

# Infinite Dimensional Lie Algebras

Songyu Ye

October 18, 2025

## Abstract

These are notes for a reading course supervised by Prof. Richard Borcherds in the Fall of 2025. The main references are [3] and [5]. There is an accompanying set of notes on loop groups, following [4].

## Contents

<b>1</b>	<b>Semisimple Lie algebras</b>	<b>3</b>
1.1	Preliminaries . . . . .	3
1.2	Highest weight modules . . . . .	6
1.3	Category $\mathcal{O}$ and the Weyl character formula . . . . .	16
<b>2</b>	<b>Introduction to infinite-dimensional Lie algebras</b>	<b>21</b>
2.1	Basic definitions . . . . .	21
2.2	Invariant bilinear form . . . . .	28
2.3	Generalized Casimir operator . . . . .	36
<b>3</b>	<b>Integrable modules</b>	<b>41</b>
3.1	Locally nilpotent operators . . . . .	43
3.2	Weyl group, invariance, and geometry . . . . .	47
<b>4</b>	<b>Classification of affine root systems</b>	<b>51</b>
<b>5</b>	<b>Real and imaginary roots</b>	<b>60</b>
5.1	Key properties . . . . .	60
5.2	Weyl group orbits of imaginary roots . . . . .	63
<b>6</b>	<b>Affine Algebras: the Normalized Invariant Form, the Root System, and the Weyl Group</b>	<b>71</b>
6.1	Distinguished elements . . . . .	71

6.2	The Root System of an Affine Kac-Moody Algebra . . . . .	75
6.3	The Affine Weyl Group . . . . .	78
6.4	The fundamental alcove . . . . .	84
<b>7</b>	<b>Affine Algebras as Central Extensions of Loop Algebras</b>	<b>90</b>
7.1	Formulas for loop algebras . . . . .	94
7.2	The root system . . . . .	97
<b>8</b>	<b>Highest weight modules over Kac-Moody algebras</b>	<b>99</b>
8.1	Category $\mathcal{O}$ . . . . .	100
8.2	Formal characters . . . . .	103
8.3	Shapovalov form . . . . .	105
8.4	Virasoro algebra . . . . .	106
8.5	Integrable highest weight modules . . . . .	108
8.6	Integrable modules by level . . . . .	111
<b>9</b>	<b>Vertex algebras</b>	<b>116</b>
9.1	Fields . . . . .	116
9.2	Composing Fields . . . . .	117
<b>10</b>	<b>Vertex algebras from Lie algebras</b>	<b>123</b>
10.1	The Fock representation . . . . .	125
10.2	Checking locality . . . . .	130
10.3	Vertex algebras from affine Kac-Moody algebras . . . . .	136
10.4	The Virasoro algebra . . . . .	137
10.5	The Segal-Sugawara construction. . . . .	141
<b>11</b>	<b>Associativity and the operator product expansion</b>	<b>142</b>
11.1	Associativity and the operator product expansion . . . . .	143
11.2	The OPE and the Reconstruction Theorem . . . . .	147
<b>12</b>	<b>Appendix 1: Theta functions and Riemann surfaces</b>	<b>155</b>
<b>13</b>	<b>Appendix 2: Lie algebra cohomology</b>	<b>157</b>
<b>14</b>	<b>References</b>	<b>158</b>

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  $\Phi$  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  $\kappa$  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

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

**Remark 1.5 (Pairing with simple roots is evaluating against simple coroot).** *If  $\phi$  is a root, then  $t_\phi$  is not quite the coroot, but they coincide if  $\phi$  has length squared 2 (because then  $(\phi|\phi)/2 = 1$ ). In particular, if  $\alpha$  is a simple root, then  $t_\alpha$  is the coroot  $h_\alpha$  and for any  $\lambda \in \mathfrak{h}^*$  we have*

$$(\lambda|\alpha) = \langle \lambda, \alpha^\vee \rangle$$

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  $\mathfrak{g}_\alpha$  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  $\alpha^\vee$  or  $h_\alpha$  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  $x_\alpha, y_\alpha$  together with  $h_\alpha$  form a standard  $\mathfrak{sl}_2$ -triple, with the eigenvalue of  $x_\alpha$  under  $h_\alpha$  equal to 2.

We will denote by  $\rho$  half the sum of the positive roots. If  $\alpha \in \Phi$  we will denote by  $r_\alpha$  the reflection

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

**Remark 1.6.** Recall that for each root  $\alpha$ , 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  $\mathfrak{g}_\beta$  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  $\text{ad}$  to the subalgebra  $\mathfrak{s}_\alpha$ , 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  $\mathfrak{g}_\beta$  is invariant under  $\mathfrak{h}$ . In particular, under  $h_\alpha$ , vectors in  $\mathfrak{g}_\beta$  have weight  $[h_\alpha, x_\beta] = \beta(h_\alpha) x_\beta$  for  $x_\beta \in \mathfrak{g}_\beta$ .

$x_\alpha$  and  $y_\alpha$  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  $\mathfrak{g}_\beta$ , repeated commutators with  $x_\alpha$  and  $y_\alpha$  move you up and down the  $\alpha$ -string through  $\beta$ :  $\beta, \beta + \alpha, \beta + 2\alpha, \dots, \beta - q\alpha$ .

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  $\beta$  is a root, then so is  $\beta - \langle \beta, \alpha^\vee \rangle \alpha$ , which is exactly the reflection of  $\beta$  across the hyperplane orthogonal to  $\alpha$ . 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  $\alpha$ .

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  $r_\alpha$  for real roots  $\alpha$  still exist, but they do not generate a group which permutes all the roots.

If  $\alpha$  is a simple root, we will also use the notation  $s_\alpha$  for  $r_\alpha$ . We have proved that  $s_\alpha$  maps  $\alpha$  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  $\alpha$ .

An element  $\lambda$  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  $\alpha$ . 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  $\alpha$ ). We call  $\lambda$  **strongly dominant** if  $(\lambda|\alpha^\vee) > 0$ . Thus the special vector  $\rho$  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  $\Phi$  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 1.7 (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 1.8.** *Let  $\lambda$  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  $V_\lambda$ .

**Proposition 1.9.** *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  $\lambda_i$  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:

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

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  $\mathfrak{g}_\alpha$  ( $\alpha \in \Phi^+$ ), and let  $\mathfrak{n}^-$  be the subalgebra generated by the  $\mathfrak{g}_\alpha$  with  $\alpha \in \Phi^-$ . Then clearly we have the **triangular decomposition**

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

**Lemma 1.10.** *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  $\ell$  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  $\lambda$  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  $\lambda$  if it is generated by a highest weight vector  $v \in V_\lambda$ . (Humphreys calls a highest weight module a **standard cyclic module**.)

**Proposition 1.11.** *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,

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

Consider the basis  $\{x_{-\alpha}\}$  ( $\alpha \in \Phi^+$ ) of  $\mathfrak{n}^-$  with  $x_{-\alpha} \in \mathfrak{g}_{-\alpha}$ . Using a fixed order on  $\Phi^+$ , the elements  $\prod_{\alpha \in \Phi^+} x_{-\alpha}^{k_\alpha}$  are a PBW basis of  $U(\mathfrak{n}^-)$ . Since  $x_{-\alpha}$  maps  $V_\mu$  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  $k_\alpha = 0$ ,  $\mu$  is strictly  $\prec \lambda$ , so  $V_\lambda$  is one-dimensional.  $\square$

**Remark 1.12.** *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 1.13** (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  $m$ :

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

Then for  $m + 1$ :

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

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 1.14.** Let  $V$  be a highest weight module with highest weight  $\lambda$ . 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  $V_\lambda$  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 1.15.** Let  $V$  be a highest weight module with highest weight  $\lambda$ . Then  $V$  has a unique maximal proper submodule. Moreover  $V$  has a unique irreducible quotient.

*Proof.* Let  $\Sigma$  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 1.16.** *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  $\lambda$  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  $H_1, \dots, H_\ell$  be a basis of  $\mathfrak{h}$  and  $x_\alpha$  ( $\alpha \in \Phi^+$ ) 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  $k_\alpha$  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  $J_\psi$  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  $Hv = \lambda(H)m$  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  $\lambda$ , 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  $\beta \in \mathfrak{b}$ ,  $J_\psi$  is in the kernel. Thus the map factors uniquely through  $U(\mathfrak{g})/J_\psi = M(\lambda)$ .  $\square$

**Corollary 1.17.** *Let  $\lambda \in \mathfrak{h}^*$ . Up to isomorphism,  $\mathfrak{g}$  has a unique irreducible highest weight module  $L(\lambda)$  with highest weight  $\lambda$ .*

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

**Remark 1.19 (Reducible highest weight modules are not unique).** *Let  $\mathfrak{g}$  have two simple roots  $\alpha_1, \alpha_2$  (e.g.  $\mathfrak{sl}_3$ ). Fix a weight  $\lambda$  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*

$$N_1 = U(\mathfrak{g}) v_{s_1 \cdot \lambda}, \quad N_2 = U(\mathfrak{g}) v_{s_2 \cdot \lambda}$$

*Now the quotients  $M(\lambda)/N_1, M(\lambda)/N_2$  are both highest weight modules of highest weight  $\lambda$ , both reducible, and not isomorphic (their composition series differ). Hence reducible highest weight modules with the same top weight are not unique.*

**Remark 1.20 (Importance of the dot action).** *The BGG theorem tells us: if  $\mu$  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  $\lambda$  and  $\mu$  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  $\lambda, \mu$  have the same central character iff they're in the same dot orbit.*

More precisely, the BGG theorem states:

**Theorem 1.21 (BGG theorem).** *If  $\lambda \in \mathfrak{h}^*$ ,  $\alpha$  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  $s_\alpha$  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 1.22 (Classification of finite-dimensional irreducible modules).** *Let  $V$  be a finite dimensional irreducible module. Then  $V \cong L(\lambda)$  where  $\lambda$  is a dominant integral weight. Conversely, if  $\lambda$  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  $\lambda$  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  $\lambda$  is a dominant integral weight, let  $\alpha$  be a simple positive root. The restriction of  $V$  to the  $\mathfrak{sl}_2$  spanned by  $x_\alpha, x_{-\alpha}$  and  $h_\alpha$  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  $\lambda$  is dominant and integral.

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

**Corollary 1.23 (Weyl).** *For  $V$  an irreducible finite-dimensional  $\mathfrak{g}$ -module, the highest weight  $\lambda$  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  $\{\gamma^j\}$  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 1.24.**  *$c_{\mathfrak{g}}$  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  $\kappa$ . 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  $\Omega$  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  $\kappa$ :

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

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 1.25.** *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  $t_{\mu}$  is that

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

First let us show that

$$t_\mu = \sum_i \mu(h_i) h^i. \quad (1)$$

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  $t_\mu$ . Since the  $h_j$  span  $\mathfrak{h}$  and  $\kappa$  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  $t_\mu$ .  $\square$

**Proposition 1.26.** *Let  $V$  be a highest weight module with highest weight  $\lambda$ . 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  $t_{\alpha}$ . 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  $t_{\alpha}$ . 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  $v_{\lambda}$  is a highest weight vector it is annihilated by  $x_{\alpha}$  and  $y_{\alpha}$  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 1.27.** *A module is in Category  $\mathcal{O}$  if it is  $\mathfrak{h}$ -diagonalizable with finite dimensional weight spaces  $V_\lambda$ , 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 1.28.** *Suppose that  $V$  is a highest weight module with weight  $\mu$  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  $\lambda$ , so

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

as desired.  $\square$

**Definition 1.29.** *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  $e^{\lambda}$  is a formal symbol for  $\lambda \in \mathfrak{h}^*$ .

**Proposition 1.30 (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  $\Phi^+$  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 1.31.** 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 1.32.** 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  $\mu$  a **primitive weight** if  $V_\mu$  contains a primitive vector.

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

**Proposition 1.33.** 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 1.34.** 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  $\mu$  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  $\mu$  that is maximal with respect to  $\preceq$ . Then clearly a primitive vector  $v$  of weight  $\mu$  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 1.35 (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  $\mu$  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  $\mu, \nu$  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 1.36.** *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  $s_{\alpha_i}$  maps the simple root  $\alpha_i$  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 1.37 (Weyl Character Formula).** *Let  $V$  be a finite-dimensional irreducible representation of  $\mathfrak{g}$ . Thus by Theorem 1.22 there is a dominant integral weight  $\lambda$  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 1.30 we may rewrite (2) in the form (note that the  $\rho$  in the exponent comes from dividing by  $\Delta$ )

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

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 1.35. So by (2), we will have terms corresponding to  $\mu$  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  $\mu$  is of the form  $w \circ \lambda$  for some  $w \in W$ .

Therefore we start with  $\mu$  such that  $c_\mu \neq 0$ . By Proposition 1.7, 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  $\lambda$  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  $\alpha$ . So  $|\lambda + \rho|^2 = |\nu + \rho|^2$  implies that  $k_\alpha = 0$  for all  $\alpha$  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

$$\Pi = \{\alpha_1, \dots, \alpha_n\} \subset \mathfrak{h}^*$$

$$\Pi^\vee = \{\alpha_1^\vee, \dots, \alpha_n^\vee\} \subset \mathfrak{h}$$

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  $\Pi$ . 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  $\beta$  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  $A$ .

**Remark 2.4 (What are coroots?).** In the finite dimensional case, coroots come from the  $\mathfrak{sl}_2$ -subalgebras attached to each root: concretely, they are the Cartan elements  $h_\alpha := [e_\alpha, f_\alpha]$  normalized with the Killing form so that root-coroot evaluation pairing are integers.

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 2.5 (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  $\lambda$  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 2.6.** *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 2.7 (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 2.5 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 2.8 (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  $\alpha$ . Note that  $\mathfrak{g}_0 = \mathfrak{h}$ . The number  $\operatorname{mult} \alpha := \dim \mathfrak{g}_\alpha$  is called the **multiplicity** of  $\alpha$ . Note that

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

by Theorem 2.5 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  $\Delta$ ,  $\Delta_+$  and  $\Delta_-$  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 2.5 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.  $\alpha < 0$ ),  $\mathfrak{g}_\alpha$  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 2.9.** *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  $\alpha_i$  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  $\beta$  is a root and not equal to  $\alpha_i$ . The only possible multiples of  $\alpha_i$  that are roots are  $\pm\alpha_i$ . So  $\beta = (r-1)\alpha_i$  is impossible unless  $\beta = \alpha_i$ .  $\square$

**Remark 2.10 (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  $\beta$ , the root space  $\mathfrak{g}_\beta$  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  $\alpha_i$ -string through  $\beta$  has finite length.*

**Lemma 2.11.** *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 2.12.** 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  $e_i, f_i, \alpha_i^\vee$  ( $i = 1, \dots, n$ ) 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 2.13.** 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  $\alpha_i$  and simple coroots  $\alpha_i^\vee$  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 2.14 (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 2.15.** A Cartan matrix  $A$  is called *symmetrizable* if there exists a diagonal matrix

$$D = \text{diag}(\varepsilon_1, \dots, \varepsilon_n)$$

with positive entries  $\varepsilon_i$  and a symmetric matrix  $B = (b_{ij})$  such that  $A = DB$ .

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:

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

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

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

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

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

**Lemma 2.16.** 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

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

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 2.17.** *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

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

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 2.18** (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

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

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  $\mathfrak{g}_\alpha$  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  $\beta_x$  vanishes on  $\ker L$ ; i.e.  $\beta_x$  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  $\beta_x$  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  $\beta_x$  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  $\beta_{x_i}$  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  $\beta_x$  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  $\beta_x$  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 2.16 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 2.19 (Uniqueness of the invariant bilinear form).** *Such a form is uniquely determined once you ask for a) and b).*

*First you can deduce that  $\mathfrak{g}_\alpha$  and  $\mathfrak{g}_\beta$  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  $\mathfrak{g}_\alpha$ ) 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  $\epsilon_i$  are **positive** rational numbers and  $(b_{ij})$  is a symmetric rational matrix.

**Lemma 2.20.** *Such a decomposition always exists.*

*Proof.* This is equivalent to a system of homogeneous linear equations and inequalities over  $\mathbb{Q}$  with unknowns  $\epsilon_i^{-1}$  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 2.21 (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  $\epsilon_j$  have the same sign as  $\epsilon_1$ . We can then scale all  $\epsilon_j$  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  $\nu$  (the identification  $\mathfrak{h} \rightarrow \mathfrak{h}^*$  using the form), we had  $\nu(\alpha_i^\vee) = \epsilon_i \alpha_i$ . Applying  $\nu^{-1}$ ,  $\alpha_i^\vee = \frac{1}{\epsilon_i} \nu^{-1}(\alpha_i)$ . But we also know  $\epsilon_i = (\alpha_i | \alpha_i)/2$ .

We deduce that:

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

Note that the coroots are not exactly the adjoint of the roots under the form, but differ by a positive scaling factor so that the entries of the Cartan matrix come out right. Hence we obtain the usual expression for the generalized Cartan matrix:

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

**Remark 2.22** (Two pairings between  $\mathfrak{h}$  and  $\mathfrak{h}^*$ ). *We have two natural pairings between  $\mathfrak{h}$  and  $\mathfrak{h}^*$ :*

- *The canonical pairing  $\langle \cdot, \cdot \rangle$ .*
- *The pairing induced by the bilinear form  $(\cdot | \cdot)$*

*The  $\nu$  map identifies these two pairings. In particular, we have*

$$\langle \alpha_j, \alpha_i^\vee \rangle = (\nu(\alpha_i^\vee) | \alpha_j) = \epsilon_i (\alpha_i | \alpha_j).$$

*essentially because we defined the invariant pairing on  $\mathfrak{h}$  by  $(\alpha_i^\vee | h) = \langle \alpha_i, h \rangle \epsilon_i$  for  $h \in \mathfrak{h}$ , so rewriting gives*

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

*They are related by*

$$a_{ij} = \langle \alpha_i, \alpha_j^\vee \rangle = \frac{2(\alpha_i | \alpha_j)}{(\alpha_i | \alpha_i)}.$$

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 2.18 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 2.18 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  $\alpha$ .

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 2.18.

Given a restricted  $\mathfrak{g}(A)$ -module  $V$ , we introduce a linear operator  $\Omega$  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  $\rho$  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  $\alpha$  we choose a basis  $\{e_{\alpha}^{(i)}\}$  of the space  $\mathfrak{g}_{\alpha}$ , 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  $\Omega_0$  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,  $\Omega_0$  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 2.23.** The generalized Casimir operator  $\Omega$  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 2.24 (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 2.25.** *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 2.26 ()**. *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  $\Omega$  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  $\alpha_i^{\vee}$  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

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

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 2.24 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 2.9) and the third equality follows from 2.24. 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 2.27 (Eigenvalue on highest weight modules).** *If, under the hypotheses of Theorem 2.26 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  $\Omega_0$  kills  $v$ , so  $\Omega_0 v = 0$ .

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

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

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 2.26 b),  $\Omega$  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  $\Omega$  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 2.28.** *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

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

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

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

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 2.29.** We define the **compact form**  $\mathfrak{t}(A)$  of  $\mathfrak{g}(A)$  as follows. Let  $\omega_0$  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  $\omega_0$ . 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 2.30.** Let  $(\cdot \mid \cdot)$  be a symmetric bilinear form on  $\mathfrak{g}(A)$ . Define a Hermitian form on  $\mathfrak{g}(A)$  by

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

Then  $(\cdot \mid \cdot)_0$  is nondegenerate on each  $\mathfrak{g}_{\alpha}$ ,  $\alpha \in \Delta \cup \{0\}$ , and satisfies  $(\mathfrak{g}_{\alpha} \mid \mathfrak{g}_{\beta})_0 = 0$  if  $\alpha \neq \beta$ . Moreover, the operators  $\text{ad } u$  and  $-\text{ad } \omega_0(u)$  are adjoint to each other with respect to  $(\cdot \mid \cdot)_0$ . In particular, the restriction of  $(\cdot \mid \cdot)$  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  $\mathfrak{sl}_2(\mathbb{C})$ :

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

**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  $v_{k+1} = 0 = v_{-1}$ .

**Example 3.2** (Deducing the Serre relations). Let  $\mathfrak{g}(A)$  be a Kac–Moody algebra,  $e_i, f_i$  ( $i = 1, \dots, n$ ) be its Chevalley generators. Set

$$\mathfrak{g}_{(i)} = \mathbb{C}e_i + \mathbb{C}\alpha_i^\vee + \mathbb{C}f_i;$$

then  $\mathfrak{g}_{(i)}$  is isomorphic to  $\mathfrak{sl}_2(\mathbb{C})$ , with standard basis  $\{e_i, \alpha_i^\vee, f_i\}$ .

We can deduce now the following relations between the Chevalley generators:

$$(\text{ad } e_i)^{1-a_{ij}}e_j = 0, \quad (\text{ad } f_i)^{1-a_{ij}}f_j = 0, \quad \text{if } i \neq j. \quad (18)$$

We prove the second relation; the first one follows by making use of the Chevalley involution  $\omega$ .

Denote  $v = f_j$ ,  $\theta_{ij} = (\text{ad } f_i)^{1-a_{ij}} f_j$ . Consider  $\mathfrak{g}(A)$  as a  $\mathfrak{g}_{(i)}$ -module by restricting the adjoint representation. We have:

$$\alpha_i^\vee(v) = -a_{ij}v, \quad e_i(v) = 0 \quad \text{if } i \neq j.$$

We can now compute

$$\begin{aligned} [e_i, \theta_{ij}] &= [e_i, \text{ad}^{1-a_{ij}}(f_i)f_j] \\ &= [e_i, f_i^{1-a_{ij}}]f_j + f_i^{1-a_{ij}}[e_i, f_j] \quad (\text{by the Leibniz rule}) \\ &= [e_i, f_i^{1-a_{ij}}]f_j \\ &= (1-a_{ij})(-a_{ij} - (1-a_{ij}) + 1)(\text{ad } f_i)^{-a_{ij}} f_j \\ &= 0 \quad \text{if } i \neq j. \end{aligned}$$

where in the second to last line, we invoked the fact that  $[e, f^k] = -k(k-1)f^{k-1} + kf^{k-1}h$  and  $h = \alpha_i^\vee$  acts on  $f_j$  by  $-a_{ij}$ .

It is also clear that  $e_k$  commutes with  $\theta_{ij}$  if  $k \neq i, k \neq j$  (by relations (1.2.1)), and also if  $k = j$  but  $a_{ij} \neq 0$ . Finally, if  $k = j$  and  $a_{ij} = 0$ , we have:

$$[e_j, \theta_{ij}] = [e_j, [f_i, f_j]] = a_{ji}f_i = 0 \quad (\text{by (C3)}).$$

So,  $[e_k, \theta_{ij}] = 0$  for all  $k$  and we apply the following lemma to conclude that  $\theta_{ij} = 0$ .

**Lemma 3.3.** *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 3.4.** *Note that in the lemma is where we invoke the fact that we have taken the quotient by the largest ideal intersecting  $\mathfrak{h}$  trivially. In particular, this is what allows us to deduce the Serre relations (18).*

### 3.1 Locally nilpotent operators

Now we need a general fact about a module  $V$  over a Lie algebra  $\mathfrak{g}$ .

**Definition 3.5.** An element  $x \in \mathfrak{g}$  is said to be **locally nilpotent on**  $V$  if for any  $v \in V$  there exists a positive integer  $N$  such that  $x^N(v) = 0$ .

**Lemma 3.6 (Local nilpotence lemma).**

- (a) Let  $y_1, y_2, \dots$  be a system of generators of a Lie algebra  $\mathfrak{g}$  and let  $x \in \mathfrak{g}$  be such that  $(\text{ad } x)^{N_i} y_i = 0$  for some positive integers  $N_i, i = 1, 2, \dots$ . Then  $\text{ad } x$  is locally nilpotent on  $\mathfrak{g}$ .
- (b) Let  $v_1, v_2, \dots$  be a system of generators of a  $\mathfrak{g}$ -module  $V$ , and let  $x \in \mathfrak{g}$  be such that  $\text{ad } x$  is locally nilpotent on  $\mathfrak{g}$  and  $x^{N_i}(v_i) = 0$  for some positive integers  $N_i, i = 1, 2, \dots$ . Then  $x$  is locally nilpotent on  $V$ .

*Proof.* Since  $\text{ad } x$  is a derivation of  $\mathfrak{g}$ , one has the Leibniz formula

$$(\text{ad } x)^k [y, z] = \sum_{i=0}^k \binom{k}{i} [(\text{ad } x)^i y, (\text{ad } x)^{k-i} z],$$

proving (a) by induction on the length of commutators in the  $y_i$ . Part (b) follows from the following formula (for  $\lambda = \mu = 0$ ):

$$(x - \lambda - \mu)^k a = \sum_{s=0}^k \binom{k}{s} ((\text{ad } x - \lambda)^s a) (x - \mu)^{k-s}, \quad k \geq 0, \lambda, \mu \in \mathbb{C}, \quad (19)$$

which holds in any associative algebra. In order to prove (19), note that  $\text{ad } x = L_x - R_x$ , where  $L_x$  and  $R_x$  are the operators of left and right multiplication by  $x$ , and that  $L_x$  and  $R_x$  commute (by associativity). Now we apply the binomial formula to  $L_x - \lambda - \mu = (\text{ad } x - \lambda) + (R_x - \mu)$ .  $\square$

Applying the binomial formula to  $\text{ad } x = L_x - R_x$ , we obtain another useful formula (in any associative algebra):

$$(\text{ad } x)^k a = \sum_{s=0}^k (-1)^s \binom{k}{s} x^{k-s} a x^s. \quad (20)$$

**Lemma 3.7 (Nilpotence of Chevalley generators).**  $\text{ad } e_i$  and  $\text{ad } f_i$  are locally nilpotent on  $\mathfrak{g}(A)$ .

*Proof.* By the Serre relations (18),

$$(\text{ad } e_i)^{|a_{ij}|+1} x = 0 = (\text{ad } f_i)^{|a_{ij}|+1} x, \quad \text{if } x = e_j \text{ or } f_j.$$

Also,

$$(\text{ad } e_i)^2 h = 0 = (\text{ad } f_i)^2 h, \quad \text{if } h \in \mathfrak{h}.$$

Now apply Lemma 3.6.  $\square$

**Definition 3.8.** A  $\mathfrak{g}(A)$ -module  $V$  is called  **$\mathfrak{h}$ -diagonalizable** if

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

As usual,  $V_\lambda$  is called a **weight space**,  $\lambda \in \mathfrak{h}^*$  is called a **weight** if  $V_\lambda \neq 0$ , and  $\dim V_\lambda$  is called the **multiplicity** of  $\lambda$  and is denoted by  $\text{mult}_V \lambda$ . Similarly, one defines an  $\mathfrak{h}'$ -diagonalizable  $\mathfrak{g}'(A)$ -module, its weights, etc.

An  $\mathfrak{h}$ - (resp.  $\mathfrak{h}'$ -) diagonalizable module over a Kac–Moody algebra  $\mathfrak{g}(A)$  (resp.  $\mathfrak{g}'(A)$ ) is called **integrable** if all  $e_i$  and  $f_i$  ( $i = 1, \dots, n$ ) are locally nilpotent on  $V$ .

Note that the underlying module of the adjoint representation of a Kac–Moody algebra is an integrable module by 3.7.

**Proposition 3.9 (Weights of integrable modules).** Let  $V$  be an integrable  $\mathfrak{g}(A)$ -module.

- (a) As a  $\mathfrak{g}_{(i)}$ -module,  $V$  decomposes into a direct sum of finite dimensional irreducible  $\mathfrak{h}$ -invariant modules (hence the action of  $\mathfrak{g}_{(i)}$  on  $V$  can be “integrated” to the action of the group  $SL_2(\mathbb{C})$ ).
- (b) Let  $\lambda \in \mathfrak{h}^*$  be a weight of  $V$  and let  $\alpha_i$  be a simple root of  $\mathfrak{g}(A)$ . Denote by  $M$  the set of all  $t \in \mathbb{Z}$  such that  $\lambda + t\alpha_i$  is a weight of  $V$ , and let  $m_t = \text{mult}_V(\lambda + t\alpha_i)$ . Then:
  - (i)  $M$  is the closed interval of integers  $[-p, q]$ , where  $p$  and  $q$  are both either nonnegative integers or  $\infty$ , and  $p - q = \langle \lambda, \alpha_i^\vee \rangle$  when both  $p$  and  $q$  are finite; if  $\text{mult}_V \lambda < \infty$ , then  $p$  and  $q$  are finite.
  - (ii)  $e_i : V_{\lambda+t\alpha_i} \rightarrow V_{\lambda+(t+1)\alpha_i}$  is an injection for  $t \in [-p, -\frac{1}{2}\langle \lambda, \alpha_i^\vee \rangle]$ ; in particular, the function  $t \mapsto m_t$  increases on this interval.
  - (iii) The function  $t \mapsto m_t$  is symmetric with respect to  $t = -\frac{1}{2}\langle \lambda, \alpha_i^\vee \rangle$ .
  - (iv) If both  $\lambda$  and  $\lambda + \alpha_i$  are weights, then  $e_i(V_\lambda) \neq 0$ .

*Proof.* We have by the Leibniz rule:

$$e_i f_i^k(v) = k(1 - k + \langle \lambda, \alpha_i^\vee \rangle) f_i^{k-1}(v) + f_i^k e_i(v), \quad v \in V_\lambda. \quad (21)$$

We deduce that the subspace

$$U = \sum_{k, m \geq 0} f_i^k e_i^m(v)$$

is  $(\mathfrak{g}_{(i)} + \mathfrak{h})$ -invariant. As  $e_i$  and  $f_i$  are locally nilpotent on  $V$ ,  $\dim U < \infty$ . By the Weyl complete reducibility theorem applied to the  $\mathfrak{g}_{(i)}$ -module  $U$ , the latter decomposes into a direct sum of finite

dimensional  $\mathfrak{h}$ -invariant irreducible  $\mathfrak{g}_{(i)}$ -modules (cf. Exercise 3.11). So, each  $v \in V$  lies in a direct sum of finite dimensional  $\mathfrak{h}$ -invariant irreducible  $\mathfrak{g}_{(i)}$ -modules, and (a) follows.

For the proof of (b) we use (a) and Lemma 3.1. Set

$$U = \sum_{k \in \mathbb{Z}} V_{\lambda + k\alpha_i};$$

this is a  $(\mathfrak{g}_{(i)} + \mathfrak{h})$ -module, which is a direct sum of finite dimensional irreducible  $(\mathfrak{g}_{(i)} + \mathfrak{h})$ -modules. Let  $p = -\inf M$ ,  $q = \sup M$ . Both  $p$  and  $q$  are nonnegative as  $0 \in M$ . Now all the statements of (b) follow from Lemma 3.2(b), as  $\langle \lambda + t\alpha_i, \alpha_i^\vee \rangle = 0$  for  $t = -\frac{1}{2}\langle \lambda, \alpha_i^\vee \rangle$ .  $\square$

**Corollary 3.10** (Consequences for weights of integrable modules).

- (a) If  $\lambda$  is a weight of an integrable  $\mathfrak{g}(A)$ -module  $V$  and  $\lambda + \alpha_i$  (resp.  $\lambda - \alpha_i$ ) is not a weight, then  $\langle \lambda, \alpha_i^\vee \rangle \geq 0$  (resp.  $\langle \lambda, \alpha_i^\vee \rangle \leq 0$ ).
- (b) If  $\lambda$  is a weight of  $V$ , then  $\lambda - \langle \lambda, \alpha_i^\vee \rangle \alpha_i$  is also a weight of the same multiplicity.

*Proof.* Suppose  $\lambda$  is a weight. This means  $0 \in M$ .

If  $\lambda + \alpha_i$  is not a weight, then  $1 \notin M$ . Since  $M$  is the integer interval  $[-p, q]$ , the only way this happens is if  $q = 0$ . Then  $p - q = p - 0 = p = \langle \lambda, \alpha_i^\vee \rangle \geq 0$ .

The multiplicities are symmetric with respect to  $t_0 = -\frac{1}{2}\langle \lambda, \alpha_i^\vee \rangle$ . At  $t = 0$ , the multiplicity is  $\text{mult}_V(\lambda)$ . Symmetry says  $m_t = m_{-\langle \lambda, \alpha_i^\vee \rangle - t}$ . Plugging  $t = 0$  gives  $\text{mult}_V(\lambda) = m_0 = m_{-\langle \lambda, \alpha_i^\vee \rangle}$ .  $\square$

**Remark 3.11.** Let  $V$  be an integrable  $\mathfrak{g}'(A)$ -module. Then, clearly the proposition and corollary with  $\mathfrak{h}$  replaced by  $\mathfrak{h}'$ , still hold. Furthermore, the local nilpotency of  $e_i$  and  $f_i$  on  $V$  guarantees that  $V$  is  $\mathfrak{h}_i$ -diagonalizable, and hence  $\mathfrak{h}'$ -diagonalizable provided that  $n < \infty$ . This follows from a general fact which will be proved later.

**Remark 3.12** (Interpretation of the number  $\langle \lambda, \alpha_i^\vee \rangle$ ). There is a lattice  $P$  of weights, called the **weight lattice**, which contains all weights of integrable modules. It is defined by

$$P = \{\lambda \in \mathfrak{h}^* \mid \langle \lambda, \alpha_i^\vee \rangle \in \mathbb{Z} \text{ for all } i = 1, \dots, n\}.$$

It is generated by the **fundamental weights**  $\{\omega_i\}_{i=1}^n$  defined by

$$\langle \omega_i, \alpha_j^\vee \rangle = \delta_{ij}.$$

In particular, the fundamental weights are the dual basis to the simple coroots  $\{\alpha_i^\vee\}_{i=1}^n$  under the evaluation pairing. Then the number  $\langle \lambda, \alpha_i^\vee \rangle$  is the coefficient of  $\omega_i$  in the expansion of  $\lambda$  in terms of the fundamental weights.

One has to ask why the weights of integrable modules lie in  $P$ . This is because restricting to each  $\mathfrak{sl}_2$ -triple forces the eigenvalues of  $\alpha_i^\vee$  to be integers, i.e.  $\langle \lambda, \alpha_i^\vee \rangle \in \mathbb{Z}$ , precisely because of the integrable condition.

The root lattice  $Q$  is defined to be the integer span of the simple roots  $\{\alpha_i\}_{i=1}^n$ . It is a sublattice of  $P$ . In the finite dimensional case,  $Q$  has finite index in  $P$  of order equal to the determinant of the Cartan matrix. In the infinite dimensional case,  $\det A = 0$ , so  $Q$  has infinite index in  $P$ .

### 3.2 Weyl group, invariance, and geometry

**Definition 3.13 (Weyl group).** For each  $i = 1, \dots, n$  we define the **fundamental reflection**  $r_i$  of the space  $\mathfrak{h}^*$  by

$$r_i(\lambda) = \lambda - \langle \lambda, \alpha_i^\vee \rangle \alpha_i, \quad \lambda \in \mathfrak{h}^*.$$

It is clear that  $r_i$  is a reflection since its fixed point set is  $T_i = \{\lambda \in \mathfrak{h}^* \mid \langle \lambda, \alpha_i^\vee \rangle = 0\}$ , and  $r_i(\alpha_i) = -\alpha_i$ .

The subgroup  $W$  of  $GL(\mathfrak{h}^*)$  generated by all fundamental reflections is called the **Weyl group** of  $\mathfrak{g}(A)$ . We will write  $W(A)$  when necessary to emphasize the dependence on  $A$ .

**Proposition 3.14 (Weyl group invariance of weights and roots).**

- (a) Let  $V$  be an integrable module for Kac–Moody algebra  $\mathfrak{g}(A)$ . Then  $\text{mult}_V \lambda = \text{mult}_V w(\lambda)$  for every  $\lambda \in \mathfrak{h}^*$  and  $w \in W$ . In particular, the set of weights of  $V$  is  $W$ -invariant.
- (b) The root system  $\Delta$  of  $\mathfrak{g}(A)$  is  $W$ -invariant, and  $\text{mult } \alpha = \text{mult } w(\alpha)$  for every  $\alpha \in \Delta$ ,  $w \in W$ .

*Proof.* The multiplicities  $m_t$  in a weight string along a simple root direction are symmetric and behave nicely under reflection. The fundamental reflection  $r_i$  exactly encodes this symmetry. It sends  $\lambda \mapsto \lambda - \langle \lambda, \alpha_i^\vee \rangle \alpha_i$   $\square$

The following lemma tells us what happens to the positive roots under a fundamental reflection. In particular, it says that all of the positive roots stay positive except for the simple root corresponding to the reflection, which gets sent to its negative.

**Lemma 3.15.** If  $\alpha \in \Delta_+$  and  $r_i(\alpha) < 0$ , then  $\alpha = \alpha_i$ . In other words,  $\Delta_+ \setminus \{\alpha_i\}$  is  $r_i$ -invariant.

*Proof.* Follows from Lemma 2.9. In particular, the simple reflection  $r_i$  reverses the  $\alpha_i$ -string through  $\alpha$  and we know that root strings don't pass through zero unless  $\alpha = \alpha_i$ .  $\square$

**Proposition 3.16 (3.9).** The restriction of the bilinear form  $(\cdot | \cdot)$  to  $\mathfrak{h}^*$  is  $W$ -invariant.

*Proof.* It is a standard fact from linear algebra that reflection with respect to a nonzero vector  $\alpha_i$  preserves the bilinear form, provided the form makes  $\alpha_i$  non-isotropic (i.e.  $(\alpha_i|\alpha_i) \neq 0$ ). As  $|r_i(\alpha_i)|^2 = |-\alpha_i|^2 = |\alpha_i|^2 \neq 0$ , it suffices to check that  $(\lambda|\alpha_i) = 0$  implies  $(r_i(\lambda)|\alpha_i) = 0$ . Compute:

$$(r_i(\lambda)|\alpha_i) = (\lambda|\alpha_i) - \langle \lambda, \alpha_i^\vee \rangle (\alpha_i|\alpha_i)$$

Now use the formula:

$$\alpha_i^\vee = \frac{2}{(\alpha_i|\alpha_i)} \nu^{-1}(\alpha_i)$$

So

$$\langle \lambda, \alpha_i^\vee \rangle = \frac{2}{(\alpha_i|\alpha_i)} (\lambda|\alpha_i)$$

Plugging back:

$$(r_i(\lambda)|\alpha_i) = (\lambda|\alpha_i) - \frac{2}{(\alpha_i|\alpha_i)} (\lambda|\alpha_i) (\alpha_i|\alpha_i) = (\lambda|\alpha_i) - 2(\lambda|\alpha_i) = -(\lambda|\alpha_i)$$

Therefore if  $(\lambda|\alpha_i) = 0$ , then  $(r_i(\lambda)|\alpha_i) = 0$ .  $\square$

We have the following technical lemma about words in the Weyl group. It says that if applying a word  $w = r_{i_1} \cdots r_{i_t}$  to a simple root makes it negative, then  $wr_i$  is a shorter word than  $w$ .

**Lemma 3.17 (Exchange lemma).** *If  $\alpha_i$  is a simple root and*

$$r_{i_1} \cdots r_{i_t}(\alpha_i) < 0,$$

*then there exists  $s$  ( $1 \leq s \leq t$ ) such that*

$$r_{i_1} \cdots r_{i_s} \cdots r_{i_t} r_i = r_{i_1} \cdots r_{i_{s-1}} r_{i_{s+1}} \cdots r_{i_t}. \quad (22)$$

**Definition 3.18.** *The expression  $w = r_{i_1} \cdots r_{i_s} \in W$  is called **reduced** if  $s$  is minimal possible among all representations of  $w \in W$  as a product of the  $r_i$ . Then  $s$  is called the **length** of  $w$  and is denoted by  $\ell(w)$ .*

Note that  $\det_{\mathfrak{h}^*} r_i = -1$  and hence

$$\det_{\mathfrak{h}^*} w = (-1)^{\ell(w)} \quad \text{for } w \in W. \quad (23)$$

The following lemma is an important corollary of Lemma 3.17



**Lemma 3.19.** *Let  $w = r_{i_1} \cdots r_{i_t} \in W$  be a reduced expression and let  $\alpha_i$  be a simple root. Then we have*

- (a)  $\ell(wr_i) < \ell(w)$  if and only if  $w(\alpha_i) < 0$ .
- (b)  $w(\alpha_{i_t}) < 0$ .
- (c) If  $\ell(wr_i) < \ell(w)$ , then there exists  $s$ ,  $1 \leq s \leq t$ , such that

$$r_{i_1} \cdots r_{i_s} \cdots r_{i_t} = r_{i_1} \cdots r_{i_{s-1}} r_i r_{i_{s+1}} \cdots r_{i_t}.$$

*Proof.* By Lemma 3.17 (applied to  $w$ ),  $w(\alpha_i) < 0$  implies that  $\ell(wr_i) < \ell(w)$ . If now  $w(\alpha_i) > 0$ , then  $wr_i(\alpha_i) < 0$  and hence  $\ell(w) = \ell(wr_i^2) < \ell(wr_i)$ , proving (a). Part (b) follows immediately from (a).

Finally, if  $\ell(wr_i) < \ell(w)$ , then (a) implies  $w(\alpha_i) < 0$  and applying Lemma 3.17 to  $w$  we deduce the exchange condition, multiplying it by  $(r_{i_1} \cdots r_{i_{s-1}})^{-1}$  on the left and by  $r_i$  on the right.  $\square$

Now we are in a position to study the geometric properties of the action of the Weyl group. Pick a real realization  $(\mathfrak{h}_{\mathbb{R}}, \Pi, \Pi^{\vee})$  of  $A$  so that  $(\mathfrak{h}_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}, \Pi, \Pi^{\vee})$  is a realization of  $A$  over  $\mathbb{C}$ . The set

$$C = \{h \in \mathfrak{h}_{\mathbb{R}} \mid \langle \alpha_i, h \rangle \geq 0 \text{ for } i = 1, \dots, n\}$$

is called the **fundamental chamber**. There is a contragradient action of  $W$  on  $\mathfrak{h}$  defined by

$$r_i(h) = h - \langle \alpha_i, h \rangle \alpha_i^{\vee}, \quad h \in \mathfrak{h}$$

which respects the pairing  $\langle \cdot, \cdot \rangle$ .

$$\langle r_i(\alpha), h \rangle = \langle \alpha, r_i(h) \rangle, \quad \alpha \in \mathfrak{h}^*, h \in \mathfrak{h}.$$

Note that  $\mathfrak{h}_{\mathbb{R}}$  is stable under  $W$  since  $Q^{\vee} \subset \mathfrak{h}_{\mathbb{R}}$ . Therefore we can talk about the sets  $w(C)$ ,  $w \in W$ , called **chambers**, and their union

$$X = \bigcup_{w \in W} w(C)$$

is called the **Tits cone**. We clearly have the corresponding dual notions of  $C^{\vee}$  and  $X^{\vee}$  in  $\mathfrak{h}_{\mathbb{R}}^*$ .

**Proposition 3.20 (Geometry of  $W$ ).**

- (a) For  $h \in C$ , the group  $W_h = \{w \in W \mid w(h) = h\}$  is generated by the fundamental reflections which it contains.
- (b) The fundamental chamber  $C$  is a fundamental domain for the action of  $W$  on  $X$ , i.e. any orbit  $W \cdot h$  of  $h \in X$  intersects  $C$  in exactly one point. In particular,  $W$  operates simply transitively on chambers.

(c)  $X = \{h \in \mathfrak{h}_{\mathbb{R}} \mid \langle \alpha, h \rangle < 0 \text{ only for a finite number of } \alpha \in \Delta_+\}$ . In particular,  $X$  is a convex cone.

(d)  $C = \{h \in \mathfrak{h}_{\mathbb{R}} \mid \text{for every } w \in W, h - w(h) = \sum_i c_i \alpha_i^\vee \text{ where } c_i \geq 0\}$ .

(e) The following conditions are equivalent:

(i)  $|W| < \infty$ ;

(ii)  $X = \mathfrak{h}_{\mathbb{R}}$ ;

(iii)  $|\Delta| < \infty$ ;

(iv)  $|\Delta^\vee| < \infty$ .

(f) If  $h \in X$ , then  $|W_h| < \infty$  if and only if  $h$  lies in the interior of  $X$ .

