see that we get *all* irreducible representations this way. The reader is recommended to look at the classical text H. Weyl: *Classical Groups and their Invariants and Representations.*

A useful tool in the tensor analysis comes from physics: The use of the algebra of bosonic or fermionic creation and annihilation operators. We shall briefly discuss this method, through examples, in the end of the section. The linear groups $SU(n)$ and $SO(n)$ appear in physics often as symmetries of many particle systems. This could be for example a nucleus exhibiting various kinds of particle interchange and combined rotational symmetries. If the symmetry is exact, that is, the group commutes with the hamiltonian, then one can classify eigenvectors of the hamiltonian belonging to the same eigenvalue using the representation theory of the symmetry group G. Even in the case when the symmetry is only approximate it might still be of advantage to classify the physical states according to representations of G ('supermultiplets').

To see how the symmetry operates on many particle systems let us assume first that G is represented in a vector space V ('single particle space') with basis vectors $v_1 \dots v_n$. A 2-particle system is then described using the tensor product space $V \otimes V$ carrying the tensor product representation of G. Tensors can be split two antisymmetric and symmetric tensors. Writing a general element of $V \otimes V$ as $t = \sum t_{ij} v_i \otimes v_j$ we can split

$$t = a + s, \qquad a_{ij} = \frac{1}{2}(t_{ij} - t_{ji}), s_{ij} = \frac{1}{2}(t_{ij} + t_{ij}),$$

where s is symmetric and a is antisymmetric in the indices.

Writing a group element $g \in G$ as a matrix g_{ij} acting on the coordinates in the v_i basis we observe that in the tensor product representation the G action is

$t'_{ij} = g_{ia}g_{jb}t_{ab}$ (sum over repeated indices) and therefore by linearity

$$a_{ij} = g_{ia}g_{jb}a_{ab}, \qquad s_{ij} = g_{ia}g_{jb}s_{ab},$$

i.e. the antisymmetric and symmetric parts transform separately. We have therefore two subrepresentations, one in the space of antisymmetric tensors and one in the space of symmetric tensors.

In general, the antisymmetric and symmetric parts can be further reduced to irreducible components. There are some exceptions, most notably the case when $G = SU(n)$ or $G = GL(n)$ acting in V through the defining representation. In these cases one can prove that the representations A and S are already irreducible.

One can go on and consider 3-, 4-,...n-particle systems. For example, in quantum mechanics a system of indistinguishable half-integer spin particles (fermions, e.g. electrons) obeys the Pauli exclusion principle: no two particles should be in the same state. Mathematically, this means that the system is described by elements in the completely antisymmetric tensor product space $\Lambda^k V$. Here k is the number of particles. The number of particles cannot exceed the number of one-particle levels n for combinatorial reasons; there are no completely antisymmetric tensors of rank $k > n$. For $k \leq n$ the number of independent antisymmetric tensors is

$$N(k, n) = \frac{n!}{k!(n-k)!}$$

This is the number of ways how one can select k *different* numbers from the sequence $1, 2, \ldots, n$. Each such selection defines a basis vector in $\Lambda^k V$ by

$$(i_1, \ldots, i_k) \mapsto \sum_\sigma \epsilon(\sigma) v_{i_1} \otimes \cdots \otimes v_{i_k}$$

where the sum is over all permutations of k letters and $\epsilon(\sigma) = \pm 1$ depending whether the permutation is a product of even or odd number of transpositions. It is clear that any antisymmetric tensor can be written uniquely as a linear combination of these elementary tensors.

In the case of integral spin particles (bosons) there is no Pauli exclusion principle; instead, the multiparticle wave function should be completely symmetric with respect to the interchange of arguments (Bose statistics). That is, the k particle states should be elements in the completely symmetrized tensor product $S^k V$. A complete basis in S^kV is obtained by symmetrizing the vectors $v_{i_1} \otimes \cdots \otimes v_{i_k}$ with $i_1 \leq i_2 \leq \cdots \leq i_k$. Now $i_1 < i_2 + 1 < i_3 + 2 \cdots < i_k + k - 1$ are different positive integers in the set $1, 2, \ldots, n + k - 1$ and therefore the dimension

$$dim(S^kV) = \frac{(n+k-1)!}{k!(n-1)!}.$$

In situations where not all of the particles are indistinguishable one has to deal with tensors of *mixed symmetry* type. For example, we could consider third rank tensors obtained from arbitrary tensors by an application of the mixed symmetry operator

$$R = (1 - (13))(1 + (12)),$$

where (ij) means the transposition of the i:th and of the j:th index; thus

$$(Rt)_{i_1 i_2 i_3} = t_{i_1 i_2 i_3} + t_{i_2 i_1 i_3} - t_{i_3 i_2 i_1} - t_{i_2 i_3 i_1}.$$

Note that the order of permutations is important. We denote tensors Rt symbolically by the Young diagram

$$\begin{array}{|c|c|} \hline i_1 & i_2 \\ \hline i_3 \\ \cline{1-1} \end{array}$$

The completely symmetric tensors are denoted by $\boxed{i_1}\boxed{i_2}\ldots\boxed{i_k}$ and the completely antisymmetric ones by

$$\begin{array}{|c|} \hline i_1 \\ \hline i_2 \\ \hline \multicolumn{1}{c}{\cdot} \\ \multicolumn{1}{c}{\cdot} \\ \hline i_k \\ \hline \end{array}$$

As another example of tensors of mixed symmetry type consider the Young diagram

$$\begin{array}{|c|c|} \hline i_1 & i_2 \\ \hline i_3 & i_4 \\ \hline \end{array}$$

The corresponding *Young symmetrizer* is $R = QP$ where

$$P = (1 + (12))(1 + (34)) \text{ and } Q = (1 - (13))(1 - (24)).$$

The general principle is the following: To each row in the Young diagram one associates a symmmetrizer in the corresponding tensor indices. Then one forms the product of all row symmetrizers; here the order is unimportant because the different rows do not mix. To each column one associates an antisymmerizer in the indices included in the column. Finally one multiplies by the product of antisymmetrizers from the left. So in the case of the above diagram one has

$$\begin{aligned}(Rt)_{i_1i_2i_3i_4} &= t_{i_1i_2i_3i_4} - t_{i_3i_2i_1i_4} - t_{i_1i_4i_3i_2} + i_{3i_4i_1i_2} \\ &+ t_{i_2i_1i_3i_4} - t_{i_2i_3i_1i_4} - t_{i_4i_1i_3i_2} + t_{i_4i_3i_1i_2} \\ &+ t_{i_1i_2i_4i_3} - t_{i_3i_2i_4i_1} - t_{i_1i_4i_2i_3} + t_{i_3i_4i_2i_1} \\ &+ t_{i_2i_1i_4i_3} - t_{i_2i_3i_4i_1} - t_{i_4i_1i_2i_3} + t_{i_4i_3i_2i_1}\end{aligned}$$

All the permutation operators R commute with the linear group transformations $g \in G$. For this reason a tensor of the type Rt is transformed into a similar tensor Rt'. Thus the space RV^k of tensors of type R carries a representation of the group G. In fact, one can show that in the case of $G = SU(n)$ or $GL(n)$ in the defining representation this is irreducible. Not so in the case of $SO(n)$. The reason is simple: For the orthogonal group there are geometric invariants formed by the partial traces $t_{jji_1i_2\dots}$ of the tensors. For example, all the tensors for which this partial trace vanishes form an invariant subspace (the orthogonal transformations preserve the real euclidean inner product).

The operators R are *idempotents* modulo a normalization factor. This means that $R^2 = n_R \cdot R$ for some integer n_R. Exercise: Prove this in the case of the 3-box Young diagram above. The idempotent property means that (the normalized) symmetrization operators R act as projectors in the space of all tensors, projecting to the various irreducible representations of $SU(n)$ (or $GL(n)$).

Example $G = SU(3)$, defining representation in $V = \mathbb{C}^3$. The Young diagram $\begin{array}{|c|c|}\hline i_1 & i_2 \\ \hline i_3 & \\ \cline{1-1}\end{array}$ gives the adjoint representation. To see this consider the tensor $u = R(e_1 \otimes e_1 \otimes e_2)$, where e_i is the standard basis in $\mathbb{C}^3$. The eigenvalues of diagonal matrices for a tensor product Lie algebra representation add up, so u is an eigenvector of h_1 (here $h_i = e_{ii} - \frac{1}{3} \cdot \mathbf{1}$) with eigenvalue $\frac{2}{3} + \frac{2}{3} - \frac{1}{3} = 1$ and the eigenvalue for h_2 is $-\frac{1}{3} - \frac{1}{3} + \frac{2}{3} = 0$ giving the heighest weight $(1, 0)$ of the adjoint representation of A_2. Furthermore, u is annihilated by e_{12} and e_{23}. For example,

$$e_{12}(e_1 \otimes e_1 \otimes e_2) = e_1 \otimes e_1 \otimes e_1$$

which is mapped to zero by R because of the antisymmetrization Q. Thus $e_{12}u = 0$.

Similarly,

$$e_{23}(e_1 \otimes e_1 \otimes e_2) = 0$$

(since $e_{23}e_1 = 0 = e_{23}e_2$) and therefore also $e_{23}u = 0$. It follows that u is a highest weight vector. Finally, one checks that $R(e_1 \otimes e_1 \otimes e_2) \neq 0$.

Creation and annihilation operator formalism

In the case of completely symmetric wave functions (bosons) there is a simple formalism to describe the many particle states. To each bases vector v_i for one-particle states one associates a **creation operator** a_i^* with the commutation relations

$$[a_i^*, a_j^*] = 0.$$

A **vacuum** (zero particle state) is denoted by $|0>$. Multiparticle states are then obtained as polynomials

$$|k_1, k_2, \dots, k_n >= (a_1^*)^{k_1} \dots (a_n^*)^{k_n} |0>$$

acting on the vacuum; here the k_i's are arbitrary nonnegative integers. The bosonic structure of the indistinguishable particles is encoded in the commutation relations: the order of factors is unimportant and therefore the states $|k_1 \dots k_n >$ can be put to correspond vectors in the completely symmetric tensor product $S^k V$, where $k = k_1 + \dots + k_n$,

$$|k_1 \dots k_n > \mapsto S(v_1 \otimes \dots v_1 \otimes v_2 \otimes \dots v_2 \otimes \dots \otimes v_n \otimes \dots \otimes v_n)$$

where S is the complete symmetrization opeator (sum over all permutations of k factors), the number of v_1's is k_1, ..., the number of v_n's is k_n.

To describe the inner product in the Hilbert space of multiparticle states (called the *bosonic Fock space* $\mathcal{F}$) it is convenient to introduce also the **annihilation operators** a_i with the commutation relations

$$[a_i, a_j] = 0, \text{ but } [a_i, a_j^*] = \delta_{ij}.$$

The inner product is now fixed uniquely by the requirement that 1) the annihilation operator a_i is the adjoint of a_i^*, 2) the vacuum is annihilated by all annihilation operators, $a_i|0>=0$, and 3) the normalization $<0|0>=1$. For example,

$$<1,1|1,1>=<0|(a_1^*a_2^*)^*(a_1^*a_2^*)|0>=<0|a_2a_1a_1^*a_2^*|0>$$
$$=<0|a_2[a_1,a_1^*]a_2^*|0>=<0|a_2a_2^*|0>=<0|[a_2,a_2^*]|0>=<0|0>=1$$

We define the operators

$$e_{ij}=a_i^*a_j.$$

It is easy to check the commutation relations

$$[e_{ij},e_{kl}]=\delta_{jk}e_{il}-\delta_{il}e_{kj}.$$

We have thus constructed the Lie algebra of the general linear group $GL(n,\mathbb{C})$ acting in the bosonic Fock space. This representation is reducible. Define the *particle number operator*

$$N=\sum_i a_i^*a_i.$$

This commutes with all the operators e_{ij} and it follows that the different eigenspaces of N are invariant under the Lie algebra $\mathbf{gl}(n)$. This corresponds to the fact that the Fock space consists of completely symmetric tensors of arbitrary rank; the symmetric tensors of fixed rank form an irreducible representation space. Let $|m>=(a_1^*)^m|0>$. This vector is of rank m and is annihilated by all e_{ij} with $i<j$. It is also an eigenvector of all elements e_{ii} in the Cartan subalgebra. (For slight technical convenience we have added also the central element $e_{11}+\cdots+e_{nn}$ and consider the Lie algebra $\mathbf{gl}(n)$ instead of the (semi)simple Lie algebra A_{n-1}.) Thus $|m>$ is a highest weight vector corresponding to the weight $\lambda(e_{ii})=m\cdot\delta_{1i}$.

As already noted before, the group $GL(n)$ acts irreducibly in the space of completely symmetric tensors; therefore a complete set of vectors in the subspace $\mathcal{F}_m=\{\psi\in\mathcal{F}|N\psi=m\psi\}$ is obtained by acting with the operators e_{ij} on the highest weight vector $\psi_m\in\mathcal{F}_m$. We can write

$$\mathcal{F}=\mathcal{F}_0\oplus\mathcal{F}_1\oplus\mathcal{F}_2\dots$$

and each $\mathcal{F}_m$ carries an irreducible representation of $GL(n)$.

In order to construct more general representations using the Fock space methods one has to increase the number of independent bosonic oscillator modes. We can prove that all finite-dimensional highest weight representations of $GL(n)$ or $SU(n)$ can be constructed using a set a_{ij}, a^*_{ij} of creation and annihilation operators with $1 \le i, j \le n$, commutation relations

$$[a_{ij}, a^*_{kl}] = \delta_{ik}\delta_{jl},$$

all other commutators being zero. The Lie algebra is constructed as

$$e_{ij} = \sum_k a^*_{ik} a_{jk}.$$

For each sequence $m = (m_1, m_2, \dots, m_n)$ of nonnegative integers we construct the vector

$$\psi(m) = \prod_k (\det(a^*_{ij})_{i,j\le k})^{m_k} |0> .$$

Using the antisymmetry of a determinant as a function of the row vectors we first observe that $e_{ij}\psi(m) = 0$ for all $i < j$. The vector $\psi(m)$ is also an eigenvector of each e_{ii}; e_{ii} acts like a number operator for the oscillator modes with first index equal to i. The determinants are homogenenous functions of order 1 in each of the rows and columns and it follows that the action of e_{ii} on $\psi(m)$ is just a multiplication by the total degree $m_n + m_{n-1} + \cdots + m_i$. Thus we get for the components $\lambda_i = \lambda(e_{ii})$ of the highest weight, $\lambda_i = m_i + m_{i+1} \cdots + m_n$. In particular

$$\lambda_1 \ge \lambda_2 \ge \dots \lambda_n \ge 0$$

and all the components are integers. Conversely, for each such a sequence λ there is a unique set of nonnegative integers m with the above relation to λ.

Working carefully out the normalization factors in the space of roots of A_ℓ one observes that the conditions $< \lambda, \alpha >= 2\frac{(\lambda,\alpha)}{(\alpha,\alpha)} = 0, 1, 2, \dots$ for each simple root $\alpha = \alpha_{i,i+1}$ of A_ℓ are essentially the conditions on the compents λ_i derived above; the only difference is that we have added the number operator N to the Cartan subalgebra, thus discarding the trace zero condition on elements of A_ℓ. We shall prove later in the next section that all the finite-dimensional irreducible representations of semisimple Lie algberas are classified by the *highest weight* . This is the weight of a vector v is the representation space which satisfies $x_\alpha v = 0$ for all root vectors x_α corresponding to positive roots α. The highest weights λ have the characteristic property that $< \lambda, \alpha >$ is a nonnegative integer for all simple roots α.

Therefore, all the finite- dimensional representations of A_{n-1} are generated by the different highest weight vectors $\psi(m)$ in the bosonic Fock space for n^2 independent oscillators. In the Young diagram notation, the representation λ corresponds to the diagram with row lengths $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n$, read from top to bottom.

The completely antisymmetric representations (only one column in the Young diagram) are best constructed using the **fermionic oscillators** b_i^*, b_i, $i = 1, 2, \ldots, n$. The defining relations are described by *anticommutators* $[A, B]_+ = AB + BA$ instead of commutators,

$$[b_i^*, b_j]_+ = \delta_{ij}$$

and all other anticommutators are zero. The Lie algebra of $GL(n)$ is now constructed as

$$e_{ij} = b_i^* b_j.$$

The commutation relations can be checked using the identity

$$[AB, CD] = A[B, C]_+ D - [A, C]_+ BD + CA[B, D]_+ - C[A, D]_+ B.$$

The fermionic Fock space consists of all creation operator polynomials acting on the vacuum $|0>$. As in the bosonic case the vacuum is defined by the relations $b_i|0>= 0$. The vacuum is again normalized, $< 0|0 >= 1$ and b_i^* is supposed to be the adjoint of b_i. These requirements fix the inner product uniquely.

The bosonic Fock space was infinite-dimenional. In the fermionic case the dimension is finite. The reason is that, because of the anticommutation relations, all the powers $(b_i^*)^k$ vanish identically for $k > 1$. The only nonzero vectors in the Fock space are of the type

$$b_{i_1}^* b_{i_2}^* \ldots b_{i_k}^* |0>,$$

where all the indices i_μ are distinct. By the anticommutation relations we can assume that $i_1 > i_2 \cdots > i_k$ (a change in the ordering corresponds just a multiplicative factor ± 1.) Thus the number of independent vectors of length k is $\binom{n}{k}$, which is equal to the number of independent components of a fully antisymmetric tensor of rank k in dimension n. We can again introduce a number operator $N = \sum_k b_k^* b_k$. The eigenvalue of N is now the rank of the antisymmetric tensor, or in other words, the number of boxes in the one-column Young diagram.

Exercise 4.4.1 Define the operators

$$e_{jk} = a_j^* a_k$$

where $a_i a_j - a_j a_i = 0 = a_i^* a_j^* - a_j^* a_i^*$ and $[a_i, a_j^*] = \delta_{ij}$ for $i, j = 1, 2, 3$. Show that these span the Lie algebra A_2 extended by the operator $c = e_{11} + e_{22} + e_{33}$, which commutes with the rest of the operators e_{jk}. Study the representations of A_2 in the Fock representation of the canonical commutation relations. In particular, find the representations of the subalgebra $A_1 \subset A_2$ which are included in a given representation of A_2.

Show that

$$S_{31} = e_{31}(e_{11} - e_{22}) + e_{32}e_{21}, \text{ and } S_{32} = e_{32}$$

are shift operators for the A_1 subalgebra, that is, they take any vector ψ satisfying the conditions $e_{12}\psi = 0$, $(e_{11} - e_{22})\psi = \lambda\psi$ to a vector satisfying the same conditions but with a different eiegenvalue of $h = e_{11} - e_{22}$. Are there other simple shift operators (at most of degree 2 in the generators)? How can one use the shift operators to construct a basis in a representation space?

