

### 3. KAC-MOODY ALGEBRAS

We are now going to switch to a completely different topic and study representations of some Lie algebras. In this chapter,  $k$  will denote an algebraically closed field of characteristic 0. It may as well be the complex numbers so since I'll get it muddled up let us just say  $k = \mathbb{C}$  throughout! My basic reference for this material is the book "Infinite dimensional Lie algebras" by Victor Kac.

It is important to appreciate that I am only telling you HALF the story and showing you HALF the important examples. This is because I am restricting attention to the simply-laced Dynkin diagrams (i.e. symmetric Cartan matrices). I think this is a reasonable thing to do in a course like this – you get a much better overview of the big picture without worrying about all the technicalities that arise when dealing with symmetrizable but not symmetric Cartan matrices.

But remember there are also Dynkin diagrams  $B_n, C_n, G_2, F_4$  out there that have two different root lengths.

**3.1. Lie algebras.** I need to spend a week or so reviewing some basic definitions about Lie algebras. This will be boring for many of you – sorry. I dont think I will type this stuff in. I'll define:

- Lie algebras by axioms;
- Give examples coming from associative algebras by commutator especially  $\mathfrak{sl}_n(\mathbb{C})$ ;
- Give examples coming from derivations of arbitrary algebras;
- Discuss ideals and simple Lie algebras;
- Prove that  $\mathfrak{sl}_n$  is a simple Lie algebra;
- Define the other classical simple Lie algebras;
- Define modules over a Lie algebras;
- Define the universal enveloping algebra and explain what it has to do with modules;
- Write down the PBW theorem.
- Discuss weight space decomposition in modules over commutative Lie algebras. The main thing we'll need that is not so obvious is the lemma below.
- Explain the finite dimensional representations of  $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ . Let me just recall the final result here since it is needed in the Exercises below. The finite dimensional irreducible representations are the modules  $L(n)$  for  $n = 0, 1, 2, \dots$ , where  $L(n)$  is the unique irreducible module generated by a vector  $v$  such that  $ev = 0, hv = nv$ . The vectors  $v, fv, f^2v, \dots, f^nv$  form a basis for  $L(n)$ , so it is of dimension  $(n + 1)$ .

**Lemma 3.1.** *Let  $\mathfrak{h}$  be an abelian Lie algebra and  $V$  a diagonalizable  $\mathfrak{h}$ -module, i.e.*

$$V = \bigoplus_{\lambda \in \mathfrak{h}^*} V_\lambda$$

where

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

Then any submodule  $U$  of  $V$  is also diagonalizable.

*Proof.* Any  $v \in V$  can be written in the form  $v = \sum_{j=1}^m v_j$  where  $v_j \in V_{\lambda_j}$  for distinct  $\lambda_1, \dots, \lambda_m \in \mathfrak{h}^*$ . Pick  $h \in \mathfrak{h}$  such that  $\lambda_j(h)$  are all distinct. Then for  $v \in U$ , we have that

$$h^k(v) = \sum_{j=1}^m \lambda_j(h)^k v_j \in U$$

for  $k = 0, 1, \dots, m-1$ . This is a system of linear equations with non-degenerate matrix (its a Vandermonde!). Hence we can solve the system in the vector space  $U$  to deduce that all  $v_j$  belong to  $U$ .  $\square$

**Exercise 7.** In the next section, we are going to define a Lie algebra by generators and relations. Of course this means the quotient of the free Lie algebra by the ideal generated by the relations. The purpose of this exercise is to construct the free Lie algebra generated by a vector space  $V$  (i.e. the vector space with basis given by the generators you have in mind).

- Write down the definition of the free Lie algebra  $F(V)$  on the vector space  $V$  by universal property.
- Here is a construction of  $F(V)$ . Let  $T(V)$  be the tensor algebra on the vector space  $V$ , so  $T(V)$  is universal amongst all associative algebras generated by the vector space  $V$ . Viewing  $T(V)$  instead as a Lie algebra, let  $F(V)$  be the Lie subalgebra of  $T(V)$  generated by the subspace  $V$ , i.e.  $F(V)$  is the intersection of all the Lie subalgebras of  $T(V)$  containing  $V$ . Prove that  $F(V)$  together with the inclusion  $V \hookrightarrow F(V)$  IS the free Lie algebra on vector space  $V$ .
- Now prove that the universal enveloping algebra  $U(F(V))$  is isomorphic to the tensor algebra  $T(V)$ .

