A CRASH COURSE IN LIE ALGEBRAS

Nils-Peter Skoruppa

$$[h, x] = 2x$$

$$[h, y] = -2y$$

$$[x, y] = h$$

Université Bordeaux 1 — UFR Mathématiques et Informatique
— Ecole doctorale 1998 —

Version: Id: lie.tex,v 1.2 2003/11/20 22:49:00 fenrir Exp

Preface

These are the notes of a course on Lie algebras which I gave at the university of Bordeaux in spring 1997. The course was a so-called "Cours PostDEA", and as such had to be held within 12 hours. Even more challenging, no previous knowledge about Lie algebras should be assumed. Nevertheless, I had the goal to reach as peak of the course the character formula for Kac-Moody algebras, and, at the same time, to give complete proofs as far as possible.

Accordingly, the present notes are a somewhat streamlined approach to Kac-Moody algebras, and some absolutely crucial topics (like root systems and Weyl groups) are only very briefly touched. Proofs are almost always completely given, with a few exceptions: some facts about root systems and Weyl groups, and the classical basic theorems of Lie, Engel and Cartan (criterion for solvability) are explained, but not proved. However, the missing proofs can be easily complemented from the literature.

The material is taken from various sources, in particular from courses of Serre, the book of Humphreys and the basic monograph of Kac on Kac-Moody algebras. Probably there is not a single theorem in these notes which cannot already be found in some text book. In that sense I did not add anything new. However, I tried to reach the main results of the theory as fast and as straight forward as possible, thereby starting from scratch. The resulting very condensed presentation may as such fill a gap and thus justify these notes. I hope that this crash course approach will be of some use for those who want or need to learn the basics of classical and Kac-Moody Lie algebras within a few afternoons.

Probably there are a lot of misplaced symbols, typographical errors or even (hopefully minor) mistakes awaiting to be found. I am grateful for any hint. I thank Jacques Martinet, who already pointed out a lot of such 'bugs' in a first draft.

N-P. Skoruppa — Bordeaux, October 1998

Contents

Chapter 1

Basics on Lie algebras

1.1 Definitions

For the following we fix once and for all an arbitrary field K. All vector spaces will be over K.

Definition. A Lie algebra over a field K is a vector space L over K together with a map $L \times L \to L$ denoted $(x, y) \mapsto [xy]$ satisfying the following axioms:

1. The bracket operation is bilinear.

2. $[xx] = 0$ for all x in L.

3. $[x[yz]] + [y[zx]] + [z[xy]] = 0$ for all x, y, z in L.

It is clear how to define the notions of *(Lie) subalgebra, morphisms and ideals* of Lie algebras. If I and J are ideals then $[IJ]$, the subspace spanned by all $[xy]$ (x in I, y in J) is an ideal (use $[L[IJ]] = -[I[JL]] - [J[LI]] \subset [IJ]$). We can define the *quotient algebra with respect to an ideal*. If I is an ideal of L, then the bracket on L/I is defined by $[x + I, y + I] = [xy] + I$. This is well-defined, since I is an ideal, and not just a subalgebra.

Moreover, we have the direct sum of a family Lie algebras L_j. This is the direct sum of the vector spaces L_j with the bracket $[\sum x_j \sum y_k] = \sum [x_j y_j]$ (x_j, y_j in L_j). Note that $[L_j L_k] = 0$ for $j \neq k$.

Trivial examples of Lie algebras are the *abelian Lie algebras* i.e. those with $[xy] = 0$ for all x, y in L. Thus any vector space can be made into a Lie algebra.

The basic nontrivial examples are gl_l, the K-vector space of square matrices of size l over K, and $\mathrm{gl}(V)$, the space of endomorphisms of the vector space V, both with $[xy] = xy - yx$. Any Lie algebra isomorphic to a subalgebra of some $\mathrm{gl}(V)$ is called a *linear Lie algebra*. By Ado's theorem any finite dimensional Lie algebra is linear.

Obvious linear algebras are the algebras t_l, n_l and d_l of upper triangular, strictly upper triangular (i.e. upper triangular and zero on the diagonal) and diagonal matrices, respectively.

More subtle examples of linear algebras are the *classical algebras*. They fall into four families, each indexed by integers $l \geq 1$.

The first family is $\mathrm{A}_l = \mathrm{sl}_{l+1}$, the subalgebra of gl_{l+1} consisting of matrices of trace 0. (That the matrices of trace 0 are closed under the bracket is due to the identity $\mathrm{tr}(xy) = \mathrm{tr}(yx)$, valid for arbitrary square matrices x and y.)

The other three families arise by the following construction: Fix any matrix s. Then the matrices x such that $sx + x^t s = 0$ form a subalgebra of gl_l. (To check that the subspace built by these matrices is closed under the bracket, note $sxy = -x^t sy = x^t y^t s$.) For

$$s = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1_l \\ 0 & 1_l & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1_l \\ -1_l & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1_l \\ 1_l & 0 \end{pmatrix}$$

we obtain the orthogonal algebra $\mathrm{B}_l = \mathrm{o}(2l+1)$, the symplectic algebra $\mathrm{C}_l = \mathrm{sp}_{2l}$, and the orthogonal algebra $\mathrm{D}_l = \mathrm{o}(2l)$, respectively.

A second method of construction is as follows. Let $\mathfrak{A}$ be an algebra over K (i.e. a vector space equipped with a bilinear form denoted by juxtaposition). Then the derivations $\mathrm{Der}(\mathfrak{A})$ form a subalgebra of $\mathrm{gl}(\mathfrak{A})$. Recall that a derivation is a linear map $\delta : \mathfrak{A} \to \mathfrak{A}$ satisfying

$$\delta(ab) = \delta(a)b + a\delta(b).$$

For a Lie algebra L we always have the *inner derivations* $\mathrm{ad}x = \mathrm{ad}_L x$ given by

$$\mathrm{ad}x(y) = [xy].$$

That these are derivations is equivalent to the Jacobi identity. Indeed, we may rewrite it (using $[xy] = -[yx]$) in the form

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

The map

$$\mathrm{ad} : L \to \mathrm{Der}L, \quad x \mapsto \mathrm{ad}x$$

is a morphism of Lie algebras, called the *adjoint representation*. That it is really a morphism is again equivalent to the Jacobi identity:

$$[\mathrm{ad}x\ \mathrm{ad}y](z) - \mathrm{ad}[xy](z) = [x[yz]] - [y[xz]] - [[xy]z]] = 0.$$

The adjoint representation plays an important role in the study of Lie algebras. Note that its kernel is the *center* $\mathrm{Z}(L)$ of L, i.e. the subalgebra of all $x \in L$ such that $[xL] = 0$. As kernel of a morphism, the center is in particular an ideal of L.

Definition. A representation of L is a morphism of L into $\mathrm{gl}(V)$ for some K-vector space V. An L-module is a K-vector space V together with a bilinear map $L \times V \to V$ (denoted by a dot or simply by juxtaposition) such that

$$[xy] \cdot m = x \cdot (y \cdot m) - y \cdot (x \cdot m)$$

for all x, y in L and m in V.

Any L-module V defines a representation ρ into $\mathrm{gl}(V)$ via $\rho(x)(m) = x \cdot m$ and vice versa.

It is clear how to define morphism and direct sum of L-modules. The tensor product of two L-modules V and W is defined as the tensor product of the vector spaces $V \otimes W$ with the bracket

$$x \cdot (m \otimes n) = x \cdot m \otimes n + m \otimes x \cdot n.$$

The dual module V^* is the dual vector space of V with

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

Note the minus sign.

In particular, if V and W are L-modules, then

$$\mathrm{Hom}(V, W) \cong V^* \otimes W \quad ((f \otimes n)(m) = f(m)n)$$

is also an L-module. The operation of L on a ϕ is given by

$$(x \cdot \phi)(m) = x \cdot \phi(m) - \phi(x \cdot m).$$

Definition. A L-module V is called simple if it has no nontrivial submodules. We call a L-module V semisimple if it decomposes into a direct sum of simple submodules. A Lie algebra is called *simple* if it has no nontrivial ideals, (i.e. if the adjoint representation is simple), and if it is not abelian.

We exclude here in the definition of 'simple Lie algebra' the one dimensional Lie algebras, which are all abelian, since we want to show later that any finite dimensional representation of a simple Lie algebra is semisimple. For a one dimensional Lie algebra this is false: indeed, consider the representation $\rho : \mathrm{gl}_(1) \to \mathrm{gl}(K^2)$ given by

$$x \cdot v = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} v$$

($x \in K = \mathrm{gl}_1$, v a row vector and usual matrix multiplication on the right hand side). Then $V = K^2$ is mapped under gl_1 onto the subspace V_0 of row vectors with second entry 0, and V_0 is mapped onto 0. Hence V contains exactly one nontrivial submodule, namely V_0, and therefore it cannot be semisimple.

Since the kernel of a representation is an ideal, we observe that any representation of a simple Lie algebra is either trivial or injective. Taking for instance the adjoint representation we note that a simple Lie algebra is always linear.

Definition. L is called *reductive*, if the adjoint representation is semisimple. We call L semisimple if, in addition, L equals its *derived algebra* $[LL]$.

Again, the second condition ensures that finite dimensional representations of a semisimple Lie algebra will be semisimple (at least if $\mathrm{char}K = 0$; see Weyl's theorem below). Indeed, we have to exclude the existence of ideals I of L such that L/I is nontrivial and abelian (see the discussion of the notion of simple algebra above), and a quotient L/I is easily seen to be abelian if and only if $[LL] \subset I$. Likewise, we can replace the condition $L = [LL]$ by the condition that L contains no abelian ideals: If I is an ideal and L reductive, we can find an ideal I' such that $L = I \oplus I'$. We have $[I\,I'] = 0$ (since the left hand side is in I and in I'). Hence the projection $L \to I$ with respect to this decomposition is an epimorphism of Lie algebras and we obtain an isomorphism of Lie algebras $I \cong L/I'$. Thus, if $L = [LL]$ and I abelian, then $I' \supseteq [LL] = L$, hence $I = 0$. Conversely, if L has no nontrivial abelian ideals and if we take $I = [LL]$, then $I' \cong L/[LL]$ is abelian, hence 0, which shows $L = [LL]$.

In particular we note that the center $\mathrm{Z}(L)$ is trivial for a semisimple L. But the center is the kernel of the adjoint representation. Thus

Proposition. *For a semisimple Lie algebra L the adjoint representation* ad *is injective. In particular, L is linear.*

Later we shall see, for characteristic 0 and for finite dimensional Lie algebras, that the condition that L contains no abelian ideals alone already suffices to ensure that L is semisimple.

Proposition. *A Lie algebra is semisimple if and only if it is isomorphic to a direct sum of simple Lie algebras.*

Proof. Suppose L is semisimple. Since L is reductive it decomposes into a direct sum (in the category of vector spaces) of minimal ideals.

$$L = \bigoplus_j I_j.$$

Clearly, in such a decomposition $[I_j, I_k] \subseteq I_j \cap I_k = 0$ for $i \neq k$. The minimal property of the I_j implies that they are Lie algebras without nontrivial ideals: If J is an ideal of I_j then, on using $[I_k J] = 0$ for $k \neq j$, we have $[LJ] = \sum_k [I_k J] \subset [I_j J] \subset J$, i.e. J is also an ideal of L. Thus the I_j are either simple as Lie algebras, or else they are abelian. The latter is impossible if L contains no nontrivial abelian ideals.

To prove the converse statement note that in a decomposition of a Lie algebra L into a direct sum of simple algebras I_j any I_j is a minimal ideal of L. Moreover, L contains no abelian ideals $\neq 0$, since any ideal I is the direct sum of certain I_j (which in turn are not abelian as simple algebras). Indeed, first of all, any ideal is the union of all minimal ones contained in it (if $x \in I$, then the intersection of all ideals containing x is a minimal ideal in I). Secondly, any minimal ideal J has to be equal to one of the I_j (since, for a minimal ideal J, one has, $J = [JL] = \sum [JI_j]$, and $[JI_j] = I_j$ if $J = I_j$ and $= 0$ otherwise.). Since the I_j are not abelian,

Alternatively, one could show directly $[LL] = \sum [I_j I_j] = \sum I_j = L$. □

1.2 The universal enveloping algebra

Sometimes it is useful to view a representation of a Lie algebra as a representation of an associative algebra. This can be done using the associated universal enveloping algebra.

By *algebra* we understand an algebra over K, i.e. a K-vector space A together with a bilinear map $A \times A \to A$, which we simply denote by juxtaposition. Note that any associative algebra A can be viewed as Lie algebra via the bracket $[x\,y] = xy - yx$ $(x, y \in A)$. (The associativity of the multiplication on A ensures the Jacobi identity). If necessary, we denote this Lie algebra by $\mathrm{Lie}(A)$. If we speak of a Lie algebra homomorphism of a Lie algebra L into A we always mean a homomorphism of L into $\mathrm{Lie}(A)$.

Definition. An associative unitary (i.e. with unit) algebra $\mathrm{U}(L)$ is called *universal enveloping algebra* of L if there is a Lie algebra homomorphism $\iota : L \to \mathrm{U}(L)$ such that for any Lie algebra homomorphism $f : L \to A$ into a unitary associative algebra A there is a unique factorization

$$L \to \mathrm{U}(L) \xrightarrow{F} A$$

(i.e. a unique F such that $f = F \circ \iota$), where F a morphism of unitary algebras.

Clearly, the universal algebra is unique up to isomorphism. Moreover, it yields a bijection

$$\mathrm{Hom}_{\mathrm{Lie}}(L, \mathrm{Lie}(A)) \cong \mathrm{Hom}_{\mathrm{unit.\ alg.}}(\mathrm{U}(L), A).$$

In particular, If V is an L-module, then the associated representation $L \mapsto \mathrm{gl}(V)$ extends to a morphism of associated algebras with units $\mathrm{U}(L) \mapsto \mathrm{gl}(V)$. We obtain in this way a Lie an isomorphism of the category of L-modules and the category of $\mathrm{U}(L)$-modules.

To prove the existence of a universal enveloping algebra let

$$\mathrm{T}(L) = \bigoplus_{n \geq 0} L^{\otimes n}$$

the tensor algebra of L (viewed as a vector space over K), and let

$$\mathrm{U}(L) = T(L)/I,$$

where I is the two-sided ideal of the tensor algebra generated by the elements of the form $x \otimes y - y \otimes x - [xy]$ (x, y in L). The map ι is given by the composition $L \subset \mathrm{T}(L) \to \mathrm{U}(L)/I$ where the map on the right is the canonical projection. It is easily shown that this $U(L)$ is in fact a universal enveloping algebra.

The main theorem is (see [Serre], pp. 14):

Theorem. *(PBW: Poincaré-Birkhoff-Witt) Let $(x_i)_{i \in I}$ be a basis of L, and choose a total order on I. Then the family of monomials*

$$\iota(x_{i_1}) \cdots \iota(x_{i_n}) \qquad (i_1 \leq i_2 \leq \cdots i_n,\ n \geq 0)$$

is a basis of $\mathrm{U}(L)$.

Corollary. *The map $\iota : L \to \mathrm{U}(L)$ is injective.*

Example. One can easily verify $\mathrm{U}(\mathrm{gl}_1) \cong K[X]$ and $\mathrm{U}(\mathrm{sl}_1) \cong K$.

Chapter 2

Tools for finite dimensional Lie algebras

In this section all vector spaces and all Lie algebras will be assumed to be finite dimensional.

2.1 Cartan, Killing

Let x, y, z in $\mathrm{gl}(V)$ and V finite dimensional. Then

$$\mathrm{tr}([xy]z) = \mathrm{tr}(xyz - yxz) = \mathrm{tr}(xyz - xzy) = \mathrm{tr}(x[yz]).$$

If ρ is a representation of L on V then we set

$$B_\rho(x, y) = \mathrm{tr}(\rho(x)\rho(y)) \qquad (x, y \in L).$$

Then B_ρ is bilinear, and it is moreover *associative* in the sense

$$B_\rho([xy], z) = B_\rho(x, [yz]),$$

or equivalently

$$-B_\rho(\mathrm{ad}y(x), z) = B_\rho(x, \mathrm{ad}y(z)).$$

The usefulness of B_ρ is due to this identity. For example, for any ideal I, its orthogonal complement

$$I^\perp = \{x \in L \mid B_\rho(x, I) = 0\}$$

is also an ideal.

If $\rho = \mathrm{ad}$ and L is finite dimensional, then we write simply B for B_ρ. This is the *Killing form* of L.

If I is an ideal of the finite dimensional Lie algebra L, and if b is the Killing form of I (viewed as Lie algebra), then

Lemma. $b = B|_{I\times I}$.

Proof. Indeed, for x, y in I we have $\mathrm{ad}_I(x)\,\mathrm{ad}_I(y) = (\mathrm{ad}_L x\,\mathrm{ad}_L y)|_I$, and, since $f := \mathrm{ad}_L x\,\mathrm{ad}_L y$ maps L into I, the trace of $f|_I$ and f coincide. □

It is natural to study the intersections $I \cap I^\perp$ (say, with respect to the Killing form). The main theorem here is provided by Cartan's criterion for solvability.