Exercise 4.4.2 Prove in the case of third rank tensors that any tensor is a sum of components corresponding to the different symmetry types defined by the complete symmetrization and antisymmetrization operators and the Young diagrams

$$\begin{array}{|c|c|}\hline i_1 & i_2 \\ \hline i_3 \\ \cline{1-1} \end{array} \text{ and } \begin{array}{|c|c|}\hline i_1 & i_3 \\ \hline i_2 \\ \cline{1-1} \end{array}$$

Next let the dimension of the underlying vector space V be equal to 3. The rotation group $SO(3)$ acts naturally on tensors in a 3 dimensional space. Determine the values λ of the angular momentum (the highest eigenvalue of h) and their multiplicities occuring in each of the representations corresponding the different symmetry types of the third rank tensors.

Exercise 4.4.3 Analyse the adjoint representation of A_2 in terms of bosonic creation and annihilaton operators a_i, a_i^* $(i = 1, 2, 3)$. It is not possible to construct the adjoint representation with a single set of bosonic operators, but it is possible if you add a new set b_i, b_i^* which commutes with the operators a_i, a_i^*. Find the polynomials in the creation operators which span the 8-dimensional adjoint representation and check that the weights indeed come out correctly, as expected in the adjoint representation.

4.5 Standard cyclic modules

Recall that a cyclic vector in a $\mathbf{g}$ module V is a vector v with the property $U(\mathbf{g})v = V$. The module V is cyclic if it has at least one cyclic vector. Note that a cyclic module does not need to be irreducible.

Let $\mathbf{g}$ be a semisimple Lie algebra and $\mathbf{h} \subset \mathbf{g}$ a Cartan subalgebra. We fix a set Δ of simple roots so that the set of roots splits to positive and negative roots, $\Phi = \Phi^+ \cup \Phi^-$ with $\Phi^- = -\Phi^+$ and any positive root is a unique linear combination of simple roots with nonnegative integral coefficients.

Since the roots span the dual $\mathbf{h}^*$ we can write any vector λ uniquely as

$$\lambda = \sum_{\alpha \in \Delta} k_\alpha \cdot \alpha,$$

where $k_\alpha \in \mathbb{F}$. In particular, if all the coefficients are nonnegative integers we set $\lambda \geq 0$. This defines a partial ordering in the dual, $\lambda \geq \mu$ if $\lambda - \mu \geq 0$.

We say that a vector $v \in V$ is *a maximal vector* if $\mathbf{g}_\alpha \cdot v = 0$ for all $\alpha \in \Phi^+$. Since the simple root subspaces $\mathbf{g}_\alpha$ generate all the root subspaces corresponding to positive roots (Theorem 2.3.11 and Corollary 3.3.5), this condition is equivalent to saying that $\mathbf{g}_\alpha \cdot v = 0$ for all simple roots.

Racall that a vector v has weight $\lambda \in \mathbf{h}^*$ if $hv = \lambda(h)v$ for all $h \in \mathbf{h}$. We call V a *a standard cyclic module of highest weight* λ if there is a cyclic maximal vector $v^+ \in V$ of weight λ. Thus

$$V = U(\mathbf{g})v^+, \qquad hv = \lambda(h)v \ \forall h \in \mathbf{h} \qquad \mathbf{g}_\alpha v = 0 \ \forall \alpha \in \Phi^+.$$

Note that any irreducible finite-dimensional $\mathbf{g}$ module is standard cyclic: The subspace $V^+ = \{v \in V | \mathbf{g}_\alpha v = 0 \forall \alpha \in \Phi^+\}$ is finite-dimensional and in cannot be equal to zero; otherwise V would be infinite-dimensional. In adddition $\mathbf{h}V^+ \subset V^+$ since for $x \in \mathbf{g}_\alpha$ and $v \in V^+$

$$xhv = hxv + [x,h]v = [x,h]v = -\alpha(h)xv = 0$$

for postive roots α. Since $\mathbf{h}$ is commutative and V^+ is finite-dimensional there must be a common eigenvector v^+ for all $h \in \mathbf{h}$. Finally, v^+ is cyclic since V is irreducible.

Theorem 4.5.1. *Let V be a (nonzero) standard cyclic $\mathbf{g}$ module with highest weight λ. Let $\Phi^- = \{\beta_1, \ldots, \beta_n\}$ and a maximal vector $v^+ \in V$. Choose $0 \neq y_i \in \mathbf{g}_{\beta_i}$. Then*

(1) *V is spanned by the vectors $y_1^{k_1} \ldots y_n^{k_n} v^+$ with $k_i = 0, 1, 2, \ldots$. In particular, V is a direct sum of the weight subspaces $V_\mu = \{v \in V | hv = \mu(h)v \forall h \in \mathbf{h}\}$*

(2) *All weights of V are of the form $\mu = \lambda - \sum_{\alpha \in \Delta} k_\alpha \cdot \alpha$ with nonnegative coefficients k_α, that is, all weights satisfy $\mu \leq \lambda$*

(3) $\dim V_\mu < \infty$ *for all weights μ and* $\dim V_\lambda = 1$

(4) *V is indecomposable and it has a unique maximal submodule W with the property $v^+ \notin W$*

(5) *If $\phi : V \to V'$ is a surjective $\mathbf{g}$ module homomorphism then also V' is a standard cyclic module of weight λ.*

Proof. (1-3) Choose a basis $x_1, \ldots, x_n$ in the root subspaces corresponding to positive roots and a basis $h_1, \ldots, h_\ell$ in the Cartan subalgebra. By the PBW theorem all elements in $U(\mathbf{g})$ are linear combinations of ordered monomials $P = y_1^{k_1} \ldots y_n^{k_n} h_1^{j_1} \ldots h_\ell^{j_\ell} x_1^{i_1} \ldots x_n^{i_n}$. But the vectors Pv^+ span the space V by the definition of a standard cyclic module. Now $x_i v^+ = 0$ for all i and v^+ is an eigenvector of any h_i. It follows that we may restrict to polynomials which do not contain any x_i, h_j factors. This proves the first statement in (1). All vectors $y_1^{k_1} \ldots y_n^{k_n} v^+$ are eigenvectors of h_j, with eigenvalue $(\lambda + k_1\beta_1 + \ldots k_n\beta_n)(h_j)$. Since the coefficients k_i are nonnegative and the roots β_i are negative we have proven the second statement in (1) and the claim (2). We obtain also (3) since there are only a finite number of sequences of nonnegative integers k_i such that $\mu = \lambda + \sum k_i \cdot \beta_i$. Clearly $\lambda = \mu$ if and only if all $k_i = 0$.

(4) To prove this we first observe that any submodule W is a sum of its weight spaces: Let $w \in W$ and write (by (1)) $w = w_1 + w_2 + \ldots w_n$ where $w_i \in V_{\mu_i}$ are the components in different weight spaces. Choose n in a minimal way such that not all w_i belong to W, so in this case no w_i belongs to W. Take any $h \in \mathbf{h}$ such that $\mu_1(h) \neq \mu_2(h)$. Then

$$(h - \mu_1(h))w = (\mu_2(h) - \mu_1(h))w_2 + \ldots (\mu_n(h) - \mu_1(h))w_n \neq 0.$$

But since $(h - a)w \in W$ for any $a \in \mathbb{F}$ and $hw \in W$ we reduce that $w_2 \in W$, by the minimality of n; but this is a contradiction.

Now let us assume in the contrary that $V = V_1 \oplus V_2$ where $V_i \subset V$ are nonzero submodules. Write $v^+ = v_i + v_2$ where $v_i \in V_i$. Since $x_i v^+ = 0$ for all i we must have $x_i v_1 = x_i v_2 = 0$ since $V_1 \cap V_2 = 0$. In addition, both v_1 and v_2 must be eigenvectors of all $h \in \mathbf{h}$. By (3) the vectors v_i must be linearly dependent, which is absurd by $V = V_1 \oplus V_2$.

Let then $W \subset V$ such that $W \neq V$. Then $v^+ \notin W$ since $V = U(\mathbf{g})v^+$. Furthermore, by the observation above, no vector in W has a nonzero projection on V_λ (since otherwise the projection would be in the submodule and then $W = V$, contradiction). It follows that the sum of all submodules not containing the vector v^+ is again a submodule not containing v^+ and thus a proper (maximal) submodule.

(5) Let $\phi : V \to V'$ be a surjective homomorphism. Now $U(\mathbf{g})\phi(v^+) = \phi(U(\mathbf{g})v^+) = \phi(V) = V'$. In addition,

$$x_i\phi(v^+) = \phi(x_iv^+) = 0 \qquad h\phi(v^+) = \phi(hv^+) = \phi(\lambda(h)v^+) = \lambda(h)\phi(v^+)$$

and so $\phi(v^+) \in V'$ is a maximal cyclic vector of weight λ.

Theorem 4.5.2. *Any two irreducible standard cyclic modules with the same highest weight are isomorphic.*

Proof. Let V, W be standard cyclic modules of highest weight λ. Consider the $\mathbf{g}$ module $X = V \oplus W$. Let $v^+ \in V$ and $w^+ \in W$ be maximal vectors and denote $x^+ = v^+ \oplus w^+ \in X$. Then x^+ is a maximal vector with weight λ. Denote $Y = U(\mathbf{g})x^+ \subset X$. Then Y is a standard cyclic module of weight λ. Let $p : Y \to V$ and $p' : Y \to W$ be the projections. Now $V = U(\mathbf{g})v^+ = p(U(\mathbf{g})(v^+ \oplus w^+)) = p(Y)$ so $p : Y \to V$ is surjective. In the same way $p' : Y \to W$ is surjective. The kernel of p' is the submodule $V \cap Y$. Since V was assumed to be irreducible this submodule must be either 0 or V. The latter is impossible since then $v^+ \in Y$ would be another maximal vector in Y of weight λ. But according to Theorem 4.5.1 (3) the vectors $v^+ \oplus 0$ and $x^+ = v^+ \oplus w^+$ would be linearly dependent, that is, $w^+ = 0$. Thus $V \cap Y = 0$ and p' is is injective. We have shown that $p' : Y \to W$ is an isomorphism. In the same way one proves that $p : Y \to V$ is an isomorphism and therefore we have the isomorphism $p' \circ p^{-1} : V \to W$.

By Theorem 4.5.1 all the other weights in a standard cyclic module of highest weight λ are strictly smaller than λ. This motivates our terminology of highest weights.

The following theorem, together with 4.5.1, tells us that we may identify the set of equivalence classes of irreducible standard cyclic modules as the dual space $\mathbf{h}^*$.

Theorem 4.5.3. *For each $\lambda \in \mathbf{h}^*$ there is an irreducible nonzero standard cyclic module V of highest weight λ.*

Proof. Choose the basis $\{y_i, h_i, x_i\}$ in $\mathbf{g}$ as in the proof of 4.5.1. Let $I_\lambda \subset U(\mathbf{g})$ be the left-ideal generated by the vectors x_i and the vectors $h_i - \lambda(h_i) \cdot \mathbf{1}$. Define the $\mathbf{g}$ module

$$Z(\lambda) = U(\mathbf{g})/I_\lambda.$$

This module does not need to be irreducible. Let $v^+ = \mathbf{1} + I_\lambda \in Z(\lambda)$. Clearly $x_i v^+ = 0$ for all i and $hv^+ = \lambda(h_i)v^+$. In addition, $Z(\lambda) = U(\mathbf{g})v^+$. Thus $Z(\lambda)$ is a standard cyclic module of highest weight λ.

The module $Z(\lambda) \neq 0$ by the PBW theorem (compare with 4.5.1 (1)). Let $W \subset Z(\lambda)$ be the unique maximal submodule of $Z(\lambda)$ given by 4.5.1 (4). Set $V = Z(\lambda)/W$. This is again a standard cyclic $\mathbf{g}$ module of highest weight λ. The highest weight vector is $v^+ + W \neq 0$. This module is irreducible. Otherwise, there would be a nonzero proper submodule $X = W'/W$ where $W' \subset Z(\lambda)$ is a strictly larger submodule than W. But this is in contradiction with the maximality of W.

Exercise 4.5.4 Let $\mathbf{g}$ be semisimple and V an irreducible $\mathbf{g}$ module. a) Assume that there is at least one nonzero weight space $V_\lambda \subset V$. Prove that V is a direct sum of weight spaces. b) Show that V has a nonzero weight space if and only if $U(\mathbf{h})v$ is finite-dimensional for every vector $v \in V$. Here $\mathbf{h} \subset \mathbf{g}$ is a Cartan subalgebra.

Exercise 4.5.5 Let $\mathbf{g} = \mathbf{sl}(2, \mathbb{C})$ and x, y, h the standard basis of $\mathbf{g}$. a) Show that the element $1 - x \in U(\mathbf{g})$ is not invertible; hence it lies in a maximal proper left ideal $I \subset U(\mathbf{g})$. b) Now $V = U(\mathbf{g})/I$ is a nonzero irreducible $\mathbf{g}$ module. Show that all the vectors $1, h, h, \dots$ represent linearly independent elements in V. Conclude that there are no nonzero weight spaces in V. (Exercise 4.5.4!) Hint: Use the fact that $(x-1)^r h^s \equiv 0 \mod I$ if $r > s$ and $(x-1)^r h^s \equiv (-2)^r r! \cdot \mathbf{1} \mod I$ if $r = s$.

Exercise 4.5.6 Calculate weights and find the maximal vectors for the defining representation of the simple Lie algebras $A_\ell - D_\ell$.

Exercise 4.5.7 Let $Z(\lambda)$ be the standard cyclic module constructed in Theorem 4.5.3. Assume that there is a maximal vector $w^+ \in Z(\lambda)$ of weight μ. Construct an *injective* module homomorphism $\phi : Z(\mu) \to Z(\lambda)$.

4.6 Finite-dimensional modules

Let $\Delta = \{\alpha_1, \dots, \alpha_\ell\} \subset \Phi$ be a set of simple roots. We set $h_i = 2h_{\alpha_i}/(\alpha_i, \alpha_i)$. With this normalization $\alpha_i(h_i) = 2$ for each i. For any $\lambda \in \mathbf{h}^*$ we denote $\lambda_i = \lambda(h_i)$.

Theorem 4.6.1. *Let V be a finite-dimensional* $\mathbf{g}$ *module with highest weight* λ. *Then $\lambda_i \in \mathbb{Z}_+$ for $i = 1, 2, \dots, \ell$.*

Proof. Let $0 \neq x_i \in \mathbf{g}_{\alpha_i}$ and $0 \neq y_i \in \mathbf{g}_{-\alpha_i}$. Then for each index i the vectors x_i, h_i, y_i span a Lie algebra isomorphic to $\mathbf{sl}(2, \mathbb{F})$. Let v be a highest weight vector. Now $U(\mathbf{sl}(2, \mathbb{F}))v$ is a cyclic $\mathbf{sl}(2, \mathbb{F})$ module, and by Weyl's theorem, it must be irreducible. But then by Theorem 4.2.2 the eigenvalue λ_i of h_i must be a nonnegative integer.

We denote by Λ the set of *integral weights,* that is, the set of all $\lambda \in \mathbf{h}^*$ with $\lambda_i \in \mathbb{Z}$. The subset Λ^+ of dominant integral weights consists of $\lambda \in \Lambda$ with $\lambda_i \geq 0$ for all i.

Lemma 4.6.2. *Let $x_i, h_i, y_i \in$* $\mathbf{g}$ *as above, with the normalization $[x_i, y_i] = h_i$. Then the following relations hold in $U(\mathbf{g})$:*

(1) $[x_j, y_i^{k+1}] = 0$ *for* $i \neq j$
(2) $[h_j, y_i^{k+1}] = -(k+1)\alpha_i(h_j)y_i^{k+1}$
(3) $[x_i, y_i^{k+1}] = (k+1)y_i^k \cdot (h_i - k \cdot \mathbf{1})$

for $k = 0, 1, 2 \dots$ *and* $i, j = 1, \dots, \ell$.

Proof. (1) This follows directly from $[x_i, y_j] = 0$ since $\alpha_i - \alpha_j$ is not a root.

(2) The case $k = 0$ is clear by $[h, y_i] = -\alpha_i(h)y_i$ for any $h \in \mathbf{h}$. Induction in k :

$$\begin{aligned}[h_j, y_i^{k+1}] &= h_j y_i^{k+1} - y_i^{k+1} h_j = (h_j y_i^k - y_i^k h_j) y_i + y_i^k (h_j y_i - y_i h_j) \\ &= -k\alpha_i(h_j) y_i^k \cdot y_i + y_i^k (-\alpha_i)(h_j) y_i = -(k+1)\alpha_i(h_j) y_i^{k+1}.\end{aligned}$$

(3) The case $k = 0$ follows from the choice of normalization. Induction in k :

$$\begin{aligned}[x_i, y_i^{k+1}] &= [x_i, y_i] y_i^k + y_i [x_i, y_i^k] = h_i y_i^k + y_i k y_i^{k-1}(h_i - (k-1)) \\ &= y_i^k h_i + k\alpha_i(h_i) y_i^k + y_i^k (k h_i - k(k-1)) \\ &= (k+1) y_i^k (h_i - k).\end{aligned}$$

Lemma 4.6.3. *Let $\alpha_1, \dots, \alpha_\ell$ be any basis in in a real inner product space such that $(\alpha_i, \alpha_j) \leq 0$ for all $i \neq j$. Then $(\alpha_i^*, \alpha_j^*) \geq 0$ for all i, j. Here the star refers to the dual basis, $(\alpha_i^*, \alpha_j) = \delta_{ij}$.*

Proof. The general case reduces to a two-dimensional problem by taking a projection to a subspace spanned by α_i, α_j for any fixed index pair. In two dimensions, denoting the angle between α_i and α_j by θ with $\pi/2 \leq \theta \leq \pi$, the angle between α_2, α_2^* is $\theta - \pi/2$, the angle between α_2, α_1^* is $\pi/2$, the angle between α_1^*, α_2^* is $\pi - \theta \leq \pi/2$. Draw a picture to convince yourself!

Lemma 4.6.4. *Let $\lambda \in \Lambda^+$. Then the set $S_\lambda = \{\mu \in \Lambda^+ | \mu < \lambda\}$ is finite.*

Proof. Again, denoting by starred vectors the elements in the dual basis, we have

$$\lambda = \sum r_i \cdot \alpha_i = \sum \frac{(\alpha_i, \alpha_i)}{2} \lambda_i \alpha_i^*$$

with $\lambda_i \in \mathbb{Z}_+$ and

$$r_i = (\alpha_i^*, \lambda) = \sum_j \frac{(\alpha_j, \alpha_j)}{2} \lambda_j (\alpha_i^*, \alpha_j^*) \geq 0.$$

If now $\mu \in S_\lambda$ then $\mu = \sum s_i \alpha_i$ with $s_i \geq 0$ and $r_i - s_i \in \mathbb{Z}_+$. For fixed set $\{r_i\}$ there is only a finite number of solutions of these inequalities.