*Proof.* Let  $w \in W$  and let  $w = r_{i_1} \cdots r_{i_s}$  be a reduced expression of  $w$ . Take  $h \in C$  and suppose that  $h' = w(h) \in C$ . We have  $\langle \alpha_i, h \rangle \geq 0$  and therefore  $\langle w(\alpha_i), h' \rangle \geq 0$  by the  $W$ -invariance of the pairing. But by lemma above,  $w(\alpha_i) < 0$ , and hence  $\langle w(\alpha_i), h' \rangle \leq 0$  because  $h' \in C$ . So  $\langle w(\alpha_i), h' \rangle = 0$  and  $\langle \alpha_i, h \rangle = 0$ . Hence  $r_i(h) = h$  and both (a) and (b) follow by induction on  $\ell(w)$ .

Set

$$X' = \{h \in \mathfrak{h}_{\mathbb{R}} \mid \langle \alpha, h \rangle < 0 \text{ only for a finite number of } \alpha \in \Delta_+\}$$

It is clear that  $C \subset X'$  and that  $X'$  is  $r_i$ -invariant. Hence  $X' \supset X$ . In order to prove the reverse inclusion, let  $h \in X'$  and set  $M_h = \{\alpha \in \Delta_+ \mid \langle \alpha, h \rangle < 0\}$ . We need to show that  $M_h$  is empty. By definition,  $|M_h|$  is finite. If  $M_h \neq \emptyset$ , then  $\alpha_i \in M_h$  for some  $i$ . But then it follows from Lemma 3.15 that  $|M_{r_i(h)}| < |M_h|$ . This process must eventually terminate with some  $h^{(k)}$  such that  $M_{h^{(k)}} = \emptyset$ . But  $M_{h^{(k)}} = \emptyset$  means  $h^{(k)} \in C$ . Thus  $h^{(k)} \in C$  and  $h = w^{-1}(h^{(k)}) \in w^{-1}(C)$ , so  $h \in W(C) = X$  as desired.

The inclusion  $\supset$  of (d) follows from the fact that

$$h - r_i(h) = \langle \alpha_i, h \rangle \alpha_i^\vee$$

so being in the RHS implies that  $\langle \alpha_i, h \rangle \geq 0$  for all  $i$ , i.e.  $h \in C$ .

We prove the reverse inclusion by induction on  $s = \ell(w)$ . For  $\ell(w) = 1$ , (d) is the definition of  $C$ . If  $\ell(w) > 1$ , let  $w = r_{i_1} \cdots r_{i_s}$ . We have

$$h - w(h) = (h - r_{i_1} \cdots r_{i_{s-1}}(h)) + r_{i_1} \cdots r_{i_{s-1}}(h - r_{i_s}(h)),$$

and we apply the inductive assumption to the first summand and part (b) of the lemma above for  $\Delta^\vee$  to the second summand. In particular, the lemma says that  $w(\alpha_{i_s}) < 0$ . But  $w(\alpha_{i_s}) < 0$  is

equivalent to  $r_{i_1} \cdots r_{i_{s-1}}(\alpha_{i_s}) > 0$  is a positive root. The second summand is

$$\begin{aligned} r_{i_1} \cdots r_{i_{s-1}}(h - r_{i_s}(h)) &= r_{i_1} \cdots r_{i_{s-1}}(\langle \alpha_{i_s}, h \rangle \alpha_{i_s}^\vee) \\ &= \langle \alpha_{i_s}, h \rangle r_{i_1} \cdots r_{i_{s-1}}(\alpha_{i_s}^\vee) \end{aligned}$$

is a nonnegative multiple of a positive coroot, as desired.

Now we prove (e). To show (i)  $\Rightarrow$  (ii). Pick  $\rho \in \mathfrak{h}_{\mathbb{R}}^*$  in the interior of  $C^\vee$ , e.g.  $\rho = \Lambda_1 + \cdots + \Lambda_n$ . Then  $\langle \rho, \alpha_i^\vee \rangle > 0$  for all  $i$ . Consider the finite set  $W \cdot h$  and choose  $h' \in W \cdot h$  that maximizes  $\langle \rho, h' \rangle$ . If  $h' \notin C$ , then for some  $i$  we have  $\langle \alpha_i, h' \rangle < 0$ . Reflect:

$$\langle \rho, r_i(h') \rangle = \langle \rho, h' - \langle \alpha_i, h' \rangle \alpha_i^\vee \rangle = \langle \rho, h' \rangle - \langle \alpha_i, h' \rangle \langle \rho, \alpha_i^\vee \rangle > \langle \rho, h' \rangle$$

since  $\langle \rho, \alpha_i^\vee \rangle > 0$  and  $\langle \alpha_i, h' \rangle < 0$ . This contradicts maximality. Hence  $\langle \alpha_i, h' \rangle \geq 0$  for all  $i$ , i.e.  $h' \in C$ . So every orbit meets  $C$ ; thus  $X = \mathfrak{h}_{\mathbb{R}}$ .

In order to show (ii)  $\Rightarrow$  (iii) take  $h$  in the interior of  $C$ , so  $\langle \alpha_i, h \rangle > 0$  for all simple  $\alpha_i$ . Then for every positive root  $\alpha$  (a nonnegative combination of simple roots) we have  $\langle \alpha, h \rangle > 0$ , hence  $\langle \alpha, -h \rangle < 0$  for all  $\alpha \in \Delta_+$ . Since (ii) says  $-h \in X$ , apply part (c) of this proposition:  $X = \{x : \langle \alpha, x \rangle < 0 \text{ for only finitely many } \alpha \in \Delta_+\}$ . But for  $x = -h$  every  $\alpha \in \Delta_+$  pairs negatively, so  $\Delta_+$  must be finite. Thus  $|\Delta| < \infty$ .

(iii)  $\Rightarrow$  (i) because The Weyl group  $W$  acts on  $\Delta$  by permuting roots, so we get a homomorphism  $\varphi : W \rightarrow \text{Sym}(\Delta)$ . To conclude  $|W| < \infty$ , it suffices to show  $\varphi$  is injective, because  $\text{Sym}(\Delta)$  is finite.

We invoke the fact that

$$\{w(\alpha) = \alpha \text{ for } w \in W \text{ and all } \alpha \in \Delta_+\} \Rightarrow w = 1. \quad (24)$$

To prove (24), note that if a reduced expression  $w = r_{i_1} \cdots r_{i_s}$  is nontrivial, then part (b) of the above lemma implies that  $w(\alpha_{i_s}) < 0$ , contradiction. The fact that (iv) is equivalent to (i) follows by using the dual root system.

Finally, to prove (f) we may assume that  $h \in C$ . Then (f) follows from (a) by applying the equivalence of (e)(i) and (e)(ii) to  $W_h$  operating on  $\mathfrak{h}/\mathbb{C}h$ .  $\square$

## 4 Classification of affine root systems

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  ${}^tA$ :*

(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.*

**Lemma 4.3** (Symmetrizability of finite and affine type matrices). *Let  $A = (a_{ij})$  be a matrix of finite or affine type such that  $a_{ii} = 2$  and  $a_{ij}a_{ji} = 0$  or  $\geq 1$ . Then  $A$  is symmetrizable.*

In particular, all generalized Cartan matrices of finite or affine type are symmetrizable. This is quite important as much of the structure theory for symmetrizable Kac-Moody algebras we have developed would not be very interesting if we could not handle these examples.

We proceed to classify all generalized Cartan matrices of finite and affine type. For this it is convenient to introduce the so-called **Dynkin diagrams**.

**Definition 4.4** (Dynkin diagram). *Let  $A = (a_{ij})_{i,j=1}^n$  be a generalized Cartan matrix. We associate with  $A$  a graph  $S(A)$ , called the **Dynkin diagram** of  $A$  as follows. If  $a_{ij}a_{ji} \leq 4$  and  $|a_{ij}| \geq |a_{ji}|$ , the vertices  $i$  and  $j$  are connected by  $|a_{ij}|$  lines, and these lines are equipped with an arrow pointing toward  $j$  if  $|a_{ij}| > 1$ . If  $a_{ij}a_{ji} > 4$ , the vertices  $i$  and  $j$  are connected by a bold-faced line equipped with an ordered pair of integers  $|a_{ij}|, |a_{ji}|$ .*

It is clear that  $A$  is indecomposable if and only if  $S(A)$  is a connected graph. Note also that  $A$  is determined by the Dynkin diagram  $S(A)$  and an enumeration of its vertices. We say that  $S(A)$  is of finite, affine, or indefinite type if  $A$  is of that type.

**Proposition 4.5** (Properties of generalized Cartan matrices of finite and affine type). *Let  $A$  be an inde-*

composable generalized Cartan matrix.

- a)  $A$  is of finite type if and only if all its principal minors are positive.
- b)  $A$  is of affine type if and only if all its proper principal minors are positive and  $\det A = 0$ .
- c) If  $A$  is of finite or affine type, then any proper subdiagram of  $S(A)$  is a union of (connected) Dynkin diagrams of finite type.
- d) If  $A$  is of finite type, then  $S(A)$  contains no cycles. If  $A$  is of affine type and  $S(A)$  contains a cycle, then  $S(A)$  is the cycle  $A_\ell^{(1)}$  from Table Aff 1.
- e)  $A$  is of affine type if and only if there exists  $\delta > 0$  such that  $A\delta = 0$ ; such a  $\delta$  is unique up to a constant factor.

In particular, an affine generalized Cartan matrix  $A$  satisfies all proper principal minors are positive,  $\det A = 0$ , and is indecomposable. This means  $A$  is “just one step” beyond the finite type: it’s rank deficient by one. That’s why one gets affine (loop) extensions.

**Theorem 4.6** (Classification of generalized Cartan matrices of finite and affine type).

- a) The Dynkin diagrams of all generalized Cartan matrices of finite type are listed in Table Fin.
- b) The Dynkin diagrams of all generalized Cartan matrices of affine type are listed in Tables Aff 1–3 (all of them have  $\ell + 1$  vertices).
- c) The numerical labels in Tables Aff 1–3 are the coordinates of the unique vector

$$\delta = {}^t(a_0, a_1, \dots, a_\ell)$$

such that  $A\delta = 0$  and the  $a_i$  are positive relatively prime integers.

**Remark 4.7.** Suppose  $S(A)$  is of affine type. It turns out that if  $S(A)$  has cycles, then it is the cycle  $A_\ell^{(1)}$  from Table Aff 1. Otherwise,  $S(A)$  has a vertex such that removing it produces a diagram of finite type. Now suppose that  $S(A)$  has no cycles. Then by Proposition 4.5(c), removing any vertex produces a diagram of finite type. Conversely, any diagram of affine type is obtained from a diagram of Table Fin by adding one vertex in such a way that any subdiagram is from Table Fin. Then one sees that the only the diagrams from Tables Aff 1–3 may be obtained in this way.

**Proposition 4.8** (Characterizations of finite type). Let  $A$  be an indecomposable generalized Cartan matrix. Then the following conditions are equivalent:

- (i)  $A$  is a generalized Cartan matrix of finite type;

- (ii)  $A$  is symmetrizable and the bilinear form  $(\cdot \mid \cdot)_{\mathfrak{h}_{\mathbb{R}}}$  is positive-definite;
- (iii)  $|W| < \infty$ ;
- (iv)  $|\Delta| < \infty$ ;
- (v)  $\mathfrak{g}(A)$  is a simple finite-dimensional Lie algebra;
- (vi) there exists  $\alpha \in \Delta_+$  such that  $\alpha + \alpha_i \notin \Delta$  for all  $i = 1, \dots, n$ .

TABLE Aff 1

$A_1^{(1)}$	$\underset{1}{\circ} \longleftrightarrow \underset{1}{\circ}$
$A_\ell^{(1)} (\ell \geq 2)$	$\begin{array}{c} \underset{1}{\circ} \quad \underset{1}{\circ} \quad \cdots \quad \underset{1}{\circ} \quad \underset{1}{\circ} \\ \diagup \quad \diagdown \\ \underset{1}{\circ} \end{array}$
$B_\ell^{(1)} (\ell \geq 3)$	$\begin{array}{c} \underset{1}{\circ} \\   \\ \underset{1}{\circ} - \underset{2}{\circ} - \underset{2}{\circ} - \cdots - \underset{2}{\circ} \Rightarrow \underset{2}{\circ} \end{array}$
$C_\ell^{(1)} (\ell \geq 2)$	$\underset{1}{\circ} \Rightarrow \underset{2}{\circ} - \cdots - \underset{2}{\circ} \Leftarrow \underset{1}{\circ}$
$D_\ell^{(1)} (\ell \geq 4)$	$\begin{array}{c} \underset{1}{\circ} \quad \quad \quad \underset{1}{\circ} \\   \quad \quad \quad   \\ \underset{1}{\circ} - \underset{2}{\circ} - \underset{2}{\circ} - \cdots - \underset{2}{\circ} - \underset{1}{\circ} \end{array}$
$G_2^{(1)}$	$\underset{1}{\circ} - \underset{2}{\circ} \Rightarrow \underset{3}{\circ}$
$F_4^{(1)}$	$\underset{1}{\circ} - \underset{2}{\circ} - \underset{3}{\circ} \Rightarrow \underset{4}{\circ} - \underset{2}{\circ}$
$E_6^{(1)}$	$\begin{array}{c} \underset{1}{\circ} \\   \\ \underset{2}{\circ} \\   \\ \underset{1}{\circ} - \underset{2}{\circ} - \underset{3}{\circ} - \underset{2}{\circ} - \underset{1}{\circ} \end{array}$
$E_7^{(1)}$	$\begin{array}{c} \underset{2}{\circ} \\   \\ \underset{1}{\circ} - \underset{2}{\circ} - \underset{3}{\circ} - \underset{4}{\circ} - \underset{3}{\circ} - \underset{2}{\circ} - \underset{1}{\circ} \end{array}$
$E_8^{(1)}$	$\begin{array}{c} \underset{3}{\circ} \\   \\ \underset{1}{\circ} - \underset{2}{\circ} - \underset{3}{\circ} - \underset{4}{\circ} - \underset{5}{\circ} - \underset{6}{\circ} - \underset{4}{\circ} - \underset{2}{\circ} \end{array}$

Figure 1: Dynkin diagrams of affine type - part 1

TABLE Aff 2

$A_2^{(2)}$	$\begin{array}{c} 2 \\ \circ \\ \alpha_0 \end{array} \Leftarrow \begin{array}{c} 1 \\ \circ \\ \alpha_1 \end{array}$
$A_{2\ell}^{(2)} (\ell \geq 2)$	$\begin{array}{c} 2 \\ \circ \\ \alpha_0 \end{array} \Leftarrow \begin{array}{c} 2 \\ \circ \\ \alpha_1 \end{array} - \dots - \begin{array}{c} 2 \\ \circ \\ \alpha_{\ell-1} \end{array} \Leftarrow \begin{array}{c} 1 \\ \circ \\ \alpha_\ell \end{array}$
$A_{2\ell-1}^{(2)} (\ell \geq 3)$	$\begin{array}{c} 1 \\ \circ \\ \alpha_1 \end{array} - \begin{array}{c} \alpha_0 \\ \circ \\ \alpha_2 \end{array} \begin{array}{c} 1 \\ \circ \\ \alpha_2 \end{array} - \begin{array}{c} 2 \\ \circ \\ \alpha_3 \end{array} - \dots - \begin{array}{c} 2 \\ \circ \\ \alpha_{\ell-1} \end{array} \Leftarrow \begin{array}{c} 1 \\ \circ \\ \alpha_\ell \end{array}$
$D_{\ell+1}^{(2)} (\ell \geq 2)$	$\begin{array}{c} 1 \\ \circ \\ \alpha_0 \end{array} \Leftarrow \begin{array}{c} 1 \\ \circ \\ \alpha_1 \end{array} - \dots - \begin{array}{c} 1 \\ \circ \\ \alpha_{\ell-1} \end{array} \Rightarrow \begin{array}{c} 1 \\ \circ \\ \alpha_\ell \end{array}$
$E_6^{(2)}$	$\begin{array}{c} 1 \\ \circ \\ \alpha_0 \end{array} - \begin{array}{c} 2 \\ \circ \\ \alpha_1 \end{array} - \begin{array}{c} 3 \\ \circ \\ \alpha_2 \end{array} \Leftarrow \begin{array}{c} 2 \\ \circ \\ \alpha_3 \end{array} - \begin{array}{c} 1 \\ \circ \\ \alpha_4 \end{array}$

TABLE Aff 3

$D_4^{(3)}$	$\begin{array}{c} 1 \\ \circ \\ \alpha_0 \end{array} - \begin{array}{c} 2 \\ \circ \\ \alpha_1 \end{array} \Leftarrow \begin{array}{c} 1 \\ \circ \\ \alpha_2 \end{array}$
-------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Figure 2: Dynkin diagrams of affine type - parts 2 and 3

**Remark 4.9 (Why the number 4?).** For a finite-dimensional semisimple Lie algebra  $\mathfrak{g}$  with Killing form  $(\cdot, \cdot)$ , the Cartan matrix entries are

$$a_{ij} = 2 \frac{(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}$$

The Killing form gives a notion of angle between roots. In particular

$$(\alpha_i, \alpha_j) = |\alpha_i| |\alpha_j| \cos \theta_{ij}$$

from which it follows

$$a_{ij} = 2 \frac{|\alpha_j|}{|\alpha_i|} \cos \theta_{ij}, \quad a_{ji} = 2 \frac{|\alpha_i|}{|\alpha_j|} \cos \theta_{ij}$$



and hence

$$a_{ij}a_{ji} = 4 \cos^2 \theta_{ij}$$

It immediately follows that  $a_{ij}a_{ji} \in \{0, 1, 2, 3\}$  (we cannot have  $a_{ij}a_{ji} = 4$  because that would imply  $\theta_{ij} = 0$  and hence  $\alpha_i, \alpha_j$  are linearly dependent), with corresponding angles  $\pi/2, 2\pi/3, 3\pi/4, 5\pi/6$ . Then we can go back and get formulas for the ratios of squares of the root lengths.

$$\frac{|\alpha_i|^2}{|\alpha_j|^2} = \frac{a_{ji}}{a_{ij}}$$

This in fact shows how to go from the Dynkin diagram to pictures of the root systems.

The only way to get  $a_{ij}a_{ji} = 4$  is considering the  $A_2^{(2)}$  which has Cartan matrix

$$A = \begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix}$$

In affine  $A_2^{(2)}$ , we fold a triple-laced diagram by an order-2 automorphism. This multiplies one root's length by 2 (since we're averaging over an orbit of size 2), squaring that gives a ratio of 4.

So one way of thinking about affineness is that the entries of the Cartan matrix closely resemble those of finite type, in which case this respects the geometry of root systems that the Killing form gives us.

**Remark 4.10.** A root of a finite root system  $\Delta$  which satisfies condition (vi) of Proposition 4.8 is called a **highest root**. It is unique and given by the formula

$$\theta = \sum_{i=1}^{\ell} a_i \alpha_i,$$

where  $a_i$  are the labels of the extended Dynkin diagram from Table Aff 1.

**Remark 4.11.** Kac separates the affine types into three families:

1. **Table Aff 1: Untwisted affine types** These come from taking a finite-dimensional simple Lie algebra  $\mathfrak{g}$  and forming the loop algebra

$$\mathfrak{g} \otimes \mathbb{C}[t, t^{-1}] \oplus \mathbb{C}c \oplus \mathbb{C}d.$$

The corresponding Cartan matrix is obtained by adding one extra node (the "affine node") to the Dynkin diagram of  $\mathfrak{g}$ . These are the "standard" affine types:  $A_{\ell}^{(1)}, B_{\ell}^{(1)}, C_{\ell}^{(1)}, \dots, G_2^{(1)}, F_4^{(1)}, E_6^{(1)}, E_7^{(1)}, E_8^{(1)}$ .

2. **Table Aff 2: Twisted affine types of order 2** These come from folding untwisted diagrams by a diagram automorphism of order 2. For example, folding  $A_{2\ell-1}^{(1)}$  by an involution produces  $C_\ell^{(1)}$ , but folding  $D_{\ell+1}^{(1)}$  or  $E_6^{(1)}$  produces new "twisted" types  $A_{2\ell-1}^{(2)}$ ,  $D_{\ell+1}^{(2)}$ ,  $E_6^{(2)}$ . These correspond to twisted loop algebras:

$$\dot{\mathfrak{g}} \otimes \mathbb{C}[t, t^{-1}]^\sigma$$

where  $\sigma$  is a diagram automorphism of order 2.

3. **Table Aff 3: Twisted affine types of order 3** These arise similarly from folding by a diagram automorphism of order 3 (only possible for  $D_4^{(1)}$ ), producing  $D_4^{(3)}$ .

**Example 4.12 (Affine  $\mathfrak{sl}_2$ ).** We consider the affine Dynkin diagram  $A_1^{(1)}$ . It turns out that this corresponds to the affine Lie algebra  $\widehat{\mathfrak{sl}}_2$ .

$$\widehat{\mathfrak{sl}}_2 = (\mathfrak{sl}_2 \otimes \mathbb{C}[t, t^{-1}]) \oplus \mathbb{C}c \oplus \mathbb{C}d$$

which has bracket

$$\begin{aligned} [x \otimes t^m, y \otimes t^n] &= [x, y] \otimes t^{m+n} + m\delta_{m,-n}\kappa(x, y)c \\ [d, x \otimes t^m] &= mx \otimes t^m \\ [c, \widehat{\mathfrak{sl}}_2] &= 0 \end{aligned}$$

where  $\kappa(x, y) = \text{tr}(xy)$  is the Killing form on  $\mathfrak{sl}_2$ . The Cartan subalgebra is

$$\begin{aligned} \widehat{\mathfrak{h}} &= (\mathfrak{h} \otimes 1) \oplus \mathbb{C}c \oplus \mathbb{C}d \\ &= \mathbb{C}h \oplus \mathbb{C}c \oplus \mathbb{C}d \end{aligned}$$

where  $\mathfrak{h} = \mathbb{C}h$  is the Cartan subalgebra of  $\mathfrak{sl}_2$ . If you study the weight spaces which appear in the adjoint representation of  $\widehat{\mathfrak{sl}}_2$ , you find that the roots are

$$\begin{aligned} \Delta &= \{\alpha + n\delta, -\alpha + n\delta, n\delta \mid n \in \mathbb{Z}, n \neq 0\} \\ &= \{\pm\alpha + n\delta \mid n \in \mathbb{Z}\} \cup \{n\delta \mid n \in \mathbb{Z}, n \neq 0\} \end{aligned}$$

where  $\alpha \in \mathfrak{h}^*$  is the positive root of  $\mathfrak{sl}_2$  and  $\delta \in \widehat{\mathfrak{h}}^*$  is the **null root** defined by

$$\begin{aligned} \delta(h) &= 0 \\ \delta(c) &= 0 \\ \delta(d) &= 1. \end{aligned}$$

The simple roots are  $\alpha_0 = \delta - \alpha$  and  $\alpha_1 = \alpha$ . One can choose corresponding Chevalley generators

$$\begin{aligned} e_0 &= f \otimes t, & f_0 &= e \otimes t^{-1}, & h_0 &= -h + K \\ e_1 &= e \otimes 1, & f_1 &= f \otimes 1, & h_1 &= h \end{aligned}$$

where  $e, f, h$  are the standard Chevalley generators of  $\mathfrak{sl}_2$ .

The positive roots are

$$\begin{aligned}\Delta_+ &= \{\alpha + n\delta, -\alpha + n\delta, n\delta \mid n \in \mathbb{Z}_{>0}\} \cup \{\alpha\} \\ &= \{\alpha_1 + n\delta, \alpha_0 + n\delta, n\delta \mid n \in \mathbb{Z}_{>0}\} \cup \{\alpha_1\}\end{aligned}$$

Note that for affine type  $A$ , one has the imaginary root  $\delta$  is always the sum of the simple roots (in particular the labels on the affine Dynkin diagram are all 1's).

Then one computes the pairings

$$\begin{aligned}a_{11} &= \langle \alpha_1, \alpha_1^\vee \rangle = 2 \\ a_{00} &= \frac{2(\alpha_0 \mid \alpha_0)}{(\alpha_0 \mid \alpha_0)} \quad \text{because there's a unique extension of } (\cdot \mid \cdot) \text{ to } \widehat{\mathfrak{sl}}_2 \\ a_{10} &= \frac{2(\alpha_1 \mid \alpha_0)}{(\alpha_1 \mid \alpha_1)} = \frac{2(\alpha \mid \delta - \alpha)}{(\alpha \mid \alpha)} = -2\end{aligned}$$

because under the extension of  $(\cdot \mid \cdot)$  to  $\widehat{\mathfrak{sl}}_2$ ,  $\delta$  is orthogonal to everything. By general uniqueness results, this is the only way to extend the Killing form so we didn't cheat. Hence the Cartan matrix is

$$A = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

which is indeed the Cartan matrix of  $A_1^{(1)}$ .

**Example 4.13 (Decomposition of  $\widehat{\mathfrak{sl}}_2$  into action for  $\mathfrak{sl}_2$  triple).** Consider the  $\mathfrak{sl}_2$ -triple  $\{e_0, f_0, h_0\}$  inside  $\widehat{\mathfrak{sl}}_2$ . Then  $\widehat{\mathfrak{sl}}_2$  decomposes into irreducible representations of this  $\mathfrak{sl}_2$  as follows:

For each  $n \in \mathbb{Z}$ , set  $W(n) := \text{span}\{e \otimes t^n, h \otimes t^{n+1}, f \otimes t^{n+2}\}$ . Then for  $n \neq -1$ ,  $W(n)$  is a 3-dim irreducible  $\mathfrak{sl}_2$ -module (spin 1) under  $\text{ad}(e_0, f_0, h_0)$ , with weights  $-2, 0, +2$  (eigenvalues of  $\text{ad } h_0$ ). Concretely,

$$\begin{aligned}[h_0, e \otimes t^n] &= -2e \otimes t^n, & [e_0, e \otimes t^n] &= -h \otimes t^{n+1}, & [f_0, e \otimes t^n] &= 0, \\ [h_0, h \otimes t^{n+1}] &= 0, & [e_0, h \otimes t^{n+1}] &= -2f \otimes t^{n+2}, & [f_0, h \otimes t^{n+1}] &= 2e \otimes t^n, \\ [h_0, f \otimes t^{n+2}] &= +2f \otimes t^{n+2}, & [e_0, f \otimes t^{n+2}] &= 0, & [f_0, f \otimes t^{n+2}] &= h \otimes t^{n+1}.\end{aligned}$$

No central term appears because the loop degrees don't sum to zero in these brackets.

For  $n = -1$ , instead we have to take

$$W(-1) = \text{span}\{e \otimes t^{-1}, -h \otimes 1 + K, f \otimes t\}$$

to be the adjoint representation of  $\mathfrak{sl}_2$ . There is also the 1-dimensional trivial submodule  $\mathbb{C}K$ .

There is another 1-dimensional trivial submodule:  $\mathbb{C}(h_0 - 2d)$ , since

$$[e_0, h_0 - 2d] = 0 = [f_0, h_0 - 2d] = [h_0, h_0 - 2d]$$

Indeed  $[d, x \otimes t^m] = m x \otimes t^m$  gives  $[e_0, d] = -e_0$ ,  $[f_0, d] = f_0$ , so the combination  $h_0 - 2d$  is fixed.

Putting these together, you can write the adjoint module as

$$\widehat{\mathfrak{sl}}_2 \cong \left( \bigoplus_{n \in \mathbb{Z}} W(n) \right) \oplus \mathbb{C}K \oplus \mathbb{C}(h_0 - 2d)$$

## 5 Real and imaginary roots

We give an explicit description of the root system of a Kac-Moody algebra  $\mathfrak{g}(A)$ . Our main instrument is the notion of an imaginary root, which has no counterpart in the finite dimensional theory.

**Definition 5.1 (Real root).** A root  $\alpha \in \Delta$  is called **real** if there exists  $w \in W$  such that  $w(\alpha)$  is a simple root. Otherwise,  $\alpha$  is called **imaginary**.

Denote by  $\Delta^{\text{re}}$  and  $\Delta_+^{\text{re}}$  the sets of all real and positive real roots respectively.

Let  $\alpha \in \Delta^{\text{re}}$ ; then  $\alpha = w(\alpha_i)$  for some  $\alpha_i \in \Pi$ ,  $w \in W$ . Define the **dual (real) root**  $\alpha^\vee \in \Delta^{\vee \text{re}}$  by  $\alpha^\vee = w(\alpha_i^\vee)$ . This is independent of the choice of the presentation  $\alpha = w(\alpha_i)$ . Indeed, we have to show that the equality  $u(\alpha_i) = \alpha_j$  implies  $u(\alpha_i^\vee) = \alpha_j^\vee$  for  $u \in W$ . But we showed this earlier. Thus we have a canonical  $W$ -equivariant bijection  $\Delta^{\text{re}} \rightarrow \Delta^{\vee \text{re}}$ . By an easy induction on  $\text{ht } \alpha$ , one shows, using Proposition 5.2e below, that  $\alpha > 0$  if and only if  $\alpha^\vee > 0$ .

We define a reflection  $r_\alpha$  with respect to  $\alpha \in \Delta^{\text{re}}$  by

$$r_\alpha(\lambda) = \lambda - \langle \lambda, \alpha^\vee \rangle \alpha, \quad \lambda \in \mathfrak{h}^*.$$

Since  $\langle \alpha, \alpha^\vee \rangle = 2$ , this is a reflection, and since  $r_\alpha = wr_iw^{-1}$  if  $\alpha = w(\alpha_i)$ , it lies in  $W$ . Note that  $r_\alpha$  is the fundamental reflection  $r_i$ .

### 5.1 Key properties

**Proposition 5.2 (Properties of real roots).** Let  $\alpha$  be a real root of a Kac-Moody algebra  $\mathfrak{g}(A)$ . Then:

- a)  $\text{mult } \alpha = 1$ .
- b)  $k\alpha$  is a root if and only if  $k = \pm 1$ .

c) If  $\beta \in \Delta$  then there exist nonnegative integers  $p$  and  $q$  related by the equation

$$p - q = \langle \beta, \alpha^\vee \rangle,$$

such that  $\beta + k\alpha \in \Delta \cup \{0\}$  if and only if  $-p \leq k \leq q$ ,  $k \in \mathbb{Z}$ .

d) Suppose that  $A$  is symmetrizable and let  $(\cdot|\cdot)$  be a standard invariant bilinear form on  $\mathfrak{g}(A)$ . Then

(i)  $(\alpha|\alpha) > 0$ .

(ii)  $\alpha^\vee = 2\nu^{-1}(\alpha)/(\alpha|\alpha)$ .

(iii) If  $\alpha = \sum_i k_i \alpha_i$ , then  $k_i(\alpha_i|\alpha_i) \in (\alpha|\alpha)\mathbb{Z}$ .

e) Provided that  $\pm\alpha \notin \Pi$ , there exists  $i$  such that

$$|\text{ht } r_i(\alpha)| < |\text{ht } \alpha|.$$

*Proof.* All the statements a)–d) are clear if  $\alpha$  is a simple root. Now a), b), and c) follow from the fact that the action of the Weyl group preserves root system and root multiplicities 3.14.

Statements d)(i) and (ii) follow from the fact that the restriction of the bilinear form to  $\mathfrak{h}^*$  is  $W$ -invariant. Statement d)(iii) follows from the fact that  $\alpha^\vee \in \sum \mathbb{Z}\alpha_i^\vee$  and the following formula:

$$\alpha^\vee = \sum_i \frac{(\alpha_i|\alpha_i)}{(\alpha|\alpha)} k_i \alpha_i^\vee$$

This formula follows from the identities

$$\begin{aligned} \nu^{-1}(\alpha_i) &= \frac{(\alpha_i|\alpha_i)}{2} \alpha_i^\vee \\ \nu^{-1}(\alpha) &= \sum_i k_i \nu^{-1}(\alpha_i). \end{aligned}$$

because  $\alpha$  is a  $\mathbb{Z}$ -linear combination of simple roots  $\alpha_i$  and apply  $\nu^{-1}$  to both sides.

Finally, suppose the contrary to e); we may assume that  $\alpha > 0$ . But then  $-\alpha \in C^\vee$ , where we recall that the dual cone  $C^\vee$  is defined by

$$C^\vee = \{\lambda \in \mathfrak{h}^* \mid \langle \lambda, \alpha_i^\vee \rangle \leq 0 \text{ for all } i\}$$

In particular, if  $\alpha > 0$ , then  $-\alpha$  is a nontrivial nonnegative linear combination of simple roots, and therefore pairs nonnegatively with all simple coroots, because the Cartan matrix has nonpositive off-diagonal entries.

By Proposition 3.20 (d) for the dual root system,  $-\alpha + w(\alpha) \geq 0$  for any  $w \in W$ . Taking  $w$  such that  $w(\alpha) \in \Pi$ , we arrive at a contradiction.  $\square$

**Lemma 5.3.** Suppose that  $A$  is symmetrizable. Then the set of all  $\alpha = \sum_i k_i \alpha_i \in Q$  such that

$$k_i(\alpha_i|\alpha_i) \in (\alpha|\alpha)\mathbb{Z} \quad \text{for all } i \quad (25)$$

is  $W$ -invariant.

*Proof.* It suffices to check that  $r_i(\alpha)$  again satisfies 25, i.e. that

$$(k_i - \langle \alpha, \alpha_i^\vee \rangle)(\alpha_i|\alpha_i) \in (\alpha|\alpha)\mathbb{Z}.$$

This is equivalent to

$$2(\alpha|\alpha_i) \in (\alpha|\alpha)\mathbb{Z} \quad (26)$$

which follows from 25

$$2(\alpha|\alpha_i) = \sum_j \frac{2(\alpha_j|\alpha_i)}{(\alpha_j|\alpha_j)} k_j(\alpha_j|\alpha_j) = \sum_j a_{ji} k_j(\alpha_j|\alpha_j) \in (\alpha|\alpha)\mathbb{Z}.$$

as desired.  $\square$

Let  $A$  be a symmetrizable generalized Cartan matrix, and let  $(\cdot|\cdot)$  be a standard invariant bilinear form. Then, given a real root  $\alpha$  we have  $(\alpha|\alpha) = (\alpha_i|\alpha_i)$  for some simple root  $\alpha_i$ .

**Definition 5.4.** We call  $\alpha$  a **short** (resp. **long**) real root if  $(\alpha|\alpha) = \min_i (\alpha_i|\alpha_i)$  (resp.  $\max_i (\alpha_i|\alpha_i)$ ). These are independent of the choice of a standard form (because any two standard forms are proportional).

Note that if  $A$  is symmetric, then all simple roots and hence all real roots have the same square length. If  $A$  is not symmetric and  $S(A)$  is equipped with  $m$  arrows pointing in the same direction, then there are simple roots of exactly  $m + 1$  different square lengths since an arrow in  $S(A)$  points to a shorter simple root. It follows that if  $A$  is a nonsymmetric matrix from Table Fin, then every root is either short or long. Furthermore, if  $A$  is a nonsymmetric matrix from Table Aff, and  $A$  is not of type  $A_{2\ell}^{(2)}$  with  $\ell > 1$ , then every real root is either short or long; for the type  $A_{2\ell}^{(2)}$  with  $\ell > 1$  there are real roots of three different lengths.

**Remark 5.5 (Short and long roots).** Recall that a root system is called **simply laced** if all roots have the same length. Equivalently

1. The Cartan matrix has all off-diagonal entries either 0 or -1.
2. The Dynkin diagrams have only single edges (no double or triple bonds).
3. If we restrict to finite or affine types, the only simply laced types are  $A_n$ ,  $D_n$ ,  $E_6$ ,  $E_7$ ,  $E_8$  and their affine versions  $A_n^{(1)}$ ,  $D_n^{(1)}$ ,  $E_6^{(1)}$ ,  $E_7^{(1)}$ ,  $E_8^{(1)}$ .

## 5.2 Weyl group orbits of imaginary roots

In this section, we shall normalize  $(\cdot|\cdot)$  such that the  $(\alpha_i|\alpha_i)$  are relatively prime positive integers for each connected component of  $S(A)$ . For example, if  $A$  is symmetric, then  $(\alpha_i|\alpha_i) = 1$  for all  $i$ , and all real roots are short (= long).

Denote by  $\Delta^{\text{im}}$  and  $\Delta_+^{\text{im}}$  the sets of imaginary and positive imaginary roots, respectively. By definition,

$$\Delta = \Delta^{\text{re}} \cup \Delta^{\text{im}} \quad (\text{disjoint union}).$$

It is also clear that

$$\Delta^{\text{im}} = \Delta_+^{\text{im}} \cup (-\Delta_+^{\text{im}}).$$

The following properties of imaginary roots are useful.

### Proposition 5.6 (Properties of imaginary roots).

- a) The set  $\Delta_+^{\text{im}}$  is  $W$ -invariant.
- b) For  $\alpha \in \Delta_+^{\text{im}}$  there exists a unique root  $\beta \in -C^\vee$  (i.e.  $\langle \beta, \alpha_i^\vee \rangle \leq 0$  for all  $i$ ) which is  $W$ -equivalent to  $\alpha$ .
- c) If  $A$  is symmetrizable and  $(\cdot|\cdot)$  is a standard invariant bilinear form, then a root  $\alpha$  is imaginary if and only if  $(\alpha|\alpha) \leq 0$ .

*Proof.* As  $\Delta_+^{\text{im}} \subset \Delta_+ \setminus \Pi$  and the set  $\Delta_+ \setminus \{\alpha_i\}$  is  $r_i$ -invariant, it follows that  $\Delta_+^{\text{im}}$  is  $r_i$ -invariant for all  $i$  and hence  $W$ -invariant, proving a). Let  $\alpha \in \Delta_+^{\text{im}}$  and let  $\beta$  be an element of minimal height in  $W \cdot \alpha \subset \Delta_+$ . Then  $-\beta \in C^\vee$ . Such a  $\beta$  is unique in the orbit  $W \cdot \alpha$  because the fundamental chamber  $C$  is a fundamental domain for the action of  $W$  on  $X$ , proving b). Let  $\alpha$  be an imaginary root; we may assume, by b), that  $-\alpha \in C^\vee$  (since  $(\cdot|\cdot)$  is  $W$ -invariant). Let  $\alpha = \sum_i k_i \alpha_i$ ,  $k_i \geq 0$ ; then

$$(\alpha|\alpha) = \sum_i k_i (\alpha|\alpha_i) = \sum_i \frac{1}{2} |\alpha_i|^2 k_i \langle \alpha, \alpha_i^\vee \rangle \leq 0$$

The converse holds by Proposition 5.2 (d).  $\square$

**Definition 5.7.** For  $\alpha = \sum_i k_i \alpha_i \in Q$  we define the **support** of  $\alpha$  (written  $\text{supp } \alpha$ ) to be the subdiagram of  $S(A)$  which consists of the vertices  $i$  such that  $k_i \neq 0$ , and of all the edges joining these vertices.

By the following lemma,  $\text{supp } \alpha$  is connected for every root  $\alpha$ .

**Lemma 5.8.** Let  $I_1, I_2 \subset \{1, \dots, n\}$  be disjoint subsets such that  $a_{ij} = a_{ji} = 0$  whenever  $i \in I_1$ ,

$j \in I_2$ . Let

$$\beta_s = \sum_{i \in I_s} k_i^{(s)} \alpha_i, \quad (s = 1, 2).$$

Suppose that  $\alpha = \beta_1 + \beta_2$  is a root of the Lie algebra  $\mathfrak{g}(A)$ . Then either  $\beta_1$  or  $\beta_2$  is zero.

*Proof.* If the submatrices of  $I_1$  and  $I_2$  have no edges, the corresponding subalgebras  $\mathfrak{g}^{(1)}$  and  $\mathfrak{g}^{(2)}$  commute. Then  $\mathfrak{g}_\alpha$ , the root space for  $\alpha$ , would have to lie entirely inside one of them.  $\square$

**Example 5.9.** In  $A_4$ , the simple roots are  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ , corresponding to the superdiagonal entries of  $5 \times 5$  traceless matrices. The positive roots correspond to the upper triangular entries, and one sees that to write a root as a sum of simple roots, one takes the sum of the simple roots corresponding to the entries in the path from one node to another. Thus indeed the support of any root is connected.

Set:

$$K = \{\alpha \in Q_+ \setminus \{0\} \mid \langle \alpha, \alpha_i^\vee \rangle \leq 0 \text{ for all } i \text{ and } \text{supp } \alpha \text{ is connected}\}.$$

**Lemma 5.10 (Key lemma on imaginary roots).** In the above notation,  $K \subset \Delta_+^{\text{im}}$ .

*Proof.* Let  $\alpha = \sum_i k_i \alpha_i \in K$ . Set

$$\Omega_\alpha = \{\gamma \in \Delta_+ \mid \gamma \leq \alpha\}.$$

The set  $\Omega_\alpha$  is finite, and it is nonempty because the simple roots which appear in the decomposition of  $\alpha$ , lie in  $\Omega_\alpha$ . Let  $\beta = \sum_i m_i \alpha_i$  be an element of maximal height in  $\Omega_\alpha$ . It follows from Corollary 3.10 that

$$\text{supp } \beta = \text{supp } \alpha.$$

This is because by construction. Suppose that  $\text{supp } \beta$  is a proper subset of  $\text{supp } \alpha$ .  $\beta$  was chosen as an element of maximal height in  $\Omega_\alpha$ , so  $\beta + \alpha_i$  cannot also be in  $\Omega_\alpha$ , i.e. it is not a root. Since  $\beta$  is a root and  $\beta + \alpha_i$  is not a root, we conclude:  $\langle \beta, \alpha_i^\vee \rangle \geq 0$ . But recall:  $i$  is in the support of  $\alpha$ , and by hypothesis on  $\alpha \in K$ , we know  $\langle \alpha, \alpha_i^\vee \rangle \leq 0$ . So we see that  $\langle \beta, \alpha_i^\vee \rangle$  is nonpositive, but we can add some positive multiple of  $\alpha_i$  to  $\beta$  to get  $\alpha$ , and the pairing with  $\alpha_i^\vee$  will decrease by twice that positive multiple, but at the same time it must remain nonnegative. This is a contradiction.

First, we prove that  $\alpha \in \Delta_+$ . Suppose the contrary; then  $\alpha \neq \beta$ . By definition:

$$\beta + \alpha_i \notin \Delta_+ \quad \text{if } k_i > m_i.$$

Let  $A_1$  be the principal submatrix of  $A$  corresponding to  $\text{supp } \alpha$ . If  $A_1$  is of finite type, then  $\{\alpha \in Q_+ \mid \langle \alpha, \alpha_i^\vee \rangle \leq 0 \text{ for all } i\} = \{0\}$  because  $A$  is positive definite, and there is nothing to prove. If  $A_1$  is not of finite type, then, by Proposition 4.9, we have

$$P := \{j \in \text{supp } \alpha \mid k_j = m_j\} \neq \emptyset.$$



Let  $R$  be a connected component of the subdiagram  $(\text{supp } \alpha) \setminus P$ . From the above and Corollary 3.6a) we deduce that

$$\langle \beta, \alpha_i^\vee \rangle \geq 0 \quad \text{if } i \in R.$$

Set  $\beta' = \sum_{i \in R} m_i \alpha_i$ . Since  $\text{supp } \alpha$  is connected, the previous relations imply

$$\langle \beta', \alpha_i^\vee \rangle \geq 0 \text{ if } i \in R, \quad \langle \beta', \alpha_j^\vee \rangle > 0 \text{ for some } j \in R.$$

Therefore, by Theorem 4.3, the diagram  $R$  is of finite type.

On the other hand, set

$$\alpha' = \sum_{i \in R} (k_i - m_i) \alpha_i.$$

Since  $\text{supp } \alpha'$  is a connected component of  $\text{supp}(\alpha - \beta)$ , we obtain that

$$\langle \alpha', \alpha_i^\vee \rangle = \langle \alpha - \beta, \alpha_i^\vee \rangle \quad \text{for } i \in R.$$

But  $\langle \alpha, \alpha_i^\vee \rangle \leq 0$  since  $\alpha \in K$ , hence by the above:

$$\langle \alpha', \alpha_i^\vee \rangle \leq 0 \quad \text{for } i \in R.$$

This contradicts the fact that  $R$  is of finite type. Thus, we have proved that  $\alpha \in \Delta_+$ .

