

Infinite Dimensional Lie Algebras

Songyu Ye

September 18, 2025

Abstract

These are notes for a reading course supervised by Prof. Richard Borcherds in the Fall of 2025 at UC Berkeley. The main references are [3]. There is an accompanying set of notes on loop groups, following [4].

Contents

1	Semisimple Lie algebras	2
1.1	Preliminaries	2
1.2	Highest weight modules	6
1.3	Category \mathcal{O} and the Weyl character formula	15
2	Introduction to infinite-dimensional Lie algebras	20
2.1	Basic definitions	20
2.2	Invariant bilinear form	27
2.3	Generalized Casimir operator	34
3	Integrable modules	40
4	Affine Lie algebras	41
5	Sept 17 meeting	46
6	Appendix A: Lie algebra cohomology	47
7	References	48

Bump [1] writes that the crown jewel of the theory of semisimple Lie algebras is the Weyl character formula, which gives a closed-form expression for the characters of all irreducible finite-dimensional representations. The proof of this formula uses the Casimir operator, which is constructed using an invariant bilinear form on the Lie algebra. We will generalize these concepts to

infinite-dimensional Lie algebras, specifically Kac-Moody algebras.

1 Semisimple Lie algebras

We recall the basic structure theory and representation theory of semisimple Lie algebras, culminating in a discussion of the Weyl character formula. Let \mathfrak{g} be a finite dimensional semisimple Lie algebra over the complex numbers \mathbb{C} . We follow [1].

1.1 Preliminaries

Definition 1.1. Recall that a Lie algebra is solvable if its derived series eventually becomes zero. The derived series is defined by $L^{(0)} = L$, $L^{(1)} = [L, L]$, and $L^{(n+1)} = [L^{(n)}, L^{(n)}]$.

Definition 1.2. A Lie algebra L is **semisimple** if and only if its radical (the largest solvable ideal) is zero.

Definition 1.3. The Killing form of a Lie algebra L is the bilinear form $\kappa(x, y) = \text{tr}(\text{ad}(x)\text{ad}(y))$.

Recall that there is an big Adjoint representation $\text{Ad} : G \rightarrow \text{GL}(\mathfrak{g})$ of a Lie group G on its Lie algebra \mathfrak{g} , defined by $\text{Ad}_g(X) = gXg^{-1}$. The differential of this at the identity is the little adjoint representation of the Lie algebra \mathfrak{g} on itself, defined by

$$\text{ad}(X)Y = \left. \frac{d}{dt} \right|_{t=0} \text{Ad}_{\exp(tX)}(Y).$$

In coordinates, $\text{ad}(X) = [X, Y]$.

Now the key idea is to consider the adjoint representation of \mathfrak{h} on \mathfrak{g} . Since \mathfrak{h} consists of simultaneously diagonalizable endomorphisms, $\text{ad}(H)$ is diagonalizable. Thus we can decompose:

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

where

$$\mathfrak{g}_{\alpha} = \{X \in \mathfrak{g} : [H, X] = \alpha(H)X \ \forall H \in \mathfrak{h}\}.$$

The nonzero functionals $\alpha \in \mathfrak{h}^*$ are the roots, and Φ is the root system. The following proposition is the beginning of the story of the geometry of root systems. The nondegeneracy of the Killing form on the Cartan subalgebra allows us to identify \mathfrak{h} with its dual \mathfrak{h}^* . In particular, we get an inner product, lengths, angles, and reflections. However, note that the notion of simple, positive, and integral roots does not come from the Killing form.

Proposition 1.4 (Humphreys 8.2). *The Killing form κ is nondegenerate on \mathfrak{h} .*

Proof. Recall that a Lie algebra is semisimple if and only if its Killing form is nondegenerate. One shows that the restriction of the Killing form to the centralizer $L_0 = C_L(\mathfrak{h})$ of \mathfrak{h} in L is nondegenerate. Then one shows that in fact $L_0 = \mathfrak{h}$. \square

Let \mathfrak{g} be a semisimple Lie algebra, \mathfrak{h} a maximal toral subalgebra, $\Phi \subset \mathfrak{h}^*$ the root system, which we partition into positive and negative roots.

Thus we will associate to $\phi \in \mathfrak{h}^*$ an element $t_\phi \in \mathfrak{h}$ such that

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

Then we will define an inner product on \mathfrak{h}^* which we will denote

$$(\lambda|\mu) = \kappa(t_\lambda, t_\mu).$$

If $\alpha \in \Phi \cup \{0\}$ we will denote

$$\mathfrak{g}_\alpha = \{x \in \mathfrak{g} \mid [h, x] = \alpha(h)x \text{ for } h \in \mathfrak{h}\}.$$

If $\alpha = 0$, then $\mathfrak{g}_\alpha = \mathfrak{h}$. On the other hand, if $\alpha \in \Phi$, then \mathfrak{g}_α is one-dimensional. We have the decomposition

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

By Humphreys Proposition 8.3(c) we have, for $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_{-\alpha}$,

$$[x, y] = \kappa(x, y)t_\alpha.$$

We will denote by

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

Thus if $x_\alpha \in \mathfrak{g}_\alpha$ we have $[h_\alpha, x_\alpha] = 2x_\alpha$. If

$$\alpha^\vee = \frac{2\alpha}{(\alpha|\alpha)}$$

then for $\lambda \in \mathfrak{h}^*$ we have $(\alpha^\vee | \lambda) = \lambda(h_\alpha)$. Either α^\vee or h_α is called a **coroot**. They are really the same thing if we identify \mathfrak{h} with its double dual \mathfrak{h}^{**} . The factor $\frac{2}{(\alpha | \alpha)}$ is chosen so that the root vectors x_α, y_α together with h_α form a standard \mathfrak{sl}_2 -triple, with the eigenvalue of x_α under h_α equal to 2.

We will denote by ρ half the sum of the positive roots. Humphreys denotes this δ , but the notation ρ is now universally used by everyone. If $\alpha \in \Phi$ we will denote by r_α the reflection

$$r_\alpha(x) = x - (x | \alpha^\vee) \alpha.$$

Remark 1.5. Recall that for each root α , we have an \mathfrak{sl}_2 -triple $\{x_\alpha, y_\alpha, h_\alpha\}$. When you restrict the adjoint representation of \mathfrak{g} to this \mathfrak{sl}_2 , every other root space \mathfrak{g}_β becomes a finite-dimensional \mathfrak{sl}_2 -representation.

For a root $\alpha \in \Phi$, we can find $x_\alpha \in \mathfrak{g}_\alpha$, $y_\alpha \in \mathfrak{g}_{-\alpha}$, and $h_\alpha = [x_\alpha, y_\alpha] \in \mathfrak{h}$ satisfying the relations of \mathfrak{sl}_2 :

$$\begin{aligned} [h_\alpha, x_\alpha] &= 2x_\alpha, \\ [h_\alpha, y_\alpha] &= -2y_\alpha, \\ [x_\alpha, y_\alpha] &= h_\alpha. \end{aligned}$$

So $\mathfrak{s}_\alpha = \langle x_\alpha, y_\alpha, h_\alpha \rangle \cong \mathfrak{sl}_2$ is a subalgebra of \mathfrak{g} . The adjoint representation is $\text{ad} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$ with $\text{ad}(z)(w) = [z, w]$. If we restrict ad to the subalgebra \mathfrak{s}_α , then \mathfrak{g} becomes an \mathfrak{sl}_2 -module.

Take another root $\beta \in \Phi$, $\beta \neq \pm\alpha$. For any $H \in \mathfrak{h}$, $[H, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_\beta$, so \mathfrak{g}_β is invariant under \mathfrak{h} . In particular, under h_α , vectors in \mathfrak{g}_β have weight $[h_\alpha, x_\beta] = \beta(h_\alpha) x_\beta$ for $x_\beta \in \mathfrak{g}_\beta$.

x_α and y_α act as “raising” and “lowering” operators:

$$\begin{aligned} [x_\alpha, \mathfrak{g}_\beta] &\subseteq \mathfrak{g}_{\beta+\alpha}, \\ [y_\alpha, \mathfrak{g}_\beta] &\subseteq \mathfrak{g}_{\beta-\alpha}. \end{aligned}$$

So if you start with a vector in \mathfrak{g}_β , repeated commutators with x_α and y_α move you up and down the α -string through β : $\beta, \beta + \alpha, \beta + 2\alpha, \dots, \beta - q\alpha$.

Now suppose \mathfrak{g} is finite dimensional. Then there are only finitely many roots, so this process stops in both directions. Thus the span $\bigoplus_k \mathfrak{g}_{\beta+k\alpha}$ is a finite-dimensional representation of \mathfrak{sl}_2 . It decomposes into irreducible \mathfrak{sl}_2 -modules, with weights given by integers

$$\beta(h_\alpha), \beta(h_\alpha) - 2, \dots, \beta(h_\alpha) - 2m$$

From this structure one proves that if β is a root, then so is $\beta - \langle \beta, \alpha^\vee \rangle \alpha$, which is exactly the reflection of β across the hyperplane orthogonal to α . The map

$$r_\alpha(x) = x - (x | \alpha^\vee) \alpha$$

is literally a reflection in the Euclidean space $V = \mathbb{R}\Phi \subset \mathfrak{h}^*$. One checks that $r_\alpha(\Phi) = \Phi$, i.e. it permutes the set of roots.

In addition, since all the finite dimensional irreducible representations of \mathfrak{sl}_2 are classified and have 1-dimensional weight spaces, it follows that $\dim(\mathfrak{g}_\alpha) = 1$ for all roots α .

Beware that this argument does not work if \mathfrak{g} is infinite-dimensional. In particular, Kac-Moody algebras have imaginary roots which do not behave like real roots. The reflections r_α for real roots α still exist, but they do not generate a group which permutes all the roots.

If α is a simple root, we will also use the notation s_α for r_α . We have proved that s_α maps α to its negative and permutes the remaining positive roots. Therefore $s_\alpha(\rho) = \rho - \alpha$ and so

$$(\rho|\alpha^\vee) = 1$$

for all simple roots α .

An element λ of \mathfrak{h}^* is called an **integral weight** if $(\lambda|\alpha^\vee) \in \mathbb{Z}$ for all $\alpha \in \Phi^+$, or equivalently, for all simple roots α . The integral weights form a lattice

$$\Lambda = \{x \in V | (\alpha^\vee|x) \in \mathbb{Z} \text{ for } \alpha \in \Phi^+\},$$

called the **weight lattice**. We call $\lambda \in \mathfrak{h}^*$ **dominant** if $(\lambda|\alpha^\vee) \geq 0$ for all $\alpha \in \Phi^+$ (or equivalently for simple roots α). We call λ **strongly dominant** if $(\lambda|\alpha^\vee) > 0$. Thus the special vector ρ is a strongly dominant integral weight.

Here are a couple of important properties of the Weyl group action. Let V be the \mathbb{R} -span of Φ in \mathfrak{h}^* . The inner product $(|)$ makes V into a Euclidean space, and $\mathfrak{h}^* = V + iV$. The set

$$C^+ = \{x \in V | (\alpha^\vee|x) > 0 \text{ for } \alpha \in \Phi^+\}$$

is called the **positive Weyl chamber**. The dominant weights are the ones in C^+ .

Proposition 1.6 (Fundamental domain). *The positive Weyl chamber is a fundamental domain for the action of the Weyl group: if $x \in V$ there is a unique element of C^+ in the W orbit of x .*

Recall that there is a partial order (known as dominance order) \succeq on \mathfrak{h}^* defined by

$$\lambda \succeq \mu \iff \lambda - \mu = \sum_{\alpha \in \Phi^+} n_\alpha \alpha \text{ with } n_\alpha \in \mathbb{Z}_{\geq 0}.$$

Proposition 1.7. *Let λ be a dominant, integral weight and let $w \in W$. Then $\lambda \succeq w\lambda$.*

1.2 Highest weight modules

Let V be a \mathfrak{g} -module. For $\lambda \in \mathfrak{h}^*$ we denote the **weight space**

$$V_\lambda = \{v \in V \mid h \cdot v = \lambda(h)v \text{ for } h \in \mathfrak{h}\}.$$

We will say that V is **\mathfrak{h} -diagonalizable** if V is the algebraic direct sum of the V_λ .

Proposition 1.8. *If V is \mathfrak{h} -diagonalizable, then so is any submodule or quotient module.*

Proof. Let $U \subset V$ be a submodule. We must show that an element of U may be expressed as a finite linear sum of $u_\lambda \in U_\lambda$. Since V has a weight space decomposition, we may write u as a sum of $u_\lambda \in V_\lambda$, and the problem is then to show that $u_\lambda \in U$. There exist a finite number of λ_i such that

$$u = \sum_{i=1}^m u_{\lambda_i},$$

and we choose $h \in \mathfrak{h}$ such that the values $\lambda_i(h)$ are all distinct. Then for $j = 0, \dots, m-1$

$$h^j \cdot u = \sum \lambda_i(h)^j u_{\lambda_i} \in \mathfrak{h}.$$

The $m \times m$ matrix $\{\lambda_i(h)^j\}$ is invertible since its determinant is a Vandermonde determinant. Consider the vectors $u, h \cdot u, h^2 \cdot u, \dots, h^{m-1} \cdot u$. Each is in U , and together they form a linear system:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1(h) & \lambda_2(h) & \cdots & \lambda_m(h) \\ \lambda_1(h)^2 & \lambda_2(h)^2 & \cdots & \lambda_m(h)^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1(h)^{m-1} & \lambda_2(h)^{m-1} & \cdots & \lambda_m(h)^{m-1} \end{bmatrix} \begin{bmatrix} u_{\lambda_1} \\ u_{\lambda_2} \\ \vdots \\ u_{\lambda_m} \end{bmatrix} = \begin{bmatrix} u \\ h \cdot u \\ \vdots \\ h^{m-1} \cdot u \end{bmatrix}.$$

This is a Vandermonde system. The matrix is invertible because the $\lambda_i(h)$ are distinct. Applying the inverse to this shows that each $u_{\lambda_i} \in U$, as required.

This proves that a submodule of a \mathfrak{h} -diagonalizable module is diagonalizable. It follows that the same is true for quotient modules, with $(V/U)_\lambda = V_\lambda/U_\lambda$. \square

We will work exclusively with diagonalizable modules with $\dim(V_\lambda) < \infty$ for all $\lambda \in \mathfrak{h}^*$. We will define the **support** $\text{supp}(V) = \{\lambda \in \mathfrak{h}^* \mid V_\lambda \neq 0\}$.

Let $U(\mathfrak{g})$ be the universal enveloping algebra. Let \mathfrak{n}^+ be the nilpotent subalgebra of \mathfrak{g} generated by the \mathfrak{g}_α ($\alpha \in \Phi^+$), and let \mathfrak{n}^- be the subalgebra generated by the \mathfrak{g}_α with $\alpha \in \Phi^-$. Then clearly we have the **triangular decomposition**

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+.$$

Lemma 1.9. We have $U(\mathfrak{g}) \cong U(\mathfrak{n}^-) \otimes U(\mathfrak{h}) \otimes U(\mathfrak{n}^+)$ in the sense that the multiplication map

$$U(\mathfrak{n}^-) \times U(\mathfrak{h}) \times U(\mathfrak{n}^+) \longrightarrow U(\mathfrak{g})$$

induces a vector space isomorphism $U(\mathfrak{n}^-) \otimes U(\mathfrak{h}) \otimes U(\mathfrak{n}^+) \longrightarrow U(\mathfrak{g})$.

Proof. This follows from the Poincaré-Birkhoff-Witt theorem (PBW) together with the triangular decomposition. Namely, if $\{x_i\}$ is a basis for \mathfrak{g} , then PBW asserts that a basis for $U(\mathfrak{g})$ consists of all elements of the form

$$x_1^{k_1} \cdots x_d^{k_d}, \quad 0 \leq k_i \in \mathbb{Z}.$$

Now we take the basis in a particular way, where its first $\frac{1}{2}|\Phi|$ elements are a basis for \mathfrak{n}^- , the next ℓ elements are a basis for \mathfrak{h} , and the last $\frac{1}{2}|\Phi|$ elements are a basis for \mathfrak{n}^+ . Then the element $x_1^{k_1} \cdots x_d^{k_d}$ factors uniquely as a product abc where a runs through a basis of $U(\mathfrak{n}^-)$, b runs through a basis of $U(\mathfrak{h})$ and c runs through a basis of $U(\mathfrak{n}^+)$. \square

We will call a vector $v \in V$ a **highest weight vector** of weight λ if $v \in V_\lambda$ and if $x_\alpha v = 0$ for $\alpha \in \Phi^+$. (Humphreys calls such v a **maximal vector**.) We will call V a **highest weight module** of weight λ if it is generated by a highest weight vector $v \in V_\lambda$. (Humphreys calls a highest weight module a **standard cyclic module**.)

Proposition 1.10. Suppose that $v \in V$ is a highest weight vector. Then the \mathfrak{g} -submodule $U(\mathfrak{g})v$ generated by v equals $U(\mathfrak{n}^-)v$. The weight space $V_\mu = 0$ unless $\mu \preceq \lambda$. We have $\dim(V_\lambda) = 1$.