Theorem 4.6.5. *Let $\lambda \in \Lambda^+$. Then the irreducible standard module $V(\lambda)$ with highest weight λ is finite-dimensional.*

Proof. (1) Let $0 \neq v^+ \in V(\lambda)$ be a maximal vector with weight λ. Using the same notation as in Lemma 4.6.2 we set

$$w = y_i^{\lambda_i + 1} v^+ \text{ with } \lambda_i = \lambda(h_i).$$

If $j \neq i$ then $x_j w = 0$ by 4.6.2 (1). In addition,

$$x_i y_i^{\lambda_i+1} v^+ = y_i^{\lambda_i+1} x_i v^+ + (\lambda_i + 1) y_i^{\lambda_i} (h_i - \lambda_i) v^+ = 0.$$

If $w \neq 0$ then w is a maximal vector with weight $\mu = \lambda - (\lambda_i + 1)\alpha_i \neq \lambda$. In an irreducible module the maximal weight is unique, so $w = 0$.

(2) By (1) the subspace spanned by the vectors $y_i^k v^+$ with $k = 0, 1, \dots, \lambda_i$ is a submodule for the subalgebra $\mathbf{g}_i = \mathbf{sl}(2, \mathbb{F})$ spanned by x_i, h_i, y_i.

(3) Let $V' \subset V(\lambda)$ be the sum of all finite-dimensional $\mathbf{g}_i$ submodules. By (2) we have $V' \neq 0$. Let $W \subset V'$ be any finite-dimensional $\mathbf{g}_i$ submodule. It is easy to see that the subspace spanned by all the vectors $\mathbf{g}_\alpha W$, with $\alpha \in \Phi$, is a finite-dimensional $\mathbf{g}_i$ submodule. It follows that $\mathbf{g}$ maps V' onto itself. But $V(\lambda)$ is irreducible and therefore $V' = V(\lambda)$.

(4) The action of both x_i and y_i is clearly nilpotent in each finite-dimensional $\mathbf{g}_i$ submodule. As we saw in (3), any vector $v \in V(\lambda)$ belongs to some finite-dimensional $\mathbf{g}_i$ submodule and thus $x_i^n v = y_i^n v = 0$ when $n \geq n_i(v)$. It follows that the element $s_i = \exp(x_i) \cdot \exp(-y_i) \cdot \exp(x_i)$ defined by power series expansion is well-defined in $V(\lambda)$.

(5) Let μ be a weight of $V(\lambda)$. Since $V(\lambda)$ is a sum of finite-dimensional $\mathbf{g}_i$ submodules and all the weight subspaces $V_\mu \subset V(\lambda)$ are finite-dimensional (Theorem 4.5.1), we have $V_\mu \subset W$ where W is a finite-dimensional $\mathbf{g}_i$ submodule. Thus s_i is well-defined in V_μ. Now

$$\begin{aligned} s_i h_i s_i^{-1} &= e^{x_i} e^{-y_i} e^{x_i} h_i e^{-x_i} e^{y_i} e^{-x_i} \\ &= e^{\mathrm{ad}_{x_i}} e^{-\mathrm{ad}_{y_i}} e^{\mathrm{ad}_{x_i}} h_i = -h_i. \end{aligned}$$

The last equation follows by a direct computation from the power series expansion and the basic commutation relations between x_i, h_i, y_i.

(6) From (5) (and from a similar calculation for $s_i h_j s_i^{-1}$) follows that $s_i V_\mu = V_{\sigma_i \mu}$, where $\sigma_i = \sigma_{\alpha_i}$ is the basic reflection $\sigma_i \mu = \mu - <\mu, \alpha_i> \alpha_i$.

(7) Let $T(\lambda)$ be the set of all weights in $V(\lambda)$. The Weyl group W is generated by the basic reflections σ_i and so by (6) the group W permutes the weights $T(\lambda)$. It follows that dim $V_\mu =$ dim $V_{\sigma(\mu)}$ for all $\sigma \in W$. By the Lemma 4.6.4 the set S_λ is finite. On the other hand, by the Theorem 4.6.7 below, $T(\lambda) \subset W S_\lambda$ and it follows that $T(\lambda)$ is finite. But $V(\lambda)$ is a sum of finite-dimensional weight subspaces and we are done.

Corollary 4.6.6. dim $V_\mu =$ dim $V_{\sigma(\mu)}$ *for all weights* μ *and for all* $\sigma \in W$.

Theorem 4.6.7. *Let* $\lambda \in \Lambda$. *Then there exists a unique* $\mu \in \Lambda^+$ *such that* $\sigma(\mu) = \lambda$ *for some* $\sigma \in W$.

Proof. Let $T(\Delta)$ be the Weyl chamber corresponding to the basis Δ. Then

$$\Lambda^+ = \Lambda \cap \overline{T(\Delta)}.$$

By Theorem 3.4.5 there exists $\sigma \in W$ such that $\sigma(\mu) \in \overline{T(\Delta)}$. The action of σ preserves the integrality property and thus $\sigma(\mu) \in \Lambda^+$.

Remark One can also prove that for any $\lambda \in \Lambda^+$ and $\sigma \in W$ we have $\sigma(\lambda) \leq \lambda$. If in addition $\lambda_i > 0$ for all i then $\sigma(\lambda) = \lambda$ only when $\sigma = 1$.

Exercise 4.6.8 *The fundamental dominant weights* $\lambda_1, \dots, \lambda_\ell$ are defined by the property $< \lambda_i, \alpha_j >= \delta_{ij}$, where $\alpha_1, \dots, \alpha_\ell$ are simple roots. Find all the weights in an irreducible A_2 module with highest weight $\lambda = 3\lambda_1$.

Exercise 4.6.9 Denote by **3** the defining three-dimensional representation of A_2 and by $\mathbf{3}^*$ the representation in the dual space. By analysing the weights of the tensor product representation show that $\mathbf{3} \otimes \mathbf{3}^*$ is equivalent to the direct sum of the trivial one-dimensional representation **1** and the adjoint representation **8**.

Exercise 4.6.10 Let $\mathbf{g} = A_\ell$ and let $\mathbf{h}$ be the standard Cartan subalgebra consisting of diagonal matrices. Define coordinates $\mu_i : \mathbf{h} \to \mathbb{C}$ by setting $\mu_i(h) =$ the i : th diagonal element in h. Then $\mu_1 + \dots + \mu_{\ell+1} = 0$. The simple roots can be written as $\alpha_i = \mu_i - \mu_{i+1}$ with $i = 1, 2, \dots, \ell$. Show that the Weyl group acts on $\mathbf{h}^*$ by permuting the coordinates μ_i. Show that the fundamental dominant weights are $\lambda_i = \mu_1 + \dots \mu_i$, with $i = 1, 2, \dots, \ell$.

Exercise 4.6.11 Let λ_k be the fundamental weight of A_ℓ discussed in 4.6.10. Show that the irreducible finite-dimensional module corresponding to this weight can be realized in the space of totally antisymmetric tensors of rank k constructed from $V = \mathbb{C}^{\ell+1}$.

CHAPTER 5 AFFINE KAC-MOODY ALGEBRAS

5.1. Affine Kac-Moody algebras from generalized Cartan matrices

In the earlier chapters we explained how simple finite-dimensional Lie algebras can be completely characterized in terms of their Cartan matrices or Dynkin diagrams. The same holds for an arbitrary semisimple finite-dimensional Lie algebra. A semisimple Lie algebra is a direct sum of simple ideals which are pairwise orthogonal with respect to the Killing form. It follows that the Cartan matrix of a semisimple Lie algebra decomposes to a block diagonal form, each block representing a simple ideal. Similarly, the Dynkin diagram is a disconnected union of Dynkin diagrams of simple Lie algebras. Next we shall study certain infinite-dimensional Lie algebras which have many similarities with the simple finite-dimensional Lie algebras. In particular, they can be described in terms of generalized Cartan matrices. These algebras were independently introduced in V. Kac and R. Moody in 1968.

A *generalized Cartan matrix* is a real $n \times n$ matrix $A = (a_{ij})$ such that

(C1) $a_{ii} = 2$ for $i = 1, 2, \dots, n$

(C2) a_{ij} is a nonpositive integer for $i \neq j$

(C3) $a_{ij} = 0$ iff $a_{ji} = 0$.

To each generalized Cartan matrix one can associate a Lie algebra using the method of generators and relations as explained in V. Kac: *Infinite Dimensional Lie Algebras,* Cambridge University Press (1985). However, we shall not take that road since we shall describe in the next section a simple method for constructing those algebras which we shall deal with in this book; however, see the exercise 5.2.7. The set of *indecomposable* matrices A, i.e., those which cannot be written in a block diagonal form by reordering the indices {1,2, ... ,n}, can be divided into three disjoint subsets:

(1) There is a vector $v \in \mathbf{N}_+^n$ such that also $Av \in \mathbb{N}_+^n$. In this case the Lie algebra $\mathbf{g}(A)$ corresponding to A is a simple finite-dimensional Lie algebra.
(2) There is $v \in \mathbb{N}_+^n$ such that $Av = 0$. The algebra $\mathbf{g}(A)$ is an *affine Lie algebra* and $\dim \mathbf{g}(A) = \infty$.

(3) There is $v \in \mathbb{N}_+^n$ such that $(Av)_i < 0 \, \forall i$.

In this chapter we shall concentrate to the theory of affine Kac-Moody algebras, which is much better understood than the Kac-Moody algebras of class (3). However, the class (3) contains the subclass of the so-called *hyperbolic* Lie algebras which seem to have interesting mathematical structures; see the discussion in Feingold and Frenkel, Math. Ann. **263,** p. 87 (1983), where the hyperbolic algebra corresponding to the Cartan matrix

$$A = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

has been studied in detail. We shall now give a list of the Dynkin diagrams of the affine Lie algebras. For the proofs see Kac [1985]. The diagrams with the upper index 1 correspond to the *untwisted affine Lie algebras* and the rest describe the *twisted affine Lie algebras* . The reason for this division will become apparent in the next section. Note that each of the Dynkin diagrams is obtained by adjoining the node labeled by 0 to a Dynkin diagram of a simple finite-dimensional algebra.

5.2. Affine Lie algebras as central extensions of loop algebras: the untwisted case

Let $\mathbf{g}$ be an arbitrary finite-dimensional complex Lie algebra and denote by $L\mathbf{g}$ the space of smooth maps (loops) $f : S^1 \to \mathbf{g}$, where S^1 is the unit circle. Consider $L\mathbf{g}$ as a vector space by pointwise addition of the loops and the natural multiplication of functions by complex numbers. Furthermore, $L\mathbf{g}$ is naturally an infinite-dimensional Lie algebra through the commutator $[\cdot,\cdot]_{(0)}$,

$$[f,g]_{(0)} = [f(z), g(z)],\ \ z \in S^1.$$

A smooth function on S^1 is always square-integrable and a basis for square-integrable functions is given by the Fourier modes. Let $\{T_1, T_2, \ldots, T_r\}$ be a basis of $\mathbf{g}$ and denote

$$T_a^n = e^{in\phi} T_a,$$

where $0 \le \phi \le 2\pi$ parametrizes the circle and $n \in \mathbb{Z}$. Define the structure constants of $\mathbf{g}$ by

$$[T_a, T_b] = \sum_{c=1}^{r} \lambda_{ab}^c T_c.$$

Then

$$[T_a^n, T_b^m]_{(0)} = \sum_c \lambda_{ab}^c T_c^{n+m}.$$

Let $(\cdot,\cdot) : \mathbf{g} \times \mathbf{g} \to \mathbb{C}$ be any *invariant* bilinear symmetric form, that means

$$([x,y],z) = (y,[z,x])\,\forall x,y,z \in \mathbf{g}.$$

Let $\hat{\mathbf{g}}$ denote the vector space $L\mathbf{g} \oplus \mathbb{C}$. We define in $\hat{\mathbf{g}}$ the following commutator:

(A) $$[(f,\alpha),(g,\beta)] = \left([f,g]_{(o)},\ \frac{k}{2\pi i}\int_0^{2\pi} (f(\phi), g'(\phi))d\phi\right).$$

Here $0 \neq k \in \mathbb{C}$ is an arbitrary constant. For brevity, we shall denote the pair $(f,0)$ by f. For the Fourier modes the equation (A) gives

(B) $$[T_a^n, T_b^m] = \sum \lambda_{ab}^c T_c^{n+m} + km\delta_{n,-m}(T_a,T_b).$$

Next let $\mathbf{g}$ be a *simple* Lie algebra. We shall show that the commutation relations (B) define an untwisted affine Lie algebra. Choose a Cartan subalgebra $\mathbf{h} \subset \mathbf{g}$. We shall identify $\mathbf{g}$ with a subalgebra of $\hat{\mathbf{g}}$ by $x \mapsto$ (the constant function $x : S^1 \to \mathbf{g}$ taking the value x). We can write

$$\hat{\mathbf{g}} = \mathbb{C} \oplus \sum_{n \in \mathbf{Z}} \mathbf{g}^{(n)},$$

where $\mathbf{g}^{(n)}$ is spanned by the vectors T_a^n, $1 \le a \le r$, and $\mathbb{C}$ stands for the *center* of $\hat{\mathbf{g}}$ spanned by the vector $k = (0, k)$. In particular, $\mathbf{g}^{(0)} = \mathbf{g}$. Let Φ be the system of roots for $(\mathbf{g}, \mathbf{h})$ and $\Delta \subset \Phi^+$ a system of simple roots. Choose $0 \neq x_\alpha \in \mathbf{g}_\alpha, 0 \neq y_\alpha \in \mathbf{g}_{-\alpha}\, \forall \alpha \in \Phi^+$. From (B) we get

$$[h, x_\alpha^n] = \alpha(h) x_\alpha^n,$$
$$[h, y_\alpha^n] = -\alpha(h) y_\alpha^n,$$
$$[k, x_\alpha^n] = [k, y_\alpha^n] = 0,$$

where we have also used Lemma 2.3.1 (1). We notice that if we define $\mathbf{h} \oplus \mathbb{C}k$ to be the Cartan subalgebra of $\hat{\mathbf{g}}$, then each of the roots α has an infinite multiplicity. For this reason we extend the algebra $\hat{\mathbf{g}}$ by an element d (and to add confusion we shall denote the new algebra also by $\hat{\mathbf{g}}$) which has the following commutation relations:

(C) $$[d, T_a^n] = n T_a^n \qquad [d, k] = 0.$$

A concrete realization for the new element is $d = -i \frac{d}{d\phi}$. We then define the Cartan subalgebra of $\hat{\mathbf{g}}$ as

$$\hat{\mathbf{h}} = \mathbf{h} \oplus \mathbb{C}k \oplus \mathbb{C}d.$$

Correspondingly, we write a root of $(\hat{\mathbf{g}}, \hat{\mathbf{h}})$ in the component form
$(\alpha, 0, n)$; this root corresponds to the root vector x_α^n. Thus the set of nonzero roots for $(\hat{\mathbf{g}}, \hat{\mathbf{h}})$ is

$$\widehat{\Phi} = \{(\pm\alpha, 0, n) \mid \alpha \in \Phi^+, n \in \mathbb{Z}\} \cup \{(0, 0, n) \mid 0 \neq n \in \mathbb{Z}\}.$$

The root subspace of the root $(0, 0, n)$, $(n \neq 0)$ is spanned by the vectors h_i^n, where $\{h_1, \ldots, h_l\}$ is an orthonormal basis of $\mathbf{h}$. Each of the roots $(\pm\alpha, 0, n)$ has multiplicity $=1$ and each of the nonzero roots $(0, 0, n)$ has multiplicity $=l$. We define a system of simple roots

$$\widehat{\Delta} = \{(\alpha, 0, 0) \mid \alpha \in \Delta\} \cup \{(-\psi, 0, 1)\},$$

where ψ is the *highest root* of $(\mathbf{g}, \mathbf{h})$, that is, ψ is the highest weight of the adjoint representation $ad_x(y) = [x, y]$ of $\mathbf{g}$. The set of positive roots is then

$$\widehat{\Phi}^{+} = \{(\alpha, 0, n) \mid \alpha \in \Phi,\ n > 0\} \cup \{(\alpha, 0, 0) \mid \alpha \in \Phi^{+}\}$$

and $\widehat{\Phi}^{-} = -\widehat{\Phi}^{+}$ as in the case of finite-dimensional semisimple Lie algebras.

Example 5.2.1. Let $\mathbf{g} = A_l$. We use the standard Cartan subalgebra of diagonal matrices and denote the root corresponding to the Lie algebra element e_{ij} by α_{ij}. . The highest weight vector in the adjoint representation is $e_{1,l+1}$ since $[x_{\alpha_{ij}}, e_{1,l+1}] = [e_{ij}, e_{1,l+1}] = 0$ for $i < j$. The highest root is thus $\alpha_{1,l+1} = \alpha_{12} + \alpha_{23} + \cdots + \alpha_{l,l+1} = \alpha_1 + \alpha_2 + \cdots + \alpha_l$.

Exercise 5.2.2. Let $\ell_n = -ie^{in\phi}\frac{d}{d\phi}$ with $n \in \mathbb{Z}$. Compute all the commutators $[\ell_n, \ell_m]$ and $[\ell_n, T^a_m]$ and show that they define a Lie algebra. Define then the new commutators

$$[\ell_n, \ell_m]_c = [\ell_n, \ell_m] - c(n^3 - n)\delta_{n,-m},$$

where c is a new operator which (by definition) commutes with everything. Show that also the new commutation relations define a Lie algebra; this is called the *Virasoro algebra.*

Exercise 5.2.3 Consider the algebra of fermionic creation and annihilation operators generated by the elements a^*_i, a_i (see Section 4.4) and with the defining relations $a_i a_j + a_j a_i = 0 = a^*_i a^*_j + a^*_j a^*_i$ and $a^*_i a_j + a_j a^*_i = \delta_{ij}$. We let i, j be arbitrary integers. Define the operators $\hat{e}_{ij} =: a^*_i a_j :$ where the dots mean *normal ordering,* $: a^*_i a_j := -a_j a^*_i$ if $i = j < 0$ and otherwise the order is unaffected by the normal ordering. Compute the commutators $[\hat{e}_{ij}, \hat{e}_{kl}]$. Let $\hat{x} = \sum_{ij} x_{ij}\hat{e}_{ij}$ where $x = (x_{ij})$ is an infinite matrix with a finite number of nonzero matrix elements.Show that

$$[\hat{x}, \hat{y}] = \widehat{[x, y]} + \frac{1}{2}\mathrm{tr}\, x[\epsilon, y],$$

where ϵ is the diagonal matrix with $\epsilon_{ii} = 1$ for $i \geq 0$ and $\epsilon_{ii} = -1$ for $i < 0$.