But  $2\alpha$  also satisfies all the hypotheses of the lemma; hence  $2\alpha \in \Delta_+$  and  $\alpha \in \Delta_+^{\text{im}}$ .  $\square$

**Theorem 5.11** (Weyl group orbit of imaginary roots).

$$\Delta_+^{\text{im}} = \bigcup_{w \in W} w(K).$$

*Proof.* Lemma 5.3 and Proposition 5.2a) prove the inclusion  $\supset$ . The reverse inclusion holds by Proposition 5.2b) and by the fact that  $\text{supp } \alpha$  is connected for every root  $\alpha$  (by Lemma 1.6).  $\square$

**Proposition 5.12** (Multiples of imaginary roots). *If  $\alpha \in \Delta_+^{\text{im}}$  and  $r$  is a nonzero (rational) number such that  $r\alpha \in Q$ , then  $r\alpha \in \Delta^{\text{im}}$ . In particular,  $n\alpha \in \Delta^{\text{im}}$  if  $n \in \mathbb{Z} \setminus \{0\}$ .*

*Proof.* By Proposition 5.2b) we can assume that  $\alpha \in -C^\vee \cap Q_+$ . Since  $\alpha \in \Delta$ , it follows that  $\text{supp } \alpha$  is connected and hence  $\alpha \in K$ . Hence  $r\alpha \in K$  for  $r > 0$  and therefore, by Lemma 5.3,  $r\alpha \in \Delta^{\text{im}}$ .  $\square$

**Theorem 5.13** (Description of imaginary roots). *Let  $A$  be an indecomposable generalized Cartan matrix.*

a) *If  $A$  is of finite type, then the set  $\Delta^{\text{im}}$  is empty.*

b) If  $A$  is of affine type, then

$$\Delta_+^{\text{im}} = \{n\delta \mid n = 1, 2, \dots\},$$

where  $\delta = \sum_{i=0}^{\ell} a_i \alpha_i$ , and the  $a_i$  are the labels of  $S(A)$  in Table Aff.

c) If  $A$  is of indefinite type, then there exists a positive imaginary root  $\alpha = \sum_i k_i \alpha_i$  such that  $k_i > 0$  and  $\langle \alpha, \alpha_i^\vee \rangle < 0$  for all  $i = 1, \dots, n$ .

*Proof.* Recall Proposition 5.15.

$$\{\alpha \in Q_+ \mid \langle \alpha, \alpha_i^\vee \rangle \leq 0, i = 1, \dots, n\}$$

is zero for  $A$  of finite type, is equal to  $\mathbb{Z}\delta$  for  $A$  of affine type, and there exists  $\alpha = \sum_i k_i \alpha_i$  such that  $k_i > 0$  and  $\langle \alpha, \alpha_i^\vee \rangle < 0$  for all  $i$  if  $A$  is of indefinite type. The theorem now follows from Theorem 5.11.  $\square$

It follows from Proposition 4.7 that if  $\alpha$  is a **null-root**, i.e.  $\alpha \in \Delta(A)$  is such that  $(\alpha|\beta) = 0$ , then  $\text{supp } \alpha$  is a diagram of affine type which is a connected component of  $S(A)$  and  $\alpha = k\delta$ ,  $k \in \mathbb{Z}$ . We proceed to describe the isotropic roots.

**Proposition 5.14 (Isotropic roots).** *Let  $A$  be symmetrizable. A root  $\alpha$  is isotropic (i.e.  $(\alpha|\alpha) = 0$ ) if and only if it is  $W$ -equivalent to an imaginary root  $\beta$  such that  $\text{supp } \beta$  is a subdiagram of affine type of  $S(A)$  (then  $\beta = k\delta$ ).*

*Proof.* Let  $\alpha$  be an isotropic root and let  $(\cdot|\cdot)$  be a standard bilinear form. We can assume that  $\alpha > 0$ . Then  $\alpha \in \Delta_+^{\text{im}}$  by Proposition 5.1d), and  $\alpha$  is  $W$ -equivalent to an imaginary root  $\beta \in K$  such that  $\langle \beta, \alpha_i^\vee \rangle \leq 0$  for all  $i$ , by Proposition 5.2b). Let  $\beta = \sum_{i \in P} k_i \alpha_i$  and  $P = \text{supp } \beta$ . Then

$$(\beta|\beta) = \sum_{i \in P} k_i (\beta|\alpha_i) = 0,$$

where  $k_i > 0$  and

$$(\beta|\alpha_i) = \frac{1}{2} |\alpha_i|^2 \langle \beta, \alpha_i^\vee \rangle \leq 0 \quad \text{for } i \in P.$$

Therefore  $\langle \beta, \alpha_i^\vee \rangle = 0$  for all  $i \in P$ , and  $P$  is a diagram of affine type. Conversely, if  $\beta = k\delta$  is an imaginary root for a diagram of affine type, then

$$(\beta|\beta) = k^2 (\delta|\delta) = k^2 \sum_i a_i (\delta|\alpha_i) = 0,$$

since  $\langle \delta, \alpha_i^\vee \rangle = 0$  for all  $i$ .  $\square$

**Proposition 5.15 (Root lengths).** *Let  $A$  be a generalized Cartan matrix of finite, affine, or hyperbolic type. Then*

a) The set of all short real roots is

$$\{\alpha \in Q \mid |\alpha|^2 = a = \min_i |\alpha_i|^2\}.$$

b) The set of all real roots is

$$\left\{ \alpha = \sum_j k_j \alpha_j \in Q \mid |\alpha|^2 > 0 \text{ and } k_j \frac{|\alpha_j|^2}{|\alpha|^2} \in \mathbb{Z} \text{ for all } j \right\}.$$

c) The set of all imaginary roots is

$$\{\alpha \in Q \setminus \{0\} \mid |\alpha|^2 \leq 0\}.$$

d) If  $A$  is affine, then there exist roots of intermediate squared length  $m$  if and only if  $A = A_{2\ell}^{(2)}$ ,  $\ell \geq 2$ . The set of such roots coincides with

$$\{\alpha \in Q \mid |\alpha|^2 = m\}.$$

**Exercise 5.16.** Show that for  $\alpha \in \Delta^{\text{re}}$ , one has

$$[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathbb{C}\alpha^\vee$$

Show that if  $A$  is a nonsymmetrizable  $3 \times 3$  matrix and  $\alpha = \alpha_1 + \alpha_2 + \alpha_3$ , then  $\alpha$  is an imaginary root with  $\dim[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] > 1$ .

**Solution 5.17 (Incomplete).** For the first statement, pick  $w \in W$  with  $w\alpha = \alpha_i$ . Let  $e_i \in \mathfrak{g}_{\alpha_i}$ ,  $f_i \in \mathfrak{g}_{-\alpha_i}$  be Chevalley generators, so  $[e_i, f_i] = \alpha_i^\vee$ . Using a representative  $n_w$  of  $w$  in the Tits group, set

$$e_\alpha := \text{Ad}(n_w^{-1})e_i \in \mathfrak{g}_\alpha$$

$$f_{-\alpha} := \text{Ad}(n_w^{-1})f_i \in \mathfrak{g}_{-\alpha}$$

Then

$$[e_\alpha, f_{-\alpha}] = \text{Ad}(n_w^{-1})[e_i, f_i] = \text{Ad}(n_w^{-1})(\alpha_i^\vee) = (w^{-1}\alpha_i)^\vee = \alpha^\vee$$

Hence  $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$  contains  $\mathbb{C}\alpha^\vee$ . On the other hand,  $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{h}$  because for  $h \in \mathfrak{h}$

$$[h, [x, y]] = [[h, x], y] + [x, [h, y]] = \alpha(h)[x, y] - \alpha(h)[x, y] = 0$$

so  $h$  acts on  $[x, y]$  with eigenvalue 0, i.e.  $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{h}$ . Moreover the map  $\mathfrak{g}_\alpha \otimes \mathfrak{g}_{-\alpha} \xrightarrow{[\cdot, \cdot]} \mathfrak{h}$  has 1-dimensional domain, so its image is at most one-dimensional.

*I'm still not so clear on what characterizes the Cartan  $\mathfrak{h}$ . I know that we have real roots and imaginary roots. Is  $\mathfrak{h}$  precisely the weight zero space? But then what about imaginary roots? Can you characterize it as being generated by all the brackets  $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$ ?*

For the second statement, we need to check that  $\alpha$  is a root and  $\dim[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] > 1$ .

To see that  $\alpha$  is a root, we recall the construction of  $\mathfrak{g}(A)$ . Work inside the positive part  $\mathfrak{n}_+$  generated by  $e_1, e_2, e_3$ . Give the free Lie algebra on  $\{e_1, e_2, e_3\}$  the multidegree grading  $\deg e_i = \epsilon_i \in \mathbb{Z}_{\geq 0}^3$ . The Serre relators are

$$(\text{ad } e_i)^{1-a_{ij}}(e_j) = 0 \quad (i \neq j)$$

Each such relator has multidegree  $(1 - a_{ij})\epsilon_i + \epsilon_j$ , i.e. it uses at least two copies of some  $e_i$  (since  $1 - a_{ij} \geq 1$ , and it is  $> 1$  unless  $a_{ij} = 0$ ). *How to interpret the meaning of  $a_{ij} = 0$ ?* In particular, no Serre relator has multidegree  $\epsilon_1 + \epsilon_2 + \epsilon_3$ .

Therefore the natural map from the  $(1, 1, 1)$ -homogeneous piece of the free Lie algebra to  $\mathfrak{n}_+$  is injective: nothing at that multidegree is killed by the ideal generated by the Serre relations. But that homogeneous piece is spanned by the two independent brackets (Jacobi gives one relation among three):  $X_1 := [e_1, [e_2, e_3]]$ ,  $X_2 := [e_2, [e_1, e_3]]$  (and  $X_3 := [e_3, [e_1, e_2]] = -X_1 - X_2$ ). Thus  $0 \neq X_i \in \mathfrak{g}_\alpha$  and hence  $\alpha$  is a root.

Now we verify that  $\dim[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] > 1$ . Pick nonzero

$$X_1 = [e_1, [e_2, e_3]], \quad X_2 = [e_2, [e_1, e_3]] \in \mathfrak{g}_\alpha$$

and nonzero

$$Y_1 = [f_1, [f_2, f_3]], \quad Y_2 = [f_2, [f_1, f_3]] \in \mathfrak{g}_{-\alpha}.$$

Set  $H_i := [X_i, Y_i] \in \mathfrak{h}$  for  $i = 1, 2$ . We claim  $H_1, H_2$  are linearly independent, so the image  $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$  has dimension  $\geq 2$ . We look at their action on the  $\mathfrak{h}$ -weight spaces  $\mathfrak{g}_{\alpha_j}$  ( $j = 1, 2, 3$ ) and check that they are not proportional.

I don't know how to quickly compute these brackets

$$[[[e_1, [e_2, e_3]], [f_1, [f_2, f_3]]], e_1]$$

and I don't know how to use a computer to do it either. *Is there a better way to do the computation or to bypass the computation to do this problem?*

If  $A$  were symmetrizable, there would be positive numbers  $d_i$  with  $d_i a_{ij} = d_j a_{ji}$ , and the two rows above would be proportional once restricted to  $\ker \alpha = \{(t_1, t_2, t_3) \mid t_1 + t_2 + t_3 = 0\}$ . But in the nonsymmetrizable case, that proportionality fails: the two eigenvalue triples are not scalar multiples on  $\ker \alpha$ . Hence  $H_1, H_2$  act differently on the three simple root spaces, so they are linearly independent elements of  $\ker \alpha$ . Therefore  $\dim[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \geq 2$  as desired (modulo the proof of the claim).

**Exercise 5.18.** Show that  $\dim[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = 1$  for all roots  $\alpha \in \Delta$  if and only if  $A$  is symmetrizable.

**Solution 5.19.** For symmetrizable Kac-Moody algebras there exists a symmetric invariant nondegenerate bilinear form  $(\cdot, \cdot)$  on  $\mathfrak{g}$  that pairs  $\mathfrak{g}_\alpha$  and  $\mathfrak{g}_{-\alpha}$  nondegenerately, and an element  $\alpha^\vee \in \mathfrak{h}$  characterized by

$$(h, \alpha^\vee) = \alpha(h) \quad (\forall h \in \mathfrak{h}).$$

Take  $x \in \mathfrak{g}_\alpha$ ,  $y \in \mathfrak{g}_{-\alpha}$ . By invariance,

$$([x, y], h) = (x, [y, h]) = (x, \alpha(h)y) = \alpha(h)(x, y) = ((x, y)\alpha^\vee, h)$$

for all  $h \in \mathfrak{h}$ . Nondegeneracy on  $\mathfrak{h}$  gives

$$[x, y] = (x, y)\alpha^\vee \in \mathbb{C}\alpha^\vee.$$

So the image of the bracket  $\mathfrak{g}_\alpha \otimes \mathfrak{g}_{-\alpha} \rightarrow \mathfrak{h}$  is exactly one-dimensional (spanned by  $\alpha^\vee$ ), for every root  $\alpha$  (real or imaginary). Conversely, if  $A$  is not symmetrizable, then there exists a root  $\alpha$  with  $\dim[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] > 1$ .

Indeed, nonsymmetrizability first appears in rank 3. Choose a nonsymmetrizable  $3 \times 3$  principal submatrix  $A_J$  (possible whenever  $A$  is nonsymmetrizable), and let  $\mathfrak{g}_J \subset \mathfrak{g}$  be the Kac-Moody subalgebra generated by  $\{e_i, f_i, h_i\}_{i \in J}$ . In that rank-3 subsystem set

$$\alpha = \alpha_i + \alpha_j + \alpha_k \quad (J = \{i, j, k\}).$$

It follows from the previous exercise that  $\alpha$  is a root of  $\mathfrak{g}_J$  (hence of  $\mathfrak{g}$ ) and  $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] \subset \mathfrak{h}_J$  has dimension at least two.

Thus for nonsymmetrizable  $A$  there is at least one  $\alpha$  with  $\dim[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] > 1$ , contradicting the hypothesis. Therefore the hypothesis forces  $A$  to be symmetrizable.

**Exercise 5.20.** If  $\dim \mathfrak{g}(A) = \infty$ , then  $|\Delta^{\text{re}}| = \infty$ .

**Solution 5.21.** By Prop 4.8, if  $\mathfrak{g}(A)$  is infinite dimensional, then  $W$  is infinite. Suppose  $|\Delta^{\text{re}}| < \infty$ . Recall that the length of  $w$  is equal to the number of positive real roots that  $w$  sends to negative roots. Since there are only finitely many real roots, the length of  $w$  is bounded above by  $|\Delta^{\text{re}}|$ . This contradicts the well known fact that an infinite Coxeter group has elements of arbitrarily large length.

**Remark 5.22 (Orbits of the  $W$  action on roots).** Recall that in for finite semisimple Lie algebras, the Weyl group acts transitively on roots of the same length. To see this, it is enough to show that any positive root  $\beta$  is in the  $W$ -orbit of some simple root. Take any positive root  $\beta = \sum_i c_i \alpha_i$  (a linear combination of simple roots with nonnegative integers). Suppose  $\beta$  is not simple, then there exists some simple root  $\alpha_j$  with  $(\beta, \alpha_j) > 0$ .

This is true because  $\Phi$  is finite type so in particular the bilinear form  $(\cdot, \cdot)$  is positive definite. Suppose  $(\beta, \alpha_j) \leq 0$  for every simple root  $\alpha_j$ . Then

$$(\beta, \beta) = \left( \beta, \sum_j c_j \alpha_j \right) = \sum_j c_j (\beta, \alpha_j) \leq 0$$

because each coefficient  $c_j \geq 0$ . But  $(\beta, \beta) > 0$  since  $\beta$  is a nonzero root in a positive definite space. Contradiction. So at least one  $j$  has  $(\beta, \alpha_j) > 0$ .

Reflecting  $\beta$  across  $\alpha_j$ :

$$s_j(\beta) = \beta - \langle \beta, \alpha_j^\vee \rangle \alpha_j$$

gives a new root of the same length but smaller height. Iterating this process, you eventually land on a simple root  $\alpha_i$ .

In affine type, the Weyl group is transitive on real roots of a fixed length. However, I don't know how to show this fact without using the classification, and this exercise comes before Kac tells us what the affine Weyl groups are in each type. This immediately shows that there are infinitely many real roots since the Weyl group is infinite, and in fact shows there are infinitely many real roots of each length. *I was thinking about showing that the stabilizer of a real root is of infinite index, but I don't know how to identify the stabilizer of a real root in the affine Weyl group, nor do I know how to do it in the finite type case (from the Dynkin diagram say).*

**Exercise 5.23.** If  $A$  is a matrix of finite or affine type,  $\beta \in \Delta$ ,  $\alpha \in \Delta^{\text{re}}$ , then the string  $\beta + k\alpha$  with  $k \in \mathbb{Z}$  contains at most five roots. Show that if  $A$  is of indefinite type, then the number of roots in a string can be arbitrarily large.

**Solution 5.24.** Fix a real root  $\alpha$ . For any root  $\beta$ , there exist integers  $p, q \geq 0$  such that

$$\{\beta + k\alpha : k \in \mathbb{Z}\} \cap \Delta = \{\beta - p\alpha, \dots, \beta, \dots, \beta + q\alpha\},$$

and

$$p - q = \langle \beta, \alpha^\vee \rangle.$$

This is the root string property, proved by embedding the span of  $\{\alpha, \beta\}$  in an  $\mathfrak{sl}_2$ -subalgebra.

So the length of the  $\alpha$ -string through  $\beta$  is

$$p + q + 1 = |\langle \beta, \alpha^\vee \rangle| + 1.$$

The entire string lives in the rank-2 subsystem generated by  $\alpha, \beta$ . Its Cartan matrix is

$$\begin{pmatrix} 2 & -m \\ -n & 2 \end{pmatrix}, \quad m = -\langle \alpha, \beta^\vee \rangle, \quad n = -\langle \beta, \alpha^\vee \rangle.$$

If  $A$  is finite or affine, every rank-2 submatrix is finite or affine. This is because every principal submatrix of a positive definite or positive semidefinite matrix is positive definite or positive semidefinite. Note that the converse is not true. In rank 2, finite/affine means  $mn \leq 4$  (essentially by the classification of rank-2 Dynkin diagrams). So the possibilities are:

- $mn = 0$ : string length = 1

- $mn = 1$  (type  $A_2$ ): string length  $\leq 2$
- $mn = 2$  (type  $B_2$ ): string length  $\leq 3$
- $mn = 3$  (type  $G_2$ ): string length  $\leq 4$
- $mn = 4$  (rank-2 affines): string length  $\leq 5$

So in all finite or affine types, every root string has at most 5 elements. Conversely,  $A$  is indefinite, then some  $2 \times 2$  principal submatrix has  $mn > 4$ . For such a pair,  $|\langle \beta, \alpha^\vee \rangle| = n$ , so the string through  $\beta$  in the direction of  $\alpha$  has length  $n + 1$ . Since  $n$  can be arbitrarily large in indefinite type, the length of root strings is unbounded.

**Remark 5.25.** In general, if  $m$  and  $n$  are arbitrary positive integers, the Kac-Moody algebra associated to the matrix

$$\begin{pmatrix} 2 & -m \\ -n & 2 \end{pmatrix}$$

is indefinite if  $mn > 4$ , affine if  $mn = 4$ , and finite if  $mn < 4$ .

## 6 Affine Algebras: the Normalized Invariant Form, the Root System, and the Weyl Group

In this section, we unwind the structure of affine Kac-Moody algebras in more detail. In particular, we describe the root system and the Weyl group of an affine Kac-Moody algebra associated to a finite dimensional simple Lie algebra.

### 6.1 Distinguished elements

Let  $A$  be a generalized Cartan matrix of affine type of order  $\ell + 1$  (and rank  $\ell$ ), and let  $S(A)$  be its Dynkin diagram from Table Aff. Let  $a_0, a_1, \dots, a_\ell$  be the numerical labels of  $S(A)$  in Table Aff. Then  $a_0 = 1$  unless  $A$  is of type  $A_{2\ell}^{(2)}$ , in which case  $a_0 = 2$ .

We denote by  $a_i^\vee$  ( $i = 0, \dots, \ell$ ) the labels of the Dynkin diagram  $S({}^t A)$  of the dual algebra which is obtained from  $S(A)$  by reversing the directions of all arrows and keeping the same enumeration of vertices. Note that in all cases

$$a_0^\vee = 1.$$

**Definition 6.1 (Coxeter numbers).** The numbers

$$h = \sum_{i=0}^{\ell} a_i, \quad h^\vee = \sum_{i=0}^{\ell} a_i^\vee$$

are called, respectively, the **Coxeter number** and the **dual Coxeter number** of the matrix  $A$ . Another important number is  $r$ , the number of the Table Aff  $r$  containing  $A$ .

The matrix  $A$  is symmetrizable. Moreover, we have

$$A = \text{diag}(a_0(a_0^\vee)^{-1}, a_1(a_1^\vee)^{-1}, \dots, a_\ell(a_\ell^\vee)^{-1}) B, \quad B = {}^t B. \quad (27)$$

Indeed, let  $\delta = {}^t(a_0, \dots, a_\ell)$  and  $\delta^\vee = {}^t(a_0^\vee, \dots, a_\ell^\vee)$ . If  $A = DB$  where  $D$  is diagonal invertible and  $B = {}^t B$ , then  $B\delta = 0$  and hence  ${}^t\delta B = 0$ . On the other hand,  ${}^t\delta^\vee A = 0$  implies  ${}^t\delta^\vee DB = 0$ , and we use the fact that  $\dim \ker B = 1$ .

Let  $\mathfrak{g} = \mathfrak{g}(A)$  be the affine algebra associated to a matrix  $A$  of affine type from Table Aff  $r$ . Let  $\mathfrak{h}$  be its Cartan subalgebra,

$$\Pi = \{\alpha_0, \dots, \alpha_\ell\} \subset \mathfrak{h}^*$$

the set of simple roots,

$$\Pi^\vee = \{\alpha_0^\vee, \dots, \alpha_\ell^\vee\} \subset \mathfrak{h}$$

the set of simple coroots,  $\Delta$  the root system,  $Q$  and  $Q^\vee$  the root and coroot lattices, etc. The center of  $\mathfrak{g}(A)$  is 1-dimensional and is spanned by

$$K = \sum_{i=0}^{\ell} a_i^\vee \alpha_i^\vee.$$

The element  $K$  is called the **canonical central element**. It is annihilated by all the simple roots, because

$$\langle \alpha_j, K \rangle = \sum_{i=0}^{\ell} a_i^\vee \langle \alpha_j, \alpha_i^\vee \rangle = \sum_i a_i^\vee a_{ij} = 0 \quad \Rightarrow \quad K \in \ker A^\top$$

by the defining property of the labels  $a_i^\vee$ .

Recall the definition of the **null root**  $\delta$ :

$$\delta = \sum_{i=0}^{\ell} a_i \alpha_i \in Q.$$

It is dual to  $K$  in the sense that  $\delta$  is annihilated by all simple coroots:

$$\langle \delta, \alpha_i^\vee \rangle = 0 \quad (i = 0, \dots, \ell), \quad \langle \delta, K \rangle = 0.$$

Fix an element  $d \in \mathfrak{h}$  which satisfies the following conditions:

$$\langle \alpha_i, d \rangle = 0 \quad \text{for } i = 1, \dots, \ell; \quad \langle \alpha_0, d \rangle = 1.$$



(Such an element is defined up to a summand proportional to  $K$ , because the coefficients of  $K$  span the kernel of  $A^\top$ .) The element  $d$  is called the **scaling element**. It is clear that the elements  $\alpha_0^\vee, \dots, \alpha_\ell^\vee, d$  form a basis of  $\mathfrak{h}$ . This is because one cannot write  $d$  as a linear combination, say with coefficients  $c$  of the  $\alpha_i^\vee$  since the linear equation  $A^\top c = e_0$  has no solution. This is because  $A^\top c = y$  has a solution if and only if  $y$  is orthogonal to the kernel of  $A$ , which is spanned by the vector of labels  $a$ . Here  $e_0$  is the first standard basis vector, so we need  $a_0 = 0$  for  $e_0$  to be orthogonal to  $a$ , which is false. Once you know that  $d$  is not in the span of the  $\alpha_i^\vee$ , it is clear that  $\{\alpha_0^\vee, \dots, \alpha_\ell^\vee, d\}$  is a basis of  $\mathfrak{h}$  since  $\dim \mathfrak{h} = 2(\ell + 1) - \ell = \ell + 2$ .

Note that

$$\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] + \mathbb{C}d.$$

We define a nondegenerate symmetric bilinear  $\mathbb{C}$ -valued form  $(\cdot | \cdot)$  on  $\mathfrak{h}$  as follows:

$$\begin{cases} (\alpha_i^\vee | \alpha_j^\vee) = a_i a_j^{-1} a_{ij}, & (i, j = 0, \dots, \ell), \\ (\alpha_i^\vee | d) = 0, & (i = 1, \dots, \ell), \\ (\alpha_0^\vee | d) = a_0, \\ (d | d) = 0. \end{cases} \quad (28)$$

By Theorem 2.2 this form can be uniquely extended to a bilinear form  $(\cdot | \cdot)$  on the whole Lie algebra  $\mathfrak{g}$  such that all the properties described by this theorem hold. From now on we fix this form on  $\mathfrak{g}$ . This is, clearly, a standard form. We call it the **normalized invariant form**.

To describe the induced bilinear form on  $\mathfrak{h}^*$ , we define an element, the **0th fundamental weight**  $\Lambda_0 \in \mathfrak{h}^*$  by

$$\langle \Lambda_0, \alpha_i^\vee \rangle = \delta_{0i} \quad \text{for } i = 0, \dots, \ell; \quad \langle \Lambda_0, d \rangle = 0.$$

Then  $\{\alpha_0, \dots, \alpha_\ell, \Lambda_0\}$  is a basis of  $\mathfrak{h}^*$  and we have

$$\begin{cases} (\alpha_i | \alpha_j) = a_i^\vee a_j^{\vee-1} a_{ij}, & (i, j = 0, \dots, \ell), \\ (\alpha_i | \Lambda_0) = 0, & (i = 1, \dots, \ell), \\ (\alpha_0 | \Lambda_0) = a_0^{-1}, \\ (\Lambda_0 | \Lambda_0) = 0. \end{cases} \quad (29)$$

The map  $\nu : \mathfrak{h} \rightarrow \mathfrak{h}^*$  defined by  $(\cdot | \cdot)$  looks as follows:

$$a_i^\vee \nu(\alpha_i^\vee) = a_i \alpha_i, \quad \nu(K) = \delta, \quad \nu(d) = a_0 \Lambda_0. \quad (30)$$

We also record some other simple formulas:

$$\begin{aligned} (\delta \mid \alpha_i) &= 0, & (\delta \mid \delta) &= 0, & (\delta \mid \Lambda_0) &= 1, \\ (K \mid \alpha_i^\vee) &= 0, & (K \mid K) &= 0, & (K \mid d) &= a_0 \end{aligned}$$

for  $i = 0, \dots, \ell$ .

Denote by  $\mathfrak{h}$  (resp.  $\mathfrak{h}_{\mathbb{R}}$ ) the linear span over  $\mathbb{C}$  (resp.  $\mathbb{R}$ ) of  $\alpha_1^\vee, \dots, \alpha_\ell^\vee$ . The dual notions  $\mathfrak{h}^*$  and  $\mathfrak{h}_{\mathbb{R}}^*$  are defined similarly. Then we have an orthogonal direct sum of subspaces:

$$\mathfrak{h} = \mathring{\mathfrak{h}} \oplus (\mathbb{C}K + \mathbb{C}d), \quad \mathfrak{h}^* = \mathring{\mathfrak{h}}^* \oplus (\mathbb{C}\delta + \mathbb{C}\Lambda_0).$$

We set

$$\mathfrak{h}_{\mathbb{R}} = \mathring{\mathfrak{h}}_{\mathbb{R}} + \mathbb{R}K + \mathbb{R}d, \quad \mathfrak{h}_{\mathbb{R}}^* = \mathring{\mathfrak{h}}_{\mathbb{R}}^* + \mathbb{R}\Lambda_0 + \mathbb{R}\delta.$$

**Remark 6.2.** For the evaluation pairing  $\langle \cdot, \cdot \rangle : \mathfrak{h}^* \times \mathfrak{h} \rightarrow \mathbb{C}$ , we have the following orthogonality relations:

1.  $K$  is orthogonal to all simple roots  $\alpha_i$  and  $\delta$ , pairing only with  $\Lambda_0$ .
2. The derivation  $d$  is orthogonal to all simple roots  $\alpha_i$  except  $\alpha_0$ , pairing only with  $\alpha_0$  and  $\delta$ .
3. The null root  $\delta$  is orthogonal to all simple coroots  $\alpha_i^\vee$  and  $K$ , pairing only with  $d$ .
4.  $\Lambda_0$  is orthogonal to all simple coroots  $\alpha_i^\vee$  except  $\alpha_0^\vee$ , pairing only with  $K$  and  $\alpha_0^\vee$ .

	$\alpha_0$	$\alpha_j \ (j > 0)$	$\delta$	$\Lambda_0$	$\Lambda_j \ (j > 0)$
$\alpha_0^\vee$	1	$a_{0j}$	0	1	0
$\alpha_i^\vee \ (i > 0)$	$a_{i0}$	$a_{ij}$	0	0	$\delta_{ij}$
$K$	0	0	0	$a_0^\vee$	$a_j^\vee$
$d$	1	0	1	0	0

Note that the restriction of the bilinear form  $(\cdot \mid \cdot)$  to  $\mathring{\mathfrak{h}}_{\mathbb{R}}^*$  and  $\mathring{\mathfrak{h}}_{\mathbb{R}}$  (resp.  $\mathring{\mathfrak{h}}_{\mathbb{R}}^* + \mathbb{R}\delta$  and  $\mathring{\mathfrak{h}}_{\mathbb{R}} + \mathbb{R}K$ ) is positive-definite (resp. positive-semidefinite with kernels  $\mathbb{R}\delta$  and  $\mathbb{R}K$ ).

**Lemma 6.3 (Useful formulas).** For a subset  $S$  of  $\mathfrak{h}^*$  denote by  $\overline{S}$  the orthogonal projection of  $S$  on  $\mathring{\mathfrak{h}}^*$ . (This should not be confused with the sign of closure in metric topology.) Then we have the following useful formula for  $\lambda \in \mathfrak{h}^*$  such that  $\lambda(K) \neq 0$ :

$$\lambda - \overline{\lambda} = \langle \lambda, K \rangle \Lambda_0 + \langle 2\lambda, K \rangle^{-1} (|\lambda|^2 - |\overline{\lambda}|^2) \delta. \quad (6.2.6)$$

We also have another useful formula from follows from pairing both sides with  $\Lambda_0$ :

$$\lambda = \overline{\lambda} + \langle \lambda, K \rangle \Lambda_0 + (\lambda \mid \Lambda_0) \delta. \quad (31)$$

*Proof.* Indeed,  $\lambda - \bar{\lambda} = b_1\Lambda_0 + b_2\delta$ . Taking inner product with  $\delta$ , we obtain that

$$\begin{aligned} (b_1\Lambda_0 + b_2\delta \mid \delta) &= b_1(\Lambda_0 \mid \delta) + b_2(\delta \mid \delta) = b_1 \\ (\lambda - \bar{\lambda} \mid \delta) &= (\lambda \mid \delta) - (\bar{\lambda} \mid \delta) = (\lambda \mid \delta) = \langle \lambda, K \rangle \end{aligned}$$

because  $(\bar{\lambda} \mid \delta) = 0$  and  $(\lambda \mid \delta)$  picks out the coefficient of  $\Lambda_0$  in  $\lambda$ , which is the same as evaluating against  $K$ .

We also have the formula  $|\lambda|^2 = |\bar{\lambda}|^2 + 2b_1b_2$  so we are done.

$$\begin{aligned} |\lambda|^2 &= (\bar{\lambda} + b_1\Lambda_0 + b_2\delta \mid \bar{\lambda} + b_1\Lambda_0 + b_2\delta) \\ &= (\bar{\lambda} \mid \bar{\lambda}) + b_1^2(\Lambda_0 \mid \Lambda_0) + b_2^2(\delta \mid \delta) \\ &\quad + 2b_1(\bar{\lambda} \mid \Lambda_0) + 2b_2(\bar{\lambda} \mid \delta) + 2b_1b_2(\Lambda_0 \mid \delta) = |\bar{\lambda}|^2 + 2b_1b_2 \end{aligned}$$

□

Define  $\rho \in \mathfrak{h}^*$  by  $\langle \rho, \alpha_i^\vee \rangle = 1$  ( $i = 0, \dots, \ell$ ) and  $\langle \rho, d \rangle = 0$  (cf. §2.5). Then

$$\rho = \bar{\rho} + h^\vee \Lambda_0. \quad (32)$$

because

$$\langle \rho, K \rangle = \sum_{i=0}^{\ell} a_i^\vee \langle \rho, \alpha_i^\vee \rangle = \sum_i a_i^\vee = h^\vee$$

and  $(\rho \mid \Lambda_0) = 0$  because  $\nu(d) = a_0\Lambda_0$  and  $\langle \rho, d \rangle = 0$ .

## 6.2 The Root System of an Affine Kac-Moody Algebra

Denote by  $\mathring{\mathfrak{g}}$  the subalgebra of  $\mathfrak{g}$  generated by the  $e_i$  and  $f_i$  with  $i = 1, \dots, \ell$ . By the results of Chapter 1 it is clear that  $\mathring{\mathfrak{g}}$  is a Kac-Moody algebra associated to the matrix  $A^\circ$  obtained from  $A$  by deleting the 0th row and column. The elements  $e_i, f_i$  ( $i = 1, \dots, \ell$ ) are the Chevalley generators of  $\mathring{\mathfrak{g}}$ , and  $\mathring{\mathfrak{h}} = \mathring{\mathfrak{g}} \cap \mathfrak{h}$  is its Cartan subalgebra;

$$\mathring{\Pi} = \{\alpha_1, \dots, \alpha_\ell\} \quad \text{is the root basis,} \quad \mathring{\Pi}^\vee = \{\alpha_1^\vee, \dots, \alpha_\ell^\vee\} \quad \text{is the coroot basis for } \mathring{\mathfrak{g}}.$$

By Proposition 4.9,  $\mathring{\mathfrak{g}} = \mathfrak{g}(A^\circ)$  is a simple finite-dimensional Lie algebra whose Dynkin diagram  $S(A^\circ)$  is obtained from  $S(A)$  by removing the 0th vertex.

The set  $\mathring{\Delta} = \Delta \cap \mathring{\mathfrak{h}}^*$  is the root system of  $\mathring{\mathfrak{g}}$ ; it is finite and consists of real roots. The set  $\mathring{\Delta}_+ = \mathring{\Delta} \cap \Delta_+$  is the set of positive roots. Denote by  $\mathring{\Delta}_s$  and  $\mathring{\Delta}_\ell$  the sets of short and long roots, respectively, in  $\mathring{\Delta}$ . Put  $\mathring{Q} = \mathbb{Z}\mathring{\Delta}$ . Let  $\mathring{W}$  be the Weyl group of  $\mathring{\Delta}$ .

Recall that the sets of imaginary and positive imaginary roots of  $\mathfrak{g}$  are as follows:

$$\Delta^{\text{im}} = \{\pm\delta, \pm 2\delta, \dots\}, \quad \Delta_+^{\text{im}} = \{\delta, 2\delta, \dots\}.$$

**Proposition 6.4 (The Real Roots of an Affine Kac-Moody Algebra).** *Let  $r$  be the number of the Table Aff containing  $A$ , which should be interpreted as the ratio of the squared lengths of long and short roots in  $\mathring{\Delta}$  (so  $r = 1, 2$ , or  $3$ ). Then the set of real roots of  $\mathfrak{g}$  is given as follows:*

- a)  $\Delta^{\text{re}} = \{\alpha + n\delta \mid \alpha \in \mathring{\Delta}, n \in \mathbb{Z}\}$  if  $r = 1$ .
- b)  $\Delta^{\text{re}} = \{\alpha + n\delta \mid \alpha \in \mathring{\Delta}_s, n \in \mathbb{Z}\} \cup \{\alpha + nr\delta \mid \alpha \in \mathring{\Delta}_\ell, n \in \mathbb{Z}\}$  if  $r = 2$  or  $3$ , but  $A$  is not of type  $A_{2\ell}^{(2)}$ .
- c)  $\Delta^{\text{re}} = \left\{ \frac{1}{2}(\alpha + (2n-1)\delta) \mid \alpha \in \mathring{\Delta}_\ell, n \in \mathbb{Z} \right\} \cup \{\alpha + n\delta \mid \alpha \in \mathring{\Delta}_s, n \in \mathbb{Z}\} \cup \{\alpha + 2n\delta \mid \alpha \in \mathring{\Delta}_\ell, n \in \mathbb{Z}\}$  if  $A$  is of type  $A_{2\ell}^{(2)}$ .
- d)  $\Delta^{\text{re}} + r\delta = \Delta^{\text{re}}$
- e)  $\Delta_+^{\text{re}} = \{\alpha \in \Delta^{\text{re}} \mid n > 0\} \cup \mathring{\Delta}_+.$

*Proof.* Recall Proposition 5.15. We want to show:  $\Delta_s^{\text{re}} = \{\alpha + n\delta \mid \alpha \in \mathring{\Delta}_s, n \in \mathbb{Z}\}$  for  $A \neq A_{2\ell}^{(2)}$ . Write any real root  $\alpha$  as a linear combination  $\alpha = \sum_{i=0}^{\ell} k_i \alpha_i$ . If  $\alpha$  is short, then its squared length is the short length  $a$ . Since  $\delta = \alpha_0 + \alpha_1 + \dots$ , you can rewrite  $\alpha = (\alpha - k_0\delta) + k_0\delta$ . But  $\alpha - k_0\delta$  lives in the finite root lattice (only involves  $\alpha_1, \dots, \alpha_\ell$ ). Moreover, its squared length is still short. So  $\alpha - k_0\delta \in \mathring{\Delta}_s$ . Thus we have shown that

$$\Delta_s^{\text{re}} \subseteq \{\alpha + n\delta \mid \alpha \in \mathring{\Delta}_s, n \in \mathbb{Z}\}$$

if  $A \neq A_{2\ell}^{(2)}$ . The reverse inclusion is clear since  $\delta$  is imaginary and adding it to a real root preserves length and therefore by Proposition 5.15 preserves realness.

If  $r = 1$  then  $\alpha_0$  is also a long root and the same argument shows that

$$\Delta_\ell^{\text{re}} = \{\alpha + n\delta \mid \alpha \in \mathring{\Delta}_\ell, n \in \mathbb{Z}\}$$

If  $r = 2$  or  $3$ , recall that

$$\alpha = \sum_j k_j \alpha_j \in Q, \quad |\alpha|^2 > 0, \quad \text{and } k_j \frac{|\alpha_j|^2}{|\alpha|^2} \in \mathbb{Z} \quad \forall j.$$

In particular there is an integrality condition on  $j = 0$

$$k_0 \frac{|\alpha_0|^2}{|\alpha|^2} = k_0 \frac{a}{ra} \in \mathbb{Z}$$

where  $|\alpha^2| = ra$  since  $\alpha$  is long. Also  $\alpha_0$  is short because the arrow in the Dynkin diagram points towards  $\alpha_0$ . Thus  $k_0$  is a multiple of  $r$ . Repeating the same argument as before, we can write

$$\alpha = (\alpha - k_0\delta) + k_0\delta$$

where  $\alpha - k_0\delta \in \mathring{\Delta}_\ell$  and  $k_0$  is a multiple of  $r$ . Thus we have shown that

$$\Delta_\ell^{\text{re}} \subseteq \{\alpha + nr\delta \mid \alpha \in \mathring{\Delta}_\ell, n \in \mathbb{Z}\}$$

if  $A \neq A_{2\ell}^{(2)}$ . The reverse inclusion is again clear.  $\square$

Introduce the following important element:

$$\theta = \delta - a_0\alpha_0 = \sum_{i=1}^{\ell} a_i\alpha_i \in \mathring{Q}.$$

One checks that

$$\begin{aligned} |\theta|^2 &= (\delta - a_0\alpha_0 \mid \delta - a_0\alpha_0) \\ &= (\delta \mid \delta) - 2a_0(\delta \mid \alpha_0) + a_0^2(\alpha_0 \mid \alpha_0) \\ &= 0 - 0 + a_0^2 \cdot \frac{2a_0^\vee}{a_0} \\ &= 2a_0^\vee a_0 \end{aligned}$$

Since  $a_0^\vee = 1$  in affine types, this gives  $|\theta|^2 = 2a_0$ .

Hence  $|\theta|^2$  is equal to the square length of a long root if  $A$  is from Table Aff 1 or is of type  $A_{2\ell}^{(2)}$ , and it is equal to the square length of a short root otherwise. One deduces now from Proposition 5.10 a), b) that in all cases  $\theta \in \mathring{\Delta}_+$ . Again, from the formulas of §6.2 we deduce:

$$\theta = a_0\nu(\theta^\vee), \quad |\theta^\vee|^2 = 2a_0^{-1}, \quad \alpha_0^\vee = \nu^{-1}(\delta - \theta) = K - a_0\theta^\vee.$$

Furthermore, one has

**Proposition 6.5 (6.4).**

a) If  $A$  is from Table Aff 1 or is of type  $A_{2\ell}^{(2)}$ , then

$$\theta \in (\mathring{\Delta}_+)_\ell \quad \text{and} \quad \theta \text{ is the unique root in } \mathring{\Delta} \text{ of maximal height } (= h - a_0).$$

Note that in the finite dimensional case, the highest root is always long.

b) If  $A$  is from Table Aff 2 or 3 and is not of type  $A_{2\ell}^{(2)}$ , then

$$\theta \in (\overset{\circ}{\Delta}_+)_s \quad \text{and} \quad \theta \text{ is the unique root in } \overset{\circ}{\Delta}_s \text{ of maximal height } (= h - 1).$$

*Proof.* It is easy to check that all simple roots of the same square length in  $\overset{\circ}{\Delta}$  are  $W^\circ$ -equivalent, hence both  $\overset{\circ}{\Delta}_s$  and  $\overset{\circ}{\Delta}_\ell$  are orbits of  $W^\circ$ . Also,

$$\langle \theta, \alpha_i^\vee \rangle = -a_0 \langle \alpha_0, \alpha_i^\vee \rangle \geq 0, \quad \text{for all } i = 1, \dots, \ell.$$

Using Proposition 3.9 we see that  $\theta + \alpha_i \notin \overset{\circ}{\Delta}$  for all  $i = 1, \dots, \ell$ . This implies that  $\theta$  is the unique root of maximal height in its  $W^\circ$ -orbit. Note that  $\theta$  lives in the fundamental co-chamber because  $\langle \theta, \alpha_i^\vee \rangle \geq 0$  for all  $i = 1, \dots, \ell$ , and that the fundamental co-chamber is a fundamental domain for the action of  $W^\circ$ . This proves uniqueness.  $\square$

Unless otherwise stated, in the case of a finite type matrix  $A$ , we shall normalize the standard invariant form  $(\cdot | \cdot)$  on  $\mathfrak{g}(A)$  by the condition

$$(\alpha | \alpha) = 2 \quad \text{if } \alpha \in \overset{\circ}{\Delta}_\ell, \quad (6.4.2)$$

and shall call it the **normalized invariant form**. We deduce from (6.4.1) and Proposition 6.4 the following:

**Corollary 6.6 (6.4).** *Let  $\mathfrak{g}$  be an affine algebra of type  $X_N^{(r)}$ . Then the ratio of the normalized invariant form of  $\mathfrak{g}$  restricted to  $\overset{\circ}{\mathfrak{g}}$  to the normalized invariant form of  $\overset{\circ}{\mathfrak{g}}$  is equal to  $r$ .*

Note that we have the following description of  $\Pi$  and  $\Pi^\vee$ :

$$\Pi = \{\alpha_0 = a_0^{-1}(\delta - \theta), \alpha_1, \dots, \alpha_\ell\}, \quad \Pi^\vee = \{\alpha_0^\vee = K - a_0\theta^\vee, \alpha_1^\vee, \dots, \alpha_\ell^\vee\}.$$

### 6.3 The Affine Weyl Group

Now we turn to the description of the Weyl group  $W$  of the affine algebra  $\mathfrak{g}$ . Recall that  $W$  is generated by fundamental reflections  $r_0, r_1, \dots, r_\ell$ , which act on  $\mathfrak{h}^*$  by

$$r_i(\lambda) = \lambda - \langle \lambda, \alpha_i^\vee \rangle \alpha_i, \quad \lambda \in \mathfrak{h}^*.$$

As  $\langle \delta, \alpha_i^\vee \rangle = 0$  ( $i = 0, \dots, \ell$ ), we have

$$w(\delta) = \delta \quad \text{for all } w \in W.$$

Recall also that the invariant standard form is  $W$ -invariant. Denote by  $\overset{\circ}{W}$  the subgroup of  $W$  generated by  $r_1, \dots, r_\ell$ . As  $r_i(\Lambda_0) = \Lambda_0$  for  $i = 1, \dots, \ell$ , we deduce that  $\overset{\circ}{W}$  operates trivially

on  $\mathbb{C}\Lambda_0 + \mathbb{C}\delta$ ; it is also clear that  $\mathfrak{h}^{\circ*}$  is  $\overset{\circ}{W}$ -invariant. We conclude that  $\overset{\circ}{W}$  operates faithfully on  $\mathfrak{h}^{\circ*}$ , and we can identify  $\overset{\circ}{W}$  with the Weyl group of the Lie algebra  $\overset{\circ}{\mathfrak{g}}$ , operating on  $\mathfrak{h}^{\circ*}$ . Hence (by Proposition 3.12 e)) the group  $\overset{\circ}{W}$  is finite.