If I is an ideal of a L such that L/I is abelian, then I contains the derived algebra $[L, L]$, and vice versa $L/[L, L]$ is abelian. Thus it is natural to study the *derived series*

$$L^{(0)} = L, \quad L^{(n+1)} = [L^{(n)}L^{(n)}].$$

Definition. We call L *solvable*, if, for some n, we have $L^{(n)} = 0$.

Theorem. *(Cartan's criterion for solvability) Let* $\mathrm{char}K = 0$*. Let L be a subalgebra of* $\mathrm{gl}(V)$ *with finite dimensional V. Suppose* $\mathrm{tr}(xy) = 0$ *for all x in L and y in $[LL]$. Then L is solvable.*

Proof. see [Humphreys] §4.3. □

Corollary. *Let* $\mathrm{char}K = 0$*. Let L be a finite dimensional Lie algebra and I an ideal of L and $I^\perp$ its dual with respect to the Killing form. Then $I \cap I^\perp$ is solvable.*

Proof. Indeed, the Killing form of $H := I \cap I^\perp$ is zero, since it equals the restriction of the Killing form of L. Thus, by Cartan's criterion applied to $\mathrm{ad}_H(H)$ we have that $H/\mathrm{Z}(H) \cong \mathrm{ad}_H(H)$ is solvable. Thus, for some n we have $H^{(n)} \subset \mathrm{Z}(H)$, and hence $H^{(n+1)} = 0$. □

As consequence of Cartan's theorem we can state several equivalent criteria for semi-simplicity. We need one more notation: It is easy to show that any finite dimensional L contains a unique maximal solvable ideal, the so-called radical $\mathrm{Rad}(L)$.

Theorem. *Let L be a finite-dimensional Lie algebra over a field of characteristic 0. Then the following statements are equivalent:*

1. *L is semisimple.*
2. *L contains no abelian ideals.*
3. *L contains no solvable ideals.*
4. $\mathrm{Rad}(L) = 0$.

5. *(Cartan's criterion for semisimplicity) The Killing form of L is non-degenerate.*

Proof. (1) implies (2): As we saw immediately after the definition of semisimplicity, a semisimple Lie algebra has no abelian ideals.

(2) equivalent (3): Indeed, an abelian ideal is solvable, and the last non-trivial factor in the derived series of a solvable ideal is abelian.

(3) equivalent (4): clear by the very definition of $\mathrm{Rad}(L)$.

(2) implies (5): Suppose L contains no abelian, and hence no solvable ideals. But the kernel $L^\perp$ of the Killing form B is solvable by the preceding corollary. Hence it is 0, i.e. B is non-degenerate.

(5) implies (2): Suppose B be non-degenerate. We prove that L contains no nonzero abelian ideals. Let I be an abelian ideal of L. Let x in I. Then, for any y in L the map $\mathrm{ad}x\mathrm{ad}y$ gives L to I, hence $(\mathrm{ad}x\mathrm{ad}y)^2 = [x[y[x[y\cdot]]]]$ gives L to $[II] = 0$, i.e. it is nilpotent. Hence $B(x, y) = \mathrm{tr}(\mathrm{ad}x\mathrm{ad}y) = 0$, and hence $x \in L^\perp = 0$.

(2) implies (1): We have to show that ad is reducible. Let I be an ideal. Then $I \cap I^\perp$ is solvable, hence 0. But then $L = I \oplus I^\perp$. □

2.2 Casimir, Weyl

We start with two propositions, which are useful in the analysis of representations $\rho : L \to \mathrm{gl}(V)$ of semisimple Lie algebras. If ρ is injective (faithful) and $\mathrm{char}K = 0$, then B_ρ is non-degenerate: Indeed, let R be its kernel, then by Cartan's criterion $\rho(R) \cong R$ is solvable, hence 0, since L is semisimple.

We can thus consider a Casimir element $c = c_\rho$ of ρ, which is defined by

$$c = \sum_i \rho(x_i)\rho(x_i^*),$$

where x_i is a basis of L and x_i^* its dual basis with respect to B_ρ. This clearly depends on the chosen basis x_i of L. Obviously

$$\mathrm{tr}(c) = \dim L$$

(since $\mathrm{tr}(\rho(x_i)\rho(x_i^*)) = B_\rho(x_i, x_i^*) = 1$).

Proposition. *c commutes with $\rho(L)$.*

Proof. For $x \in L$ set $\mathrm{ad}x x_i = \sum_j a_{ij}x_j$. From the associativity of B_ρ one obtains $\mathrm{ad}x x_i^* = -\sum_j a_{ji}x_j^*$. But then, using $[x, yz] = [x, y]z + y[x, z]$ (in gl(V)), we find

$$[\rho(x)c] = \sum_i ([\rho(x)\rho(x_i)]\rho(x_i^*) + \rho(x_i)[\rho(x)\rho(x_i^*)]$$
$$= \sum_i ([\rho([xx_i])\rho(x_i^*) + \rho(x_i)\rho([xx_i^*]) = 0.$$

□

Note as a corollary

Corollary. *(Schur's lemma) Let W be an irreducible L-submodule of V. Then $c(W) \subset W$ and $K[c|_W] \subset \mathrm{gl}(W)$ is a field. (in particular, if K is algebraically closed then $c|_W$ is a homothety.)*

Proof. Let p be a polynomial (over K). Since $x := p(c|_W)$ is a sum of product of elements of $\rho(L)$ it leaves W invariant. Its kernel and image are L-stable by the preceding proposition. Hence $x = 0$ or either x is an isomorphism of W. Thus $K[c|_W]$ is an integral domain, hence a field, since it is finite dimensional over K. □

Proposition. *Let $\rho : L \mapsto \mathrm{gl}(V)$ be a finite dimensional representation of a semisimple Lie algebra L. Then $\rho(L) \subset \mathrm{sl}(V)$. In particular, $\rho = 0$ for any 1-dimensional V.*

Proof. Since $L = [LL]$ we find $\rho(L) \subset [\mathrm{gl}(V)\,\mathrm{gl}(V)] = \mathrm{sl}(V)$. The latter identity follows from $\mathrm{tr}(xy - yx) = 0$. □

Theorem. *(Weyl) Any finite dimensional module of a semisimple Lie algebra is semisimple.*

Proof. It suffices to show that any submodule A of a finite dimensional L-module V has a complement which is invariant under L.

To begin with assume that A has codimension 1 and is irreducible. Denote by ρ the associated representation and c a Casimir element of ρ. Clearly $H := \ker(c)$ is an L-submodule of V. We shall show that H is a complement of A.

Since L acts trivially on V/A (since the latter is one-dimensional by the preceding proposition) the endomorphism c (as sum of products of elements of $\rho(L)$) does so too. Hence $\dim\ker c \geq 1$. On the other hand, the restriction of c to A is 0 or an isomorphism of A (since A is an irreducible L-submodule of V and by Schur's lemma). It is not 0 since then trace $\dim L = \mathrm{tr}(c) = \mathrm{tr}(c, V/A) + \mathrm{tr}(c, A) = 0$. Thus H is one dimensional and $H \cap A = 0$, which implies the desired result.

The case of an arbitrary A, but still of codimension 1, follows now by induction on $\dim A$: Choose a proper submodule B of A. Then A/B is irreducible and of codimension 1 in V/B. By induction hypothesis we find a submodule A' of V such that $V/B = A/B \oplus A'/B$, and then a submodule A'' such that $A' = B \oplus A''$. Clearly $V = A \oplus A''$.

The case of a general submodule A can finally be treated as follows: We are looking for a projection $V \to A$, which commutes with the action of L. Namely, if π is such a projection, then $V = A \oplus \ker(\pi)$, and $\ker(\pi)$ is stable under L. Let $\mathcal{V}$ and $\mathcal{A}$ denote the spaces of linear maps $V \to A$ whose restriction to A is a homothety of A respectively 0. Any element of $\mathcal{V}$ with non-zero restriction to A is a scalar multiple of a projection of V onto A. The space $\mathcal{V}$ is mapped into $\mathcal{A}$ under the action of L on $\mathrm{Hom}(V, A)$, since

$$(x \cdot \phi)(a) = x \cdot \phi(a) - \phi(x \cdot a) = 0$$

for any a in A and any ϕ which is a homothety on A. The last identity also shows that those ϕ in $\mathcal{V}$ such that $K\phi$ is stable under L and their restriction to A are not zero are exactly the multiples of the projections $V \to A$ which commute with L. Now $\mathcal{A}$ is a submodule of codimension 1 in $\mathcal{V}$. Hence, by our assumption, we can indeed find a submodule $K\phi$ such that $\mathcal{V} = \mathcal{A} \oplus K\phi$. Clearly, the restriction of ϕ to $\mathcal{A}$ is nonzero. □

2.3 Jordan

An important role in the representation theory and classification theory of Lie algebras is played by nilpotent and semisimple elements. Recall that an endomorphism f of a vector space V is called semisimple if V possesses a basis consisting of eigenvectors of f. It is nilpotent if $f^n = 0$ for some integer $n > 0$, or equivalently, if 0 is its only eigenvalue.

Theorem. *(Jordan decomposition) Let f be an endomorphism of the finite dimensional vector space V. Suppose that K contains all roots of the characteristic polynomial of f. There exist unique endomorphisms s and n such that $f = s + n$, s is semisimple, n is nilpotent, and s and n commute.*

The proof will actually show that s and n are polynomials in f.

Proof. Existence: Let α_j be the distinct roots of the characteristic polynomial of f, and let m_j be their multiplicities. Then V is the direct sum of the spaces $\ker(f - \alpha_j)^{m_j}$. Choose a polynomial p $K[T]$ such that

$$p \equiv \alpha_j (T - \alpha_j)^{m_j},$$

set $s = p(f)$ and $n = f - p(f)$.

Uniqueness: Let n' and s' another pair of endomorphisms with the given properties. Then $n - n' = s' - s$. Now n' commutes with f, hence also with $n = p(f)$, and similarly s' commutes with s. But then $n - n'$ is semisimple, and $s' - s$ is nilpotent. But the only endomorphism which is semisimple and nilpotent is the trivial one. □

Lemma. *Let A be a finite dimensional algebra over K. Suppose K is algebraically closed. Then* $\mathrm{Der}(A)$ *contains the semisimple and nilpotent parts of its elements.*

Proof. Let δ a derivation of A, and let $\delta = \sigma + \nu$ its Jordan decomposition. We want to show that, say σ is a derivation. Now, for a root λ of the characteristic polynomial of δ, say with multiplicity n, let $A_\lambda = \ker(f - \lambda)^n$. Then A is the direct sum of all A_λ. It suffices to check the derivation property of σ for x in A_λ and y in A_μ. Note, first of all that

$$A_\lambda A_\mu \subset A_{\lambda+\mu}.$$

Indeed

$$(\delta - (\lambda + \mu)^n(xy) = \sum \binom{n}{k} (\delta - \lambda)^k x (\delta - \sigma)^{n-k} y$$

(use induction on n). Note secondly, that σ is multiplication by λ on A_λ (exercise). Thus we find $\sigma(xy) = (\lambda + \mu)xy$, since xy in $A_{\lambda+\mu}$, and $\sigma(x)y + x\sigma(y) = (\lambda + \mu)xy$, since x in A_λ and y in A_μ. Hence σ is a derivation. □

Proposition. *Assume* $\mathrm{char} K = 0$ *and L semisimple. Then the adjoint map*

$$\mathrm{ad} : L \to \mathrm{Der}(L)$$

is an isomorphism. In particular, if K is algebraically closed, then $\mathrm{Der}(L)$ *contains the nilpotent and semisimple parts of its elements.*

Proof. The kernel of ad is the center of L, hence an abelian ideal, hence 0.

Set $M := \mathrm{ad}L$. Clearly $M \subseteq D := \mathrm{Der}L$. Moreover, M is even an ideal of D: indeed, for $\delta \in \mathrm{Der}(L)$ and $x \in L$ one has $[\delta, \mathrm{ad}x] = \mathrm{ad}\delta(x)$ since $[\delta, \mathrm{ad}x]y = \delta([xy]) - [x, \delta(y)] = [\delta(x)y]$. Let $I := M^\perp$ with respect to the Killing form B on D. We want to show $I = 0$.

Since the restriction of B to $M \times M$ is the Killing form of M, hence non-degenerate (since $M \cong L$ is semisimple), we have $I \cap M = 0$. But then $[I, M] \subset I \cap M = 0$. In other words, if $\delta \in I$ then for all x one has $\mathrm{ad}\delta(x) = 0$, whence $\delta(x) = 0$, whence $\delta = 0$. □

A useful proposition is the following.

Proposition. *Let L be a Lie-subalgebra of* $\mathrm{gl}(V)$*. If $x \in L$ is semisimple (nilpotent) as an element of* $\mathrm{End}(V)$*, then* $\mathrm{ad}_L x$ *is semisimple (nilpotent).*

Proof. Let $M := \mathrm{gl}(V)$. The restriction of a semisimple endomorphism to an invariant subspace is semisimple. Thus it suffices to show that $\mathrm{ad}_M x$ is semisimple. For this let v_j be a basis of V of eigenvectors of x with eigenvalues a_j. For fixed i, j let $E_{i,j}$ be the endomorphism of V such that $v_i \mapsto v_j$ and $v_{i'} \mapsto 0$ for $i' \neq i$. Then $\mathrm{ad}_M x(E_{i,j}) = xE_{i,j} - E_{i,j}x = (a_j - a_i)E_{i,j}$.

In the case x nilpotent, write $\mathrm{ad}x = \lambda - \rho$, where λ and ρ are right and left multiplication by x. If x is nilpotent, then λ and ρ are nilpotent too. Both commute, hence $\mathrm{ad}x$ as sum of two commuting nilpotent endomorphisms is nilpotent. □

Thus, if we have a Lie subalgebra L of $\mathrm{gl}(V)$, and if $x \in L$ has Jordan decomposition $x = s + n$ (as an endomorphism of V), say with $s, n \in L$, then $\mathrm{ad}x = \mathrm{ad}s + \mathrm{ad}n$ is the Jordan decomposition of $\mathrm{ad}x$. Indeed, by the foregoing proposition $\mathrm{ad}s$ and $\mathrm{ad}n$ are semisimple and nilpotent, respectively, and $[\mathrm{ad}s, \mathrm{ad}n] = \mathrm{ad}[s, n] = 0$.

2.4 Representations of sl_2

We shall see in the next section that finite-dimensional semisimple Lie algebras over algebraically closed fields of characteristic 0 are built in a sense by gluing together several copies of sl_2. Thus it is useful to study sl_2 first.

Let V be a (not necessarily finite dimensional) sl_2 module. For sl_2 we take the basis

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

Note that h is semisimple, and x, y are nilpotent. One has the multiplication table

$$[xy] = h, \quad [hx] = 2x, \quad [hy] = -2y.$$

From this it is easy to show that sl_2 is solvable for $\mathrm{char}K = 2$, and simple otherwise. Indeed, for the second statement let I be an ideal of sl_2. If $t = ax + by + ch \in I$, then $I \ni \mathrm{ad}h(t) = 2ax + 2by$. Hence, if $2 \neq 0$, we obtain $ax, by \in I$. Thus, if $0 \neq t \in I$, then either $a = b = 0$, and hence $h \in I$, or else $x \in I$ or $y \in I$. In any case we deduce $I = \mathrm{sl}_2$.

For $\lambda \in K$ let

$$V_\lambda = \{v \in V \,:\, hv = \lambda v\}.$$

One has

$$xV_\lambda \subset V_{\lambda+2}, \quad yV_\lambda \subset V_{\lambda-2}$$

(since e.g. $hxv = 2xv + xhv$ etc.). If $V_\lambda \neq 0$ then λ is called a *weight* and V_λ a weight space. A highest weight λ is a weight such that $V_{\lambda+2} = 0$, and the elements of V_λ are called *maximal weight vectors*.

For any given integer $N \geq 0$ the space $V(N) := K[X,Y]_N$ of homogeneous polynomials over K in two variables and of degree N is an sl_2-module via

$$x \cdot p = X\partial_Y p, \quad y \cdot p = Y\partial_X p, \quad h \cdot p = X\partial_X - Y\partial_Y,$$

where ∂_X and ∂_Y means partial derivation. The weight spaces are the spaces KX^aY^b (of weight $a-b$).

For the rest of the section we assume $\mathrm{char}K = 0$. Then the unique maximal weight of $V(N)$ is N, with maximal weight vector X^N. Moreover, $V(N)$ is then simple.

Now, assume conversely that V is a simple, finite dimensional sl_2 module. If we tensor by the algebraic closure $\overline{K}$, then there exists at least one weight of $\overline{K} \otimes V$ with respect to V. The following paragraphs with $\overline{K}$ instead with K, would then show that this weight is an integer. Thus, for any K there exists at least one weight.