There is an *invariant symmetric bilinear form* on $\hat{\mathbf{g}}$ given by

(B1) $$(f, g) = \frac{1}{2\pi}\int_0^{2\pi} (f(\phi), g(\phi))d\phi$$

(B2) $$(k, f) = (d, f) = 0 \qquad f \in L\mathbf{g}$$

(B3) $$(k, k) = (d, d) = 0$$

(B4) $$(k, d) = 1$$

where the form under the integral sign is the Killing form of $\mathbf{g}$.

Proposition 5.2.3. *Up to a multiplicative constant any invariant symmetric bilinear form on* $\hat{\mathbf{g}}$ *is obtained from the form above by replacing* $(d,d) = 0$ *by* $(d,d) = s$*, where* $s \in \mathbb{C}$ *is an arbitrary constant.*

Proof. If $(\cdot,\cdot)$ is invariant we have

$$\begin{aligned}([d, T_a^n], T_b^m) &= n(T_a^n, T_b^m)\\ &= (T_a^n, [T_b^m, d]) = -m(T_a^n, T_b^m)\end{aligned}$$

and so $(T_a^n, T_b^m) = 0$ if $n \neq -m$. For a fixed n, write $\eta_{ab} = (T_a^n, T_b^{-n})$. Using the invariance of the Killing form of $\mathbf{g}$ we get $\lambda_{ab}^c = -\lambda_{ac}^b$ in the orthonormal basis $\{T_a\}$; $(T_a, T_b) = -\delta_{ab}$. Comparing with

$$\begin{aligned}([T_c, T_a^n], T_b^{-n}) &= \sum_e \lambda_{ca}^e (T_e^n, T_b^{-n}) = \sum \lambda_{ca}^e \eta_{eb}\\ &= (T_a^n, [T_b^{-n}, T_c]) = -\sum \lambda_{cb}^e (T_a^n, T_e^{-n})\\ &= -\sum \lambda_{cb}^e \eta_{ae}\end{aligned}$$

we conclude that the matrix η commutes with each of the matrices $\lambda_a = (\lambda_{a,bc})$, $\lambda_{a,bc} = -\lambda_{a,cb} = \lambda_{ac}^e g_{eb}$, and $g_{ab} = (T_a, T_b)$. [The antisymmetry of λ_a follows from the invariance of the form $(\cdot,\cdot)$ on $\mathbf{g}$.] The adjoint representation is irreducible for any simple Lie algebra and thus by Schur's lemma the matrix η has to be proportional to the identity,

$$(T_a^n, T_b^{-n}) = \xi(n)\delta_{ab}.$$

From

$$([T_a^1, T_b^n], T_c^{-n-1}) = (T_b^n, [T_c^{-n-1}, T_a^1])$$

we conclude that $\xi(n) = \xi(n+1)\, \forall n$. On the other hand,

$$\frac{1}{2\pi}\int (T_a^n, T_b^m) d\phi = \delta_{ab}\delta_{n,-m},$$

so after a renormalization the inner product takes the form (B1). We leave as an exercise for the reader to complete the proof by showing that with this normalization (B2),(B4) and $(k,k) = 0$ holds.

Exercise 5.2.4 Complete the proof of Prop. 5.2.3

Let $\{h_1, h_2, \dots, h_l\}$ be an orthonormal basis of $\mathbf{h}$. Then

$$\{h_1, \dots, h_l, k, d\}$$

is a basis of $\hat{\mathbf{h}}$ and the restriction of the invariant form (B) to $\hat{\mathbf{h}}$ is described by the matrix

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

It is nondegenerate but *Lorentzian* in signature: In the basis

$$\left\{h_1, \dots, h_l, \frac{1}{\sqrt{2}}(k+d), \frac{1}{\sqrt{2}}(k-d)\right\}$$

it takes the form $\mathrm{diag}(+1, \dots, +1, -1)$. If $\mu, \mu' \in \hat{\mathbf{h}}^*$ are arbitrary linear forms, then the dual of the inner product (B) on $\hat{\mathbf{h}}^* \times \hat{\mathbf{h}}^*$ is

$$(\mu, \mu') = \sum_{i=1}^{l} \mu(h_i)\mu'(h_i) + \mu(d)\mu'(k) + \mu(k)\mu'(d).$$

We can now compute the scalar products between the simple roots $\widehat{\Delta}$. We shall work only through the case $\mathbf{g} = A_l$; the other cases are handled in the same way. (All we need to know is the highest root ψ as a linear combination of the simple roots and the Dynkin diagram or the Cartan matrix of $\mathbf{g}$.)

Example 5.2.5. $\mathbf{g} = A_l$. Now $\psi = \alpha_1 + \alpha_2 + \dots + \alpha_l$, where the α_i's are the simple roots. If $2 \le i \le l-1$, then

$$(\psi, \alpha_i) = (\alpha_{i-1} + \alpha_i + \alpha_{i+1}, \alpha_i) = -1 + 2 + (-1) = 0.$$

The only nonzero products involving ψ are $(\psi, \alpha_1) = (\alpha_1 + \alpha_2, \alpha_1) = 1$ and $(\psi, \alpha_l) = (\alpha_{l-1} + \alpha_l, \alpha_l) = 1$. Denoting the simple root $(-\psi, 0, 1)$ of $\hat{\mathbf{g}}$ by α_0 we obtain the Dynkin diagram of $\hat{\mathbf{g}}$ from that of $\mathbf{g}$ by adjoining the node labeled by 0 and connecting the new node to the nodes 1 and l.

Exercise 5.2.6. Show that the Dynkin diagram of $\hat{\mathbf{g}}$ is equal to the diagram $G_2^{(1)}$ in the list Aff1 when $\mathbf{g} = G_2$.

Exercise 5.2.7 Let e_i with $i = 1, 2, \dots, \ell$ be the root vectors in a simple Lie algebra $\mathbf{g}$ corresponding to simple roots $\alpha_1, \dots, \alpha_\ell$ and let f_i be the root vectors corresponding to the negative roots $-\alpha_i$, normalized such that $(e_i, f_i) = \frac{2}{(\alpha_i,\alpha_i)}$. Let $\{h_i\}$ be a basis in the Cartan subalgebra with $[e_i, f_i] = h_i$. Now consider the central extension $\hat{\mathbf{g}}$ (with central element k) of the loop algebra $L\mathbf{g}$. Set $e_0 = e^{i\phi} y_\beta$, where β is the highest root and y_β is the corresponding root vector belonging to $-\beta$ such that $(x_\beta, y_\beta) = \frac{2}{(\beta,\beta)}$. Set $f_0 = e^{-i\phi} x_\beta$ and $h_0 = [y_\beta, x_\beta] - k(y_\beta, x_\beta)$. Show that the Serre relations hold in the algebra $\hat{\mathbf{g}}$:

$$[e_i, f_j] = h_i \delta_{ij} \tag{S1}$$

$$[h_i, e_j] = a_{ji} e_j \tag{S2}$$

$$[h_i, f_j] = -a_{ji} f_j \tag{S3}$$

$$(\mathrm{ad}_{e_j})^{1-a_{kj}} e_k = 0 \text{ for } k \neq j \tag{S4}$$

$$(\mathrm{ad}_{f_j})^{1-a_{kj}} f_k = 0. \text{ for } k \neq j \tag{S5}$$

Here (a_{ij}), with $i, j = 0, 1, \ldots, \ell$, is the Cartan matrix of the Lie algebra $\hat{\mathbf{g}}$ and (a_{ij}) with $i, j = 1, \ldots, \ell$ the Cartan matrix of $\mathbf{g}$.

Exercise 5.2.8 Using the notation of Exercise 5.2.3, define the operators $L_n = \sum_j j : a^*_{n+j} a_j :$. Compute the commutators $[L_n, L_m]$ and compare with the algebra in 5.2.2.

5.3. Affine Lie algebras as central extensions of loop algebras: the twisted case

Let again $\mathbf{g}$ be a simple complex finite-dimensional Lie algebra and let $\sigma : \mathbf{g} \to \mathbf{g}$ be an automorphism such that $\sigma^N = 1$ for some integer $N > 0$; let N be the smallest positive integer for which this holds. Set $\epsilon = e^{2\pi i/N}$. Let $L_\sigma \mathbf{g}$ consist of the loops $f : S^1 \to \mathbf{g}$ such that

$$f(\epsilon^{-1} z) = \sigma f(z).$$

Clearly $L_\sigma \mathbf{g} \subset L\mathbf{g}$ is a linear subspace and

$$\begin{aligned} [f, g](\epsilon_{-1} z) &= [f(\epsilon^{-1} z), g(\epsilon^{-1} z)] = [\sigma f(z), \sigma g(z)] \\ &= \sigma[f(z), g(z)] = \sigma[f, g](z), \end{aligned}$$

so $L_\sigma \mathbf{g}$ is closed under commutation. We define

$$\hat{\mathbf{g}}(\sigma) = S^1_\sigma \mathbf{g} \oplus \mathbb{C}k \oplus \mathbb{C}d$$

as a vector space and we define the commutator by (A) and (C) as before. When $\sigma = 1$ we have $\hat{\mathbf{g}}(\sigma) = \hat{\mathbf{g}}$.

Example 5.3.1. Let $\mathbf{g} = A_2$. Define $\sigma : \mathbf{g} \to \mathbf{g}$ by

$$\sigma(e_{12}) = e_{23},\ \sigma(e_{23}) = e_{12},\ \sigma(e_{31}) = -e_{31}.$$

From the commutation relations follows then that $\sigma(e_{13}) = -e_{13}$, $\sigma(e_{32}) = e_{21}$, $\sigma(e_{21}) = e_{32}$, $\sigma(e_{11} - e_{22}) = e_{22} - e_{33}$, and $\sigma(e_{22} - e_{33}) = e_{11} - e_{22}$. Now $\sigma^2 = 1$ and $\epsilon = -1$. A basis for the polynomial loops in $S^1_\sigma \mathbf{g}$ is defined by

$$(e_{11} - e_{33})z^{2n},\ (e_{12} + e_{23})z^{2n},\ (e_{21} + e_{32})z^{2n},\ e_{13}z^{2n+1}$$
$$e_{31}z^{2n+1},\ (e_{12} - e_{23})z^{2n+1},\ (e_{21} - e_{32})z^{2n+1},\ (e_{11} + e_{33} - 2e_{22})z^{2n+1},$$

where $n \in \mathbb{Z}$. The coefficients of z^{2n} span the eigenspace $\mathbf{g}(1) \subset \mathbf{g}$ corresponding to the eigenvalue $+1$ of σ and the coefficients of z^{2n+1} correspond to the eigenvalue -1. A Cartan subalgebra of $\hat{\mathbf{g}}(\sigma)$ is spanned by the vectors k, d, and $h = e_{11} - e_{33}$. In the ordered basis $\{h, k, d\}$ the nonzero roots are

$$\{(0,0,n) \mid 0 \neq n \in \mathbb{Z}\} \cup \{(\pm 1, 0, n) \mid n \in \mathbb{Z}\} \cup \{\pm 2, 0, 2n+1) \mid n \in \mathbb{Z}\}.$$

A system of simple roots is then

$$\widehat{\Delta}(\sigma) = \{(1,0,0), (-2,0,1)\} = \{\alpha_0, \alpha_1\}$$

and the positive roots are $\{(1,0,n-1), (-2,0,2n-1), (0,0,n) \mid n > 0\}$. Now $\langle \alpha_0, \alpha_1 \rangle = -1$ and $\langle \alpha_1, \alpha_0 \rangle = -4$ so that the Dynkin diagram is $A_2^{(2)}$ in the list Aff2.

In general, given an automorphism $\sigma : \mathbf{g} \to \mathbf{g}$ with $\sigma^N = 1$ (N minimal) one can write $\mathbf{g}$ as direct sum of eigenspaces

$$\mathbf{g} = \bigoplus_{j=0}^{N-1} \mathbf{g}(\epsilon^j).$$

Since $[\mathbf{g}(\epsilon^j), \mathbf{g}(\epsilon^i)] \subset \mathbf{g}(\epsilon^{i+j})$, only the subspace $\mathbf{g}(1)$ is a subalgebra. In the above example, $\mathbf{g}(1) \cong A_1$. One has a grading for $L_\sigma \mathbf{g}$,

$$L_\sigma \mathbf{g} = \bigoplus_{j=0}^{N-1} (\mathbf{g}(\epsilon^j) \otimes V_j(z)),$$

where $V_j(z)$ consists of linear combinations of the monomials z^{nN-j}, $n \in \mathbb{Z}$. The Cartan subalgebra of $\hat{\mathbf{g}}(\sigma)$ consists of the Cartan subalgebra of $\mathbf{g}(1)$ and the elements k and d. One can show that with respect to this Cartan subalgebra a system of simple roots of $\hat{\mathbf{g}}(\sigma)$ consists of the roots $(\alpha, 0, 0)$, where α goes through the simple roots of $\mathbf{g}(1)$, and the root $(-\psi, 0, 1)$, where ψ is a certain root of $\mathbf{g}(1)$. We are not going to study the twisted algebras in detail; see [Kac, 1985] for more information.

Exercise 5.3.2. The Dynkin diagram of D_4 is

○

○

○ ○

where the three external dots are connected to the central dot by simple lines (not visible in this TeX version!). Rotations by the angles $k \cdot 2\pi/3$ are symmetries of the diagram. Corresponding to the rotation $2\pi/3$ construct an automorphism of D_4 which permutes the root subspaces $\mathbf{g}_{\alpha_1}, \mathbf{g}_{\alpha_1}$, and $\mathbf{g}_{\alpha_3}$. Construct the affine Lie algebra $D_4^{(3)}$ using this automorphism (of order 3). Show that the Dynkin diagram is $D_4^{(3)}$ in the list Aff2.

5.4. The highest weight representations of affine Lie algebras

Let $\mathbf{a}$ be an affine Lie algebra, $\mathbf{h} \subset \mathbf{a}$ a Cartan subalgebra, $\Delta \subset \mathbf{h}^*$ a system of simple roots, and $\Phi^+ \supset \Delta$ the set of positive roots. There is a splitting

$$\mathbf{a} = \mathbf{n}_- \oplus \mathbf{h} \oplus \mathbf{n}_+,$$

where the subalgebra $\mathbf{n}_+$ (respectively, $\mathbf{n}_-$) is spanned by the root subspaces $\mathbf{a}_\alpha$ corresponding to positive (respectively, negative) roots. Let $\lambda \in \mathbf{h}^*$ be arbitrary and define the *Verma module* as in the finite-dimensional case,

$$V_\lambda = \mathcal{U}(\mathbf{a})/I_\lambda,$$

where the left ideal is generated by $\mathbf{n}_+$ and the elements $h - \lambda(h)$, $h \in \mathbf{h}$. As in Section 4.5, the space V_λ is a direct sum of its weight subspaces $V_\lambda(\mu)$; this and the other assertions of Theorem 4.5.1 are proved exactly in the same way as for a finite-dimensional semisimple Lie algebra:

Theorem 5.4.1. *The Verma module V_λ contains a unique maximal proper submodule M_λ (i.e., a proper invariant subspace $M \subset V_\lambda$ such that if $M' \subset V_\lambda$ is an invariant subspace containing M, then $M' = M$ or $M' = V_\lambda$) and $L_\lambda = V_\lambda/M_\lambda$ carries an irreducible highest weight representation of* $\mathbf{a}$ *with the highest weight = λ.*

Before studying the irreducible modules L_λ in more detail, we need some more information about the structure of affine Lie algebras. Let A be a linear operator in a vector space V. We say that A is *locally nilpotent* if for any $x \in V$ there is an integer $n = n(x) \in \mathbb{N}$ such that $A^n x = 0$. Let $0 \neq e_i \in \mathbf{a}_{\alpha_i}$ and $0 \neq f_i \in \mathbf{a}_{-\alpha_i}$ for $i = 0, 1, \ldots, l$, where $\{\alpha_0, \alpha_1, \ldots, \alpha_l\}$ is a set of simple roots; we shall normalize the vectors such that $[e_i, f_i] = h_{\alpha_i}$, $(e_i, f_i) = 1$. In the case of a finite-dimensional semisimple Lie algebra it is obvious that the operators ad_{e_i} and ad_{f_i} are locally nilpotent. By inspecting the root systems of the untwisted affine Lie algebras one can see that if β is a root then $\beta + n\alpha_i$ is a root only for *finitely* many values of $n \in \mathbb{Z}$. We state without proof that the same remains true for the twisted algebras. In conclusion:

Theorem 5.4.2. *The operators ad_{e_i} and ad_{f_i} are locally nilpotent in any affine Lie algebra.*

In general, we call an $\mathbf{a}$-module V *integrable*, if e_i and f_i are locally nilpotent for $0 \leq i \leq l$ and if V is a direct sum of weight subspaces. In particular by 5.4.2 the space $\mathbf{a}$ considered as an $\mathbf{a}$-module through the adjoint action is an integrable $\mathbf{a}$-module.

Theorem 5.4.3. *Let V be an integrable* $\mathbf{a}$*-module. If λ is a weight of V and if $\lambda + \alpha_i$ (respectively, $\lambda - \alpha_i$) is not a weight of V, then $(\lambda, \alpha_i) \geq 0$ [respectively, $(\lambda, \alpha_i) \leq 0$]. If λ is any weight of V, then $\lambda' = \lambda - \langle \lambda, \alpha_i \rangle \alpha_i$ is also a weight and $dim V(\lambda) = dim V(\lambda')$.*

Proof. In the finite-dimensional case we proved that if α is any root, then the vectors x_α, y_α, and h_α span a subalgebra isomorphic to A_1. From our construction of the root systems in the untwisted case it is not difficult to see that the same holds for the simple roots of an affine Lie algebra. [It is *not* true for the nonsimple roots $(0, 0, n)$.] One can show that this result is valid also for the twisted affine algebras.