Recall that for a real root  $\alpha$  we have a reflection  $r_\alpha \in W$  defined by

$$r_\alpha(\lambda) = \lambda - \langle \lambda, \alpha^\vee \rangle \alpha, \quad \lambda \in \mathfrak{h}^*.$$

**Lemma 6.7 (6.5).** *Let  $\alpha \in \Delta_+^{\text{re}}$  be such that  $\beta := \delta - \alpha \in \Delta_+^{\text{re}}$  for some  $\alpha$ . Then*

$$r_\alpha r_\beta(\lambda) = \lambda + \langle \lambda, K \rangle \nu(\beta^\vee) - (\langle \lambda, \beta^\vee \rangle + \frac{1}{2} |\beta^\vee|^2 \langle \lambda, K \rangle) \delta, \quad \lambda \in \mathfrak{h}^*.$$

*Proof.* Recall that for a real root  $\gamma$ , the reflection on  $\mathfrak{h}^*$  is

$$r_\gamma(\lambda) = \lambda - \langle \lambda, \gamma^\vee \rangle \gamma, \quad \gamma^\vee = \frac{2\nu^{-1}(\gamma)}{(\gamma|\gamma)}, \quad \nu : \mathfrak{h} \xrightarrow{\sim} \mathfrak{h}^*$$

Equivalently

$$\langle \lambda, \gamma^\vee \rangle = \frac{2(\lambda|\gamma)}{(\gamma|\gamma)}$$

In the lemma  $\beta = \delta - \alpha$  with  $\alpha \in \Delta_{\text{re}}^+$  and  $\beta \in \Delta_{\text{re}}^+$ . Using the affine form:

$$(\delta|\delta) = 0, \quad (\delta|\alpha) = 0, \quad (\beta|\beta) = (\alpha|\alpha)$$

The null/central pair:

$$\nu(K) = \delta, \quad (\lambda|\delta) = \langle \lambda, K \rangle$$

Also

$$\nu(\beta^\vee) = \frac{2\beta}{(\beta|\beta)} = \frac{2(\delta - \alpha)}{(\alpha|\alpha)} \equiv -\frac{2\alpha}{(\alpha|\alpha)} \pmod{\mathbb{C}\delta}$$

Start from the reflection formula:

$$r_\beta(\lambda) = \lambda - \langle \lambda, \beta^\vee \rangle \beta = \lambda - \frac{2(\lambda|\beta)}{(\beta|\beta)} \beta.$$

Use  $\beta = \delta - \alpha$ ,  $(\lambda|\beta) = (\lambda|\delta) - (\lambda|\alpha) = \langle \lambda, K \rangle - (\lambda|\alpha)$ , and  $(\beta|\beta) = (\alpha|\alpha)$ :

$$\langle \lambda, \beta^\vee \rangle = \frac{2(\lambda|\beta)}{(\alpha|\alpha)} = \frac{2\langle \lambda, K \rangle}{(\alpha|\alpha)} - \langle \lambda, \alpha^\vee \rangle.$$

Hence

$$r_\beta(\lambda) = \lambda - \left( \frac{2\langle \lambda, K \rangle}{(\alpha|\alpha)} - \langle \lambda, \alpha^\vee \rangle \right) (\delta - \alpha).$$

So observing that  $(\lambda|\delta) = \langle \lambda, K \rangle$ , we get that

$$r_\beta(\lambda) \cong \lambda + (\lambda|\delta) \frac{2\alpha}{(\alpha|\alpha)} - \langle \lambda, \alpha^\vee \rangle \alpha \pmod{\mathbb{C}\delta}. \quad (33)$$

First, we compute  $\text{mod } \mathbb{C}\delta$ :

$$r_\alpha r_\beta(\lambda) \text{ mod } \mathbb{C}\delta = r_\alpha(\lambda + (\lambda|\delta) 2\alpha|\alpha|^{-2} - \langle \lambda, \alpha^\vee \rangle \alpha) \text{ mod } \mathbb{C}\delta$$

Now apply  $r_\alpha$  to each term:

$$\begin{aligned} r_\alpha(\lambda) &= \lambda - \langle \lambda, \alpha^\vee \rangle \alpha \\ r_\alpha((\lambda|\delta) \frac{2\alpha}{|\alpha|^2}) &= (\lambda|\delta) r_\alpha(\frac{2\alpha}{|\alpha|^2}) = (\lambda|\delta) \left( -\frac{2\alpha}{|\alpha|^2} \right) \equiv \langle \lambda, K \rangle \nu(\beta^\vee) \pmod{\mathbb{C}\delta} \\ r_\alpha(-\langle \lambda, \alpha^\vee \rangle \alpha) &= -\langle \lambda, \alpha^\vee \rangle r_\alpha(\alpha) = +\langle \lambda, \alpha^\vee \rangle \alpha \end{aligned}$$

Adding these three terms, we get

$$r_\alpha r_\beta(\lambda) \text{ mod } \mathbb{C}\delta = \lambda + \langle \lambda, K \rangle \nu(\beta^\vee) \pmod{\mathbb{C}\delta}.$$

Assume

$$r_\alpha r_\beta(\lambda) = \lambda + \langle \lambda, K \rangle \nu(\beta^\vee) + c(\lambda) \delta.$$

Pair both sides with  $d$  (recall  $\langle \delta, d \rangle = 1$ ):

$$\langle r_\alpha r_\beta(\lambda), d \rangle = \langle \lambda, d \rangle + \langle \lambda, K \rangle \underbrace{\langle \nu(\beta^\vee), d \rangle}_{(*)} + c(\lambda).$$

So to get  $c(\lambda)$  we need  $\langle r_\alpha r_\beta(\lambda), d \rangle$  and  $(*)$ . Use the reflection formula  $r_\gamma(\mu) = \mu - \langle \mu, \gamma^\vee \rangle \gamma$  and then pair with  $d$ :

$$\langle r_\gamma(\mu), d \rangle = \langle \mu, d \rangle - \langle \mu, \gamma^\vee \rangle \langle \gamma, d \rangle.$$

Apply first  $\gamma = \beta$  to  $\mu = \lambda$ :

$$\langle r_\beta(\lambda), d \rangle = \langle \lambda, d \rangle - \langle \lambda, \beta^\vee \rangle \underbrace{\langle \beta, d \rangle}_{=\langle \delta - \alpha, d \rangle = 1 - 0 = 1} = \langle \lambda, d \rangle - \langle \lambda, \beta^\vee \rangle.$$

Now apply  $\gamma = \alpha$ ,  $\mu = r_\beta(\lambda)$ ; since  $\langle \alpha, d \rangle = 0$ , the second term vanishes:

$$\langle r_\alpha r_\beta(\lambda), d \rangle = \langle r_\beta(\lambda), d \rangle = \langle \lambda, d \rangle - \langle \lambda, \beta^\vee \rangle.$$

Now we compute the second term  $(*)$ : By definition of  $\nu$ ,  $\langle \nu(h), d \rangle = (h|d)$  for  $h \in \mathfrak{h}$ . So  $\langle \nu(\beta^\vee), d \rangle = (\beta^\vee|d)$ . Using  $\nu(\beta^\vee) = \frac{2\beta}{(\beta|\beta)} = \frac{2(\delta-\alpha)}{(\alpha|\alpha)}$  and  $\langle \delta, d \rangle = 1$ ,  $\langle \alpha, d \rangle = 0$ :

$$\langle \nu(\beta^\vee), d \rangle = \frac{2}{(\alpha|\alpha)}.$$



It's often written in the compact form

$$\langle \nu(\beta^\vee), d \rangle = \frac{1}{2} |\beta^\vee|^2$$

because (for any real root)  $|\beta^\vee|^2 = (\beta^\vee | \beta^\vee) = \frac{4}{(\beta | \beta)} = \frac{4}{(\alpha | \alpha)}$  using  $\langle \beta, \beta^\vee \rangle = 2$ .

This follows from the definition:

$$|\beta^\vee|^2 = \left( \frac{2\nu^{-1}(\beta)}{(\beta | \beta)} \mid \frac{2\nu^{-1}(\beta)}{(\beta | \beta)} \right)$$

But  $(\nu^{-1}(\beta) | \nu^{-1}(\beta)) = (\beta | \beta)$ , because  $\nu$  is the identification of  $\mathfrak{h}$  with  $\mathfrak{h}^*$  via the bilinear form.

So:

$$|\beta^\vee|^2 = \frac{4}{(\beta | \beta)}$$

Since  $\beta = \delta - \alpha$ , we have  $(\delta | \delta) = 0$  and  $(\delta | \alpha) = 0$ , so

$$(\beta | \beta) = (\delta - \alpha | \delta - \alpha) = (\alpha | \alpha).$$

Therefore,

$$|\beta^\vee|^2 = \frac{4}{(\alpha | \alpha)}.$$

Hence  $\frac{1}{2} |\beta^\vee|^2 = 2/(\alpha | \alpha)$ , matching the computation above and the formula for  $c(\lambda)$  follows.  $\square$

Motivated by this formula, we introduce the following endomorphism  $t_\alpha$  of the vector space  $\mathfrak{h}^*$  for  $\alpha \in \mathfrak{h}^*$ :

$$t_\alpha(\lambda) = \lambda + \langle \lambda, K \rangle \alpha - ((\lambda | \alpha) + \frac{1}{2} |\alpha|^2 \langle \lambda, K \rangle) \delta. \quad (34)$$

In the case when  $m := \langle \lambda, K \rangle \neq 0$  we can rewrite this as follows using the formulas in Lemma 6.3

$$t_\alpha(\lambda) = m\Lambda_0 + (\bar{\lambda} + m\alpha) + \frac{1}{2m} (|\lambda|^2 - |\bar{\lambda} + m\alpha|^2) \delta. \quad (35)$$

Letting  $\lambda = \Lambda_0$  so that  $m = 1$  and  $|\lambda|^2 = 0$ , we get

$$t_\alpha(\Lambda_0) = \Lambda_0 + \alpha - \frac{1}{2} |\alpha|^2 \delta. \quad (36)$$

Note also that (34) implies

$$t_\alpha(\lambda) = \lambda - (\lambda | \alpha) \delta, \quad \text{if } \langle \lambda, K \rangle = 0. \quad (37)$$

**Proposition 6.8 (Additivity of  $t_\alpha$ ).** *For all  $\alpha, \beta \in \mathfrak{h}^*$  we have*

$$t_\alpha t_\beta = t_{\alpha+\beta}. \quad (38)$$

*Proof.* We want to check that  $t_\alpha t_\beta = t_{\alpha+\beta}$ . Pick any  $\lambda$  and let  $m = \langle \lambda, K \rangle$ .

First apply  $t_\beta$ :

$$t_\beta(\lambda) = \lambda + m\beta - \left( (\lambda|\beta) + \frac{1}{2}m|\beta|^2 \right) \delta.$$

Now apply  $t_\alpha$ . Since  $\langle t_\beta(\lambda), K \rangle = \langle \lambda, K \rangle = m$ , the level stays the same:

$$t_\alpha t_\beta(\lambda) = t_\beta(\lambda) + m\alpha - \left( (t_\beta(\lambda)|\alpha) + \frac{1}{2}m|\alpha|^2 \right) \delta.$$

Expand the  $(t_\beta(\lambda)|\alpha)$  term. Since the  $\delta$ -term in  $t_\beta(\lambda)$  drops out (as  $(\delta|\alpha) = 0$ ), we have:

$$(t_\beta(\lambda)|\alpha) = (\lambda|\alpha) + m(\beta|\alpha).$$

Therefore:

$$t_\alpha t_\beta(\lambda) = \lambda + m(\alpha + \beta) - \left( (\lambda|\alpha) + (\lambda|\beta) + m(\alpha|\beta) + \frac{1}{2}m(|\alpha|^2 + |\beta|^2) \right) \delta.$$

On the other hand:

$$t_{\alpha+\beta}(\lambda) = \lambda + m(\alpha + \beta) - \left( (\lambda|\alpha + \beta) + \frac{1}{2}m|\alpha + \beta|^2 \right) \delta.$$

We have  $(\lambda|\alpha + \beta) = (\lambda|\alpha) + (\lambda|\beta)$  and  $|\alpha + \beta|^2 = |\alpha|^2 + |\beta|^2 + 2(\alpha|\beta)$ . Therefore:

$$\frac{1}{2}m|\alpha + \beta|^2 = \frac{1}{2}m(|\alpha|^2 + |\beta|^2) + m(\alpha|\beta).$$

Plugging this in, we see the two formulas agree exactly.  $\square$

**Remark 6.9.** Kac says that it is enough to just check linearity on the level of  $\Lambda_0$  but I didn't understand why. It turns out that one can check it directly without too much trouble.

We also have the compatibility with the finite Weyl group:

$$t_{w(\alpha)} = wt_\alpha w^{-1}, \quad w \in \overset{\circ}{W}. \quad (39)$$

Indeed,

$$wt_\alpha(w^{-1}(\lambda)) = w(w^{-1}(\lambda) + \langle w^{-1}(\lambda), K \rangle \alpha - ((w^{-1}(\lambda)|\alpha) + \frac{1}{2}|\alpha|^2 \langle w^{-1}(\lambda), K \rangle) \delta),$$

Since  $w(w^{-1}(\lambda)) = \lambda$  and since  $K$  and  $\delta$  are fixed by the finite Weyl group  $w$  (i.e.  $w(K) = K$ ,  $w(\delta) = \delta$ ), and since  $w$  preserves the bilinear form  $(\cdot|\cdot)$ , we have:

$$wt_\alpha(w^{-1}(\lambda)) = \lambda + \langle \lambda, K \rangle w(\alpha) - \left( (\lambda|w(\alpha)) + \frac{1}{2}|\alpha|^2 \langle \lambda, K \rangle \right) \delta.$$

This is exactly  $t_{w(\alpha)}(\lambda)$ .

Now we introduce the following important lattice  $M \subset \mathfrak{h}^*$ . Let  $\mathbb{Z}(\overset{\circ}{W} \cdot \theta^\vee)$  denote the lattice in  $\mathfrak{h}^*$  spanned over  $\mathbb{Z}$  by the (finite) set  $\overset{\circ}{W} \cdot \theta^\vee$ , and set

$$M = \nu(\mathbb{Z}(\overset{\circ}{W} \cdot \theta^\vee)).$$

Here is a description of the lattice  $M$ :

$$M = \overline{Q} = \overset{\circ}{Q} \quad \text{if } A \text{ is symmetric or } r > a_0; \quad M = \nu(\overline{Q}^\vee) = \nu(\overset{\circ}{Q}^\vee) \quad \text{otherwise.} \quad (40)$$

The point is that  $\nu$  rescales directions differently depending on whether they come from long or short roots. In simply laced types (i.e. symmetric  $A$ ) all roots have the same length, so  $\nu$  acts as a uniform scaling on the root lattice  $\overset{\circ}{Q}$ .

Indeed, if  $r = 1$ , then  $\theta^\vee$  is a short root of  $\overset{\circ}{\Delta}^\vee$ , hence

$$\overset{\circ}{W} \cdot \theta^\vee = (\overset{\circ}{\Delta}^\vee)_s.$$

It is well known that in the finite type case the root lattice is spanned over  $\mathbb{Z}$  by the short roots, hence in this case

$$M = \nu(\overset{\circ}{Q}^\vee),$$

giving (40) for  $r = 1$ . Equivalently,

$$M = \overset{\circ}{Q} \quad (\text{resp. } M = \mathbb{Z}\Delta_\ell) \quad \text{if } r = 1 \text{ and } A \text{ is symmetric (resp. nonsymmetric).}$$

Similarly, if  $a_0 r = 2$  or  $3$ , then  $\theta^\vee$  is a long root of  $\overset{\circ}{\Delta}^\vee$ , hence

$$\overset{\circ}{W} \cdot \theta^\vee = (\overset{\circ}{\Delta}^\vee)_\ell,$$

and we get  $M = \overset{\circ}{Q}$ . Finally, for  $A_{2\ell}^{(2)}$  one has  $\nu(\theta^\vee) = \frac{1}{2}\theta$ , hence

$$M = \frac{1}{2}\mathbb{Z}\overset{\circ}{\Delta}_\ell,$$

which is equivalent to (6.5.8) in this case also (see (6.3.6)).

Note also the following useful fact:

$$\overset{\circ}{Q} \supset \nu(\overset{\circ}{Q}^\vee) \quad \text{if } r = 1, \quad \overset{\circ}{Q} \subset \nu(\overset{\circ}{Q}^\vee) \quad \text{if } r > 1. \quad (41)$$

The lattice  $M$ , considered as an abelian group, operates faithfully on  $\mathfrak{h}^*$  by the formula for  $t_\alpha$ . We denote the corresponding subgroup of  $GL(\mathfrak{h}^*)$  by  $T$ , and call it the **group of translations**.

**Proposition 6.10 (6.5).**

$$W = \overset{\circ}{W} \ltimes T.$$

*Proof.* By (6.5.1),  $t_\alpha \in W$  for  $\alpha = \nu(\theta^\vee)$ , hence  $t_{w(\alpha)} \in W$  for  $w \in \overset{\circ}{W}$  by (6.5.7). Now it follows from (6.5.6) that  $t_\alpha \in W$  for all  $\alpha \in M$ . Since  $\overset{\circ}{W}$  is finite and  $M$  is a free abelian group,  $\overset{\circ}{W} \cap T = 1$ . We checked that  $T$  is a normal subgroup in  $W$ . Finally  $r_{\alpha_0} = t_{\nu(\theta^\vee)} r_\theta$ , and therefore the subgroup of  $W$  generated by  $T$  and  $\overset{\circ}{W}$  contains all  $r_i$  ( $i = 0, \dots, \ell$ ) and hence coincides with  $W$ .  $\square$

**Remark 6.11 (Memory Aid).** The operators  $t_\alpha$  are the **translations** in the affine Weyl group. They commute with each other and form a free abelian group  $T$  generated by a lattice  $M$  (essentially a coroot lattice, depending on the type). The finite Weyl group  $\overset{\circ}{W}$  normalizes  $T$ , so the full affine Weyl group splits as a semidirect product:

$$W \cong \overset{\circ}{W} \ltimes T.$$

## 6.4 The fundamental alcove

For  $s \in \mathbb{R}$  set

$$\mathfrak{h}_s^* = \{\lambda \in \mathfrak{h}_{\mathbb{R}}^* \mid \langle \lambda, K \rangle = s\}.$$

Note that

$$\mathfrak{h}_0^* = \sum_{i=0}^{\ell} \mathbb{R}\alpha_i$$

and that the hyperplanes  $\mathfrak{h}_s^*$  are  $W$ -invariant. Furthermore, the action of  $W$  on  $\mathfrak{h}_0^*$  is faithful.

Consider now the affine space  $\mathfrak{h}_1^* \bmod \mathbb{R}\delta$ . We should do this because  $\delta$  is in the radical of the invariant form and therefore, quotienting by  $\delta$  removes the degenerate direction. Since the action of  $W$  on  $\mathfrak{h}_0^*$  is faithful, its action on  $\mathfrak{h}_{\mathbb{R}}^*/\mathbb{R}\delta \simeq \mathfrak{h}_0^*$  and thus on  $\mathfrak{h}_1^* \bmod \mathbb{R}\delta$  is also faithful. The affine action of  $W$  on the affine space  $\mathfrak{h}_1^* \bmod \mathbb{R}\delta$  has the following simple geometrical meaning. We identify  $\mathfrak{h}_1^* \bmod \mathbb{R}\delta$  with  $\overset{\circ}{\mathfrak{h}}^*$  by projection.

Concretely, any  $\lambda \in \mathfrak{h}_{\mathbb{R}}^*$  can be written uniquely as  $\lambda = \bar{\lambda} + m\Lambda_0 + s\delta$ , where  $\bar{\lambda} \in \mathfrak{h}_0^*$  and  $m, s \in \mathbb{R}$ . Then  $\langle \lambda, K \rangle = m$  (because  $\langle \bar{\lambda}, K \rangle = 0$  and  $\langle \delta, K \rangle = 0$ ). So the level-1 hyperplane is  $\mathfrak{h}_1^* = \{\lambda : \langle \lambda, K \rangle = 1\} = \Lambda_0 + \mathfrak{h}_0^* + \mathbb{R}\delta$ . Now mod out by  $\mathbb{R}\delta$ . In the quotient, every class has a unique representative of the form  $\Lambda_0 + \bar{\lambda}$  with  $\bar{\lambda} \in \mathfrak{h}_1^*$ . That gives a canonical identification  $\mathfrak{h}_1^*/\mathbb{R}\delta \xrightarrow{\circ} \mathfrak{h}_1^*$ , sending  $[\Lambda_0 + \bar{\lambda} + s\delta]$  to  $\bar{\lambda}$ .

We thus obtain an isomorphism from  $W$  onto a group of affine transformations  $W_{\text{af}}$  of  $\mathfrak{h}_0^*$ . We denote this isomorphism by  $\text{af}$ , so that

$$\text{af}(w)(\bar{\lambda}) = \overline{w(\lambda)} \quad \text{for } \lambda \in \mathfrak{h}_1^*.$$

Now we describe the action of  $W_{\text{af}}$  on  $\mathfrak{h}_0^*$ . It is clear that

$$\text{af}(w) = w \quad \text{for } w \in \overset{\circ}{W}. \quad (42)$$

Furthermore,

$$\text{af}(r_{\alpha_0})(\lambda) = r_{\theta}(\lambda) + \nu(\theta^{\vee}), \quad \lambda \in \mathfrak{h}_1^*. \quad (43)$$

Indeed, if  $\mu \in \mathfrak{h}_1^*$ , then  $\langle \mu, K \rangle = 1$ , hence

$$\langle \mu, \alpha_0^{\vee} \rangle = \langle \mu, K - a_0\theta^{\vee} \rangle = 1 - a_0\langle \mu, \theta^{\vee} \rangle,$$

and

$$r_{\alpha_0}(\mu) = \mu - (1 - a_0\langle \mu, \theta^{\vee} \rangle)\alpha_0 = \mu + (1 - a_0\langle \mu, \theta^{\vee} \rangle)a_0^{-1}\theta = \mu - \langle \mu, \theta^{\vee} \rangle\theta + \nu(\theta^{\vee}) \bmod \mathbb{R}\delta.$$

So,  $\text{af}(r_{\alpha_0})$  is a reflection in the hyperplane

$$\theta = 1 \quad (\text{i.e., } \{\lambda \in \mathfrak{h}_1^* \mid (\lambda|\theta) = 1\}).$$

$r_{\theta}$  is reflection in the hyperplane  $H_{\theta} = \{\lambda \in \mathfrak{h}_0^* \mid (\lambda|\theta) = 0\}$ .  $r_{\alpha_0}$  is an affine reflection in the hyperplane  $H_{\alpha_0} = \{\lambda \in \mathfrak{h}_1^* \mid (\lambda|\theta) = 1\}$ .

These two hyperplanes are parallel (both perpendicular to  $\theta$ ), but one is shifted by 1 unit in the  $\theta$ -direction. The product of two reflections in parallel hyperplanes is a translation by twice the perpendicular distance vector between the hyperplanes. Therefore the product  $r_{\alpha_0}r_{\theta}$  is a translation by  $\nu(\theta^{\vee})$  since  $(\nu(\theta^{\vee})|\theta) = \langle \theta^{\vee}, \theta \rangle = 2$ . Thus,

$$t_{\nu(\theta^{\vee})} = r_{\alpha_0}r_{\theta}$$

Therefore

$$\text{af}(t_{\nu(\theta^{\vee})})(\lambda) = \text{af}(r_{\alpha_0})(r_{\theta}(\lambda)) = r_{\theta}(r_{\alpha_0}(\lambda)) + \nu(\theta^{\vee}) = \lambda + \nu(\theta^{\vee}), \quad \lambda \in \mathfrak{h}_1^*.$$

Hence, by additivity and invariance, we obtain

$$\text{af}(t_\alpha)(\lambda) = \lambda + \alpha, \quad \lambda \in \mathfrak{h}^*, \alpha \in M. \quad (44)$$

So, the group  $W_{\text{af}}$  is none other than the so-called affine Weyl group of  $\mathfrak{g}$ .

**Definition 6.12 (Fundamental Alcove).** *We have the **fundamental alcove (at level 1)***

$$C_{\text{af}} = \{\lambda \in \mathfrak{h}_{\mathbb{R}}^* \mid (\lambda|\alpha_i) \geq 0 \text{ for } 1 \leq i \leq \ell, \text{ and } (\lambda|\theta) \leq 1\}.$$

Another way of writing this is that we took the fundamental Weyl chamber and looked at the level 1 slice  $H_k = \{\lambda : \langle \lambda, K \rangle = 1\}$ .

$$C_{\text{af}} = \{\lambda \in H_k \mid (\lambda|\alpha_i) \geq 0 \text{ for } 1 \leq i \leq \ell, (\lambda|\theta) \leq 1\}$$

Note that the condition  $(\lambda|\theta) \leq 1$  is equivalent to  $(\lambda|\alpha_0) \geq 0$  so that the fundamental alcove is cut out by the affine simple roots  $\alpha_0, \alpha_1, \dots, \alpha_\ell$ , just as the fundamental Weyl chamber is cut out by the finite simple roots  $\alpha_1, \dots, \alpha_\ell$  in the finite case.

The affine Weyl group  $W$  acts faithfully on  $\mathfrak{h}_{\mathbb{R}}^*$ . So in that space, you can cut out alcoves just as in the finite case, but now you need one extra hyperplane  $(\lambda|\theta) \leq 1$ , corresponding to the affine simple root  $\alpha_0 = \delta - \theta$ . The following proposition describes the action of  $W$  on  $\mathfrak{h}^*$ . The reason we care about alcoves is that the weight lattice modulo affine Weyl group is tiled by alcoves, and this gives us an understanding of representation theory.

**Proposition 6.13 (6.6).**

- a) Every point of  $\mathfrak{h}_{\mathbb{R}}^*$  is  $\bar{W}$ -equivalent mod  $M$  to a unique point of  $C_{\text{af}}$ .
- b) The stabilizer of every point of  $C_{\text{af}}$  under the action of  $\bar{W}$  on  $\mathfrak{h}_{\mathbb{R}}^*/M$  is generated by its intersection with  $\{r_\theta, r_{\alpha_1}, \dots, r_{\alpha_\ell}\}$ .
- c) For every  $\lambda \in \mathfrak{h}_1^*$  one has

$$\text{af}(W_\lambda) = (W_{\text{af}})_{\bar{\lambda}},$$

and  $W_\lambda \cap T = e$ , where  $W_\lambda$  denotes the stabilizer of  $\lambda$ .

*Proof.* Consider the projection map  $\pi : \mathfrak{h}_1^* \rightarrow \mathfrak{h}_{\mathbb{R}}^*$ . Then it is clear that  $\pi$  is surjective and that

$$\text{af}(w) \circ \pi = \pi \circ w \quad \text{for } w \in W.$$

Furthermore:

$$\pi^{-1}(C_{\text{af}}) = C^\vee \cap \mathfrak{h}_1^*.$$

But by Proposition 3.12b),  $C^\vee \cap \mathfrak{h}_1^*$  is a fundamental domain for the action of  $W$  on  $\mathfrak{h}_1^* \times V$ . This together with (6.6.1 and 3) proves part a). We also deduce from the above that

$$\text{af}(W_\lambda) = (W_{\text{af}})_{\bar{\lambda}}.$$

Now part b) follows from Proposition 3.12a). Finally  $W_\lambda \cap T = 1$ , since  $\text{af}(W_\lambda) = (W_{\text{af}})_{\bar{\lambda}}$  contains no nontrivial translation.  $\square$

**Example 6.14 (Affine  $\mathfrak{sl}_2$ ).** Recall that for  $\widehat{\mathfrak{sl}}_2$ , we have  $\theta = \alpha_1$ ,  $\alpha_0 = \delta - \alpha_1$ , and  $\mathfrak{h}_{\mathbb{R}}^{\circ*} = \mathbb{R}$  with coordinate given by  $\lambda \mapsto (\lambda|\alpha_1) = \langle \lambda, \alpha_1^\vee \rangle$ . Note that this equality is because evaluating against simple coroots is the same as pairing against simple roots, since  $\alpha_1^\vee = \frac{2h_{\alpha_1}}{(\alpha_1|\alpha_1)}$  since all roots have the same square length which is 2. The fundamental alcove is then

$$\begin{aligned} C_{\text{af}} &= \{\lambda \in \mathbb{R} \mid (\lambda|\alpha_1) \geq 0, (\lambda|\theta) \leq 1\} \\ &= \{\lambda \in \mathbb{R} \mid 0 \leq (\lambda|\alpha_1) \leq 1\} \end{aligned}$$

i.e. the interval  $[0, 1]$  in our coordinate  $(\cdot|\alpha)$ . The affine Weyl group  $W = \overset{\circ}{W} \ltimes T$  corresponds to reflections across affine hyperplanes

$$H_{\alpha, k} = \{\lambda \in \mathfrak{h}_{\mathbb{R}}^* \mid (\lambda|\alpha) = k\}$$

which is generated by the simple reflections  $r_{\alpha_1}$  (reflection across  $H_{\alpha_1, 0}$ ) and  $r_{\alpha_0}$  (reflection across  $H_{\alpha_1, 1}$ ). The group of translations  $T$  is generated by  $t_{\nu(\alpha_1^\vee)} = r_{\alpha_0} r_{\alpha_1}$ .

The reflections are indexed by real roots  $\alpha + n\delta$  for  $\alpha \in \overset{\circ}{\Delta}$  and  $n \in \mathbb{Z}$ . In particular, they have the formula

$$\begin{aligned} r_{\alpha+n\delta}(\lambda) &= \lambda - \langle \lambda, (\alpha + n\delta)^\vee \rangle (\alpha + n\delta) \\ &= \lambda - \langle \lambda, \alpha^\vee \rangle (\alpha + n\delta) \\ &= \lambda - (\lambda|\alpha)\alpha - n\langle \lambda, K \rangle \alpha + n(\lambda|\alpha)\delta \\ &= \lambda - ((\lambda|\alpha) + n)\alpha \end{aligned}$$

where we used the fact that we are on the level-1 hyperplane  $\langle \lambda, K \rangle = 1$  and projected away the  $\delta$ -term to get an isomorphism

$$\mathfrak{h}_1^*/\mathbb{R}\delta \xrightarrow{\sim} \overset{\circ}{\mathfrak{h}}^*.$$

Now we can check on the level of coordinates that

$$\begin{aligned} \langle r_{\alpha_1+n\delta}(\lambda), \alpha_1^\vee \rangle &= \langle \lambda - ((\lambda|\alpha_1) + n)\alpha_1, \alpha_1^\vee \rangle \\ &= (\lambda|\alpha_1) - ((\lambda|\alpha_1) + n)(\alpha_1|\alpha_1^\vee) \\ &= (\lambda|\alpha_1) - 2((\lambda|\alpha_1) + n) \\ &= -(\lambda|\alpha_1) - 2n \end{aligned}$$

So we can identify the reflection associated with  $\alpha_1 - n\delta$  as the reflection across the hyperplane  $H_{\alpha_1, n} = \{\lambda \mid (\lambda|\alpha_1) = n\}$ , in coordinates the map is given by

$$x \mapsto -x + 2n.$$

The translations are indexed by

$$M = \nu(\mathbb{Z}(\overset{\circ}{W}) \cdot \theta^\vee)$$

which in this case is just  $\nu(\overset{\circ}{Q}^\vee) = \mathbb{Z}\nu(\alpha_1^\vee) = \mathbb{Z}\alpha_1$ . In particular, there are  $\mathbb{Z}\alpha_1$  worth of translations given by

$$t_{m\alpha_1}(\lambda) = \lambda + m\alpha_1$$

which in our coordinate is just  $x \mapsto x + 2m$ .

All together, the affine Weyl group is generated  $r_0 : x \mapsto -x$  and  $r_1 : x \mapsto -x + 2$ . They satisfy the relations  $r_0^2 = r_1^2 = e$ . Also  $t_m = (r_0 r_1)^m$  for  $m \in \mathbb{Z}$  and conjugation by  $r_0$  inverts the translation  $t_m$ . Therefore, the structure is precisely the infinite dihedral group  $D_\infty \cong \mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z}$ .

**Example 6.15 (Fundamental chamber of  $\widehat{\mathfrak{sl}}_2$ ).** Let  $d, \alpha_1^\vee, K$  be the basis of  $\mathfrak{h}$  dual to  $\delta, \omega_1, \Lambda_0$ . Recall that in general, the fundamental weights are dual to the simple coroots, and the fundamental coweights are dual to the simple roots.

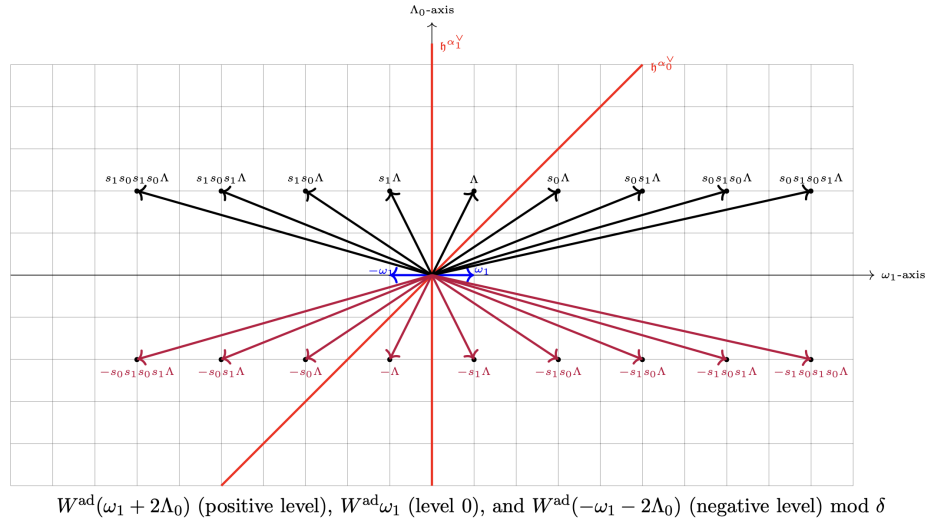


Figure 3: Fundamental chamber of  $\widehat{\mathfrak{sl}}_2$

The vector space is 3-dimensional but in fact we have projected onto the plane  $\delta = 0$ . The figure shows the dominant chamber  $C$  in  $\mathfrak{h}_\mathbb{R}^*$  for  $\widehat{\mathfrak{sl}}_2$  cut out by the hyperplanes  $\alpha_0^\vee \geq 0$  and  $\alpha_1^\vee \geq 0$ . Another thing



to note is that the lattice points where the fundamental chamber meets the hyperplane  $\langle \lambda, K \rangle = k$  are precisely the dominant integral weights of level  $k$ . The reflections  $r_0, r_1$  are reflections across the coroot hyperplanes, and looking at the  $\langle \lambda, K \rangle = 1$  hyperplane we see that the fundamental alcove is the intersection of the dominant chamber with this hyperplane. The affine Weyl group acts on the level-1 hyperplane and tiles it with alcoves.

The  $W$ -action fixes the level (the coefficient of  $\Lambda_0$ ) but it actually changes the  $\delta$  coordinate significantly.

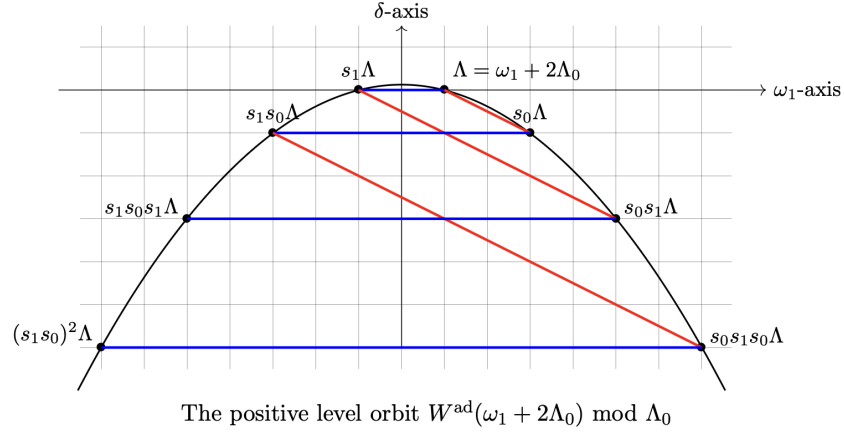


Figure 4: Different projection of the chamber

We compute what this action looks like. Any weight  $\lambda \in \widehat{\mathfrak{h}}^*$  can be written as  $\lambda = x\omega_1 + k\Lambda_0 + y\delta$ . Recall that  $\langle \omega_1, h \rangle = 1$ ,  $\langle \omega_1, K \rangle = 0$ ,  $\langle \Lambda_0, K \rangle = 1$ ,  $\langle \Lambda_0, h \rangle = 0$ ,  $\langle \delta, \text{anything in } \mathfrak{h} \oplus \mathbb{C}K \rangle = 0$ .

Then:

$$\begin{aligned}\langle \lambda, \alpha_1^\vee \rangle &= \langle x\omega_1 + k\Lambda_0 + y\delta, h \rangle = x, \\ \langle \lambda, \alpha_0^\vee \rangle &= \langle x\omega_1 + k\Lambda_0 + y\delta, K - h \rangle = k - x.\end{aligned}$$

We compute that

$$s_1(\lambda) = \lambda - \langle \lambda, \alpha_1^\vee \rangle \alpha_1 = (x\omega_1 + k\Lambda_0 + y\delta) - x(2\omega_1) = (-x)\omega_1 + k\Lambda_0 + y\delta$$

so in coordinates,  $s_1$  acts by  $(x, y) \mapsto (-x, y)$ . Similarly,

$$\begin{aligned}s_0(\lambda) &= \lambda - \langle \lambda, \alpha_0^\vee \rangle \alpha_0 \\ &= (x\omega_1 + k\Lambda_0 + y\delta) - (k - x)(\delta - 2\omega_1) \\ &= x\omega_1 + k\Lambda_0 + y\delta - (k - x)\delta + 2(k - x)\omega_1 \\ &= (x + 2(k - x))\omega_1 + k\Lambda_0 + (y - (k - x))\delta \\ &= (2k - x)\omega_1 + k\Lambda_0 + (y - k + x)\delta.\end{aligned}$$

so in coordinates,  $s_0$  acts by  $(x, y) \mapsto (2k - x, y - k + x)$ .

The translation  $t_{\alpha_1^\vee} = s_1 s_0$  acts by

$$t(x, y) = (x - 2k, y - k + x)$$

and induction shows that

$$t^n(x, y) = (x - 2nk, y + nx - n^2k)$$

for  $n \in \mathbb{Z}$ . This agrees with the general formula of Kac:

$$t_\alpha(\lambda) = \lambda + \langle \lambda, K \rangle \alpha - ((\lambda|\alpha) + \frac{1}{2}|\alpha|^2 \langle \lambda, K \rangle) \delta.$$

We're translating by  $\beta = -n\alpha_1$  at level  $k = \lambda(K)$ . Using  $(\alpha_1, \alpha_1) = 2$  and  $(\omega_1, \alpha_1) = 1$ , and writing  $\lambda = x_0\omega_1 + k\Lambda_0 + y_0\delta$ , we get  $(\lambda|\alpha_1) = x_0$  and  $\langle \lambda, K \rangle = k$ . The formula gives

$$t^n(\lambda) = \lambda + k(-n\alpha_1) - (-n(\lambda|\alpha_1) + \frac{1}{2}kn^2(\alpha_1|\alpha_1))\delta = \lambda - 2nk\omega_1 + nx_0\delta - n^2k\delta$$

which matches our formula exactly in coordinates.

If we eliminate the parameter, we get an equation for the parabola:

$$y_n = y_0 - \frac{1}{4k}(x_n - x_0)^2 + \frac{x_0^2}{4k},$$

In the figure above,  $k = 2$  and  $\omega_1 + 2\Lambda_0$  has coordinates  $(x_0, y_0) = (0, 1)$ , we get a downward opening parabola as the orbit of  $\omega_1 + 2\Lambda_0$  under the affine Weyl group.

## 7 Affine Algebras as Central Extensions of Loop Algebras

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:

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

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  $\omega$  is a generator. In particular, any central extension of  $\mathfrak{g}[z, z^{-1}]$  is a multiple of the one defined by  $\omega$ . 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  $\iota_\xi$ . Hence  $\mathfrak{g}$  acts trivially on cohomology so every class has a  $\mathfrak{g}$ -invariant representative (Lemma 13.1).

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

**Proposition 7.1 (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  $\varphi$  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  $\omega, \omega'$  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  $k, \ell$ . 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  $\omega$  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 7.2.** *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 7.3.**  $\gamma_{n,-n} = n\gamma_{1,-1} = n\gamma$  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 7.4 (Derived affine Kac-Moody Lie algebra).** Any central extension of  $\mathfrak{g}[z, z^{-1}]$  has the form

$$0 \rightarrow \mathbb{C}c \rightarrow \widehat{\mathfrak{g}}' \rightarrow \mathfrak{g}[z, z^{-1}] \rightarrow 0, \quad [xz^n, yz^m]_{\widehat{\mathfrak{g}}'} = [xz^n, yz^m]_{\mathfrak{g}} + \omega(xz^n, yz^m)c$$

where  $\omega$  is any 2-cocycle on  $\mathfrak{g}[z, z^{-1}]$ . Suppose we make two choices of  $\omega$  and are given an isomorphism  $\phi : \mathfrak{g}_1 \rightarrow \mathfrak{g}_2$  which preserves the central element  $c$ .

Choosing splittings  $\iota_1 : \mathbb{C}c \oplus \mathfrak{g} \rightarrow \mathfrak{g}_1$  and  $\iota_2 : \mathbb{C}c \oplus \mathfrak{g} \rightarrow \mathfrak{g}_2$  for these extensions, we obtain the two-cocycles  $\omega_1$  and  $\omega_2$ . Define a linear map  $f : \mathfrak{g} \rightarrow \mathbb{C}$  as the composition of  $\iota_1$ ,  $\phi$  and the projection  $\mathfrak{g}_2 \rightarrow \mathbb{C}c$  induced by  $\iota_2$ . A straightforward calculation gives  $\omega_2(x, y) = \omega_1(x, y) + f([x, y])$ . In particular, central extensions are classified up to isomorphism by  $H^2(\mathfrak{g}[z, z^{-1}])$ , which we showed is one-dimensional and representatives can be chosen to be  $\mathfrak{g}$ -invariant.

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  $\mathfrak{g}[z, z^{-1}]$ .