Since V is finite dimensional there is at least one maximal one. Let v be a maximal weight vector with, say weight λ. Then there exist a maximal N such that $y^N \neq 0$. Since $v, yv, y^2v, \ldots, y^Nv$ have different weights, they are linearly independent. Moreover, the space $V' = \langle v, yv, y^2v, \ldots, y^Nv \rangle$ is not only stable under h and y, but also under x. One has

$$xy^kv = k(\lambda - (k-1))y^{k-1}v,$$

as can easily be seen using induction on k: $xv = 0$, and $xy^kv = [xy]y^{k-1}v + yxy^{k-1}v = (\lambda - 2(k-1))y^{k-1}v + y(k-1)(\lambda-(k-2))y^{k-2}v$. We note that for $k = N+1$ we obtain $(\lambda - N))y^Nv = 0$, hence $\lambda = N$. It is easily checked that V' is thus a submodule, hence equal to V. In fact, it is isomorphic as an sl_2-module to $V(N)$ via $X^N \mapsto v$, and $y^kX^N = N(N-1)\cdots(N-k+1)Y^kX^{N-k} \mapsto y^kv$

Summing up we can state:

Theorem. *Assume* $\mathrm{char}K = 0$. *Then any simple finite dimensional* sl_2-*module is isomorphic to* $V(N)$ *for suitable* N, *and, conversely, the* $V(N)$ *are simple* sl_2-*modules.*

As consequence, if V is a finite dimensional semisimple sl_2-module, then it decomposes into a direct sum of highest weight representations $V(N)$.

For a finite dimensional sl_2 module V we define its *character* as the formal Laurent polynomial

$$\mathrm{ch}_V = \mathrm{tr}(q^h, V) := \sum_{n\in\mathbb{Z}} \dim V_n \, q^n.$$

It is easy to show that V is determined (up to isomorphism) by its character. This follows easily from

$$\mathrm{ch}_{V(N)} = \frac{q^{N+1} - q^{-(N+1)}}{q - q^{-1}}$$

and the following proposition.

Proposition. *Any Laurent polynomial $p(q)$ satisfying $p(q^{-1}) = p(q)$ is a unique linear combination of the $\frac{q^{N+1}-q^{-(N+1)}}{q-q^{-1}}$ $(N \in \mathbb{Z}_{\geq 0})$.*

Proof. Exercise. □

2.5 Appendix: Engel, Lie

There are two further basic theorems in the theory of Lie algebras. They are actually needed as lemmas in the next section. We state them here without giving their proofs.

Theorem. *(Engel's Theorem.) Let L be a subalgebra of $\mathrm{gl}(V)$ for a finite dimensional V. Assume that L is nilpotent, i.e. that all elements of L are nilpotent. Then there exists a flag (V_i) in V, stable under L, and such that $L(V_i) \subseteq V_{i-1}$ for all i.*

Proof. [Humphreys], pp. 12. □

Recall that a flag (V_i) in V is a sequence of subspaces

$$0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V,$$

where $n = \dim V$, such that $\dim V_i = i$. The typical example of an nilpotent L is n_l, and the theorem states that any other nilpotent Lie algebra is isomorphic to a subalgebra of n_l.

Theorem. *(Lie's Theorem.) Assume that K is algebraically closed and of characteristic 0. Let L be a solvable subalgebra of $\mathrm{gl}(V)$ for a finite dimensional V. Then L stabilizes some flag in V.*

Proof. [Humphreys], pp. 16. □

The typical example of a solvable Lie algebra is t_l, and by Lie's theorem any linear solvable Lie algebra is isomorphic to a subalgebra of t_l.

Chapter 3

Finite dimensional semisimple Lie algebras

In this section we assume throughout that L is of finite dimension and semisimple, and that K is of characteristic 0 and algebraically closed.

3.1 The root space decomposition

Recall that any $x \in L$ has a unique decomposition $x = x_s + x_n$ with semisimple and nilpotent $\mathrm{ad}x_s$ and $\mathrm{ad}x_n$ which commute. In particular, L contains ad-semisimple elements $\neq 0$, since otherwise all elements of L would be ad-nilpotent, and hence $L \cong \mathrm{ad}(L)$ could not be semi-simple (use Engel's theorem). But then L contains also *toral* subalgebras, i.e. nonzero subalgebras consisting of ad-semisimple elements: e.g. Kx for $x \neq 0$ and semisimple.

Proposition. *Let H be a maximal toral subalgebra. Then H is abelian, and, even more, $H = \{x \in L \mid [x, H] = 0\}$.*

Proof. Let $x \in H$, $x \neq 0$. We want to show that $[xy] = 0$ for all $y \in H$. Since $\mathrm{ad}x$ is semisimple, we can assume that y is an eigenvector of $\mathrm{ad}x$, say $[xy] = ay$. On the other hand $\mathrm{ad}y$ is semisimple too. Hence we may write $x = \sum_j x_j$ with linearly independent eigenvectors x_j of $\mathrm{ad}y$, say $[yx_j] = a_j x_j$ for a suitable $a_j \in K$. But then $-ay = [yx] = \sum_j a_j x_j$, and applying once more $\mathrm{ad}y$ gives $0 = \sum_j a_j^2 x_j$. This implies $a_j = 0$ for all j, hence $a = 0$.

The proof of the second property, namely that H coincides with its centralizer in L, is somewhat painful. The idea is essentially to show that the centralizer is nilpotent on using that L contains the ad-semisimple and ad-nilpotent parts of its elements, and that H is maximal toral. Then one applies Engel's theorem to obtain thereby a description of the centralizer which is sufficient to deduce the desired result. For details we refer to [Humphreys], §8.2. □

Example. As an example we consider sl_l. Here a maximal toral subalgebra is built by the set H_l of all diagonal matrices with trace 0. Indeed, H_l is obviously toral. It is maximal as such since any element commuting with H_l has to be diagonal.

For the following we fix a maximal toral subalgebra H of L.

Since all elements of H are ad-semisimple and commute, L decomposes under $\mathrm{ad}(H)$ as direct sum of all simultaneous eigenspaces

$$L_\lambda := \{x \in L \,|\, \forall h \in H : \mathrm{ad}h(x) = \lambda(h)x\},$$

where λ runs through H^*. We set

$$R := \{\alpha \in H^* \,|\, L_\alpha \neq 0,\, \alpha \neq 0\}.$$

These are the *roots of* L *relative to* H, and the L_α with α a root are called *the root spaces*. Using the preceding lemma we have $L_0 = H$. Thus, we obtain the so-called *root space decomposition of* L:

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

Note that

$$[L_\lambda, L_\mu] \subset L_{\lambda+\mu}, \qquad (\lambda, \mu \in H^*).$$

This follows immediately from the fact that the elements of $\mathrm{ad}_L(H)$ are derivations of L. This identity has several useful applications. As a first one we note $[L_\lambda, L_{-\lambda}] \subset H$ for all $\lambda \in H^*$. Secondly, for $\lambda, \mu \in H^*$, set

$$S_\lambda := L_\lambda + L_{-\lambda} + [L_\lambda, L_{-\lambda}], \qquad N(\mu, \lambda) := \bigoplus_{j \in \mathbb{Z}} L_{\mu + j\lambda}.$$

Then $N(\mu, \lambda)$ is stable under $\mathrm{ad}_L S_\lambda$. We shall prove in a moment that S_α, for a root α, is a subalgebra of L, which is isomorphic to sl_2. The study of the sl_2-modules $N(\beta, \alpha)$ will show that R obeys a very rigid structure.

Example. Again we consider sl_l. Let a_i denote the element of H_l^* such that $a_i(h) = i$-th diagonal element of h. The elements $\alpha_{i,j} := a_i - a_j$ $(i \neq j)$ are the roots of sl_l relative to H_l. Indeed, let $E_{i,j}$ be the matrix with a 1 at the (i, j)-th place and 0 otherwise. Then, for $h \in H_l$ one has

$$\mathrm{ad}h(E_{i,j}) = \alpha_{i,j}(h)E_{i,j},$$

and the root space decomposition is

$$\mathrm{sl}_l = H_l \oplus \bigoplus_{i \neq j} K \cdot E_{i,j}.$$

Using $E_{i,j}E_{p,q} = \delta_{j,p}E_{i,q}$ one verifies that $[E_{i,j}, E_{j,i}] = E_{i,i} - E_{j,j}$, and that

$$E_{i,j} \mapsto x, \quad E_{j,i} \mapsto y, \quad E_{i,i} - E_{j,j} \mapsto h$$

defines an isomorphism of $S_{i,j} := S_{\alpha_{i,j}}$ with sl_2 (with x, y, h denoting the standard basis of sl_2).

Proposition. *Let $\alpha, \beta \in H^*$ such that $\alpha + \beta \neq 0$. Then $L_\alpha \perp L_\beta$ with respect to the Killing form B of L. In particular, the restriction of B to H is non-degenerate.*

Note that the first part implies that L_α, for a root α, is not perpendicular to $L_{-\alpha}$, since otherwise $L_\alpha \perp L$, contrary to the non-degeneracy of B. In particular, if $\alpha \in R$ the $-\alpha \in R$.

Proof. Let $h \in H$ with $(\alpha + \beta)(h) \neq 0$. Then, for $x \in L_\alpha$ and $y \in L_\beta$, we find

$$\alpha(h)B(x, y) = B(\mathrm{ad}h(x), y) = -B(x, \mathrm{ad}h(y)) = -\beta(h)B(x, y),$$

which implies the first statement.

If $h \in H \cap H^\perp$, then $h \perp L_0$, and from the first statement $h \perp L_\alpha$ for all $\alpha \neq 0$, hence $h \in L^\perp$, whence $h = 0$. □

Using this proposition, we have the isomorphism

$$H^* \to H,\ \lambda \mapsto t_\lambda,\ \text{where } \forall h \in H : B(t_\lambda, h) = \lambda(h).$$

We transfer B to a scalar product on H^* by setting

$$(\lambda, \mu) := B(t_\lambda, t_\mu) = \lambda(t_\mu) = \mu(t_\lambda).$$

Example. Again, for $L = \mathrm{sl}_l$ we calculate

$$\mathrm{ad}h_1 \mathrm{ad}h_2(E_{i,j}) = \alpha_{i,j}(h_1)\, \alpha_{i,j}(h_2) E_{i,j}.$$

Hence,

$$B(h_1, h_2) = \mathrm{tr}(\mathrm{ad}_L h_1 \mathrm{ad}_L h_2) = \sum_{i \neq j} \alpha_{i,j}(h_1)\alpha_{i,j}(h_2).$$

In particular, $t_{p,q} := t_{\alpha_{p,q}}$ is the diagonal matrix with $1/2(l-1)$ at the p-th and $-1/2(l-1)$ at the q-th place on the diagonal.

Theorem. *Let $\alpha \in R$. Then one has*

1. $[L_\alpha, L_{-\alpha}] = K \cdot t_\alpha$.
2. $L_\alpha + L_{-\alpha} + [L_\alpha, L_{-\alpha}] \cong \mathrm{sl}_2$.

Remark. Note that the theorem implies in particular that L_α is one dimensional for all roots α.

In particular, if α and β are linearly independent roots, then $N(\beta, \alpha) = \sum_{j\in\mathbb{Z}} L_{\beta+j\alpha}$ as module over $L_\alpha + L_{-\alpha} + [L_\alpha, L_{-\alpha}] \cong \mathrm{sl}_2$ is isomorphic to $V(N)$ for a suitable N. Indeed, the nontrivial $L_{\beta+j\alpha}$ are weight spaces with pairwise different weights.

Note also that the theorem implies

$$B(h_1, h_2) = \sum_{\alpha\in R} \alpha(h_1)\alpha(h_2).$$

Indeed, $\mathrm{ad}h_1\mathrm{ad}h_2$ acts on the one dimensional spaces L_α as multiplication by $\alpha(h_1)\alpha(h_2)$ and trivial on H.

Proof. ad 1.: Let $x \in L_\alpha$ and $y \in L_{-\alpha}$. Then $[x, y] \in L_{\alpha+(-\alpha)} = H$. Now let $h \in H$. We have

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

Hence $[x, y] - B(x, y)t_\alpha$ is orthogonal to H, hence 0. This, together with the fact that $B(L_\alpha, L_{-\alpha}) \neq 0$, implies 1.

ad 2.: We show first of all that $(\alpha, \alpha) \neq 0$. Choose $x \in L_\alpha$ and $y \in L_{-\alpha}$ such that $B(x, y) = 1$. From 2. we have $[x_\alpha y_\alpha] = t_\alpha$. If $(\alpha, \alpha) = \alpha(t_\alpha)$ were equal to 0, then $[t_\alpha, x_\alpha] = [t_\alpha, y] = 0$. But this would mean that the span S of x_α, y_α and t_α is a solvable algebra, as is then too $\mathrm{ad}_L S \cong S$. By Lie's theorem, $\mathrm{ad}_L S$ is then isomorphic to a subalgebra of the algebra of upper triangular matrices. In particular, $[SS]$ consists of nilpotent elements. Hence $\mathrm{ad}t_\alpha \in [SS]$ is nilpotent and semisimple, hence 0, hence $t_\alpha \in \mathrm{Z}(L) = 0$. Since $\alpha \neq 0$ and the restriction of B to H is nondegenerate, this is impossible.

Let $x_\alpha \in L_\alpha$, choose a $y \in L_{-\alpha}$ such that $B(x_\alpha, y_\alpha) = 2/(\alpha, \alpha)$. Then $[x_\alpha y_\alpha] = 2t_\alpha/(\alpha, \alpha) =: h_\alpha$. One has $[h_\alpha, x_\alpha] = \alpha(h_\alpha)x_\alpha = 2x_\alpha$, and similarly $[h_\alpha, y_\alpha] = -2y_\alpha$. Thus $S_\alpha := L_\alpha + L_{-\alpha} + [L_\alpha, L_{-\alpha}]$ contains in any case a subalgebra Σ_α isomorphic to sl_2.

Let $M := \bigoplus L_{a\alpha}$, where a runs though all $a \in K$ such that $a\alpha$ is a root (or 0). As the $N(\lambda, \alpha)$ introduced above, M is an Σ_α-module with respect to the adjoint representation. (In fact, in a moment it will turn out that $\pm\alpha$

are the only multiples of α which are roots; hence M is nothing else than $N(0,\alpha)$.)

Thus, by Weyl's theorem, M decomposes

$$M \cong \bigoplus_{N\geq 0} V(N)^{\mu(N)},$$

with multiplicities $\mu(N)$. Note that h_α acts on $L_{a\alpha}$ as multiplication by $2a$ (since $\alpha(h_\alpha) = 2a$). In particular, $L_{a\alpha} = 0$, unless $a \in \frac{1}{2}\mathbb{Z}$. Moreover, the weight space M_0 of h_α equals H. Hence

$$\sum_{N \text{ even}} \mu(N) = \dim H.$$

But M contains $\dim H - 1$ copies of $V(0)$ (namely $\ker\alpha$, where L, and, in particular, Σ_α acts trivially on), and one copy of $V(2)$ (namely Σ_α). Thus, if we let M_{ev} be the part of M belonging to the even N, we find

$$M_{\text{ev}} = H \oplus \Sigma_\alpha.$$

Hence the only even weights occurring in M are 0 and ± 2. In particular, 2α is not a root.

Since this is true for all roots α, we see that $\frac{1}{2}\alpha$ cannot be a root (since α is one). Hence 1 is not a weight of h_α in M.

This shows $M = M_{\text{ev}}$, thus $M = H \oplus \Sigma_\alpha$, and hence completes the proof. □

Let α be a root. As we have seen in the preceding proof $B(t_\alpha, t_\alpha) \neq 0$. We set in the sequel

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

Note that

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

The next goal is to describe the structure of R equipped with the scalar product $(\ ,\ )$.

Example. For sl_l the $\mathbb{Z}$-submodule of H_l^* generated by the roots

$$R_l := \{\alpha_{i,j} \mid i \neq j\}$$

has rank equal to $l-1 = \dim H_l - 1$. As basis one may take $\Delta_l := \{\alpha_{i,i+1} \mid 1 \leq i \leq l-1\}$. Moreover, one has

$$(\alpha_{i,j}, \alpha_{p,q}) = \frac{1}{2(l-1)}(\delta_{i,p} - \delta_{j,p} - \delta_{i,q} + \delta_{j,q}).$$

In particular, $(\ ,\ )$ is rational and positive definite on $\mathbb{Z}[R_l]$.

Definition. A subset R of an euclidean space E is called root system of E if the following conditions are satisfied:

1. R is finite, spans E and does not contain 0.
2. $\forall \alpha \in R \forall a \in \mathbb{R} : (a\alpha \in R \Longleftrightarrow a = \pm 1)$.
3. $\forall \alpha \in R : \sigma_\alpha(R) = R$. Here $\sigma_\alpha(x) = x - \frac{2(x,\alpha)}{(\alpha,\alpha)}\alpha$ is the reflection with respect to $\alpha^\perp$.
4. $\forall \alpha, \beta \in R : \frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$.

The number $l = \dim E$ is called the rank of R.

Theorem. *Let R be the set of roots of L with respect to H. Then*

1. $\mathbb{Z}[R] := \mathrm{span}_{\mathbb{Z}} R \subset H^*$ *is a complete lattice, i.e.* $a \otimes \alpha \mapsto a\alpha$ *defines an isomorphism* $K \otimes \mathbb{Z}[R] \cong H^*$.
2. *For* $\lambda, \mu \in \mathbb{Z}[R]$ *one has* $(\lambda, \mu) \in \mathbb{Q}$, *and* $(\lambda, \lambda) > 0$. *In particular,* $E := \mathbb{Z}[R] \otimes \mathbb{R}$ *equipped with the scalar product induced by* $(\lambda, \mu) = B(t_\lambda, t_\mu)$ *is an euclidean space*
3. R *is a root system of* E.