For any fixed i, let A_1 be the Lie algebra spanned by e_i, f_i, and h_{α_i}. Let $0 \neq v$ be a vector of weight λ in V. Because V is integrable, $\mathcal{U}(A_1)v$ is a finite-dimensional A_1-module. If $\lambda + \alpha_i$ is not a weight, then $e_i v = 0$ and so $\langle \lambda, \alpha_i \rangle$ is a non-negative integer by our earlier analysis of A_1-modules in Section 4.2. If $\lambda - \alpha_i$ is not a weight, then $f_i v = 0$ and so v is the lowest weight vector for a finite-dimensional A_1-module. The lowest weight of an A_1-module is minus the highest weight; thus in this case $\langle \lambda, \alpha_i \rangle \leq 0$ and $(\lambda, \alpha_i) \leq 0$. If $0 \neq v \in V(\lambda)$, then

$$h_{\alpha_i} v = \lambda(h_{\alpha_i})v = (\lambda, \alpha_i)v$$

and similarly $h_{\alpha_i} v' = (\lambda', \alpha_i)v'$ if there is $0 \neq v' \in V(\lambda')$. But

$$(\lambda', \alpha_i) = (\lambda, \alpha_i) - \langle \lambda, \alpha_i \rangle (\alpha_i, \alpha_i) = -(\lambda, \alpha_i).$$

Since in a finite-dimensional A_1-module the weights appear symmetrically (i.e., μ is a weight iff $-\mu$ is a weight) we can conclude that also λ' is a weight.

As in the case of semisimple Lie algebras, for each $0 \leq i \leq l$ we define the linear map

$$\sigma_i : \mathbf{h}^* \to \mathbf{h}^*, \ \sigma_i(\lambda) = \lambda - \langle \lambda, \alpha_i \rangle \alpha_i.$$

Let $W = W(\mathbf{a}, \mathbf{h})$ be the group generated by the *fundamental reflections* σ_i; W is called the *Weyl group* of $(\mathbf{a}, \mathbf{h})$. Note that $< \lambda, \alpha_i >= 2(\lambda, \alpha_i)/(\alpha_i, \alpha_i)$ is well-defined for the simple roots because of $(\alpha_i, \alpha_i) \neq 0$. In the case of a finite-dimensional semisimple algebra the Weyl group can equivalently be defined as the group generated by *all* reflections σ_α, corresponding to an arbitrary nonzero root, because in that case the inner product is positive definite. From the Theorem 5.4.3 follows immediately that the Weyl group maps in an integrable representation the weight system onto itself. In particular, the set of roots Φ is mapped onto itself by W as a consequence of the fact that the adjoint representation is integrable. As in the finite-dimensional case, we define for the affine algebras

$$\Lambda = \{\lambda \in \mathbf{h}^* \mid < \lambda, \alpha_i > \in \mathbb{Z} \, \forall i\}$$

$$\Lambda^+ = \{\lambda \in \Lambda \mid < \lambda, \alpha_i > \geq 0 \, \forall i\}.$$

Let $\lambda \in \Lambda^+$. Using the fact that $(\alpha_0, \alpha_0) = \psi^2$ we observe that $\lambda(k) = \frac{\psi^2}{2} x$, where x is a positive integer called the *level* of λ. Note that $\alpha_0(d) = 1$ and

$$\lambda(k) = \frac{\psi^2}{2}(x' + < \lambda, \psi >) = \frac{\psi^2}{2} x$$

with

$$< \lambda, \alpha_0 >= \frac{2}{\psi^2}(\lambda, \alpha_0) = \frac{2}{\psi^2}[\lambda(k)\alpha_0(d) - (\lambda, \psi)] = x',$$

a nonnegative integer.

Theorem 5.4.4. *The irreducible highest weight module L_λ is integrable if and only if $\lambda \in \Lambda^+$.*

Proof. 1) Let L_λ be integrable and let $v \neq 0$ be the vector of highest weight. Then there exists a smallest non-negative integer n_i such that $(f_i)^{n_i+1}v = 0$. Consequently

$$0 = e_i(f_i)^{n_i+1}v = (n_i + 1)[\lambda(h_i) - \tfrac{1}{2}n_i\alpha_i(h_i)]f_i^{n_i}v,$$

where $h_i = [e_i, f_i] = \frac{2}{(\alpha_i,\alpha_i)}h_{\alpha_i}$. Thus

$$0 = \lambda(h_{\alpha_i}) - \frac{1}{2}n_i\alpha_i(h_{\alpha_i}) = (\lambda, \alpha_i) - \frac{1}{2}n_i(\alpha_i, \alpha_i),$$

so that $< \lambda, \alpha_i >= n_i$ is a non-negative integer.

2) Let $\lambda \in \Lambda^+$. By the same formula as above,

$$(f_i)^{<\lambda,\alpha_i>+1}v = 0, \ 0 \le i \le l.$$

Let U be the maximal subspace of L_λ where the action of $\mathbf{a}$ is locally nilpotent; $U \neq 0$ because of $v \in U$. We shall show that U is invariant under the action of $\mathbf{a}$. Let $u \in U$ and $x \in \mathbf{a}$. Now for any $y \in \mathbf{a}$,

$$y^n xu = \sum_{j=0}^{n} \binom{n}{j} [(\mathrm{ad}_y)^j x] y^{n-j} u,$$

which is proven by induction on n. For large enough j, $(\mathrm{ad}_y)^j x = 0$ for $y = e_i$ or $y = f_i$. On the other hand, $y^{n-j}u = 0$ for large enough $n - j$ when $y = e_i, f_i$. Thus it follows that $y^n xu = 0$ for some n, when $y = e_i$ or $y = f_i$. Because of the irreducibility of L_λ we must have $U = L_\lambda$.

Let $\lambda, \mu \in \mathbf{h}^*$ and $\lambda' = \lambda - < \lambda, \alpha > \alpha$, $\mu' = \mu - < \mu, \alpha > \alpha$, where α is any simple root. Then

$$\begin{aligned}(\lambda', \mu') =&(\lambda - < \lambda, \alpha > \alpha, \mu - < \mu, \alpha > \alpha) = (\lambda, \mu) - (\lambda, \alpha) < \mu, \alpha > \\ &- < \lambda, \alpha > (\alpha, \mu) + < \lambda, \alpha >< \mu, \alpha > (\alpha, \alpha) = (\lambda, \mu)\end{aligned}$$

by using $< \lambda, \alpha > (\alpha, \alpha) = 2(\lambda, \alpha)$. Thus the inner product in $\mathbf{h}^*$ is *invariant under the action of the Weyl group.* As a consequence, also the brackets $< \lambda, \alpha >$ are invariant under W.

Lemma 5.4.5. *Let $\lambda \in \Lambda_+$ and let μ be a weight of L_λ. Then $(\lambda, \lambda - \mu) \geq 0$ and the equality holds if and only if $\lambda = \mu$.*

Proof. Let $\lambda \neq \mu$. Let $\mathbf{m}$ be the subalgebra of $\mathbf{n}_-$ generated by those elements f_i for which $< \lambda - \mu, \alpha_i > \neq 0$ (denote this set of indices i by S). Now

$$L_\lambda(\mu) \subset \mathcal{U}(\mathbf{n}_-)\mathbf{m}v$$

where $v \neq 0$ is the highest weight vector. We can write $\lambda - \mu = \sum n_j \alpha_j$, where the n_j's are non-negative integers. By 5.4.3, $(\lambda, \alpha_i) \neq 0$ and $n_i > 0$ at least for one index $i \in S$ (otherwise $\mathbf{m}v = 0$ and thus $L_\lambda(\mu) = 0$). Now $(\lambda, \lambda - \mu) = \sum (\lambda, \alpha_j) n_j$. Each term is non-negative and in the case $\mu \neq \lambda$ at least one is positive.

Lemma 5.4.6. *Let $\lambda \in \Lambda^+$ and let μ be a weight of L_λ. Then there is $w \in W$ such that $w \cdot \mu \in \Lambda^+$.*

Proof. Writing $\mu = \sum k_i \alpha_i$ we set $\mathrm{ht}\mu = \sum k_i$. Choose $w \in W$ such that $\mathrm{ht}(\lambda - w \cdot \mu)$ is minimal. Now $< w \cdot \mu, \alpha_i > \geq 0$; otherwise $\mathrm{ht}(\lambda - \sigma_{\alpha_i} w \cdot \mu) < \mathrm{ht}(\lambda - w \cdot \mu)$.

We define $\rho \in \mathbf{h}^*$ by $\rho(h_{\alpha_i}) = \frac{1}{2}(\alpha_i, \alpha_i)$, $0 \leq i \leq l$ and $\rho(d) = 0$.

Proposition 5.4.7. *Let $\lambda \in \Lambda^+$ and let μ, ν be weights of L_λ. Then*

(1) $(\lambda, \lambda) - (\mu, \nu) \geq 0$*; the equality holds if and only if $\mu = \nu$ and $\mu \in W \cdot \lambda$.*
(2) $|\lambda + \rho|^2 - |\mu + \rho|^2 \geq 0$*; the equality holds only if $\mu = \lambda$.*

Proof. (1) Using the invariance of the inner product under the Weyl group action and Lemma 5.4.6 we can assume that $\mu \in \Lambda^+$. We can write $(\lambda, \lambda) - (\mu, \nu) = (\lambda, \lambda - \mu) + (\mu, \lambda - \nu)$, both terms being non-negative (see the proof of Lemma 5.4.5). If the equality holds then $(\lambda, \lambda - \mu) = 0 = (\mu, \lambda - \nu)$ and from 5.4.5 follows that $\lambda = \mu$ and thus also $\mu = \nu$.

(2) Write

$$(\lambda + \rho, \lambda + \rho) - (\mu + \rho, \mu + \rho) = [(\lambda, \lambda) - (\mu, \mu)] + 2(\rho, \lambda - \mu).$$

The first term is non-negative by (1) and the second by the definition of ρ and the fact that $\lambda - \mu$ is a nonnegative linear combination of the α_i's. Since $(\rho, \alpha_i) = \frac{1}{2}(\alpha_i, \alpha_i) > 0 \, \forall i$, in the case of equality sign we must have $(\lambda - \mu, \alpha_i) = 0 \, \forall i$.

There is one more property of the Weyl group which we shall need in the next section but which we state without proof:

Lemma 5.4.8. *Let $w \in W$ and $\lambda \in \Lambda^+$ such that $(\lambda, \alpha) > 0$ for all $\alpha \in \Delta$. Then $w\lambda = \lambda$ implies $w = 1$.*

We shall define *an antilinear antiautomorphism* θ of $\mathbf{a}$ by

$$\theta(e_i) = f_i,\ \theta(f_i) = e_i,\ \theta(h_{\alpha_i}) = h_{\alpha_i},\ \theta(d) = d,\ \theta(k) = k$$

for all $0 \leq i \leq l$. These relations determine θ uniquely, since all the vectors corresponding to positive (respectively, negative) roots are obtained by taking commutators of the elements e_i (respectively, f_i), and the vectors h_{α_i} form a basis of $\mathbf{h}$. Antilinearity means that $\theta(ax + by) = \overline{a}\theta(x) + \bar{b}\theta(y)$ for $x, y \in \mathbf{a}$ and $a, b \in \mathbb{C}$ and the antiautomorphism property is $\theta([x, y]) = -[\theta(x), \theta(y)]$. The antiautomorphism θ can be extended to an antilinear antiautomorphism of the enveloping algebra of $\mathbf{a}$ by setting $\theta(x_1 x_2 \ldots x_n) = \theta(x_n) \ldots \theta(x_2)\theta(x_1)$, $x_i \in \mathbf{a}$. It satisfies $\theta(uv) = \theta(v)\theta(u)$ for $u, v \in \mathcal{U}(\mathbf{a})$.

Exercise 5.4.9. Show that θ is really an antiautomorphism.

Example 5.4.10. Let $\mathbf{a} = A_l^{(1)}$. A basis for $\mathbf{a}$ is given by the following elements: $e_{ij}z^n$, $n \in \mathbb{Z}$ and $1 \leq i \neq j \leq l+1$; $h_i z^n$, $n \in \mathbb{Z}$ and $1 \leq i \leq l$ with $h_i = e_{ii} - e_{i+1,i+1}$; the elements d, k. Now

$$\begin{aligned}
\theta(e_{ij}z^n) &= \frac{1}{2(l+1)} e_{ji} z^{-n},\ i < j \\
\theta(e_{ji}z^n) &= 2(l+1) e_{ij} z^{-n},\ i < j \\
\theta(h_i z^n) &= h_i z^{-n},\ \theta(d) = d,\ \theta(k) = k.
\end{aligned}$$

Note that the restriction of θ to root subspaces gives a linear isomorphism $\theta : \mathbf{g}_\alpha \to \mathbf{g}_{-\alpha}$. On the other hand, from the defining formula (B1) we observe that the restriction $\mathbf{g}_\alpha \times \mathbf{g}_{-\alpha} \to \mathbb{C}$ of the invariant bilinear form is nondegenerate; it is also positive definite in the sense that

$$(x, \theta(x)) \geq 0\ \forall x \in \mathbf{g}_\alpha.$$

It follows that we can define a basis $\{x_\alpha^{(i)}\}$ in $\mathbf{g}_\alpha$ for each $\alpha \in \Phi^-$ such that $(x_\alpha^{(i)}, \theta(x_\alpha^{(j)})) = \delta_{ij}$. The multiplicity label i is really necessary only for the roots $(0, 0, n)$; see Section 5.2. We set $x_\alpha^{(i)} = \theta(x_{-\alpha}^{(i)})$, $\alpha \in \Phi^+$. Fix also a basis $\{h^i\}$ of $\mathbf{h}$ dual to the basis $\{h^i\}$, $(h_i, h^j) = \delta_{ij}$.

In the case of a finite-dimensional semisimple Lie algebra one defines a *Casimir operator* Ω' by

$$\Omega' = \sum h_i h^i + \sum_{\alpha \in \Phi^+} (x_\alpha x_{-\alpha} + x_{-\alpha} x_\alpha).$$

No multiplicity label is needed here because the root subspaces are one-dimensional; compare with c in Section 4.3! In the infinite-dimensional case we cannot use this formula because the sum will in general diverge. However, we can apply a "normal ordering" prescription to make the sum finite. We set

$$\Omega = \sum h_i h^i + 2 \sum_{\alpha \in \Phi^+} \sum_i x_{-\alpha}^{(i)} x_\alpha^{(i)} + 2h_\rho.$$

Ω is a well-defined linear operator in any highest weight representation of $\mathbf{a}$. Namely, any vector in the representation space can be written as a polynomial in the generators of $\mathbf{n}_-$ acting on the highest weight vector. It follows that the action of the second term in Ω reduces to a finite polynomial.

Proposition 5.4.11. *The element $\Omega \in \mathcal{U}(\mathbf{a})$ commutes with $\mathbf{a}$, and thus the action in a highest weight representation reduces to a multiplication with a scalar. The value of the scalar is $|\lambda + \rho|^2 - |\rho|^2$, where λ is the highest weight.*

Proof. Denote by Ω_0 the part of Ω involving the x's. Let α, β be roots and $z \in \mathbf{a}_\beta$. Then ad_z maps $\mathbf{a}_\alpha$ into $\mathbf{a}_{\alpha+\beta}$ and $\mathbf{a}_{-\beta-\alpha}$ into $\mathbf{a}_{-\alpha}$. By the invariance of the bilinear form, $([z,x],y) = -(x,[z,y])$, the former map is (-1) times the transpose of the latter. Let now β be a simple root. We obtain

$$\begin{aligned}
[z, \Omega_0] &= 2 \sum_{\alpha \in \Phi^+, i} ([z, x_{-\alpha}^{(i)}] x_\alpha^{(i)} + x_{-\alpha}^{(i)} [z, x_\alpha^{(i)}]) \\
&= 2[z, x_{-\beta}] x_\beta + 2 \sum_{\beta \neq \alpha \in \Phi^+, i} ([z, x_{-\alpha}^{(i)}] x_\alpha^{(i)} + x_{-\alpha+\beta}^{(i)} [z, x_{\alpha-\beta}^{(i)}]),
\end{aligned}$$

where we have done a simple renaming of the summation index in the last term. We have dropped the multiplicity index in the first term, since the simple roots have multiplicity =1. By the remark above, the second and the third term cancel on the right-hand side. Thus we get

$$[z, \Omega_0] = 2[z, x_{-\beta}] x_\beta = 2(z, x_{-\beta}) h_\beta x_\beta = 2h_\beta z.$$

On the other hand,

$$\begin{aligned}
[z, \sum h_i h^i] &= -\sum \beta(h_i) z h^i - \sum h_i \beta(h^i) z \\
&= -2 \sum \beta(h_i) h^i z + \sum \beta(h_i) \beta(h^i) z = -2h_\beta z + (\beta, \beta) z.
\end{aligned}$$

Finally $[z, 2h_\rho] = -2\beta(h_\rho)z = -2(\beta,\rho)z = -(\beta,\beta)z$ and combining this with the results above we get $[z,\Omega] = 0$. In the same way one can show that $[z,\Omega] = 0$ when β is *minus* a simple root. Taking commutators of vectors belonging to simple roots or to minus simple roots one can generate the whole algebra $\mathbf{a}$. Thus $[z,\Omega] = 0$ for all $z \in \mathbf{a}$. Next we evaluate Ω by applying it to the highest weight vector v in a highest weight representation. We get

$$\Omega v = (\sum h_i h^i + 2h_\rho)v = [\lambda(h_i)\lambda(h^i) + 2\lambda(h_\rho)]v = [(\lambda,\lambda) + 2(\lambda,\rho)]v.$$

The coefficient in front of v is easily seen to be equal to $|\lambda+\rho|^2 - |\rho|^2$.

A Hermitian form H on a $\mathbf{a}$-module V is *contravariant* if

$$H(xu, v) = H(u, \theta(x)v),\ \forall u, v \in V,\ x \in \mathbf{a}.$$

We use the convention that a Hermitian form is linear in the first and antilinear in the second argument. If V is a highest weight module, we define a contravariant Hermitian form in V as follows. Let v be a highest weight vector (unique up to a multiplicative constant) and set $H(v, v) = 1$. If $v_1, v_2 \in V$ are arbitrary, we can write $v_i = u_i \cdot v$, where $u_i \in \mathcal{U}(\mathbf{n}_-)$. Define

$$H(v_1, v_2) = H(u_1 v, u_2 v) = H(v, \theta(u_1)u_2 v).$$

Next we can write $\theta(u_1)u_2 v = uv$ for some $u \in \mathcal{U}(\mathbf{n}_-)$. Now we have

$$H(v_1, v_2) = H(v, uv) = \overline{H(uv, v)} = \overline{H(v, \theta(u)v)} = H(\theta(u)v, v).$$

Since $\theta(u) \in \mathcal{U}(\mathbf{n}_+)$, we obtain $\theta(u)v = a \cdot v$ for some $a \in \mathbb{C}$. Thus the value $H(v_1, v_2) = a$ has been uniquely determined by the contravariantness of the Hermitian form and by the normalization $H(v, v) = 1$.