We will see 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 7.5 (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 7.6 (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

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

## 7.1 Formulas for loop algebras

Let  $\mathcal{L} = \mathbb{C}[t, t^{-1}]$  be the algebra of Laurent polynomials in  $t$ . Recall that the residue of a Laurent polynomial

$$P = \sum_{k \in \mathbb{Z}} c_k t^k$$

(where all but a finite number of  $c_k$  are 0) is defined by

$$\text{Res } P = c_{-1}.$$

This is a linear functional on  $\mathcal{L}$  defined by the properties

$$\text{Res } t^{-1} = 1, \quad \text{Res } \frac{dP}{dt} = 0.$$

Define a bilinear  $\mathbb{C}$ -valued function  $\varphi$  on  $\mathcal{L}$  by

$$\varphi(P, Q) = \text{Res} \frac{dP}{dt} Q.$$

Then it is easy to check the following two properties:

$$\varphi(P, Q) = -\varphi(Q, P), \quad (7.1.1)$$

$$\varphi(PQ, R) + \varphi(QR, P) + \varphi(RP, Q) = 0, \quad (P, Q, R \in \mathcal{L}). \quad (7.1.2)$$

Consider the **loop algebra**

$$\mathcal{L}(\overset{\circ}{\mathfrak{g}}) := \mathcal{L} \otimes_{\mathbb{C}} \overset{\circ}{\mathfrak{g}}.$$

This is an infinite-dimensional complex Lie algebra with the bracket  $[\cdot, \cdot]_0$  defined by

$$[P \otimes x, Q \otimes y]_0 = PQ \otimes [x, y] \quad (P, Q \in \mathcal{L}; x, y \in \overset{\circ}{\mathfrak{g}}).$$

It may be identified with the Lie algebra of regular rational maps  $\mathbb{C}^\times \rightarrow \overset{\circ}{\mathfrak{g}}$ , so that the element  $\sum_i (t^i \otimes x_i)$  corresponds to the mapping

$$z \mapsto \sum_i z^i x_i.$$

Fix a nondegenerate invariant symmetric bilinear  $\mathbb{C}$ -valued form  $(\cdot | \cdot)$  on  $\overset{\circ}{\mathfrak{g}}$ ; such a form exists (e.g. by Theorem 2.2) and is unique up to a constant multiple. We extend this form by linearity to an  $\mathcal{L}$ -valued bilinear form  $(\cdot | \cdot)_t$  on  $\mathcal{L}(\overset{\circ}{\mathfrak{g}})$  by

$$(P \otimes x | Q \otimes y)_t = PQ(x | y).$$

Also, we extend every derivation  $D$  of the algebra  $\mathcal{L}$  to a derivation of the Lie algebra  $\mathcal{L}(\overset{\circ}{\mathfrak{g}})$  by

$$D(P \otimes x) = D(P) \otimes x.$$

Now we can define a  $\mathbb{C}$ -valued 2-cocycle on the Lie algebra  $\mathcal{L}(\overset{\circ}{\mathfrak{g}})$  by

$$\psi(a, b) = \text{Res} \left( \frac{da}{dt} \middle| b \right)_t$$

Recall that a  $\mathbb{C}$ -valued 2-cocycle on a Lie algebra  $\mathfrak{g}$  is a bilinear  $\mathbb{C}$ -valued function  $\psi$  satisfying two conditions:

$$(\text{Co } 1) \quad \psi(a, b) = -\psi(b, a),$$

$$(\text{Co } 2) \quad \psi([a, b], c) + \psi([b, c], a) + \psi([c, a], b) = 0 \quad (a, b, c \in \mathfrak{g}).$$

It is sufficient to check these conditions for  $a = P \otimes x, b = Q \otimes y, c = R \otimes z$ , where  $P, Q, R \in \mathcal{L}$  and  $x, y, z \in \dot{\mathfrak{g}}$ . We have

$$\psi(a, b) = (x \mid y) \varphi(P, Q).$$

Hence, (Co 1) follows from (7.1.1) and the symmetry of  $(\cdot \mid \cdot)$ . Property (Co 2) follows from (7.1.2) and the symmetry and invariance of  $(\cdot \mid \cdot)$ .

Denote by  $\tilde{\mathcal{L}}(\dot{\mathfrak{g}})$  the extension of the Lie algebra  $\mathcal{L}(\dot{\mathfrak{g}})$  by a 1-dimensional center, associated to the cocycle  $\psi$ . Explicitly,

$$\tilde{\mathcal{L}}(\dot{\mathfrak{g}}) = \mathcal{L}(\dot{\mathfrak{g}}) \oplus \mathbb{C}K$$

(direct sum of vector spaces), and the bracket is given by

$$[a + \lambda K, b + \mu K] = [a, b]_0 + \psi(a, b)K, \quad (a, b \in \mathcal{L}(\dot{\mathfrak{g}}); \lambda, \mu \in \mathbb{C}). \quad (7.2.1)$$

Finally, denote by  $\hat{\mathcal{L}}(\dot{\mathfrak{g}})$  the Lie algebra that is obtained by adjoining to  $\tilde{\mathcal{L}}(\dot{\mathfrak{g}})$  a derivation  $d$  which acts on  $\mathcal{L}(\dot{\mathfrak{g}})$  as  $t \frac{d}{dt}$  and which kills  $K$ . In other words,  $\hat{\mathcal{L}}(\dot{\mathfrak{g}})$  is a complex vector space

$$\hat{\mathcal{L}}(\dot{\mathfrak{g}}) = \mathcal{L}(\dot{\mathfrak{g}}) \oplus \mathbb{C}K \oplus \mathbb{C}d$$

with the bracket defined as follows ( $x, y \in \dot{\mathfrak{g}}; \lambda, \mu, \lambda_1, \mu_1 \in \mathbb{C}$ ):

$$\begin{aligned} & [t^m \otimes x \oplus \lambda K \oplus \mu d, t^n \otimes y \oplus \lambda_1 K \oplus \mu_1 d] \\ &= (t^{m+n} \otimes [x, y] + \mu n t^n \otimes y - \mu_1 m t^m \otimes x) \oplus m \delta_{m, -n} (x \mid y) K. \end{aligned} \quad (7.2.2)$$

We shall prove that  $\hat{\mathcal{L}}(\dot{\mathfrak{g}})$  is an affine algebra associated to the affine matrix  $A$  of type  $X_\ell^{(1)}$ .

Here we check that  $d$  is a derivation of the Lie algebra  $\tilde{\mathcal{L}}(\dot{\mathfrak{g}})$ . More generally, denote by  $d_s$  the endomorphism of the space  $\tilde{\mathcal{L}}(\dot{\mathfrak{g}})$  defined by

$$d_s|_{\mathcal{L}(\dot{\mathfrak{g}})} = -t^{s+1} \frac{d}{dt}, \quad d_s(K) = 0, \quad (7.3.1)$$

so that  $d_0 = -d$ .

**Proposition 7.7 (7.3).**  $d_s$  is a derivation of  $\tilde{\mathcal{L}}(\dot{\mathfrak{g}})$ .

*Proof.* Since  $D := d_s$  is a derivation of  $\mathcal{L}(\dot{\mathfrak{g}})$ , we deduce that

$$D([a + \lambda K, b + \mu K]) = D([a, b]_0) = [D(a), b]_0 + [a, D(b)]_0.$$

But

$$[D(a), b] = [D(a), b]_0 + \psi(D(a), b)K.$$



Hence, one has to check that

$$\psi(D(a), b) + \psi(a, D(b)) = 0. \quad (7.3.2)$$

Set  $a = P \otimes x, b = Q \otimes y$ . Then the left-hand side of (7.3.2) is

$$\begin{aligned} (x | y)(\varphi(D(P), Q) + \varphi(P, D(Q))) &= (x | y)(-\varphi(Q, D(P)) + \varphi(P, D(Q))) \\ &= (x | y) \operatorname{Res} \left( -\frac{dQ}{dt} t^{s+1} \frac{dP}{dt} + \frac{dP}{dt} t^{s+1} \frac{dQ}{dt} \right) = 0 \end{aligned}$$

as desired.  $\square$

Note that

$$\mathfrak{v} := \bigoplus_{j \in \mathbb{Z}} \mathbb{C} d_j$$

is a  $\mathbb{Z}$ -graded subalgebra in  $\operatorname{Der} \tilde{\mathcal{L}}(\mathring{\mathfrak{g}})$  with the following commutation relations:

$$[d_i, d_j] = (i - j) d_{i+j}.$$

This is the Lie algebra of regular vector fields on  $\mathbb{C}^\times$  (= derivations of  $\mathcal{L}$ ). The Lie algebra  $\mathfrak{v}$  has a unique (up to isomorphism) nontrivial central extension by a 1-dimensional center, say  $\mathbb{C}c$ , called the **Virasoro algebra**  $\operatorname{Vir}$ , which is defined by the following commutation relations:

$$[d_i, d_j] = (i - j) d_{i+j} + \frac{1}{12} (i^3 - i) \delta_{i,-j} c, \quad (i, j \in \mathbb{Z}). \quad (7.3.3)$$

## 7.2 The root system

Let  $\mathring{\Delta} \subset \mathring{\mathfrak{h}}^*$  be the root system of the Lie algebra  $\mathring{\mathfrak{g}}$ , let  $\{\alpha_1, \dots, \alpha_\ell\}$  be the root basis,  $\{H_1, \dots, H_\ell\}$  the coroot basis, and  $E_i, F_i$  ( $i = 1, \dots, \ell$ ) the Chevalley generators. Let  $\theta$  be the highest root of the finite root system  $\mathring{\Delta}$ . Let

$$\mathring{\mathfrak{g}} = \bigoplus_{\alpha \in \mathring{\Delta} \cup \{0\}} \mathring{\mathfrak{g}}_\alpha$$

be the root space decomposition of  $\mathring{\mathfrak{g}}$ .

Recall that  $(\alpha | \alpha) \neq 0$  and  $\dim \mathring{\mathfrak{g}}_\alpha = 1$  for  $\alpha \in \mathring{\Delta}$  (there are no imaginary roots). Let  $\mathring{\omega}$  be the Chevalley involution of  $\mathring{\mathfrak{g}}$ . We choose  $F_0 \in \mathring{\mathfrak{g}}_{-\theta}$  such that

$$(F_0 | \mathring{\omega}(F_0)) = -\frac{2}{(\theta | \theta)}, \quad \text{and set } E_0 = -\mathring{\omega}(F_0).$$

Then, due to Theorem 2.2(e), we have

$$[E_0, F_0] = -\theta^\vee. \quad (7.4.1)$$

The elements  $E_i$  ( $i = 0, \dots, \ell$ ) generate the Lie algebra  $\mathring{\mathfrak{g}}$ , since in the adjoint representation we have

$$\mathring{\mathfrak{g}} = U(\mathring{\mathfrak{g}})(E_0) = U(\mathring{\mathfrak{n}}_+)(E_0) \quad (\text{recall that } \mathring{\mathfrak{g}} \text{ is simple}).$$

Now we turn to the Lie algebra  $\widehat{\mathcal{L}}(\mathring{\mathfrak{g}})$ . It is clear that  $\mathbb{C}K$  is the one-dimensional center of  $\widehat{\mathcal{L}}(\mathring{\mathfrak{g}})$ , and that the centralizer of  $\mathfrak{v}$  in  $\widehat{\mathcal{L}}(\mathring{\mathfrak{g}})$  is a direct sum of Lie algebras:

$$\mathbb{C}K \oplus \mathbb{C}d \oplus (1 \otimes \mathring{\mathfrak{g}}).$$

In particular,  $1 \otimes \mathring{\mathfrak{g}}$  is a subalgebra of  $\widehat{\mathcal{L}}(\mathring{\mathfrak{g}})$ ; we identify  $\mathring{\mathfrak{g}}$  with this subalgebra by  $x \mapsto 1 \otimes x$ .

Furthermore, set

$$\mathfrak{h} := \mathring{\mathfrak{h}} \oplus \mathbb{C}K \oplus \mathbb{C}d,$$

which is an  $(\ell + 2)$ -dimensional commutative subalgebra of  $\widehat{\mathcal{L}}(\mathring{\mathfrak{g}})$ .

We extend  $\lambda \in \mathring{\mathfrak{h}}^*$  to a linear function on  $\mathfrak{h}$  by setting

$$(\lambda, K) = \langle \lambda, K \rangle = 0, \quad \langle \lambda, d \rangle = 0,$$

so that  $\mathring{\mathfrak{h}}^*$  is identified with a subspace in  $\mathfrak{h}^*$ . We denote by  $\delta$  the linear function on  $\mathfrak{h}$  defined by

$$\delta|_{\mathring{\mathfrak{h}} + \mathbb{C}K} = 0, \quad \langle \delta, d \rangle = 1.$$

Finally, set

$$e_0 = t \otimes E_0, \quad f_0 = t^{-1} \otimes F_0, \quad e_i = 1 \otimes E_i, \quad f_i = 1 \otimes F_i \quad (i = 1, \dots, \ell).$$

We deduce from (7.4.1) that

$$[E_0, F_0] = \frac{2}{(\theta | \theta)} K - \theta^\vee. \quad (7.4.2)$$

Now we describe the root system and the root space decomposition of  $\widehat{\mathcal{L}}(\mathring{\mathfrak{g}})$  with respect to  $\mathfrak{h}$ :

$$\Delta = \{ j\delta + \gamma \mid j \in \mathbb{Z}, \gamma \in \mathring{\Delta} \} \cup \{ j\delta \mid j \in \mathbb{Z} \setminus \{0\} \},$$

$$\widehat{\mathcal{L}}(\mathring{\mathfrak{g}}) = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \widehat{\mathcal{L}}(\mathring{\mathfrak{g}})_\alpha,$$

where

$$\widehat{\mathcal{L}}(\mathring{\mathfrak{g}})_{j\delta + \gamma} = t^j \otimes \mathring{\mathfrak{g}}_\gamma, \quad \widehat{\mathcal{L}}(\mathring{\mathfrak{g}})_{j\delta} = t^j \otimes \mathring{\mathfrak{h}}.$$

We set

$$\Pi = \{\alpha_0 := \delta - \theta, \alpha_1, \dots, \alpha_\ell\}, \quad \Pi^\vee = \{\alpha_0^\vee := \frac{2}{(\theta|\theta)} K - 1 \otimes \theta^\vee, \alpha_i^\vee := 1 \otimes H_i \ (i = 1, \dots, \ell)\}.$$

Then Proposition 6.4(a) implies

$$A = ((\alpha_j, \alpha_i^\vee))_{i,j=0}^\ell. \quad (7.4.3)$$

In other words,  $(\mathfrak{h}, \Pi, \Pi^\vee)$  is a realization of the affine matrix  $A$  we started with. Indeed,  $\Pi$  and  $\Pi^\vee$  are linearly independent, i.e. (1.1.1) holds, and

$$2n - \text{rank } A = 2(\ell + 1) - \ell = \ell + 2 = \dim \mathfrak{h},$$

**Theorem 7.8 (7.4).** *Let  $\mathring{\mathfrak{g}}$  be a complex finite-dimensional simple Lie algebra, and let  $A$  be its extended Cartan matrix. Then  $\widehat{\mathcal{L}}(\mathring{\mathfrak{g}})$  is the affine Kac–Moody algebra associated to the affine matrix  $A$ ,  $\mathfrak{h}$  is its Cartan subalgebra,  $\Pi$  and  $\Pi^\vee$  are the root and coroot bases, and  $e_0, \dots, e_\ell, f_0, \dots, f_\ell$  are the Chevalley generators. In other words, the quadruple*

$$(\widehat{\mathcal{L}}(\mathring{\mathfrak{g}}), \mathfrak{h}, \Pi, \Pi^\vee)$$

*is the quadruple associated to  $A$ .*

**Corollary 7.9.** *Let  $\mathfrak{g}$  be an untwisted affine Kac–Moody algebra with Cartan matrix of rank  $\ell + 1$ . Then every imaginary root space of  $\mathfrak{g}$  is of rank  $\ell$ .*

*Proof.* The root space of the imaginary root  $n\delta$  is  $t^n \otimes \mathring{\mathfrak{h}}$ , which is of rank  $\ell$ . To see this, we write  $h = h' + \lambda K + \mu d \in \mathfrak{h}$  and  $x \in \mathring{\mathfrak{h}}$ . Then, for  $a = t^n \otimes x$ , the action of  $h' \in \mathring{\mathfrak{h}}$  is given by

$$\begin{aligned} [h', t^n \otimes x] &= t^n \otimes [h', x] = 0 \text{ since } \mathring{\mathfrak{h}} \text{ is abelian.} \\ [K, t^n \otimes x] &= 0 \text{ since } K \text{ is central.} \\ [d, t^n \otimes x] &= n t^n \otimes x \end{aligned}$$

Putting these together:  $[h, t^n \otimes x] = \mu n t^n \otimes x$ .

That is, the functional  $\alpha \in \mathfrak{h}^*$  corresponding to this weight satisfies  $\alpha|_{\mathring{\mathfrak{h}} + \mathbb{C}K} = 0$ ,  $\alpha(d) = n$ . The functional above is just  $n\delta$ .  $\square$

## 8 Highest weight modules over Kac–Moody algebras

Recall from the finite dimensional theory, we introduced the notion of highest weight modules, the universal such module (Verma module), and the unique irreducible quotient of the Verma module. We also introduced the notion of a primitive vector. The notions generalize to Kac–Moody algebras.

## 8.1 Category $\mathcal{O}$

One might ask if every nonzero module in the category  $\mathcal{O}$  is a highest weight module. The answer is no because there are easy counterexamples like the sum of two irreducible highest weight modules. However, it turns out that these are the only counterexamples. More precisely, every irreducible module in the category  $\mathcal{O}$  is a highest weight module, and every module in the category  $\mathcal{O}$  is generated by its primitive vectors.

**Proposition 8.1 (9.3).** *Let  $V$  be a nonzero module from the category  $\mathcal{O}$ . Then:*

- (a)  *$V$  contains a nonzero weight vector  $v$  such that  $\mathfrak{n}_+(v) = 0$ .*
- (b) *The following conditions are equivalent:*
  - (i)  *$V$  is irreducible;*
  - (ii)  *$V$  is a highest-weight module and any primitive vector of  $V$  is a highest-weight vector;*
  - (iii)  *$V \simeq L(\Lambda)$  for some  $\Lambda \in \mathfrak{h}^*$ .*
- (c)  *$V$  is generated by its primitive vectors as a  $\mathfrak{g}(A)$ -module.*

*Proof.* To prove (a), take a maximal  $\lambda \in P(V)$  (with respect to the ordering  $\geq$ ). Then one can take  $v$  to be a weight vector of weight  $\lambda$ .

Let  $V$  be an irreducible module; then a weight vector  $v$  is primitive if and only if  $\mathfrak{n}_+(v) = 0$ . Take a primitive vector  $v$  of weight  $\lambda$ . Then  $U(\mathfrak{g})(v)$  is a submodule of  $V$ , hence  $V = U(\mathfrak{g})(v)$  and  $V$  is a module with highest weight  $\lambda$ . In particular,  $P(V) \leq \lambda$  and  $\dim V_\lambda = 1$ . Hence every primitive vector is proportional to  $v$ , which proves the implication (i)  $\Rightarrow$  (ii) of (b).

If  $V$  is a highest-weight module and  $U \subset V$  is a proper submodule, then  $U$  contains a primitive vector by (a). This proves the implication (ii)  $\Rightarrow$  (iii) and the assertion (b).

Let  $V'$  be the submodule in  $V$  generated by all primitive vectors. If  $V' \neq V$ , then the  $\mathfrak{g}(A)$ -module  $V/V'$  contains a primitive vector  $v$  by (a). But a weight vector in  $V$  which is a preimage of  $v$  is a primitive vector.  $\square$

The following statement is a key technical result which is used in establishing the complete reducibility in many cases.

**Lemma 8.2 (9.5).** *Let  $V$  be a  $\mathfrak{g}(A)$ -module from the category  $\mathcal{O}$ . If for any two primitive weights  $\lambda$  and  $\mu$  of  $V$  the inequality  $\lambda \geq \mu$  implies  $\lambda = \mu$ , then the module  $V$  is completely reducible (i.e.  $V$  decomposes into a direct sum of irreducible modules).*

*Proof.* Set

$$V^0 = \{ v \in V \mid \mathfrak{n}_+(v) = 0 \}.$$

This is  $\mathfrak{h}$ -invariant, hence we have the weight-space decomposition

$$V^0 = \bigoplus_{\lambda \in L} V_\lambda^0,$$

where all elements from  $L$  are primitive weights.

Let  $\lambda \in L$  and  $v \in V_\lambda^0$ ,  $v \neq 0$ . Then the  $\mathfrak{g}(A)$ -module  $U(\mathfrak{g})(v)$  is irreducible (and hence isomorphic to  $L(\lambda)$ ). Indeed, if this is not the case, then by Proposition 9.3(b) we have  $U(\mathfrak{n}_-)(v) \cap V_\mu^0 \neq 0$  for some  $\mu < \lambda$ . This contradicts the assumption of the lemma. Therefore, the  $\mathfrak{g}(A)$ -submodule  $V'$  of  $V$  generated by  $V^0$  is completely reducible.

It remains to show that  $V' = V$ . If this is not the case, we consider the  $\mathfrak{g}(A)$ -module  $V/V'$ . Then there exists a weight vector  $v \in V$  of weight  $\mu$  such that  $v \notin V'$  but  $e_i(v) \in V'$  and  $\neq 0$  for some  $i$ . But then, since  $V$  is from the category  $\mathcal{O}$ , there exists  $\lambda \in L$  such that  $\lambda \geq \mu + \alpha_i$ , and hence  $\lambda > \mu$ , which contradicts the assumption of the lemma.  $\square$

The following lemma establishes the existence of a "bounded from below" filtration of any module in the category  $\mathcal{O}$ .

**Lemma 8.3 (9.6).** *Let  $V \in \mathcal{O}$  and  $\lambda \in \mathfrak{h}^*$ . Then there exists a filtration by a sequence of submodules*

$$V = V_t \supset V_{t-1} \supset \cdots \supset V_1 \supset V_0 = 0$$

*and a subset  $J \subset \{1, \dots, t\}$  such that:*

- (i) *if  $j \in J$ , then  $V_j/V_{j-1} \simeq L(\lambda_j)$  for some  $\lambda_j \geq \lambda$ ;*
- (ii) *if  $j \notin J$ , then  $(V_j/V_{j-1})_\mu = 0$  for every  $\mu \geq \lambda$ .*

*Proof.* Define

$$a(V, \lambda) = \sum_{\mu \geq \lambda} \dim V_\mu.$$

We prove the lemma by induction on  $a(V, \lambda)$ .

If  $a(V, \lambda) = 0$ , then there are no weights  $\mu \geq \lambda$ , hence  $V = 0$  and the required filtration is  $0 = V_0 = V_1 = V$  with  $J = \emptyset$ .

Assume  $a(V, \lambda) > 0$ . Choose a maximal element  $\mu \in P(V)$  such that  $\mu \geq \lambda$ , and pick a weight vector  $v \in V_\mu$ . Let  $U = U(\mathfrak{g})(v)$ . Then  $U$  is a highest-weight module. By Proposition 9.3(c),  $U$  contains a maximal proper submodule  $\overline{U}$ , and we have

$$0 \subset \overline{U} \subset U \subset V, \quad U/\overline{U} \simeq L(\mu), \quad \mu \geq \lambda.$$

Since the top weight  $\mu$  occurs in  $V$  but not in  $\overline{U}$  or  $V/U$ , we have

$$a(\overline{U}, \lambda) < a(V, \lambda), \quad a(V/U, \lambda) < a(V, \lambda).$$

Indeed,  $a(V, \lambda)$  counts the total dimension of all weight spaces  $V_\nu$  with  $\nu \geq \lambda$ . Passing to  $\overline{U}$  or  $V/U$  removes at least the contribution from the weight  $\mu$ , so the inequalities are strict.

By the induction hypothesis, both  $\overline{U}$  and  $V/U$  possess filtrations satisfying the conditions of the lemma. Say

$$0 = U_0 \subset U_1 \subset \cdots \subset U_p = \overline{U}, \quad 0 = W_0 \subset W_1 \subset \cdots \subset W_q = V/U.$$

Combine them to obtain a filtration of  $V$ :

$$0 = U_0 \subset \cdots \subset U_p = \overline{U} \subset U_{p+1} := U \subset \pi^{-1}(W_1) \subset \cdots \subset \pi^{-1}(W_q) = V,$$

where  $\pi : V \rightarrow V/U$  is the natural projection. Each successive quotient in this sequence is either one of the quotients from the filtration of  $\overline{U}$ , the factor  $U/\overline{U} \simeq L(\mu)$ , or one of the quotients from the filtration of  $V/U$ .

This gives the required filtration of  $V$  satisfying (i) and (ii).  $\square$

**Remark 8.4 (Different than semisimple Category  $\mathcal{O}$ ).** When  $\mathfrak{g}$  is finite semisimple, the filtration of Lemma 9.6 can be refined to a Jordan-Hölder series, i.e. a filtration where every successive quotient is irreducible, i.e. every module in Category  $\mathcal{O}$  has finite length.

This is because the BGG theorem says that a Verma module  $M(\Lambda)$  has a primitive vector of weight  $\mu$  iff  $\mu = w \cdot \Lambda$  for some  $w \in W$ , where  $W$  is the Weyl group and  $w \cdot \Lambda = w(\Lambda + \rho) - \rho$ . Since  $W$  is finite, so the only possible singular weights are a finite set. Thus  $M(\Lambda)$  has finitely many singular vectors (possibly none if  $(\Lambda + \rho, \beta^\vee) \notin \mathbb{Z}_{>0}$  for all  $\beta \in \Delta_+$ ).

However, when  $\mathfrak{g}$  is an infinite-dimensional Kac-Moody algebra, the situation is more complicated. For example, if  $\mathfrak{g}$  is an affine Kac-Moody algebra, then there are infinitely many positive roots (the imaginary roots), and hence infinitely many possible singular vectors in  $M(\Lambda)$ .  $(\Lambda + \rho, (\alpha + n\delta)^\vee) = (\Lambda + \rho, \alpha^\vee)$ . So if  $(\Lambda + \rho, \alpha^\vee) = m \in \mathbb{Z}_{>0}$  for one  $\alpha$ , you get singular vectors at  $\Lambda - m(\alpha + n\delta)$  ( $n = 0, 1, 2, \dots$ ).

Thus, a Verma module  $M(\Lambda)$  may have infinitely many submodules, and hence may not have finite length. Therefore, the filtration of Lemma 9.6 cannot always be refined to a Jordan-Hölder series in the infinite-dimensional case. However, it is good enough to define multiplicities because we can essentially truncate from below. If you fix a weight cutoff  $\lambda$ , then above that cutoff (weights  $\geq \lambda$ ), every  $V \in \mathcal{O}$  has only finitely many relevant weights.

Let  $[V : L(\lambda)]$  be the multiplicity of  $L(\lambda)$  in a module  $V \in \mathcal{O}$ , i.e. the number of times  $\lambda$  appears as a  $\lambda_j$  in the filtration of Lemma 9.6. This number is independent of the choice of filtration by

standard Jordan-Hölder arguments (in particular, Category  $\mathcal{O}$  is a finite-length abelian category and every simple object is isomorphic to some  $L(\Lambda)$ ).

Moreover,  $[V : L(\mu)] \neq 0$  if and only if  $\mu$  is a primitive weight of  $V$ . This follows because if  $\mu$  is a primitive weight of  $V$ , then we say that  $V$  has a subquotient isomorphic to  $L(\mu)$ , and conversely, if  $[V : L(\mu)] \neq 0$ , then  $V$  has a subquotient isomorphic to  $L(\mu)$ , which contains a primitive vector of weight  $\mu$ .

## 8.2 Formal characters

Let  $\mathcal{E}$  be the vector space of all formal sums

$$\sum_{\lambda \in \mathfrak{h}^*} c_\lambda e^\lambda, \quad c_\lambda \in \mathbb{C},$$

so that  $c_\lambda = 0$  for  $\lambda$  outside a finite union of sets of the form  $D(\mu) = \{\nu \in \mathfrak{h}^* \mid \nu \leq \mu\}$ . Let  $V$  be a module in the category  $\mathcal{O}$ . The **formal character** of  $V$  is defined to be

$$\text{ch } V = \sum_{\lambda \in \mathfrak{h}^*} \dim V_\lambda e^\lambda \in \mathcal{E}.$$

**Proposition 8.5 (9.7).** *Let  $V$  be a  $\mathfrak{g}(A)$ -module from the category  $\mathcal{O}$ . Then*

$$\text{ch } V = \sum_{\lambda \in \mathfrak{h}^*} [V : L(\lambda)] \text{ch } L(\lambda). \quad (45)$$

*Proof.* Denote by  $\phi$  the map which associates to each  $V \in \mathcal{O}$  the difference

$$\phi(V) \in \mathcal{E}$$

between the left- and right-hand sides of (45). Then  $\phi(L(\lambda)) = 0$ , and given an exact sequence of modules

$$0 \rightarrow V_1 \rightarrow V_2 \rightarrow V_3 \rightarrow 0,$$

we have  $\phi(V_2) = \phi(V_1) + \phi(V_3)$ .  $\phi$  is additive because both the character function and the composition multiplicities are additive on short exact sequences of modules in category  $\mathcal{O}$ .

Using Lemma 9.6, we deduce that given  $\lambda \in \mathfrak{h}^*$ , there exist modules  $M_1, \dots, M_r \in \mathcal{O}$  such that  $(M_i)_\mu = 0$  when  $\mu \geq \lambda$ , and

$$\phi(V) = \sum_{i=1}^r \phi(M_i).$$

In particular, for every  $\lambda \in \mathfrak{h}^*$ , the coefficient at  $e(\lambda)$  in  $\phi(V)$  is zero.  $\square$

By the PBW theorem, the character of a Verma module is given by

$$\text{ch } M(\lambda) = e^\lambda \prod_{\alpha \in \Delta_+} (1 + e^{-\alpha} + e^{-2\alpha} + \dots)^{\dim \mathfrak{g}_\alpha}.$$

or equivalently,

$$\text{ch } M(\lambda) = \frac{e^\lambda}{\prod_{\alpha \in \Delta_+} (1 - e^{-\alpha})^{\dim \mathfrak{g}_\alpha}}.$$

**Lemma 8.6 (9.8).**

(a) *If  $V$  is a  $\mathfrak{g}(A)$ -module with highest weight  $\Lambda$ , then*

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

(b) *If  $V$  is a module from the category  $\mathcal{O}$  and  $v$  is a primitive vector with weight  $\lambda$ , then there exists a submodule  $U \subset V$  such that  $v \notin U$  and*

$$\Omega(v) = (|\lambda + \rho|^2 - |\rho|^2)v \pmod{U}.$$

*Proof.* Follows immediately from Corollary 2.6.  $\square$

**Proposition 8.7 (9.8).** *Let  $V$  be a  $\mathfrak{g}(A)$ -module with highest weight  $\Lambda$ . Then*

$$\text{ch } V = \sum_{\substack{\lambda \leq \Lambda \\ |\lambda + \rho|^2 = |\Lambda + \rho|^2}} c_\lambda \text{ch } M(\lambda), \quad \text{where } c_\lambda \in \mathbb{Z}, \ c_\Lambda = 1. \quad (46)$$

*Proof.* Using (45), it suffices to prove (46) for  $V = L(\Lambda)$ . Set

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

and order the elements of this set  $\lambda_1, \lambda_2, \dots$  so that the inequality  $\lambda_i \geq \lambda_j$  implies  $i \leq j$ . Then the proof of Proposition 9.7 and Lemma 9.8 imply the following system of equations:

$$\text{ch } M(\lambda_i) = \sum_{\lambda_j \in B(\Lambda)} c_{ij} \text{ch } L(\lambda_j).$$

The matrix  $(c_{ij})$  of this system is triangular with ones on the diagonal. Solving this system proves (46).  $\square$

Note that this proposition and the previous are exactly the same as in the finite-dimensional case.



**Proposition 8.8 (9.9).** *Let  $A$  be a symmetrizable matrix.*

- (a) *If  $2(\Lambda + \rho \mid \beta) \neq (\beta \mid \beta)$  for every  $\beta \in Q_+$ ,  $\beta \neq 0$ , then the  $\mathfrak{g}(A)$ -module  $M(\Lambda)$  is irreducible.*
- (b) *If  $V$  is a  $\mathfrak{g}(A)$ -module from the category  $\mathcal{O}$  such that, for any two primitive weights  $\lambda$  and  $\mu$  of  $V$  with  $\lambda - \mu = \beta > 0$ , one has  $2(\lambda + \rho \mid \beta) \neq (\beta \mid \beta)$ , then  $V$  is completely reducible.*

*Proof.* Proposition 9.3(b) implies that if  $M(\Lambda)$  is not irreducible, then there exists a primitive weight  $\lambda = \Lambda - \beta$ , where  $\beta > 0$ . But then Lemma 9.8(a) gives

$$|\Lambda + \rho|^2 = |\Lambda - \beta + \rho|^2$$

which is equivalent to

$$2(\Lambda + \rho \mid \beta) = (\beta \mid \beta)$$

To prove (b), we may assume that the  $\mathfrak{g}(A)$ -module  $V$  is indecomposable. Since, clearly,  $\Omega$  is locally finite on  $V$ , i.e. every  $v \in V$  lies in a finite-dimensional  $\Omega$ -invariant subspace, we obtain that there exists  $a \in \mathbb{C}$  such that  $\Omega - aI$  is locally nilpotent on  $V$ . Hence Lemma 9.8(b) implies  $|\lambda + \rho|^2 = |\mu + \rho|^2$  for any two primitive weights  $\lambda$  and  $\mu$ . Now (b) follows from Lemma 9.5 because in fact we see that  $\lambda \geq \mu$  implies  $\lambda = \mu$ .  $\square$

### 8.3 Shapovalov form

Let  $L(\Lambda)^*$  be the  $\mathfrak{g}(A)$ -module contragredient to  $L(\Lambda)$ . Then

$$L(\Lambda)^* = \prod_{\lambda} (L(\Lambda)_{\lambda})^*.$$

The subspace

$$L^*(\Lambda) := \bigoplus_{\lambda} (L(\Lambda)_{\lambda})^*$$

is a submodule of the  $\mathfrak{g}(A)$ -module  $L(\Lambda)^*$ . It is clear that the module  $L^*(\Lambda)$  is irreducible and that for  $v \in (L(\Lambda)_{\Lambda})^*$  one has:

$$\mathfrak{n}_-(v) = 0, \quad h(v) = -\Lambda(h)v \quad \text{for } h \in \mathfrak{h}.$$

Such a module is called an **irreducible module with lowest weight**  $-\Lambda$ . As in §9.3 we have a bijection between  $\mathfrak{h}^*$  and irreducible lowest-weight modules:  $\Lambda \mapsto L^*(-\Lambda)$ .

Denote by  $\pi_{\Lambda}$  the action of  $\mathfrak{g}(A)$  on  $L(\Lambda)$ , and introduce the new action  $\pi_{\Lambda}^*$  on the space  $L(\Lambda)$  by

$$\pi_{\Lambda}^*(g)v = \pi_{\Lambda}(\omega(g))v, \tag{9.4.1}$$

where  $\omega$  is the Chevalley involution of  $\mathfrak{g}(A)$ . It is clear that  $(L(\Lambda), \pi_\Lambda^*)$  is an irreducible  $\mathfrak{g}(A)$ -module with lowest weight  $-\Lambda$ . By the uniqueness theorem, this module can be identified with  $L^*(\Lambda)$ , and the pairing between  $L(\Lambda)$  and  $L^*(\Lambda)$  gives us a nondegenerate bilinear form  $B$  on  $L(\Lambda)$  such that

$$B(gx, y) = -B(x, \omega(g)(y)) \quad \text{for all } g \in \mathfrak{g}(A), x, y \in L(\Lambda). \quad (9.4.2)$$

A bilinear form on  $L(\Lambda)$  which satisfies (9.4.2) is called a **contravariant bilinear form**.

**Proposition 8.9 (9.4).** *Every  $\mathfrak{g}(A)$ -module  $L(\Lambda)$  carries a unique up to constant factor nondegenerate contravariant bilinear form  $B$ . This form is symmetric, and  $L(\Lambda)$  decomposes into an orthogonal direct sum of weight spaces with respect to this form.*

*Proof.* The existence of  $B$  was proved above, the uniqueness follows from Lemma 9.3. The symmetry follows from the uniqueness. The fact that  $B(L(\Lambda)_\lambda, L(\Lambda)_\mu) = 0$  if  $\lambda \neq \mu$  follows from (9.4.2) for  $g \in \mathfrak{h}$ ,  $x \in L(\Lambda)_\lambda$ ,  $y \in L(\Lambda)_\mu$ .  $\square$

A more explicit way to introduce the contravariant bilinear form is the following. Let  $V$  be a highest-weight  $\mathfrak{g}(A)$ -module with a fixed highest-weight vector  $v_\Lambda$ . Given  $v \in V$ , we define its **expectation value**  $\langle v \rangle \in \mathbb{C}$  by

$$v = \langle v \rangle v_\Lambda + \sum_{\alpha \in Q_+ \setminus \{0\}} v_{\Lambda-\alpha}, \quad v_{\Lambda-\alpha} \in V_{\Lambda-\alpha}.$$

## 8.4 Virasoro algebra

Recall that the Virasoro algebra  $\text{Vir}$  is the Lie algebra with basis  $\{d_n \mid n \in \mathbb{Z}\} \cup \{c\}$  and relations

$$[d_m, d_n] = (m - n)d_{m+n} + \delta_{m,-n} \frac{m^3 - m}{12} c, \quad [d_n, c] = 0.$$

Define the **triangular decomposition** of  $\text{Vir}$  as follows:

$$\text{Vir} = \text{Vir}_- \oplus \text{Vir}_0 \oplus \text{Vir}_+, \quad \text{Vir}_\pm = \bigoplus_{j>0} \mathbb{C}d_{\pm j}, \quad \text{Vir}_0 = \mathbb{C}c \oplus \mathbb{C}d_0.$$

Given  $c, h \in \mathbb{C}$ , define a  $\text{Vir}$ -module  $V$  with highest weight  $(c, h)$  by the requirement that there exists a nonzero vector  $v = v_{c,h}$  such that

$$\text{Vir}_+(v) = 0, \quad U(\text{Vir}_-)v = V, \quad d_0(v) = hv, \quad c(v) = cv.$$

We denote the eigenvalue of a central element by the same letter.

The definition of the Verma module  $M(c, h)$  over  $\text{Vir}$ , as well as the proof of its existence and uniqueness, are given in the same way as in for general Kac–Moody algebras. It is clear that  $c$  acts on  $M(c, h)$  as  $cI$ . The number  $c$  is called the **conformal central charge**. By the Poincaré–Birkhoff–Witt theorem, the elements

$$d_{-j_n} \cdots d_{-j_2} d_{-j_1}(v_{c,h}), \quad 0 < j_1 \leq j_2 \leq \cdots, \quad (47)$$

form a basis of  $M(c, h)$ . Since  $[d_0, d_{-n}] = nd_{-n}$ , we see that  $d_0$  is diagonalizable on  $M(c, h)$  with spectrum  $h + \mathbb{Z}_+$  and with the eigenspace decomposition

$$M(c, h) = \bigoplus_{j \in \mathbb{Z}_+} M(c, h)_{h+j}, \quad (48)$$

where  $M(c, h)_{h+j}$  is spanned by elements of the form (47) with  $j_1 + \cdots + j_n = j$ . It follows that

$$\dim M(c, h)_{h+j} = p(j), \quad (49)$$

where  $p(j)$  is the classical partition function. Thus, the Kostant partition function for the Virasoro algebra is just the classical partition function. Equation (49) can be rewritten as follows:

$$\text{tr}_{M(c,h)} q^{d_0} := \sum_{\lambda} \dim M(c, h)_{\lambda} q^{\lambda} = q^h \prod_{j=1}^{\infty} (1 - q^j)^{-1}.$$

The series  $\text{tr}_V q^{d_0}$  is defined for any diagonalizable operator  $d_0$  on a vector space  $V$  with eigenvalues  $\lambda_i$  (written with multiplicity) by the formula

$$\text{tr}_V q^{d_0} := \sum_i q^{\lambda_i}$$

is called the **formal character** of the  $\text{Vir}$ -module  $V$ .

As with Kac–Moody algebras, one shows that there exists a unique irreducible  $\text{Vir}$ -module  $L(c, h)$  with highest weight  $(c, h)$ .

The Chevalley involution  $\omega$  of  $\text{Vir}$  is defined by

$$\omega(d_n) = -d_{-n}, \quad \omega(c) = -c. \quad (50)$$

The contravariant bilinear form  $B$  on a  $\text{Vir}$ -module is defined as in §9.4. In other words, this is a symmetric bilinear form with respect to which  $d_n$  and  $d_{-n}$  are adjoint operators. Its existence, uniqueness, and construction are established in the same way.

## 8.5 Integrable highest weight modules

Let  $\mathfrak{g}(A)$  be a Kac–Moody algebra of rank  $n$  and let  $\mathfrak{h}$  be its Cartan subalgebra. Set

$$\begin{aligned} P &= \{\lambda \in \mathfrak{h}^* \mid \langle \lambda, \alpha_i^\vee \rangle \in \mathbb{Z} \ (i = 1, \dots, n)\}, \\ P_+ &= \{\lambda \in P \mid \langle \lambda, \alpha_i^\vee \rangle \geq 0 \ (i = 1, \dots, n)\}, \\ P_{++} &= \{\lambda \in P \mid \langle \lambda, \alpha_i^\vee \rangle > 0 \ (i = 1, \dots, n)\}. \end{aligned}$$

The set  $P$  is called the **weight lattice**; elements from  $P$  (resp.  $P_+$  or  $P_{++}$ ) are called **integral weights** (resp. **dominant** or **regular dominant integral weights**). Note that  $P$  contains the root lattice  $Q$ .

Let  $V$  be a highest-weight module over  $\mathfrak{g}(A)$ , and  $v$  a highest-weight vector. Recall a highest-weight module  $V$  with highest-weight vector  $v$  is integrable if for every simple root  $\alpha_i$ ,  $e_i, f_i$  act locally nilpotently — that is,  $f_i^{N_i}(v) = 0$  for some  $N_i > 0$ . This means the subalgebra generated by  $e_i, f_i, h_i$  (an  $\mathfrak{sl}_2$  inside  $\mathfrak{g}$ ) acts as a finite-dimensional representation of  $\mathfrak{sl}_2$ .

It follows from Lemmas 3.4(b) and 3.5 that the module  $V$  is integrable if and only if  $f_i^{N_i}(v) = 0$  for some  $N_i > 0$  ( $i = 1, \dots, n$ ).

**Lemma 8.10 (Integrable highest-weight modules).** *The  $\mathfrak{g}(A)$ -module  $L(\Lambda)$  is integrable if and only if  $\Lambda \in P_+$ .*

*Proof.* We have the formula from the representation theory of  $\mathfrak{sl}_2$ :

$$[e_i, f_i^m] = m f_i^{m-1} h_i - m(m-1) f_i^{m-2} e_i$$

If we apply this to the highest-weight vector  $v$  with  $h_i v = \langle \Lambda, \alpha_i^\vee \rangle v$  and  $e_i v = 0$ , we get

$$e_i f_i^m v = m(\langle \Lambda, \alpha_i^\vee \rangle - m + 1) f_i^{m-1} v.$$

$N_i$  is the smallest integer such that  $f_i^{N_i} v = 0$ . So from the formula, when  $m = N_i = \langle \Lambda, \alpha_i^\vee \rangle + 1$  we have  $e_i f_i^{N_i} v = 0$ . In particular,  $f_i^{N_i} v$  is annihilated by  $e_i$  and also by  $e_j$  for  $j \neq i$  since  $e_j$  commutes with  $f_i$ . Thus,  $f_i^{N_i} v$  is a primitive vector. Since  $L(\Lambda)$  is irreducible, the only primitive vector is the highest-weight vector  $v$ , so we must have  $f_i^{N_i} v = 0$ .  $N$  must be a positive integer, so  $\langle \Lambda, \alpha_i^\vee \rangle \geq 0$  for all  $i$ , i.e.  $\Lambda \in P_+$ .

Conversely, if  $\Lambda \in P_+$ , then  $\langle \Lambda, \alpha_i^\vee \rangle$  is a nonnegative integer for all  $i$ . So we can set  $N_i = \langle \Lambda, \alpha_i^\vee \rangle + 1$  and the same calculation shows that  $f_i^{N_i} v$  is a primitive vector. Since  $L(\Lambda)$  is irreducible, we must have  $f_i^{N_i} v = 0$ . Thus,  $L(\Lambda)$  is integrable.  $\square$

**Corollary 8.11 (Classification of integrable highest-weight modules for untwisted affine algebras).** *For an untwisted affine  $\widehat{\mathfrak{g}}$  at level  $k$ , the integrable highest-weight modules are in bijection with  $\bar{\lambda} \in \bar{P}_+$  such that  $\langle \bar{\lambda}, \theta^\vee \rangle \leq k$ , i.e. the integral points of the closed level- $k$  affine alcove*

$${}^k C_+^{af} = \{x \in \bar{\mathfrak{h}}_\mathbb{R}^* : \langle x, \alpha_i^\vee \rangle \geq 0 \ (i = 1, \dots, \ell), \ \langle x, \theta^\vee \rangle \leq k\}.$$

Given such a  $\bar{\lambda}$ , the highest weight is (choose the energy so the highest vector has grade 0)

$$\Lambda = k \Lambda_0 + \bar{\lambda}.$$

*Proof.* The integrable highest-weight condition is  $\langle \Lambda, \alpha_i^\vee \rangle \in \mathbb{Z}_{\geq 0}$  for  $i = 0, 1, \dots, \ell$ . Since  $\alpha_0^\vee = K - \theta^\vee$  and  $\langle \delta, \alpha_i^\vee \rangle = 0$ , we get

$$\langle \Lambda, \alpha_i^\vee \rangle = \begin{cases} \langle \bar{\lambda}, \alpha_i^\vee \rangle & (i \geq 1) \\ k - \langle \bar{\lambda}, \theta^\vee \rangle & (i = 0) \end{cases}$$

so  $s$  never appears. Thus, at fixed level  $k$ , integrable highest weights are classified by  $\bar{\lambda} \in \bar{P}_+$  satisfying  $\langle \bar{\lambda}, \theta^\vee \rangle \leq k$ .