Proof. We note that any element of $U(\mathfrak{n}^+)$ may be written as a constant times an element of the left ideal $U(\mathfrak{n}^+)\mathfrak{n}^+$, but this ideal annihilates v , so $U(\mathfrak{n}^+)v = \mathbb{C}v$. Similarly $U(\mathfrak{h})v = \mathbb{C}v$ since $v \in V_\lambda$. By Lemma 4,

$$U(\mathfrak{g}) = U(\mathfrak{n}^-)U(\mathfrak{h})U(\mathfrak{n}^+)v = U(\mathfrak{n}^-)v.$$

Consider the basis $\{x_{-\alpha}\}$ ($\alpha \in \Phi^+$) of \mathfrak{n}^- with $x_{-\alpha} \in \mathfrak{g}_{-\alpha}$. Using a fixed order on Φ^+ , the elements $\prod_{\alpha \in \Phi^+} x_{-\alpha}^{k_\alpha}$ are a PBW basis of $U(\mathfrak{n}^-)$. Since $x_{-\alpha}$ maps V_μ to $V_{\mu-\alpha}$,

$$\prod_{\alpha \in \Phi^+} x_{-\alpha}^{k_\alpha} v \in V_\mu, \quad \mu = \lambda - \sum_{\alpha \in \Phi^+} k_\alpha \alpha,$$

so $\mu \preceq \lambda$. Unless all $k_\alpha = 0$, μ is strictly $\prec \lambda$, so V_λ is one-dimensional. \square

Remark 1.11. Beware that highest weight modules need not be irreducible. It is true that if a highest weight module is finite dimensional, then it is irreducible and uniquely determined by its highest weight, which must be a dominant integral weight. But infinite-dimensional highest weight modules need not be irreducible, and even if they are irreducible, they need not be uniquely determined by their highest weight.

In particular, if you consider $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$, then basis vectors of $M(\lambda)$ are $\{f^k v_\lambda : k \geq 0\}$. We compute the action of e :

$$ef^m v_\lambda = m(\lambda - m + 1)f^{m-1}v_\lambda.$$

At $m = \lambda + 1$: This becomes

$$ef^{\lambda+1}v_\lambda = 0,$$

but the vector $f^{\lambda+1}v_\lambda$ is nonzero in $M(\lambda)$. So $f^{\lambda+1}v_\lambda$ is a new singular vector, generating a proper submodule. $L(\lambda)$ is defined as

$$L(\lambda) = M(\lambda) / \langle f^{\lambda+1}v_\lambda \rangle,$$

i.e. we quotient out the submodule generated by that new singular vector. In $L(\lambda)$, the vector $f^{\lambda+1}v_\lambda$ is identically zero, so we get no new singular vectors.

Remark 1.12 (Computing the product ef^m in \mathfrak{sl}_2). Using the relation $[e, f] = h$, we can induct: For $m = 1$:

$$ef = fe + h.$$

Suppose for m :

$$ef^m = f^m e + mf^{m-1}(h - m + 1).$$

Then for $m + 1$:

$$\begin{aligned} ef^{m+1} &= e(f^m f) = (ef^m)f \\ &= (f^m e + mf^{m-1}(h - m + 1))f \\ &= f^m ef + mf^{m-1}(h - m + 1)f. \end{aligned}$$

Now expand each part:

$$\begin{aligned} f^m ef &= f^m(fe + h) = f^{m+1}e + f^m h, \\ f^{m-1}(h - m + 1)f &= f^{m-1}(fh - 2f - (m - 1)f) = f^m h - (m + 1)f^m. \end{aligned}$$

where in the second line we used $[h, f] = -2f$. So altogether:

$$\begin{aligned} ef^{m+1} &= f^{m+1}e + f^m h + m(f^m h - (m + 1)f^m) \\ &= f^{m+1}e + (m + 1)f^m h - m(m + 1)f^m. \end{aligned}$$

Proposition 1.13. Let V be a highest weight module with highest weight λ . A submodule U of V is proper if and only if $U \cap V_\lambda = 0$.

Proof. Since $\dim(V_\lambda) = 1$, if $U \cap V_\lambda \neq 0$ then $V_\lambda \subseteq U$ and then since V_λ generates V , it is clear that $U = V$. On the other hand if $U \cap V_\lambda = 0$ then clearly U is proper. \square

Proposition 1.14. Let V be a highest weight module with highest weight λ . Then V has a unique maximal proper submodule. Moreover V has a unique irreducible quotient.

Proof. Let Σ be the set of proper submodules of V , and let

$$W = \sum_{U \in \Sigma} U.$$

By Proposition 3 each $U \in \Sigma$ is diagonalizable, so evidently for $\mu \in \mathfrak{h}^*$

$$W_\mu = \sum_{U \in \Sigma} U_\mu.$$

We apply this with $\mu = \lambda$. Since $U \in \Sigma$ is proper, $U_\lambda = 0$ by the previous proposition, and so $W_\lambda = 0$. This shows that W is proper. We have proved that W is the unique maximal proper submodule of V , and consequently V/W is the unique irreducible quotient. \square

Theorem 1.15. *Let $\lambda \in V^*$. There is a highest weight module $M = M(\lambda)$ with highest weight vector $m \in M_\lambda$ with the following universal property. If V is another highest weight module with highest weight λ and if $v \in V_\lambda$, then there is a unique \mathfrak{g} -module homomorphism $M \rightarrow V$ mapping $m \mapsto v$. The map $\xi \mapsto \xi \cdot v$ is vector space isomorphism $U(\mathfrak{n}^-) \rightarrow M$.*

Proof. Note that since \mathfrak{h} normalizes \mathfrak{n}^+ , $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}^+$ is a subalgebra of \mathfrak{g} , the “Borel subalgebra.” As in Lemma 4, $U(\mathfrak{g}) \cong U(\mathfrak{n}^-) \otimes U(\mathfrak{b})$, that is, the multiplication map $U(\mathfrak{n}^-) \times U(\mathfrak{b}) \rightarrow U(\mathfrak{g})$ induces a vector space isomorphism $U(\mathfrak{n}^-) \otimes U(\mathfrak{b}) \rightarrow U(\mathfrak{g})$. This result is a simple consequence of this fact.

To elaborate, regarding \mathbb{C} as a one-dimensional abelian Lie algebra, we have a Lie algebra homomorphism $\theta_\lambda : \mathfrak{b} \rightarrow \mathbb{C}$ that maps $H \in \mathfrak{h}$ to $\lambda(H)$, and \mathfrak{n}^+ to zero. Thus let H_1, \dots, H_ℓ be a basis of \mathfrak{h} and x_α ($\alpha \in \Phi^+$) be a basis of \mathfrak{n}^+ . By the PBW theorem, the elements

$$H_1^{k_1} \cdots H_\ell^{k_\ell} \prod_{\alpha \in \Phi^+} x_\alpha^{k_\alpha}$$

with k_i and k_α nonnegative integers are a basis for $U(\mathfrak{b})$. It is understood that in the product $\prod x_\alpha^{k_\alpha}$ the roots $\alpha \in \Phi^+$ are taken in a fixed definite order. We then have

$$\theta_\lambda(H_1^{k_1} \cdots H_\ell^{k_\ell} \prod_{\alpha} x_\alpha^{k_\alpha}) = \begin{cases} \prod \lambda(H_i)^{k_i} & \text{if all } k_\alpha = 0, \\ 0 & \text{if any } k_\alpha > 0. \end{cases}$$

Now let J_ψ be the left ideal generated by $\xi - \theta_\lambda(\xi)$ for $\xi \in \mathfrak{b}$. Let

$$M = M(\lambda) = U(\mathfrak{g})/J_\psi,$$

and let m be the image of $1 \in U(\mathfrak{g})$ in $M(\lambda)$.

It is clear from the PBW theorem that $Hv = \lambda(H)v$ for $H \in \mathfrak{h}$, while $\mathfrak{n}^+v = 0$, and moreover from $U(\mathfrak{g}) \cong U(\mathfrak{n}^-) \otimes U(\mathfrak{b})$, it is clear that every element of $M(\lambda)$ may be written uniquely as $\eta \cdot v$ for $\eta \in U(\mathfrak{n}^-)$.

Now let us verify the universal property. Let V be a highest weight module with weight λ , and let $v_\lambda \in V_\lambda$ be a generator. Then we have a surjective $U(\mathfrak{g})$ -module homomorphism

$$U(\mathfrak{g}) \rightarrow V, \quad \xi \mapsto \xi \cdot v_\lambda,$$

and since $\beta \cdot v = \theta_\lambda(\beta)v$ for $\beta \in \mathfrak{b}$, J_ψ is in the kernel. Thus the map factors uniquely through $U(\mathfrak{g})/J_\psi = M(\lambda)$. \square

Corollary 1.16. *Let $\lambda \in \mathfrak{h}^*$. Up to isomorphism, \mathfrak{g} has a unique irreducible highest weight module $L(\lambda)$ with highest weight λ .*

Proof. Every highest weight module is a quotient of $M(\lambda)$. Since $M(\lambda)$ has a unique irreducible quotient, there is a unique irreducible highest weight module. \square

Remark 1.17. *The irreducible quotient $L(\lambda)$ might be finite or infinite dimensional. Recall that λ is called **integral** if $\langle \alpha^\vee, \lambda \rangle \in \mathbb{Z}$ for all coroots α^\vee , and **dominant** if $\langle \alpha^\vee, \lambda \rangle \geq 0$. If λ is a dominant integral weight, then $L(\lambda)$ is finite-dimensional. On the other hand if λ is not integral, $L(\lambda)$ will be infinite dimensional, and unless $\langle \alpha^\vee, \lambda \rangle \in \mathbb{Z}$ for some coroot α^\vee , we will actually have $M(\lambda)$ irreducible, and $L(\lambda) = M(\lambda)$.*

Remark 1.18 (Reducible highest weight modules are not unique). *Let \mathfrak{g} have two simple roots α_1, α_2 (e.g. \mathfrak{sl}_3). Fix a weight λ such that both integers $\langle \lambda + \rho, \alpha_1^\vee \rangle, \langle \lambda + \rho, \alpha_2^\vee \rangle$ are positive. Then the Verma module $M(\lambda)$ contains two distinct singular vectors (i.e. highest weight vectors inside $M(\lambda)$ below the top) of weights $s_1 \cdot \lambda$ and $s_2 \cdot \lambda$ (dot action). They generate two different submodules $N_1 = U(\mathfrak{g})v_{s_1 \cdot \lambda}, N_2 = U(\mathfrak{g})v_{s_2 \cdot \lambda}$. Now the quotients $M(\lambda)/N_1, M(\lambda)/N_2$ are both highest weight modules of highest weight λ , both reducible, and not isomorphic (their composition series differ). Hence reducible highest weight modules with the same top weight are not unique.*

Remark 1.19 (Importance of the dot action). *The BGG theorem tells us: if μ is a weight such that $\langle \lambda + \rho, \alpha^\vee \rangle \in \mathbb{Z}_{>0}$, then there exists a singular vector in $M(\lambda)$ of weight $s_\alpha \cdot \lambda := s_\alpha(\lambda + \rho) - \rho$.*

Additionally, in category \mathcal{O} , irreducible highest weight modules $L(\lambda)$ can only appear as composition factors of Verma modules $M(\mu)$ if λ and μ are in the same dot-orbit under the Weyl group. The dot action partitions the weight lattice into blocks inside which the category decomposes.

One can also give an interpretation via Harish-Chandra isomorphism. The center $Z(U(\mathfrak{g}))$ acts on a Verma module $M(\lambda)$ by a character. Harish-Chandra's isomorphism says these central characters are Weyl group invariant under the dot action. In other words, two highest weights λ, μ have

the same central character iff they're in the same dot orbit.

More precisely, the BGG theorem states:

Theorem 1.20 (BGG theorem). *If $\lambda \in \mathfrak{h}^*$, α a positive root, and $m = \langle \lambda + \rho, \alpha^\vee \rangle \in \mathbb{Z}_{>0}$, then there exists a nonzero homomorphism of Verma modules $M(s_\alpha \cdot \lambda) \hookrightarrow M(\lambda)$, where s_α is the reflection in the Weyl group, and the dot action is $w \cdot \lambda = w(\lambda + \rho) - \rho$.*

Concretely: inside $M(\lambda)$, there is a singular vector of weight $s_\alpha \cdot \lambda$, which generates a highest weight submodule isomorphic to $M(s_\alpha \cdot \lambda)$. If $w \leq w'$ in Bruhat order, then $M(w \cdot \lambda) \hookrightarrow M(w' \cdot \lambda)$.

Theorem 1.21 (Classification of finite-dimensional irreducible modules). *Let V be a finite dimensional irreducible module. Then $V \cong L(\lambda)$ where λ is a dominant integral weight. Conversely, if λ is a dominant integral weight, then $L(\lambda)$ is finite-dimensional.*

Proof. Assume that V is finite-dimensional. Choose a vector $v \in V_\lambda$ where λ is a weight of V that is maximal with respect to \succ . If $\alpha \in \Phi^+$ then $x_\alpha v \in V_{\lambda+\alpha}$ so $x_\alpha v = 0$. Therefore v is a highest weight vector. Then $V = U(\mathfrak{g})v$ since V is irreducible. We have proved that V is a highest weight module; it is irreducible so $V \cong L(\lambda)$.

To show that λ is a dominant integral weight, let α be a simple positive root. The restriction of V to the \mathfrak{sl}_2 spanned by $x_\alpha, x_{-\alpha}$ and h_α is finite-dimensional, and $x_\alpha v = 0$. From the classification of finite-dimensional \mathfrak{sl}_2 -modules, this means that $(\alpha^\vee | \lambda) = \lambda(h_\alpha) \in \mathbb{Z}$ is a nonnegative integer. Therefore λ is dominant and integral.

We will omit the slightly tedious proof of the converse, that if λ is a dominant integral weight then $L(\lambda)$ is finite-dimensional. For a proof of this see Kac, Lemma 10.1. \square

Corollary 1.22 (Weyl). *For V an irreducible finite-dimensional \mathfrak{g} -module, the highest weight λ is a dominant integral weight, and*

$$V \longleftrightarrow \lambda$$

is a bijection between the irreducible highest weight modules and the dominant integral weights.

Now let \mathfrak{g} be finite dimensional. The Casimir element of the universal enveloping algebra $U(\mathfrak{g})$ may be defined as follows. Let $\{\gamma_i\}$ be a basis of \mathfrak{g} and $\{\gamma^i\}$ the dual basis with respect to the Killing form, so $\kappa(\gamma_i, \gamma^j) = \delta_{ij}$. Then

$$c_{\mathfrak{g}} = \sum_{i=1}^{\dim(\mathfrak{g})} \gamma_i \gamma^i$$

Proposition 1.23. $c_{\mathfrak{g}}$ is independent of the choice of basis $\{\gamma_i\}$. It lies in the center of $U(\mathfrak{g})$.

Proof. For any finite-dimensional vector space V , there is a canonical iso $V \otimes V^* \cong \text{End}(V)$, $v \otimes \phi \mapsto (w \mapsto \phi(w)v)$. Under this isomorphism, the element $\sum_i v_i \otimes \phi_i$, where $\{v_i\}$ is a basis and $\{\phi_i\}$ the dual basis, corresponds to the identity map on V . This element is independent of the chosen basis (it's just the coordinate expression of the identity endomorphism). Now take $V = \mathfrak{g}$ with the Killing form κ . So the canonical element $\sum_i \gamma_i \otimes \gamma^i \in \mathfrak{g} \otimes \mathfrak{g}$ corresponds to the identity operator $\text{id}_{\mathfrak{g}}$. Finally push Ω into $U(\mathfrak{g})$ using multiplication

$$c_{\mathfrak{g}} = m(\Omega) = \sum_i \gamma_i \gamma^i$$

This shows that $c_{\mathfrak{g}}$ is independent of the choice of basis. To see that $c_{\mathfrak{g}}$ is central, let $x \in \mathfrak{g}$. Then

Write

$$[x, \gamma_i] = \sum_j a_{ij} \gamma_j.$$

Use ad-invariance of κ :

$$0 = \kappa([x, \gamma_i], \gamma^k) + \kappa(\gamma_i, [x, \gamma^k]) = a_{ik} + \kappa(\gamma_i, [x, \gamma^k]).$$

If we expand $[x, \gamma^k] = \sum_j b_{kj} \gamma^j$, the relation above gives

$$b_{ki} = -a_{ik},$$

i.e.

$$[x, \gamma^k] = - \sum_i a_{ik} \gamma^i.$$

Now compute in $U(\mathfrak{g})$:

$$\begin{aligned} [x, c_{\mathfrak{g}}] &= \sum_i [x, \gamma_i] \gamma^i + \sum_i \gamma_i [x, \gamma^i] \quad \text{by the Leibniz rule} \\ &= \sum_{i,j} a_{ij} \gamma_j \gamma^i - \sum_{i,k} a_{ki} \gamma_i \gamma^k. \end{aligned}$$

so $[x, c_{\mathfrak{g}}] = 0$. \square

Proposition 1.24. Let h_i be a basis of \mathfrak{h} and let h^i be the dual basis of \mathfrak{h} with respect to the Killing form, so $\kappa(h_i, h^j) = \delta_{ij}$. Then if $\lambda, \mu \in \mathfrak{h}^*$ we have

$$(\lambda|\mu) = \sum_i \lambda(h^i) \mu(h_i).$$

Proof. Recall that the defining property of t_μ is that

$$\kappa(t_\mu, h) = \mu(h) \quad \text{for all } h \in \mathfrak{h}.$$

First let us show that

$$t_\mu = \sum_i \mu(h_i) h_i. \tag{1}$$

To check this, we pair both sides with h_j . We have

$$\kappa(t_\mu, h_j) = \mu(h_j) = \kappa\left(\sum_i \mu(h_i) h_i, h_j\right) = \mu(h_j)$$

is exactly the defining property of t_μ . Since the h_j span \mathfrak{h} and κ restricted to \mathfrak{h} is nondegenerate, this proves (1).

Now (1) implies

$$(\lambda|\mu) = \kappa(t_\lambda, t_\mu) = \sum_i \mu(h_i) \kappa(t_\lambda, h_i) = \sum_i \mu(h_i) \lambda(h_i).$$

again using the defining property of t_μ . \square

Proposition 1.25. *Let V be a highest weight module with highest weight λ . Then the Casimir element $c_{\mathfrak{g}}$ acts by the scalar*

$$|\lambda + \rho|^2 - |\rho|^2$$

on V .