Theorem 5.4.12. *The Hermitian form H is positive definite in all integrable irreducible highest weight modules.*

Proof. From the definitions follows at once that the different weight subspaces $L_\lambda(\mu)$ in L_λ are pairwise orthogonal. Thus it is sufficient to show that the restriction of H to any of these subspaces is positive definite. We prove it by induction on

$n = \mathrm{ht}(\lambda - \mu)$. The case $n = 0$ is clear by $H(v,v) = 1$. Using Theorem 5.4.11 we get

$$
\begin{aligned}
(|\lambda+\rho|^2 - &|\rho|^2)H(w,w) \\
&= H(\Omega w, w) \\
&= (|\mu|^2 + 2\mu(h_\rho))H(w,w) + \sum_{\alpha\in\Phi^+, i} H(x_\alpha^{(i)} w, x_\alpha^{(i)} w),
\end{aligned}
$$

where we have also used $\theta(x_{-\alpha}^{(i)}) = x_\alpha^{(i)}$. If we subtract the first term on the right from the left-hand side we get $(|\lambda+\rho|^2 - |\mu+\rho|^2)H(w,w)$. The factor multiplying $H(w,w)$ is positive by (5.4.7) when $\mu \neq \lambda$. On the other hand, the height of the weight of $x_\alpha^{(i)} w$ is smaller than n and so by induction assumption each term in the sum on the right-hand side is also non-negative. To complete the proof we still have to show that the form H is nondegenerate. Because the representation is irreducible, we can choose $u_w \in \mathcal{U}(\mathbf{n}_+)$ such that $u_w \cdot w = v$. Now $H(w, \theta(u_w)v) = H(u_w \cdot w, v) = H(v,v) \neq 0$ and thus H is non-degenerate.

Exercise 5.4.13 Let H be the space of square-integrable functions $\psi : S^1 \to \mathbb{C}^N$. We write $H = H_+ \oplus H_-$, where H_+ is spanned by the Fourier modes $e^{in\phi}v$, with $n = 0, 1, 2, \ldots$ and $v \in \mathbb{C}^N$. The space H_- is the orthogonal complement of H_+, spanned by the negative Fourier modes. The inner product in H is defined by

$$(\psi, \psi') = \int_0^{2\pi} \sum_{i=1}^N \overline{\psi(\phi)}\psi(\phi) d\phi.$$

Let X be a smooth $N \times N$ traceless matrix valued function on S^1. Define the linear operator $T(X) : H \to H$ by $(T(X)\psi)(\phi) = X(\phi)\psi(\phi)$. Next introduce the CAR algebra generated by the standard generators $a_{n,i}$ and $a^*_{n,i}$ where $n \in \mathbb{Z}$ and $i = 1, 2, \ldots N$. (Compare with exercise 5.2.3.) Define the operators

$$\hat{X} = \sum_{m,n,i,j} X_{ij}(n) : a^*_{m+n,i} a_{m,j} :$$

where $X_{ij}(n)$ denotes the $n : th$ Fourier component of the matrix valued function $X = (X_{ij})$ and the normal ordering is defined with respect to the Fourier index. Following exercise 5.2.3, show that

$$[\hat{X}, \hat{Y}] = \widehat{[X,Y]} + \frac{1}{2\pi i}\int_0^{2\pi} \mathrm{tr}\, X(\phi)\frac{d}{d\phi}Y(\phi) d\phi.$$

In the Fock space where the CAR algebra is operating the vacuum vector v_0 is characterized by $a_{n,i}v_0 = 0 = a^*_{m,i}v_0$ for $n \geq 0$ and $m < 0$. Show that this vector is a *lowest weight vector* for the affine Kac-Moody algebra generated by the operators $\hat{X}$.

Exercise 5.4.14 We use the notation of exercise 5.2.2. The element ℓ_0 in the Virasoro algebra plays the role of a Cartan subalgebra. Consider a highest weight representation of the Virasoro algebra in a vector space V with a highest weight vector v_0 which has the property $\ell_n v_0 = 0$ for $n < 0$ and $\ell_0 v_0 = hv_0$ where h is a constant; the element c which commutes with everything is assumed to take a constant value in the whole representation. Show by PBW theorem that all the weight spaces $V_\lambda = \{v \in V | \ell_0 v = \lambda v\}$ are finite-dimensional and that $\lambda - h$ is a nonnegative integer when $V_\lambda \neq 0$. Assume that we have an inner product in V such that $\ell^*_n = \ell_{-n}$. Show that $h \geq 0$ and $c \geq 0$. Hint: Study the norm of the vector $\ell_n v_0$ for $n > 0$.

5.5. The character formula

If V carries a finite-dimensional representation T of a semisimple Lie group G one can define the *character* of the representation by

$$ch(g) = \operatorname{tr} T(g).$$

Thus the character is a complex valued function on G. Let H be a Cartan subgroup, $\mathbf{h}$ the corresponding Cartan subalgebra and denote by $V(\mu)$ the weight subspace belonging to the weight $\mu \in \mathbf{h}^*$. Then for $x \in \mathbf{h}$ and $h = e^x \in H$,

$$ch(h) = \sum_{\mu \in \Lambda} e^{\mu(x)} \cdot \dim V(\mu) \tag{5.5.1}$$

where the sum is over the set Λ of weights and $V(\lambda) \subset V$ are the weight subspaces.

In an infinite-dimensional case one has to proceed in a more formal way since the sum (5.5.1) does not converge in general. We can still define the *formal character* by

$$ch\, V = \sum_{\mu \in \Lambda} e(\mu) \cdot \dim V(\mu), \tag{5.5.2}$$

where the symbols $e(\mu)$ are now formal exponentials; they are the generators of a commutative algebra subject to the defining relations $e(\mu)\cdot e(\nu) = e(\mu+\nu)$. The element $e(0)$ is the neutral element with respect to multiplication and we write $e(0) = 1$. In this section we shall compute the formal characters of the highest weight representations of affine Lie algebras.

The formal characters of the Verma modules V_λ are easely computed. Let $x_{-\beta_i,p_i}$ be a basis of the root subspace $\mathbf{g}_{-\beta_i}$, where $1 \le p_i \le m(i) = mult\beta_i$ (the multiplicity of the root β_i) and $\{\beta_1, \beta_2, \dots, \beta_\ell\} = \Phi^+$ is the set of positive roots. Then the vectors

$$(x_{-\beta_1,1})^{n(1,1)}(x_{-\beta_1,2})^{n(1,2)}\dots$$
$$\dots(x_{-\beta_1,m(1)})^{n(1,m(1))}\dots(x_{-\beta_\ell,m(\ell)})^{n(\ell,m(\ell))}v$$

form a basis of the subspace $V_\lambda(\mu)$, where $[n(1,1)+\dots n(1,m(1))]\beta_1 + [n(2,1)+\dots n(2,m(2))]\beta_2+\dots+[n(\ell,1)+\dots+n(\ell,m(\ell))]\beta_\ell = \lambda-\mu$ and $n(i,j)$'s are non-negative integers. Thus

$$\begin{aligned} ch\, V_\lambda &= e(\lambda)\prod_{\beta\in\Phi^+}[1+e(-\beta)+e(-2\beta)+\dots]^{mult\beta} \\ &= e(\lambda)\prod_{\beta\in\Phi^+}(1-e(-\beta))^{-mult\beta}. \end{aligned} \tag{5.5.3}$$

If V and V' are a pair of modules for a given Lie algebra and W is a submodule of V then

$$ch\, V/W = ch\, V - ch\, W,\;\; ch(V\oplus V') = ch\, V + ch\, V'.$$

A highest weight module can always be thought as a quotient of a Verma module by a submodule; for this reason it is natural that the character of a highest weight module can be expanded as

$$ch\, V = \sum_\lambda c(\lambda) ch\, V_\lambda \tag{5.5.4}$$

where the $c(\lambda)$'s are integers. The proof is not completely trivial but we shall skip it here because it is not very illuminating [Kac, 1985]. Taking account that the value of the Casimir operator in a highest weight module with highest weight λ is equal to $|\lambda+\rho|^2 - |\rho|^2$ one can show that the only possible nonzero terms in the sum above are those which satisfy $|\lambda+\rho|^2 = |\Lambda+\rho|^2$ and $\lambda \leq \Lambda$ where Λ is the highest weight of V.

Next we let the Weyl group W act on the formal exponentials by $w[e(\lambda)] = e(w(\lambda))$. From Theorem 5.4.3 follows that

$$ch\, L_\lambda = w(ch\, L_\lambda)\ \forall w \in W. \tag{5.5.5}$$

For any $w \in W$ we can write $w = \sigma_1\sigma_2\ldots\sigma_s$ where σ_i is the fundamental reflection in the plane orthogonal to the simple root α_i, $1 \leq i \leq \ell$. Clearly the determinant of the linear transformation $\sigma_i : \mathbf{h}^* \to \mathbf{h}^*$ is equal to -1 and therefore the determinant of w is $\epsilon(w) = (-1)^s$. Define the formal character

$$R = \prod_{\alpha\in\Phi^+}[1-e(-\alpha)]^{mult\,\alpha}.$$

We shall need the following fact: The action of a fundamental reflection σ_i in $\Phi^+ \setminus \{\alpha_i\}$ permutes the elements among themselves. This is a consequence of the fact that any positive root is a sum of simple roots and that $\langle\alpha_j,\alpha_i\rangle \leq 0$ for $i \neq j$..

Lemma 5.5.6. *$w[e(\rho)R] = \epsilon(w)e(\rho)R$ for all $w \in W$.*

Proof. It is sufficient to prove the lemma in the case $w = \sigma_i$ for any i. Now $mult\alpha = mult\, w(\alpha)$ for any $\alpha \in \Phi^+$ and $\Phi^+ \setminus \{\alpha_i\}$ is invariant under w. Therefore,

$$\begin{aligned} w[e(\rho)R] &= e(\rho-\alpha_i)[1-e(\alpha_i)]\sigma_i \prod_{\alpha\in\Phi^+\setminus\{\alpha_i\}}[1-e(-\alpha)]^{mult\alpha} \\ &= e(\rho)e(-\alpha_i)[1-e(\alpha_i)] \prod_{\alpha\in\Phi^+\setminus\{\alpha_i\}}[1-e(-\alpha)]^{mult\alpha} \\ &= -e(\rho)R = \epsilon(w)e(\rho)R. \end{aligned}$$

Theorem 5.5.7. *Let $\lambda \in \Lambda^+$ and L_λ the irreducible module for an affine Lie algebra with highest weight λ. Then*

$$ch\, L_\lambda = \frac{\sum_{w\in W}\epsilon(w)e[w(\lambda+\rho)-\rho]}{\prod_{\alpha\in\Phi^+}[1-e(-\alpha)]^{mult\alpha}}.$$

Proof. From (5.5.3) and (5.5.4) we obtain

$$e(\rho)\, R\, ch\, L_\lambda = \sum_{\mu\in B} c(\mu)e(\mu+\rho)$$

where B is the set of weights consisting of those $\mu \in \Lambda$ for which $\mu \leq \lambda$ and $|\mu+\rho|^2 = |\lambda+\rho|^2$. From (5.5.5) and (5.5.6) follows that

$$c(\mu) = \epsilon(w)c(\nu) \text{ if } w(\mu+\rho) = \nu+\rho$$

for some $w \in W$. It follows that $c(\mu) \neq 0$ if and only if $c(w(\mu+\rho)-\rho) \neq 0$ and so $w(\mu+\rho) \leq \lambda+\rho$ if $c(\mu) \neq 0$. Assuming $c(\mu) \neq 0$ choose a weight $\nu \in \{w(\mu+\rho)-\rho \mid w \in W\}$ such that $\mathrm{ht}(\lambda-\nu)$ is minimal. Then $\nu+\rho \in \Lambda_+$ and $|\nu+\rho|^2 = |\mu+\rho|^2 = |\lambda+\rho|^2$. Applying 5.4.7 we conclude that $\nu = \lambda$ and therefore $w(\mu+\rho) = \lambda+\rho$. Thus $c(\mu) = \epsilon(w^{-1}) = \epsilon(w)$.

Since $(\lambda+\rho, \alpha) > 0$ for all $\alpha \in \Delta$ we get from 5.4.8 that $w(\lambda+\rho) = \lambda+\rho$ only if $w = 1$. Clearly $c(\lambda) = 1$ and therefore we have

$$e(\rho)R\, ch\, L_\lambda = \sum_{w\in W} \epsilon(w)e(w(\lambda+\rho)),$$

which gives the asserted formula for $ch\, L_\lambda$.

If $\lambda = 0$ then L_λ is the trivial one-dimensional representation and so $ch\, L_0 = e(0) = 1$. From the character formula we obtain the identity

$$\prod_{\alpha\in\Phi^+} [1-e(-\alpha)]^{mult\alpha} = \sum_{w\in W} \epsilon(w)e(w(\rho)-\rho). \tag{5.5.8}$$

We can now write 5.5.7 alternatively as

$$ch\, L_\lambda = \frac{\sum_{w\in W} \epsilon(w)e[w(\lambda+\rho)-\rho]}{\sum_{w\in W} \epsilon(w)e(w(\rho)-\rho)}. \tag{5.5.9}$$

In the case of a finite-dimensional semisimple Lie algebra this is the classical Weyl character formula.

In the finite-dimensional case the multiplicities of the weights can be also obtained from the *Kostant multiplicity formula*

$$\dim L_\lambda(\mu) = \sum_{w\in W} \epsilon(w)K[(\mu+\rho)-w(\lambda+\rho)] \tag{5.5.10}$$

where K is the Kostant partition function obtained from the expansion

$$\prod_{\alpha\in\Phi^+} [1-e(-\alpha)]^{-mult\alpha} = \sum_{\beta\in\mathbf{h}^*} K(\beta)e(\beta). \tag{5.5.11}$$

Expanding $[1-e(-\alpha)]^{-1}$ as a power series we can write the left-hand side of (5.5.11) also as

$$\prod_{\alpha\in\Phi^+}[1+e(-\alpha)+e(-2\alpha)+\dots]^{mult\alpha}$$

and therefore $K(\beta)$ is equal to the number of partitions of β into a sum of negative roots, where each root is counted as many times as is its multiplicity. Clearly $K(0)=1$ and in general, $K(\beta)=\dim V_\lambda(\lambda+\beta)$ according to (5.5.3).

Exercise 5.5.12. Prove the formula (5.5.10) in the case of an affine Lie algebra starting from 5.5.7 and the definition (5.5.11).

We define a homomorphism F from the polynomial algebra generated by the formal exponentials $e(-\alpha)$, $\alpha\in\Delta$, to the polynomial algebra in one variable q by

$$F(e(-\alpha))=q,\ \forall\alpha\in\Delta.$$

Since all weights of L_λ are of the form λ minus a sum of simple roots we can define the formal power series $\dim_q L_\lambda = F(e(-\lambda)ch\, L_\lambda)$. The coefficient of the monomial q^n is equal to the sum of the dimensions $\dim L_\lambda(\mu)$ where $\mathrm{ht}(\lambda-\mu)=n$, where ht is defined as in the proof of 5.4.6.

Let $a_{ij}=2\frac{(\alpha_i,\alpha_j)}{(\alpha_j,\alpha_j)}$ be the Cartan matrix of an affine Lie algebra. The transposed matrix $b_{ij}=a_{ji}$ defines also an affine Lie algebra. The simple roots of the transposed algebra $\mathbf{g}^t$ are $\beta_i=2\alpha_i/(\alpha_i,\alpha_i)$. Let $\rho^*\in\mathbf{h}^*$ be the weight such that $(\alpha_i,\rho^*)=1$ for all simple roots. Then ρ^*, considered as a weight for $\mathbf{g}^t$, corresponds to the weight ρ of $\mathbf{g}$. Let Φ^* be the set of roots for the transposed Lie algebra.

Theorem 5.5.13.

$$dim_q L_\lambda = \prod_{\alpha\in\Phi^{*+}}\left(\frac{1-q^{(\lambda+\rho,\alpha)}}{1-q^{(\rho,\alpha)}}\right)^{mult\alpha}$$

Proof. For any positive dominant weight μ define

$$N(\mu)=\sum_{w\in W}\epsilon(w)e(w(\mu)-\mu).$$

Now $\mathrm{ht}(\mu-w(\mu))=(\mu-w(\mu),\rho^*)$ and so

$$F(e(w(\mu)-\mu))=q^{(\mu-w(\mu),\rho^*)}. \tag{5.5.14}$$

Applying the homomorphism F to both sides of (5.5.8) we get

$$\prod_{\alpha\in\Phi^+}(1-q^{(\alpha,\rho^*)})^{mult\alpha}=\sum_{w\in W}\epsilon(w)q^{(\rho-w(\rho),\rho^*)}$$

and combining this with (5.5.14) we get

$$\begin{aligned}F(N(\mu))&=\sum_{w\in W}\epsilon(w)q^{(\mu-w(\mu),\rho^*)}\\&=\sum_{w\in W}\epsilon(w)q^{(\mu,\rho^*-w(\rho^*))}\\&=F'\left(\sum_{w\in W}\epsilon(w)e(w(\rho^*)-\rho^*)\right),\end{aligned}$$

where the homomorphism F' is defined by the relations $F'(e(-\alpha))=q^{(\mu,\alpha^*)}$ for $\alpha\in\Delta$ with $\alpha^*=2\alpha/(\alpha,\alpha)$. Applying the identity (5.5.8) to the transposed Lie algebra we get

$$F(N(\mu))=F'\left(\prod_{\alpha\in\Phi^*_+}(1-e(-\alpha))^{mult\,\alpha}\right)$$

where Φ^* is the root system of the transposed algebra. Thus

$$F(N(\mu))=\prod_{\alpha\in\Phi^*_+}(1-q^{(\mu,\alpha)})^{mult\,\alpha}.$$

Combining this with (5.5.9) we obtain

$$F(e(-\lambda)ch\,L(\lambda))=\prod_{\alpha\in\Phi^*_+}\left(\frac{1-q^{(\lambda+\rho,\alpha)}}{1-q^{(\rho,\alpha)}}\right)^{mult\,\alpha}$$

which implies the theorem.

Exercise 5.5.15 The analytic character of the group $SU(2)$ is obtained from the formal character $ch\,V(\lambda)=e(-\lambda)+e(-\lambda+2)+\dots e(\lambda)$ by replacing the formal exponential $e(\mu)$ by the exponential function $e^{i\mu x}$. Prove directly from this analytic formula the decomposition formula for tensor products, $V(\lambda)\otimes V(\lambda')=V(|\lambda-\lambda'|)\oplus\cdots\oplus V(\lambda+\lambda'-2)\oplus V(\lambda+\lambda')$.

Exercise 5.5.16 Apply the Kostant multiplicity formula to the case of the finite-dimensional simple Lie algebra A_2. In particular, compute the the weight multiplicities in the case of the finite-dimensional irreducible module of highest weight $\lambda=2\lambda_1+\lambda_2$, where λ_1,λ_2 are the fundamental weights of A_2.