The remaining exercises are to do with the representation theory of the Lie algebra  $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ , on standard basis  $e, h, f$  with relations  $[e, f] = h$ ,  $[h, e] = 2e$ ,  $[h, f] = -2f$ . Recall the finite dimensional irreducible  $L(n)$  of dimension  $n+1$  described in detail above.

**Exercise 8.** If  $\mathfrak{g}$  is any Lie algebra and  $V, W$  are  $\mathfrak{g}$ -modules, there is a natural way to make the tensor product  $V \otimes W$  into a  $\mathfrak{g}$ -module:  $x(v \otimes w) := (xv) \otimes w + v \otimes (xw)$ .

- Let  $V$  be a  $\mathfrak{g}$ -module. Recall the symmetric algebra  $S(V)$  is the quotient of  $T(V)$  by the ideal generated by  $\{x \otimes y - y \otimes x \mid x, y \in V\}$ .

Verify that this ideal is invariant under the action of  $\mathfrak{g}$ , hence  $S(V)$  is a  $\mathfrak{g}$ -module. (So is  $\bigwedge(V)$ ).

- (b) In the case  $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ , let  $V$  be the natural 2 dimensional module on standard basis  $v_1, v_2$ . Prove that  $S^n(V) \cong L(n)$ , the irreducible module of dimension  $(n + 1)$ .

**Exercise 9.** If  $\mathfrak{g}$  is any Lie algebra and  $V$  is a finite dimensional  $\mathfrak{g}$ -module, there is a natural way to make the dual space  $V^*$  into a  $\mathfrak{g}$ -module:  $(xf)(v) = -f(xv)$  for  $f \in V^*$ ,  $x \in \mathfrak{g}$  and  $v \in V$ . There is also always a trivial  $\mathfrak{g}$ -module, namely the one dimensional vector space  $\mathbb{C}$  on which each  $x \in \mathfrak{g}$  acts as zero.

- (a) Suppose that  $V$  is a finite dimensional  $\mathfrak{g}$ -module. Prove that  $V \cong V^*$  as  $\mathfrak{g}$ -modules if and only if there is a non-degenerate bilinear form  $(\cdot, \cdot)$  on  $V$  which is *invariant* in the sense that  $(xv, w) + (v, xw) = 0$  for all  $v, w \in V$  and  $x \in \mathfrak{g}$ .
- (b) Suppose  $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ . Prove that  $L(n)^* \cong L(n)$  as  $\mathfrak{g}$ -modules.
- (c) In particular consider the case that  $V = L(2)$  is the natural two dimensional representation of  $\mathfrak{sl}_2(\mathbb{C})$ . Write down explicitly a non-degenerate invariant bilinear form on  $V$ , hence deduce that  $\mathfrak{sl}_2(\mathbb{C}) \cong \mathfrak{sp}_2(\mathbb{C})$ .
- (d) Do the same if  $V = L(3)$  and hence show that  $\mathfrak{sl}_2(\mathbb{C}) \cong \mathfrak{so}_3(\mathbb{C})$ .

Note in fancy language, the observations made in the preceding two exercises show that  $U(\mathfrak{g})$  is a Hopf algebra. The comultiplication  $\Delta : U(\mathfrak{g}) \rightarrow U(\mathfrak{g}) \otimes U(\mathfrak{g})$  is defined by  $\Delta(x) = x \otimes 1 + 1 \otimes x$  for each  $x \in \mathfrak{g}$  and then extended to  $U(\mathfrak{g})$  by the universal property. The counit is the map  $x \mapsto 0$  and the antipode is the map  $x \mapsto -x$ , all written for  $x \in \mathfrak{g}$ .