Proof. Since $c_{\mathfrak{g}}$ is central in $U(\mathfrak{g})$ it commutes with the action of \mathfrak{g} on any module V . Because V is generated by a highest weight vector $v_\lambda \in V_\lambda$, it is sufficient to show that

$$c_{\mathfrak{g}} v = (|\lambda + \rho|^2 - |\rho|^2) v$$

We need to choose dual bases of \mathfrak{g} with respect to the Killing form. For one basis, we choose a basis h_i of $\mathfrak{h} = \mathfrak{g}_0$, and vectors $x_\alpha \in \mathfrak{g}_\alpha$.

Now we describe the dual basis. We know that the Killing form is nondegenerate on \mathfrak{h} , so we find h^i such that $\kappa(h_i, h^j) = \delta_{ij}$. Then we define another set of representatives $y_\alpha \in \mathfrak{g}_\alpha$ so that

$$y_\alpha = \frac{x_\alpha}{\kappa(x_\alpha, x_{-\alpha})}$$

so that

$$\kappa(x_\alpha, y_{-\beta}) = \delta_{\alpha\beta}$$

Thus we have dual bases $\{h_i, x_\alpha\}$ and $\{h^i, y_{-\alpha}\}$. Then

$$c_{\mathfrak{g}} = \sum_i h_i h^i + \sum_{\alpha \in \Phi} x_\alpha y_{-\alpha}.$$

We want to rewrite this slightly. We write this as

$$c_{\mathfrak{g}} = \sum_i h_i h^i + \sum_{\alpha \in \Phi^+} x_\alpha y_{-\alpha} + \sum_{\alpha \in \Phi^+} x_{-\alpha} y_\alpha.$$

Now observe that $[x_\alpha, y_{-\alpha}] = t_\alpha$. Certainly $[x_\alpha, y_{-\alpha}] \in \mathfrak{h}$ so write it as t_α . Then for $h \in \mathfrak{h}$, we check that $\kappa([x_\alpha, y_{-\alpha}], h) = \alpha(h)$.

$$\kappa([x_\alpha, y_{-\alpha}], h) = \kappa(x_\alpha, [y_{-\alpha}, h])$$

But since h acts on the root vector $y_{-\alpha}$ by $[h, y_{-\alpha}] = -\alpha(h)y_{-\alpha}$, we get $[y_{-\alpha}, h] = \alpha(h)y_{-\alpha}$. So $\kappa([x_\alpha, y_{-\alpha}], h) = \kappa(x_\alpha, \alpha(h)y_{-\alpha}) = \alpha(h)\kappa(x_\alpha, y_{-\alpha})$.

And by construction of $y_{-\alpha}$, $\kappa(x_\alpha, y_{-\alpha}) = 1$. Thus $\kappa([x_\alpha, y_{-\alpha}], h) = \alpha(h)$.

That is exactly the defining property of t_α . Hence $[x_\alpha, y_{-\alpha}] = t_\alpha$.

So in the enveloping algebra

$$x_\alpha y_{-\alpha} = t_\alpha + y_{-\alpha} x_\alpha.$$

Thus

$$c_{\mathfrak{g}} = \sum_i h_i h^i + \sum_{\alpha \in \Phi^+} t_\alpha + \sum_{\alpha \in \Phi^+} (y_{-\alpha} x_\alpha + x_{-\alpha} y_\alpha).$$

Since v_λ is a highest weight vector it is annihilated by x_α and y_α when $\alpha \in \Phi^+$. On the other hand, $Hv_\lambda = \lambda(H)v_\lambda$ for $H \in \mathfrak{h}$, and so

$$c_{\mathfrak{g}} v_\lambda = \sum_i \lambda(h_i) \lambda(h^i) + \sum_{\alpha \in \Phi^+} \lambda(t_\alpha).$$

The first expression equals $(\lambda|\lambda)$ by the previous proposition, while

$$\sum_{\alpha \in \Phi^+} \lambda(t_\alpha) = \sum_{\alpha \in \Phi^+} \langle \lambda, \alpha \rangle = 2(\lambda|\rho).$$

Thus

$$c_{\mathfrak{g}} v_\lambda = ((\lambda|\lambda) + 2(\lambda|\rho)) v_\lambda = ((\lambda + \rho|\lambda + \rho) - (\rho|\rho)) v_\lambda,$$

as desired. \square

1.3 Category \mathcal{O} and the Weyl character formula

We will now prove the Weyl character formula following Kac. It will be useful to work in the following category of representations, Category \mathcal{O} , introduced by Bernstein, Gelfand and Gelfand.

Definition 1.26. *A module is in Category \mathcal{O} if it is \mathfrak{h} -diagonalizable with finite dimensional weight spaces V_λ , such that there exists a finite set of weights $\{\lambda_1, \dots, \lambda_N\}$ such that $V_\mu = 0$ unless $\mu \preceq \lambda_i$ for some i .*

This category contains all highest weight modules, is closed under finite direct sums, and it contains all submodules and quotient modules of a Category \mathcal{O} module. In particular it is an abelian category with enough projectives and injectives and has a good homological theory. The Verma modules $M(\lambda)$ may or may not be irreducible. We will say a module V is a **subquotient** of a module W if there are submodules $U \supset Q$ of W such that $U/Q \cong V$. Thus either a submodule or a quotient module is a subquotient.

Proposition 1.27. *Suppose that V is a highest weight module with weight μ and V is a subquotient of $M(\lambda)$. Then*

$$|\lambda + \rho|^2 = |\mu + \rho|^2.$$

Proof. Since c commutes with the action of \mathfrak{g} it must act as a scalar on $M(\lambda)$, and we computed that scalar to be $|\lambda + \rho|^2 - |\rho|^2$. So it acts by the same scalar on any submodule, quotient module or subquotient. Also c acts by the scalar $|\mu + \rho|^2 - |\rho|^2$ on any highest weight module V with highest weight λ , so

$$|\lambda + \rho|^2 - |\rho|^2 = |\mu + \rho|^2 - |\rho|^2.$$

as desired. \square

Definition 1.28. *Let V be a module in Category \mathcal{O} . We define the **character** of V to be the formal expression*

$$\chi_V = \sum_{\lambda} \dim(V_{\lambda}) e^{\lambda}$$

where e^{λ} is a formal symbol for $\lambda \in \mathfrak{h}^*$.

Proposition 1.29 (Character of Verma modules). *The character of $M(\lambda)$ is*

$$e^{\lambda} \prod_{\alpha \in \Phi^+} (1 - e^{-\alpha})^{-1}.$$

Proof. Let v be the highest weight vector. We recall from Theorem 8 that the map

$$\xi \mapsto \xi \cdot v$$

from $U(\mathfrak{n}^-)$ to $M(\lambda)$ is a vector space isomorphism. So by the PBW theorem a basis of $M(\lambda)$ consists of the vectors

$$\left(\prod_{\alpha \in \Phi^+} x_{-\alpha}^{k_\alpha} \right) v, \quad k_\alpha \geq 0,$$

where the positive roots Φ^+ are taken in some fixed definite order. The weight of this vector is

$$\lambda - \sum_{\alpha \in \Phi^+} k_\alpha \alpha,$$

so

$$\chi_V = e^\lambda \prod_{\alpha \in \Phi^+} e^{-k_\alpha \alpha} = e^\lambda \prod_{\alpha \in \Phi^+} (1 - e^{-\alpha})^{-1}.$$

as desired. \square

Remark 1.30. Note that we get a geometric series at the end because $M(\lambda)$ is a Verma module: it has no relations among the negative root vectors beyond the Lie algebra relations themselves. This is what makes Verma modules universal highest weight modules: you can push down indefinitely.

Definition 1.31. Let V be a module in Category \mathcal{O} . A nonzero vector $v \in V$ is called **primitive** if there exists a proper submodule $U \subset V$ such that $v \notin U$ but $x_\alpha v \in U$ for all $\alpha \in \Phi^+$ (or equivalently, for all simple roots). We can take $U = 0$, so if $x_\alpha v = 0$ then v is primitive. In other words, a highest weight vector is a primitive vector. More generally, v being primitive means that the image of v in V/U is a highest weight vector for some proper submodule U of V . We will call μ a **primitive weight** if V_μ contains a primitive vector.

A primitive vector is like a “hidden” highest weight vector, but visible only in a quotient.

Proposition 1.32. Let V be a module in Category \mathcal{O} . Then V is generated by its primitive vectors.

Proof. If not, consider the submodule U generated by the primitive vectors. Then $Q = V/U$ would be a nonzero submodule. If we choose a nonzero vector in Q whose weight is maximal with respect to \preceq , then its preimage in V would be a primitive vector, which is a contradiction. \square

Proposition 1.33. Let V be a module in Category \mathcal{O} . Assume that V has only a finite number of weights. Then V has finite length. That is, it has a composition series

$$V = V_m \supset V_{m-1} \supset \cdots \supset V_0 = 0$$

such that each quotient V_i/V_{i-1} is irreducible, isomorphic to $L(\mu)$, where μ is a primitive weight of V . (The quotients V_i/V_{i-1} are called **composition factors**, and they are independent of the composition series, by the Jordan–Hölder theorem.)

Proof. We argue by induction on the number of linearly independent primitive vectors.

Choose a primitive weight μ that is maximal with respect to \preceq . Then clearly a primitive vector v of weight μ must be a highest weight vector, so $W = U(\mathfrak{g}) \cdot v = U(\mathfrak{n}^-)v$ is a highest weight module. It has a maximal submodule W' and the quotient $Q = W/W'$ is irreducible. Both V/W and W' have fewer independent primitive vectors than V (note that there are finitely many weights and each weight space is finite dimensional since we are in Category \mathcal{O}), so by induction they have finite length. Since V/W , W' and the irreducible quotient W/W' all have finite length, it follows that V has finite length. \square

Proposition 1.34 (Character of irreducible highest weight modules). *Let $\lambda \in \mathfrak{h}^*$. Then the character $\chi_{L(\lambda)}$ is of the form*

$$\chi_{L(\lambda)} = \sum_{\substack{\mu \preceq \lambda \\ |\mu + \rho|^2 = |\lambda + \rho|^2}} c_\mu \chi_{M(\mu)} \quad (2)$$

where $c_\lambda = 1$.

Proof. The weight μ of a primitive vector must satisfy $\mu \preceq \lambda$ and $|\mu + \rho|^2 = |\lambda + \rho|^2$.

Since the inner product is positive definite, this implies that there are only a finite number of possible weights for primitive vectors (because $|\mu + \rho|^2 = |\lambda + \rho|^2$ cuts out a sphere, and the lattice of weights intersected with that sphere is finite). $M(\mu)$ has finite length because only finitely many irreducibles can appear as factors, and modules in Category \mathcal{O} have finite dimensional weight spaces. Also the composition factors of $M(\mu)$ must be $L(\nu)$ where $|\nu + \rho|^2 = |\mu + \rho|^2 = |\lambda + \rho|^2$. This is because every composition factor is a highest weight module and every irreducible highest weight module is of the form $L(\nu)$.

Let $d(\mu, \nu)$ be the multiplicity of such $L(\nu)$. Then

$$\chi_{M(\mu)} = \sum_{\substack{\nu \preceq \mu \\ |\nu + \rho|^2 = |\lambda + \rho|^2}} d(\mu, \nu) \chi_{L(\nu)}.$$

Now the matrix $d(\mu, \nu)$ indexed by pairs μ, ν is triangular since $d(\mu, \mu) = 1$ and $d(\mu, \nu) = 0$ unless $\nu \preceq \mu$. So it is invertible and we may write

$$\chi_{L(\mu)} = \sum_{\substack{\nu \preceq \mu \\ |\nu + \rho|^2 = |\lambda + \rho|^2}} d'(\mu, \nu) \chi_{M(\nu)}.$$

Applying this to $\mu = \lambda$ gives (2). \square

We will define the **Weyl denominator**

$$\Delta = e^\rho \prod_{\alpha \in \Phi^+} (1 - e^{-\alpha}).$$

Lemma 1.35. *Let $w \in W$ (the Weyl group). Then*

$$w(\Delta) = \text{sgn}(w)\Delta.$$

Proof. It is sufficient to check this if $w = s_{\alpha_i}$ is a simple reflection. We recall that s_{α_i} maps the simple root α_i to $-\alpha_i$ and it permutes the remaining positive roots. Moreover $s_{\alpha_i}(\rho) = \rho - \alpha_i$. So if we write

$$\Delta = e^\rho (1 - e^{-\alpha_i}) \prod_{\substack{\alpha \in \Phi^+ \\ \alpha \neq \alpha_i}} (1 - e^{-\alpha}),$$

then s_i maps $e^\rho(1 - e^{-\alpha_i})$ to

$$e^{\rho - \alpha_i}(1 - e^{\alpha_i}) = -e^\rho(1 - e^{-\alpha_i}),$$

and it fixes the product. Hence $s_i(\Delta) = -\Delta$. \square

Theorem 1.36 (Weyl Character Formula). *Let V be a finite-dimensional irreducible representation of \mathfrak{g} . Thus by Theorem 1.21 there is a dominant integral weight λ such that $V = L(\lambda)$. We have*

$$\chi_V = \Delta^{-1} \sum_{w \in W} \text{sgn}(w) e^{w(\lambda + \rho)}$$

where W is the Weyl group and

$$\Delta = e^\rho \prod_{\alpha \in \Phi^+} (1 - e^{-\alpha}).$$

The following argument is due to Kac, improving the proof of BGG. As an application, Kac extended the applicability of the Weyl character formula for characters of integrable representations of infinite-dimensional Kac-Moody Lie algebras. We will discuss this in more detail in the section, following Chapter 10 of Kac [3].

Proof. Using Proposition 1.29 we may rewrite (2) in the form (note that the ρ in the exponent comes from dividing by Δ)

$$\chi_{L(\lambda)} = \sum_{\substack{\mu \preceq \lambda \\ |\mu + \rho|^2 = |\lambda + \rho|^2}} c_\mu e^{\mu + \rho} \Delta^{-1}$$

It may be simpler to write this as

$$\chi_{L(\lambda)} = \sum_{\mu \in P^+} c_\mu e^{\mu+\rho} \Delta^{-1}$$

and remember that $c_\mu = 0$ unless $\mu \preceq \lambda$ and $|\mu + \rho|^2 = |\lambda + \rho|^2$. We claim that if $w \in W$, then

$$c_\mu = \text{sgn}(w) c_{w \circ \mu}. \quad (3)$$

Indeed, since $\chi_{L(\lambda)}$ is invariant under the action of W , and since $w(\Delta) = \text{sgn}(w)\Delta$, we have an identity

$$\sum_{\mu \in P^+} c_\mu e^{\mu+\rho} \Delta^{-1} = \sum_{\mu \in P^+} \text{sgn}(w) c_\mu e^{w(\mu+\rho)} \Delta^{-1}$$

and comparing the coefficients of $e^{w \circ \mu} = e^{w(\mu+\rho)-\rho}$ on both sides of this equation gives (2).

We know that $c_\lambda = 1$, since this is part of Proposition 1.34. So by (2), we will have terms corresponding to μ of the form $w \circ \lambda$ and the sum of these terms is

$$\Delta^{-1} \sum_{w \in W} c_{w \circ \lambda} e^{w(\lambda+\rho)-\rho} e^\rho = \Delta^{-1} \sum_{w \in W} \text{sgn}(w) e^{w(\lambda+\rho)}.$$

This is the right hand side of the Weyl character formula, so our task is to show that these are the **only terms**. That is, we must show that $c_\mu = 0$ unless μ is of the form $w \circ \lambda$ for some $w \in W$.

Therefore we start with μ such that $c_\mu \neq 0$. By Proposition 1.6, there exists $w \in W$ such that $w(\mu + \rho)$ is dominant. Let $\nu = w \circ \mu = w(\mu + \rho) - \rho$. We will show that $\nu = \lambda$. In any case by (2), $c_\nu \neq 0$ and so $\nu \preceq \lambda$ and $|\lambda + \rho|^2 = |\nu + \rho|^2$. We write

$$\lambda - \nu = \sum_{\alpha \in \Phi^+} k_\alpha \alpha,$$

where since $\nu \preceq \lambda$ we have $k_\alpha \geq 0$. We note the identity, for $a, b \in \mathfrak{h}^*$:

$$|a|^2 - |b|^2 = (a + b|a - b).$$

We apply this and learn that

$$|\lambda + \rho|^2 - |\nu + \rho|^2 = \left(\lambda + \nu + 2\rho \mid \sum_{\alpha \in \Phi^+} k_\alpha \alpha \right).$$

Now λ and $\nu + \rho = w(\mu + \rho)$ are both dominant, so $\lambda + \nu + 2\rho$ is **strongly dominant** meaning

$$(\alpha^\vee \mid \lambda + \nu + 2\rho) > 0$$

for all positive roots α . So $|\lambda + \rho|^2 = |\nu + \rho|^2$ implies that $k_\alpha = 0$ for all α and so $\nu = \lambda$. \square

2 Introduction to infinite-dimensional Lie algebras

We begin with some definitions and constructions that will allow us to define Kac-Moody Lie algebras. Then we introduce key tools for studying their representations, such as an invariant bilinear form and the generalized Casimir operator. This will enable us to formulate and prove a version of the Weyl character formula. This section follows [3].

2.1 Basic definitions

Definition 2.1. A *Cartan matrix* is a square integer matrix $A = (a_{ij})$ of rank l such that

- $a_{ii} = 2$ for all i ,
- $a_{ij} \leq 0$ for $i \neq j$,
- $a_{ij} = 0$ if and only if $a_{ji} = 0$.