Exercise 5.5.17 Apply the Theorem 5.5.13 to the case $\mathbf{g}=A_1^{(1)}$ and work out a more explicite expression for the q-character formula.

CHAPTER 6 QUANTUM GROUPS

6.1 Algebras, coalgebras, and Hopf algebras

Recall the definition of an associative algebra A : It is a vector space (over a field k) with a bilinear product map $m : A \times A \to A$ such that $m(a, m(b, c)) = m(m(a, b), c)$ for all $a, b, c \in A$. Most of the time we write $m(a, b) = a \cdot b = ab$.

Since m is linear in each argument we may as well think of m as a map

$$m : A \otimes A \to A.$$

If the algebra A has a unit 1 then $m(1, a) = m(a, 1) = a$ for all $a \in A$.

Next we define *a coalgebra.* A coalgebra is a vector space A with a linear map

$$\Delta : A \to A \otimes A,$$

called the *coproduct,* such that the coassociativity condition

$$(\Delta \otimes id) \circ \Delta = (id \otimes \Delta) \circ \Delta$$

is satisfied. Some words about notation. We can write

$$\Delta(a) = \sum_i a_i^{(1)} \otimes a_i^{(2)}$$

but often this is abreviated as

$$\Delta(a) = \sum_{(a)} a^{(1)} \otimes a^{(2)}$$

or

$$\Delta(a) = \sum_{(a)} a' \otimes a''.$$

A coalgebra A has a *counit* ϵ if $\epsilon : A \to k$ is a linear map with the property $(id \otimes \epsilon) \circ \Delta =$ the natural isomorphism $A \simeq A \otimes k$. Likewise, $(\epsilon \otimes id) \circ \Delta$ is the similar natural isomorphism $A \simeq k \otimes A$. Using the Sweedler's sigma notation,

$$\sum_{(a)} a'\epsilon(a'') = a = \sum_{(a)} \epsilon(a')a''$$

for $\Delta(a) = \sum a' \otimes a''$.

Using Sweedler's sigma notation the coassociativity can be written as

$$\sum_{(a)} \left(\sum_{(a')} (a')' \otimes (a')'' \right) \otimes a'' = \sum_{(a)} a' \otimes \left(\sum_{(a'')} (a'')' \otimes (a'')'' \right)$$

which we shall simply write as

$$\sum_{(a)} a' \otimes a'' \otimes a'''.$$

We can apply the coproduct once more to identify the following three expressions,

$$\sum_{(a)} \Delta(a') \otimes a'' \otimes a''', \ \sum_{(a)} a' \otimes \Delta(a'') \otimes a''', \ \sum_{(a)} a' \otimes a'' \otimes \Delta(a''')$$

which we agree to write as

$$\sum_{(a)} a' \otimes a'' \otimes a''' \otimes a''''$$

or as

$$\sum_{(a)} a^{(1)} \otimes a^{(2)} \otimes a^{(3)} \otimes a^{(4)}.$$

Example 6.1.1 Let $A = M_n(k)$ be the algebra of $n \times n$ matrices over k and let A^* be the dual vector space of A. Define the basis $x_{ij} : A \to k$, $x_{ij}(a) = a_{ij}$. The map $\Delta : A^* \to A^* \otimes A^*$ defined by

$$\Delta(x_{ij}) = \sum_k x_{ik} \otimes x_{kj}$$

satifies the coassociavity relation; this follows from the associativity of the matrix product rule $(ab)_{ij} = \sum_k a_{ik} b_{kj}$. Furthermore, there is a counit ϵ defined by $\epsilon(x_{ij}) = \delta_{ij}$.

Exercise 6.1.2 Show that the dual vector space A^* of any finite-dimensional associative algebra A is a coalgebra with a coproduct defined by

$$(\Delta(f))(a \otimes b) = f(ab) \text{ where } a, b \in A.$$

Hint: Use the isomorphism $A^* \otimes A^* \simeq (A \otimes A)^*$. Show that A^* has a counit if A has a unit.

The dual A^* of a coalgebra is always an algebra. The multiplication $m : A^* \otimes A^* \to A^*$ is defined by

$$m(a \otimes b)(x) = \sum_{(x)} a(x')b(x'')$$

for $x \in A$. If A has a counit ϵ then the unit in A^* is the map $f : A \to k$ given by $f(a) = \epsilon(a)$.

A coalgebra A is *cocommutative* if

$$\Delta(a) = \sum_{(a)} a' \otimes a'' = \sum_{(a)} a'' \otimes a'.$$

A linear map $\phi : A \to B$ of coalgebras is a homomorphism if $\Delta_B \circ \phi = (\phi \otimes \phi) \circ \Delta_A$ and if $\epsilon_A = \epsilon_B \circ \phi$.

For any coalgebra A there is the *opposite coalgebra* A^{op} defined by the opposite coproduct $\Delta^{op}(a) = \sum_{(a)} a'' \otimes a'$.

Next we define a *bialgebra.* A bialgebra is an algebra A with unit which in addition has a coalgebra structure, with a counit ϵ, such that

(1) the multiplication and the unit (viewed as a map $k \to A$) are homomorphisms of coalgebras
(2) the coproduct Δ is a homomorphism from the algebra A to the algebra $A \otimes A$ and the counit is an algebra homomorphism $A \to k$.

In particular, $\epsilon(\mathbf{1}) = 1 \in k$ by the algebra homomorphism property of the map $\epsilon : A \to k$.

Exercise 6.1.3 Show that the above two conditions are in fact equivalent.

Example 6.1.4 Let $M_n(k)$ be the polynomial algebra over k in the independent variables x_{ij} with $i, j = 1, 2, \dots, n$. Define the coproduct by

$$\Delta(x_{ij}) = \sum_k x_{ik} \otimes x_{kj}.$$

The counit is defined by $\epsilon(x_{ij}) = \delta_{ij}$. Then $M_n(k)$ is a bialgebra.

Example 6.1.5 Let G be a finite group and let A be the algebra of k-valued functions on G. The product of functions is defined as usual. In addition, we define a coproduct Δ by

$$(\Delta(f))(a, b) = f(ab),$$

where $a, b \in G$. We have identified the tensor product $A \otimes A$ as the the space of functions of two variables a, b. It is easy to see that the map from $A \otimes A$ to the space of functions in a, b defined by

$$\sum_i f_i \otimes g_i \mapsto f, \text{ with } f(a,b) = \sum_i f_i(a) g_i(b)$$

is an isomorphism. The counit is $\epsilon(f) = f(e)$, where $e \in G$ is the neutral element. Now A is a bialgebra.

Exercise 6.1.6 Check the bialgebra axioms in the above example.

Example 6.1.7 Let V be a vector space over k and let $T(V)$ be the tensor algebra over V. As a vector space $T(V)$ is a direct sum of vector spaces $V^n = V \otimes V \otimes \cdots \otimes V$ (n times) with $n = 0, 1, 2, \dots$. The product is defined by the tensor product, $(x_1 \otimes \cdots \otimes x_n) \cdot (x_{n+1} \otimes \cdots \otimes x_{n+m}) = x_1 \otimes \cdots \otimes x_{n+m}$. The algebra $T(V)$ is by construction generated by the elements in V and the unit element in $k = V^0$. The coproduct is then uniquely defined by

$$\Delta(v) = v \otimes' 1 + 1 \otimes' v \text{ for } v \in V$$

and by the requirement that the coproduct is an algebra homomorphism. For example,

$$\Delta(v \otimes w) = \Delta(v) \cdot \Delta(w) = (v \otimes w) \otimes' 1 + v \otimes' w + w \otimes' v + 1 \otimes' (v \otimes w) \in T(V) \otimes, T(V).$$

We have used $\otimes'$ sign for the tensor product $T(V) \otimes' T(V)$ and the unprimed tensor product for the product *inside* $T(V)$. The counit is the map $\epsilon : T(V) \to k$ defined by $\epsilon(u) = 0$ for $u \in V^n$ with $n > 0$ and $\epsilon(1) = 1$.

Let now $(A, m, 1, \Delta, \epsilon)$ be a bialgebra. We say that a linear map $S : A \to A$ is an *antipode* if

$$\sum_{(a)} a' S(a'') = \epsilon(a) \cdot 1 = \sum_{(a)} S(a') a''$$

for any $a \in A$ where $\Delta(a) = \sum_{(a)} a' \otimes a''$.

If a bialgebra has an antipode S then it is uniquely defined: Let S' be another antipode. Then

$$\begin{aligned} S(a) &= S(\sum_{(a)} a' \epsilon(a'')) = \sum_{(a)} S(a') \epsilon(a'') \cdot 1 = \sum_{(a)} S(a') a'' S'(a''') \\ &= \sum_{(a)} \epsilon(a') S'(a'') = \sum_{(a)} S'[(\epsilon(a') a''] = S'(a). \end{aligned}$$

A bialgebra equipped with an antipode is a *Hopf algebra.*

Theorem 6.1.8. *Let H be a finite-dimensional Hopf algebra. Then the dual bialgebra H^* is a Hopf algebra with an antipode $S^* : H^* \to H^*$ defined as the dual linear map to S, $(S^*(f))(a) = f(S(a))$.*

Proof. Let $f \in H^*$ and $a \in H$. Then

$$\left(\sum_{(f)} f' S^*(f'')\right)(a) = \sum_{(f),(a)} f'(a')(S^*(f''))(a'')$$
$$= \sum_{(f),(a)} f'(a') f''(S(a'')) = f(\sum_{(a)} a' S(a'')),$$

where in the last equation we have used the definition of the coproduct in the dual bialgebra. By the defining relations of an antipode, the last expression is equal to

$$f(\epsilon(a)\cdot 1) = \epsilon(a) f(1) = \epsilon(a)\epsilon^*(f) = (\epsilon^*(f)\cdot 1^*)(a).$$

The second of the axioms for S^* is proven in a similar way.

Theorem 6.1.9. *In a Hopf algebra H, $S(ab) = S(b)S(a)$ for all $a, b \in H$.*

Proof. First we note by $\Delta(xy) = \Delta(x)\cdot\Delta(y)$ that

$$\sum_{(xy)} (xy)' \otimes (xy)'' = \sum_{(x),(y)} x'y' \otimes x''y''.$$

Since the antipode is uniquely defined, it is sufficient to prove that the function $f(x, y) = S(y)S(x)$ satisfies the defining relations

$$\sum_{(xy)} (xy)' S((xy)'') = \sum_{(xy)} S((xy)')(xy)'' = \epsilon(xy)\cdot 1$$

when we replace $S(xy)$ by $f(x, y)$. But

$$\sum_{(x),(y)} x'y'S(y'')S(x'') = \sum_{(x)} x'(\epsilon(y)\cdot 1)S(x'') = \epsilon(y)\sum_{(x)} x'S(x'') = \epsilon(y)\epsilon(x)\cdot 1 = \epsilon(xy)\cdot 1.$$

A similar calculation can be carried through for the second relation.

Example 6.1.10 The bialgebra in the example 6.1.5 is a Hopf algebra with the antipode $(S(f))(g) = f(g^{-1})$ where $g \in G$. Indeed,

$$\left(\sum_{(f)} f'S(f'')\right)(g) = \sum f'(g)S(f'')(g) = \sum f'(g)f''(g^{-1}) = f(gg^{-1}) = f(e) = \epsilon(f)\cdot 1$$

and likewise for $\sum S(f')f''$.

An element $a \neq 0$ in a coalgebra is said to be *group like* if

$$\Delta(a) = a \otimes a.$$

If a, b is a pair of group like elements then $\Delta(ab) = \Delta(a) \cdot \Delta(b) = (a \otimes a) \cdot (b \otimes b) = ab \otimes ab$. Thus also ab is group like. Since $\Delta(1) = 1 \otimes 1$, the unit is also group like. In the special case when $S : H \to H$ is invertible, the set $G(H)$ of group like elements is a group: The inverse of a is then $S(a)$ because of

$$\sum_{(a)} a'S(a'') = aS(a) = \epsilon(a) \cdot 1.$$

On the other hand, $a \otimes 1 = (id \otimes \epsilon)\Delta(a) = (id \otimes \epsilon)(a \otimes a) = a \otimes \epsilon(a)$ and so $\epsilon(a) = 1$. This completes the proof of $aS(a) = 1$. The relation $S(a)a = 1$ is proven in a similar way.

Example 6.1.11 The tensor bialgebra $T(V)$ in example 6.1.7 becomes a Hopf algebra with the antipode $S : T(V) \to T(V)$ defined by $S(1) = 1$ and $S(v) = -v$ for $v \in V$. For a generic element $v_1 v_2 \dots v_n \in V^n$ we have then

$$S(v_1 v_2 \dots v_n) = (-1)^n v_n \dots v_2 v_1.$$

Exercise 6.1.12 Let H be the algebra $k[t,x]/I$, where $k[t,x]$ is the free (non-commutative) algebra with unit and with two generators t and x, and I is the ideal generated by the polynomials $t^2 - 1, x^2, xt + tx$. Show that H is finite-dimensional as a vector space and that $\Delta(t) = t \otimes t, \Delta(x) = 1 \otimes x + x \otimes t$ extend to a coproduct on H. Show also that $\epsilon(t) = 1, \epsilon(x) = 0$ and $S(t) = t, S(x) = tx$ define a Hopf algebra structure on H.

6.2 The Hopf algebra $SL_q(2)$

We have already met the bialgebra of $n \times n$ matrix coordinates x_{ij} in 6.1.4. This is not a Hopf algebra; this is related to the fact that a general $n \times n$ matrix does not have an inverse. But we can define the bialgebra $SL(n)$ as the quotient of the algebra $M_n(k)$ by the ideal generated by the single element $det = det(x_{ij}) - 1$. That is, every time we see the polynomial $det(x_{ij})$ we replace it with the unit element 1. This is now a Hopf algebra. The antipode is defined as

$$S(x_{ij}) = (-1)^{i+j} X_{ji}$$

where X_{ij} is the determinant of the submatrix obtained by deleting the i:th row and the j:th column from the matrix (x_{ij}). We can immediately check that for $z = x_{ij}$

$$\sum_{(z)} z' S(z'') = \sum_k x_{ik} S(x_{kj}) = \sum_k x_{ik} (-1)^{j+k} X_{kj} = \det(x_{ij}) \delta_{ij} = \epsilon(z) \cdot 1.$$

The antipode S is then extended to an arbitrary product of the generators x_{ij} using the condition $S(z_1 z_2 \dots z_p) = S(z_p) \dots S(z_2) S(z_1)$ and by linearity to all elements in $SL(n)$.

In the following we shall concentrate to the case $n = 2$ and we shall use the notation

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The antipode applied to the generators is then

$$S(g) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

The only relation in this commutative algebra is the determinant relation

$$ad - bc = 1.$$

The coproduct can be written in the matrix notation as

$$\Delta(g) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

That is, for example $\Delta(b) = a \otimes b + b \otimes d$. The counit satisfies $\epsilon(a) = \epsilon(d) = 1$ and $\epsilon(b) = \epsilon(c) = 0$. In this case S is invertible.

The Hopf algebra $SL(2)$ is commutative but not cocommutative. Next we shall construct a 1-parameter family $SL_q(2)$ of Hopf algebras which are both noncommutative and noncocommutative, except in the limiting case $q = 1$ when the algebra becomes the classical Hopf algebra $SL(2)$. In general, q is a complex number and we consider here all algebras over $k = \mathbb{C}$.

We start with the defining algebra relations:

$$\begin{aligned} ba &= qab & db &= qbd \\ ca &= qac & dc &= qcd \\ bc &= cb & ad - da &= (q^{-1} - q)bc. \end{aligned}$$

Thus in the case $q = 1$ this algebra is commutative. We denote the algebra by $M_q(2)$. The element

$$det_q = ad - q^{-1}bc$$

commutes with every element in $M_q(2)$ (prove this!). As in the commutative case we define the algebra $SL_q(2) = M_q(2)/I$, where I is the two-sided ideal generated by $det_q - 1$.

Next we define the coproduct Δ in $SL_q(2)$ exactly the same way as in the commutative case $q = 1$. Also the counit ϵ is defined by the same formulas as before. However, the definition of the antipode must be modified:

$$\begin{pmatrix} S(a) & S(b) \\ S(c) & S(d) \end{pmatrix} = \begin{pmatrix} d & -qb \\ -q^{-1}c & a \end{pmatrix}.$$

Exercise 6.2.1 Check that the antipode satisfies the relations $\sum x'S(x'') = \epsilon(x) \cdot 1 = \sum S(x')x''$.

Exercise 6.2.3 Show that the antipode of $SL_q(2)$ satisfies

$$\begin{pmatrix} S^{2n}(a) & S^{2n}(b) \\ S^{2n}(c) & S^{2n}(d) \end{pmatrix} = \begin{pmatrix} q^n & 0 \\ 0 & q^{-n} \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} q^{-n} & 0 \\ 0 & q^n \end{pmatrix}$$

for any integer n. Thus if n is a n:th root of identity then S^{2n} is the identity transformation.

Exercise 6.2.4 (The quantum plane) Let $A = \mathbb{C}_q[x, y]$ be the complex algebra with unit and the generators x, y subject to the relations $yx - qxy = 0$. Here $q \in \mathbb{C}$ is a constant. Show that the algebra $SL_q(2)$ acts in $\mathbb{C}_q[x, y]$ in the following sense: Set

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Then also $y'x' - qx'y' = 0$. Define further $\Delta_A : A \to SL_q(2) \otimes A$ as an algebra homomorphism such that $\Delta_A(x) = a \otimes x + b \otimes y$ and $\Delta_A(y) = c \otimes x + d \otimes y$. Show that Δ_A satisfies the *comodule relations* $(\Delta \otimes id) \circ \Delta_A = (id \otimes \Delta_A) \circ \Delta_A$.

Next we define a $^*-$algebra structure in the Hopf algebra $SL_q(2)$ over complex numbers. The $^*-$ operation should be thought of as taking the adjoint of linear operators. We require that it is antilinear, $(\alpha x+\beta y)^* = \overline{\alpha}x^* + \overline{\beta}y^*$ for all $x, y \in SL_q(2)$ and $\alpha, \beta \in \mathbb{C}$. Furthermore, $(xy)^* = y^*x^*$, and $(x^*)^* = x$. For Hopf algebras we require in addition

(1) $\Delta(x^*) = \Delta(x)^*$,
(2) $[S(S(x)^*)]^* = x$
(3) $1^* = 1$ and $\epsilon(x^*) = \overline{\epsilon(x)}$

for all x.