**Exercise 10.** Let  $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ , on standard basis  $e, h, f$  with relations  $[e, f] = h, [h, e] = 2e, [h, f] = -2f$ .

- (a) Prove that the element  $c := fe + \frac{1}{4}h(h + 2)$  belongs to the center of the universal enveloping algebra  $U(\mathfrak{g})$ . (Hint: you need to show it commutes with each of the generators  $e, h, f$  of  $U(\mathfrak{g})$ . It is useful to note that  $[x, yz] = [xy]z + y[xz]$  for  $x \in \mathfrak{g}$  and  $y, z \in U(\mathfrak{g})$ , i.e.  $\text{ad } x$  acts on  $U(\mathfrak{g})$  as a derivation.)
- (b) Show that  $c$  acts on the irreducible module  $L(n)$  as the scalar  $\frac{1}{4}n(n + 2)$ . Deduce that any short exact sequence

$$0 \longrightarrow L(n) \longrightarrow V \longrightarrow L(m) \longrightarrow 0$$

of  $\mathfrak{g}$ -modules splits for  $m \neq n$ .

- (c) Prove that the short exact sequence also splits in the case  $m = n$  (hint: think about the  $h$ -weight space of eigenvalue  $n$  first).
- (d) Deduce that any finite dimensional  $\mathfrak{g}$ -module is completely reducible.

**3.2. Kac-Moody algebras.** Now we are ready to define the Kac-Moody Lie algebras. Fix a graph  $\Gamma$  with no loops, like in §2.10. Remember the vertices are labelled  $1, \dots, n$  and there are  $n_{i,j} = n_{j,i}$  edges from  $i$  to  $j$ . The *root lattice* is

$$R = \mathbb{Z}\epsilon_1 \oplus \dots \oplus \mathbb{Z}\epsilon_n$$

where  $\epsilon_1, \dots, \epsilon_n$  are the simple roots. There is a bilinear form  $(\cdot, \cdot)$  on  $R$  defined by  $(\epsilon_i, \epsilon_i) = 2, (\epsilon_i, \epsilon_j) = -n_{i,j}$  for  $i \neq j$ . We have also defined the Weyl group  $W < Isom(R)$  generated by the simple reflections  $s_1, \dots, s_n$  in the hyperplanes orthogonal to the simple roots.

We are now going to associate a Lie algebra  $\mathfrak{g} = \mathfrak{g}(\Gamma)$ . Using this we will define the roots  $\Delta$  in general – and this will be the same as the set of roots  $\Delta$  already defined in the Dynkin and Euclidean cases. To start with, let

$$\mathfrak{h}' = \mathbb{C} \otimes_{\mathbb{Z}} R.$$

This is a  $\mathbb{C}$ -vector space on basis  $\epsilon_1, \dots, \epsilon_n$  and the bilinear form  $(\cdot, \cdot)$  extends to  $\mathfrak{h}'$ . Let  $\mathfrak{c}$  be the radical of the bilinear form on  $\mathfrak{h}'$ . Let

$$\mathfrak{h} = \mathfrak{h}' \oplus \mathfrak{c}^*$$

where  $\mathfrak{c}^*$  is the dual space to  $\mathfrak{c}$ . We wish to extend the bilinear form  $(\cdot, \cdot)$  on  $\mathfrak{h}'$  to a non-degenerate symmetric bilinear form on all of  $\mathfrak{h}$ . To do this we need to make a choice, but all choices lead to isomorphic spaces so it doesn't matter: pick a complement to  $\mathfrak{c}$  in  $\mathfrak{h}'$  and define  $(f, g) = 0$  for  $f, g \in \mathfrak{c}^*$ , that  $(f, c) = f(c)$  for  $f \in \mathfrak{c}^*$  and  $c \in \mathfrak{c}$ , and that  $(f, d) = 0$  for  $f \in \mathfrak{c}^*$  and  $d$  in the chosen complement to  $\mathfrak{c}$  in  $\mathfrak{h}'$ .