Proof. For a root α, we studied in the preceding theorem the action of $S_\alpha = L_\alpha + L_{-\alpha} + [L_\alpha, L_{-\alpha}]$. on $\sum_{a \in K} L_{j\alpha}$, and we obtained that the only multiples of α which are roots are $\pm\alpha$. This is axiom 2. of a root system.

We now let β be a root different from $\pm\alpha$. Then S_α acts on $N := N(\beta, \alpha) = \sum_j L_{\beta + j\alpha}$. Since h_α acts as multiplication by $\beta(h_\alpha) + 2j$ on $L_{\beta+j\alpha}$, we find that the nonzero $L_{\beta+j\alpha}$ are the weight spaces of h_α, and that $\beta + j\alpha$ is a root if and only if $\beta(h_\alpha) + 2j$ is a weight of h_α.

Since weights are integral we deduce $\beta(h_\alpha) = 2(\alpha, \beta)/(\alpha, \alpha) \in \mathbb{Z}$, whence axiom 4.

Moreover, β is a root. Hence $\beta(h_\alpha)$ is a weight of h_α, hence $-\beta(h_\alpha)$ is a weight. But $-\beta(h_\alpha) = (\beta - 2\beta(h_\alpha)\alpha)(h_\alpha)$. In other words, $\beta - \beta(h_\alpha)\alpha$ is a root, whence axiom 3.

To show 1. we remark first of all that R spans H^*. In fact, otherwise there would be a $0 \neq h \in H$ such that $R(h) = 0$. But then $\mathrm{ad}_L h = 0$, i.e. $h \in \mathrm{Z}(L) = 0$. Thus the map in 1., defined by $a \otimes \alpha \mapsto a\alpha$, is surjective.

It is injective since the rank of $\mathbb{Z}[R]$ is $l := \dim_K H^*$. For this it suffices to show that

$$\mathbb{Z}[R] \subseteq \frac{1}{N}(\mathbb{Z}\alpha_1 + \cdots + \mathbb{Z}\alpha_l),$$

with a suitable integer N, where $\alpha_1, \ldots, \alpha_l \in R$ is a basis of H^*. For this it suffices to show that each $\beta \in R$ is a rational linear combination of the α_i.

For proving this write $\beta = \sum a_j \alpha_j$. Then

$$(\beta, \alpha_k)/(\alpha_k, \alpha_k) = \sum a_j (\alpha_j, \alpha_k)/(\alpha_k, \alpha_k).$$

But the matrix of this system of linear equations (in the unknowns a_j) is regular (since it is essentially the Gram matrix of B on H, which is nondegenerate), By what we have already proved, the coefficients of these linear equations are integers. Hence the a_j are rational numbers.

Finally, for proving 2. let $\lambda, \mu \in H^*$. We already saw that

$$(\lambda, \mu) = B(t_\lambda, t_\mu) = \sum_{\alpha \in R} \alpha(t_\lambda)\alpha(t_\mu) = \sum_{\alpha \in R} (\alpha, \lambda)(\alpha, \mu).$$

For $\beta \in R$ we thus have

$$(\beta|, \beta|)^{-1} = \sum_{\alpha \in R} (\alpha, \beta)^2/(\beta, \beta)^2.$$

Using axiom 4, we obtain $1/(\beta, \beta) \in \mathbb{Q}$. But then, the form $(\ ,\ )$ takes values in $\mathbb{Q}$ on $\mathbb{Z}[R]$, and again, by the last but not least formula, each (λ, λ) $(\lambda \in \mathbb{Z}[R])$ is a sum of squares of rational numbers, hence positive. □

3.1.1 Appendix: Cartan subalgebras

To sum up, we have associated to L the root system R in $E = \mathbb{R} \otimes \mathbb{Z}[R]$. A priori this construction depends on the choice of the maximal toral subalgebra H. In this appendix we explain how much two root systems associated via two different maximal toral subalgebras differ.

A maximal toral subalgebra H is a *Cartan subalgebra*, i.e. a nilpotent subalgebra which equals its normalizer in L (i.e. a subalgebra C such that $\mathrm{ad}_C C$ consists of nilpotent elements only, and such that for all $x \in L$ the inclusion $[xC] \subseteq C$ implies $x \in C$). Note that this definition makes sense for an arbitrary finite dimensional Lie algebra, not only for the ones considered in this section.

Clearly, H is nilpotent, since it is abelian. To see that H equals its normalizer in L let $x = h + \sum z_\alpha \in L$ ($h \in H$ and $z_\alpha \in L_\alpha$) such that $[xH] \subset H$. But for $k \in H$ one has $[xk] = \sum \alpha(k) z_\alpha$, which is in H only if all $z_\alpha = 0$.

As it turns out, any two Cartan subalgebras are conjugate by a suitable automorphism of L. Thus the notions of maximal toral and Cartan subalgebra coincide under the general assumptions on L and K listed at the beginning.

In fact, any two Cartan subalgebras are even conjugate under inner automorphisms of L. We digress to explain this in more detail:

If $x \in L$ is ad$x$-nilpotent, then

$$\exp(\mathrm{ad}x) := 1 + \mathrm{ad}x + \frac{(\mathrm{ad}x)^2}{2} + \cdots$$

makes sense. In fact, $\exp(\mathrm{ad}x)$ defines even an automorphism of L. This is true for any nilpotent derivation δ of any algebra A, since one has

$$(\exp\delta)(y) \cdot (\exp\delta)(z) = (\exp\delta)(yz) \quad (x, y \in A), \exp(-\delta)(\exp\delta) = 1.$$

These formulas can be deduced by a straight forward calculation. If L is a subalgebra of $\mathrm{gl}(V)$ and x nilpotent, then one can easily verify

$$(\exp \mathrm{ad}x)(y) = \exp(x)\, y\, \exp(x)^{-1}.$$

Theorem. *Any two Cartan subalgebras C and C' of L are conjugate by an inner automorphism of L (i.e. there exist an ad-nilpotent $x \in L$ such that $C' = (\exp \mathrm{ad}x)(C)$).*

Proof. [Humphreys] p. 84 □

As indicated above we have as immediate consequence

Proposition. *Any maximal toral subalgebra of L is a Cartan subalgebra and vice versa.*

Let now H and H' two Cartan subalgebras of L (or equivalently, two maximal toral subalgebras), and R and R' their associated root systems. Then there is an ad-nilpotent x such that $\phi = \exp(\mathrm{ad}x)$ maps H to H'. By a straightforward calculation one verifies that R' is bijectively mapped to R by the dual map $\phi^* : (H')^* \to H^*$. Moreover, using that the Killing form of L is invariant under ϕ (use that for any automorphism ϕ of L one has $\mathrm{ad}\phi(y) = \phi\, \mathrm{ad}y\, \phi^{-1}$) we see that ϕ^* leaves invariant the numbers $2(\alpha, \beta)/(\beta, \beta)$. We thus define

Definition. Two root systems R of E and and R' of E' are equivalent, if there is an isomorphism of vector spaces $\sigma : E \to E'$ which leaves invariant the *Cartan integers*, i.e. such that

$$\frac{2(\sigma(\alpha), \sigma(\beta))}{(\sigma(\beta), \sigma(\beta))} = \frac{2(\alpha, \beta)}{(\beta, \beta)} \qquad (\alpha, \beta \in R).$$

As a result of the discussion in this section we have

Theorem. *The root systems associated to L by all different choices of Cartan subalgebras are pairwise equivalent.*

3.2 Root systems

Let R be a root system in the Euclidean space E. For a pair of roots α, β write $(\alpha, \beta) = |\alpha| \cdot |\beta| \cos\theta$ with a suitable $0 \le \theta \le \pi$. By axiom 4. we see that the numbers

$$4\cos^2\theta = 2\frac{(\alpha,\beta)}{|\alpha|^2} \cdot 2\frac{(\alpha,\beta)}{|\beta|^2}$$

are all integers. Since $\cos^2\theta \le 1$, we thus have only 5 possibilities for the value of $4\cos^2\theta$. Assume $|\alpha| \ge |\beta|$ and $\alpha \ne \pm\beta$. Then the only possibilities for $4\cos^2\theta$ are given by the following table:

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

Definition. The Weyl group W of R is the subgroup of $\mathrm{GL}(E)$ generated by the reflections

$$\sigma_\alpha(x) = x - \frac{2(x,\alpha)}{(\alpha,\alpha)}\alpha \qquad (\alpha \in R).$$

A subset Δ of R is called a root basis, if it is a basis for E, and if each root is an integral linear combination of roots in Δ with all coefficients negative or positive. The elements of Δ are called simple (or fundamental) roots.

Clearly, W can be identified with a subgroup of the group of permutations of R. In particular, it is finite.

Proposition. *R possesses a basis. The Weyl group acts transitively on all basis. Let Δ be a basis. The Weyl group is generated by the reflections σ_α with $\alpha \in \Delta$, and*

$$R = \{\sigma_\alpha(\beta) \,|\, \alpha, \beta \in \Delta\}.$$

Moreover, each root β can be written as $\beta = \epsilon(\alpha_1 + \cdots + \alpha_k)$ with $\epsilon = \pm 1$ and with (not necessarily distinct) roots $\alpha_i \in \Delta$ such that each partial sum $\beta = \epsilon(\alpha_1 + \cdots + \alpha_s)$ $(1 \le s \le k)$ is a root.

Proof. [Humphreys],§10. □

Definition. Let $\Delta = \{\alpha_1, \ldots, \alpha_l\}$ be a basis of R. The matrix

$$\left(\frac{2(\alpha_i,\alpha_j)}{(\alpha_j,\alpha_j)}\right)_{1\le i,j\le l} = \begin{pmatrix} 2\frac{(\alpha_1,\alpha_1)}{|\alpha_1|^2} & 2\frac{(\alpha_1,\alpha_2)}{|\alpha_2|^2} & 2\frac{(\alpha_1,\alpha_3)}{|\alpha_3|^2} & \cdots \\ 2\frac{(\alpha_2,\alpha_1)}{|\alpha_1|^2} & 2\frac{(\alpha_2,\alpha_2)}{|\alpha_2|^2} & \cdots & \\ \vdots & & & \end{pmatrix}$$

is called the Cartan matrix of R.

Proposition. *Two Cartan matrices C and C' associated to equivalent root systems are equal up to a permutation of indices (i.e. $C = PC'P^{-1}$ with a suitable permutation matrix P).*

Proof. This is clear from the notion of equivalence of root systems, since the Weyl group acts transitively on basis, and since the elements of the Weyl group are isometric mappings. □

Example. For sl_l we can take $\Delta_l = \{\alpha_{i,i+1} \,|\, 1 \le i \le l-1\}$ as basis. Then

$$(\alpha_{i,i+1}, \alpha_{j,j+1}) = \frac{1}{l-1}(\delta_{i,j} - \frac{1}{2}\delta_{i,j+1} - \frac{1}{2}\delta_{i+1,j}).$$

Thus the corresponding Cartan matrix has 2's on the diagonal, -1's on the first side diagonals, and 0 at all other places.

On the diagonal of a Cartan matrix we always find 2. It can easily be shown that all other entries are ≤ 0, hence equal to $0, -1, -2$ or -3.

Proposition. *If $\alpha \ne \beta$ are roots of some root basis of R, then $(\alpha, \beta) < 0$.*

Proof. If $(\alpha, \beta) \ge 0$, then a look at the table shows that $b := 2(\alpha,\beta)/|\beta|^2$ or $a := 2(\beta,\alpha)/|\alpha|^2$ equal 1. Say $b = 1$. Then $\sigma_\beta(\alpha) = \alpha - \beta$ is a root, contradicting the axioms of a basis of a root system. Similarly, if $a = 1$, one deduces the contradiction $\beta - \alpha \in R$. □

Proposition. *The Cartan matrix uniquely determines R up to isomorphism.*

Proof. Let R and R' two root systems in E and E' with basis α_i and α'_i such that the associated Cartan matrices coincide. Let ϕ be the isomorphism of E with E' defined by $\alpha_i \mapsto \alpha'_i$ (for all i). Then $\sigma_{\phi(\alpha_i)} = \phi \circ \sigma_{\alpha_i} \circ \phi^{-1}$. Since R is the set $\sigma_{\alpha_i}(\alpha_j)$ (all i and j) (and similarly for R') we see that ϕ maps R bijectively to R', and that $\sigma_{\phi(\alpha_i)} = \phi \circ \sigma_{\alpha_i} \circ \phi^{-1}$ for any root α. The latter formula also implies that ϕ leaves the numbers $2(\beta,\alpha)/(\alpha,\alpha)$ invariant. □

One often describes the Cartan matrix of a given root system R by a certain graph, the so-called Dynkin diagram of R. To each simple root corresponds a vertex. Two vertices are connected by $4\cos^2\theta \in \{0,1,2,3\}$ lines, where θ is their angle. If i, j are vertices such that $|\alpha_j| > |\alpha_i|$ one puts an arrow on the lines pointing from j to i. It is clear how to discover the *Cartan integers* (i.e. the elements of the Cartan matrix)

$$c := \frac{2(\alpha,\beta)}{|\beta|^2}$$

from the Dynkin diagram : if $\alpha = \beta$ then $c = 2$; if $\alpha \neq \beta$ and they are not connected then $c = 0$; if $\alpha \neq \beta$, if they are connected and if $|\alpha| \geq |\beta|$, then $c = -1$; if $\alpha \neq \beta$, if they are connected by i edges, and if $|\alpha| \leq |\beta|$, then $c = -i$ (cf. the table above).

We call a root system (and its Dynkin diagram and Cartan matrix) *simple* if its Dynkin diagram is connected. Equivalently, R is simple if its Cartan matrix is a simple matrix, i.e. if it cannot be written (after permutation of indices) as direct sum

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

with square matrices A_1 and A_2 (having a positive number of lines each). Finally, R is simple if and only if R cannot be written as disjoint union $R = R_1 \cup R_2$ of nonempty subsets $R_1 \perp R_2$. Note that in such a decomposition R_1 and R_2 are root systems in the subspaces of E spanned by them.

Thus any root system can be written as union of simple ones which are pairwise orthogonal. Or else, any Cartan matrix can be decomposed as direct sum of simple ones. Thus, for classifying all root systems, it suffices to describe the simple ones.

Let FIN be the following set

$$\begin{aligned} \mathrm{FIN} := \{C = (a_{i,j}) \text{ simple square matrix} \,|\, \forall i : a_{i,i} = 2, \\ \forall i \neq j : a_{i,j} = 0, -1, -2, -3,\ a_{i,j} = 0 \Longrightarrow a_{j,i} = 0, \\ \exists D \text{diagonal matrix} : CD \text{ is symmetric and positive definite}\}. \end{aligned}$$

Clearly, the simple Cartan matrices belong to FIN. One can inductively determine FIN. As it turns out, all matrices in this resulting list are Cartan matrices. Hence FIN coincides with the set of simple Cartan matrices of root systems. This leads to a classification of all simple root systems. See the table section for the complete list.

The first (simple) specimen in FIN can immediately be written down:

$$\mathrm{A}_1 = \begin{pmatrix} 2 \end{pmatrix}, \quad \mathrm{A}_2 = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad \mathrm{B}_2 = \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}, \quad \mathrm{G}_2 = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}.$$

(We put the classical names in front).

3.3 Serre's Theorem and consequences

In the preceding two sections we have constructed a map

$$\{\text{semisimple Lie algebras}\}/\cong \quad \to \quad \{\text{Cartan matrices of root systems}\}/\cong,$$

where the second $\cong$ means "up to permutation of indices". This map is in fact a bijection. This can handily be seen by a construction known as Serre's theorem. We shall not exactly prove Serre's theorem, but instead we shall prove in the broader context of Kac-Moody algebras a slightly weaker theorem, which however suffices to deduce the bijectivity of the above map.

Theorem. *Let L be a semisimple Lie algebra, H a Cartan subalgebra, R the root system with respect to H, and Δ be a basis of R. For a root α fix $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$ such that $[x_\alpha\, y_\alpha] = h_\alpha$. Then L is generated (as a Lie algebra) by x_α, y_α ($\alpha \in \Delta$), and, for $\alpha, \beta \in \Delta,$, $\alpha \neq \beta$, one has*

1. $[h_\alpha\, h_\beta] = 0.$
2. $[x_\alpha\, y_\alpha] = h_\alpha$ *and* $[x_\alpha\, y_\beta] = 0.$
3. $[h_\alpha\, x_\beta] = \beta(h_\alpha)\, x_\beta$ *and* $[h_\alpha\, y_\beta] = -\beta(h_\alpha)\, y_\beta.$
4. $(\mathrm{ad} x_\alpha)^{-\beta(h_\alpha)+1}(x_\beta) = 0.$
5. $(\mathrm{ad} y_\alpha)^{-\beta(h_\alpha)+1}(y_\beta) = 0.$

In the following we shall refer to the relations 1.–5. as Weyl relations. Note that they depend only on the Cartan matrix

$$C := \big(a_{\beta,\alpha}\big)_{\beta,\alpha\in\Delta} \qquad (a_{\beta,\alpha} = \beta(h_\alpha) = \frac{2(\beta,\alpha)}{(\alpha,\alpha)}).$$

The $x_\alpha, y_\alpha, h_\alpha$ ($\alpha \in \Delta$) are called the Chevalley generators of L (with respect to Δ).