The reason why we can split by setting  $s = 0$  is as follows. Consider the automorphism of the affine algebra:

$$\phi_c : d \mapsto d - cK, \quad K \mapsto K$$

with  $\bar{\mathfrak{g}} \otimes \mathbb{C}[t, t^{-1}]$  fixed. Precomposing the representation by  $\phi_c$  changes the highest weight by  $\Lambda \mapsto \Lambda - ck\delta$  (since  $\langle \Lambda, d - cK \rangle = s - ck$ ). Choosing  $c = s/k$  makes the new highest weight have grade 0. This doesn't alter  $k$  or  $\bar{\lambda}$ , so it preserves integrability and the finite labels. At level 0 you cannot shift  $s$  this way (the trick needs  $k \neq 0$ ). But the only integrable highest-weight module at  $k = 0$  is the trivial one, where  $s$  is irrelevant anyway.  $\square$

Note that the adjoint representation of an affine Kac-Moody algebra is level 0 but it is not a highest weight module.

Denote by  $P(\Lambda)$  the set of weights of  $L(\Lambda)$ . It is clear that  $P(\Lambda) \subset P$  if  $\Lambda \in P$ . The following proposition follows from Lemma 10.1 and Proposition 3.7(a).

**Proposition 8.12 (10.1).** *If  $\Lambda \in P_+$ , then*

$$\text{mult}_{L(\Lambda)} \lambda = \text{mult}_{L(\Lambda)} w(\lambda) \quad \text{for all } w \in W.$$

*In particular,  $P(\Lambda)$  is  $W$ -invariant.*

**Corollary 8.13 (10.1).** *If  $\Lambda \in P_+$ , then any  $\lambda \in P(\Lambda)$  is  $W$ -equivalent to a unique  $\mu \in P_+ \cap P(\Lambda)$ .*

*Proof.* Take  $\mu \in W \cdot \lambda$  such that  $\text{ht}(\Lambda - \mu)$  is minimal; Lemma 3.12(b) implies the uniqueness.  $\square$

We let the Weyl group  $W$  act on the complex vector space  $\tilde{\mathcal{E}}$  of all (possibly infinite) linear combinations of formal exponentials by

$$w \left( \sum_{\lambda} c_{\lambda} e(\lambda) \right) = \sum_{\lambda} c_{\lambda} e(w(\lambda)) \quad (w \in W).$$

The space  $\tilde{\mathcal{E}}$  contains  $\mathcal{E}$  as a subspace. However, the product of two elements  $P_1, P_2 \in \tilde{\mathcal{E}}$  doesn't always make sense, but if it does, then  $w(P_1 P_2) = w(P_1)w(P_2)$ . Proposition 10.1 says that

$$w(\text{ch } L(\Lambda)) = \text{ch } L(\Lambda) \quad (w \in W, \Lambda \in P_+). \quad (51)$$

Consider now the element (cf. §9.7)

$$R = \prod_{\alpha \in \Delta_+} (1 - e(-\alpha))^{\text{mult } \alpha} \in \mathcal{E}.$$

Fix an element  $\rho \in \mathfrak{h}^*$  such that (cf. §2.5)

$$\langle \rho, \alpha_i^\vee \rangle = 1 \quad (i = 1, \dots, n).$$

For  $w \in W$  set  $\varepsilon(w) = (-1)^{\ell(w)}$ . By (3.11.1) we have

$$\varepsilon(w) = \det_{\mathfrak{h}} w.$$

Furthermore, one has

$$w(e(\rho)R) = \varepsilon(w) e(\rho)R \quad (w \in W). \quad (52)$$

Indeed, it is sufficient to check (52) for each fundamental reflection  $r_i$ . Recall that by Lemma 3.7, the set  $\Delta_+ \setminus \{\alpha_i\}$  is  $r_i$ -invariant and, by Proposition 3.7, we have  $\text{mult } r_i(\alpha) = \text{mult } \alpha$  for  $\alpha \in \Delta_+$ . Hence,

$$\begin{aligned} r_i(e(\rho)R) &= e(\rho - \alpha_i) r_i(1 - e(-\alpha_i)) r_i \prod_{\alpha \in \Delta_+ \setminus \{\alpha_i\}} (1 - e(-\alpha))^{\text{mult } \alpha} \\ &= e(\rho) e(-\alpha_i) (1 - e(\alpha_i)) \prod_{\alpha \in \Delta_+ \setminus \{\alpha_i\}} (1 - e(-\alpha))^{\text{mult } \alpha} \\ &= -e(\rho) R = \varepsilon(r_i) e(\rho) R. \end{aligned}$$

Now we can prove the following fundamental result of our representation theory.

**Theorem 8.14 (Weyl–Kac character formula).** *Let  $\mathfrak{g}(A)$  be a symmetrizable Kac–Moody algebra, and let  $L(\Lambda)$  be an irreducible  $\mathfrak{g}(A)$ -module with highest weight  $\Lambda \in P_+$ . Then*

$$\text{ch } L(\Lambda) = \frac{\sum_{w \in W} \varepsilon(w) e(w(\Lambda + \rho) - \rho)}{\prod_{\alpha \in \Delta_+} (1 - e(-\alpha))^{\text{mult } \alpha}}. \quad (53)$$

*Proof.* The proof is the same as in the finite-dimensional case, except one should note that the action of the Weyl group  $W$  also preserves the multiplicities of roots.  $\square$

## 8.6 Integrable modules by level

Recall that the center of an affine algebra  $\mathfrak{g}(A)$  is 1-dimensional and is spanned by the canonical central element:

$$K = \sum_{i=0}^{\ell} a_i^{\vee} \alpha_i^{\vee}.$$

It is clear that  $K$  operates on a  $\mathfrak{g}(A)$ -module  $L(\Lambda)$  by the scalar operator  $\langle \Lambda, K \rangle I_{L(\Lambda)}$ . In particular,  $\langle \Lambda, K \rangle = \langle \lambda, K \rangle$  for every  $\lambda \in P(\Lambda)$ . The number

$$k := \langle \Lambda, K \rangle = \sum_{i=0}^{\ell} a_i^{\vee} \langle \Lambda, \alpha_i^{\vee} \rangle \quad (54)$$

is called the **level** of  $\Lambda \in \mathfrak{h}^*$ , or of the module  $L(\Lambda)$ .

If  $\Lambda \in P_+$ , then the level of  $\Lambda$  is a nonnegative integer; it is zero if and only if all the labels  $\langle \Lambda, \alpha_i^{\vee} \rangle$  of  $\Lambda$  are zero. Hence, an integrable  $\mathfrak{g}(A)$ -module  $L(\Lambda)$  has level 0 if and only if  $\dim L(\Lambda) = 1$ ; if  $L(\Lambda)$  is integrable and  $\dim L(\Lambda) \neq 1$ , then the level of  $L(\Lambda)$  is a positive integer.

Note that the level of  $\rho$  is equal to the **dual Coxeter number**

$$\langle \rho, K \rangle = \sum_{i=0}^{\ell} a_i^{\vee} = h^{\vee}, \quad (55)$$

Let  $\Lambda_i$  ( $i = 0, \dots, \ell$ ) be the **fundamental weights**:

$$\langle \Lambda_i, \alpha_j^{\vee} \rangle = \delta_{ij}, \quad j = 0, \dots, \ell, \quad \text{and} \quad \langle \Lambda_i, d \rangle = 0.$$

These conditions force us to take (recall that we wrote  $\bar{\lambda}$  for the restriction of  $\lambda \in \mathfrak{h}^*$  to  $\bar{\mathfrak{h}}^*$ , in particular throw away the  $\Lambda_0$  and  $\delta$  components)

$$\Lambda_i = \bar{\Lambda}_i + a_i^{\vee} \Lambda_0, \quad (56)$$

where  $\bar{\Lambda}_0 = 0$  and  $\bar{\Lambda}_1, \dots, \bar{\Lambda}_{\ell}$  are the fundamental weights of  $\bar{\mathfrak{g}}$ . Note also that

$$P_+ = \sum_{i=0}^{\ell} \mathbb{Z}_+ \Lambda_i + \mathbb{C} \delta. \quad (57)$$

It is clear that the level of  $\Lambda$  from  $P_+$  is 1 if and only if  $\Lambda \equiv \Lambda_i \pmod{\mathbb{C} \delta}$  and  $i$  is such that  $a_i^{\vee} = 1$ ; in particular, the level of  $\Lambda_0$  is always 1. A glance at Table Aff gives that if  $A$  is symmetric or  $r > 1$ , then

$$\text{level}(\Lambda_i) = 1 \quad \text{if and only if} \quad i \in (\text{Aut } S(A)) \cdot 0. \quad (58)$$

Finally, the following observation, which follows from the description of the fundamental chamber will be useful. It says that there is a unique dominant element in each  $W$ -orbit at positive level; for integral  $\lambda$  that unique element is in  $P_+$ .

**Lemma 8.15 (12.4).** *If  $\operatorname{Re}\langle\lambda, K\rangle > 0$ , then*

$$(W \cdot \lambda) \cap \{\mu \in \mathfrak{h}^* \mid \operatorname{Re}\langle\mu, \alpha_i^\vee\rangle \geq 0 \text{ for all } i\}$$

*consists of a single element. In particular, if  $\lambda \in P$  has a positive level, then the set  $P_+ \cap W \cdot \lambda$  consists of a single element.*

**Remark 8.16 (Relationship to the fundamental chamber and alcove).** *In the full affine Cartan dual  $\mathfrak{h}_{\mathbb{R}}^* = \bar{\mathfrak{h}}_{\mathbb{R}}^{\oplus} \mathbb{R}\Lambda_0 \oplus \mathbb{R}\delta$ , recall that we have the fundamental chamber*

$$C_+^{\text{aff}} := \{\lambda \mid \langle\lambda, \alpha_i^\vee\rangle \geq 0 \ (i = 1, \dots, \ell), \ \langle\lambda, K - \theta^\vee\rangle \geq 0\}$$

*which is a fundamental domain for the Weyl group action.*

*Intersect with the level- $k$  hyperplane  $H_k = \{\langle\lambda, K\rangle = k\}$ . The fundamental alcove at level  $k$  is:*

$$A_+^{(k)} = \{\lambda \in H_k \mid 0 \leq \langle\lambda, \alpha_i^\vee\rangle \ (i = 1, \dots, \ell), \ \langle\lambda, \theta^\vee\rangle \leq k\}$$

*At level  $k = 1$ , this is exactly the fundamental chamber defined earlier.*

*This is the fundamental domain for the action of the affine Weyl group  $W^{\text{aff}}$  on  $H_k$  modulo  $\mathbb{R}\delta$ . On the level slice  $H_k = \{\langle\lambda, K\rangle = k\}$ , the affine Weyl group  $W^{\text{aff}}$  acts trivially along the  $\delta$ -direction (reflections don't change the  $\delta$ -component).*

*Therefore a true fundamental domain on  $H_k \subset \mathfrak{h}_{\mathbb{R}}^*$  is the alcove times the whole  $\delta$ -line  $A_+^{(k)} + \mathbb{R}\delta$ . Every  $W^{\text{aff}}$ -orbit in  $H_k$  meets  $\mathcal{F}_k$  in exactly one point (one point per fixed  $\delta$ -fiber), and the action is free on alcove interiors. If you mod out  $\mathbb{R}\delta$  (or equivalently fix  $\langle\lambda, d\rangle = 0$  to pick a transverse slice), then  $\delta$  disappears and the bounded alcove itself  $A_+^{(k)}$  is the fundamental domain for the induced  $W^{\text{aff}}$ -action.*

We let

$$P^k \text{ (resp. } P_+^k) = \{\lambda \in P \text{ (resp. } P_+) \mid \langle\lambda, K\rangle = k\}$$

Note that the action of  $W$  preserves the level and therefore  $P^k$  and  $P_+^k$  are  $W$ -stable.

We collect here some facts about the weight system  $P(\Lambda)$  of an integrable module  $L(\Lambda)$  over an affine algebra  $\mathfrak{g}(A)$ , proved earlier in the general context of Kac-Moody algebras. First we recall an important proposition which gives conditions on when  $\mu$  is a weight of  $L(\Lambda)$  in terms of norms.

**Proposition 8.17 (Norm bounds on weights).** *Let  $\Lambda \in P_+$  and  $\lambda, \mu \in P(\Lambda)$ . Then*

1.  $\langle\Lambda, \Lambda\rangle - \langle\lambda, \mu\rangle \geq 0$ , and equality holds if and only if  $\lambda = \mu \in W \cdot \Lambda$ .
2.  $|\Lambda + \rho|^2 - |\lambda + \rho|^2 \geq 0$ , and equality holds if and only if  $\lambda = \Lambda$ .



*Proof.* Since both the bilinear form  $(\cdot, \cdot, |\cdot, \cdot)$  and the weight set  $P(\Lambda)$  are  $W$ -invariant, we can assume in the proof of (a) that  $\lambda \in P_+$ . Let  $\beta := \Lambda - \lambda$  and  $\beta_1 := \Lambda - \mu$ , so that  $\beta, \beta_1 \in Q_+$ . Then

$$(\Lambda|\Lambda) - (\lambda|\mu) = (\Lambda|\beta) + (\lambda|\beta_1) \geq 0$$

In the case of equality we have  $(\Lambda|\beta) = (\lambda|\beta_1) = 0$  because dominant weights always pair nonnegatively with positive roots. Since  $\lambda$  is nondegenerate with respect to  $\Lambda$  (by Lemma 11.2), we deduce that  $\beta = 0$ , i.e.  $\lambda = \Lambda$ . But then  $(\Lambda|\beta_1) = 0$  and, by the same argument,  $\mu = \Lambda$ , proving (a).

To prove (b), we compute:

$$(\Lambda + \rho|\Lambda + \rho) - (\lambda + \rho|\lambda + \rho) = ((\Lambda|\Lambda) - (\lambda|\lambda)) + 2(\Lambda - \lambda|\rho) \geq 0$$

since  $(\Lambda|\Lambda) - (\lambda|\lambda) \geq 0$  by (a) and  $\Lambda - \lambda \in Q_+$ . Equality occurs if and only if  $\Lambda = \lambda$ .  $\square$

The following proposition generalizes the paraboloid picture that we worked out in detail for  $\widehat{\mathfrak{sl}}_2$ . In particular  $P(\Lambda) = (\Lambda + Q) \cap \text{conv}(W \cdot \Lambda)$  inside the level- $k$  hyperplane, and inside a paraboloid whose boundary is  $W \cdot \Lambda$ .  $P(\Lambda)$  is  $W$ -stable, and each  $W$ -orbit has a unique dominant element  $\leq \Lambda$ . Along  $\delta$ , weights extend indefinitely downward (subtracting  $\delta$ ) with nondecreasing multiplicities.

**Proposition 8.18 (Weights of integrable modules).** *Let  $L(\Lambda)$  be an integrable module of positive level  $k$  over an affine algebra. Then:*

- (a)  $P(\Lambda) = W \cdot \{\lambda \in P_+ \mid \lambda \leq \Lambda\}$ .
- (b)  $P(\Lambda) = (\Lambda + Q)$  in the convex hull of  $W \cdot \Lambda$ .
- (c) If  $\lambda, \mu \in P(\Lambda)$  and  $\mu$  lies in the convex hull of  $W \cdot \lambda$ , then  $\text{mult}_{L(\Lambda)}(\mu) \geq \text{mult}_{L(\Lambda)}(\lambda)$ .
- (d)  $P(\Lambda)$  lies in the paraboloid

$$\{\lambda \in \mathfrak{h}_{\mathbb{R}}^* \mid \|\lambda\|^2 + 2k\langle \lambda | \Lambda_0 \rangle \leq \|\Lambda\|^2, \langle \lambda, K \rangle = k\}$$

the intersection of  $P(\Lambda)$  with the boundary of this paraboloid is  $W \cdot \Lambda$ .

- (e) For  $\lambda \in P(\Lambda)$ , the set of integers  $t \in \mathbb{Z}$  such that  $\lambda - t\delta \in P(\Lambda)$  is an interval  $[-p, +\infty)$  with  $p \geq 0$ , and the function

$$t \longmapsto \text{mult}_{L(\Lambda)}(\lambda - t\delta)$$

is nondecreasing on this interval. Moreover, if  $x \in \mathfrak{g}_{-\delta}$ ,  $x \neq 0$ , then the map

$$x : L(\Lambda)_{\lambda - t\delta} \longrightarrow L(\Lambda)_{\lambda - (t+1)\delta}$$

is injective.

(f) Setting  $\mathfrak{n}^{(\delta)} = \bigoplus_{n > 0} \mathfrak{g}_{-n\delta}$ , one has that  $L(\Lambda)$  is a free  $U(\mathfrak{n}^{(\delta)})$ -module.

*Proof.* To be filled in, but I have some remarks.

Part (d) follows from the norm bounds in the previous proposition. For an integrable highest-weight module  $L(\Lambda)$  one has

$$\|\Lambda + \rho\|^2 - \|\lambda + \rho\|^2 \geq 0 \quad (\lambda \in P(\Lambda)),$$

with equality if and only if  $\lambda \in W \cdot \Lambda$ .

Expand, cancel the  $\rho$ -constants, and rewrite in terms of  $\|\lambda\|^2$  using the identity above. This is exactly the paraboloid inequality:

$$\|\lambda\|^2 + 2k\langle \lambda, \Lambda_0 \rangle \leq \|\Lambda\|^2, \quad \langle \lambda, K \rangle = k.$$

The point about (c) is that  $\lambda$  is on the boundary of the convex hull of  $W \cdot \lambda$ , and if  $\mu$  is in the interior, then that corresponds to an  $\mathfrak{sl}_2$ -submodule where the weight  $\lambda$  is on the boundary and  $\mu$  is in the interior, so the multiplicity at  $\mu$  is larger. The statement of (a) is very familiar in the finite dimensional case.  $\square$

**Definition 8.19 (Maximal weights).** A weight  $\lambda \in P(\Lambda)$  is called **maximal** if  $\lambda + \delta \notin P(\Lambda)$ .

Denote by  $\max(\Lambda)$  the set of all maximal weights of  $L(\Lambda)$ . It is clear that  $\max(\Lambda)$  is a  $W$ -invariant set (since  $P(\Lambda)$  is  $W$ -invariant and  $\delta$  is  $W$ -fixed) and a maximal weight is  $W$ -equivalent to a unique dominant maximal weight.

On the other hand, it follows from Proposition 12.5(e) that for every  $\mu \in P(\Lambda)$  there exists a unique  $\lambda \in \max(\Lambda)$  and a unique nonnegative integer  $n$  such that

$$\mu = \lambda - n\delta$$

i.e.

$$P(\Lambda) = \bigsqcup_{\lambda \in \max(\Lambda)} \lambda - n\delta \mid n \in \mathbb{Z}_+, \quad (59)$$

where the union is disjoint. We can spell this out explicitly using the paraboloid inequality.

Write any level- $k$  weight as  $\lambda = \bar{\lambda} + k\Lambda_0 + a\delta$ , with  $\bar{\lambda} \in \bar{\mathfrak{h}}^*$ . Then the norm squared is computed using the bilinear form on  $\mathfrak{h}^*$ , which satisfies

$$(\Lambda_0, \Lambda_0) = 0, \quad (\delta, \delta) = 0, \quad (\delta, \Lambda_0) = 1, \quad (\bar{\lambda}, \Lambda_0) = (\bar{\lambda}, \delta) = 0.$$

Hence  $\|\lambda\|^2 = \|\bar{\lambda}\|^2 + 2ka$  and  $(\lambda, \Lambda_0) = a$ .

For every weight  $\lambda \in P(\Lambda)$  with  $\langle \lambda, K \rangle = k$ ,  $\|\lambda\|^2 + 2k(\lambda, \Lambda_0) \leq \|\Lambda\|^2$ . For  $\lambda + N\delta$  we have  $a \mapsto a + N$ , so

$$\begin{aligned} \|\lambda + N\delta\|^2 + 2k(\lambda + N\delta, \Lambda_0) &= (\|\bar{\lambda}\|^2 + 2k(a + N)) + 2k(a + N) \\ &= \|\bar{\lambda}\|^2 + 4ka + 4kN. \end{aligned}$$

If one increases  $N$ , this quantity increases without bound. Therefore there is a largest  $N$  such that  $\lambda + N\delta$  satisfies the paraboloid inequality, i.e. is a weight. That largest  $N$  is characterized by  $\lambda + (N + 1)\delta$  failing the inequality, i.e.  $\lambda + N\delta$  is maximal.

**Proposition 8.20 (12.6).** *The map*

$$\lambda \longmapsto \bar{\lambda}$$

*defines a bijection from  $\max(\Lambda) \cap P_+$  onto*

$$kC^{\text{af}} \cap (\bar{\Lambda} + Q)$$

*where  $kC^{\text{af}}$  is the fundamental alcove of level  $k$  and  $Q$  is the root lattice. In particular, the set of dominant maximal weights of  $L(\Lambda)$  is finite.*

*Proof.* Every weight  $\lambda$  of an affine algebra splits as  $\lambda = \bar{\lambda} + k\Lambda_0 + s\delta$ , where  $\bar{\lambda} \in \bar{\mathfrak{h}}^*$  is the "finite part" (projection that kills  $K$  and  $\delta$ ),  $k = \langle \lambda, K \rangle$  is the level, and  $s \in \mathbb{R}$  is the  $\delta$ -coordinate.

Under the projection map, the image  $\subseteq$  because if  $\lambda \in \max(\Lambda) \cap P_+$  then  $\langle \lambda, \alpha_i^\vee \rangle \geq 0$  for  $i \geq 1$ , and the affine dominance against  $\alpha_0^\vee = K - \theta^\vee$  gives  $\langle \bar{\lambda}, \theta^\vee \rangle \leq k$ . Also  $\bar{\lambda} - \bar{\Lambda} \in Q$ .

The map is onto because given  $\bar{\mu} \in kC_+^{\text{af}} \cap (\bar{\Lambda} + Q)$ , pick any weight of  $L(\Lambda)$  whose finite part is  $\bar{\mu}$ , then go up along the  $\delta$ -string to its maximal element; that maximal element is dominant and projects to  $\bar{\mu}$ .  $\square$

**Remark 8.21 (The point of maximal dominant weights).** *Observe that adding a multiple of  $\delta$  does not change dominance.*

$$\langle \delta, \alpha_i^\vee \rangle = 0 \quad (i = 0, 1, \dots, \ell),$$

*because  $\delta$  vanishes on  $\mathfrak{h}'$  (the span of the simple coroots). Hence*

$$\langle \lambda + n\delta, \alpha_i^\vee \rangle = \langle \lambda, \alpha_i^\vee \rangle + n \underbrace{\langle \delta, \alpha_i^\vee \rangle}_0 = \langle \lambda, \alpha_i^\vee \rangle.$$

*So the inequalities defining dominance (against all simple coroots, including  $\alpha_0^\vee$ ) are unchanged. More generally, level and the finite part  $\bar{\lambda}$  are also unchanged by adding  $\delta$ -multiples. Without a "maximal" condition you'd get infinitely many dominant weights - once  $\lambda$  is dominant, so are all  $\lambda - n\delta$  below it.*

*So you restrict to maximal weights (the tops of the  $\delta$ -strings). Those are finite in number: they correspond to the lattice points*

$$kC_+^{\text{af}} \cap (\bar{\Lambda} + Q)$$

The dominant maximal weights of  $L(\Lambda)$

**Example 8.22 (Maximal dominant weights for  $\widehat{\mathfrak{sl}}_2$ ).** Recall in an earlier example, we introduced the coordinate  $x(\lambda) = (\lambda|\alpha) = \langle \lambda, \alpha^\vee \rangle$  on the fundamental alcove of  $\widehat{\mathfrak{sl}}_2$ . We can also write this coordinate as projecting to the fundamental weight  $\omega = \frac{\alpha}{2}$  of  $\mathfrak{sl}_2$  (which is characterized by  $\langle \omega, \alpha^\vee \rangle = 1$ ). Then  $Q = 2\mathbb{Z}\omega$ .

The level- $k$  alcove is the segment

$$kC_+^{af} = \{j\omega \mid 0 \leq j \leq k\}.$$

Fix  $\Lambda$  of level  $k$  with  $\bar{\Lambda} = r\omega$  (some integer  $r$ ). Then

$$kC_+^{af} \cap (\bar{\Lambda} + Q) = \{j\omega \mid 0 \leq j \leq k, j \equiv r \pmod{2}\}.$$

Those  $j$ 's (with the parity constraint) are in bijection with the dominant maximal weights of  $L(\Lambda)$ . Each gives a  $\delta$ -string  $\{\lambda_j - n\delta\}_{n \geq 0}$ ; the whole weight set is the disjoint union of these strings. For a fixed representative in the fundamental alcove, the coefficient of  $\delta$  in the expansion

$$\lambda = \bar{\Lambda} + k\Lambda_0 + s\delta$$

is controlled by the paraboloid inequality, which reduces to

$$s \leq \frac{\|\Lambda\|^2 - \|\bar{\Lambda}\|^2}{2k}$$

## 9 Vertex algebras

### 9.1 Fields

Let  $V$  be a vector space over  $\mathbb{C}$ . Denote by  $\text{End } V$  the algebra of linear operators on  $V$ . A formal power series

$$A(z) = \sum_{j \in \mathbb{Z}} A_j z^{-j} \in \text{End } V[[z^{\pm 1}]] \quad (60)$$

is called a **field** on  $V$  if for any  $v \in V$  we have  $A_j v = 0$  for large enough  $j$ . In other words,  $A(z)v$  is an element of  $V((z))$ , the space of formal Laurent series with coefficients in  $V$  (i.e. only finitely many terms with negative powers of  $z$ ). Fields on  $V$  form a vector space denoted by  $\mathcal{F}(V)$ .

Now suppose that  $V$  is  $\mathbb{Z}$ -graded,

$$V = \bigoplus_{n \in \mathbb{Z}} V_n.$$

Recall that in this case a linear operator  $\phi : V \rightarrow V$  is called **homogeneous of degree  $m$**  if  $\phi(V_n) \subseteq V_{n+m}$  for all  $n$ .

A (homogeneous) field of **conformal dimension**  $\Delta \in \mathbb{Z}$  is by definition a field (60) on  $V$ , where each  $A_j$  is a homogeneous linear operator on  $V$  of degree  $-j + \Delta$ . Suppose that  $V_k = 0$  for all  $k$  less than a fixed integer  $K$ . Then if  $v \in V_n$  is a vector of degree  $n$ , we have

$$\deg(A_j v) = n - j + \Delta,$$

so  $A_j v = 0$  for all  $j > n + K + \Delta$ . It follows that for an arbitrary  $v \in V$  there is an integer  $N$  such that  $A_j v = 0$  for  $j > N$ . Thus

$$A(z) v = \sum_{j \in \mathbb{Z}} (A_j v) z^{-j} \in V((z)), \quad (61)$$

so  $A(z)$  automatically satisfies the condition of being a field.

If  $A(z)$  is a field of conformal dimension  $\Delta$ , then

$$\partial_z A(z) = \sum_{j \in \mathbb{Z}} (-j) A_j z^{-j-1}$$

is a field of conformal dimension  $\Delta + 1$ .

If we substitute for the formal variable  $z$  a non-zero complex number (also denoted by  $z$ ), then (61) yields an infinite sum of homogeneous components lying in the subspaces  $V_n$ ,  $n \in \mathbb{Z}$ . In other words, we obtain a vector in the direct product

$$\overline{V} = \prod_{n \in \mathbb{Z}} V_n.$$

Suppose that  $V_n = 0$  for  $n < K$ , i.e.

$$V = \bigoplus_{n=K}^{\infty} V_n.$$

Then the space  $\overline{V}$  is the completion of  $V$  in the topology with a basis of open neighborhoods of 0 of the form

$$V_{\geq k} = \bigoplus_{n=k}^{\infty} V_n, \quad k \geq K.$$

Thus, for any nonzero complex number  $z \in \mathbb{C}^\times$ ,  $A(z)$  can be considered as a linear operator from  $V$  to  $\overline{V}$ .

## 9.2 Composing Fields

We would like to formulate a reasonable condition of commutativity (which we will refer to as **locality**) of two fields  $A(z)$  and  $B(w)$  acting on  $V$ .

Observe that for any vector  $v \in V$  and any linear functional  $\varphi : V \rightarrow \mathbb{C}$ , the matrix element  $\langle \varphi, A(z)v \rangle$  of a field  $A(z)$  is a Laurent power series (the degrees in  $z$  are bounded from below).

Given another field  $B(w)$ , we consider the composition  $A(z)B(w)$  as an  $\text{End } V$ -valued formal power series in  $z, w$ . Given  $v \in V$  and  $\varphi \in V^*$  (where  $V^*$  denotes the space of all linear functionals on  $V$ ), consider the matrix element

$$\langle \varphi, A(z)B(w)v \rangle \in \mathbb{C}[[z^{\pm 1}, w^{\pm 1}]].$$

From the definition of a field, this formal power series actually belongs to  $\mathbb{C}((z))((w))$ . Indeed, the degrees of  $w$  in this series are bounded from below, and if we write  $B(w) = \sum_{j \in \mathbb{Z}} B_j w^{-j}$ , then the coefficient in front of  $w^{-j}$  equals  $\langle \varphi, A(z)B_j v \rangle$  and hence belongs to  $\mathbb{C}((z))$ . Similarly,  $\langle \varphi, B(w)A(z)v \rangle$  belongs to  $\mathbb{C}((w))((z))$ .

If we impose the condition that the fields  $A(z)$  and  $B(w)$  commute, i.e. that these two matrix elements are equal for all  $v, \varphi$ , then they both belong to the intersection

$$\mathbb{C}((z))((w)) \cap \mathbb{C}((w))((z)) = \mathbb{C}[[z, w]][z^{-1}, w^{-1}]$$

This condition is too strong, so we relax it. It is more reasonable to require that the two expressions be expansions of one and the same rational function of  $z$  and  $w$ , with possible poles only at  $z = 0$ ,  $w = 0$ , and  $z = w$ . This motivates the following definition.

**Definition 9.1.** *Two fields  $A(z)$  and  $B(w)$  acting on a vector space  $V$  are said to be **local** with respect to each other if for every  $v \in V$  and  $\varphi \in V^*$ , the matrix elements*

$$\langle \varphi, A(z)B(w)v \rangle \quad \text{and} \quad \langle \varphi, B(w)A(z)v \rangle$$

*are expansions of one and the same element*

$$f_{\varphi, v} \in \mathbb{C}[[z, w]][z^{-1}, w^{-1}, (z - w)^{-1}]$$

*in  $\mathbb{C}((z))((w))$  and  $\mathbb{C}((w))((z))$ , respectively, and the order of the pole of  $f_{\varphi, v}$  in  $(z - w)$  is uniformly bounded for all  $v, \varphi$ .*

The last condition may be reformulated as saying that there exists  $N \in \mathbb{Z}_+$  such that

$$(z - w)^N f_{v, \varphi} \in \mathbb{C}[[z, w]][z^{-1}, w^{-1}]$$

for all  $v, \varphi$ . But then the expansions of  $(z - w)^N f_{v, \varphi}$  in  $\mathbb{C}((z))((w))$  and  $\mathbb{C}((w))((z))$  are equal to each other. Therefore if  $A(z)$  and  $B(w)$  are local with respect to each other, then

$$(z - w)^N A(z)B(w) = (z - w)^N B(w)A(z),$$

or equivalently,

$$(z - w)^N [A(z), B(w)] = 0,$$

The following proposition shows that the converse is also true.

**Proposition 9.2.** *Two fields  $A(z), B(w)$  are local if and only if there exists  $N \in \mathbb{Z}_+$  such that*

$$(z - w)^N[A(z), B(w)] = 0 \quad (62)$$

*as a formal power series in  $\text{End } V[[z^{\pm 1}, w^{\pm 1}]]$ .*

We also give an analytic interpretation of locality.

**Proposition 9.3.** *Two fields  $A(z), B(w)$  are local with respect to each other if and only if:*

1. *For any  $z, w \in \mathbb{C}^\times$  with  $|z| > |w|$ , the composition  $A(z)B(w)$  exists and can be analytically continued to an operator-valued meromorphic function  $R(A(z)B(w))$  on  $\mathbb{C}^2$ , with singularities at  $\{z = 0\}$ ,  $\{w = 0\}$ , and  $\{z = w\}$ , such that the order of the pole at  $z = w$  is uniformly bounded.*
2. *For  $|w| > |z|$  the composition  $B(w)A(z)$  exists and can be analytically continued to an operator-valued meromorphic function  $R(B(w)A(z))$  on  $\mathbb{C}^2$ , with singularities at  $\{z = 0\}$ ,  $\{w = 0\}$ , and  $\{z = w\}$ , such that the order of the pole at  $z = w$  is uniformly bounded.*
3.  $R(A(z)B(w)) = R(B(w)A(z))$ .

**Remark 9.4 (Delta function identities).** *Consider the expansion maps from rational functions into two different completed rings:*

$$\iota_{z,w} : \mathbb{C}(z, w)[(z - w)^{-1}] \longrightarrow \mathbb{C}((z))((w)), \quad \iota_{w,z} : \mathbb{C}(z, w)[(z - w)^{-1}] \longrightarrow \mathbb{C}((w))((z)).$$

where  $\mathbb{C}((z))((w))$  means *Laurent series in  $z$  whose coefficients are Laurent series in  $w$  with finitely many negative powers of  $w$  in each coefficient, etc.* A function like  $\frac{1}{z-w}$  has two different expansions in these two rings:

$$\begin{aligned} \iota_{z,w} \frac{1}{z-w} &= \frac{1}{z} \sum_{n \geq 0} \left(\frac{w}{z}\right)^n = \sum_{n \geq 0} z^{-n-1} w^n \\ \iota_{w,z} \frac{1}{z-w} &= \frac{1}{w} \sum_{n \geq 0} \left(\frac{z}{w}\right)^n = \sum_{n \geq 0} w^{-n-1} z^n \end{aligned}$$

Define the formal delta

$$\delta(z - w) = \sum_{n \in \mathbb{Z}} z^{-n-1} w^n.$$

Then the fundamental identities is

$$\iota_{z,w} \frac{1}{z-w} - \iota_{w,z} \frac{1}{z-w} = \delta(z - w)$$

The point is that in the space of formal distributions  $\mathbb{C}[[z, w]][z^{-1}, w^{-1}]$ ,

$$\delta(z - w) = \sum_{n \in \mathbb{Z}} z^{-n-1} w^n \quad \text{satisfies} \quad (z - w) \delta(z - w) = 0.$$

So  $\delta$  is  $(z - w)$ -torsion. Derivatives give higher-order torsion:

$$(z - w)^{k+1} \partial_w^k \delta(z - w) = 0 \quad (k \geq 0),$$

and, generically, no lower power will kill it. So  $\partial_w^k \delta$  is  $(z - w)^{k+1}$ -torsion. The two expansion maps differ by exactly this torsion:

$$\iota_{z,w} \frac{1}{(z - w)^{k+1}} - \iota_{w,z} \frac{1}{(z - w)^{k+1}} = \frac{1}{k!} \partial_w^k \delta(z - w).$$

They are the same rational function, meaning the mismatch between completions is a delta-torsion piece on  $z = w$ . Locality is "the commutator is  $(z - w)$ -power torsion."

From an OPE

$$A(z)B(w) = \sum_{k=0}^{N-1} \frac{C_k(w)}{(z - w)^{k+1}} + :A(z)B(w):,$$

you get

$$[A(z), B(w)] = \sum_{k=0}^{N-1} \frac{1}{k!} C_k(w) \partial_w^k \delta(z - w),$$

hence  $(z - w)^N [A(z), B(w)] = 0$ . The normal-ordered term is regular (torsion-free) at  $z = w$ ; all the singular/"support on the diagonal" part is encoded by  $\delta$  and its derivatives.

**Definition 9.5 (Vertex Algebras).** A *vertex algebra* is a collection of data:

- (*space of states*) a vector space  $V$ ;
- (*vacuum vector*) a vector  $|0\rangle \in V$ ;
- (*translation operator*) a linear operator  $T : V \rightarrow V$ ;
- (*vertex operators*) a linear operation

$$Y(\cdot, z) : V \rightarrow \text{End } V[[z^{\pm 1}]],$$

taking each  $A \in V$  to a field

$$Y(A, z) = \sum_{n \in \mathbb{Z}} A_{(n)} z^{-n-1}$$

acting on  $V$ .



These data are subject to the following axioms:

- **(vacuum axiom)**

$$Y(|0\rangle, z) = \text{id}_V.$$

Furthermore, for any  $A \in V$  we have

$$Y(A, z)|0\rangle \in V[[z]],$$

so that  $Y(A, z)|0\rangle$  has a well-defined value at  $z = 0$ , and

$$Y(A, z)|0\rangle \Big|_{z=0} = A.$$

In other words,  $A_{(n)}|0\rangle = 0$  for  $n \geq 0$ , and  $A_{(-1)}|0\rangle = A$ .

- **(translation axiom)** For any  $A \in V$ ,

$$[T, Y(A, z)] = \partial_z Y(A, z), \quad T|0\rangle = 0.$$

- **(locality axiom)** All fields  $Y(A, z)$  are local with respect to each other.

A vertex algebra is called  $\mathbb{Z}$ -graded if  $V$  is a  $\mathbb{Z}$ -graded vector space, the vacuum  $|0\rangle$  has degree 0, the operator  $T$  has degree 1, and for  $A \in V_m$  the field  $Y(A, z)$  has conformal dimension  $m$  (i.e.  $\deg A_{(n)} = -n + m - 1$ ).

We begin with some basic remarks about the axioms.

**Remark 9.6** ( $T$  is determined by  $Y$ ). Recall the translation axiom:

$$[T, Y(A, z)] = \partial_z Y(A, z), \quad T|0\rangle = 0$$

We know  $Y(A, z)|0\rangle \in V[[z]]$ :

$$Y(A, z)|0\rangle = \sum_{n \geq 0} A_{(-n-1)}|0\rangle z^n$$

The coefficient of  $z^0$  is  $A$ , the coefficient of  $z^1$  is  $A_{(-2)}|0\rangle$ , and so on. Apply  $[T, Y(A, z)] = \partial_z Y(A, z)$  to  $|0\rangle$  and compute both sides:

$$[T, Y(A, z)]|0\rangle = T(Y(A, z)|0\rangle) - Y(A, z)T|0\rangle = T(Y(A, z)|0\rangle)$$

since  $T|0\rangle = 0$ . The lefthand side gives  $\partial_z Y(A, z)|0\rangle = \sum_{n \geq 0} (n+1)A_{(-n-2)}|0\rangle z^n$ . Comparing coefficients of  $z^0$ :

$$TA = TA_{(-1)}|0\rangle = A_{(-2)}|0\rangle$$

**Remark 9.7** (*Y is injective*). This comes directly from the vacuum axiom:

$$Y(A, z)|0\rangle \in V[[z]], \quad Y(A, z)|0\rangle|_{z=0} = A$$

Expanding we get that  $Y(A, z)|0\rangle = \sum_{n \geq 0} A_{(-n-1)}|0\rangle z^n$ . Then the coefficient of  $z^0$  is exactly  $A_{(-1)}|0\rangle$ . The vacuum axiom says this equals  $A$ :  $A = A_{(-1)}|0\rangle$ .

Hence, if  $A_{(-1)} = 0$ , then  $A_{(-1)}|0\rangle = 0$ , so  $A = 0$ . That shows the map

$$V \rightarrow \text{End}(V), \quad A \mapsto A_{(-1)}$$

is injective and this implies the vertex operator map  $Y$  is injective.

**Definition 9.8** (Morphisms, Subalgebras, Ideals, Quotients). A **vertex algebra homomorphism**  $\rho$  between vertex algebras

$$(V, |0\rangle, T, Y) \longrightarrow (V', |0'\rangle, T', Y')$$

is a linear map  $V \rightarrow V'$  mapping  $|0\rangle$  to  $|0'\rangle$ , intertwining the translation operators, and satisfying

$$\rho(Y(A, z)B) = Y(\rho(A), z)\rho(B).$$

A **vertex subalgebra**  $V' \subset V$  is a  $T$ -invariant subspace containing the vacuum vector, and satisfying  $Y(A, z)B \in V'((z))$  for all  $A, B \in V'$  (with the induced vertex algebra structure).

A **vertex algebra ideal**  $I \subset V$  is a  $T$ -invariant subspace satisfying  $Y(A, z)B \in I((z))$  for all  $A \in I$  and  $B \in V$ . The skew-symmetry property proved in Proposition 3.2.5 below implies that then  $Y(B, z)A \in I((z))$  as well (with  $A, B$  as before), so that  $I$  is automatically a “two-sided” ideal. It follows that for any proper ideal  $I$ , the quotient  $V/I$  inherits a natural quotient vertex algebra structure.

**Definition 9.9** (Tensor Product of Vertex Algebras). For two vertex algebras  $(V_1, |0\rangle_1, T_1, Y_1)$  and  $(V_2, |0\rangle_2, T_2, Y_2)$ , the data

$$(V_1 \otimes_{\mathbb{C}} V_2, |0\rangle_1 \otimes |0\rangle_2, T_1 \otimes 1 + 1 \otimes T_2, Y)$$

where

$$Y(A_1 \otimes A_2, z) = Y_1(A_1, z) \otimes Y_2(A_2, z),$$

defines a vertex algebra called the **tensor product** of  $V_1$  and  $V_2$ .

**Example 9.10** (Commutative Vertex Algebras). A vertex algebra is called **commutative** if all vertex operators  $Y(A, z)$ ,  $A \in V$ , commute with each other (i.e. we have  $N = 0$  in formula (1.2.3)).

Suppose we are given a commutative vertex algebra  $V$ . Then for any  $A, B \in V$  we have

$$Y(A, z)B = Y(A, z)Y(B, w)|0\rangle|_{w=0} = Y(B, w)Y(A, z)|0\rangle|_{w=0}.$$

But by the vacuum axiom, the last expression has no negative powers of  $z$ . Therefore  $Y(A, z)B \in V[[z]]$  for all  $A, B \in V$ , so  $Y(A, z) \in \text{End } V[[z]]$  for all  $A \in V$ .

Conversely, suppose that we are given a vertex algebra  $V$  in which  $Y(A, z) \in \text{End } V[[z]]$  for all  $A \in V$ . Observe that if the equality

$$(z - w)^N f_1(z, w) = (z - w)^N f_2(z, w)$$

holds for  $f_1, f_2 \in \mathbb{C}[[z, w]]$  and  $N \in \mathbb{Z}_+$ , then necessarily  $f_1(z, w) = f_2(z, w)$ . Therefore we obtain that  $[Y(A, z), Y(B, w)] = 0$  for all  $A, B \in V$ , so  $V$  is commutative.

Thus, a commutative vertex algebra may alternatively be defined as one in which all vertex operators  $Y(A, z)$  are regular at  $z = 0$ .

Denote by  $Y_A$  the endomorphism of a commutative vertex algebra  $V$  which is the constant term of  $Y(A, z)$ ,  $A \in V$ , and define a bilinear operation  $\circ$  on  $V$  by setting  $A \circ B = Y_A B$ . By construction,  $Y_A Y_B = Y_B Y_A$ . As explained in §1.3.3, this implies both commutativity and associativity of  $\circ$ . Hence we obtain a commutative and associative product on  $V$ . Furthermore, the vacuum vector  $|0\rangle$  is a unit, and the operator  $T$  is a derivation with respect to this product. Thus, we have the structure of a commutative algebra with a derivation on  $V$ .