So now we have constructed a  $\mathbb{C}$ -vector space  $\mathfrak{h}$  equipped with a non-degenerate symmetric bilinear form  $(\cdot, \cdot)$ . The subspace  $\mathfrak{h}'$  has basis  $\epsilon_1, \dots, \epsilon_n$ ,  $(\epsilon_i, \epsilon_i) = 2$  and  $(\epsilon_i, \epsilon_j) = -n_{i,j}$ . Finally the radical of the restriction of  $(\cdot, \cdot)$  to  $\mathfrak{h}'$  is denoted  $\mathfrak{c}$ . Important note: since we have a non-degenerate bilinear form on  $\mathfrak{h}$  we can identify  $\mathfrak{h}$  canonically with the dual vector space  $\mathfrak{h}^*$ . Thus, any linear functional on  $\mathfrak{h}$  can be realized as the function  $(\alpha, \cdot)$  for some element  $\alpha \in \mathfrak{h}$ . We'll use this always from now on.

OKAY, enough said about building  $\mathfrak{h}$ . Returning to the general  $\Gamma$  with vertices  $1, \dots, n$ , let  $\tilde{\mathfrak{g}} = \tilde{\mathfrak{g}}(\Gamma)$  be the Lie algebra with generators  $\mathfrak{h}, e_i, f_i$  ( $i = 1, \dots, n$ ) subject only to the relations

$$\begin{aligned} [e_i, f_j] &= \delta_{i,j} \epsilon_i, \\ [h, h'] &= 0, \\ [h, e_i] &= (\epsilon_i, h) e_i, \\ [h, f_i] &= -(\epsilon_i, h) f_i \end{aligned}$$

for  $i, j = 1, \dots, n$  and  $h, h' \in \mathfrak{h}$ . Let  $\tilde{\mathfrak{n}}^+$  (resp.  $\tilde{\mathfrak{n}}^-$ ) be the subalgebra of  $\tilde{\mathfrak{g}}$  generated by  $e_1, \dots, e_n$  (resp.  $f_1, \dots, f_n$ ).

**Example 3.2.** Take the trivial case when  $\Gamma$  is just one vertex, no edges. Then  $\tilde{\mathfrak{g}}$  has generators  $h, e, f$  subject to the relations  $[h, h] = 0, [e, f] = h, [h, e] = 2e, [h, f] = -2f$ . This is of course the Lie algebra  $\mathfrak{sl}_2(\mathbb{C})$  – you should know what  $2 \times 2$  traceless matrices  $e, h, f$  correspond to!

**Theorem 3.3.** (a)  $\tilde{\mathfrak{g}} = \tilde{\mathfrak{n}}^- \oplus \mathfrak{h} \oplus \tilde{\mathfrak{n}}^+$  (vector space direct sum).

(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  $e_i \mapsto f_i, f_i \mapsto e_i, h \mapsto h$  extends uniquely to an antiautomorphism  $\tilde{\omega}$  of  $\tilde{\mathfrak{g}}$  of order 2.

(d) With respect to  $\mathfrak{h}$ , one has the root space decomposition:

$$\tilde{\mathfrak{g}} = \bigoplus_{0 < \alpha \in R} \tilde{\mathfrak{g}}_{-\alpha} \oplus \mathfrak{h} \oplus \bigoplus_{0 < \alpha \in R} \tilde{\mathfrak{g}}_{\alpha},$$

where  $\tilde{\mathfrak{g}}_{\alpha} = \{x \in \tilde{\mathfrak{g}} \mid [h, x] = (\alpha, h)x \text{ for all } h \in \mathfrak{h}\}$  (a finite dimensional vector space).

(e) There exists a unique maximal ideal  $\mathfrak{r}$  in  $\mathfrak{g}$  that intersects  $\mathfrak{h}$  trivially. Moreover,  $\mathfrak{r} = (\mathfrak{r} \cap \tilde{\mathfrak{n}}^-) \oplus (\mathfrak{r} \cap \tilde{\mathfrak{n}}^+)$  (direct sum of ideals).