Proof. The Lie algebra L' generated by the x_α, y_α contains H since it contains the basis $h_\alpha = [x_\alpha\, y_\alpha]$ of H. It also contains $L_\alpha = \mathbb{C}\cdot x_\alpha$ and $L_{-\alpha} = \mathbb{C}\cdot y_\alpha$ for any fundamental α. Let β be an arbitrary root. Then $\beta = \epsilon(\alpha_1+\cdots+\alpha_s)$ with simple roots α_i, with $\epsilon = \pm 1$, and such that each partial sum $\epsilon(\alpha_1+\cdots+\alpha_k)$ ($k \le s$) is a root. If α, β are roots, then $[L_\alpha, L_\beta] = [x_\alpha, L_\beta] = L_{\alpha+\beta}$ since

$$N(\alpha,\beta) = \sum_{j\in\mathbb{Z}} L_{\beta+j\alpha}$$

as module over $S_\alpha := \mathrm{span}\{x_\alpha, y_\alpha, h_\alpha\} \cong \mathrm{sl}_2$, is isomorphic to $V(N)$ with a suitable N, and since the nontrivial $L_{\beta+j\alpha}$ are the weight spaces of h_α. Thus we find inductively $[L_1[L_2[\dots L_k]\dots] = L_\beta$, where we have put $L_i = L_{\epsilon\alpha_i}$.

(1) holds true since H is abelian. The second identity of (2) is true since $[x_\alpha\, y_\beta] \in L_{\alpha-\beta}$, and since $\alpha - \beta$ is not a root (recall that any root is linear combination of the α with all coefficients either negative or either positive). (3) is clear from the definition of the $x_\alpha, y_\alpha, h_\alpha$. (4) follows from the fact that the element on the left is in L_γ, where $\gamma = \beta + (-\beta(h_\alpha)+1)\alpha = \sigma_\alpha(\beta - \alpha)$. with respect to H. But $\sigma_\alpha(\beta-\alpha)$ is not a root since $\beta-\alpha$ is not. (5) follows similarly. □

Let now R be an arbitrary root system with Cartan matrix A. Let Δ be a basis. Denote by $L(A)$ the Lie algebra with generators $X_\alpha, Y_\alpha, H_\alpha$ ($\alpha \in \Delta$) and defining relations 1.–5. (with $x_\alpha, \dots$ replaced by $X_\alpha, \dots$). We explain in the next section what this precisely means. Then one has

Theorem. *(Serre's Theorem.) The Lie algebra $L(A)$ is semisimple, and its Cartan matrix is A (up to permutation of indices).*

Proof. [Serre], Chapitre VI, Appendice □

Serre's theorem states that the map "Lie algebra (modulo $\cong$) $\mapsto$ Cartan matrix (modulo $\cong$)" considered at the beginning is surjective. But it implies also the injectivity.

Corollary. *A semisimple Lie algebra is uniquely determined up to isomorphism by its associated Cartan matrices.*

Proof. Let the notations as in the first theorem of this paragraph, and let A the Cartan matrix of L. Let $L(A)$ be the Lie algebra from Serre's theorem. From the definition of the latter (cf. next section for details) it will be clear then that there is a unique Lie algebra morphism $L(A) \to L$ mapping the $X_\alpha, Y_\alpha, H_\alpha$ to $x_\alpha, y_\alpha, h_\alpha$, respectively. This map is clearly surjective. On the other hand, both Lie algebras have equivalent root systems (since the have the same Cartan matrices). Thus if r is the number of roots and l the number of lines of the Cartan matrix, then both Lie algebras have dimension $r + l$. Hence the map $L(A) \to L$ is actually an isomorphism of Lie algebras, which proves the corollary. □

Without proof we note what one might expect: under the bijection of the beginning of this section simple Cartan matrices correspond to simple Lie algebras. This follows from the fact that, for any two Cartan matrices A_1 and A_2, one has $L(A_1 \oplus A_2) \cong L(A_1) \oplus L(A_2)$. This in turn can be easily deduced from the precise definition of $L(A)$.

Chapter 4

Kac-Moody algebras

4.1 Definitions

Serre's theorem gives a description of a semisimple Lie algebra in terms of its Cartan matrix. It lies at hand to take this description as a starting point for a larger class of Lie algebras by admitting more general matrices. This leads in general to infinite dimensional Lie algebras, which, however, show many analogies with the semisimple ones. The ground field K is from now on the field of complex numbers $\mathbb{C}$ as in [Kac], though almost all considerations which follow could be done over or generalized to an arbitrary field of characteristic 0.

To start with let $A = (a_{i,j})$ be an arbitrary complex $n \times n$ matrix. Let l be the rank of A. We assume that the first l columns of A are linearly independent. Denote the matrix formed by these columns by A_1, and the matrix of the remaining columns by A_2. Let

$$C = \begin{pmatrix} A_1 & A_2 \\ 0 & \mathrm{I}_{n-l} \end{pmatrix}$$

We set

$$H := \mathbb{C}^{2n-l} \qquad (= \text{ column vectors}),$$

we denote by h_j the j-th column of C, and we let $\alpha_i \in H^*$ the i-th coordinate function. Clearly one has

$$a_{i,j} = \alpha_i(h_j) \qquad (1 \le i, j \le n).$$

Note that the α_i and the h_j are linearly independent. We set

$$\Delta = \{\alpha_1, \dots, \alpha_n\}.$$

The α_i are supposed to play the analogue of a root basis of a semisimple Lie algebra, the h_j the role of the corresponding h_{α_j} from that theory, and A corresponds to the Cartan matrix. (When reading the important book [Kac] note that the corresponding A in that book are the transposes of ours.) Accordingly, we shall set $h_{\alpha_j} := h_j$.

Note also that, for having the h_j and α_i linearly independent, we had to define them in a $2n-l$-dimensional space H. (Indeed, A being of rank l, the spaces S and Σ generated by the h_j and α_i, respectively, being of dimension n, means that $\dim S \cap \Sigma^\perp = n-l$; the formula $\dim \Sigma + \dim \Sigma^\perp = \dim H$ thus implies $\dim H \geq 2n-l$; hence our H is the minimal choice). If $\det A \neq 0$, as for Cartan matrices, then the h_j span H.

Complete the set of the h_α to a basis B of H. Denote by $\widetilde{L}(A)$ the Lie algebra generated by the elements of B and elements X_α, Y_α ($\alpha \in \Delta$), and with defining relations ($\alpha, \beta \in \Delta$, $h, h' \in B$):

1. $[X_\alpha, Y_\beta] = \delta_{\alpha,\beta} h_\alpha$.
2. $[h, h'] = 0$
3. $[h, X_\alpha] = \alpha(h) X_\alpha$.
4. $[h, Y_\alpha] = -\alpha(h) Y_\alpha$.

This Lie algebra is uniquely characterized by the following property: For every Lie algebra M and any family ϕ of elements $\phi(h)$, $\phi(X_\alpha)$, $\phi(Y_\alpha)$ of M ($h \in B$, $\alpha \in \Delta$) which satisfy relations 1.–4. (with the symbol s replaced by $\phi(s)$) there exists a unique extension to a homomorphism of Lie algebras $\widetilde{L}(A) \to M$. Obviously, this characterization, determines $\widetilde{L}(A)$ uniquely up to isomorphism.

A special but important case is $M = \mathrm{gl}(V)$: More precisely, if $\rho(s)$ ($s \in S$) is a family of endomorphisms of a vector space V satisfying relations 1.–4., then there is a unique extension of ρ to a representation $\widetilde{L}(A) \to \mathrm{gl}(V)$.

The existence of $\widetilde{L}(A)$ is ensured by the following construction: Let

$$T := \bigoplus_{n \geq 0} \mathbb{C}[S]^{\otimes n}$$

be the tensor algebra of $\mathbb{C}[S]$, where $\mathbb{C}[S]$ is the vector space of formal sums of elements from the set S of pairwise different symbols κ_h, ξ_α, η_α ($h \in B$, $\alpha \in \Delta$). Let L be the Lie subalgebra of T generated by the elements of S, and let I be the ideal of L generated by the relations 1.–4. (with h replaced by κ_h etc.). Set $\widetilde{L}(A) = L/I$.

It is easily shown that $\widetilde{L}(A) = L/I$ satisfies the property explained in the preceding paragraph, where h (by abuse of notation), X_α, Y_α denote the images of κ_h, ξ_α, η_α under the canonical map, respectively.

For stating a first theorem concerning the structure of $\widetilde{L}(A)$ we need two more notions.

We say that a Lie algebra L is freely generated by a subset S, if any map from S into a Lie algebra M can be uniquely extended to a Lie algebra homomorphism $L \to M$. This is equivalent to saying that L is isomorphic to the Lie subalgebra generated by S of the tensor algebra of $\mathbb{C}[S]$ (viewed as Lie algebra via $[x\,y] = xy - yx$).

Lemma. *(Diagonalization Lemma) Let H be an abelian Lie algebra and V be a diagonalizable H-module, i.e. $V = \bigoplus_{\lambda \in H^*} V_\lambda$, with*

$$V_\lambda = \{x \in V \mid \forall h \in H : h \cdot x = \lambda(h)x\}.$$

Then for any submodule U of V one has $U = \bigoplus_{\lambda \in H^*}(U \cap V_\lambda)$.

Note that the lemma implies that quotients V/U of diagonalizable modules are diagonalizable. Indeed, using the lemma,

$$V/U \cong \bigoplus_{\lambda \in H^*} V_\lambda/(U \cap V_\lambda).$$

Proof. Let $u \in U$, write $u = \sum_{j=1}^m v_j$ with $v_j \in V_{\lambda_j}$. We have to show $v_j \in U$. For this choose $h \in H$ such that the $\lambda_j(h)$ are pairwise distinct. Then $h^k \cdot u = \sum_j \lambda_j(h)^k v_j \in U$ $(0 \le k \le m)$. The matrix $(\lambda_j(h)^k)$ is nonsingular, hence we may write the v_j as linear combinations of the $h^k u$. □

Theorem. *The subalgebras $\widetilde{N}_+$ and $\widetilde{N}_-$ of $\widetilde{L}(A)$ generated by the X_α and Y_α $(\alpha \in \Delta)$, respectively, are freely generated by these elements. The vector subspace of $\widetilde{L}(A)$ spanned by the elements of the basis B of H has dimension n (and will henceforth be identified with H). One has*

$$\widetilde{L}(A) = \widetilde{N}_- \oplus H \oplus \widetilde{N}_+$$

(as direct sum of vector subspaces). Moreover, $\widetilde{L}(A)$ is diagonalizable as H-module. For each $\lambda \in H^*$ the space $L(A)_\lambda$ is finite dimensional. One has

$$\widetilde{N}_- = \bigoplus_{\alpha \in Q_+} L_{-\alpha}, \qquad \widetilde{N}_+ = \bigoplus_{\alpha \in Q_+} L_\alpha,$$

where $Q_+ = \left(\sum_{\alpha \in \Delta} \mathbb{Z}_{\ge 0}\, \alpha\right) \setminus \{0\}$.

Proof. We consider the following tensor algebra

$$T_- := \bigoplus_{n\geq 0} \mathbb{C}[\eta_\alpha | \alpha \in \Delta]^{\otimes n},$$

where η_α ($\alpha \in \Delta$) are pairwise different symbols.

We define an action of $\widetilde{L}(A)$ on T_-. For this it suffices to define the action of the generators of $\widetilde{L}(A)$. Choose a $\lambda \in H^*$. The actions of Y_α, $h \in B$ and X_β are inductively defined by

$$\begin{aligned} Y_\alpha \cdot v &:= \eta_\alpha \otimes v, \\ h \cdot 1 := \lambda(h), \quad h \cdot (Y_\alpha \otimes v) &:= -\alpha(h) Y_\alpha \otimes v + Y_\alpha \otimes h \cdot v \\ X_\beta \cdot 1 := 0, \quad X_\beta \cdot (Y_\alpha \otimes v) &:= \delta_{\alpha,\beta} h_\alpha \cdot v + Y_\alpha \otimes X_\beta \cdot v \end{aligned}$$

It is easily checked that the relations 1.–4. hold with X_α, Y_α, h replaced by the corresponding endomorphisms of T_- defined by the action of the X_α, Y_α, h. Hence, these actions can be uniquely extended to an action of $\widetilde{L}(A)$ on T_-.

In particular, we find that $n \mapsto n \cdot 1$ defines a morphism of Lie algebras $\widetilde{N}_- \to T_-$. But T_-, as tensor algebra, is freely generated by the $\eta_\alpha = Y_\alpha \cdot 1$ (as associative algebra with unit), i.e. any family of elements $\psi(\eta_\alpha)$ from an associative algebra with unit $\mathcal{A}$ can be extended to an homomorphism of associative algebras $L \to \mathcal{A}$ mapping the unit 1 to 1. Thus, for any morphism of Lie algebras $\phi : \widetilde{N}_- \to \mathcal{A}$ into an associative algebra $\mathcal{A}$ there exists a unique homomorphism of algebras $\underline{\phi} : T_- \to \mathcal{A}$ such that $\underline{\phi}(n \cdot 1) = \phi(n)$. This is equivalent to saying that T_- is the universal enveloping algebra of $\widetilde{N}_-$. Thus $n \mapsto n \cdot 1$ is injective, and $\widetilde{N}_-$ is isomorphic to the Lie subalgebra of T_- generated by the η_α, whence freely generated by the Y_α.

Denote for a moment the subspace of $\widetilde{L}(A)$ spanned by the element of B by $\widetilde{H}$. Clearly, $h \mapsto h$ ($h \in B$) defines a homomorphism of vector spaces $f : H \to \widetilde{H}$. But $\widetilde{H} \ni \widetilde{h} \mapsto \widetilde{h} \cdot 1$ defines an element of $\widetilde{H}^*$, which is a continuation of $B \ni h \mapsto \lambda(h)$. This is possible only if $\ker f \subseteq \ker \lambda$. Since $\lambda \in H^*$ is arbitrary, we deduce $\ker f = 0$. In the sequel we identify H with its image in $\widetilde{L}(A)$.

By induction on s and using the relation 1.–4. one can show that a (Lie) product of s generators of $\widetilde{L}(A)$ is in $\widetilde{N}_- + H + \widetilde{N}_+$. To show that these sum is direct let $x = n_- + h + n_+ = 0$ with $n_- \in \widetilde{N}_-, \ldots$. Then $x \cdot 1 = n_- \cdot 1 + \lambda(h) = 0$. But $n_- \cdot 1$ is an element of the Lie subalgebra of T_- generated by the η_α, hence lies not in K (use that T_- is a graded algebra). Thus $n_- \cdot 1 = 0$, whence $n_- = 0$, and then also $\lambda(h) = 0$. Since λ is arbitrary also $h = 0$, hence finally $n_+ = 0$.

To prove that $\widetilde{N}_+$ is freely generated by the X_α one can check that $X_\alpha \mapsto -Y_\alpha$, $Y_\alpha \mapsto -X_\alpha$ and $h \mapsto -h$ defines an Lie algebra homomorphism (an involution) of $\widetilde{L}(A)$, which maps $\widetilde{N}_-$ to $\widetilde{N}_+$.

From the relations 4. it follows that

$$[Y_{\alpha_1}, [Y_{\alpha_2}[\dots Y_{\alpha_n}]\dots] \in \big(\widetilde{N}_-\big)_{-\alpha_1-\alpha_1-\dots-\alpha_n}.$$

Moreover, any element of $\widetilde{N}_-$ is a sum of such elements (use the antisymmetry of the Lie bracket to write an arbitrary Lie product of n elements in this form). Hence $\widetilde{N}_-$ is H-diagonalizable (with respect to the adjoint action of H) with finite dimensional weight spaces $(\widetilde{N}_-)_\lambda$, and $(\widetilde{N}_-)_\lambda = 0$ unless $-\lambda \in Q_+$.

The analogue statement for $\widetilde{N}_+$ (and with X_α instead of Y_α) then shows that $\widetilde{L}(A)$ is H-diagonalizable with

$$\dim\big(\widetilde{L}(A)\big)_\lambda \leq \deg(\lambda),$$

where we use $\deg(\sum_{\alpha\in\Delta} c_\alpha \alpha) = \sum c_\alpha$. □

Let $I(A)$ be the sum of all ideals of $\widetilde{L}(A)$ intersecting H trivially. Since every ideal of $\widetilde{L}(A)$ is H-diagonalizable, we easily deduce that $I(A)$ intersects H trivially.

On the other hand,

Proposition. *If L is a finite dimensional semisimple Lie algebra and H a Cartan subalgebra, then any nonzero ideal I of L intersects H nontrivially.*

Proof. Indeed, applying the diagonalization lemma to the root space decomposition of L with respect to H, one finds $I \cap L_\alpha \neq 0$ for some root α of L. Either $L_\alpha = H$, or else L_α is one dimensional, hence $L\alpha \subset I$. But in the latter case the one dimensional subspace $[L_\alpha, L_{-\alpha}] \subset H$ is contained in I. □

It is thus reasonable to remove the ideals I with $I \cap H = 0$ from $\widetilde{L}(A)$ by setting

Definition. $L(A) := \widetilde{L}(A)/I(A)$.