A **realization** of a Cartan matrix A is a triple $(\mathfrak{h}, \Pi, \Pi^\vee)$ where \mathfrak{h} is a complex vector space, $\Pi = \{\alpha_1, \dots, \alpha_n\} \subset \mathfrak{h}^*$ and $\Pi^\vee = \{\alpha_1^\vee, \dots, \alpha_n^\vee\} \subset \mathfrak{h}$ are linearly independent sets such that $\alpha_j(\alpha_i^\vee) = a_{ij}$ for all i, j and $\dim(\mathfrak{h}) = 2n - l$.

Remark 2.2 (Finite-dimensional Cartan matrices). In the finite-dimensional case, the Cartan matrix is invertible and positive definite and $l = n$, so $\dim(\mathfrak{h}) = n$. The set of simple roots $\Pi = \alpha_1, \dots, \alpha_n$ is a basis of the real vector space spanned by the roots $E = \mathbb{R}\Phi \subseteq \mathfrak{h}^*$. Similarly, the set of simple coroots $\Pi^\vee = \alpha_1^\vee, \dots, \alpha_n^\vee$ is a basis of $E^\vee = \mathbb{R}\Phi^\vee \subseteq \mathfrak{h}$.

Denote by Q the root lattice, i.e. the integer span of the simple roots Π . Let $Q^+ = \sum_{i=1}^n \mathbb{Z}_{\geq 0} \alpha_i$ be the positive cone in Q . We write $\beta \geq 0$ if $\beta \in Q^+$ and $\beta > 0$ if $\beta \in Q^+ \setminus \{0\}$. We define a partial order on \mathfrak{h}^* by $\lambda \preceq \mu$ if and only if $\mu - \lambda \geq 0$. The sum of the coefficients of $\beta = \sum_i k_i \alpha_i$ is called the **height** of β and denoted $\text{ht}(\beta) = \sum_i k_i$.

Definition 2.3 (Universal Lie algebra associated to a Cartan matrix). Let $A = (a_{ij})$ be an $n \times n$ -matrix over \mathbb{C} , and let $(\mathfrak{h}, \Pi, \Pi^\vee)$ be a realization of A . We introduce an auxiliary Lie algebra $\tilde{\mathfrak{g}}(A)$ with the generators e_i, f_i ($i = 1, \dots, n$) and \mathfrak{h} , and the following defining relations:

$$\begin{aligned} [e_i, f_j] &= \delta_{ij} \alpha_i^\vee & (i, j = 1, \dots, n), \\ [h, h'] &= 0 & (h, h' \in \mathfrak{h}), \\ [h, e_i] &= \langle \alpha_i, h \rangle e_i, \\ [h, f_i] &= -\langle \alpha_i, h \rangle f_i & (i = 1, \dots, n; h \in \mathfrak{h}). \end{aligned}$$

By the uniqueness of the realization of A it is clear that $\tilde{\mathfrak{g}}(A)$ depends only on A .

Denote by $\tilde{\mathfrak{n}}_+$ (resp. $\tilde{\mathfrak{n}}_-$) the subalgebra of $\tilde{\mathfrak{g}}(A)$ generated by e_1, \dots, e_n (resp. f_1, \dots, f_n).

Theorem 2.4 (Properties of the universal Lie algebra associated to a Cartan matrix). *Let $\tilde{\mathfrak{g}}(A)$ be the Lie algebra associated to a Cartan matrix A with realization $(\mathfrak{h}, \Pi, \Pi^\vee)$.*

- a) $\tilde{\mathfrak{g}}(A) = \tilde{\mathfrak{n}}_- \oplus \mathfrak{h} \oplus \tilde{\mathfrak{n}}_+$ (direct sum of vector spaces).
- b) $\tilde{\mathfrak{n}}_+$ (resp. $\tilde{\mathfrak{n}}_-$) is freely generated by e_1, \dots, e_n (resp. f_1, \dots, f_n).
- c) The map $e_i \mapsto -f_i$, $f_i \mapsto -e_i$ ($i = 1, \dots, n$), $h \mapsto -h$ ($h \in \mathfrak{h}$), can be uniquely extended to an involution $\tilde{\omega}$ of the Lie algebra $\tilde{\mathfrak{g}}(A)$.
- d) With respect to \mathfrak{h} one has the root space decomposition:

$$\tilde{\mathfrak{g}}(A) = \left(\bigoplus_{\substack{\alpha \in Q_+ \\ \alpha \neq 0}} \tilde{\mathfrak{g}}_{-\alpha} \right) \oplus \mathfrak{h} \oplus \left(\bigoplus_{\substack{\alpha \in Q_+ \\ \alpha \neq 0}} \tilde{\mathfrak{g}}_{\alpha} \right),$$

where

$$\tilde{\mathfrak{g}}_{\alpha} = \{x \in \tilde{\mathfrak{g}}(A) \mid [h, x] = \alpha(h)x \text{ for all } h \in \mathfrak{h}\}.$$

Furthermore, $\dim \tilde{\mathfrak{g}}_{\alpha} < \infty$, and $\tilde{\mathfrak{g}}_{\alpha} \subset \tilde{\mathfrak{n}}_{\pm}$ for $\pm\alpha \in Q_+$, $\alpha \neq 0$.

- e) Among the ideals of $\tilde{\mathfrak{g}}(A)$ intersecting \mathfrak{h} trivially, there exists a unique maximal ideal \mathfrak{r} . Furthermore,

$$\mathfrak{r} = (\mathfrak{r} \cap \tilde{\mathfrak{n}}_-) \oplus (\mathfrak{r} \cap \tilde{\mathfrak{n}}_+) \quad (\text{direct sum of ideals}).$$

Proof. Let V be the n -dimensional complex vector space with a basis v_1, \dots, v_n and let λ be a linear function on \mathfrak{h} . We define an action of the generators of $\tilde{\mathfrak{g}}(A)$ on the tensor algebra $T(V)$ over V by

- $\alpha)$ $f_i(a) = v_i \otimes a$ for $a \in T(V)$;
- $\beta)$ $h(1) = \langle \lambda, h \rangle 1$, and inductively on s ,
 $h(v_j \otimes a) = -\langle \alpha_j, h \rangle v_j \otimes a + v_j \otimes h(a)$ for $a \in T^{s-1}(V)$, $j = 1, \dots, n$;
- $\gamma)$ $e_i(1) = 0$, and inductively on s ,
 $e_i(v_j \otimes a) = \delta_{ij} \alpha_i^\vee(a) + v_j \otimes e_i(a)$ for $a \in T^{s-1}(V)$, $j = 1, \dots, n$.

This defines a representation of the Lie algebra $\tilde{\mathfrak{g}}(A)$ on the space $T(V)$. To see that, we have to check all of the relations. Provided one does that, the statements of the theorem quickly follow.

Using the relations it is easy to show by induction on s that a product of s elements from the set $\{e_i, f_i (i = 1, \dots, n); h\}$ lies in $\tilde{\mathfrak{n}}_- + \mathfrak{h} + \tilde{\mathfrak{n}}_+$. Let now $u = n_- + h + n_+ = 0$, where $n_{\pm} \in \tilde{\mathfrak{n}}_{\pm}$ and

$h \in \mathfrak{h}$. Then in the representation $T(V)$ we have

$$u(1) = n_-(1) + \langle \lambda, h \rangle = 0.$$

It follows that $\langle \lambda, h \rangle = 0$ for every $\lambda \in \mathfrak{h}^*$ and hence $h = 0$.

Furthermore, using the map $f_i \mapsto v_i$, we see that the tensor algebra $T(V)$ is an associative enveloping algebra of the Lie algebra $\tilde{\mathfrak{n}}_-$. Since $T(V)$ is a free associative algebra, we conclude that $T(V)$ is automatically the universal enveloping algebra $U(\tilde{\mathfrak{n}}_-)$ of $\tilde{\mathfrak{n}}_-$, the map $n_- \mapsto n_-(1)$ being the canonical embedding $\tilde{\mathfrak{n}}_- \hookrightarrow U(\tilde{\mathfrak{n}}_-)$. Hence $n_- = 0$ and we obtain the triangular decomposition of $\tilde{\mathfrak{g}}(A)$, proving a). Moreover, by the Poincaré–Birkhoff–Witt theorem, $\tilde{\mathfrak{n}}_-$ is freely generated by f_1, \dots, f_n . The statement c) is obvious. Now applying $\tilde{\omega}$ we deduce that $\tilde{\mathfrak{n}}_+$ is freely generated by e_1, \dots, e_n , proving b).

The relations make e_i, f_i weight vectors, $\text{ad } h$ acts diagonally, eigenvectors with distinct eigenvalues are independent. Thus we get the decomposition d). The bound on the weight space dimension comes from the fact that each root space $\tilde{\mathfrak{g}}_\alpha$ is generated by commutators of $\text{ht}(\alpha)$ simple generators. There are at most $n^{\text{ht}(\alpha)}$ such brackets, so $\dim \tilde{\mathfrak{g}}_\alpha \leq n^{\text{ht}(\alpha)}$.

To prove e), note that for any ideal i of $\tilde{\mathfrak{g}}(A)$ one has (by the proposition to follow)

$$i = \bigoplus_{\alpha} (\tilde{\mathfrak{g}}_{\alpha} \cap i).$$

Hence the sum of ideals which intersect \mathfrak{h} trivially, itself intersects \mathfrak{h} trivially, and the sum of all ideals with this property is the unique maximal ideal \mathfrak{r} which intersects \mathfrak{h} trivially. In particular, we obtain that (e) is a direct sum of vector spaces. But, clearly,

$$[f_i, \mathfrak{r} \cap \tilde{\mathfrak{n}}_+] \subset \tilde{\mathfrak{n}}_+.$$

Hence

$$[\tilde{\mathfrak{g}}(A), \mathfrak{r} \cap \tilde{\mathfrak{n}}_+] \subset \mathfrak{r} \cap \tilde{\mathfrak{n}}_+;$$

similarly,

$$[\tilde{\mathfrak{g}}(A), \mathfrak{r} \cap \tilde{\mathfrak{n}}_-] \subset \mathfrak{r} \cap \tilde{\mathfrak{n}}_-.$$

This shows that (e) is a direct sum of ideals. \square

Proposition 2.5. *Let \mathfrak{h} be a commutative Lie algebra, V a diagonalizable \mathfrak{h} -module, i.e.*

$$V = \bigoplus_{\lambda \in \mathfrak{h}^*} V_{\lambda}, \quad V_{\lambda} = \{v \in V \mid h(v) = \lambda(h)v \text{ for all } h \in \mathfrak{h}\}. \quad (4)$$

Then any submodule U of V is graded with respect to the gradation (4).

Proof. Any $v \in V$ can be written in the form

$$v = \sum_{j=1}^m v_j, \quad v_j \in V_{\lambda_j},$$

and there exists $h \in \mathfrak{h}$ such that $\lambda_j(h)$ ($j = 1, \dots, m$) are distinct. We have for $v \in U$:

$$h^k(v) = \sum_{j=1}^m \lambda_j(h)^k v_j \in U \quad (k = 0, 1, \dots, m-1).$$

This is a system of linear equations with a nondegenerate matrix. Hence all v_j lie in U . This also shows that the sum in (4) is direct because if $v = 0$ then all $h^k(v) = 0$ and we can apply the invertible matrix to conclude that all $v_j = 0$. \square

Given a complex $n \times n$ -matrix A , we can now define the main object of our study: the Lie algebra $\mathfrak{g}(A)$.

Definition 2.6 (Kac-Moody algebra). Let $(\mathfrak{h}, \Pi, \Pi^\vee)$ be a realization of A and let $\tilde{\mathfrak{g}}(A)$ be the Lie algebra on generators e_i, f_i ($i = 1, \dots, n$) and \mathfrak{h} , and the defining relations (1.2.1). By Theorem 2.4 the natural map $\mathfrak{h} \rightarrow \tilde{\mathfrak{g}}(A)$ is an embedding. Let \mathfrak{r} be the maximal ideal in $\tilde{\mathfrak{g}}(A)$ which intersects \mathfrak{h} trivially. We set:

$$\mathfrak{g}(A) = \tilde{\mathfrak{g}}(A)/\mathfrak{r}.$$

The matrix A is called the **Cartan matrix** of the Lie algebra $\mathfrak{g}(A)$, and n is called the **rank** of $\mathfrak{g}(A)$. The Lie algebra $\mathfrak{g}(A)$ whose Cartan matrix is a generalized Cartan matrix is called a **Kac-Moody algebra**.

Remark 2.7 (Interpreting the maximal ideal which meets the Cartan subalgebra trivially). It is true but not obvious that the maximal ideal \mathfrak{r} which meets the Cartan subalgebra \mathfrak{h} trivially is generated by the so-called **Serre relations**:

For $i \neq j$,

$$(\text{ad } e_i)^{1-a_{ij}}(e_j) = 0, \quad (\text{ad } f_i)^{1-a_{ij}}(f_j) = 0,$$

where a_{ij} are entries of the Cartan matrix. These relations are what turn the free Lie algebras $\tilde{\mathfrak{n}}_\pm$ into the correct nilpotent subalgebras.

The Serre relations can be understood from the representation theory of \mathfrak{sl}_2 . Inside $\mathfrak{g}(A)$, consider the subalgebra

$$\mathfrak{sl}_2(i) = \langle e_i, f_i, h_i \rangle \cong \mathfrak{sl}_2.$$

For fixed i , every other generator e_j or f_j is a weight vector for this copy of \mathfrak{sl}_2 . The Cartan matrix entry $a_{ij} = \langle \alpha_j, \alpha_i^\vee \rangle$ tells you the weight of e_j relative to $\mathfrak{sl}_2(i)$. Thus, e_j generates an $\mathfrak{sl}_2(i)$ -submodule.

But in an \mathfrak{sl}_2 -representation, if a vector has weight m , then applying e_i more than m times kills it. This is exactly what the Serre relation enforces:

$$(\operatorname{ad} e_i)^{1-a_{ij}}(e_j) = 0$$

is the statement that e_j generates an $\mathfrak{sl}_2(i)$ -submodule of dimension $(-a_{ij}) + 1$.

The quadruple $(\mathfrak{g}(A), \mathfrak{h}, \Pi, \Pi^\vee)$ is called the **quadruple associated to the matrix A** . Two quadruples $(\mathfrak{g}(A), \mathfrak{h}, \Pi, \Pi^\vee)$ and $(\mathfrak{g}(A_1), \mathfrak{h}_1, \Pi_1, \Pi_1^\vee)$ are called **isomorphic** if there exists a Lie algebra isomorphism $\varphi : \mathfrak{g}(A) \rightarrow \mathfrak{g}(A_1)$ such that $\varphi(\mathfrak{h}) = \mathfrak{h}_1$, $\varphi(\Pi^\vee) = \Pi_1^\vee$ and $\varphi^*(\Pi_1) = \Pi$.

We keep the same notation for the images of e_i, f_i, \mathfrak{h} in $\mathfrak{g}(A)$. The subalgebra \mathfrak{h} of $\mathfrak{g}(A)$ is called the **Cartan subalgebra**. The elements e_i, f_i ($i = 1, \dots, n$) are called the **Chevalley generators**. In fact, they generate the **derived subalgebra** $\mathfrak{g}'(A) = [\mathfrak{g}(A), \mathfrak{g}(A)]$. Furthermore,

$$\mathfrak{g}(A) = \mathfrak{g}'(A) + \mathfrak{h}$$

with $\mathfrak{g}(A) = \mathfrak{g}'(A)$ if and only if $\det A \neq 0$.

We set $\mathfrak{h}' = \sum_{i=1}^n \mathbb{C}\alpha_i^\vee$. Then $\mathfrak{g}'(A) \cap \mathfrak{h} = \mathfrak{h}'$; $\mathfrak{g}'(A) \cap \mathfrak{g}_\alpha = \mathfrak{g}_\alpha$ if $\alpha \neq 0$.

It follows from (1.2.2) that we have the following **root space decomposition** with respect to \mathfrak{h} :

$$\mathfrak{g}(A) = \bigoplus_{\alpha \in Q} \mathfrak{g}_\alpha. \quad (5)$$

Here,

$$\mathfrak{g}_\alpha = \{x \in \mathfrak{g}(A) \mid [h, x] = \alpha(h)x \text{ for all } h \in \mathfrak{h}\}$$

is the **root space** attached to α . Note that $\mathfrak{g}_0 = \mathfrak{h}$. The number $\operatorname{mult} \alpha := \dim \mathfrak{g}_\alpha$ is called the **multiplicity** of α . Note that

$$\operatorname{mult} \alpha \leq n^{|\operatorname{ht} \alpha|} \quad (6)$$

by Theorem 2.4 d).

An element $\alpha \in Q$ is called a **root** if $\alpha \neq 0$ and $\operatorname{mult} \alpha \neq 0$. A root $\alpha > 0$ (resp. $\alpha < 0$) is called **positive** (resp. **negative**). It follows from the root space decomposition that every root is either positive or negative. Denote by Δ , Δ_+ and Δ_- the sets of all roots, positive and negative roots respectively. Then

$$\Delta = \Delta_+ \dot{\cup} \Delta_- \quad (\text{a disjoint union}).$$

Sometimes we will write $\Delta(A), Q(A), \dots$ in order to emphasize the dependence on A .

Let \mathfrak{n}_+ (resp. \mathfrak{n}_-) denote the subalgebra of $\mathfrak{g}(A)$ generated by e_1, \dots, e_n (resp. f_1, \dots, f_n). By Theorem 2.4 e) and the definition of $\mathfrak{g}(A)$, we have the *triangular decomposition*

$$\mathfrak{g}(A) = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+ \quad (\text{direct sum of vector spaces}).$$

because the ideal \mathfrak{r} is graded and hence respects the triangular decomposition of $\tilde{\mathfrak{g}}(A)$.