*Proof.* Let  $V$  be an  $n$  dimensional vector space with basis  $v_1, \dots, v_n$ . Let  $\lambda \in \mathfrak{h}$ . Define an action of the generators of  $\tilde{\mathfrak{g}}$  on  $T(V)$  as follows:

- ( $\alpha$ )  $f_i(a) = v_i \otimes a$  for  $a \in T(V)$ ;  
 ( $\beta$ )  $h(1) = (\lambda, h)1$  and then inductively,

$$h(v_j \otimes a) = -(\epsilon_j, h)v_j \otimes a + v_j \otimes h(a)$$

for  $a \in T^{d-1}V$  and  $j = 1, \dots, n$ ;

- ( $\gamma$ )  $e_i(1) = 0$  and then inductively,

$$e_i(v_j \otimes a) = \delta_{i,j}\epsilon_i(a) + v_j \otimes e_i(a)$$

for  $s \in T^{d-1}V$ .

Now we want to check the relations to see that this is a well-defined representation of  $\tilde{\mathfrak{g}}$  on  $T(V)$ .

The second relation  $[h, h'] = 0$  is obvious since  $\mathfrak{h}$  acts diagonally according to ( $\alpha$ ).

For the first relation  $[e_i, f_j] = \delta_{i,j}\epsilon_i$ , we have

$$\begin{aligned} (e_i f_j - f_j e_i)(a) &= e_i(v_j \otimes a) - v_j \otimes e_i(a) \\ &= \delta_{i,j}\epsilon_i(a) + v_j \otimes e_i(a) - v_j \otimes e_i(a) = \delta_{i,j}\epsilon_i(a) \end{aligned}$$

as required. For the fourth relation  $[h, f_i] = -(\epsilon_i, h)f_i$ , we have

$$\begin{aligned} (h f_j - f_j h)(a) &= h(v_j \otimes a) - v_j \otimes h(a) \\ &= -(\epsilon_j, h)v_j \otimes a + v_j \otimes h(a) - v_j \otimes h(a) \\ &= -(\epsilon_j, h)f_j(a) \end{aligned}$$

as required. Finally for the third relation  $[h, e_i] = (\epsilon_i, h)e_i$  proceed by induction on  $d$ . For  $d = 0$  it is clear. For  $d > 0$  take  $a = v_k \otimes a_1$  with  $a_1 \in T^{d-1}(V)$ . We have

$$\begin{aligned} (h e_j - e_j h)(v_k \otimes a_1) &= h(\delta_{j,k}\epsilon_j(a_1)) + h(v_k \otimes e_j(a_1)) \\ &\quad - e_j(-(\epsilon_k, h)(v_k \otimes a_1) + v_k \otimes h(a_1)) \\ &= \delta_{j,k}\epsilon_j((h(a_1))) - (\epsilon_k, h)v_k \otimes e_j(a_1) + v_k \otimes h e_j(a_1) \\ &\quad + (\epsilon_k, h)\delta_{j,k}\epsilon_j(a_1) \\ &\quad + (\epsilon_k, h)v_k \otimes e_j(a_1) - \delta_{j,k}\epsilon_j h(a_1) - v_k \otimes e_j h(a_1) \\ &= (\epsilon_j, h)\delta_{j,k}\epsilon_j(a_1) + v_k \otimes (h e_j - e_j h)(a_1). \end{aligned}$$

Now apply the inductive assumption to the second summand to complete the proof.

Now we prove the theorem. First using the first relation it is easy to see using the relations and induction on  $s$  that a product of  $s$  elements from the set  $\{e_i, f_i\}_{i=1, \dots, n} \cup \mathfrak{h}$  lies in  $\tilde{\mathfrak{n}}^- + \mathfrak{h} + \tilde{\mathfrak{n}}^+$ . Now let  $u = n^- + h + n^+ = 0$ . Then in the representation  $T(V)$  we have that  $u(1) = n^-(1) + (\lambda, h) = 0$ . It follows that  $(\lambda, h) = 0$  for every  $\lambda \in \mathfrak{h}$ , hence  $h = 0$  since the form is non-degenerate. Furthermore, using the map  $f_i \mapsto v_i$ , we see that the tensor algebra  $T(V)$  is an enveloping algebra of the Lie algebra  $\tilde{\mathfrak{n}}^-$ . Hence since  $T(V)$  is free  $T(V)$  is automatically the universal enveloping algebra of  $\tilde{\mathfrak{n}}^-$ . In particular, the map  $n^- \mapsto n^-(1)$  is an embedding. So we get that  $n^- = 0$  too and (a) is proven. The PBW theorem implies moreover that  $\tilde{\mathfrak{n}}^-$  is freely generated by  $f_1, \dots, f_n$  and – since (c) is obvious – we get the same thing for  $\tilde{\mathfrak{n}}^+$ . This proves (b).