Note that the kernel of the canonical map $\pi : \widetilde{L}(A) \to L(A)$ intersects H trivially. Thus we shall henceforth identify H with its image under π.

We finally have the easy

Proposition. *Let L be a Lie algebra generated by elements x_α, y_α ($\alpha \in \Delta$) and h ($h \in B$) which satisfy the relations 1.–4. (with X_α, Y_α replaced by x_α, y_α), and such that* $\dim \operatorname{span}_L B = \dim H$. *Assume that any nonzero ideal intersects* $\operatorname{span}_L B$ *nontrivially. Then $L \cong L(A)$.*

Proof. By assumption we can identify H with the subspace of L spanned by B. We have the obvious surjective homomorphism of Lie algebras $\phi : \widetilde{L} \to L$ mapping $X_\alpha \mapsto x_\alpha$, $Y_\alpha \mapsto y_\alpha$, and being the identity on H. The latter implies $\ker\phi \subset I(A)$. Conversely, $\phi(I(A))$ is an ideal of L having trivial intersection with H (if h is in this intersection, then $h + k \in I(A)$ for some $k \in \ker\phi \subset I(A)$, whence $h \in I(A)$). Hence $\phi(I(A)) = 0$ by assumption on L. □

As an immediate corollary we obtain

Theorem. *Let A be the Cartan matrix of a semisimple finite dimensional Lie algebra L, then $L \cong L(A)$.*

We can now give the definition of a Kac-Moody algebra:

Definition. A Kac-Moody algebra is a Lie algebra isomorphic to a Lie algebra of the form $L(A)$, where A is a *generalized Cartan matrix* (GCM), i.e. an $n \times n$-matrix matrix $A = (a_{i,j})$ which satisfies for all $1 \leq i, j \leq n$

1. $a_{i,i} = 2$.
2. For $i \neq j$ the element $a_{i,j}$ is an integer ≤ 0.
3. If $a_{i,j} = 0$ then $a_{j,i} = 0$.

Note that a Cartan matrix A of a semisimple Lie algebra L is a GCM. Thus finite dimensional semisimple Lie algebras are Kac-Moody algebras.

For the following we introduce some notations, which generalize the corresponding notions for finite dimensional semisimple Lie algebras. We call the elements of Δ *fundamental roots*. We shall call

$$Q = \sum_{\alpha \in \Delta} \mathbb{Z}\alpha$$

the *root lattice*. By the diagonalization lemma, $L = L(A)$, as quotient of the diagonalizable H-module $\widetilde{L}(A)$, is diagonalizable, i.e. we have a *root space decomposition*

$$L(A) = H \oplus \bigoplus_{\lambda \in Q} L_\lambda.$$

Note that L_λ is finite dimensional, and $L_\lambda = 0$ unless $\lambda \in Q$. If L_λ is nontrivial we call λ a *root*, and we shall use R for the set of all roots. In general the dimension of L_λ will be greater than 1 (see the section on affine Kac-Moody algebras).

Moreover, we also have the *triangular decomposition*

$$L(A) = N_- \oplus H \oplus N_+$$

(sum of vector spaces), where $N_\pm$ denotes the subalgebras generated by the x_α and y_α ($\alpha \in \Delta$), respectively. Here and in the following we use x_α and y_α for the images of X_α and Y_α under the canonical map $\widetilde{L}(A) \to L(A)$. We call x_α, y_α, h_α ($\alpha \in \Delta$) the Chevalley generators associated to α. Clearly, any L_λ is contained in N_+ or in N_-, accordingly as $\lambda > 0$ or $\lambda < 0$. It follows that L_λ for $\lambda > 0$ is the span of all elements of the form $[x_{\alpha_1}, [x_{\alpha_2}[\ldots x_{\alpha_n}]\!] \ldots]$ with $\alpha_j \in \Delta$ such that $\sum \alpha_j = \lambda$, and a similar statement holds for $\lambda < 0$. This shows in particular that L_α, for $\alpha \in \Delta$, is one dimensional, and that $s\alpha$ is a root if and only if $s = \pm 1$.

We put a partial ordering on H^* by writing

$$\lambda \geq \nu :\Longleftrightarrow \lambda - \nu = \sum_{\alpha \in \Delta} c_\alpha \, \alpha \text{ with all } c_\alpha \geq 0.$$

One can easily deduce from the corresponding decomposition of $\widetilde{L}(A)$ that

$$N_+ = \bigoplus_{\lambda > 0} L(A)_\lambda, \qquad N_- = \bigoplus_{\lambda < 0} L(A)_\lambda.$$

The algebra $L(A)$ is in general not semisimple. In fact, if $h \in H$ is annihilated by all $\alpha \in \Delta$, then h is in the center $\mathrm{Z}(L(A))$, as is clear from the defining relations. Thus the center of $L(A)$ has at least dimension equal to $n - \mathrm{rank}(A) = n - l$. In fact, it is easy to show that it is always contained in H, hence its dimension is exactly $n - l$ ([Kac], Proposition 1.6).

4.2 Remarks on a classification

Without proof we note that, for any two generalized Cartan matrices (GCM) A and A' one has $L(A) \cong L(A')$ if and only if A and A' are equal up to a permutation of indices ([Peterson-Kac]). It is an open problem to decide whether the same holds true or not for general matrices A and A'.

One calls a GCM A simple if it cannot be written as direct sum of two square matrices (with $n \geq 1$ lines each) after suitable permutation of indices.

Clearly any GCM is the unique sum of simple ones. Moreover, the sum of GCM is compatible with the sum of the associated Lie algebras.

The simple GCMs fall into 3 classes. The first class, noted FIN in [Kac] (and being identical to our FIN introduced in earlier sections) is the class of Cartan matrices of simple root systems. Recall that the GCM of class FIN can be characterized as those simple GCMs A which are symmetrizable and positive definite (i.e. AD is symmetric and positive definite for a suitable diagonal matrix D). Note that these GCMs have nonzero determinant, and that any minor of such a matrix is again a direct sum of Cartan matrices of simple root systems, as is immediate from the definition of FIN. By minor we mean a matrix of the form $(a_{i,j})_{i,j\in S}$, where $S \subseteq \{1, 2, \dots, n\}$, and where $A = (a_{i,j})_{1\le i,j\le n}$. The associated Kac-Moody algebras are the classical simple Lie algebras discussed in the preceding sections.

The second class is the class of simple GCMs A such that $\det(A) = 0$, but any proper minor of A is a direct sum of Cartan matrices of simple root systems. Note that such an A is necessarily of corank 1. The associated Lie algebras are called *affine Lie algebras*. This class is denoted AFF in [Kac]. It is further subdivided into 3 subclasses AFF1 – AFF3. Their Lie algebras can be described in terms of the classical simple ones. The GCM of class AFF can all "inductively" determined (cf. [Kac], pp. 51). They are listed in the table section.

The third class of simple GCMs comprises the rest, and is denoted by IND (indefinite type) in [Kac]. No natural construction is known for a single algebra associated to its the specimen. However, in the last years they have been studied in several papers by Gritsenko and Nikulin.

Lie algebras of type AFF and IND are always infinite dimensional ([Kac], Proposition 4.9).

4.3 Affine Kac-Moody algebras

In this section we give a natural description of the Lie algebras of class AFF1, also known as *nontwisted affine algebras*.

We start with a simple finite dimensional Lie algebra L. The *loop space* of $\mathcal{L}(L)$ is the set of all regular algebraic maps $\mathbb{C}^* \to L$. It carries a natural structure of Lie algebra via $[f, g](z) = [f(z), g(z)]_L$ $(z \in \mathbb{C}^*)$. It is convenient to use a slightly different description of the loop space, namely

$$\mathcal{L}(L) := \mathbb{C}[t, t^{-1}] \otimes L.$$

Here a $P \otimes x$ may be viewed as the map $z \mapsto P(z)\, z$. Thus the bracket is given by

$$[P \otimes x, Q \otimes y]_{\mathcal{L}(L)} = PQ\,[x, y].$$

In the next step we consider a certain central extension of $\mathcal{L}(L)$. Let C be a symbol and set

$$\widetilde{\mathcal{L}}(L) := \mathcal{L}(L) \oplus \mathbb{C} \cdot C,$$

$$[f, C] := 0, \quad [f, g] := [f, g]_{\mathcal{L}(L)} + \operatorname{Res} B(\frac{d}{dt} f, g) \cdot C \qquad (f, g \in \mathcal{L}(L)).$$

Here B is the Killing form on L, extended linearly to $\mathcal{L}(L)$ with values in the ring of Laurent polynomials $\mathbb{C}[t, t^{-1}]$. The residue is as usual defined as the coefficient of t^{-1}. Thus, for integers m and n and $x, y \in L$ we have

$$[t^m \otimes x, t^n \otimes y] = [x\,y]_{\mathcal{L}(L)} + m\,\delta_{m+n,0}\,B(x, y) \cdot C.$$

It is not evident that $\widetilde{\mathcal{L}}(L)$ is really a Lie algebra. Of course, this can be verified by a straightforward, though lengthy computation. Essentially, one has to check that the term in front of C is bilinear, antisymmetric and satisfies the so-called *cocycle property*.

Finally, we adjoin the derivation $d = t\frac{d}{dt}$. Thus we set

$$\widehat{L} := \widetilde{\mathcal{L}}(L) \oplus \mathbb{C} \cdot d \qquad [d, f] = d\,f, \quad [f, d] = -d\,f \quad (f \in \widetilde{\mathcal{L}}(L)).$$

Here d acts on elements of the loop algebra in the obvious manner, i.e. by $d(P \otimes x) = (t\frac{d}{dt}P) \otimes x$, and it acts trivially on C. It is easy to show that "adjoining a derivation of a Lie algebra to the Lie algebra" always yields again a Lie algebra.

Theorem. *$\widehat{L}$ is an affine Lie algebra.*

Proof. For this let H be a Cartan subalgebra of the simple Lie algebra L, and let R its root system with respect to H. We identify L with a subalgebra of $\widehat{L}$ by $x \mapsto 1 \otimes x$. Let

$$\widehat{H} := H + \mathbb{C} \cdot C + \mathbb{C}d$$

Clearly, $\widehat{H}$ is an abelian subalgebra of $\widehat{L}$. We view H^* as subspace of $\widehat{H}^*$ by extending a $\lambda \in H^*$ to $\widehat{H}^*$ by $\lambda(C) = \lambda(d) = 0$. Finally, we let $\delta \in \widehat{H}^*$ be the element such that $\delta(x) = 0$ for $x \in H + \mathbb{C} \cdot C$ and $\delta(d) = 1$.

With respect to the adjoint action of $\widehat{H}$ the Lie algebra $\widetilde{L}$ is then diagonalizable, i.e. we then have the following root space decomposition

$$\widehat{R} = \widehat{H} \oplus \bigoplus_{n \in \mathbb{Z},\, \alpha \in R} \big(\mathbb{C} \cdot t^n \otimes L_\alpha\big) \oplus \bigoplus_{n \in \mathbb{Z},\, n \neq 0} \big(\mathbb{C} \cdot t^n \otimes H\big)$$

with roots $n\delta + \alpha$ and $n\delta$ ($n \neq 0$), respectively.

Let Δ be a root basis of the root system R of L, and let $x_\alpha, y_\alpha, h_\alpha$ ($\alpha \in \Delta$) be the corresponding Chevalley generators of L. Moreover, let θ be the highest root of Δ, i.e. let θ be maximal with respect to the partial ordering on R induced by the choice of Δ. It can be shown that θ is unique — of course, it depends on the choice of root basis Δ.

Then $\widehat{L}$ is generated by $\widehat{H}$, the elements $x_\alpha, y_\alpha, h_\alpha$ ($\alpha \in \Delta$), and the elements

$$x_{\delta-\theta} := t \otimes y_\theta, \qquad y_{\delta-\theta} := t^{-1} \otimes x_\theta,$$
$$h_{\delta-\theta} := [t \otimes y_\theta, t^{-1} \otimes x_\theta] = -h_\theta + \frac{2}{(\theta,\theta)} \cdot C.$$

Indeed, let M be the Lie algebra generated by these elements. Clearly $L \subseteq M$. But also $[t \otimes y_\theta, L] = t \otimes [y_\theta, L] = t \otimes L \subset M$ (here we used that L is simple, and hence the nonzero ideal $[y_\theta, L]$ is equal to L), and then $[t \otimes y_\theta, t \otimes L] = t^2 \otimes L \subset M$, and so forth. Similarly, taking $t \otimes x_\theta$ instead of $t \otimes y_\theta$ we find $t^{-n} \otimes L \subset M$ for all $n \geq 0$. Since $H \subset M$ we conclude $M = \widehat{L}$.

Let $A = \big(\alpha(h_\beta)\big)$, where α, β run through

$$\widehat{\Delta} := \{\delta - \theta\} \cup \Delta$$

(for some ordering of $\widehat{\Delta}$). The generators of $\widetilde{L}(A)$ satisfy the same relations as the corresponding ones of $\widetilde{L}$. For showing $[x_{\delta-\theta}, y_\alpha] = t \otimes [x_\theta, y_\alpha] = 0$ we use that θ is maximal. Indeed, $[x_\theta, y_\alpha] \subset L_{\theta+\alpha}$, and the latter space is 0 since $\theta + \alpha$ is not a root (since θ is the highest root).

Hence we have the obvious surjective homomorphism of Lie algebras $\widetilde{L}(A) \mapsto \widehat{L}$. As explained in the first section on Kac-Moody algebras it suffices, for concluding that this map is actually an isomorphism, to show that any ideal I of $\widehat{L}$ intersecting $\widehat{H}$ trivially is 0. Indeed, by the diagonalization lemma such an ideal I, if it were nonzero, would have nonzero intersection with $(\widehat{L})_\sigma$ for some root $\sigma = n\delta + \alpha$ (α a root or 0), hence would contain an element $t^n \otimes x$ for some $0 \neq x \in L_\alpha$. But then $I \ni [t^n \otimes x, t^{-n} \otimes y] = [x,y] + nB(x,y)C$ for all $y \in L$. Choosing $y \in L_{-\alpha}$ and such that $B(x,y) \neq 0$ gives $[x,y] + nB(x,y)C \in \widehat{H} \cap I = 0$. Hence $[x,y] = 0$ and $n = 0$. But $n = 0$ implies $\alpha \neq 0$, hence $[x,y]$ is a nonzero multiple of h_α, a contradiction.

It remains to show that A is a GCM of class AFF. We note

$$(\delta-\theta)(h_{\delta-\theta}) = 2, \quad (\delta-\theta)(h_\alpha) = -\theta(h_\alpha), \quad \alpha(h_{\delta-\theta}) = -\alpha(h_\theta) \quad (\alpha \in \Delta).$$

Hence, if we let D be the diagonal matrix with $\frac{1}{2}(\theta,\theta)$ and $\frac{1}{2}(\alpha,\alpha)$ ($\alpha \in \Delta$) on the diagonal (the scalar product is the natural one on the root system of

L), then

$$AD = \big((\alpha,\beta)\big)_{\alpha,\beta\in\Delta_0} \qquad \Delta_0 = \Delta \cup \{-\theta\}.$$

Now $(\alpha,-\theta) \le 0$ for any fundamental α and $(\alpha,-\theta) < 0$ for at least one $\alpha \in \Delta$ ([Humphreys], §10.4). Hence A is a GCM, and it is simple. Moreover AD is singular (as Gram matrix of $l+1$ vectors in a lattice of rank l), and any proper minor of AD is positive definite (here one has to use that, for any fundamental root α of L the roots θ and $\Delta \setminus \{\alpha\}$ form a basis of $\mathbb{R} \otimes \mathbb{Z}[\mathbb{R}]$, since $\theta = \sum_\alpha a_\alpha \alpha$ with all $a_\alpha > 0$ [Humphreys], §10.4.). Thus A is of class AFF. □

One can show that A is actually of class AFF1. Moreover, the other affine Kac-Moody algebras can be obtained by a similar construction ([Kac], Chapter 8). (Roughly spoken, one has to consider equivariant maps with respect to the action of certain cyclic groups instead of the full loop algebra).

4.4 Highest weight representations

Let $L = L(A)$ with a general complex matrix A. We keep the notations of the first section on Kac-Moody algebras. If V is a L-module we set

$$V_\lambda = \{v \in V \,|\, \forall h \in H : hv = \lambda(h)v\} \qquad (\lambda \in H^*).$$

We call λ a weight of V if $V_\lambda \neq 0$.

Definition. A L-module V is called highest weight module with highest weight $\Lambda \in H^*$ (or simply L-module with highest weight Λ), if there exists a nonzero $v_\Lambda \in V$ such that

$$U(L)v_\Lambda = V, \quad N_+ v_\Lambda = 0, \quad \forall h \in H : hv_\Lambda = \Lambda(h)v_\Lambda.$$

Clearly $V = U(N_-)v_\Lambda$. Since $U(N_-)$ is generated by the y_α with $\alpha \in \Delta$, we find that any element of V is a linear combination of elements of the form $y_{\alpha_1} y_{\alpha_2} \cdots y_{\alpha_n} v_\Lambda$ ($\alpha_i \in \Delta$). But these are elements of weight $\Lambda - (\alpha_1 + \cdots + \alpha_n)$. Hence we deduce that V is H-diagonalizable with finite dimensional weight spaces:

$$V = \bigoplus_{\lambda \le \Lambda} V_\Lambda \quad \forall \lambda \in H : \dim V_\lambda < \infty.$$

Note also that $V_\Lambda = \mathbb{C} \cdot v_\Lambda$. We call v_Λ a *highest weight vector of V*.