Note that $\mathfrak{g}_\alpha \subset \mathfrak{n}_+$ if $\alpha > 0$ and $\mathfrak{g}_\alpha \subset \mathfrak{n}_-$ if $\alpha < 0$. In other words, for $\alpha > 0$ (resp. $\alpha < 0$), \mathfrak{g}_α is the linear span of the elements of the form

$$[\dots [[e_{i_1}, e_{i_2}], e_{i_3}] \dots e_{i_s}] \quad (\text{resp. } [\dots [[f_{i_1}, f_{i_2}], f_{i_3}] \dots f_{i_s}]),$$

such that $\alpha_{i_1} + \dots + \alpha_{i_s} = \alpha$ (resp. $= -\alpha$). It follows immediately that

$$\mathfrak{g}_{\alpha_i} = \mathbb{C}e_i, \quad \mathfrak{g}_{-\alpha_i} = \mathbb{C}f_i, \quad \mathfrak{g}_{s\alpha_i} = 0 \quad \text{if } |s| > 1. \quad (7)$$

because for example the $2\alpha_i$ root space is spanned by $[e_i, e_i] = 0$.

Since every root is either positive or negative, (7) implies the following important fact:

Lemma 2.8. *If $\beta \in \Delta_+ \setminus \{\alpha_i\}$, then*

$$(\beta + \mathbb{Z}\alpha_i) \cap \Delta \subset \Delta_+$$

Proof. Suppose $\beta \neq \alpha_i$ is positive, but $\beta - q\alpha_i$ is negative for some q . Then the string must pass through $\beta - r\alpha_i = 0$ or $-\alpha_i$ at some step $r \leq q$. But the only multiples of α_i that are roots are $\pm\alpha_i$. So the only way to hit a negative root is if the string actually reaches $-\alpha_i$. If $\beta - r\alpha_i = -\alpha_i$, then $\beta = (r-1)\alpha_i$. But β is a root and not equal to α_i . The only possible multiples of α_i that are roots are $\pm\alpha_i$. So $\beta = (r-1)\alpha_i$ is impossible unless $\beta = \alpha_i$. \square

Remark 2.9 (Finiteness of root strings). *Using the interpretation of the Serre relations from the representation theory of \mathfrak{sl}_2 , one sees that these root strings are in fact finite. Look at the subalgebra $\mathfrak{sl}_2(i) = \langle e_i, f_i, h_i \rangle$. For each root β , the root space \mathfrak{g}_β is a weight space of $\mathfrak{sl}_2(i)$ with weight $\langle \beta, \alpha_i^\vee \rangle$. Acting with $\text{ad } e_i$ and $\text{ad } f_i$ generates a finite-dimensional \mathfrak{sl}_2 -module, because the Serre relations*

$$(\text{ad } e_i)^{1-a_{ij}}(e_j) = 0, \quad (\text{ad } f_i)^{1-a_{ij}}(f_j) = 0$$

kill sufficiently long strings.

Thus the α_i -string through β has finite length.

Lemma 2.10. *Let $a \in \mathfrak{n}_+$ be such that $[a, f_i] = 0$ for all $i = 1, \dots, n$. Then $a = 0$. Similarly, for $a \in \mathfrak{n}_-$, if $[a, e_i] = 0$ for all $i = 1, \dots, n$, then $a = 0$.*

Proof. Let $a \in \mathfrak{n}_+$ be such that $[a, \mathfrak{g}_{-1}(1)] = 0$. Then it is easy to see that

$$\sum_{i,j \geq 0} (\text{ad } \mathfrak{g}_1(1))^i (\text{ad } \mathfrak{h})^j a$$

is a subspace of $\mathfrak{n}_+ \subset \mathfrak{g}(A)$, which is invariant with respect to $\text{ad } \mathfrak{g}_1(1)$, $\text{ad } \mathfrak{h}$ and $\text{ad } \mathfrak{g}_{-1}(1)$ (the condition on a is used only in the last case). Hence if $a \neq 0$, we obtain a nonzero ideal in $\mathfrak{g}(A)$ which intersects \mathfrak{h} trivially. This contradicts the definition of $\mathfrak{g}(A)$. \square

Remark 2.11. Sometimes it is useful to consider the Lie algebra $\mathfrak{g}'(A)$ instead of $\mathfrak{g}(A)$. Let us give a more direct construction of $\mathfrak{g}'(A)$. Denote by $\tilde{\mathfrak{g}}'(A)$ the Lie algebra on generators e_i, f_i, α_i^\vee ($i = 1, \dots, n$) and defining relations

$$[e_i, f_j] = \delta_{ij} \alpha_i^\vee, \quad [\alpha_i^\vee, \alpha_j^\vee] = 0, \quad [\alpha_i^\vee, e_j] = a_{ij} e_j, \quad [\alpha_i^\vee, f_j] = -a_{ij} f_j.$$

Let Q be a free abelian group on generators $\alpha_1, \dots, \alpha_n$. Introduce a Q -gradation

$$\tilde{\mathfrak{g}}'(A) = \bigoplus_{\alpha} \tilde{\mathfrak{g}}'_\alpha$$

setting

$$\deg e_i = \alpha_i = -\deg f_i, \quad \deg \alpha_i^\vee = 0.$$

There exists a unique maximal Q -graded ideal $\mathfrak{r} \subset \tilde{\mathfrak{g}}'(A)$ intersecting $\tilde{\mathfrak{g}}'_0 (= \sum_i \mathbb{C} \alpha_i^\vee)$ trivially. Then

$$\mathfrak{g}'(A) = \tilde{\mathfrak{g}}'(A) / \mathfrak{r}.$$

Note that this definition works for an infinite n as well.

Remark 2.12. In the presentation of $\mathfrak{g}(A)$, you start with a Cartan subalgebra \mathfrak{h} large enough so that you can realize the simple roots α_i and simple coroots α_i^\vee as linear maps. In general,

$$\dim \mathfrak{h} = 2n - \text{rank}(A).$$

So if A is singular (affine/indefinite type), then \mathfrak{h} strictly contains $\mathfrak{h}' = \text{span}\{\alpha_i^\vee\}$. In the presentation of $\mathfrak{g}'(A)$, you only build in the “minimal Cartan” generated by the simple coroots:

$$\mathfrak{h}' = \sum_i \mathbb{C} \alpha_i^\vee.$$

If A is invertible (finite type): then $\mathfrak{h} = \mathfrak{h}'$, so $\mathfrak{g}(A) = \mathfrak{g}'(A)$. If A is singular (e.g. affine type): then \mathfrak{h} has more dimensions than \mathfrak{h}' , and these extra directions give rise to central elements and sometimes a degree derivation. In this case, $\mathfrak{g}(A) = \mathfrak{g}'(A) \oplus (\mathfrak{h}/\mathfrak{h}')$.

Proposition 2.13 (Center of a Kac-Moody algebra). The center of the Lie algebra $\mathfrak{g}(A)$ or $\mathfrak{g}'(A)$ is equal to

$$\mathfrak{c} := \{h \in \mathfrak{h} \mid \langle \alpha_i, h \rangle = 0 \text{ for all } i = 1, \dots, n\}.$$

Furthermore, $\dim \mathfrak{c} = n - \ell$.

Proof. Let c lie in the center; write $c = \sum_i c_i$ with respect to the principal gradation. Then $[c, \mathfrak{g}_{-1}(1)] = 0$ implies $[c_i, \mathfrak{g}_{-1}(1)] = 0$ and hence, by Lemma 1.5, $c_i = 0$ for $i > 0$. Similarly, $c_i = 0$ for $i < 0$. Hence $c \in \mathfrak{h}$ and $[c, e_i] = \langle \alpha_i, c \rangle e_i = 0$ implies that $\langle \alpha_i, c \rangle = 0$ ($i = 1, \dots, n$). Conversely, if $c \in \mathfrak{h}$ and the latter condition holds, c commutes with all Chevalley generators and, therefore, lies in the center. The simple roots $\alpha_1, \dots, \alpha_n$ are linear functionals on \mathfrak{h} . They span a subspace of \mathfrak{h}^* of dimension $\ell = \text{rank}(A)$. Therefore, the common kernel

$$\{h \in \mathfrak{h} : \alpha_i(h) = 0 \ \forall i\}$$

has dimension $n - \ell$.

Finally, $\mathfrak{c} \subset \mathfrak{h}'$ since in the contrary case, then there would exist some extra element $c \in \mathfrak{h} \setminus \mathfrak{h}'$ that is annihilated by every simple root. That would mean the simple roots $\{\alpha_i\}$ vanish on a larger subspace of \mathfrak{h} than expected, so they would not be linearly independent in \mathfrak{h}^* , contradicting the axioms of a realization. \square

2.2 Invariant bilinear form

Definition 2.14. A Cartan matrix A is called *symmetrizable* if there exists a diagonal matrix $D = \text{diag}(d_1, \dots, d_n)$ with positive entries d_i such that DA is symmetric.

Let A be a symmetrizable matrix with a fixed decomposition and let $(\mathfrak{h}, \Pi, \Pi^\vee)$ be a realization of A . Fix a complementary subspace \mathfrak{h}'' to $\mathfrak{h}' = \sum \mathbb{C}\alpha_i^\vee$ in \mathfrak{h} , and define a symmetric bilinear \mathbb{C} -valued form $(\cdot \mid \cdot)$ on \mathfrak{h} by the following two equations:

$$(\alpha_i^\vee \mid h) = \langle \alpha_i, h \rangle \epsilon_i, \quad \text{for } h \in \mathfrak{h}, \ i = 1, \dots, n \quad (8)$$

$$(h' \mid h'') = 0, \quad \text{for } h', h'' \in \mathfrak{h}'' \quad (9)$$

Since $\alpha_1^\vee, \dots, \alpha_n^\vee$ are linearly independent and since

$$(\alpha_i^\vee \mid \alpha_j^\vee) = b_{ij} \epsilon_i \epsilon_j, \quad (i, j = 1, \dots, n) \quad (10)$$

there is no ambiguity in the definition of $(\cdot \mid \cdot)$.

Lemma 2.15. Let $\mathfrak{g}(A)$ be the Kac-Moody algebra associated to a symmetrizable matrix A . Then the following holds:

1. The kernel of the restriction of the bilinear form $(\cdot \mid \cdot)$ to \mathfrak{h}' coincides with \mathfrak{c} .
2. The bilinear form $(\cdot \mid \cdot)$ is nondegenerate on \mathfrak{h} .

Proof. (a) follows from Proposition 1.6. If now for all $h \in \mathfrak{h}$ we have

$$0 = \left(\sum_{i=1}^n c_i \alpha_i^\vee \mid h \right) = \left\langle \sum_{i=1}^n c_i \epsilon_i \alpha_i, h \right\rangle,$$

then

$$\sum_{i=1}^n c_i \epsilon_i \alpha_i = 0$$

and hence $c_i = 0, i = 1, \dots, n$, proving (b). \square

Remark 2.16. *If A is symmetric, you are in the “simply-laced” world (types A, D, E or untwisted affine). If A is symmetrizable but not symmetric, you are in the “multiply-laced” world (types B, C, F, G or twisted affine).*

Every symmetrizable GCM gives rise to a Kac-Moody algebra that has:

- *A symmetric, invariant bilinear form on \mathfrak{g} .*
- *A Weyl group that acts as isometries with respect to this form.*
- *A root system with well-behaved reflection geometry.*

If A were not symmetrizable, these structures might not exist at all (the theory gets pathological).

In fact, there is this tension between being symmetrizable and insisting that the Cartan matrix have 2s on the diagonal. If you drop the condition that the diagonal entries are 2, then you could avoid worrying about symmetrizability.

Since the bilinear form $(\cdot \mid \cdot)$ is nondegenerate, we have an isomorphism

$$\nu : \mathfrak{h} \rightarrow \mathfrak{h}^*$$

defined by

$$\langle \nu(h), h_1 \rangle = (h \mid h_1), \quad h, h_1 \in \mathfrak{h},$$

and the induced bilinear form $(\cdot \mid \cdot)$ on \mathfrak{h}^* .

We had defined the bilinear form on \mathfrak{h} by $(\alpha_i^\vee \mid h) = \langle \alpha_i, h \rangle \epsilon_i$ for $h \in \mathfrak{h}$, so rewriting gives

$$\nu(\alpha_i^\vee) = \epsilon_i \alpha_i, \quad i = 1, \dots, n. \quad (11)$$

Now observe that $(\alpha_i \mid \alpha_j) := (\nu^{-1}(\alpha_i) \mid \nu^{-1}(\alpha_j))$. We know that $\nu(\alpha_i^\vee) = \epsilon_i \alpha_i$, so $\nu^{-1}(\alpha_i) = \frac{1}{\epsilon_i} \alpha_i^\vee$.

Therefore,

$$\begin{aligned}
(\alpha_i \mid \alpha_j) &= \left(\frac{1}{\epsilon_i} \alpha_i^\vee \mid \frac{1}{\epsilon_j} \alpha_j^\vee \right) \\
&= \frac{1}{\epsilon_i \epsilon_j} (\alpha_i^\vee \mid \alpha_j^\vee) \\
&= \frac{1}{\epsilon_i \epsilon_j} (b_{ij} \epsilon_i \epsilon_j) \\
&= b_{ij}.
\end{aligned}$$

where we invoke equation (8) in the last line.

Theorem 2.17 (Invariant bilinear form on a symmetrizable Kac-Moody algebra). *Let $\mathfrak{g}(A)$ be a symmetrizable Lie algebra. Since A is symmetrizable, fix a symmetrization $A = DB$ as above. Then there exists a nondegenerate symmetric bilinear \mathbb{C} -valued form $(\cdot \mid \cdot)$ on $\mathfrak{g}(A)$ such that:*

a) $(\cdot \mid \cdot)$ is invariant, i.e.

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

b) $(\cdot \mid \cdot)|_{\mathfrak{h}}$ is nondegenerate and defined by

$$\begin{aligned}
(\alpha_i^\vee \mid h) &= \langle \alpha_i, h \rangle \epsilon_i, \quad \text{for } h \in \mathfrak{h}, i = 1, \dots, n \\
(h' \mid h'') &= 0, \quad \text{for } h', h'' \in \mathfrak{h}''
\end{aligned}$$

c) $(\mathfrak{g}_\alpha \mid \mathfrak{g}_\beta) = 0$ if $\alpha + \beta \neq 0$.

d) $(\cdot \mid \cdot)|_{\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}}$ is nondegenerate for $\alpha \neq 0$, and hence \mathfrak{g}_α and $\mathfrak{g}_{-\alpha}$ are nondegenerately paired by $(\cdot \mid \cdot)$.

e) $[x, y] = (x \mid y) \nu^{-1}(\alpha)$ for $x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_{-\alpha}, \alpha \in \Delta$.

Proof. Consider the principal \mathbb{Z} -gradation

$$\mathfrak{g}(A) = \bigoplus_{j \in \mathbb{Z}} \mathfrak{g}_j, \quad \mathfrak{g}(N) = \bigoplus_{j=-N}^N \mathfrak{g}_j \quad \text{for } N = 0, 1, \dots$$

where \mathfrak{g}_j is the subspace of roots of height j .

Define a bilinear symmetric form $(\cdot \mid \cdot)$ on $\mathfrak{g}(0) = \mathfrak{h}$ by (2.1.2) and (2.1.3) and extend it to $\mathfrak{g}(1)$ by

$$(e_i \mid f_j) = \delta_{ij} \epsilon_i \quad (i, j = 1, \dots, n), \quad (12)$$

$$(\mathfrak{g}_0 \mid \mathfrak{g}_{\pm 1}) = 0, \quad (\mathfrak{g}_{\pm 1} \mid \mathfrak{g}_{\pm 1}) = 0. \quad (13)$$

Then the form $(\cdot \mid \cdot)$ on $\mathfrak{g}(1)$ satisfies invariance as long as both $[x, y]$ and $[y, z]$ lie in $\mathfrak{g}(1)$. Indeed every bracket between e_i, f_j and h remains in $\mathfrak{g}(1)$ and the only nontrivial check is

$$([e_i, f_j] \mid h) = (e_i \mid [f_j, h]) \quad \text{for } h \in \mathfrak{h},$$

or, equivalently,

$$\delta_{ij}(\alpha_i^\vee \mid h) = \delta_{ij}\epsilon_i\langle \alpha_j, h \rangle,$$

which is indeed true.

Now we extend $(\cdot \mid \cdot)$ to a bilinear form on the space $\mathfrak{g}(N)$ by induction on $N \geq 1$ so that $(\mathfrak{g}_i \mid \mathfrak{g}_j) = 0$ if $|i|, |j| \leq N$ and $i + j \neq 0$, and also condition a) is satisfied as long as both $[x, y]$ and $[y, z]$ lie in $\mathfrak{g}(N)$. Suppose this is already defined on $\mathfrak{g}(N - 1)$; then we have only to define $(x \mid y)$ for $x \in \mathfrak{g}_{\pm N}, y \in \mathfrak{g}_{\mp N}$. We can write $y = \sum_i [u_i, v_i]$, where u_i and v_i are homogeneous elements of nonzero degree which lie in $\mathfrak{g}(N - 1)$. Then $[x, u_i] \in \mathfrak{g}(N - 1)$ and we set

$$(x \mid y) = \sum_i ([x, u_i] \mid v_i).$$

To show that this is well defined, we prove that if $i, j, s, t \in \mathbb{Z}$ are such that $|i| + |j| = |s| + |t| = N$, $i + j + s + t = 0$, $|i|, |j|, |s|, |t| < N$ and $x_i \in \mathfrak{g}_i, x_j \in \mathfrak{g}_j, x_s \in \mathfrak{g}_s, x_t \in \mathfrak{g}_t$, then we have (on $\mathfrak{g}(N - 1)$)

$$([x_i, x_j], x_s \mid x_t) = (x_i \mid [[x_j, x_s], x_t]).$$

To explain why this is what we need to check, fix $x \in \mathfrak{g}_{\pm N}$. Define a bilinear map

$$\beta_x : \bigoplus_{p+q=\mp N} \mathfrak{g}_p \otimes \mathfrak{g}_q \longrightarrow \mathbb{C}, \quad \beta_x(u \otimes v) := ([x, u] \mid v),$$

where u, v are homogeneous, $|p|, |q| < N$.