Using the last two relations and considering ordered monomials in the  $e_i$  resp.  $f_i$  we have that

$$\tilde{\mathfrak{n}}^\pm = \bigoplus_{0 < \alpha \in R} \tilde{\mathfrak{g}}_{\pm\alpha}.$$

We also have the obvious estimate (how many sequences  $(\epsilon_{i_1}, \dots, \epsilon_{i_N})$  are there summing to  $\alpha$ )

$$\dim \tilde{\mathfrak{g}}_\alpha \leq n^{|\text{ht}(\alpha)|},$$

giving the finite dimensionality of  $\tilde{\mathfrak{g}}_\alpha$ . These together with (a) prove (d).

Finally, for (e), we need Lemma 3.1. This shows that for any ideal  $\mathfrak{r}$  of  $\tilde{\mathfrak{g}}$ ,

$$\mathfrak{r} = \bigoplus_{\alpha \in R} (\tilde{\mathfrak{g}}_\alpha \cap \mathfrak{r}).$$

It follows at once that the sum of ideals that intersects  $\mathfrak{h}$  trivially itself intersects  $\mathfrak{h}$  trivially, hence this is the unique  $\mathfrak{r}$  we are after. The same fact shows that

$$\mathfrak{r} = (\mathfrak{r} \cap \tilde{\mathfrak{n}}^-) \oplus (\mathfrak{r} \cap \tilde{\mathfrak{n}}^+).$$

So we are done.  $\square$

Now define the Kac-Moody Lie algebra  $\mathfrak{g} = \mathfrak{g}(\Gamma)$  associated to the graph  $\Gamma$ :  $\mathfrak{g} := \tilde{\mathfrak{g}}/\mathfrak{r}$ . Also let  $\mathfrak{n}^\pm = \tilde{\mathfrak{n}}^\pm/(\mathfrak{r} \cap \tilde{\mathfrak{n}}^\pm) \subseteq \mathfrak{g}$ . We also write simply  $e_i, f_i$  for the canonical images of these elements in  $\mathfrak{g}$ . Note since  $[e_i, f_i] = \epsilon_i \neq 0$  in  $\mathfrak{g}$  these elements are definitely still non-zero. Let us now list the properties of  $\mathfrak{g}$  that follow directly from the theorem.

- (Triangular decomposition)  $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$ .
- (Root space decomposition)  $\mathfrak{g} = \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha$ .
- (Finite dimensionality of weight spaces)  $\dim \mathfrak{g}_\alpha \leq n^{|\text{ht}(\alpha)|}$  for each  $0 \neq \alpha \in R$ ;  $\mathfrak{g}_0 = \mathfrak{h}$  of dimension  $n + \dim \mathfrak{c}$ ;  $\mathfrak{g}_{\epsilon_i} = \mathbb{C}e_i$  and  $\mathfrak{g}_{-\epsilon_i} = \mathbb{C}f_i$  for each  $i = 1, \dots, n$ .