Definition. A Verma module $M(\Lambda)$ is a L-module with highest weight Λ such that for every L-module V with highest weight Λ we have a surjective homomorphism of L-modules $M(\Lambda) \to V$.

Proposition. *For each $\Lambda \in H^*$ there exists one and, up to isomorphism of L-modules, only one Verma module $M(\Lambda)$ with highest weight Λ.*

Proof. If we have two Verma modules M and M', then we have a map $\phi : M \to M'$ mapping M_λ surjectively onto M'_λ for all $\lambda \leq \Lambda$. In particular, $\dim M_\lambda \geq \dim M'_\lambda$. By symmetry we have actually equality, and hence ϕ is an isomorphism.

Set $M(\Lambda) := U(L)/J$ where J is the left ideal generated by N_+ and $h - \Lambda(h)$ ($h \in H$). This is clearly a Verma module with highest weight vector v_Λ =image of $1 \in U(L)$. □

The sum of any proper submodules of $M(\Lambda)$ is again a proper submodule. Indeed, any submodule is H-diagonalizable, and is proper if and only if it does not contain $M(\Lambda)_\Lambda$. Thus there exist a unique maximal proper submodule $M'(\Lambda)$, namely the sum of all proper submodules. But then there exist also a unique simple module with highest weight Λ, namely

Definition. $L(\Lambda) := M(\Lambda)/M'(\Lambda)$.

Proposition. *Any module with highest weight Λ contains $L(\Lambda)$ as quotient.*

Proof. Let V be a module with highest weight Λ, and let V' be a maximal submodule not containing the highest weight vector v_Λ (whose existence is easily be shown by using that every submodule is H-diagonalizable). Then V/V' is simple with highest weight Λ, hence isomorphic to $L(\Lambda)$. □

We now associate characters to highest weight modules. For this let $\mathcal{E}$ be the vector space of all formal sums

$$\sum_{\lambda \in H} c_\lambda \, \mathrm{e}(\lambda)$$

such that $c_\lambda = 0$ for λ outside a finite union of sets of the form

$$H^*_{\leq \mu} := \{\lambda \in H^* \mid \lambda \leq \mu\}.$$

We multiply two such series on using the rule $\mathrm{e}(\lambda)\mathrm{e}(\mu) = \mathrm{e}(\lambda + \mu)$. Thus $\mathcal{E}$ is an algebra over $\mathbb{C}$.

Definition. For a module V with highest weight Λ we set

$$\mathrm{ch}_V := \sum_{\lambda \in H} \dim V_\lambda \, \mathrm{e}(\lambda).$$

This clearly defines an element of $\mathcal{E}$.

As a first application we calculate the character of the Verma module $M(\Lambda)$. Let R_+ be the set of positive roots of L, and, for $\alpha \in R_+$ let $y(\alpha)_1, \ldots, y(\alpha)_{d(\alpha)}$ be a basis of $L_{-\alpha}$. Let J be the left ideal as in the construction of the Verma modules, i.e. the ideal generated by N_+ and $h - \Lambda(h)$ $(h \in H)$. Let $\alpha_1, \alpha_2, \ldots$ be an enumeration of R_+. Using the PBW theorem we then see that the images of

$$y(\alpha_1)_1^{n_{1,1}} \cdots y(\alpha_1)_{d(\alpha_1)}^{n_{1,d(\alpha_1)}} y(\alpha_2)_1^{n_{2,1}} \ldots y(\alpha_2)_{d(\alpha_2)}^{n_{2,d(\alpha_2)}} \cdots$$

under the canonical map $U(L) \to U(L)/J$ form a basis of $M(L)$. Moreover, every such element has weight

$$\Lambda - (n_{1,1} + \cdots + n_{1,d(\alpha_1)})\alpha_1 - (n_{2,1} + \cdots + n_{2,d(\alpha_2)})\alpha_2 - \cdots .$$

This shows

Proposition. *The Verma module $M(\Lambda)$ has character*

$$\mathrm{ch}_{M(\Lambda)} = e(\Lambda) \prod_{\alpha > 0} (1 - e(-\alpha))^{-\dim L_\alpha}.$$

4.5 Character Formula and consequences

The goal of this section is to obtain a formula for the character of $L(\Lambda)$, the unique simple module with highest weight Λ. For this we have to assume that $L = L(A)$ is a symmetrizable Kac-Moody algebra. Here symmetrizable means that AD is symmetric for some regular diagonal matrix D. The classical simple Lie algebras are obvious symmetrizable. To state the character formula we need to explain two more ingredients of the theory.

We let W denote the Weyl group of L. Generalizing the classical case this is by definition the subgroup of $\mathrm{GL}(H^*)$ generated by the reflections

$$\sigma_\alpha(\lambda) = \lambda - \lambda(h_\alpha)\alpha,$$

where α runs through the set of fundamental roots Δ. In contrast to the case of the finite dimensional semisimple Lie algebras this is in general an infinite group.

We assume that A is a GCM (so that, in particular, the root lattice $Q = \sum_{\alpha \in \Delta} \mathbb{Z}\alpha$ is invariant under W). The sets

$$C := \{\lambda \in H_{\mathbb{R}}^* \mid \lambda(h_\alpha) \geq 0\}, \qquad X := \bigcup_{w \in W} w(C)$$

are called the *fundamental chamber* and the *Tits cone* respectively. Here $H_{\mathbb{R}} := \mathbb{R}^{2n-l}$ ($l = \operatorname{rank}(A)$). Recall that the h_α lie in $H_{\mathbb{R}}$, since A, as GCM, is a real matrix. Without proof we cite

Proposition. *The set C is a fundamental domain for the action of W on X (i.e. it contains exactly one point from each orbit $W \cdot \lambda$ with $\lambda \in X$.) Moreover, for each $\lambda \in C$, one has $w(\lambda) - \lambda = \sum_{\alpha \in \Delta} c(\alpha)\, \alpha$ with all $c(\alpha) \leq 0$.*

Proof. [Kac], Proposition 3.12. □

Fix an element $\rho \in H^*$ such that

$$\rho(h_\alpha) = 1 \qquad (\alpha \in \Delta).$$

If A is degenerate, then ρ is not unique. We simply pick one. If A is of type FIN then $\rho = \frac{1}{2}\sum_{\alpha>0} \alpha$, the sum being over all positive roots (not necessarily fundamental). In fact, each σ_β ($\beta \in \Delta$) permutes the set $R_+ \setminus \{\beta\}$ of all positive roots with Δ omitted (see step 7 of the proof of the character formula). Hence $\sigma_\beta(\rho) = \rho - \beta$, and, since by definition of σ_β we have $\sigma_\beta(\rho) = \rho - \rho(h_\beta)\beta$, we deduce $\rho(h_\beta) = 1$.

Theorem. *(Character Formula) Let $\Lambda \in H^*$ such that $\Lambda(h_\alpha) \in \mathbb{Z}_{\geq 0}$ for all $\alpha \in \Delta$. Then the character of the simple highest weight module $L(\Lambda)$ is given by the formula*

$$\operatorname{ch}_{L(\Lambda)} = \frac{\displaystyle\sum_{w \in W} \det(w)\, e(w(\Lambda + \rho))}{\displaystyle e(\rho) \prod_{\alpha>0} (1 - e(-\alpha))^{\dim L_\alpha}}$$

Example. Let $L = L((2)) = \mathrm{sl}_2$. Write $q = \mathrm{e}(-\frac{1}{2}\alpha)$, α denoting the single element of Δ. Clearly $W = \{\pm 1\}$. For a $\Lambda = N \cdot \frac{\alpha}{2}$ with integral $N \geq 0$ the theorem states

$$\operatorname{ch}_{L(N \cdot \frac{\alpha}{2})} = \frac{q^N - q^{-(N+2)}}{1 - q^2}.$$

This is the formula which we found in the section about the representations of sl_2 ($L(N \cdot \frac{\alpha}{2})$ was there denoted by $V(N)$).

If we apply the theorem to $\Lambda = 0$, then using $\dim L(0) = 1$ we find

Theorem. *(Weyl's Denominator Identity)*

$$\prod_{\alpha>0} (1 - e(-\alpha))^{\dim L_\alpha} = \sum_{w \in W} \det(w)\, e(w(\rho) - \rho).$$

Replacing the denominator in the character formula for $L(\Lambda)$ by the right hand side of the denominator identity, we find

Theorem. *(Weyl's Character Formula)*

$$\mathrm{ch}_{L(\Lambda)} = \frac{\sum\limits_{w\in W} \det(w)\, e(w(\Lambda+\rho))}{\sum\limits_{w\in W} \det(w)\, e(w(\rho))}$$

The Weyl character formula, for finite dimensional semisimple Lie algebras, is classical. The denominator formula, specialized to affine Kac-Moody algebras, gives the so-called Macdonald identities. The simplest case is the affine algebra associated to

$$\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix},$$

i.e. the affine algebra $\widehat{\mathrm{sl}_2}$. Here the denominator identity becomes the classical Jacobi triple product identity ([Kac], Exercise 10.9).

4.6 Proof of the Character Formula

This section is devoted to an outline of the proof of the character formula, as it can be found in [Kac]. However, we tried here to put this proof into a compact and self-contained form. We shall use for the the first time that we really deal with Kac-Moody algebras, and not simply with arbitrary matrix algebras $L(A)$. However, many of the following considerations are still valid for arbitrary complex A. We shall always point out when a conclusion really needs that A is a GCM.

It is obvious that we can write the character of $L(\Lambda)$ in the form

$$\mathrm{ch}_{L(\Lambda)} = \frac{\sum\limits_{\lambda\le\Lambda} c(\lambda)\,\mathrm{e}(-\lambda)}{\prod\limits_{\alpha>0}(1-\mathrm{e}(-\alpha))^{\dim L_\alpha}}$$

with suitable integers $c(\lambda)$. However, it is quite difficult to calculate the desired formula for the numerator. This calculation will be carried out in 8 steps.

Step 1: For every $V = M(\Lambda)$ $(\Lambda \in H^*)$ one has

$$\mathrm{ch}_V = \sum_{\lambda\in H^*} m(\lambda)\,\mathrm{ch}_{L(\lambda)}$$

with suitable nonnegative integers $m(\lambda)$.

For this let $\mathcal{O}$ be the category of modules V over $L = L(A)$ which are H-diagonalizable with finite dimensional weight spaces V_λ, and such that $V_\lambda = 0$ for λ outside a finite union of sets of the form $H^*_{\leq\mu}$. Clearly, all highest weight modules belong to this category. Moreover, for each such module we can define its character by exactly the same formula as for the highest weight modules. We shall show that the claimed identity is in fact true for any V in $\mathcal{O}$.

If $0 \to V_1 \to V_2 \to V_3 \to 0$ is an exact sequence of modules over $L = L(A)$, then by the very definition of characters we have $\mathrm{ch}_{V_2} = \mathrm{ch}_{V_1} + \mathrm{ch}_{V_3}$. From this it is then immediate that $\mathrm{ch}_V = \sum_i \mathrm{ch}_{V_i}$ for any filtration by L-modules $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V$ of a given module V.

Let $\mu \in H^*$ and V from $\mathcal{O}$. We shall show in a moment that there exists then a filtration of V by L-modules (V_i) such that either $V_i/V_{i-1} \cong L(\lambda)$ for some $\lambda \geq \mu$, or else V_i/V_{i-1} contains no vectors with a weight $\geq \mu$. For a given $\lambda \geq \mu$ denote by $m_\mu(\lambda)$ the number of i with $V_i/V_{i-1} \cong L(\lambda)$. If we take $m(\lambda) = m_\mu(\lambda)$ for any μ with $\lambda \geq \mu$, the claimed identity holds then above μ, i.e. the coefficients of all $\mathrm{e}(\lambda)$ with $\lambda \geq \mu$ on both sides of the identity coincide.

But $m_\mu(\lambda)$ does not depend on the choice of μ (or the choice of the filtration) Indeed, if (W_j) is a filtration belonging to some $\mu' \leq \lambda$, then we consider the filtrations F and F' given by

$$V_{i-1} + (V_i \cap W_j) \quad \text{and} \quad W_{j-1} + (W_j \cap V_i) \qquad (i, j),$$

respectively. Since $L(\lambda)$ is simple, the number of factors isomorphic to $L(\lambda)$ in F and F' equal $m_\mu(\Lambda)$ and $m_{\mu'}(\lambda)$, respectively. But for all i, j one has

$$\frac{V_{i-1} + (V_i \cap W_j)}{V_{i-1} + (V_i \cap W_{j-1})} \cong \frac{W_{j-1} + (W_j \cap V_i)}{W_{j-1} + (W_j \cap V_{i-1})}$$

by the Zassenhaus Butterfly Lemma.

Thus, if, for all λ, we finally set $m(\lambda) := m_\lambda(\lambda)$, then the claimed identity holds above μ for any $\mu \in H^*$, whence is a true identity.

Note that $m(\lambda) \neq 0$ only if $L(\lambda)$ occurs as quotient of some submodule of V.

It remains to show the existence of the filtrations (V_i) for a given V and μ. This is done by induction on

$$n(V) := \sum_{\lambda \geq \mu} \dim V_\lambda.$$

If $n = 0$ we can take $0, V$ as filtration. Otherwise choose a maximal weight $\lambda \geq \mu$ of V, a nonzero $v \in V_\lambda$, set $U = U(L)v$ and let $\overline{U}$ a maximal proper submodule of U (which exists by the diagonalization lemma). Then $0, \overline{U}, U, V$ is a filtration such that $U/\overline{U} \cong L(\lambda)$. But $n(\overline{U}), n(V/U) < n(V)$, and applying the induction hypothesis to V/U and $\overline{U}$ yields than the desired filtration.

Step 2: Since A is symmetrizable, there is a regular diagonal matrix D such that $AD = (b_{\alpha,\beta})_{\alpha,\beta}$ is symmetric. Thus one can define a scalar product on the root lattice by setting $(\alpha, \beta) := b_{\alpha,\beta}$ for simple roots α, β. In general this is degenerate. Now, if A is a GCM, then the Weyl group leaves the root lattice Q invariant, and also the scalar product (α, β) (easy exercise). Moreover, since for a GCM, all elements $a_{i,j}$ $(i \neq j)$ of A are ≤ 0, we can and will choose D such that $(\alpha, \alpha) > 0$ for all fundamental roots α ([Kac], §2.3 or easy exercise).

Without proof we shall use the following: For a symmetrizable Kac-Moody algebra there exists always a universal operator, the *generalized Casimir operator* Ω (depending on the choice of D), which commutes with the action of L on modules in $\mathcal{O}$ ([Kac], Theorem 2.6).

If L is a finite-dimensional semisimple Lie algebra then we take for D the matrix with $B(t_\alpha, t_\alpha)/2$ as diagonal elements so that the above scalar product is the usual one which we considered in the the section on semisimple algebras (notations as loc.cit.). Set

$$\Omega = \sum_{i=1}^{d} x_i x^i \in U(L),$$

where x_i is basis of L, and where x^i is the dual basis with respect to the Killing form. It is easily checked that Ω is independent of the choice of basis x_i, and that it commutes with L on any module V. If now V is a highest weight module with highest weight Λ and highest weight vector $v = v_\Lambda$, then

$$\Omega \cdot v = (\Lambda + \rho, \rho) v.$$

Indeed, take as basis x_i the elements h_α and elements $z_\alpha \in L_\alpha$. Recall $z^\alpha \in L_{-\alpha}$, and $[z_\alpha, z^\alpha] = t_\alpha$. Then, if $\alpha < 0$, we have $z_\alpha z^\alpha \cdot v = 0$, and if $\alpha > 0$ we have

$$z_\alpha z^\alpha \cdot v = [z_\alpha, z^\alpha] \cdot v + z^\alpha z_\alpha \cdot v = t_\alpha \cdot v = (\Lambda, \alpha) v.$$

Thus,

$$\Omega \cdot v = \big(\sum_\alpha \Lambda(h_\alpha)\, \Lambda(h^\alpha) + \sum_\alpha (\Lambda, \alpha) \big) v = (\Lambda + 2\rho, \Lambda)\, v.$$

As it turns out this formula holds true in general ([Kac], Corollary 2.6).

Since Ω commutes with L and since $V = U(L)\cdot v_\Lambda$, we find that Ω acts on a highest weight module with highest weight Λ as multiplication by $(\Lambda, \Lambda+2\rho)$.

But than it acts also as multiplication by $(\Lambda, \Lambda+2\rho)$ on any submodule, and on any quotient of such. Accordingly we see that

$$M(\Lambda) = \sum_{\lambda\in B} m(\lambda)\mathrm{ch}_{L(\lambda)},$$

where the summation is over

$$B = \{\lambda \leq \Lambda \mid (\lambda, \lambda+2\rho) = (\Lambda, \Lambda+2\rho)\}.$$

Step 3: From the last formula in step 2 we deduce that

$$\mathrm{ch}_{L(\Lambda)} = \sum_{\lambda\in B} c(\lambda)\,\mathrm{ch}_{M(\lambda)},$$

for suitable integers $c(\lambda)$. Note that the summation is again over B.