Conversely, let  $V$  be a commutative algebra with a unit and finite-dimensional homogeneous components. Suppose that  $V$  is equipped with a derivation  $T$  of degree 1. Then  $V$  carries a canonical vertex algebra structure. Namely, we take the unit of  $V$  as the vacuum vector  $|0\rangle$ , and define  $Y(A, z)$  to be the operator of multiplication by

$$\sum_{n \geq 0} \frac{z^n}{n!} (T^n A) = e^{zT} A.$$

It is straightforward to check that all axioms of a commutative vertex algebra are satisfied.

Thus, we see that the notion of commutative vertex algebra is essentially equivalent to that of a commutative algebra with a unit and a derivation.

## 10 Vertex algebras from Lie algebras

Consider the vector space  $\mathbb{C}((t))$  of formal Laurent series in one variable as a commutative Lie algebra.

**Definition 10.1.** We define the **Heisenberg Lie algebra**  $\mathcal{H}$  as the central extension

$$0 \longrightarrow \mathbb{C}1 \longrightarrow \mathcal{H} \longrightarrow \mathbb{C}((t)) \longrightarrow 0 \quad (63)$$

with the cocycle

$$c(f, g) = -\text{Res}_{t=0} f dg. \quad (64)$$

There is another version of the Heisenberg Lie algebra: the one-dimensional central extension  $\mathcal{H}'$  of the commutative Lie algebra of Laurent polynomials  $\mathbb{C}[t, t^{-1}]$  defined by the two-cocycle (64). This Lie algebra has an obvious basis  $b_n = t^n$ ,  $n \in \mathbb{Z}$ , and the central element 1. The Lie algebra  $\mathcal{H}$  does not possess such a simple basis. However,  $\mathcal{H}$  is the completion of  $\mathcal{H}'$  with respect to the topology in which the basis of open neighborhoods of 0 is formed by the subspaces  $t^N \mathbb{C}[t]$ ,  $N \in \mathbb{Z}$ . Hence it makes sense to say that  $\mathcal{H}$  is **topologically** generated by  $b_n = t^n$ ,  $n \in \mathbb{Z}$ , and 1. Note that the Lie bracket is continuous with respect to the above topology, and so  $\mathcal{H}$  is a complete topological Lie algebra.

By a representation of a topological Lie algebra like  $\mathcal{H}$  we will understand (unless noted otherwise) a continuous representation in a vector space  $V$  equipped with the **discrete** topology. In the case of  $\mathcal{H}$ , this is equivalent to the requirement that for any  $v \in V$ , we have  $t^N \mathbb{C}[[t]] \cdot v = 0$  for some  $N \in \mathbb{Z}$ . Giving such a representation of  $\mathcal{H}$  is the same as giving a representation of  $\mathcal{H}'$  such that for any  $v \in V$ , we have  $t^N \mathbb{C}[t] \cdot v = 0$  for some  $N \in \mathbb{Z}$ , because then the action of  $\mathcal{H}'$  on  $V$  may be extended to  $\mathcal{H}$  by continuity.

*The advantage of  $\mathcal{H}$  over  $\mathcal{H}'$  is that  $\mathcal{H}$  is preserved under changes of the coordinate  $t$ , while  $\mathcal{H}'$  is not. This property will become important later on when we develop a coordinate-independent approach to vertex algebras. However, when studying concrete representations, it is more convenient to view them as representations of the Lie algebra  $\mathcal{H}'$ .*

From formula (64) we find that

$$\begin{aligned} c(t^n, t^m) &= -\operatorname{Res}_{t=0} t^n dt^m \\ &= -m \operatorname{Res}_{t=0} t^{n+m-1} dt \\ &= -m \delta_{n, -m} = n \delta_{n, -m}. \end{aligned}$$

The definition (64) of the cocycle using residues of one-forms rather than coefficients of power series is clearly independent of the choice of local coordinate  $t$ . Thus we may define a Heisenberg Lie algebra canonically as a central extension of the space of functions on a punctured disc  $D^\times$  which is not endowed with a specific choice of formal coordinate  $t$ .

**Why is the coordinate free definition important? Has it been useful for anything related to automorphic forms and representation theory?**

The universal enveloping algebra  $U(\mathcal{H}')$  of  $\mathcal{H}'$  is an associative algebra with generators  $b_n$ ,  $n \in \mathbb{Z}$ , and 1, and relations (for  $n, m \in \mathbb{Z}$ )

$$b_n b_m - b_m b_n = n \delta_{n, -m} 1, \quad b_n 1 - 1 b_n = 0.$$

Introduce a topology on  $U(\mathcal{H}')$ , in which the basis of open neighborhoods of 0 is formed by the left ideals of the subspaces  $t^N \mathbb{C}[t] \subset \mathcal{H}' \subset U(\mathcal{H}')$ ,  $N \in \mathbb{Z}$ . The completion of  $U(\mathcal{H}')$  with respect to this topology will be denoted by  $\tilde{U}(\mathcal{H})$ . In more concrete terms, elements of  $\tilde{U}(\mathcal{H})$  may be

described as (possibly infinite) series of the form

$$R_0 + \sum_{n \geq 0} P_n b_n,$$

where  $R_0$  and the  $P_n$ 's are arbitrary (finite) elements of  $U(\mathcal{H}')$ . It is easy to check that  $\tilde{U}(\mathcal{H})$  is an associative algebra with a unit.

Note that  $\tilde{U}(\mathcal{H})$  contains as a subalgebra (but is not equal to) the universal enveloping algebra  $U(\mathcal{H})$ . The latter consists of finite linear combinations of products of elements of  $\mathcal{H}$ . Moreover,  $\tilde{U}(\mathcal{H})$  is the completion of  $U(\mathcal{H})$  with respect to the topology in which the basis of open neighborhoods of 0 is formed by the left ideals of the subspaces  $t^N \mathbb{C}[[t]] \subset \mathcal{H} \subset U(\mathcal{H})$ ,  $N \in \mathbb{Z}$ .

It is clear that any representation of  $\mathcal{H}$  of the type discussed above is automatically a  $\tilde{U}(\mathcal{H})$ -module.

**Definition 10.2 (Weyl Algebra).** *The **Weyl algebra**  $\tilde{\mathcal{H}}$  is defined as the quotient of  $\tilde{U}(\mathcal{H})$  by the two-sided ideal generated by  $(1-1)$ . Here  $1$  stands for the central element of  $\mathcal{H} \subset \tilde{U}(\mathcal{H})$ , and  $1$  is the unit element of  $\tilde{U}(\mathcal{H})$ .*

## 10.1 The Fock representation

We wish to construct a representation of the Weyl algebra  $\tilde{\mathcal{H}}$  which is “as small as possible”. Unlike commutative algebras,  $\tilde{\mathcal{H}}$  does not possess one-dimensional representations. Indeed, in  $\tilde{\mathcal{H}}$  the commutator of  $b_n$ ,  $n \neq 0$ , and  $b_{-n}$  gives a non-zero multiple of the unit element,

$$b_n b_{-n} - b_{-n} b_n = n, \tag{65}$$

so it is impossible to have a representation on which both  $b_n$  (for  $n \neq 0$ ) and  $b_{-n}$  act by 0.

Consider the subalgebra  $\tilde{\mathcal{H}}_+ \subset \tilde{\mathcal{H}}$  generated by  $b_0, b_1, b_2, \dots$ . It is a commutative subalgebra, and hence has a trivial one-dimensional representation (in fact,  $\tilde{\mathcal{H}}_+$  is a maximal commutative subalgebra of  $\tilde{\mathcal{H}}$ ). We may then define a representation of  $\tilde{\mathcal{H}}$  as the induced representation from this representation of  $\tilde{\mathcal{H}}_+$ :

$$\pi \stackrel{\text{def}}{=} \text{Ind}_{\tilde{\mathcal{H}}_+}^{\tilde{\mathcal{H}}} \mathbb{C} = \tilde{\mathcal{H}} \otimes_{\tilde{\mathcal{H}}_+} \mathbb{C},$$

(note that  $\tilde{\mathcal{H}}_+$  acts on  $\tilde{\mathcal{H}}$  from the right). It is called the **Fock representation** of  $\tilde{\mathcal{H}}$  (or of  $\mathcal{H}$ ).

Let  $\tilde{\mathcal{H}}_-$  be the commutative subalgebra of  $\mathcal{H}$  generated by  $b_n$ ,  $n < 0$ . It follows from the Poincaré–Birkhoff–Witt theorem that

$$\tilde{\mathcal{H}} \simeq \tilde{\mathcal{H}}_+ \otimes \tilde{\mathcal{H}}_-.$$

Therefore

$$\pi \simeq \tilde{\mathcal{H}}_- = \mathbb{C}[b_{-1}, b_{-2}, \dots].$$

Under this isomorphism, the generators  $b_n$ ,  $n < 0$ , simply act on  $\pi$  by multiplication.

To find the action of the  $b_n$ ,  $n \geq 0$ , we consider a monomial  $M = b_{-i_1} \cdots b_{-i_r} |0\rangle$  with all  $i_j > 0$ . Move  $b_n$  to the right, one factor at a time, using

$$b_n b_{-m} = b_{-m} b_n + [b_n, b_{-m}] = b_{-m} b_n + n \delta_{n,m}$$

Each time  $b_n$  passes a  $b_{-m}$  with  $m \neq n$ , nothing happens; when it passes a  $b_{-n}$ , you pick up a scalar  $n$ . After commuting past all  $r$  factors you get

$$b_n M = \sum_{\substack{1 \leq j \leq r \\ i_j = n}} (b_{-i_1} \cdots \widehat{b_{-i_j}} \cdots b_{-i_r}) |0\rangle \cdot n + b_{-i_1} \cdots b_{-i_r} b_n |0\rangle$$

The last term vanishes because  $n \geq 0$  implies  $b_n |0\rangle = 0$ . So

$$b_n (b_{-i_1} \cdots b_{-i_r} |0\rangle) = n \sum_{j: i_j = n} b_{-i_1} \cdots \widehat{b_{-i_j}} \cdots b_{-i_r} |0\rangle$$

Under the polynomial identification, that is exactly  $b_n = n \frac{\partial}{\partial b_{-n}}$  on  $\mathbb{C}[b_{-1}, b_{-2}, \dots]$ .

The operators  $b_n$  with  $n < 0$  are known in this context as **creation operators**, since they “create the state  $b_n$  from the vacuum 1,” while the operators  $b_n$  with  $n \geq 0$  are the **annihilation operators**, repeated application of which will “kill” any vector in  $\pi$ .

We will endow  $\pi$  with the structure of a  $\mathbb{Z}^+$ -graded vertex algebra. Define

$$|0\rangle \stackrel{\text{def}}{=} 1 \in \pi$$

and the translation operator  $T$  by the rules

$$\begin{aligned} T|0\rangle &\stackrel{\text{def}}{=} 0 \\ [T, b_n] &\stackrel{\text{def}}{=} -n b_{n-1}, \quad n \in \mathbb{Z}. \end{aligned}$$

The  $\mathbb{Z}^+$ -grading is given by assigning  $b_n$  degree  $-n$  and extending multiplicatively.

Now we consider the vertex operator map  $Y(\cdot, z) : \pi \rightarrow \text{End } \pi[[z^{\pm 1}]]$ . To the vacuum vector  $|0\rangle = 1$ , we are required to assign  $Y(|0\rangle, z) = \text{Id}$ , so there is no ambiguity here. The key definition is that of the field  $Y(b_{-1}, z)$ , which as we will see generates  $\pi$  in an appropriate sense. Let us denote  $Y(b_{-1}, z)$  by  $b(z)$  for brevity. Since  $\deg b_{-1} = 1$ , the field  $b(z)$  needs to have conformal dimension one. We now set

$$b(z) = \sum_{n \in \mathbb{Z}} b_n z^{-n-1},$$

where  $b_n$  is considered as an endomorphism of  $\pi$ . Since  $\deg b_n = -n$ ,  $b(z)$  is indeed a field of conformal dimension 1. Note that  $b(z)$  is nothing but a generating function for the generators  $b_n$  of the Heisenberg Lie algebra  $\mathcal{H}$ .

Next, we proceed to define  $Y(b_{-2}, z)$  as the field

$$\partial_z b(z) = \sum_{n \in \mathbb{Z}} (-n - 1) b_n z^{-n-2}.$$

It has conformal dimension 2, as required, since  $\deg b_{-2} = 2$ . This formula may be motivated by the fact that  $b_{-2} \sim t^{-2}$  equals  $-\partial_t \cdot t^{-1}$ . Following the same route, we obtain by induction

$$Y(b_{-k}, z) = \frac{1}{(k-1)!} \partial_z^{k-1} b(z).$$

Besides  $b_{-2}$ , the other generator in degree 2 is  $b_{-1}^2$ , to which one is tempted to assign

$$b(z)^2 = \sum_{n \in \mathbb{Z}} \left( \sum_{k+l=n} b_k b_l \right) z^{-n-2}.$$

Let us check the behavior of this expression. Consider, for instance, the sum

$$\sum_{k+l=10} b_k b_l \tag{66}$$

appearing as the  $z^{-12}$  coefficient in  $b(z)^2$ . Given any vector  $x \in \pi$ , we have  $b_n x = 0$  for all but finitely many annihilation operators  $b_n$  ( $n \geq 0$ ) (since these are differentiation operators and  $x$  is a polynomial). Thus in (66), except for finitely many terms, either  $b_k \cdot x = 0$  or  $b_l \cdot x = 0$ .

The Heisenberg relations (2.1.4) imply that  $b_k$  and  $b_l$  commute whenever  $k+l \neq 0$ ; hence we may rearrange the terms in our sum so that an annihilation operator always appears to the right of a creation operator. It follows that the operator (66), when applied to any  $x \in \pi$ , yields a finite sum. Thus, the  $z^{-12}$  coefficient in  $b(z)^2$  is a well-defined operator on  $\pi$ .

The same argument applies to the  $z^i$  coefficient of  $b(z)^2$  for any  $i \neq -2$ . But what about the coefficient

$$\sum_{k+l=0} b_k b_l ? \tag{67}$$

In front of  $z^{-2}$ ? In this case  $b_k$  and  $b_l$  do not commute, so the above argument does not apply. We rewrite (2.2.2) as

$$\sum_{k \in \mathbb{Z}} b_k b_{-k} = \sum_{k < 0} (-k) b_k \frac{\partial}{\partial b_k} + \sum_{k > 0} k \frac{\partial}{\partial b_{-k}} b_{-k}$$

(recall that  $b_0$  acts by 0 on  $\pi$ ). The first sum again becomes a finite sum when applied to any vector in  $\pi$ . But each term in the second sum first multiplies by  $b_{-k}$  and then differentiates with respect to the same  $b_{-k}$ , producing an infinite sum  $\sum_{k > 0} k$  when applied to the vector  $1 \in \pi$  (or a similar

infinite sum for any other vector of  $\pi$ ). This sum obviously does not converge. Thus, the  $z^{-2}$  coefficient of  $b(z)^2$  is not well-defined, and hence  $b(z)^2$  does not make sense as a field.

There is a standard way to remove the infinities in products of fields. Namely, we rearrange the factors in each “problematic” term in such a way that the annihilation operators always act before the creation operators. Then the infinite sums arising from repeatedly creating and annihilating the same state will not occur. This procedure is known as **normal ordering**.

In the case at hand, we define the normally ordered product of  $b(z)$  with itself, denoted by  $:b(z)b(z):$ , as the following formal power series in  $z$ :

$$:b(z)b(z): = \sum_{n \in \mathbb{Z}} (:b_k b_l:) z^{-n-2},$$

where we set

$$:b_k b_l: \stackrel{\text{def}}{=} \begin{cases} b_l b_k, & \text{if } l = -k, k \geq 0, \\ b_k b_l, & \text{otherwise.} \end{cases}$$

Thus, taking the normally ordered product simply amounts to the instruction to replace everywhere the terms  $b_k b_{-k}$  for  $k > 0$  by  $b_{-k} b_k$ .

This procedure does not change the coefficients of  $b(z)^2$  in front of  $z^n$ ,  $n \neq -2$ , but the coefficient in front of  $z^{-2}$  now becomes

$$\sum_{k \in \mathbb{Z}} :b_k b_{-k}: = 2 \sum_{k < 0} (-k) b_k \frac{\partial}{\partial b_k},$$

which is a well-defined linear operator on  $\pi$ .

In general, one must take the normally ordered product of fields.

**Definition 10.3 (Normally Ordered Product).** *The **normally ordered product** of the fields*

$$A(z) = \sum_{n \in \mathbb{Z}} A_{(n)} z^{-n-1}, \quad B(w) = \sum_{m \in \mathbb{Z}} B_{(m)} w^{-m-1}$$

*is defined as the formal power series*

$$\begin{aligned} :A(z)B(w): &:= \sum_{n \in \mathbb{Z}} \left( \sum_{m < 0} A_{(m)} B_{(n)} z^{-m-1} + \sum_{m \geq 0} B_{(n)} A_{(m)} z^{-m-1} \right) w^{-n-1} \\ &= A(z)_+ B(w) + B(w) A(z)_-, \end{aligned}$$

*where for a formal power series  $f(z) = \sum_{n \in \mathbb{Z}} f_n z^n$ , we write*

$$f(z)_+ = \sum_{n \geq 0} f_n z^n, \quad f(z)_- = \sum_{n < 0} f_n z^n.$$



The proof of the following lemma is left to the reader.

**Lemma 10.4 (2.2.3).**

1. For any  $v \in V$  and  $\varphi \in V^*$ , the matrix element

$$\langle \varphi, : A(z)B(w) : v \rangle$$

belongs to  $\mathbb{C}[[z, w]][z^{-1}, w^{-1}]$ . For any  $z, w \in \mathbb{C}^\times$ ,  $: A(z)B(w) :$  is a well-defined linear operator  $V \rightarrow \overline{V}$ .

2. The specialization of  $: A(z)B(w) :$  at  $w = z$  is a well-defined field, which is called the **normally ordered product** of  $A(z)$  and  $B(z)$ . If  $A(z)$  and  $B(z)$  are homogeneous of conformal dimensions  $\Delta_A$  and  $\Delta_B$ , respectively, then  $: A(z)B(z) :$  is also homogeneous of conformal dimension  $\Delta_A + \Delta_B$ .

3. We have the formula

$$: A(w)B(w) := \text{Res}_{z=0} (\delta(z-w)_- A(z)B(w) + \delta(z-w)_+ B(w)A(z))$$

where

$$\begin{aligned} \delta(z-w)_+ &\stackrel{\text{def}}{=} z^{-1} \sum_{n \geq 0} (w/z)^n = \frac{1}{z-w}, \quad |z| > |w| \\ \delta(z-w)_- &\stackrel{\text{def}}{=} -z^{-1} \sum_{n < 0} (w/z)^n = -\frac{1}{z-w}, \quad |z| < |w|. \end{aligned}$$

**Remark 10.5.** In general, the operation of normally ordered product is neither commutative nor associative. We follow the convention that the normal ordering is read from right to left, so that by definition

$$: A(z)B(z)C(z) : := : A(z)(: B(z)C(z) :) :$$

In general we define

$$Y(b_{j_1} b_{j_2} \cdots b_{j_k}, z) = \frac{1}{(-j_1 - 1)! \cdots (-j_k - 1)!} : \partial_z^{-j_1-1} b(z) \cdots \partial_z^{-j_k-1} b(z) : . \quad (68)$$

It is not clear why we should define the vertex operation  $Y(\cdot, z)$  in this way. It turns out that the vertex operation  $Y(\cdot, z)$  is completely determined by the assignment  $Y(b_{-1}, z) = b(z)$ . In particular there exists (at most) one vertex algebra structure on  $\pi$  such that  $Y(b_{-1}, z) = b(z)$ , namely the one we defined. The reason is that the Fourier coefficients of the field  $b(z)$  generate  $\pi$  from the vacuum vector 1. Thus, the operation of normal ordering is not an arbitrary convention, but a natural consequence of the structure of a vertex algebra.

## 10.2 Checking locality

General tools will let us check locality for our “generating” field  $b(z)$  and then derive locality in an inductive fashion from some simple properties of generating fields.

We must check that the fields  $b(z)$  are mutually local. We start out by writing

$$b(z)b(w) = \sum_{n,m} b_n b_m z^{-n-1} w^{-m-1} = \underbrace{\sum_{N \in \mathbb{Z} \setminus \{0\}} \sum_{n+m=N} b_n b_m z^{-n-1} w^{-m-1}}_{\Sigma_{\neq 0}} + \underbrace{\sum_{n \in \mathbb{Z}} b_{-n} b_n z^{-n-1} w^{-n-1}}_{\Sigma_0}.$$

As we observed in calculating the ill-fated expression  $b(z)^2$ , the terms of  $\Sigma_{\neq 0}$  are products  $b_n b_m$  in which  $b_n, b_m$  commute with each other. Thus we may switch the order of the factors with impunity to a normally ordered form, and hence  $\Sigma_{\neq 0} =: \Sigma_{\neq 0} :.$  Again the difficulty lies with  $\Sigma_0$ .

Let us write  $\Sigma_0 = -\Sigma_{0-} + (\Sigma_{0+} - \Sigma_{0-})$ . Normal ordering does not change the positive part of the sum, which already has annihilation operators on the right:

$$\sum_{n \geq 0} b_{-n} b_n z^{-n-1} w^{-n-1} = \sum_{n \geq 0} : b_{-n} b_n : z^{-n-1} w^{-n-1}.$$

In the negative part, however, applying normal ordering means switching the order of the factors:

$$\sum_{n < 0} b_{-n} b_n z^{-n-1} w^{-n-1} = \sum_{n < 0} b_n b_{-n} z^{-n-1} w^{-n-1}.$$

Hence

$$\Sigma_0 =: \Sigma_0 : + \sum_{n < 0} [b_{-n}, b_n] z^{-n-1} w^{-n-1} =: \Sigma_0 : + \sum_{m > 0} m z^{-m-1} w^{m-1}.$$

Summarizing, we obtain

$$b(z)b(w) =: b(z)b(w) : + \sum_{m > 0} m z^{-m-1} w^{m-1}.$$

Now observe that the series

$$\sum_{m > 0} m z^{-m-1} w^{m-1} \tag{69}$$

is the expansion in  $\mathbb{C}((z))((w))$  of the rational function  $1/(z-w)^2$ . Hence we write

$$b(z)b(w) = \frac{1}{(z-w)^2} + : b(z)b(w) :, \tag{70}$$

where  $1/(z-w)^2$  stands for its expansion in  $\mathbb{C}((z))((w))$ , i.e. in positive powers of  $w/z$ .

To compute the product in the opposite order,  $b(w)b(z)$ , we simply switch  $z$  and  $w$  in the above formulas. We then obtain the series

$$\sum_{m>0} m w^{-m-1} z^{m-1}, \quad (71)$$

which is the expansion of the same rational function  $1/(z-w)^2$ , but now in  $\mathbb{C}((w))((z))$ , i.e. in positive powers of  $z/w$ . Hence we write

$$b(w)b(z) = \frac{1}{(z-w)^2} + :b(w)b(z):, \quad (72)$$

where  $1/(z-w)^2$  now stands for its expansion in  $\mathbb{C}((w))((z))$ .

In addition, we have

$$:b(z)b(w): = :b(w)b(z):$$

(as elements of  $\text{End } V[[z^{\pm 1}, w^{\pm 1}]]$ ). Indeed, almost all generators  $b_k, b_l$  commute, and in the only remaining case  $k = -l$ , the factor with positive number appears to the right after ordering, regardless of how they started out.

By Lemma 2.2.3 (1), for any  $v \in V$ ,  $\varphi \in V^*$ , the matrix elements

$$\langle \varphi, b(z)b(w)v \rangle = \langle \varphi, b(w)b(z)v \rangle$$

are well-defined elements of  $\mathbb{C}[[z, w]][z^{-1}, w^{-1}]$ .

From the computation right above, you have

$$b(z)b(w) = \iota_{z,w} \left( \frac{1}{(z-w)^2} \right) + :b(z)b(w):, \quad (2.3.3)$$

$$b(w)b(z) = \iota_{w,z} \left( \frac{1}{(z-w)^2} \right) + :b(w)b(z):, \quad (2.3.4)$$

where  $\iota_{z,w}$  means the expansion in positive powers of  $w/z$  (think  $|z| > |w|$ ), and  $\iota_{w,z}$  means the expansion in positive powers of  $z/w$  (think  $|w| > |z|$ ). Subtract the two expansions to get

$$[b(z), b(w)] = \iota_{z,w} \left( \frac{1}{(z-w)^2} \right) - \iota_{w,z} \left( \frac{1}{(z-w)^2} \right)$$

Multiply by  $(z-w)^2$ . Since both  $\iota_{z,w}$  and  $\iota_{w,z}$  are just two formal expansions of the same rational function, you get

$$(z-w)^2 [b(z), b(w)] = 1 - 1 = 0 \quad \text{in } \text{End}(V)[[z^{\pm 1}, w^{\pm 1}]].$$

This is exactly the locality condition for the fields  $b(z)$  and  $b(w)$ .

In order to establish that other fields in the Heisenberg vertex algebra are local, we have a general result that shows that if  $A(z), B(z), C(z)$  are fields which are pairwise mutually local, then the normally ordered product  $:A(z)B(z):$  is local with respect to  $C(z)$ .

**Lemma 10.6 (2.3.4, Dong's Lemma).** *If  $A(z)$ ,  $B(z)$ ,  $C(z)$  are three mutually local fields, then the fields  $A(z)B(z)$  and  $C(z)$  are also mutually local.*

2.3.5. *Proof.* By assumption we may find  $r$  so that for all  $s \geq r$ ,

$$\begin{aligned}(w - z)^s A(z)B(w) &= (w - z)^s B(w)A(z), \\ (u - z)^s A(z)C(u) &= (u - z)^s C(u)A(z), \\ (u - w)^s B(w)C(u) &= (u - w)^s C(u)B(w).\end{aligned}$$

We wish to find an integer  $N$  such that

$$(w - u)^N : A(w)B(w) : C(u) = (w - u)^N C(u) : A(w)B(w) : .$$

By part (3) of Lemma 2.2.3, this will follow from the statement

$$\begin{aligned}(w - u)^N (\delta(z - w)_- A(z)B(w) + \delta(z - w)_+ B(w)A(z)) C(u) \\ = (w - u)^N C(u) (\delta(z - w)_- A(z)B(w) + \delta(z - w)_+ B(w)A(z))\end{aligned} \tag{73}$$

Let us take  $N = 3r$ . Writing

$$(w - u)^{3r} = (w - u)^r \sum_{s=0}^{2r} \binom{2r}{s} (w - z)^s (z - u)^{2r-s},$$

we study each summand

$$T_s := (w - u)^r (w - z)^s (z - u)^{2r-s} [\delta_- A(z)B(w) + \delta_+ B(w)A(z)] C(u).$$

For  $0 \leq s \leq r$ , notice that  $2r - s \geq r$ . Using locality  $(u - z)^r [A(z), C(u)] = 0$  we can move  $C(u)$  past  $A(z)$  (note that  $\delta$ 's are central):

$$(z - u)^{2r-s} A(z)C(u) = (z - u)^{2r-s} C(u)A(z)$$

Then using  $(w - u)^r [B(w), C(u)] = 0$  we can move  $C(u)$  past  $B(w)$ :

$$(w - u)^r B(w)C(u) = (w - u)^r C(u)B(w)$$

The net effect is:

$$T_s = (w - u)^r (w - z)^s (z - u)^{2r-s} C(u) [\delta_- A(z)B(w) + \delta_+ B(w)A(z)]$$

which is exactly the  $s$ -term from the right-hand side of (2.3.6). So the contributions for  $s \leq r$  match on both sides.

For  $r < s \leq 2r$ , here there are more than  $r$  powers of  $(w - z)$ . First use  $(w - z)^r [A(z), B(w)] = 0$  to swap  $B(w)A(z)$  to  $A(z)B(w)$  inside the  $\delta_+$ -term, keeping at least one extra  $(w - z)$  factor:

$$(w - z)^s \delta_+ B(w)A(z) = (w - z)^{s-r} \delta_+ (w - z)^r B(w)A(z) = (w - z)^{s-r} \delta_+ (w - z)^r A(z)B(w).$$

Thus in  $T_s$  we may replace

$$\delta_- A(z)B(w) + \delta_+ B(w)A(z) \rightsquigarrow (\delta_- + \delta_+) A(z)B(w) = \delta(z - w) A(z)B(w).$$

But now there is at least one leftover  $(w - z)$  factor ( $s - r \geq 1$ ), and

$$(w - z) \delta(z - w) = 0,$$

so  $T_s = 0$ . In words: after you use the available  $(w - z)^r$  to commute  $A$  through  $B$ , the remaining  $(w - z)$  kills the delta-torsion.

The same analysis applies symmetrically to the right-hand side of (2.3.6). Hence, after summing over  $s$ , both sides agree.  $\square$

**Theorem 10.7 (2.3.7).** *The Fock representation  $\pi$  with the structures*

- **Vacuum vector:**  $|0\rangle = 1$ .
- **Translation operator:**  $T$  is defined by the rules

$$T \cdot 1 = 0, \quad [T, b_i] = -i b_{i-1}.$$

- **Vertex operators:**

$$Y(b_{j_1} b_{j_2} \cdots b_{j_k}, z) = \frac{1}{(-j_1 - 1)! \cdots (-j_k - 1)!} : \partial_z^{-j_1-1} b(z) \cdots \partial_z^{-j_k-1} b(z) : .$$

- **$\mathbb{Z}_+$ -gradation:**

$$\deg(b_{j_1} \cdots b_{j_k}) = - \sum_i j_i.$$

satisfies the axioms of a  $\mathbb{Z}_+$ -graded vertex algebra.

*Proof.* First we verify that the field  $Y(b_{j_1} b_{j_2} \cdots b_{j_k}, z)$  defined above is homogeneous of conformal dimension  $-(j_1 + \cdots + j_k)$ , using Lemma 2.2.3(2). Now we check the axioms of a vertex algebra.

The statement  $Y(|0\rangle, z) = \text{Id}$  follows from our definition. The remainder of the vacuum axiom,

$$\lim_{z \rightarrow 0} Y(A, z)|0\rangle = A, \tag{74}$$

follows by induction. We start with the case  $A = b_{-1}$ , where

$$Y(b_{-1}, z)|0\rangle = \sum_{n \in \mathbb{Z}} b_n z^{-n-1} |0\rangle.$$

But all nonnegative  $b_n$  annihilate the vacuum, so the limit is well defined, while the constant coefficient is indeed  $b_{-1}$ .

Next, according to the above definition, the vertex operator associated to each polynomial in the  $b_i$ 's is a normally ordered product of derivatives of our basic field  $b(z)$ . Thus we just need to check that if (74) holds for the field

$$Y(A, z) = \sum_{n \in \mathbb{Z}} A_{(n)} z^{-n-1},$$

then it also holds for

$$Y(b_{-k}A, z) = \frac{1}{(k-1)!} : \partial_z^{k-1} b(z) Y(A, z) :, \quad k > 0.$$

By definition of the normally ordered product,

$$\begin{aligned} \frac{1}{(k-1)!} : \partial_z^{k-1} b(z) Y(A, z) : &:= \frac{1}{(k-1)!} \sum_{m \in \mathbb{Z}} \left( \sum_{n \leq -k} (-n-1)(-n-2) \cdots (-n-k+1) b_n A_{(m-n)} \right. \\ &\quad \left. + \sum_{n \geq 0} (-n-1)(-n-2) \cdots (-n-k+1) A_{(m-n)} b_n \right) z^{-m-k-1}. \end{aligned}$$

where the inner sums come from the normal ordering rule splitting the modes into “creation” and “annihilation.” The second sum kills  $|0\rangle$ , and by our inductive assumption, the first sum gives a power series with positive powers of  $z$  only, with the constant term

$$b_{-k} A_{(-1)} |0\rangle = b_{-k} A.$$

To check the translation axiom, we first note that  $T|0\rangle = 0$  by our definition. Next, since  $[T, b_j] = -j b_{j-1}$ , we have  $[T, b(z)] = \partial_z b(z)$ . Continuing, we derive  $[T, \partial_z^n b(z)] = \partial_z^{n+1} b(z)$ . Then one checks explicitly that the Leibniz rule holds for the normally ordered product:

$$\partial_z : A(z) B(z) : = : \partial_z A(z) B(z) : + : A(z) \partial_z B(z) :,$$

using the residue definition of the normal ordering given in Lemma 2.2.3(3). This implies that if  $[T, \cdot]$  acts as  $\partial_z$  on two fields, it will act this way on their normally ordered product. By induction, this implies the full translation axiom.

Finally, locality of any two fields follows by a similar induction from the locality of the fields  $\partial_z^n b(z)$  and  $\partial_w^m b(w)$  with arbitrary  $n, m \geq 0$  using Dong's Lemma. To see that  $\partial_z^n b(z)$  and

$\partial_w^m b(w)$  are mutually local, recall that  $b(z)$  and  $b(w)$  are local. But if  $A(z), B(w)$  are local, then  $(z - w)^N [A(z), B(w)] = 0$  for some  $N$ .

Differentiating this formula with respect to  $z$ , and multiplying the result by  $(z - w)$ , we obtain

$$(z - w)^{N+1} [\partial_z A(z), B(w)] = 0,$$

so  $\partial_z A(z)$  and  $B(w)$  are mutually local. Hence, by induction,  $\partial_z^n b(z)$  and  $\partial_w^m b(w)$  are indeed local for any  $n, m \geq 0$ , and this completes the proof.  $\square$

The above proof is easily generalized to yield the following result, which provides a “generators–and–relations” approach to the construction of vertex algebras.

Let  $V$  be a vector space,  $|0\rangle$  a nonzero vector, and  $T$  an endomorphism of  $V$ . Let  $S$  be a countable ordered set and  $\{a^\alpha\}_{\alpha \in S}$  a collection of vectors in  $V$ . Suppose we are also given fields

$$a^\alpha(z) = \sum_{n \in \mathbb{Z}} a_{(n)}^\alpha z^{-n-1}$$

such that the following conditions hold:

1. For all  $\alpha$ ,  $a^\alpha(z)|0\rangle = a^\alpha + z(\dots)$ .
2.  $T|0\rangle = 0$  and  $[T, a^\alpha(z)] = \partial_z a^\alpha(z)$  for all  $\alpha$ .
3. All fields  $a^\alpha(z)$  are mutually local.
4.  $V$  has a basis of vectors

$$a_{(j_1)}^{\alpha_1} \cdots a_{(j_m)}^{\alpha_m} |0\rangle, \quad j_1 \leq j_2 \leq \cdots \leq j_m < 0,$$

and if  $j_i = j_{i+1}$ , then  $\alpha_i \preceq \alpha_{i+1}$  with respect to the given order on the set  $S$ .

**This basis should be compared with the PBW basis.**

**Theorem 10.8 ((Preliminary) Reconstruction Theorem).** *Under the above assumptions, the assignment*

$$Y(a_{(j_1)}^{\alpha_1} \cdots a_{(j_m)}^{\alpha_m} |0\rangle, z) = \frac{1}{(-j_1 - 1)! \cdots (-j_m - 1)!} : \partial_z^{-j_1-1} a^{\alpha_1}(z) \cdots \partial_z^{-j_m-1} a^{\alpha_m}(z) : \quad (75)$$

*defines a vertex algebra structure on  $V$ . Moreover, if  $V$  is a  $\mathbb{Z}$ -graded vector space,  $|0\rangle$  has degree 0, the vectors  $a^\alpha$  are homogeneous,  $T$  has degree 1, and the fields  $a^\alpha(z)$  have conformal dimension  $\deg a^\alpha$ , then  $V$  is a  $\mathbb{Z}$ -graded vertex algebra.*

It is practically identical to the proof in the Heisenberg algebra case, where we had

$$V = \pi_0 \simeq \mathbb{C}[b_{-1}, b_{-2}, \dots] |0\rangle, a^\alpha = b_{-1}, \text{ and } a^\alpha(z) = b(z) = \sum_{n \in \mathbb{Z}} b_n z^{-n-1}.$$

### 10.3 Vertex algebras from affine Kac-Moody algebras

Let  $\hat{\mathfrak{g}}$  be the (derived) affine Kac-Moody algebra associated to a simple Lie algebra  $\mathfrak{g}$ . In particular, it fits into the exact sequence

$$0 \rightarrow \mathbb{C}K \rightarrow \hat{\mathfrak{g}} \rightarrow L\mathfrak{g} \rightarrow 0.$$

with bracket

$$[x \otimes f(z), y \otimes g(z)] = [x, y] \otimes f(z)g(z) - (x, y) \text{Res}_{z=0} f dg \cdot K$$

where the bilinear form  $(\cdot, \cdot)$  is the invariant form on  $\mathfrak{g}$ , normalized so that the highest root has length squared 2.

Inside  $\hat{\mathfrak{g}}$  we have the positive Lie subalgebra  $\mathfrak{g}[[z]] \oplus \mathbb{C}K$ . Consider the one dimensional representation  $\mathbb{C}_k$  of this subalgebra, where  $\mathfrak{g}[[z]]$  acts by 0 and  $K$  acts by  $k \in \mathbb{C}$ . We define the vacuum representation of  $\hat{\mathfrak{g}}$  at level  $k$  to be the induced representation

$$V_k(\mathfrak{g}) = U(\hat{\mathfrak{g}}) \otimes_{U(\mathfrak{g}[[z]] \oplus \mathbb{C}K)} \mathbb{C}_k.$$

By the PBW theorem,  $V_k(\mathfrak{g}) \simeq U(z^{-1}\mathfrak{g}[z^{-1}])$  as vector spaces. We will now put a vertex algebra structure of  $V_k(\mathfrak{g})$ . Let  $\{J^a\}$  be an ordered basis of  $\mathfrak{g}$  and for  $A \in \mathfrak{g}, n \in \mathbb{Z}$  let  $A_n = A \otimes z^n \in \hat{\mathfrak{g}}$ .

- **Vacuum vector:**  $|0\rangle = v_k$ .
- **Translation operator:**  $Tv_k = 0, \quad [T, J_n^a] = -nJ_{n-1}^a$ .
- **Vertex operators:**

$$Y(v_k, z) = \text{Id}, \quad Y(J_{-1}^a v_k, z) = J^a(z) = \sum_{n \in \mathbb{Z}} J_n^a z^{-n-1}.$$

In general,

$$Y(J_{n_1}^{a_1} \cdots J_{n_m}^{a_m} v_k, z) = \frac{1}{(-n_1 - 1)! \cdots (-n_m - 1)!} : \partial_z^{-n_1-1} J^{a_1}(z) \cdots \partial_z^{-n_m-1} J^{a_m}(z) : . \quad (76)$$

- **$\mathbb{Z}_+$ -gradation:**

$$\deg(J_{n_1}^{a_1} \cdots J_{n_m}^{a_m} v_k) = - \sum_{i=1}^m n_i.$$

**Theorem 10.9 (Affine Kac-Moody vertex algebra).**  $V_k(\mathfrak{g})$  is a  $\mathbb{Z}_+$ -graded vertex algebra.

*Proof.* We check the conditions of the reconstruction theorem. We need  $J^a(z)|0\rangle = J_{-1}^a|0\rangle + z(\dots)$ . This follows because  $J_n^a v_k = 0$  for  $n \geq 0$ .



The translation operator  $T$  satisfies

$$[T, J_n^a] = -nJ_{n-1}^a$$

which immediately implies  $[T, J^a(z)] = \partial_z J^a(z)$ .

To see that fields  $J^a(z)$  are mutually local, we compute

$$\begin{aligned} [J^a(z), J^b(w)] &= \sum_{n \in \mathbb{Z}} \sum_{m \in \mathbb{Z}} [J_n^a, J_m^b] z^{-n-1} w^{-m-1} \\ &= \sum_{n, m \in \mathbb{Z}} [J^a, J^b]_{n+m} z^{-n-1} w^{-m-1} + \sum_{n \in \mathbb{Z}} nk(J^a, J^b) z^{-n-1} w^{n-1} \\ &= \sum_{\ell \in \mathbb{Z}} [J^a, J^b]_{\ell} \left( \sum_{n \in \mathbb{Z}} z^{-n-1} w^{-\ell+n} \right) w^{-\ell-1} + (J^a, J^b)k \sum_{n \in \mathbb{Z}} nz^{-n-1} w^{n-1} \\ &= [J^a, J^b](w) \delta(z-w) + (J^a, J^b)k \partial_w \delta(z-w). \end{aligned}$$

It then follows that for any  $a, b$ ,

$$(z-w)^2 [J^a(z), J^b(w)] = 0,$$

The final condition is that  $V_k(\mathfrak{g})$  has a basis of vectors

$$J_{n_1}^{a_1} \cdots J_{n_m}^{a_m} v_k, \quad n_1 \leq n_2 \leq \cdots \leq n_m < 0,$$

which follows from the PBW theorem.  $\square$

## 10.4 The Virasoro algebra

From now on we will use the notation  $\mathcal{K}$  for  $\mathbb{C}((t))$  and  $\mathcal{O}$  for  $\mathbb{C}[[t]]$ . Consider the Lie algebra  $\text{Der } \mathcal{K} = \mathbb{C}((t))\partial_t$  of (continuous) derivations of  $\mathcal{K}$ . The **Virasoro algebra** is by definition the central extension of  $\text{Der } \mathcal{K}$ :

$$0 \rightarrow \mathbb{C}C \rightarrow \text{Vir} \rightarrow \text{Der } \mathcal{K} \rightarrow 0,$$

defined by the commutation relations

$$[f(t)\partial_t, g(t)\partial_t] = (fg' - f'g)\partial_t - \frac{1}{12}(\text{Res}_{t=0} fg''' dt)C. \quad (77)$$

It is known that this extension is universal. It has (topological) generators  $C$  and

$$L_n = -t^{n+1}\partial_t, \quad n \in \mathbb{Z},$$

satisfying the relations that  $C$  is central and

$$[L_n, L_m] = (n - m)L_{n+m} + \frac{n^3 - n}{12} \delta_{n,-m} C. \quad (78)$$

The last term comes from

$$\text{Res}_{t=0} t^{n+1} (m+1)m(m-1)t^{m-2} dt.$$

Unlike the Heisenberg or Kac–Moody cocycles, the cocycle used to define the Virasoro algebra is **not coordinate-independent**. Thus, it is not clear that a coordinate-independent Virasoro algebra exists. Fortunately, it does exist, as explained in § 19.6.3 below. However, there is no canonical splitting of the corresponding extension, even as a vector space. This explains why the above formula depends on the choice of the coordinate  $t$ .

We will say that a module  $M$  over the Virasoro algebra has **central charge**  $c \in \mathbb{C}$  if  $C$  acts on  $M$  by multiplication by  $c$ .

We may write the Virasoro relations succinctly using a generating function

$$T(z) = \sum_{n \in \mathbb{Z}} L_n z^{-n-2}$$

(the convention on the power of  $z$  will become natural when we realize  $T(z)$  as a field of conformal dimension 2).

**Lemma 10.10 (2.5.4).**

$$[T(z), T(w)] = \frac{C}{12} \partial_w^3 \delta(z - w) + 2T(w) \partial_w \delta(z - w) + \partial_w T(w) \cdot \delta(z - w)$$

as formal power series in  $z^{\pm 1}, w^{\pm 1}$ .

*Proof of Lemma 2.5.4.* We have

$$\begin{aligned} [T(z), T(w)] &= \sum_{n,m} (n - m) L_{n+m} z^{-n-2} w^{-m-2} + C \sum_n \frac{n^3 - n}{12} z^{-n-2} w^{n-2} \\ &= \sum_{j,l} 2L_j w^{-j-2} z^{-l-1} w^{l-1} + \sum_{j,l} (-j - 2) L_j w^{-j-3} z^{-l-1} w^l \\ &\quad + \frac{C}{12} \sum_l l(l-1)(l-2) z^{-l-1} w^{l-3} \\ &= 2T(w) \partial_w \delta(z - w) + \partial_w T(w) \cdot \delta(z - w) + \frac{C}{12} \partial_w^3 \delta(z - w), \end{aligned}$$

where we have made the substitution  $j = n + m, l = n + 1$ .  $\square$

Let  $U(\text{Vir})$  be the universal enveloping algebra of  $\text{Vir}$ . For each  $c \in \mathbb{C}$  we define the induced representation

$$\text{Vir}_c = \text{Ind}_{\text{Der } \mathcal{O} \oplus \mathbb{C}C}^{\text{Vir}} \mathbb{C}_c = U(\text{Vir}) \otimes_{U(\text{Der } \mathcal{O} \oplus \mathbb{C}C)} \mathbb{C}_c,$$

where  $C$  acts as multiplication by  $c$  and  $\text{Der } \mathcal{O}$  acts by zero on the one-dimensional module  $\mathbb{C}_c$ . Note that  $\text{Vir}_c$  has central charge  $c$  as a module over the Virasoro algebra.