- (Roots) We call  $\alpha \in R$  a *root* if  $\alpha \neq 0$  and  $\mathfrak{g}_\alpha \neq 0$ . We know every root is either positive or negative. Let  $\Delta, \Delta^+, \Delta^-$  denote the sets of roots, positive roots, negative roots, so

$$\Delta = \Delta^+ \sqcup \Delta^-.$$

- (Chevalley involution) Note  $\tilde{\omega}$  induces an antiautomorphism  $\omega$  of  $\mathfrak{g}$  interchanging the  $e_i, f_i$  and  $\mathfrak{g}_\alpha$  with  $\mathfrak{g}_{-\alpha}$ . It follows that  $\Delta^- = -\Delta^+$  and  $\dim \mathfrak{g}_{-\alpha} = \dim \mathfrak{g}_\alpha$ .
- (commutator subalgebra) Let  $\mathfrak{g}'$  be the commutator subalgebra  $[\mathfrak{g}, \mathfrak{g}]$ . It is obvious that this is the subalgebra generated by  $e_i, f_i$  for  $i = 1, \dots, n$ . Hence,  $\mathfrak{g}' = \mathfrak{n}^- \oplus \mathfrak{h}' \oplus \mathfrak{n}^+$ .
- (direct sums) One more remark that is easy to see from the definition: if  $\Gamma = \Gamma' \sqcup \Gamma''$  then  $\mathfrak{g}(\Gamma) \cong \mathfrak{g}(\Gamma') \oplus \mathfrak{g}(\Gamma'')$ . So we can always restrict our attention to connected graphs.

I want to prove two more basic facts about  $\mathfrak{g}(\Gamma)$ . Note Lemma 3.1 is crucial in these proofs too.

**Lemma 3.4.** (i) Let  $a \in \mathfrak{n}^+$  be such that  $[a, f_i] = 0$  for all  $i$ . Then  $a = 0$ . Similarly for  $a \in \mathfrak{n}^-$ .

(ii) The center of the Lie algebra  $\mathfrak{g}$  (or of  $\mathfrak{g}'$ ) is  $\mathfrak{c}$ , the radical of the form  $(\cdot, \cdot)$  on  $\mathfrak{g}'$ .

(iii) If the graph  $\Gamma$  is connected, then every ideal of  $\mathfrak{g}$  either contains  $\mathfrak{g}'$  or is contained in the center  $\mathfrak{c}$ .

(iv)  $\mathfrak{g}$  is a simple Lie algebra if and only if the graph  $\Gamma$  is connected and  $\mathfrak{c} = 0$ .

*Proof.* (i) Consider the subspace  $I$  of  $\mathfrak{n}^+$  generated by all elements obtained from  $a$  by commuting with elements of  $\mathfrak{h}$  and with the  $e_i$ 's. This subspace is invariant under the action of  $\mathfrak{h}$  and  $\mathfrak{n}^+$  by definition. Now take  $v \in I$  and consider  $[f_i, v]$ . Using the Jacobi identity and the fact that  $[f_i, a] = 0$  we see that  $[f_i, v] \in I$  again. Hence,  $I$  is a non-trivial ideal intersecting  $\mathfrak{h}$  trivially, contradicting the definition of  $\mathfrak{g}$ .

(ii) Let  $c$  lie in the center of  $\mathfrak{g}$ , i.e.  $[c, x] = 0$  for all  $x \in \mathfrak{g}$ . Using Lemma 3.1, we may assume  $c \in \mathfrak{g}_\alpha$  for some  $\alpha \in R$ . By (i) we then get that  $c \in \mathfrak{h}$ . Say  $c = h' + h''$  with  $h' \in \mathfrak{h}', h'' \in \mathfrak{h}''$ . Now,  $[c, e_i] = (\epsilon_i, c)e_i = 0$  implies first that  $[c, c] = [c, h''] = 0$ , whence  $h'' = 0$ , then that  $c$  lies in the radical of the form, i.e.  $c \in \mathfrak{c}$ . Conversely, if  $c \in \mathfrak{c}$ , then  $c$  commutes with all the  $e_i$  and  $f_i$  and with  $\mathfrak{h}$  so  $c$  does belong to the center.