Indeed, let $\lambda_1, \lambda_2, \ldots$ be an enumeration of B such that $\lambda_i \geq \lambda_j \Longrightarrow i \leq j$. We then have

$$\mathrm{ch}_{M(\lambda_i)} = \sum_{\lambda_j\in B,\ \lambda_j\leq\lambda_i} c_{ij}\,\mathrm{ch}_{L(\lambda_j)}$$

with nonnegative integers $c_{i,j}$. This (infinite) system of linear equations can be solved since its matrix is diagonal, and it can be solved in integers since all $c_{j,j} = 1$. We obtain thus the claimed formula.

Step 4: Inserting the formula for the character of $M(\lambda)$ which we calculated in the previous chapter, into the formula of step 3, we obtain

$$\mathrm{ch}_{L(\Lambda)} = \frac{\displaystyle\sum_{\lambda\in B} c(\lambda)\,\mathrm{e}(\lambda)}{\displaystyle\prod_{\alpha>0}\big(1-\mathrm{e}(-\alpha)\big)^{\dim L_\alpha}}.$$

This formula is so far true for any $L(A)$ with diagonalizable A. Note also that the only nontrivial ingredient in this formula is the restriction of the summation to the set B, which in turn is due to the existence of the universal Casimir operator Ω. We now use the Weyl group to analyze the numerator N on the right hand side.

Step 5: We now use the assumption $\alpha(h_\alpha) = 2$ for fundamental roots α. Then $S_\alpha := \mathbb{C}\cdot x_\alpha + \mathbb{C}\cdots y_\alpha + \mathbb{C}\cdot h_\alpha$ is isomorphic to sl_2. We shall show in a moment that $L(\Lambda)$, under the assumption $\Lambda(h_\alpha) \in \mathbb{Z}_{\geq 0}$ and the assumption that A is a GCM, decomposes into a direct sum of finite dimensional irreducible S_α modules, hence into a direct sum of $V(N)$ with with integral $N \geq 0$. This implies that $\mathrm{ch}_{L(\Lambda)}$ is invariant under the Weyl group.

More generally, if V is any L-module which decomposes as direct sum of $V(N)$ ($N \geq 0$) under S_α, and such that the weight spaces V_λ with respect to H are finite dimensional, then $\dim V_\lambda = \dim V_{\sigma_\alpha(\lambda)}$ for all λ.

For showing this note that the nontrivial V_λ are exactly the weight spaces of h_α with weight $\lambda(h_\alpha)$. Let λ be an H-weight of V. Then V_β with $\beta(h_\alpha) = -\lambda(h_\alpha)$, the weight space of h_α with weight $-\lambda(h_\alpha)$, has the same dimension as V_λ. But $\beta = \lambda - \lambda(h_\alpha)\alpha = \sigma_\alpha(\lambda)$ with the reflection σ_α of the Weyl group.

For $L(\Lambda)$ this may be reformulated as $\sigma_\alpha \mathrm{ch}_V = \sigma_\alpha$. Here the action of the Weyl group on (possibly infinite) formal linear combinations of the $\mathrm{e}(\lambda)$ is defined by $\sigma \sum a(\lambda)\,\mathrm{e}(\lambda) = \sum a(\lambda)\,\mathrm{e}(\sigma\lambda)$.

Since the Weyl group W is by generated by the σ_α with α running through the fundamental roots, $\mathrm{ch}_{L(\Lambda)}$ is invariant under W.

Step 6: Showing that an sl_2-module is isomorphic to a direct sum of finite dimensional $V(N)$, is in fact equivalent to showing that $x := x_\alpha$ and $y := y_\alpha$ are locally nilpotent on $L(\Lambda)$. One says, that z is locally nilpotent on an S_α-module V, if for each $v \in V$ there exist a nonnegative integer N such that $z^N \cdot v = 0$.

Indeed, x, y being locally nilpotent on V, implies that each $v \in V$ is contained in a finite dimensional S_α-submodule of V (and hence V is the sum of modules which are all isomorphic to $V(N)$ with suitable integers N, and a minimal such representation is a direct sum). Namely, let $v \in V$ and consider

$$U := \sum_{i,j \geq 0} y^i x^j \cdot v$$

Clearly, U is finite dimensional. Moreover, the formula

$$[x, y^k] = -k(k-1)y^{k-1} + ky^{k-1}h, \qquad (\text{in } U(\mathrm{sl}_2)),$$

shows that U is invariant under sl_2.

For showing that x_α and y_α are locally nilpotent on $L(\Lambda)$, consider first $v = v_\Lambda$. Then v has h_α-weight $N := \Lambda(h_\alpha)$. By assumption this is an integer $N \geq 0$. But this implies $x_\alpha y_\alpha^{N+1} \cdot v = 0$ (recall from the section on representations of sl_2: for all v of weight N such that $x \cdot v = 0$, and for all k, one has $xy^k \cdot v = k(N+1-k) \cdot y^{k-1}$). Since $[x_\beta, y_\alpha] = 0$, for $\beta \neq \alpha$, we have $x_\beta y_\alpha^{N+1} \cdot v = 0$. Thus $U(L)y_\alpha^{N+1} \cdot v$ is a proper submodule of V. Since $V = L(\Lambda)$ is irreducible, we deduce $y_\alpha^{N+1} \cdot v = 0$ (and hence, as corollary, $S_\alpha \cdot v \cong V(N)$).

Now, $L(\lambda) = U(L) \cdot v$. We show in a moment that x_α and y_α are locally nilpotent on L. That they are locally nilpotent on $L(\Lambda)$ follows then on using the formula

$$z^N a = \sum_{k=0}^{N} \binom{N}{k} (\mathrm{ad}x)^k a \cdot x^{N-k}, \qquad (z \in L,\ a \in U(L)),$$

and applying it to v.

For showing that x_α and y_α are locally nilpotent on L (via the adjoint representation) we now have to use also that $\beta(h_\alpha)$ is an integer ≤ 0 for any fundamental root β.

Namely, this implies for any pair of fundamental roots $\alpha \neq \beta$

$$(\mathrm{ad}x_\alpha)^{1-\beta(h_\alpha)}\, x_\beta = 0,$$

and similarly with x_α, x_β replaced by y_α, y_β (we saw this relation for the finite dimensional algebras already in the section on Serre's theorem). We leave this as an exercise (or else [Kac], §3.3).

That x_α is locally nilpotent on all of L now follows from the fact that suitable powers of x_α kill the generators of L. On applying Leibniz formula to the derivation ad$x$, one easily deduces that every element in L is killed by some power of ad$x$.

Step 7: Denote by D the denominator of the formula for $\mathrm{ch}_{L(\Lambda)}$ obtained in step 4. We show that it is invariant under W.

Let R_+ be the set of all positive roots of L, (i.e. the set of all $\beta > 0$ with nontrivial L_β). We remark first of all that, for any fundamental root α, the set $R_+ \setminus \{\alpha\}$ is invariant under σ_α. Indeed, $\beta \in R_+$ implies that $\sigma_\alpha(\beta)$ is a root, hence either negative or else positive. If $\sigma_\alpha(\beta) < 0$, then $\sigma_\alpha(\beta) = \beta - k\alpha$ for an integer k ($= \beta(h_\alpha)$) together with $\beta > 0$ imply that β is a multiple of α. But the only positive multiple of a fundamental root α which is a root, is α itself.

Using that L_β and $L_{\sigma_\alpha(\beta)}$ have same dimension, we now calculate

$$\sigma_\alpha D = \big[1 - \mathrm{e}\big(\sigma_\alpha(-\alpha)\big)\big]\, \sigma_\alpha \left(\frac{D}{1 - \mathrm{e}(-\alpha)}\right) = \frac{1 - \mathrm{e}(\alpha)}{1 - \mathrm{e}(-\alpha)}\, D = -\mathrm{e}(\alpha)\, D.$$

Step 8: The numerator N of the formula in step 4 must then have the same behavior under the Weyl group as the denominator, i.e. $\sigma_\alpha N = -\mathrm{e}(\alpha) N$. But $\sigma_\alpha(\rho) = \rho - \rho(\alpha)\alpha = \rho - \alpha$. Thus, if we multiply the numerator by $\mathrm{e}(\rho)$, i.e. if we set

$$M := \sum_{\mu \leq \Pi,\ |\mu|^2 = |\Pi|^2} d(\mu)\, \mathrm{e}(\mu), \quad \Pi = \Lambda + \rho,\ d(\mu) = c(\mu - \rho),$$

then $w(M) = \det(w) M$ for all $w \in W$.

Let $d(\mu) \neq 0$, and let $w \in W$. Then $\det(w)\, d(w\mu) = d(\mu)$. In particular, $d(w\mu) \neq 0$, hence $\nu := w(\mu) \leq \Pi$. Choose w such that $\mathrm{ht}(\nu)$ is maximal. Here we use $\mathrm{ht}(\nu) = \sum_\alpha a_\alpha$ if $\nu = \sum_\alpha a_\alpha \alpha$, where α runs through the fundamental roots. Note that such a minimizing w exists, since, for any $\lambda_0 \leq \Pi$ the set of λ such that $\lambda_0 \leq \lambda \leq \Pi$ is finite.

Now, setting $\Pi - \nu = \sum_\alpha b_\alpha \alpha$, we find

$$0 = |\Pi|^2 - |\nu|^2 = (\Pi + \nu, \Pi - \nu) = \sum_\alpha (\Pi + \nu)(h_\alpha)\, \frac{(\alpha, \alpha)}{2}\, b_\alpha.$$

But $b_\alpha \geq 0$, $(\alpha, \alpha) > 0$ (since A is GCM; cf. the remark in step 2), and $(\Pi + \nu)(h_\alpha) > 0$. Hence all $b_\alpha = 0$, whence $\Pi = \nu$.

Thus, we have shown that $d(\mu) \neq 0$ implies $\mu = w(\Pi)$ for some w, and then $d(\mu) = \det(w) c(\Pi)$. Now clearly $c(\Pi) = 1$. Hence, for proving the character formula, it remains to show that the stabilizer of Π in W is trivial. But this follows immediately on writing $\Pi = \Pi_0 + \sqrt{-1}\Pi_1$ with $\Pi_0, \Pi_1 \in H^*_{\mathbb{R}}$, on noticing that Π_0 is an element of the fundamental chamber C, and on finally applying the proposition of the preceding section.

Appendix A

Dynkin Diagrams

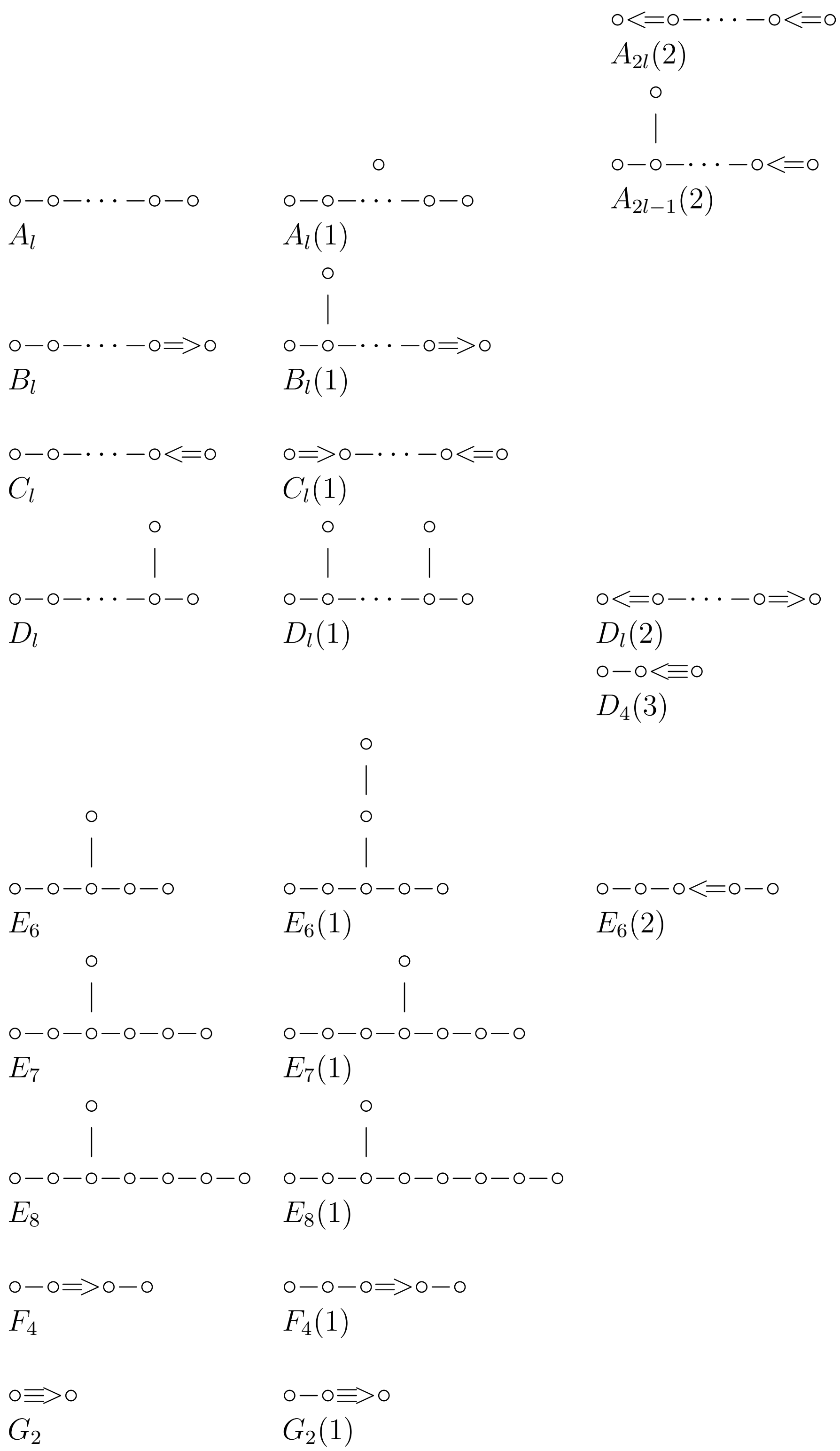
In this appendix we list the Dynkin diagrams of *all* simple generalized Cartan matrices (GCM) of type FIN and AFF. A symbol X_n indicates a GCM of type FIN. Recall that these are the Cartan matrices of the simple finite dimensional Lie algebras (over $\mathbb{C}$). The symbols $X_n(j)$ indicate GCMs of class AFF j. The range of l and the number of vertices for the diagrams $X_l(j)$ are given by the following table. For more details see [Kac], Chapter 4.

Symbol	$l \geq$	Number of vertices	Symbol	$l \geq$	Number of vertices
A_l	1	l	$C_l(1)$	2	l+1
B_l	2	l	$D_l(1)$	4	l+1
C_l	2	l	$A_{2l}(2)$	1	l+1
D_l	4	l	$A_{2l-1}(2)$	3	l+1
$A_l(1)$	2	l+1	$D_l(2)$	3	l+1
$B_l(1)$	3	l+1			

Recall how to discover a GCM from its graph: label the vertices by numbers from 1 to l, set $a_{ii} = 2$. If n lines are pointing from vertex j to vertex i, set $a_{j,i} = -n$, otherwise set $a_{j,i} = -1$; the case that i and j are not connected does not occur for simple GCMs. (Note that the resulting GCM is the transpose of the one associated to a graph as in [Kac]; however, since the Cartan matrices associated to Lie algebras as in [Kac] are also the transposes of the ones usually considered in the literature, the Lie algebras associated to a graph, as described here or in [Kac], coincide.) In the following list the special cases $A_1(1)$ and $A_2(2)$ have to be interpreted as

$$A_1(1) : \circ \overset{<=}{\underset{=>}{}} \circ = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \text{ and } A_2(2) : \circ \overset{<=}{\underset{<=}{}} \circ = \begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix},$$

respectively.



$A_{2l}(2)$
$A_{2l-1}(2)$
A_l
$A_l(1)$
B_l
$B_l(1)$
C_l
$C_l(1)$
D_l
$D_l(1)$
$D_l(2)$
$D_4(3)$
E_6
$E_6(1)$
$E_6(2)$
E_7
$E_7(1)$
E_8
$E_8(1)$
F_4
$F_4(1)$
G_2
$G_2(1)$

Bibliography

[Kac] V.G. Kac, Infinite dimensional Lie algebras, third edition, Cambridge University Press 1990 31, 32, 37, 38, 41, 44, 45, 47, 48, 50, 53

[Humphreys] J.E. Humphreys, Introduction to Lie Algebras and representation Theory, Graduate texts in Mathematics 9, Springer-Verlag 1972 8, 15, 18, 24, 25, 41

[Peterson-Kac] D.H. Peterson and V. Kac, Infinite flag varieties and conjugacy theorems, Proc. Natl. Acad. Sci. USA 80 (1983), 1778–1782 37

[Serre] J-P. Serre, Lie Algebras and Lie Groups, Lecture Notes in Mathematics 1500, Springer-Verlag 1992 6

[Serre] J-P. Serre, Algèbres de Lie semi-simples complexes, W.A. Benjamin, Inc. 1966 29