By the Poincaré–Birkhoff–Witt theorem,  $\text{Vir}_c$  has a PBW basis consisting of monomials of the form

$$L_{j_1} \cdots L_{j_m} v_c, \quad j_1 \leq j_2 \leq \cdots \leq j_m \leq -2. \quad (79)$$

Here  $v_c$  is the image of  $1 \otimes 1 \in U(\text{Vir}) \otimes \mathbb{C}_c$  in the induced representation, and we take it to be the vacuum vector of the vertex algebra. We define a  $\mathbb{Z}_+$ -gradation on  $\text{Vir}_c$  by the formulas  $\deg L_n = -n$ ,  $\deg v_c = 0$ .

As the translation operator we take  $T = L_{-1}$ . Finally, we use the Reconstruction Theorem to define the vertex operators. Namely, we set

$$Y(L_{-2}v_c, z) \stackrel{\text{def}}{=} T(z) = \sum_{n \in \mathbb{Z}} L_n z^{-n-2}.$$

This is the generating field of  $\text{Vir}_c$ . It has conformal dimension 2, because  $\deg L_{-2}v_c = 2$  by the above definition of gradation on  $\text{Vir}_c$ .

Next, we define the vertex operators  $Y(A, z)$  for the PBW monomials of the form (79) according to the prescription of Theorem 2.3.11:

$$Y(L_{j_1} \cdots L_{j_m} v_c, z) = \frac{1}{(-j_1 - 2)! \cdots (-j_m - 2)!} : \partial_z^{-j_1-2} T(z) \cdots \partial_z^{-j_m-2} T(z) : .$$

Lemma 2.5.4 implies that

$$(z - w)^4 [T(z), T(w)] = 0.$$

Hence the generating field  $T(z)$  is local with itself. The Reconstruction Theorem 2.3.11 then implies that  $\text{Vir}_c$  with the above structures is a vertex algebra for any  $c \in \mathbb{C}$ . Note that the gradation on  $\text{Vir}_c$  coincides with the gradation by the eigenvalues of  $L_0$ .

**Definition 10.11.** A  $\mathbb{Z}$ -graded vertex algebra  $V$  is called **conformal**, of central charge  $c \in \mathbb{C}$ , if we are given a non-zero **conformal vector**  $\omega \in V_2$  such that the Fourier coefficients  $L_n^V$  of the corresponding vertex operator

$$Y(\omega, z) = \sum_{n \in \mathbb{Z}} L_n^V z^{-n-2} \quad (80)$$

satisfy the defining relations of the Virasoro algebra with central charge  $c$ , and in addition we have  $L_{-1}^V = T$ ,  $L_0^V|_{V_n} = n \text{Id}$ .

In particular, being conformal means it contains a vertex subalgebra generated by a field  $T(z)$  that is isomorphic to a Virasoro vertex algebra. Note that according to our notation scheme,  $\omega_{(n)} = L_{n-1}^V$ .

**Example 10.12.** *The Virasoro vertex algebra  $\text{Vir}_c$  is clearly conformal, with central charge  $c$  and conformal vector  $\omega = L_{-2}v_c$ .*

**Example 10.13.** *The Heisenberg vertex algebra  $\pi$  comes with a natural one-parameter family of conformal structures. Given  $\lambda \in \mathbb{C}$ , set*

$$\omega_\lambda = \frac{1}{2}b_{-1}^2v_0 + \lambda b_{-2}v_0.$$

*Then  $\omega_\lambda$  is a conformal vector with central charge  $c_\lambda = 1 - 12\lambda^2$ . In particular, we obtain a family of representations of the Virasoro algebra on  $\pi$  (thus, if  $\lambda = 0$  we obtain a representation with central charge 1, which explains our normalization of the two-cocycle defining the Virasoro algebra). To see this we need to check that the Fourier coefficients of the field*

$$Y(\omega_\lambda, z) = \frac{1}{2} : b(z)^2 : + \lambda \partial_z b(z) = \sum_{n \in \mathbb{Z}} \omega_{\lambda, n} z^{-n-2}$$

*satisfy the Virasoro relations, that  $\omega_{\lambda, -1} = T$ , and that  $\omega_{\lambda, 0}$  is the degree operator.*

*The other two conditions are easy to verify. Indeed, we have (where integration is formal and means taking the residue)*

$$\omega_{\lambda, -1} = \int \omega_\lambda(z) dz = \frac{1}{2} \sum_{i+j=-1} b_i b_j,$$

*We have*

$$\omega_{\lambda, -1} = \frac{1}{2} \sum_{i+j=-1} b_i b_j, \quad [b_i, b_j] = i \delta_{i, -j}.$$

*Then for any  $k \in \mathbb{Z}$ ,*

$$\begin{aligned} [\omega_{\lambda, -1}, b_k] &= \frac{1}{2} \sum_{i+j=-1} [b_i b_j, b_k] \\ &= \frac{1}{2} \sum_{i+j=-1} (b_i [b_j, b_k] + [b_i, b_k] b_j) \\ &= \frac{1}{2} \sum_{i+j=-1} (b_i (j \delta_{j, -k}) + (i \delta_{i, -k}) b_j). \end{aligned}$$

*Only the terms with  $(i, j) = (k-1, -k)$  and  $(i, j) = (-k, k-1)$  survive. Therefore*

$$[\omega_{\lambda, -1}, b_k] = \frac{1}{2}(-k)b_{k-1} + \frac{1}{2}(-k)b_{k-1} = -kb_{k-1}.$$

*Therefore  $\omega_{\lambda, -1} = T$ .*

As for the degree operator, we have

$$\omega_\lambda(z) = \frac{1}{2} : b(z)^2 : + \lambda \partial_z b(z), \quad \omega_{\lambda,0} = \text{Res}_z (z \omega_\lambda(z)).$$

The derivative term gives no residue because  $b_0 = 0$ , so

$$\omega_{\lambda,0} = \frac{1}{2} \text{Res}_z (z : b(z)^2 :).$$

Expanding  $b(z) = \sum n \in \mathbb{Z} b_n z^{-n-1}$  gives

$$: b(z)^2 : = \sum_{n \in \mathbb{Z}} \left( \sum_{i+j=n} : b_i b_j : \right) z^{-n-2}$$

Multiplying by  $z$  and taking the residue selects the coefficient of  $z^{-2}$ , i.e.  $n = 0$ :

$$\omega_{\lambda,0} = \frac{1}{2} \sum_{i+j=0} : b_i b_j : = \frac{1}{2} \sum_{n \in \mathbb{Z}} : b_n b_{-n} :.$$

For  $n > 0$  the normal ordering gives  $: b_{-n} b_n : = b_{-n} b_n$ , while  $b_0$  acts as 0. Hence

$$\omega_{\lambda,0} = \sum_{n>0} b_{-n} b_n.$$

In the Fock space  $\pi \simeq \mathbb{C}[b_{-1}, b_{-2}, \dots]$ , one has  $b_{-n}$  acts by multiplication by  $b_{-n}$  and  $b_n$  acts as  $n \frac{\partial}{\partial b_{-n}}$ . Thus

$$\omega_{\lambda,0} = \sum_{n>0} n b_{-n} \frac{\partial}{\partial b_{-n}}$$

This is the Euler (degree) operator, acting by the grading on  $\pi$ . Therefore  $\omega_{\lambda,0}$  plays the role of  $L_0$ , and together with  $\omega_{\lambda,-1} = T$  confirms that  $\omega_\lambda$  is a conformal vector.

It turns out, proved in Lemma 3.4.5, is that the above two conditions are actually sufficient for proving that  $\omega_\lambda$  is a conformal vector (albeit not sufficient to determine the central charge  $c$ ).

## 10.5 The Segal–Sugawara construction.

The vertex algebra  $V_k(\mathfrak{g})$  has a natural conformal vector, called the **Segal–Sugawara vector**, when  $k \neq -h^\vee$ , where  $h^\vee$  denotes the dual Coxeter number of  $\mathfrak{g}$ . Recall from § 2.4.2 that we have an isomorphism of vector spaces

$$V_k(\mathfrak{g}) \simeq U(\mathfrak{g} \otimes t^{-1} \mathbb{C}[t^{-1}]) v_k.$$

Pick a basis  $\{J_a\}_{a=1}^d$  of  $\mathfrak{g}$  (where  $d = \dim \mathfrak{g}$ ), and let  $\{J_a\}_{a=1}^d$  be its dual basis with respect to the invariant bilinear form  $(\cdot, \cdot)$ . We write

$$J^a(z) = \sum_{n \in \mathbb{Z}} J_n^a z^{-n-1}, \quad J_a(z) = \sum_{n \in \mathbb{Z}} J_{a,n} z^{-n-1}.$$

Set

$$S = \frac{1}{2} \sum_{a=1}^d J_{a,-1} J_{-1}^a v_k \quad (81)$$

(clearly, it is independent of the choice of basis  $\{J^a\}$ ). Then for  $k \neq -h^\vee$ ,  $S/(k + h^\vee)$  is a conformal vector in  $V_k(\mathfrak{g})$ , of central charge

$$c(k) = \frac{k \dim \mathfrak{g}}{k + h^\vee}.$$

We call it the **Segal–Sugawara conformal vector**. Thus, the Fourier coefficients of the field

$$\frac{1}{k + h^\vee} Y(S, z) = \frac{1}{2(k + h^\vee)} \sum_{a=1}^d : J_a(z) J^a(z) : \quad (82)$$

generate an action of the Virasoro algebra with central charge  $c(k)$  on  $V_k(\mathfrak{g})$ .

Checking the Virasoro relations directly is rather difficult, so we will do it later in § 3.4.8 using the operator product expansion (OPE). Further, the resulting Virasoro operators satisfy the following commutation relations with the elements of  $\widehat{\mathfrak{g}}$ :

$$[L_n, J_m^a] = -m J_{n+m}^a. \quad (83)$$

Formula (83) will be proved in § 3.4.8 using the formalism of operator product expansion.

## 11 Associativity and the operator product expansion

In this section, we will discuss the operator product expansion (OPE) and its role in establishing the associativity of the vertex operator algebra structure on  $V_k(\mathfrak{g})$ . The OPE is a powerful tool that allows us to express the product of two vertex operators in terms of a sum of vertex operators with shifted arguments.

**Theorem 11.1 (Goddard’s Uniqueness Theorem).** *Let  $V$  be a vertex algebra, and let  $A(z)$  be a field on  $V$ . Suppose there exists a vector  $a \in V$  such that*

$$A(z) |0\rangle = Y(a, z) |0\rangle$$

*and that  $A(z)$  is local with respect to  $Y(b, z)$  for all  $b \in V$ . Then  $A(z) = Y(a, z)$ .*

*Proof.* Take any  $b \in V$ . Using the formulation of locality from Proposition 1.2.5, we obtain the following equalities in  $V[[z^{\pm 1}, w^{\pm 1}]]$  for large enough  $N$ :

$$\begin{aligned} (z - w)^N A(z) Y(b, w) |0\rangle &= (z - w)^N Y(b, w) A(z) |0\rangle && \text{by locality} \\ &= (z - w)^N Y(b, w) Y(a, z) |0\rangle && \text{by assumption} \\ &= (z - w)^N Y(a, z) Y(b, w) |0\rangle && \text{by locality} \end{aligned}$$

By the vacuum axiom, both  $(z - w)^N Y(a, z)Y(b, w)|0\rangle$  and  $(z - w)^N A(z)Y(b, w)|0\rangle$  are well-defined at  $w = 0$ . Setting  $w = 0$ , and using  $Y(b, w)|0\rangle|_{w=0} = b$ , we obtain

$$Y(a, z)b = A(z)b \quad \text{for all } b \in V,$$

which is what we wanted to show.  $\square$

**Lemma 11.2.** *We have the formula*

$$Y(a, z)|0\rangle = e^{zT}a$$

**Corollary 11.3.** *For all  $a \in V$ , we have*

$$Y(Ta, z) = \partial_z Y(a, z).$$

*Proof.* We have

$$\partial_z^2 Y(a, z)|0\rangle = \partial_z([T, Y(a, z)]|0\rangle) = \partial_z(TY(a, z)|0\rangle) = T\partial_z Y(a, z)|0\rangle,$$

and

$$\partial_z Y(a, z)|0\rangle|_{z=0} = [T, Y(a, z)]|0\rangle|_{z=0} = TY(a, z)|0\rangle|_{z=0} = Ta.$$

In view of Remark 3.1.5, the corollary now follows from Theorem 3.1.1 applied to the field

$$A(z) = \partial_z Y(a, z).$$

$\square$

## 11.1 Associativity and the operator product expansion

Recall that there are two formal maps:

$$i_{z,w} : V[[z, w]][z^{-1}, w^{-1}, (z - w)^{-1}] \longrightarrow V((z))((w))$$

and

$$i_{w,z} : V[[z, w]][z^{-1}, w^{-1}, (z - w)^{-1}] \longrightarrow V((w))((z)).$$

which differ only in the expansion of  $(z - w)^{-1}$ . The locality axiom states that

$$Y(A, z)Y(B, w)C \in V((z))((w)) \text{ and } Y(B, w)Y(A, z)C \in V((w))((z))$$

are the expansions of the same element of

$$V[[z, w]][z^{-1}, w^{-1}, (z - w)^{-1}]$$

Our associativity theorem introduces a third expression which satisfies a similar property.

**Theorem 11.4 (Associativity).** Any vertex algebra  $V$  satisfies the following associativity property: for all  $A, B, C \in V$ , the three expressions

$$\begin{aligned} Y(A, z)Y(B, w)C &\in V((z))((w)), \\ Y(B, w)Y(A, z)C &\in V((w))((z)), \\ Y(Y(A, z - w)B, w)C &\in V((w))((z - w)) \end{aligned}$$

are the expansions, in their respective domains, of the same element of

$$V[[z, w]][z^{-1}, w^{-1}, (z - w)^{-1}].$$

The proof follows from the = following lemmas.

**Lemma 11.5.**

$$e^{wT}Y(A, z)e^{-wT} = Y(A, z + w)$$

in  $\text{End } V[[z^{\pm 1}, w^{\pm 1}]]$ , where by negative powers of  $(z + w)^{-1}$  we understand their expansions in  $\mathbb{C}((z))((w))$ , i.e. in positive powers of  $w/z$ .

*Proof.* By the translation axiom,  $[T, Y(A, z)] = \partial_z Y(A, z)$ . Also recall the adjoint exponential identity

$$e^X Y e^{-X} = e^{\text{ad } X} \cdot Y = \sum_{n=0}^{\infty} \frac{1}{n!} (\text{ad } X)^n \cdot Y$$

which is a standard Baker-Campbell-Hausdorff expansion, obtained by differentiating with respect to a parameter  $t$ :

$$\frac{d}{dt} (e^{tX} Y e^{-tX}) = e^{tX} [X, Y] e^{-tX}$$

and integrating from  $t=0$  to  $t=1$ . Hence

$$e^{wT}Y(A, z)e^{-wT} = \sum_{n=0}^{\infty} \frac{1}{n!} (\text{ad } wT)^n \cdot Y(A, z) = \sum_{n=0}^{\infty} \frac{w^n}{n!} \partial_z^n Y(A, z),$$

where we use the notation  $\text{ad } B \cdot C = [B, C]$ . To complete the proof, we use the formal-variable version of the Taylor formula: for any algebra  $R$  and a formal power series  $f(z) \in R[[z^{\pm 1}]]$ , we have the identity

$$\sum_{n=0}^{\infty} \frac{w^n}{n!} \partial_z^n f(z) = f(z + w)$$

in  $R[[z^{\pm 1}, w^{\pm 1}]]$ .  $\square$



**Proposition 11.6 (Skew-symmetry).** *The identity*

$$Y(A, z)B = e^{zT}Y(B, -z)A \quad (84)$$

holds in  $V((z))$ .

*Proof.* First, note that for any field  $\phi(z)$  on  $V$ , the product  $e^{zT}\phi(z)$  is a well-defined element of  $\text{End } V[[z^{\pm 1}]]$ , so the statement of the lemma makes sense.

By locality,

$$(z - w)^N Y(A, z)Y(B, w)|0\rangle = (z - w)^N Y(B, w)Y(A, z)|0\rangle,$$

for large enough  $N$ . Using Lemma 3.1.3, we obtain

$$(z - w)^N Y(A, z)e^{wT}B = (z - w)^N Y(B, w)e^{zT}A.$$

Applying Lemma 11.5, we find that

$$(z - w)^N Y(A, z)e^{wT}B = (z - w)^N e^{zT}Y(B, w - z)A,$$

where by  $(w - z)^{-1}$  we understand its expansion in  $\mathbb{C}((w))((z))$ . We may always choose  $N$  large enough so that the right-hand side of this formula does not contain any negative powers of  $w - z$ . Then it becomes an identity in  $V((z))[[w]]$ , in which we can set  $w = 0$ , and then divide by  $z^N$ . This gives formula (84).  $\square$

Consider the expression  $Y(A, z)Y(B, w)C$ . By locality, it is the expansion in  $V((z))((w))$  of an element of the space

$$V[[z, w]][z^{-1}, w^{-1}, (z - w)^{-1}].$$

Let us denote this element by  $f_{A,B,C}$ . By Theorem 3.2.1, the expansion of  $f_{A,B,C}$  in  $V((w))((z - w))$  is equal to  $Y(Y(A, z - w)B, w)C$ . The corresponding embedding

$$V[[z, w]][z^{-1}, w^{-1}, (z - w)^{-1}] \rightarrow V((w))((z - w))$$

is obtained by sending  $z$  to  $w + (z - w)$  and  $z^{-1}$  to  $(w + (z - w))^{-1}$ , considered as a power series in positive powers of  $(z - w)/w$ . Thus, we find that applying this procedure to  $f_{A,B,C}$  we obtain  $Y(Y(A, z - w)B, w)C$ . By abusing notation, we may write this as

$$Y(A, z)Y(B, w)C = Y(Y(A, z - w)B, w)C, \quad A, B, C \in V,$$

or as

$$Y(A, z)Y(B, w)C = \sum_{n \in \mathbb{Z}} \frac{Y(A_{(n)} \cdot B, w)}{(z - w)^{n+1}} C, \quad A, B, C \in V, \quad (85)$$

where the left-hand side should be understood in the above sense.

**Definition 11.7.** Formula (89) is called the **operator product expansion** (or **OPE** for short).

**Proposition 11.8** ([?, §3.3.1]). Let  $\phi(z)$  and  $\psi(w)$  be arbitrary fields on a vector space  $V$ . Then the following are equivalent:

(1) There is an identity in  $\text{End } V[[z^{\pm 1}, w^{\pm 1}]]$ :

$$[\phi(z), \psi(w)] = \sum_{j=0}^{N-1} \frac{1}{j!} \gamma_j(w) \partial_w^j \delta(z-w), \quad (86)$$

where  $\gamma_j(w)$ ,  $j = 0, \dots, N-1$ , are some fields.

(2)  $\phi(z)\psi(w)$  (resp.  $\psi(w)\phi(z)$ ) equals

$$\sum_{j=0}^{N-1} \frac{\gamma_j(w)}{(z-w)^{j+1}} + : \phi(z)\psi(w) :, \quad (87)$$

where  $1/(z-w)$  is expanded in positive powers of  $w/z$  (resp.  $z/w$ ).

(3)  $\phi(z)\psi(w)$  converges for  $|z| > |w|$  to the expression in (87), and  $\psi(w)\phi(z)$  converges for  $|w| > |z|$  to the same expression.

*Proof.* First we derive (2) from (1). Given a formal power series  $f(z) = \sum_{n \in \mathbb{Z}} f_n z^n$ , we will write, as before,

$$f(z)_+ = \sum_{n \geq 0} f_n z^n, \quad f(z)_- = \sum_{n < 0} f_n z^n.$$

On the left-hand side of (86),  $z$  appears only in the term  $\phi(z)$ . Thus, taking the negative Fourier coefficients with respect to  $z$ , we obtain

$$[\phi(z), \psi(w)]_- = (\text{R.H.S. of (86)})_- = \sum_{j=0}^{N-1} \frac{1}{j!} \gamma_j(w) \partial_w^j \delta(z-w)_-.$$

and similarly for  $(\cdot)_+$ . But since (by definition of normal ordering, since we are supposed to put all “creation operators” to the left) so

$$: \phi(z)\psi(w) := \psi(w)\phi(z)_- + \phi(z)_+\psi(w)$$

while

$$\phi(z)\psi(w) = \phi(z)_-\psi(w) + \phi(z)_+\psi(w),$$

we see that

$$[\phi(z), \psi(w)]_- = \phi(z)\psi(w)_- - : \phi(z)\psi(w) :$$

and plugging back in we find that

$$\phi(z)\psi(w) = \sum_{j=0}^{N-1} \frac{1}{j!} \gamma_j(w) \partial_w^j \delta(z-w)_- + : \phi(z)\psi(w) :$$

Since

$$\frac{1}{j!} \partial_w^j \delta(z-w)_- \text{ is the expansion of } \frac{1}{(z-w)^{j+1}} \text{ in } \mathbb{C}((z))((w)),$$

we obtain formula (87). The formula for  $\psi(w)\phi(z)$  is obtained in the same way by taking the positive part of the commutator  $[\psi(w), \phi(z)]$  (with respect to  $z$ ), which equals

$$[\psi(w), \phi(z)]_+ = \psi(w)\phi(z)_- - : \phi(z)\psi(w) :.$$

To derive part (1) from part (2), observe that when we take the commutator between  $\phi(z)$  and  $\psi(w)$ , the normally ordered terms cancel each other. Explicitly, we have

$$\begin{aligned} : \phi(z)\psi(w) : &= \phi(z)_- \psi(w) + \phi(z)_+ \psi(w) \\ : \psi(w)\phi(z) : &= \psi(w)_- \phi(z) + \psi(w)_+ \phi(z) \\ : \phi(z)\psi(w) : - : \psi(w)\phi(z) : &= (\phi(z)_- \psi(w) + \phi(z)_+ \psi(w)) - (\psi(w)_- \phi(z) + \psi(w)_+ \phi(z)) \\ &= \phi(z)_- \psi(w) + \phi(z)_+ \psi(w) - \psi(w)_- \phi(z) - \psi(w)_+ \phi(z) \\ &= \phi(z)_- (\psi(w)_- + \psi(w)_+) + \phi(z)_+ (\psi(w)_- + \psi(w)_+) \\ &\quad - \psi(w)_- (\phi(z)_- + \phi(z)_+) - \psi(w)_+ (\phi(z)_- + \phi(z)_+) \\ &= 0 \end{aligned}$$

We are left with the combination

$$\sum_{j=0}^{N-1} \frac{1}{j!} \gamma_j(w) (\partial_w^j \delta(z-w)_- + \partial_w^j \delta(z-w)_+) = \text{R.H.S. of (86)}.$$

The equivalence of (2) and (3) is clear.  $\square$

## 11.2 The OPE and the Reconstruction Theorem

Using the OPE we can now explain the formula for the vertex operators in the Reconstruction Theorem. Let  $Y(A, z)$  and  $Y(B, w)$  be two vertex operators local with each other. Then we are in the setup of OPE because of the following lemma, which says that the  $\delta$  function witnesses all of the torsion between two local fields.

**Lemma 11.9.** *Let  $f(z, w)$  be a formal power series in  $R[[z^{\pm 1}, w^{\pm 1}]]$  satisfying  $(z-w)^N f(z, w) = 0$  for a positive integer  $N$ . Then  $f(z, w)$  can be written uniquely as a sum*

$$\sum_{i=0}^{N-1} g_i(w) \partial_w^i \delta(z-w), \quad g_i(w) \in R[[w^{\pm 1}]].$$

Using the locality, we obtain

$$[Y(A, z), Y(B, w)] = \sum_{j=0}^{N-1} \frac{1}{j!} \gamma_j(w) \partial_w^j \delta(z - w), \quad (88)$$

where  $\gamma_j(w)$ ,  $j = 0, \dots, N - 1$ , are some fields.

Using Proposition 3.3.1, we find that

$$Y(A, z)Y(B, w) = \sum_{j=0}^{N-1} \frac{\gamma_j(w)}{(z - w)^{j+1}} + : Y(A, z)Y(B, w) :, \quad (89)$$

where by  $(z - w)^{-1}$  we understand its expansion in positive powers of  $w/z$ . Hence for any  $C \in V$ , the series  $Y(A, z)Y(B, w)C \in V((z))((w))$  is an expansion of

$$\left( \sum_{j=0}^{N-1} \frac{\gamma_j(w)}{(z - w)^{j+1}} + : Y(A, z)Y(B, w) : \right) C \in V[[z, w]][z^{-1}, w^{-1}, (z - w)^{-1}].$$

Using the Taylor formula, we obtain that the expansion of this element of  $V[[z, w]][z^{-1}, w^{-1}, (z - w)^{-1}]$  in  $V((w))((z - w))$  is equal to

$$\sum_{j=0}^{N-1} \frac{\gamma_j(w)}{(z - w)^{j+1}} + \sum_{m \geq 0} \frac{(z - w)^m}{m!} \partial_w^m Y(A, w) Y(B, w). \quad (90)$$

Recall that the OPE formula (89) states that

$$Y(A, z)Y(B, w)C = \sum_{n \in \mathbb{Z}} \frac{Y(A_{(n)} \cdot B, w)}{(z - w)^{n+1}} C.$$

The coefficient in front of  $(z - w)^k$ ,  $k \in \mathbb{Z}$ , in the right-hand side of (90) must be equal to the corresponding term in the right-hand. Let us first look at the terms with  $k \geq 0$ . Comparison of the two formulas gives

$$Y(A_{(n)} \cdot B, z) = \frac{1}{(-n - 1)!} \partial_z^{-n-1} Y(A, z) \cdot Y(B, z), \quad n < 0. \quad (91)$$

Since  $Y(B, z)|0\rangle = e^{zT}B$ , by Lemma 3.1.3 we find that

$$B_{(-n-1)}|0\rangle = \frac{1}{n!} T^n B.$$

Applying Corollary 3.1.6 several times, we find that

$$Y(B_{(-n-1)}|0\rangle, z) = \frac{1}{n!} \partial_z^n Y(B, z). \quad (92)$$

Using formulas (91) and (92), we obtain the following corollary by induction on  $k$  (recalling that our convention is that the normal ordering is nested from right to left):

**Corollary 11.10 (3.3.5).** *For any  $A^1, \dots, A^k \in V$  and  $n_1, \dots, n_k < 0$ , we have*

$$Y(A_{(n_1)}^1 \cdots A_{(n_k)}^k | 0\rangle, z) = \frac{1}{(-n_1 - 1)!} \cdots \frac{1}{(-n_k - 1)!} : \partial_z^{-n_1-1} Y(A^1, z) \cdots \partial_z^{-n_k-1} Y(A^k, z) : .$$

*This gives us formula (2.3.8) under the assumptions of Theorem 2.3.11.*

Now we compare the coefficients in front of  $(z - w)^k$ ,  $k < 0$ , in formulas (3.3.1) and (3.3.7). We find that

$$\gamma_j(w) = Y(A_{(j)} \cdot B, w), \quad j \geq 0,$$

and so formula (3.3.6) can be rewritten as

$$Y(A, z)Y(B, w) = \sum_{n \geq 0} \frac{Y(A_{(n)} \cdot B, w)}{(z - w)^{n+1}} + : Y(A, z)Y(B, w) : . \quad (93)$$

Note that unlike formula (3.3.1), this identity makes sense in  $\text{End } V[[z^{-1}, w^{-1}]]$  if we expand  $(z - w)^{-1}$  in positive powers of  $w/z$ .

Now Proposition 3.3.1 implies the following commutation relations:

$$[Y(A, z), Y(B, w)] = \sum_{n \geq 0} \frac{1}{n!} Y(A_{(n)} \cdot B, w) \partial_w^n \delta(z - w). \quad (94)$$

Expanding both sides of (94) as formal power series and using the equality

$$\frac{1}{n!} \partial_w^n \delta(z - w) = \sum_{m \in \mathbb{Z}} \binom{m}{n} z^{-m-1} w^{m-n},$$

we obtain the following identity for the commutators of Fourier coefficients of arbitrary vertex operators:

$$[A_{(m)}, B_{(k)}] = \sum_{n \geq 0} \binom{m}{n} (A_{(n)} \cdot B)_{(m+k-n)}. \quad (95)$$

Here, by definition, for any  $m \in \mathbb{Z}$ ,

$$\binom{m}{n} = \frac{m(m-1) \cdots (m-n+1)}{n!}, \quad n \in \mathbb{Z}_{\geq 0}; \quad \binom{m}{0} = 1.$$

In particular, we see that only the terms in the OPE (3.3.1) which are singular on the diagonal  $z = w$  contribute to the commutator. Thus, we will often write formula (93) in the form

$$Y(A, z)Y(B, w) = \sum_{n \geq 0} \frac{Y(A_{(n)} \cdot B, w)}{(z-w)^{-n-1}} + \text{reg.},$$

emphasizing the singular terms in the OPE.

Formula (95) also shows that the commutator of Fourier coefficients of two fields is a linear combination of Fourier coefficients of other fields (namely, the ones corresponding to the vectors  $A_{(n)} \cdot B$ ). Therefore we obtain the following result.

**Proposition 11.11 (3.3.7).** *Let  $V$  be a vertex algebra. Then the span in  $\text{End } V$  of all Fourier coefficients of all vertex operators in  $V$  is a Lie algebra.*

An easy corollary of formula (95) is the construction of vertex algebra automorphisms using residues of fields.

**Corollary 11.12 (3.3.8).** *For any  $A \in V$ , the residue*

$$A_{(0)} = \int Y(A, z) dz$$

*of  $Y(A, z)$  is an infinitesimal automorphism of the vertex algebra  $V$ , that is,*

$$[A_{(0)}, Y(B, w)] = Y(A_{(0)} \cdot B, w).$$

Examples of residues of fields are the operator  $T = L_{-1} = \int T(z) dz$  in a conformal vertex algebra, and the operators

$$J_0^a = \int J^a(z) dz$$

in the vertex algebra  $V_k(\mathfrak{g})$ , giving the action of  $\mathfrak{g}$  on  $V_k(\mathfrak{g})$ . An important class of residue operators consists of the differentials in the BRST construction (see §5.7.3).

## Meetings

### Questions (Sept 18)

- A. I've been thinking about the hierarchy of objects:  $L\mathfrak{g} \rightarrow \text{central extensions} \rightarrow \text{full affine Kac-Moody with } d$ . 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  $[d, x \otimes z^n] = nx \otimes z^n$ . 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  $\text{Gr}_G$ . 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  $\text{Gr}_G$  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 (Sept 17)

Professor Borcherds 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.

## Teleman remarks

It turns out that the numerator and denominator in the Kac character formula for representations of  $LSU(2)$  are theta functions. In particular, the Jacobi triple product formula is the Kac character formula for the trivial representation of  $LSU(2)$ , in particular the character should be 1, the Kac numerator is the sum and the Kac denominator is the product.

$$\prod_{m=1}^{\infty} (1 - x^{2m}) (1 + x^{2m-1} y^2) \left( 1 + \frac{x^{2m-1}}{y^2} \right) = \sum_{n=-\infty}^{\infty} x^{n^2} y^{2n}.$$

So theta functions can be interpreted as particular sections of line bundles on the moduli stack of elliptic curves. Where do elliptic curves come from?

One starts by thinking of 2D quantum field theory on a cylinder  $S^1 \times [0, T]$ . The "time" direction is the interval, and the circular direction is the "space." The partition function of the theory on this cylinder is the matrix element of the evolution operator: it's like

$$Z_{\text{cyl}} = \langle \psi_{\text{out}} | e^{-TH} | \psi_{\text{in}} \rangle$$

where  $H$  is the Hamiltonian. So a cylinder corresponds to propagation in time.

If you glue the two circular ends together, you get a torus (an elliptic curve). In quantum field theory, "gluing the ends" means instead of fixing in/out states, you sum over them. That is exactly taking a trace:

$$Z_{\text{torus}} = \text{Tr}(e^{-TH} \cdot (\text{other insertions})).$$

So the torus partition function is a trace over Hilbert space. That's why the Kac character formula, which is a trace, naturally gives functions on elliptic curves (theta functions). If your QFT has a global symmetry by  $G$  (say  $SU(2)$ ), then you can act pointwise by  $G$  on the space circle.

Then you can upgrade this to a loop group symmetry: maps  $S^1 \rightarrow G$ , acting on the circle direction. In physics language: the current algebra (affine Lie algebra) is the symmetry algebra you get when you "localize" the group symmetry along the circle. This is why  $LSU(2)$  naturally acts on the Hilbert space of the theory on the circle.

## Questions (Oct 9)

1. I roughly understand that the product in the Weyl denominator formula can be Poisson re-summed to give a sum over the root lattice, and this is giving the relationship between weight multiplicities and theta functions. Kac describes one method of calculating weight multiplicities using Peterson's formula, and I was wondering if there has been any success in using this formula to better understand generalized Kac-Moody algebras and their denominator formulas?



2. In your papers you mention that many automorphic product forms correspond to denominator formulas of generalized Kac-Moody algebras, and that one open problem is to construct these algebras explicitly — beyond their generators and relations. As an example, you wrote the fake monster Lie algebra is the Lie algebra of chiral strings on a 26-dimensional torus. Could you say more about how this arises? For example, how exactly do the physical states or vertex operators correspond to roots and multiplicities in the BKM presentation?”

You also wrote that many automorphic forms that are modular products can be interpreted as the denominator formulas of generalized Kac-Moody algebras. How does one get a hold of this correspondence in this particular example? Are there other examples where one can see this correspondence explicitly?

3. Vertex algebras are described as “meromorphic” generalizations of commutative algebras with an action of the formal additive group. From this point of view, how do the singular terms in the operator product expansion encode root multiplicities (say in the motivating example of Frenkel’s construction of the Leech lattice Lie algebra)?
4. How does your original idea of vertex algebras as singular commutative rings on the formal additive group relate to the later coordinate-free, geometric definitions (e.g. chiral algebras or factorization algebras)?
5. You describe a vertex algebra as a kind of “commutative ring in 1 dimension,” analogous to how commutative algebras describe 0-dimensional field theories. What would the algebraic structure corresponding to a 2- or higher-dimensional analogue look like? Would the right framework for that now be factorization algebras or  $E_n$ -algebras, or is that a different direction from what you had in mind originally?
6. Is there a character formula for representations of vertex algebras, analogous to the Weyl or Kac character formulas for Lie algebras and affine Kac-Moody algebras? I came across Zhu’s modular invariance theorem for vertex algebras, which roughly says that characters of certain representations of vertex operator algebras are modular functions.

## Answers (Oct 9)

1. He gave me an exercise to convert chapter 14 of Kac’s book into vertex algebra language.
2. He said that historically people have used automorphic forms to read off simple roots of generalized Kac-Moody algebras and their multiplicities. But the problem of classification remains open and in progress.
3. He said that simplest example of generalized Kac Moody algebra is the fake monster Lie algebra in rank 24, and that the ones in lower rank are in fact more complicated. He said that people have plenty of examples in rank 2 and some in rank 3 but that there are significantly

more examples in even rank. Also the examples in rank 24 all come from automorphisms of the Leech lattice.

4. I think I will read Kac's book on vertex algebras in order to continue this reading course. I think that I can also ask him about this exercise in chapter 11 of Kac.
5. He said that the issue with reading anything written by Beilinson is that Beilinson is too smart.
6. He also made this remark that he hasn't thought about this stuff for over 30 years. He said if I want to do vertex algebras and algebraic curves, I should talk to Frenkel who would be more than happy to talk about his book with me. I didn't tell him that talking to Frenkel is most likely impossible.

## October 16

1. I told him that within the structure theory of vertex algebras, the simplest examples I had encountered were the Fock representation coming from the Heisenberg algebra, the vacuum representations of affine Kac-Moody algebras, and the ones induced from the Virasoro algebra. He said that these are not the simplest examples. The simplest examples are the lattice vertex algebras, which are constructed from even lattices. He recommended that I unwind everything like the OPEs of the vertex operators in the case of a 1-dimensional lattice.
2. He said that the vertex algebra coming from a free abelian monoid is the simplest example, in the sense that it is as free of a vertex algebra as possible. One cannot get free vertex algebras because that would require poles of infinite order, but if you ask for "free vertex algebra with poles of order at most  $N$ " then you get the ones coming from free abelian monoids.
3. He said that one needs a very good example in order to understand the structure theory of vertex algebras. There are many ways one could imagine generalizing the notion of a vertex algebra, such as from one dimensional formal groups to higher dimensional formal groups. But one needs a good example to test the theory on, and in some since there might not be any more room because the Monster group is the largest sporadic simple group, and vertex algebras were introduced to understand the fake Monster Lie algebra.
4. He said its not recommended to study vertex algebras for their own sake, but rather to study them in order to understand to solve problems. I said that my motivation was to understand the relationship between vertex algebras and algebraic curves, and I asked if he thinks that there is anything promising in this direction. He said yes, because Frenkel is a smart guy, but that he didn't know anything about it personally.
5. I asked him about his application of these results to the moduli space of Enriques surfaces. He said it works because there is a crazy coincidence that some lattice associated to the

moduli space happens to coincide with the Leech lattice, which is the root lattice of the fake Monster Lie algebra. The question in his mind is given an isomorphism class of Enriques surface, how to evaluate the corresponding automorphic form there (interpreted as the denominator of some character formula for the lattice). He gave me this beat up paper and remarked that the original proofs in the paper are quite clumsy and displeased algebraic geometers. In particular, they rewrite his argument about asymptotics of Bessel functions using Grothendieck Riemann Roch.

## 12 Appendix 1: Theta functions and Riemann surfaces

Recently in Professor Teleman's class on complex manifolds, we have been studying the classical theory of elliptic functions in order to construct nonconstant meromorphic functions on a genus 1 Riemann surface  $\mathbb{C}/L$ , where  $L = \mathbb{Z}\omega_1 \oplus \mathbb{Z}\omega_2$  is a full rank lattice in  $\mathbb{C}$ , that is  $\omega_1/\omega_2 \notin \mathbb{R}$ .

The classical story begins with the Weierstrass  $\wp$ -function, defined by

$$\wp(z; L) = \frac{1}{z^2} + \sum_{\omega \in L \setminus \{0\}} \left( \frac{1}{(z - \omega)^2} - \frac{1}{\omega^2} \right)$$

which has the properties that it is an  $L$ -periodic meromorphic function on  $\mathbb{C}$  with double poles at the lattice points, and that it satisfies the differential equation

$$(\wp'(z))^2 = 4\wp(z)^3 - g_2\wp(z) - g_3 = 4(z - e_1)(z - e_2)(z - e_3)$$

where  $g_2, g_3$  are constants depending on  $L$ , given explicitly by

$$g_2 = 60 \sum_{\omega \in L \setminus \{0\}} \frac{1}{\omega^4}$$

$$g_3 = 140 \sum_{\omega \in L \setminus \{0\}} \frac{1}{\omega^6}$$

and  $e_i$  are the values of  $\wp$  at the half-lattice points  $\omega_1/2, \omega_2/2, (\omega_1 + \omega_2)/2$ . The  $e_i$  are distinct as we will show in Prop 12.2. The convergence is uniform on any compact subset  $K \subset \mathbb{C}$ , once the terms with poles in  $K$  are set aside.

Uniform convergence implies that the series can be differentiated term-by-term, so we get a formula for  $\wp'(z)$  given by

$$\wp'(z) = -2 \sum_{\omega \in L} \frac{1}{(z - \omega)^3}$$

is an doubly periodic meromorphic function with triple poles at the lattice points. Moreover, one can see directly from the series expansion that  $\wp$  is even and  $\wp'$  is odd.

The oddness implies that  $\wp'(z)$  vanishes at the half-lattice points. Moreover, one can check that these are simple zeros of  $\wp'$ , and moreover the only zeros of  $\wp'$  modulo  $L$ . Thus  $\wp'$  has only poles at lattice points, each of order 3. In a fundamental parallelogram there is exactly one pole (mod  $L$ ), of total multiplicity 3. This implies the following proposition.

**Proposition 12.1.**  $\wp(z)$  and  $\wp'(z)$  define holomorphic maps  $\mathbb{C}/L \rightarrow \mathbb{P}^1$  of degree 2 and 3 respectively.

We conclude that each of the half-lattice points must be a simple zero of  $\wp'$  and moreover that these are all of the zeros, because any meromorphic function has divisor of degree 0.

**Proposition 12.2** (Properties of the  $\wp$ -map).

- (i) The numbers  $e_1, e_2, e_3$  are all distinct.
- (ii) For any  $a \in \mathbb{C}$  with  $a \neq e_1, e_2, e_3$ , the equation  $\wp(u) = a$  has two simple roots in a fundamental period parallelogram. For the three exceptional values  $a = e_i$ , it has a single double root.

*Proof.*

- (ii) General theory of meromorphic functions on a torus shows that we either have two simple roots or one double root. Since a double root corresponds to a zero of the derivative  $\wp'$ , the claim follows. Note that the two simple roots always differ by a sign modulo  $L$ , by the parity of  $\wp$ .
- (i) Suppose, for contradiction, that  $e_1 = e_2$ . Then  $\wp(u) = e_1$  would have a double root at  $\frac{\omega_1}{2}$  and another double root at  $\frac{\omega_2}{2}$ . This would give too many roots (multiplicity 4 in a fundamental parallelogram), contradicting the fact that  $\wp$  is a double covering of  $\mathbb{P}^1$ . Hence the  $e_i$  are distinct.

□

**Remark 12.3.** Kac writes that this quadratic term which appears in the definition of  $t_\alpha$  34 "explains" the appearance of theta functions in the theory of affine algebras. This is because when you compute the characters of highest-weight representations of affine Kac-Moody algebras, you sum over the affine Weyl group:

$$\chi(\lambda) = \sum_{w \in W} \det(w) e^{w(\lambda + \rho) - \rho}$$

and theta functions arise precisely when you sum exponentials of the form

$$\Theta(\tau, z) = \sum_{\alpha \in \text{lattice}} \exp\left(-\frac{1}{2}|\alpha|^2\tau + \langle \alpha, z \rangle\right).$$

Continuing with the discussion of theta functions, we have the following theorem about genus 1 Riemann surfaces.

**Theorem 12.4.** *Let  $\theta_1, \dots, \theta_4$  be the four Jacobi theta functions. Then there is a map*

$$E/L \rightarrow \mathbb{CP}^3, \quad z \mapsto [\theta_1(z, \tau) : \theta_2(z, \tau) : \theta_3(z, \tau) : \theta_4(z, \tau)]$$

*which is a smooth embedding of the complex torus  $E = \mathbb{C}/L$  into projective space. It is a degree 4 map and its image is the intersection of two quadrics.*

When we were proving this theorem, we wrote down the first theta function  $\theta_1$  and defined the others by half period shifts of  $\theta_1$ . Is there a reason why we didn't write down more? Borchers made a remark last time that there are  $2^{2g}$  theta divisors on a genus  $g$  Riemann surface.

For example, there's a classical theorem that the vector space of modular forms of weight  $k$  for  $\mathrm{SL}_2(\mathbb{Z})$  has dimension  $k/12$ . Should I be thinking of some result along these lines, as to why these are the only four we care about in the genus 1 case?

So the theta functions are sections of the pullback of  $\mathcal{O}(1)$  to  $E$ . Is there a way to see the affine characters as coming from elliptic curves in some natural way?

## 13 Appendix 2: Lie algebra cohomology

**Lemma 13.1 (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  $c^{\mathrm{av}} \in C^p$  with  $dc^{\mathrm{av}} = 0$  and  $[c^{\mathrm{av}}] = [c]$ . 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  $\mu$  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  $\mu$ . 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  $c^{\text{av}}$  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.  $\square$

## 14 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](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.
5. Frenkel, E., and Ben-Zvi, D. **Vertex Algebras and Algebraic Curves**. Second edition. Mathematical Surveys and Monographs, 88. American Mathematical Society, Providence, RI; 2004.
6. McGlade, F., Ram, A., and Yang, Y. **Positive level, negative level and level zero**.