Note that by the second equation in a star Hopf algebra the antipode S has always an inverse. Sometimes one writes $* = S\gamma$, where $\gamma = S^{-1}*$. Then γ is antilinear and it is an automorphism of real algebras and coalgebras. It is also an involution, $\gamma^2 = 1$, by

$$\gamma^2(x) = S^{-1}(S^{-1}x^*)^* = S^{-1}[(S(x)^*)]^* = S^{-1}S(x) = x.$$

In the case of $H = SL_q(2)$ $(q \in \mathbb{R})$ we set

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

so that

$$a^* = d,\ d^* = a,\ b^* = -qc,\ c^* = -q^{-1}b.$$

The action of $*$ on an arbitrary element in $SL_q(2)$ is then uniquely defined by the property that $*$ is an antilinear antiautomorphism.

The motivation for introducing $*$ is the following. In the classical case $q = 1$ of a commutative algebra $SL(2)$ the functions on the subgroup $SU(2) \subset SL(2,\mathbb{C})$ can be taken as (z_1, z_2) with $a = z_1, b = -\overline{z_2}, c = z_2, d = \overline{z_1}$. Then the coordinate functions satisfy the star relations above (with $q = 1$). So we can define the *quantum group* $SU_q(2)$ as the star algebra above in the case of general $q \in \mathbb{R}$. (We need to take q real in order that the axioms for the star operation are satisfied.)

Let us also mention that there is a generalization of the algebra $SL_q(2)$ to the $N \times N$ matrix case $SL_q(N)$. The algebra commutation relations are given as

$$T_i^m T_i^k = q T_i^k T_i^m, \qquad T_j^m T_i^m = q T_i^m T_j^m$$
$$T_i^m T_j^k = T_j^k T_i^m,\ T_i^k T_j^m - T_j^m T_i^k = (q^{-1} - q) T_i^m T_j^k,$$

where $i < j$ and $k < m$. The coproduct is defined by

$$\Delta(T_i^j) = \sum_k T_i^k \otimes T_k^j$$

and $\epsilon(T_i^j) = \delta_{ij}$. The quantum determinant is

$$det_q = \sum_{\sigma \in S_N} (-q)^{\ell(\sigma)} T_1^{\sigma(1)} \dots T_N^{\sigma(N)},$$

where $\ell(\sigma)$ is the length of the minimal decomposition of σ into a product of transpositions. In the algebra $SL_q(N)$ one sets $\det_q \equiv 1$.

6.3 The quantum enveloping algebra $U_q(\mathbf{sl}(2))$

We define the quantum enveloping algebra $U_q = U_q(\mathbf{sl}(2))$ as the associative algebra with unit and generators E, F, K, K^{-1} subject to the defining relations

$$KK^{-1} = K^{-1}K = 1$$
$$KEK^{-1} = q^2 E,\ KFK^{-1} = q^{-2} F$$
$$EF - FE = \frac{K - K^{-1}}{q - q^{-1}}.$$

Here $\pm 1 \neq q \in \mathbb{C}$.

Lemma 6.3.1. *Let $m \geq 0$ and n be integers. Then*

$$E^m K^n = q^{-2mn} K^n E^m,\ F^m K^n = q^{2mn} K^n F^m$$
$$[E, F^m] = [m] F^{m-1} \frac{q^{-m+1}K - q^{m-1}K^{-1}}{q - q^{-1}}$$
$$= [m] \frac{q^{m-1}K - q^{-m+1}K^{-1}}{q - q^{-1}} F^{m-1}$$
$$[E^m, F] = [m] \frac{q^{-m+1}K - q^{m-1}K^{-1}}{q - q^{-1}} E^{m-1}$$
$$= [m] E^{m-1} \frac{q^{m-1}K - q^{-m+1}K^{-1}}{q - q^{-1}}.$$

Here $[n] = \frac{q^n - q^{-n}}{q - q^{-1}} = q^{n-1} + q^{n-3} + \dots + q^{-n+1}$ when $n > 0$. Note that $[-n] = -[n]$.

Proof. The first row of relations follows immediately from the defining relations. The second (and the rest) relation is proven by induction on m. The case $m = 1$ follows from the defining relations and the induction step follows from

$$\begin{aligned}[E, F^m] &= [E, F^{m-1}]F + F^{m-1}[E, F] \\ &= [m-1]F^{m-2}\frac{q^{-m+2}K - q^{m-2}K^{-1}}{q - q^{-1}}F + F^{m-1}\frac{K - K^{-1}}{q - q^{-1}} \\ &= [m-1]F^{m-1}\frac{q^{-m+2}q^{-2}K - q^{m-2}q^{2}K^{-1}}{q - q^{-1}} + F^{m-1}\frac{K - K{-1}}{q - q^{-1}}.\end{aligned}$$

Both terms on the right contain F^{m-1} as the first factor. Combining the polynomials in K, K^{-1} one easily sees that they give the required factor in the case of $[E, F^m]$.

We have defined the quantum algebra U_q for $q \neq \pm 1$. However, in certain sense the enveloping algebra $U(\mathbf{sl}(2))$ is the limit of U_q as $q \to 1$. Let us think of K as the element $q^h = e^{h \log q}$, where h is the standard element in the Cartan subalgebra of $\mathbf{sl}(2)$. Then

$$\lim_{q \to 1} \frac{q^h - q^{-h}}{q - q^{-1}} = h.$$

So in this limit we get $[E, F] = h$ and the relation $KEK^{-1} = q^2 E$ leads to $[h, E] = 2E$. Thus we recover the standard commutation relations of $\mathbf{sl}(2)$.

We have a more rigorous relation between U_q and $U(\mathbf{sl}(2))$ using the following observation. Add a generator L to the algebra U_q such that

$$[E, F] = L,\ (q - q^{-1})L = K - K^{-1}$$

$$[L, E] = q(EK + K^{-1}E),\ [L, F] = -q^{-1}(FK + K^{-1}F).$$

This defines a new associative algebra U_q' but it is straightforward to prove that actually $U_q' \simeq U_q$. The advantage with U_q' is that it is defined for all values of q. In particular, when $q = 1$ we have $U_1'/(K - 1) \simeq U(\mathbf{sl}(2))$.

Exercise 6.3.2 Prove the last isomorphism above.

Next we study the finite-dimensional representations of $U_q(\mathbf{sl}(2))$ when $q \neq 0$ is not a root of unity.

If V is a U_q module we denote by $V(\lambda)$ the weight subspace of V defined as the space of vectors v for which $Kv = \lambda v$.

By the Lemma 6.3.1, if $v \in V(\lambda)$ then $Ev \in V(q^2\lambda)$ and $Fv \in V(q^{-2}v.)$

The vector $0 \neq v \in V(\lambda)$ is said to be a highest weight vector if $Ev = 0$. In an irreducible highest weight module all other vectors are linear combinations of the vectors $v_n = F^n v, v_0 = v$. The nonzero vectors in this sequence are linearly independent, since the eigenvalues of K are $q^{-2n}\lambda$ and they are all different when $\lambda \neq 0, q \neq \pm 1$. In the case $\lambda = 0$ we have

$$EFv = [E,F]v = \frac{K - K^{-1}}{q - q^{-1}} v = 0$$

and so Fv generates an invariant subspace. In an irreducible module we must then have $Fv = 0$ and so the module becomes the trivial one-dimensional module where E, F, K are represented by the zero operator.

In general, for $\lambda \neq 0$, let n be the smallest integer for which $F^{n+1}v = 0$. Then

$$0 = EF^{n+1}v = [n+1]\frac{q^n K - q^{-n}K^{-1}}{q - q^{-1}} F^n v = [n+1]\frac{q^{-n}\lambda - q^n \lambda^{-1}}{q - q^{-1}} v_n.$$

Since $v_n \neq 0$, we must have $q^{-n}\lambda - q^n\lambda^{-1} = 0$, that is, $\lambda^2 = q^{2n}$ or $\lambda = \pm q^n$. The dimension of V is equal to $n+1$. As in the case of $\mathbf{sl}(2)$, the space V is a direct sum of one-dimensional weight spaces $V(\mu)$ where now $\mu = q^{-2k}\lambda$ with $k = 0, 1, \ldots, n$. Taking account $\lambda = \epsilon q^n$, we see that the spectrum of K consists of the numbers $\epsilon q^n, \epsilon q^{n-2}, \ldots, \epsilon q^{-n}$ with $\epsilon = \pm 1$.

A difference to the classical situation is that for a given dimension $n+1$ we have two different irreducible highest weight modules labelled by $\epsilon = \pm 1$.

In the case of $\mathbf{sl}(2)$ we defined a Casimir element $c = yx + h^2 + h$ which commutes with the whole algebra. Here we can define the quantum Casimir element

$$c_q = EF + \frac{q^{-1}K + qK^{-1}}{(q - q^{-1})^2}.$$

Exercise 6.3.3 Show that c_q commutes with E, F and K.

The value of the Casimir c_q in an irreducible highest weight module is

$$\frac{q^{-1}\lambda^{-1} + q\lambda}{(q - q^{-1})^2}.$$

The case when $q \neq \pm 1$ is a root of unity is more tricky. So let us assume that N is the smallest positive integer for which $q^N = 1$.

Lemma 6.3.4. *The elements E^N, F^N, K^N commute with the algebra U_q.*

Proof. Follows from Lemma 6.3.1 since $[N] = 0$.

Let λ be any 1-dimensional representation of the center of U_q. Denote by J_λ the ideal in U_q generated by the elements $c - \lambda(c) \cdot 1$ where c belongs to the center. Since the elements $E^i F^j K^\ell$ span the algebra U_q (by PBW theorem) the quotient algebra U_q/J_λ is finite-dimensional when $q^N = 1$.

Theorem 6.3.5. *There are no finite-dimensional irreducible modules of dimension $> N$.*

Proof. a) Assume first that there is a weight vector $0 \neq v \in V$ such that $Fv = 0$. Now the subspace spanned by the vectors $v, Ev, E^2v, \ldots E^{N-1}v$ is invariant under the action of E, F, K, K^{-1} by the defining relations of U_q and by the fact that E^N commutes with everything, so by Schur's lemma $E^N v = av$ for some $a \in \mathbb{C}$. So if the dimension of V is bigger than N we have a proper submodule, a contradiction, since the module is irreducible.

b) Next we assume that there is no weight vector v such that $Fv = 0$. Because the module is finite-dimensional there is at least one nonzero eigenvector v for K. The subspace W spanned by the vectors $v, Fv, F^2, \ldots, F^{N-1}v$ is clearly invariant under F. It is also invariant under K, K^{-1} by Lemma 6.3.1. In addition,

$$
\begin{aligned}
E(F^p v) &= EF(F^{p-1}v) \\
&= \left(c_q - \frac{q^{-1}K + qK^{-1}}{(q-q^{-1})^2} \right) (F^{p-1}v) \\
&= c_q F^{p-1} v - \frac{q^{-1}K + qK^{-1}}{(q-q^{-1})^2} (F^{p-1}v).
\end{aligned}
$$

This shows that $E(F^p v)$ belongs to W when $p > 0$. The case $p = 0$ can be treated using the observation $v =$ *const.* $\times F^N v$. But since the module is irreducible we must have $W = V$ so that $\dim V \leq N$.

6.4 The Hopf algebra structure of $U_q(\mathbf{sl}(2))$

We define a comultiplication and counit in U_q using the generators E, F, K and K^{-1}.

$$\Delta(E) = 1 \otimes E + E \otimes K, \ \Delta(F) = K^{-1} \otimes F + F \otimes 1$$

$$\Delta(K) = K \otimes K, \ \ \Delta(K^{-1}) = K^{-1} \otimes K^{-1}$$

$$\epsilon(E) = \epsilon(F) = 0, \ \ \epsilon(K) = \epsilon(K^{-1}) = 1.$$

The antipode is defined by

$$S(E) = -EK^{-1},\ S(F) = -KF,\ S(K) = K^{-1},\ S(K^{-1}) = K.$$

Once the operations are fixed for the generators they are uniquely defined on all elements in U_q by the homomorphism property of ϵ, Δ and by the antiautomorphism property of S. The only thing to check is that the mappings satisfy the axioms on the generators and preserve the defining relations among the generators. We give a couple of typical computions and leave the rest to the reader.

First, let us take a look at Δ. Let us show that Δ preserves the relation

$$[E, F] = \frac{K - K^{-1}}{q - q^{-1}}.$$

Starting from the left-hand-side we obtain

$$\begin{aligned}
&[\Delta(E), \Delta(F)] \\
&\quad = (1 \otimes E + E \otimes K)(K^{-1} \otimes F + F \otimes 1) \\
&\quad\quad - (K^{-1} \otimes F + F \otimes 1)(1 \otimes E + E \otimes K) \\
&\quad = K^{-1} \otimes EF + F \otimes E + EK^{-1} \otimes KF + EF \otimes K \\
&\quad\quad - K^{-1} \otimes FE - K^{-1}E \otimes FK - F \otimes E - FE \otimes K \\
&\quad = K^{-1} \otimes [E, F] + [E, F] \otimes K \\
&\quad = \frac{K^{-1} \otimes (K - K^{-1}) + (K - K^{-1}) \otimes K}{q - q^{-1}} \\
&\quad = \frac{\Delta(K) - \Delta(K^{-1})}{q - q^{-1}} = \Delta(\frac{K - K^{-1}}{q - q^{-1}})
\end{aligned}$$

An example of a calculation to show that Δ is coassociative:

$$\begin{aligned}
(\Delta \otimes id)\Delta(E) &= (\Delta \otimes id)(1 \otimes E + E \otimes K) = 1 \otimes 1 \otimes E + 1 \otimes E \otimes K + E \otimes K \otimes K \\
&= (id \otimes \Delta)(1 \otimes E + E \otimes K) = (id \otimes \Delta)\Delta(E).
\end{aligned}$$

The axioms for the antipode: We give a sample calculation concerning the relation $KEK^{-1} = q^2 E$.

$$S(K^{-1})S(E)S(K) = K(-EK^{-1})K^{-1} = -q^2 EK^{-1} = q^2 S(E).$$

Exercise 6.4.1 Check the relations $\sum_{(x)} x' S(x'') = \sum_{(x)} S(x')x'' = \epsilon(x) \cdot 1$ when x is any of the generators E, F, K, K^{-1}.

Theorem 6.4.2. *We have $S^2(u) = KuK^{-1}$ for any $u \in U_q$.*

Proof. It suffices to check this for generators:

$$\begin{aligned} S^2(E) &= S(-EK^{-1}) = -S(K^{-1})S(E) = KEK^{-1} \\ S^2(F) &= S(-KF) = -S(F)S(K) = KFK^{-1} \\ S^2(K) &= K = K(K)K^{-1}. \end{aligned}$$

The classical Lie algebra $\mathbf{sl}(2)$ acts on the polynomial algebra $\mathbb{C}[x,y]$ of two commuting variables. Explicitly, we have

$$E = x\partial_y, \;\; F = y\partial_x, \;\; H = x\partial_x - y\partial_y$$

and it is easy to verify the Lie algebra commutation relations $[E,F] = H$, $[H,E] = 2E$, $[H,F] = -2F$. In the case of the quantum algebra U_q we construct an action in the quantum plane $A = \mathbb{C}_q[x,y]$ of example 6.2.4. The generators are now

$$E(u) = x\partial_y^{(q)}u, \;\; F(u) = (\partial_x^{(q)}u)y$$

where the *quantum derivations* are defined by

$$\partial_x^{(q)}(x^m y^n) = [m]x^{m-1}y^n, \;\; \partial_y^{(q)}(x^m y^n) = [n]x^m y^{n-1}.$$

The action of K is given by

$$K(x^m y^n) = q^{m-n}x^m y^n$$

and K^{-1} is the inverse action.

We can check the commutation relations by a direct computation. For example,

$$\begin{aligned} [E,F](x^m y^n) &= E([m]x^{m-1}y^{n+1}) - F([n]x^{m+1}y^{n-1}) \\ &= ([m][n+1] - [n][m+1])x^m y^n = \frac{K - K^{-1}}{q - q^{-1}}(x^m y^n). \end{aligned}$$

It is clear from the definitions that the homogeneous polynomials of order n form an invariant subspace V_n. There is a highest weight vector $v^+ = x^n$ with the property $Ev^+ = 0$ and $Kv^+ = q^n v^+$. The dimension of V_n is equal to $n+1$. Thus the representation of U_q in V_n is equivalent to the highest weight representation with highest weight $\lambda = \epsilon q^n$ with $\epsilon = +1$.

Let us consider in detail the 2-dimensional representation in V_1. In a basis where K is diagonal we can write

$$K = \begin{pmatrix} q & 0 \\ 0 & q^{-1} \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Next we define elements $A, B, C, D \in U_q^*$ by

$$u = \begin{pmatrix} A(u) & B(u) \\ C(u) & D(u) \end{pmatrix}$$

for $u \in U_q$. We denote by the same symbol the element of U_q and the 2×2 matrix representing in in V_1.

Let H be the algebra generated by the elements A, B, C, D with the multiplication defined by the coproduct in U_q. For example, AB is the element in U_q^* defined by

$$(AB)(u) = \sum_{(u)} A(u')B(u'').$$

The coproduct in H is defined as in the case of the quantum matrix algebra $M_q(2)$,

$$\Delta \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \otimes \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

It is straightforward to check that the commutation relations of A, B, C, D which follow from coproduct in U_q are exactly the same as the commutation relations of a, b, c, d in $SL_q(2)$. In addition, we can define the counit as in $SL_q(2)$. To define the antipode, we need to go to the quotient algebra $H/(det_q)$ where $det_q = AD - q^{-1}BC$. The antipode satisfies

$$S(h)(u) = h(S(u))$$

for all $u \in U_q$ and $h \in H$.

Exercise 6.4.3 Compute $S(A)$ from the definition above.

Exercise 6.4.4 Show that $BA = qAB$. Hint: It is sufficient to evaluate both sides for the basis elements $u = E^i F^j K^\ell$ for $i, j = 0, 1, 2$. (Why?)

One can also check that we have the duality relations

$$h(uv) = \Delta(h)(u, v)$$

for $h \in H$ and $u, v \in U_q$. Thus in some sense H is the dual algebra U_q^*; this statement is not completely precise, since U_q is infinite dimensional and the dual

U_q^* is not strictly speaking a Hopf algebra since one cannot identify the algebraic tensor product $U_q^* \otimes U_q^*$ as the space of bilinear functions on U_q. The latter space contains the former, but is larger.

Exercise 6.4.5 Show that an element $u \in U_q(\mathbf{sl}(2))$ is group-like if and only if $u = K^n$ for some $n \in \mathbb{Z}$.

Exercise 6.4.6 Let q be real and positive. Show that $E^* = KF, F^* = EK^{-1}, K^* = K$ determine a star algebra structure on U_q. What happens if q is complex?