There is a bracket map

$$L : \bigoplus_{p+q=\mp N} \mathfrak{g}_p \otimes \mathfrak{g}_q \longrightarrow \mathfrak{g}_{\mp N}, \quad L(u \otimes v) = [u, v].$$

Our definition says $(x \mid \cdot)$ on $\mathfrak{g}_{\mp N}$ should be the linear functional that satisfies

$$(x \mid [u, v]) = \beta_x(u \otimes v).$$

This is well defined iff β_x vanishes on $\ker L$; i.e. β_x depends only on $[u, v]$, not on the particular decomposition. In particular, a choice of decomposition $y = \sum [u_i, v_i]$ corresponds to choosing a preimage of y in V . If \tilde{y}_1, \tilde{y}_2 are two different preimages of the same y , then their difference lies in the kernel of L : $\tilde{y}_1 - \tilde{y}_2 \in \ker L$. So we need to show that β_x vanishes on $\ker L$. The kernel of L is generated by elements of two types:

- $u \otimes v + v \otimes u$ (skew-symmetry)
- $[u, v] \otimes w + [v, w] \otimes u + [w, u] \otimes v$ (Jacobi)

We get skew symmetry from the invariance of the form on $\mathfrak{g}(N-1)$:

$$\begin{aligned}\beta_x(u \otimes v) + \beta_x(v \otimes u) &= ([x, u] \mid v) + ([x, v] \mid u) \\ &= (x \mid [u, v]) + (x \mid [v, u]) = 0,\end{aligned}$$

using invariance of the form on $\mathfrak{g}(N-1)$ (true by induction) and $[v, u] = -[u, v]$. So β_x vanishes on $u \otimes v + v \otimes u$.

To check Jacobi, consider homogeneous $x_i \in \mathfrak{g}_i, x_j \in \mathfrak{g}_j, x_s \in \mathfrak{g}_s, x_t \in \mathfrak{g}_t$ with $|i|+|j| = |s|+|t| = N$, $i+j+s+t=0$, and $|i|, |j|, |s|, |t| < N$. The identity quoted in the text,

$$([x_i, x_j], x_s \mid x_t) = (x_i \mid [[x_j, x_s], x_t]),$$

implies that β_{x_i} kills the Jacobi generator:

$$\beta_{x_i}([x_j, x_s] \otimes x_t) + \beta_{x_i}([x_s, x_t] \otimes x_j) + \beta_{x_i}([x_t, x_j] \otimes x_s) = 0.$$

Indeed, if we had the identity, then we would have

$$\begin{aligned}\beta_{x_i}([x_j, x_s] \otimes x_t) &= ([x_i, [x_j, x_s]] \mid x_t) = (x_i \mid [[x_j, x_s], x_t]), \\ \beta_{x_i}([x_s, x_t] \otimes x_j) &= ([x_i, [x_s, x_t]] \mid x_j) = (x_i \mid [[x_s, x_t], x_j]), \\ \beta_{x_i}([x_t, x_j] \otimes x_s) &= ([x_i, [x_t, x_j]] \mid x_s) = (x_i \mid [[x_t, x_j], x_s]).\end{aligned}$$

and adding these three equations gives

$$\beta_{x_i}(J) = (x_i \mid [[x_j, x_s], x_t] + [[x_s, x_t], x_j] + [[x_t, x_j], x_s]) = (x_i \mid 0) = 0,$$

Thus β_x vanishes on the Jacobi-type tensors.

Now we check the identity using the invariance of $(\cdot \mid \cdot)$ on $\mathfrak{g}(N-1)$ and the Lie algebra axioms, we have

$$\begin{aligned}([x_i, x_j], x_s \mid x_t) &= ([x_i, x_j], x_s \mid x_t) - ([x_j, x_s], x_i \mid x_t) \\ &= ([x_i, x_j] \mid [x_s, x_t]) + (x_i \mid [[x_j, x_s], x_t]) \\ &= (x_i \mid [x_s, [x_j, x_t]]) + ([x_j, x_s] \mid [x_i, x_t]) \\ &= (x_i \mid [[x_j, x_s], x_t]),\end{aligned}$$

as required. So the identity holds, and hence β_x vanishes on $\ker L$. This shows that $(x \mid y)$ is well defined.

If now $x = \sum_i [u'_i, v'_i]$, then by definition and by the relation above we have

$$(x \mid y) = \sum_i ([x, u_i] \mid v_i) = \sum_i (u'_i \mid [v'_i, y]).$$

Hence this is independent of the choice of the expressions for x and y .

It is clear from the definition that a) holds on $\mathfrak{g}(N)$ whenever $[x, y]$ and $[y, z]$ lie in $\mathfrak{g}(N)$. Hence we have constructed a bilinear form $(\cdot \mid \cdot)$ on \mathfrak{g} such that a) and b) hold. Its restriction to \mathfrak{h} is nondegenerate by Lemma 2.15 b).

The form $(\cdot \mid \cdot)$ satisfies c) since for $h \in \mathfrak{h}$, $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_\beta$ we have, by the invariance property:

$$0 = ([h, x] \mid y) + (x \mid [h, y]) = (\langle \alpha, h \rangle + \langle \beta, h \rangle)(x \mid y).$$

For $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_{-\alpha}$ where $\alpha \in \Delta$, and $h \in \mathfrak{h}$, we have

$$([x, y] - (x \mid y)\nu^{-1}(\alpha) \mid h) = (x \mid [y, h]) - (x \mid y)\langle \alpha, h \rangle = 0.$$

which combined with b) gives e).

It follows from b), c) and e) that the bilinear form $(\cdot \mid \cdot)$ is symmetric. If d) fails, then by c) the form $(\cdot \mid \cdot)$ is degenerate. Let $\mathfrak{i} = \ker(\cdot \mid \cdot)$ is an ideal by invariance, and by b) we have $\mathfrak{i} \cap \mathfrak{h} = 0$, which contradicts the definition of $\mathfrak{g}(A)$. Therefore d) holds as well. \square

Remark 2.18 (Uniqueness of the invariant bilinear form). *Such a form is uniquely determined once you ask for a) and b).*

First you can deduce that \mathfrak{g}_α and \mathfrak{g}_β are orthogonal unless $\alpha + \beta = 0$. Take $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_\beta$, and $h \in \mathfrak{h}$. Recall the weight-space relation: $[h, x] = \alpha(h)x$, $[h, y] = \beta(h)y$. By invariance,

$$([h, x] \mid y) = (h \mid [x, y]).$$

Since $[h, x] = \alpha(h)x$ and $[x, y] \in \mathfrak{g}_{\alpha+\beta}$, we obtain

$$\alpha(h)(x \mid y) = (h \mid z), \quad z := [x, y] \in \mathfrak{g}_{\alpha+\beta}. \quad (14)$$

Now take $h' \in \mathfrak{h}$. Using invariance again and $[\mathfrak{h}, \mathfrak{h}] = 0$,

$$0 = ([h', h] \mid z) = (h' \mid [h, z]) = (\alpha + \beta)(h)(h' \mid z) \quad \text{for all } h, h' \in \mathfrak{h}.$$

If $\alpha + \beta \neq 0$, choose h with $(\alpha + \beta)(h) \neq 0$; then $(h' \mid z) = 0$ for all $h' \in \mathfrak{h}$, so in particular $(h \mid z) = 0$ for every $h \in \mathfrak{h}$. Returning to (14), we get $\alpha(h)(x \mid y) = 0$ for all $h \in \mathfrak{h}$; choosing h with $\alpha(h) \neq 0$ (since $\alpha \neq 0$ on \mathfrak{g}_α) yields $(x \mid y) = 0$. Hence $(\mathfrak{g}_\alpha \mid \mathfrak{g}_\beta) = 0$ unless $\alpha + \beta = 0$.

Now, take $h \in \mathfrak{h}$. Invariance gives $(h \mid [e_i, f_i]) = ([h, e_i] \mid f_i)$. On the LHS: $[e_i, f_i] = \alpha_i^\vee$, so $(h \mid [e_i, f_i]) = (h \mid \alpha_i^\vee)$. On the RHS: $[h, e_i] = \langle \alpha_i, h \rangle e_i$, so $([h, e_i] \mid f_i) = \langle \alpha_i, h \rangle (e_i \mid f_i)$.

Thus $(h \mid \alpha_i^\vee) = \langle \alpha_i, h \rangle (e_i \mid f_i)$ which determines how e_i and f_i pair against each other. Using invariance and inducting on height, you can determine how any two root vectors pair against each other.

Suppose that $A = (a_{ij})$ is a symmetrizable generalized Cartan matrix. Fix a decomposition

$$A = \text{diag}(\epsilon_1, \dots, \epsilon_n) (b_{ij})_{i,j=1}^n$$

where ϵ_i are **positive** rational numbers and (b_{ij}) is a symmetric rational matrix.

Lemma 2.19. *Such a decomposition always exists.*

Proof. This is equivalent to a system of homogeneous linear equations and inequalities over \mathbb{Q} with unknowns ϵ_i^{-1} and b_{ij} :

$$\epsilon_i^{-1} \neq 0, \quad \text{diag}(\epsilon_1^{-1}, \dots, \epsilon_n^{-1}) A = (b_{ij}), \quad b_{ij} = b_{ji}.$$

By definition, it has a solution over \mathbb{C} . Hence, it has a solution over \mathbb{Q} . We can assume that A is indecomposable, meaning that A is not a direct sum of two smaller Cartan matrices. Then for any $1 < j \leq n$ there exists a sequence

$$1 = i_1 < i_2 < \dots < i_{k-1} < i_k = j$$

such that $a_{i_s, i_{s+1}} < 0$. We have:

$$a_{i_s, i_{s+1}} \epsilon_{i_{s+1}} = a_{i_{s+1}, i_s} \epsilon_{i_s} \quad (s = 1, \dots, k-1).$$

Hence $\epsilon_j \epsilon_1 > 0$ for all j because $a_{i_s, i_{s+1}}$ and a_{i_{s+1}, i_s} have the same sign. \square

Remark 2.20 (Dynkin diagram interpretation). *Given a Cartan matrix, put an edge between two indices if $a_{ij} \neq 0$. Then A is indecomposable iff the graph is connected. Then it is obvious every index j has a path to 1, and the above argument shows that all ϵ_j have the same sign as ϵ_1 . We can then scale all ϵ_j by a constant to make them all positive.*

From this we also deduce that if A is indecomposable, then the matrix $\text{diag}(\epsilon_1, \dots, \epsilon_n)$ is uniquely determined up to a constant factor.

Fix a nondegenerate bilinear symmetric form $(\cdot \mid \cdot)$ associated to the decomposition above as defined above. Recall that

$$(\alpha_i \mid \alpha_j) = b_{ij} = \frac{a_{ij}}{\epsilon_i}.$$

so $i = j$ gives $(\alpha_i | \alpha_i) = 2\epsilon_i^{-1} > 0$. If $i \neq j$, then $a_{ij} \leq 0$ and hence $(\alpha_i | \alpha_j) \leq 0$. By definition of ν (the identification $\mathfrak{h} \rightarrow \mathfrak{h}^*$ using the form), we had $\nu(\alpha_i^\vee) = \epsilon_i \alpha_i$. Applying ν^{-1} , $\alpha_i^\vee = \frac{1}{\epsilon_i} \nu^{-1}(\alpha_i)$. But we also know $\epsilon_i = (\alpha_i | \alpha_i)/2$.

We deduce that:

$$\begin{aligned} (\alpha_i | \alpha_i) &> 0 \quad (i = 1, \dots, n), \\ (\alpha_i | \alpha_j) &\leq 0 \quad (i \neq j), \\ \alpha_i^\vee &= \frac{2}{(\alpha_i | \alpha_i)} \nu^{-1}(\alpha_i). \end{aligned}$$

Hence we obtain the usual expression for the generalized Cartan matrix:

$$A = \left(\frac{2(\alpha_i | \alpha_j)}{(\alpha_i | \alpha_i)} \right)_{i,j=1}^n.$$

We extend the bilinear form $(\cdot | \cdot)$ from \mathfrak{h} to an invariant symmetric bilinear form on the entire Kac–Moody algebra $\mathfrak{g}(A)$. By Theorem 2.17 such a form exists and satisfies all the properties stated there. It is an exercise to show that such a form is unique. The bilinear form $(\cdot | \cdot)$ on the Kac–Moody algebra $\mathfrak{g}(A)$ provided by Theorem 2.17 and satisfying the above is called a **standard invariant form**.

2.3 Generalized Casimir operator

Let $\mathfrak{g}(A)$ be a Lie algebra associated to a matrix A , \mathfrak{h} the Cartan subalgebra, $\mathfrak{g} = \bigoplus_{\alpha} \mathfrak{g}_{\alpha}$ the root space decomposition with respect to \mathfrak{h} . A $\mathfrak{g}(A)$ -module (resp. $\mathfrak{g}'(A)$ -module) V is called **restricted** if for every $v \in V$, we have $\mathfrak{g}_{\alpha}(v) = 0$ for all but a finite number of positive roots α .

It is clear that every submodule or quotient of a restricted module is restricted, and that the direct sum or tensor product of a finite number of restricted modules is also restricted. Examples of restricted modules will be constructed later (see Exercise 2.9 and Chapter 9).

Assume now that A is symmetrizable and that $(\cdot | \cdot)$ is a bilinear form provided by Theorem 2.17.

Given a restricted $\mathfrak{g}(A)$ -module V , we introduce a linear operator Ω on the vector space V , called the (generalized) **Casimir operator**, as follows.

First, introduce a linear function $\rho \in \mathfrak{h}^*$ by equations

$$\langle \rho, \alpha_i^\vee \rangle = \frac{1}{2} a_{ii} \quad (i = 1, \dots, n).$$

If $\det A = 0$, this does not define ρ uniquely, and we pick any solution. It follows from (2.1.5) and (2.1.6) that

$$(\rho | \alpha_i) = \frac{1}{2} (\alpha_i | \alpha_i), \quad (i = 1, \dots, n). \quad (15)$$

Further, for each positive root α we choose a basis $\{e_\alpha^{(i)}\}$ of the space \mathfrak{g}_α , and let $\{e_{-\alpha}^{(i)}\}$ be the dual basis of $\mathfrak{g}_{-\alpha}$. Note that the root spaces need not be 1-dimensional. We define an operator Ω_0 on V by

$$\Omega_0 = 2 \sum_{\alpha \in \Delta_+} \sum_i e_{-\alpha}^{(i)} e_\alpha^{(i)}.$$

One could easily check that this is independent of the choice of bases. Since for each $v \in V$, only a finite number of summands $e_{-\alpha}^{(i)} e_\alpha^{(i)}(v)$ are nonzero, Ω_0 is well defined on V .

Let u_1, u_2, \dots and u^1, u^2, \dots be dual bases of \mathfrak{h} . The generalized Casimir operator is defined by

$$\Omega = 2\nu^{-1}(\rho) + \sum_i u^i u_i + \Omega_0.$$

Remark 2.21. *The generalized Casimir operator Ω was introduced by Kac. The idea of its definition is borrowed from physics. We take the usual definition of the Casimir operator:*

$$\Omega = \sum_{\alpha > 0} \sum_i \left(e_{-\alpha}^{(i)} e_\alpha^{(i)} + e_\alpha^{(i)} e_{-\alpha}^{(i)} \right) + \sum_i u_i u^i,$$

we rewrite it by using commutation relations:

$$\Omega = \sum_{\alpha > 0} \nu^{-1}(\alpha) + 2 \sum_{\alpha > 0} \sum_i e_{-\alpha}^{(i)} e_\alpha^{(i)} + \sum_i u_i u^i,$$

and then replace the first summand, which makes no sense, by a finite quantity $2\nu^{-1}(\rho)$.

We record the following simple formula:

$$\sum_i \langle \lambda, u^i \rangle \langle \mu, u_i \rangle = (\lambda \mid \mu), \quad (16)$$

which is clear from

$$\lambda = \sum_i \langle \lambda, u^i \rangle \nu(u_i) = \sum_i \langle \lambda, u_i \rangle \nu(u^i).$$

We make one more simple computation. For $x \in \mathfrak{g}_\alpha$ one has

$$\left[\sum_i u^i u_i, x \right] = \sum_i \langle \alpha, u^i \rangle x u_i + \sum_i u^i \langle \alpha, u_i \rangle x = \sum_i \langle \alpha, u^i \rangle \langle \alpha, u_i \rangle x + x \left(\sum_i u^i \langle \alpha, u_i \rangle + u_i \langle \alpha, u^i \rangle \right).$$

Hence, we have

$$\left[\sum_i u^i u_i, x \right] = x((\alpha \mid \alpha) + 2\nu^{-1}(\alpha)), \quad x \in \mathfrak{g}_\alpha. \quad (17)$$

Lemma 2.22 (Important calculation). *If $\alpha, \beta \in \Delta$ and $z \in \mathfrak{g}_{\beta-\alpha}$, then in $\mathfrak{g}(A) \otimes \mathfrak{g}(A)$ we have*

$$\sum_i e_{-\alpha}^{(i)} \otimes [z, e_{\alpha}^{(i)}] = \sum_i [e_{-\beta}^{(i)}, z] \otimes e_{\beta}^{(i)}.$$

Proof. We define the bilinear form $(\cdot \mid \cdot)$ on $\mathfrak{g}(A) \otimes \mathfrak{g}(A)$ by

$$(x \otimes y \mid x_1 \otimes y_1) = (x \mid x_1)(y \mid y_1).$$

Pick $e \in \mathfrak{g}_{\alpha}$ and $f \in \mathfrak{g}_{-\beta}$. It suffices to check that pairing both sides of the above identity with $e \otimes f$ gives the same result. We have:

$$\begin{aligned} \sum_s (e_{-\alpha}^{(s)} \otimes [z, e_{\alpha}^{(s)}] \mid e \otimes f) &= \sum_s (e_{-\alpha}^{(s)} \mid e) ([z, e_{\alpha}^{(s)}] \mid f) \\ &= \sum_s (e_{-\alpha}^{(s)} \mid e) (e_{\alpha}^{(s)} \mid [f, z]) \\ &= (e \mid [f, z]), \end{aligned}$$

by invariance and the dual basis property. Similarly,

$$\begin{aligned} \sum_s ([e_{-\beta}^{(s)}, z] \otimes e_{\beta}^{(s)} \mid e \otimes f) &= \sum_s (e_{-\beta}^{(s)} \mid [z, e]) (e_{\beta}^{(s)} \mid f) \\ &= ([z, e] \mid f). \end{aligned}$$

Applying again invariance gives the result. \square

Corollary 2.23. *In the notation of the lemma, we have*

$$\sum_i [e_{-\alpha}^{(i)}, [z, e_{\alpha}^{(i)}]] = - \sum_i [[z, e_{-\beta}^{(i)}], e_{\beta}^{(i)}] \quad \text{in } \mathfrak{g}(A),$$

and

$$\sum_i e_{-\alpha}^{(i)} [z, e_{\alpha}^{(i)}] = - \sum_i [z, e_{-\beta}^{(i)}] e_{\beta}^{(i)} \quad \text{in } U(\mathfrak{g}(A)).$$

Proof. Apply to the lemma the linear maps from $\mathfrak{g}(A) \otimes \mathfrak{g}(A)$ to $\mathfrak{g}(A)$ and to $U(\mathfrak{g}(A))$, defined by

$$x \otimes y \mapsto [x, y], \quad x \otimes y \mapsto xy,$$

respectively. \square

Consider the root space decomposition of $U(\mathfrak{g}(A))$ with respect to \mathfrak{h} :

$$U(\mathfrak{g}(A)) = \bigoplus_{\beta \in Q} U_{\beta}, \quad U_{\beta} = \{x \in U(\mathfrak{g}(A)) \mid [h, x] = \langle \beta, h \rangle x \text{ for all } h \in \mathfrak{h}\}.$$

Recall that $\mathfrak{g}'(A) = [\mathfrak{g}(A), \mathfrak{g}(A)]$ is the derived algebra of $\mathfrak{g}(A)$. Put $U'_\beta = U(\mathfrak{g}'(A)) \cap U_\beta$, so that

$$U(\mathfrak{g}'(A)) = \bigoplus_{\beta} U'_\beta.$$

Theorem 2.24 (). *Let $\mathfrak{g}(A)$ be a symmetrizable Lie algebra.*

a) *If V is a restricted $\mathfrak{g}'(A)$ -module and $u \in U'_\alpha$, then*

$$[\Omega_0, u] = -u \left(2(\rho|\alpha) + (\alpha|\alpha) + 2\nu^{-1}(\alpha) \right).$$

b) *If V is a restricted $\mathfrak{g}(A)$ -module, then Ω commutes with the action of $\mathfrak{g}(A)$ on V .*

Proof. Part b) follows immediately from a) and the earlier formula 17. Note that we only have to check the formula for the derived algebra $\mathfrak{g}'(A)$ since

$$\mathfrak{g}(A) = \mathfrak{g}'(A) \oplus \mathfrak{h}'',$$

where \mathfrak{h}'' is the orthogonal complement of the subspace generated by the coroots α_i^\vee in \mathfrak{h} , and in particular $[\mathfrak{h}'', \mathfrak{g}(A)] = 0$, i.e. \mathfrak{h}'' is central. The following computation shows that (a) implies (b): Let $u \in U'_\alpha$. Then

$$\begin{aligned} [\Omega_0, u] &= -u(2(\rho|\alpha) + (\alpha|\alpha) + 2\nu^{-1}(\alpha)) \\ \left[\sum_i u^i u_i, u \right] &= u((\alpha|\alpha) + 2\nu^{-1}(\alpha)) \\ [2\nu^{-1}(\rho), u] &= 2(\rho|\alpha)u \quad \text{for } h \in \mathfrak{h} \end{aligned}$$

If a) holds for $u \in U'_\alpha$ and $u_1 \in U'_\beta$, then it holds for $uu_1 \in U'_{\alpha+\beta}$. Indeed:

$$\begin{aligned} [\Omega_0, uu_1] &= [\Omega_0, u]u_1 + u[\Omega_0, u_1] \\ &= -u(2(\rho|\alpha) + (\alpha|\alpha) + 2\nu^{-1}(\alpha))u_1 \\ &\quad - uu_1(2(\rho|\beta) + (\beta|\beta) + 2\nu^{-1}(\beta)) \\ &= -uu_1(2(\rho|\alpha) + (\alpha|\alpha) + 2\nu^{-1}(\alpha)) \\ &\quad + 2(\alpha|\beta)uu_1 + 2(\rho|\beta)uu_1 + (\beta|\beta)uu_1 + 2\nu^{-1}(\beta)uu_1 \\ &= -uu_1(2(\rho|\alpha + \beta) + (\alpha + \beta|\alpha + \beta) + 2\nu^{-1}(\alpha + \beta)). \end{aligned}$$

Hence, since $e_{\alpha_i}, e_{-\alpha_i}$ ($i = 1, \dots, n$) generate $\mathfrak{g}'(A)$, it suffices to check the formula for $u = e_{\alpha_i}$

or $u = e_{-\alpha_i}$. Applying 2.22 to $z = e_{\alpha_i}$, we have:

$$\begin{aligned}
[\Omega_0, e_{\alpha_i}] &= 2 \sum_{\alpha \in \Delta_+} \sum_s ([e_{-\alpha}^{(s)}, e_{\alpha_i}] e_{\alpha}^{(s)} + e_{-\alpha}^{(s)} [e_{\alpha}^{(s)}, e_{\alpha_i}]) \\
&= 2[e_{-\alpha_i}, e_{\alpha_i}] e_{\alpha_i} + 2 \sum_{\alpha \in \Delta_+ \setminus \{\alpha_i\}} \left(\sum_s [e_{-\alpha}^{(s)}, e_{\alpha_i}] e_{\alpha}^{(s)} + \sum_s e_{-\alpha_i + \alpha}^{(s)} [e_{\alpha - \alpha_i}^{(s)}, e_{\alpha_i}] \right) \\
&= -2\nu^{-1}(\alpha_i) e_{\alpha_i} = -2(\alpha_i | \alpha_i) e_{\alpha_i} - 2e_{\alpha_i} \nu^{-1}(\alpha_i).
\end{aligned}$$

where the second equality follows from reindexing the second summand (note that we have 2.8) and the third equality follows from 2.22. Thanks to 15 this is exactly the formula in part a) for $u = e_{\alpha_i}$. Similarly,

$$[\Omega_0, e_{-\alpha_i}] = 2e_{-\alpha_i} [e_{\alpha_i}, e_{-\alpha_i}] = 2e_{-\alpha_i} \nu^{-1}(\alpha_i),$$

which, by 15, is the same formula for $u = e_{-\alpha_i}$. \square

Corollary 2.25 (Eigenvalue on highest weight modules). *If, under the hypotheses of Theorem 2.24 b), there exists $v \in V$ such that $e_i(v) = 0$ for all $i = 1, \dots, n$, and $h(v) = \langle \Lambda, h \rangle v$ for some $\Lambda \in \mathfrak{h}^*$ and all $h \in \mathfrak{h}$, then*

$$\Omega(v) = (\Lambda + 2\rho | \Lambda) v.$$

If, furthermore, $U(\mathfrak{g}(A))v = V$, then

$$\Omega = (\Lambda + 2\rho | \Lambda) I_V.$$

Proof. Recall that

$$\Omega = 2\nu^{-1}(\rho) + \sum_i u^i u_i + \Omega_0, \quad \Omega_0 = \sum_{\alpha \in \Delta_+} \sum_s e_{-\alpha}^{(s)} e_{\alpha}^{(s)}.$$

Let v be a highest weight vector, i.e. $e_i v = 0$ for all i and $h \cdot v = \langle \Lambda, h \rangle v$ for all $h \in \mathfrak{h}$.

For every $\alpha \in \Delta_+$ we have $e_{\alpha}^{(s)} v = 0$, hence each term in Ω_0 kills v , so $\Omega_0 v = 0$.

Since $u_i, u^i \in \mathfrak{h}$ and v has weight Λ ,

$$u_i v = \langle \Lambda, u_i \rangle v, \quad u^i(u_i v) = \langle \Lambda, u_i \rangle \langle \Lambda, u^i \rangle v.$$

Summing over i and using the duality of $\{u_i\}$ and $\{u^i\}$ for $(\cdot | \cdot)$, we obtain (invoking 16):

$$\left(\sum_i u^i u_i \right) v = \sum_i \langle \Lambda, u_i \rangle \langle \Lambda, u^i \rangle v = (\Lambda | \Lambda) v.$$

Again by the weight action,

$$(2\nu^{-1}(\rho)) \cdot v = 2\langle \Lambda, \nu^{-1}(\rho) \rangle v = 2(\Lambda | \rho) v.$$

This proves the first formula. If moreover $V = U(\mathfrak{g}(A))v$, then by Theorem 2.24 b), Ω commutes with the $\mathfrak{g}(A)$ -action on V . Thus for any $x \in U(\mathfrak{g}(A))$,

$$\Omega(xv) = x \Omega v = (\Lambda + 2\rho|\Lambda) xv.$$

Since such xv span V , it follows that Ω acts as the scalar $(\Lambda + 2\rho|\Lambda)$ on all of V , i.e.

$$\Omega = (\Lambda + 2\rho|\Lambda) I_V.$$

as desired. \square

We introduce another central element defined for \mathfrak{g} -modules V which have a certain finiteness property. **What is the relation to the generalized Casimir operator?**

Proposition 2.26. *Let \mathfrak{g} be a Lie algebra with an invariant non-degenerate bilinear form $(\cdot | \cdot)$, let $\{x_i\}$ and $\{y_i\}$ be dual bases (i.e. $(x_i | y_j) = \delta_{ij}$), and let V be a \mathfrak{g} -module such that for every pair of elements $u, v \in V$, $x_i(u) = 0$ or $y_i(v) = 0$ for all but a finite number of i . Then the operator*

$$\Omega_2 := \sum_i x_i \otimes y_i$$

is defined on $V \otimes V$ and commutes with the action of \mathfrak{g} .

Proof. We have to show that for every $z \in \mathfrak{g}$,

$$\sum_i ([z, x_i] \otimes y_i + x_i \otimes [z, y_i]) = 0.$$

Write

$$[z, x_i] = \sum_j \alpha_{ij} x_j, \quad [z, y_i] = \sum_j \beta_{ij} y_j.$$

Taking the inner product of the first equation with y_j and of the second with x_j , we obtain

$$\alpha_{ij} = ([z, x_i] | y_j), \quad \beta_{ji} = ([z, y_i] | x_j).$$

Using invariance of $(\cdot | \cdot)$, we deduce

$$\alpha_{ij} = (z | [x_i, y_j]), \quad \beta_{ji} = (z | [y_i, x_j]).$$

Hence

$$\alpha_{ij} = -\beta_{ij}.$$

This proves the proposition. \square

To conclude this section, we define the compact form of a Kac-Moody algebra. Let A be an $n \times n$ matrix over \mathbb{R} . Let $(\mathfrak{h}_{\mathbb{R}}, \Pi, \Pi^{\vee})$ be a realization of the matrix A over \mathbb{R} , i.e. $\mathfrak{h}_{\mathbb{R}}$ is a real vector space of dimension $2n - \text{rank}(A)$, so that $(\mathfrak{h} = \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{h}_{\mathbb{R}}, \Pi, \Pi^{\vee})$ is the realization of A over \mathbb{C} .

Definition 2.27. We define the **compact form** $\mathfrak{t}(A)$ of $\mathfrak{g}(A)$ as follows. Let ω_0 be the antilinear automorphism of $\mathfrak{g}(A)$ determined by

$$\omega_0(e_i) = -f_i, \quad \omega_0(f_i) = -e_i \quad (i = 1, \dots, n), \quad \omega_0(h) = -h \quad (h \in \mathfrak{h}_{\mathbb{R}}).$$

Then $\mathfrak{t}(A)$ is defined as the fixed point set of ω_0 . This is a real Lie algebra whose complexification is $\mathfrak{g}(A)$. In the finite-dimensional case, this definition coincides with the usual compact real form.

Remark 2.28. Let $(\cdot | \cdot)$ be a symmetric bilinear form on $\mathfrak{g}(A)$. Define a Hermitian form on $\mathfrak{g}(A)$ by

$$(x|y)_0 := -(\omega_0(x) | y).$$

Then $(\cdot | \cdot)_0$ is nondegenerate on each \mathfrak{g}_{α} , $\alpha \in \Delta \cup \{0\}$, and satisfies $(\mathfrak{g}_{\alpha} | \mathfrak{g}_{\beta})_0 = 0$ if $\alpha \neq \beta$. Moreover, the operators $\text{ad } u$ and $-\text{ad } \omega_0(u)$ are adjoint to each other with respect to $(\cdot | \cdot)_0$. In particular, the restriction of $(\cdot | \cdot)$ to $\mathfrak{t}(A)$ is a nondegenerate invariant \mathbb{R} -bilinear form.

This is the generalization of the familiar fact that a complex semisimple Lie algebra has a compact real form, and the Killing form restricted there is negative definite.

3 Integrable modules

We begin this section by recalling the representation theory of $\mathfrak{sl}_2(\mathbb{C})$. Let

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

be the standard basis of $\mathfrak{sl}_2(\mathbb{C})$. Then

$$[e, f] = h, \quad [h, e] = 2e, \quad [h, f] = -2f.$$

By an easy induction on k we deduce the following relations in the universal enveloping algebra of $\mathfrak{sl}_2(\mathbb{C})$:

$$\begin{aligned} [h, f^k] &= -2k f^k, \\ [h, e^k] &= 2k e^k \\ [e, f^k] &= -k(k-1) f^{k-1} + k f^{k-1} h. \end{aligned}$$

Lemma 3.1 (Classification of irreducible $\mathfrak{sl}_2(\mathbb{C})$ -modules).

(a) Let V be an $\mathfrak{sl}_2(\mathbb{C})$ -module and let $v \in V$ be such that

$$h(v) = \lambda v \quad \text{for some } \lambda \in \mathbb{C}.$$

Set $v_j = (j!)^{-1} f^j(v)$. Then

$$h(v_j) = (\lambda - 2j)v_j. \quad (3.2.3)$$

If, in addition, $e(v) = 0$, then

$$e(v_j) = (\lambda - j + 1)v_{j-1}. \quad (3.2.4)$$

(b) For each integer $k \geq 0$ there exists a unique, up to isomorphism, irreducible $(k+1)$ -dimensional $\mathfrak{sl}_2(\mathbb{C})$ -module. In some \mathbb{C} -basis $\{v_j\}_{j=0}^k$ of the space of this module, the action of $\mathfrak{sl}_2(\mathbb{C})$ looks as follows:

$$h(v_j) = (k - 2j)v_j, \quad f(v_j) = (j+1)v_{j+1}, \quad e(v_j) = (k+1-j)v_{j-1}.$$

Here $j = 0, \dots, k$ and we assume that $v_{k+1} = 0 = v_{-1}$.

Where does one invoke the fact that we quotiented out by the maximal ideal intersecting \mathfrak{h} trivially, when deriving the Serre relations?

How to interpret this number $\langle \lambda, \alpha_i^\vee \rangle$? I feel like I see it a lot.

4 Affine Lie algebras

In order to develop the theory of root systems of Kac-Moody algebras we need to know some properties of generalized Cartan matrices. It is convenient to work in a slightly more general situation. Let $A = (a_{ij})$ which satisfies the following three properties:

(m1) A is indecomposable;

(m2) $a_{ij} \leq 0$ for $i \neq j$;

(m3) $a_{ij} = 0$ implies $a_{ji} = 0$.

Kac proves the following classification theorem in [3].

Theorem 4.1 (Classification of matrices satisfying (m1), (m2), and (m3)). *Let A be a real $n \times n$ matrix satisfying (m1), (m2), and (m3). Then one and only one of the following three possibilities holds for both A and ${}^t A$:*

(Fin) $\det A \neq 0$; there exists $u > 0$ such that $Au > 0$; $Av \geq 0$ implies $v > 0$ or $v = 0$;

(Aff) $\text{corank } A = 1$; there exists $u > 0$ such that $Au = 0$; $Av \geq 0$ implies $Av = 0$;

(Ind) there exists $u > 0$ such that $Au < 0$; $Av \geq 0, v \geq 0$ imply $v = 0$.

Remark 4.2. The matrices of type (Fin) are precisely the Cartan matrices of finite-dimensional semisimple Lie algebras. The matrices of type (Aff) are precisely the Cartan matrices of affine Kac-Moody algebras. The matrices of type (Ind) are called indefinite type.

Borcherds made the remark that nobody has really found a use for the indefinite type matrices.

We continue with another perspective of affine Lie algebras.

Central extensions of $\mathfrak{g}[z, z^{-1}]$

Let \mathfrak{g} be a simple Lie algebra over \mathbb{C} , e.g. $\mathfrak{sl}_n, \mathfrak{so}_n$ for $n \geq 5$, \mathfrak{sp}_{2n} for $n \geq 4$, exceptional ones, etc.

Fix an invariant inner product $\langle \cdot, \cdot \rangle$ on \mathfrak{g} satisfying

$$\langle x, [y, z] \rangle = \langle [x, y], z \rangle;$$

since \mathfrak{g} is simple, this inner product is unique up to scalar.

Now consider the *loop algebra* $\mathfrak{g}[z, z^{-1}]$. It's called loop because it's "equal" to $T_e LG$, where

$$LG = \{\text{analytic maps } S^1 \rightarrow G\}$$

(not precisely; we should really take some completion of $\mathfrak{g}[z, z^{-1}]$, but it's a dense subspace inside of $T_e LG$, which is enough for our purposes). This algebra $\mathfrak{g}[z, z^{-1}]$ is graded: the n th graded component is $\mathfrak{g}z^n$, hence the grading is given by operator $z\partial_z$. Hence

$$\mathfrak{g}[z, z^{-1}] = \bigoplus_{n \in \mathbb{Z}} \mathfrak{g}z^n$$

is a graded Lie algebra, with

$$[xz^n, yz^m] = [x, y]z^{n+m}.$$

It carries an invariant inner product defined by

$$\langle xz^n, yz^m \rangle = \delta_{n, -m} \langle x, y \rangle$$

where $\langle x, y \rangle$ is the invariant inner product on \mathfrak{g} . There is also the equivalent residue formulation of this inner product:

$$\langle x \otimes f(z), y \otimes g(z) \rangle = \langle x, y \rangle \text{Res}_{z=0} \left(f(z)g(z)/zdz \right).$$

which is more suitable for generalizations. The choice of inner product on \mathfrak{g} also gives rise to a 2-cocycle on $\mathfrak{g}[z, z^{-1}]$:

$$\omega(xz^n, yz^m) = n\delta_{n,-m}\langle x, y \rangle.$$

This is a 2-cocycle because it is antisymmetric and satisfies the cocycle condition (Jacobi):

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

Recall that central extensions are classified by $H^2(\mathfrak{g}[z, z^{-1}])$, so we want to compute this cohomology group. It turns out that $H^2(\mathfrak{g}[z, z^{-1}]) \cong \mathbb{C}$, and the class of ω is a generator. In particular, any central extension of $\mathfrak{g}[z, z^{-1}]$ is a multiple of the one defined by ω . We have the Chevalley complex

$$C^\bullet(\mathfrak{g}[z, z^{-1}]) = \Lambda^\bullet(\mathfrak{g}[z, z^{-1}]^*),$$

and it carries an action of $\mathfrak{g}[z, z^{-1}]$ which acts trivially on cohomology. In particular, \mathfrak{g} acts trivially on the cohomology. To see this, observe that the subalgebra $\mathfrak{g} \subset L$ of constant loops acts on $C^\bullet(L)$ by the usual Lie derivative $\mathcal{L}_\xi = \iota_\xi d + d \iota_\xi$ (Cartan's homotopy formula), which is chain-homotopic to 0 via the contraction ι_ξ . Hence \mathfrak{g} acts trivially on cohomology so every class has a \mathfrak{g} -invariant representative (Lemma 6.1).

Thus, to classify $H^2(L)$, we may restrict to \mathfrak{g} -invariant 2-cocycles, modulo coboundaries.

Proposition 4.3 (Second cohomology of loop algebra). *Any cohomology class from $H^2(\mathfrak{g}[z, z^{-1}])$ is uniquely represented by a \mathfrak{g} -invariant cocycle.*

Any \mathfrak{g} -invariant 2-cocycle is of the form

$$\sum_{m \neq n} \gamma_{m,n} \omega_{m,n}, \quad \gamma_{m,n} = -\gamma_{n,m} \in \mathbb{C}, \quad \omega_{m,n}(xz^k, yz^\ell) = \delta_{m=k, n=\ell} \langle x, y \rangle,$$

Proof. A 1-cochain is a linear functional $\varphi \in L^*$. A coboundary is

$$(d\varphi)(xz^k, yz^\ell) = \varphi([xz^k, yz^\ell]) = \varphi([x, y]z^{k+\ell}).$$

If φ is \mathfrak{g} -invariant, then its restriction to each $\mathfrak{g}_r \simeq \mathfrak{g}$ is a \mathfrak{g} -invariant linear functional on the adjoint module \mathfrak{g} , which must be zero.

This is because there are no nonzero invariants in \mathfrak{g}^* for a simple \mathfrak{g} . To see this, recall that a linear functional $\lambda \in \mathfrak{g}^*$ is \mathfrak{g} -invariant iff

$$\lambda([x, y]) = 0 \quad \forall x, y \in \mathfrak{g}.$$

But $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$ for simple \mathfrak{g} , so $\lambda = 0$.

Therefore $L^{*\mathfrak{g}} = 0$ so invariant 2-cocycles have no nontrivial invariant coboundaries.

Now suppose you have two invariant 2-cocycles ω, ω' that represent the same cohomology class. That means $\omega - \omega' = d\varphi$ for some $\varphi \in C^1(L)$. Then their difference $d\varphi$ is also \mathfrak{g} -invariant, so $d\varphi$ is an invariant coboundary, which must be zero by the above argument. Hence $\omega = \omega'$, i.e. once we restrict to \mathfrak{g} -invariant representatives, the class determines the representative uniquely. This proves the first statement.

Decompose L as a \mathfrak{g} -module:

$$L = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_k, \quad \mathfrak{g}_k \simeq \mathfrak{g} \text{ (adjoint)}.$$

Let $\omega \in C^2(L)$ be \mathfrak{g} -invariant. In particular, being a 2-cocycle it is alternating. Fix degrees k, ℓ . The bilinear map

$$\omega_{k,\ell}: \mathfrak{g}_k \times \mathfrak{g}_\ell \rightarrow \mathbb{C}, \quad (xz^k, yz^\ell) \mapsto \omega(xz^k, yz^\ell)$$

is \mathfrak{g} -invariant for the diagonal adjoint action, i.e.

$$\omega_{k,\ell}([\xi, x]z^k, yz^\ell) + \omega_{k,\ell}(xz^k, [\xi, y]z^\ell) = 0 \quad \forall \xi \in \mathfrak{g}.$$

Since \mathfrak{g} is simple, the space of \mathfrak{g} -invariant bilinear forms on $\mathfrak{g} \times \mathfrak{g}$ is one-dimensional and spanned by the fixed invariant form $\langle \cdot, \cdot \rangle$. Therefore there exist scalars $c_{k,\ell} \in \mathbb{C}$ with

$$\omega(xz^k, yz^\ell) = c_{k,\ell} \langle x, y \rangle \quad (k, \ell \in \mathbb{Z})$$

Skew-symmetry of ω gives $c_{k,\ell} = -c_{\ell,k}$. So every \mathfrak{g} -invariant 2-cochain is encoded by an antisymmetric function $c: \mathbb{Z}^2 \rightarrow \mathbb{C}$. \square

This still consists of infinitely many parameters. We will whittle it down using the cocycle condition.

Proposition 4.4. *We have*

$$\gamma_{n,m+p} + \gamma_{m,p+n} + \gamma_{p,n+m} = 0.$$

Proof. We have

$$\begin{aligned} 0 &= d\omega(x_1 z^n, x_2 z^m, x_3 z^p) \\ &= \omega([x_1, x_2] z^{n+m}, x_3 z^p) - \omega([x_1, x_3] z^{n+p}, x_2 z^m) + \omega([x_2, x_3] z^{m+p}, x_1 z^n) \\ &= \gamma_{m+p,n} \langle [x_1, x_2], x_3 \rangle - \gamma_{n+p,m} \langle [x_1, x_3], x_2 \rangle + \gamma_{n+m,p} \langle [x_2, x_3], x_1 \rangle. \end{aligned}$$

Since $\langle \cdot, \cdot \rangle$ is totally antisymmetric, all of the $\langle \cdot, \cdot \rangle$ equal the same constant up to sign (note that

$$\langle [x_1, x_3], x_2 \rangle = \langle x_1, [x_3, x_2] \rangle = -\langle x_1, [x_2, x_3] \rangle = -\langle [x_1, x_2], x_3 \rangle),$$

hence the above expression is equal to some constant times

$$\gamma_{p,n+m} + \gamma_{m,p+n} + \gamma_{n,m+p},$$

and choosing the constant to be nonzero we have that this is zero. \square

Corollary 4.5. $\gamma_{n,-n} = n\gamma_{1,-1} = n\gamma$ and $\gamma_{m,n} = 0$ otherwise. In particular,

$$\dim H^2(\mathfrak{g}[z, z^{-1}]) = 1.$$

Proof. We have $\gamma_{n,s-n} + \gamma_{m,s-m} = \gamma_{n+m,s-n-m} \implies \gamma_{0,s} = 0$ for all s . By induction,

$$-(s-n)\gamma_{1,s-1} = \gamma_{n,s-n} = n \cdot \gamma_{1,s-1} \implies (s-n)\gamma_{1,s-1} = -n\gamma_{1,s-1} \implies s \cdot \gamma_{1,s-1} = 0,$$

hence $\gamma_{1,s-1} = 0$ for all $s \neq 0$. This implies the result. \square

Definition 4.6 (Derived affine Kac-Moody Lie algebra). Any central extension of $\mathfrak{g}[z, z^{-1}]$ has the form

$$0 \rightarrow \mathbb{C}c \rightarrow \widehat{\mathfrak{g}}' \rightarrow \mathfrak{g}[z, z^{-1}] \rightarrow 0, \quad [xz^n, yz^m]_{\widehat{\mathfrak{g}}'} = [x, y]_{\mathfrak{g}} z^{n+m} + n \cdot \delta_{n+m=0} \cdot \langle x, y \rangle \cdot c.$$

Equivalently (terrible abuse of notation since they mean apply the bracket coefficient-wise in the Laurent series),

$$[x(z), y(z)]_{\widehat{\mathfrak{g}}'} = [x(z), y(z)]_{\mathfrak{g}} + \text{Res}_{z=0} \langle dx(z), y(z) \rangle \cdot c.$$

Therefore, we define $\widehat{\mathfrak{g}}'$ associated to a simple (finite-dimensional) Lie algebra \mathfrak{g} to be the unique (up to rescaling) nontrivial central extension of the loop space $\mathfrak{g}[z, z^{-1}]$.

Remark 4.7. We will prove that this is the derived subalgebra of an affine Kac-Moody algebra $\widehat{\mathfrak{g}}$ which is obtained by adjoining a derivation d to $\widehat{\mathfrak{g}}'$.

Example 4.8 (Affine $\widehat{\mathfrak{sl}}_2$). Let $\mathfrak{g} = \mathfrak{sl}_2$. We may consider $\widehat{\mathfrak{sl}}_2'$ as a bigraded Lie algebra, with one grading coming from $z\partial_z$, and the other grading coming from $\text{ad } h$.

Then $\widehat{\mathfrak{sl}}_2'$ is generated by h, c, e, f, fz, ez^{-1} . To see this observe that

$$[fz, e] = hz, \quad [hz, e] = 2ez$$

and so we have generated ez , even though it looks like fz only moves down the \mathfrak{h} -grading. In particular, I'm thinking of z as a left translation and the coefficient $\in \mathfrak{g}$ as vertical translation.

In particular we have a triangular decomposition

$$\widehat{\mathfrak{sl}}_2' = \widehat{\mathfrak{n}}_- \oplus \widehat{\mathfrak{h}} \oplus \widehat{\mathfrak{n}}_+$$

where

$$\begin{aligned} \widehat{\mathfrak{h}} &= \mathbb{C}h \oplus \mathbb{C}c \\ \widehat{\mathfrak{n}}_+ &= \mathbb{C}e \oplus z\mathfrak{sl}_2[z] \\ \widehat{\mathfrak{n}}_- &= \mathbb{C}f \oplus z^{-1}\mathfrak{sl}_2[z^{-1}]. \end{aligned}$$

Definition 4.9 (Affine Lie algebra). The *affine Lie algebra* $\widehat{\mathfrak{g}}$ associated to \mathfrak{g} is obtained by adjoining a derivation d , defined by

$$[d, x \otimes z^n] = n x \otimes z^n, \quad [d, c] = 0.$$

Thus,

$$\widehat{\mathfrak{g}} = L\mathfrak{g} \oplus \mathbb{C}c \oplus \mathbb{C}d.$$

and

$$0 \rightarrow \widehat{\mathfrak{g}} \rightarrow \widetilde{\mathfrak{g}} \rightarrow \mathbb{C}d \rightarrow 0.$$

5 Sept 17 meeting

Questions

- A. I've been thinking about the hierarchy of objects: $L\mathfrak{g} \rightarrow$ central extensions \rightarrow full affine Kac–Moody with d . Each step seems to be motivated by making the structure better behaved for representation theory.
- a) The loop algebra $L\mathfrak{g}$ is already a perfectly good infinite-dimensional Lie algebra, realized geometrically as a dense subalgebra of the Lie algebra of the loop group LG . Why isn't it sufficient to study its representation theory directly? In what sense are central extensions 'forced upon us' if we want a rich representation theory?
 - b) I see that the affine Kac–Moody algebra arises as the universal central extension of the loop algebra. Is there a structural reason why the extension must be central and abelian, rather than by some nonabelian ideal? In other words, what uniquely characterizes the central extension as the 'correct' enlargement of $L\mathfrak{g}$ for representation theory?
 - c) Once we accept the central extension, we typically adjoin the degree derivation d , defined by $[d, x \otimes z^n] = n x \otimes z^n$. Should I think of d as a geometric object, namely the vector field $z \frac{d}{dz}$ on the circle? Is the introduction of d mathematically necessary (e.g. for a nondegenerate invariant bilinear form, or for grading representations), or is it mainly motivated by geometric/physical considerations?
- B. I know that algebraically the invariant form defines the 2-cocycle, hence the central extension and the line bundle on Gr_G . But is there a more geometric explanation of why the bilinear form is the key ingredient? For instance, can one see directly from the geometry of G -bundles or determinant line bundles on curves that the form (\cdot, \cdot) is exactly what controls the Picard group of the affine Grassmannian? For example, what exactly does the level of a line bundle correspond to geometrically?

- C. I understand that the level (the eigenvalue of the central element) is a key invariant of representations of affine Kac-Moody algebras. But where exactly does it start to matter? The line bundle on Gr_G has a level, and this level determines which representation we get. But is there a more algebraic way to see why the level is important? For instance, if we just look at highest weight modules for $\hat{\mathfrak{g}}$, how does the level affect their structure?

Answers

Professor Borchers tends to give short answers. First one sees from the combinatorics of Dynkin diagrams that the central extension of the loop algebra is forced upon us.

He had a lot to say about modular forms and theta functions, in particular the denominator of affine Kac-Moody algebras associated to \mathfrak{g} is a theta function, which is something similar to a modular form. In fact, he said "theta functions are nothing but modular forms." I asked about theta divisors which can be thought of as square roots of the canonical bundle, and he made a funny joke. He said "Ah yes, God introduced square roots because there's a lot of sign errors floating around, 2^n many of them." Then he let me borrow his copy of a book and referred me to Mumford's three volumes on theta functions and abelian varieties.

6 Appendix A: Lie algebra cohomology

Lemma 6.1 (Invariant representative via averaging). *Let K be a compact Lie group acting on a cochain complex (C^\bullet, d) by cochain maps (so $k \cdot d = d(k \cdot)$ for all $k \in K$). Suppose the induced K -action on cohomology $H^\bullet(C^\bullet)$ is trivial. Then for every cohomology class $[c] \in H^p(C^\bullet)$ there exists a K -invariant cocycle $c^{\text{av}} \in C^p$ with $dc^{\text{av}} = 0$ and $[c^{\text{av}}] = [c]$. Equivalently, the inclusion $i : (C^\bullet)^K \hookrightarrow C^\bullet$ induces a surjection $H^\bullet((C^\bullet)^K) \twoheadrightarrow H^\bullet(C^\bullet)$.*

Proof. Let μ be the Haar probability measure on K . For $c \in C^p$ define the averaging operator

$$P(c) := \int_K k \cdot c \, d\mu(k) \in C^p.$$

This Bochner integral is well-defined because C^p is a (Hausdorff) complex vector space and $k \mapsto k \cdot c$ is continuous.

(1) P is a cochain map and projects onto invariants. For any $c \in C^p$,

$$dP(c) = \int_K d(k \cdot c) \, d\mu(k) = \int_K k \cdot (dc) \, d\mu(k) = P(dc),$$

so $dP = Pd$. Moreover, for any $k_0 \in K$,

$$k_0 \cdot P(c) = \int_K (k_0 k) \cdot c \, d\mu(k) = \int_K k \cdot c \, d\mu(k) = P(c),$$

using left-invariance of μ . Hence $P(c) \in (C^p)^K$. If c is already K -invariant then $P(c) = c$, so P is a projection onto $(C^\bullet)^K$.

(2) P preserves cocycles. If $dc = 0$ then $dP(c) = P(dc) = 0$, so $c^{\text{av}} := P(c)$ is again a cocycle.

(3) P preserves cohomology classes when the K -action on cohomology is trivial. By assumption, $[k \cdot c] = [c]$ in $H^p(C^\bullet)$ for every $k \in K$. Using linearity of cohomology,

$$[c^{\text{av}}] = \left[\int_K k \cdot c \, d\mu(k) \right] = \int_K [k \cdot c] \, d\mu(k) = \int_K [c] \, d\mu(k) = [c].$$

Thus c^{av} is a K -invariant cocycle representing the same class as c .

Therefore every class in $H^p(C^\bullet)$ has a K -invariant representative, and the inclusion of invariants induces a surjection on cohomology. \square

7 References

1. Bump, D. **Weyl Character Formula notes**, available at <http://sporadic.stanford.edu/Math210C/WCF.pdf>.
2. Cai, M. **Infinite-dimensional Lie algebras**, available at https://merrickcai.com/pdfs_notes/747.pdf.
3. Kac, V. G. **Infinite Dimensional Lie Algebras**. Cambridge University Press, 1990.
4. Pressley, A., and Segal, G. **Loop Groups**. Oxford University Press, 1986.