(iii) Let  $I$  be a non-zero ideal not contained in  $\mathfrak{c}$ . If  $I \subseteq \mathfrak{h}$ , then we can pick  $h \in I$  with  $(h, \epsilon_i) \neq 0$  for some  $i$ . But then  $[h, e_i] = (h, \epsilon_i)e_i \in I$  too, a contradiction. Hence we can find some  $0 \neq x \in \mathfrak{g}_\alpha \cap I$  for some  $0 \neq \alpha \in \Delta$ . WLOG  $\alpha > 0$ . By (i), there exists some  $i$  such that  $[f_i, x] \neq 0$ . This belongs to the  $\alpha - \epsilon_i$  root space. Repeating the argument if necessary, we see that there is some non-zero element  $x$  lying in the  $\epsilon_i$  root space for some  $i$ , i.e.  $e_i$  lies in  $I$ . Then,  $\epsilon_i = [e_i, f_i]$  lies in  $I$  too. Now let  $j$  be a neighbour to  $i$  in the graph. Then,  $[\epsilon_i, e_j]$  is a non-zero multiple of  $e_j$ , so we get that  $e_j$  lies in

$I$  too. Similarly,  $f_j$  lies in  $I$ . Repeating the argument using connectedness of the graph shows that all  $e_i, f_i$  lie in  $I$ , hence  $I \supseteq \mathfrak{g}'$ .

(iv) By (iii) if these conditions hold then  $\mathfrak{g}$  is simple. It is easy to see that these conditions are necessary.  $\square$

**Example 3.5.** Let us show that the Lie algebra  $\mathfrak{g} = \mathfrak{sl}_{n+1}(\mathbb{C})$  is the Kac-Moody Lie algebra associated to the graph of type  $A_n$ . Recall first how we constructed  $\mathfrak{h}$  explicitly in that case: we start from the complex vector space with orthonormal basis  $v_1, \dots, v_{n+1}$  and let  $\mathfrak{h}$  be the subspace spanned by the elements  $\epsilon_i = v_i - v_{i+1}$ . This can be seen in matrices:  $v_i$  is the  $i$ th diagonal matrix unit then  $\epsilon_i$  is of trace zero,  $\mathfrak{h}$  is the diagonal trace zero matrices, and the bilinear form is the *trace form*  $(x, y) = \text{tr}(xy)$ .

For  $i \neq j$ , let  $e_{i,j}$  denote the  $ij$  matrix unit. Let  $e_i = e_{i,i+1}$  and  $f_i = e_{i+1,i}$ . Then it is easy to check that the relations

$$[e_i, f_i] = \epsilon_i, [\epsilon_i, \epsilon_j] = 0, [\epsilon_i, e_j] = (\epsilon_i, \epsilon_j)e_j, [\epsilon_i, f_j] = -(\epsilon_i, \epsilon_j)f_j$$

are satisfied. Moreover,  $\mathfrak{g}$  has no non-zero ideal intersecting  $\mathfrak{h}$ -trivially: given such an ideal one can project it onto some weight component  $ce_{i,j}$ , but then commuting with  $e_{j,i}$  gives a non-zero element of  $\mathfrak{h}$ . Hence,  $\mathfrak{g}$  is a quotient of the corresponding Kac-Moody Lie algebra. Since the latter is known to be simple by the lemma below, it follows that  $\mathfrak{g}$  is EQUAL to the corresponding KM Lie algebra and we're done.

Now we can see the root space decomposition explicitly:

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+ = \bigoplus_{i>j} \mathbb{C}e_{i,j} \oplus \mathfrak{h} \oplus \bigoplus_{i<j} \mathbb{C}e_{i,j}$$

and  $[h, e_{i,j}] = (v_i - v_j, h)e_{i,j}$ , i.e.  $e_{i,j}$  is of weight  $v_i - v_j = \epsilon_i + \epsilon_{i+1} + \dots + \epsilon_{j-1}$  with respect to  $\mathfrak{h}$  if  $i < j$ . Thus the roots  $\Delta$  defined abstractly from the non-zero root spaces coincide with the roots  $\Delta$  we defined graph theoretically before!

**Example 3.6.** Consider the graphs  $\Gamma$  with just two vertices. We've just seen that the one with just one connecting edge is  $\mathfrak{sl}_3(\mathbb{C})$ . All its non-zero root spaces are therefore 1 dimensional. We will look quite soon at the Euclidean case when there are two connecting edges, and see that all non-zero roots spaces there are 1 dimensional two. On the other hand as soon as there are three connecting edges, you can show that there are non-zero root spaces of dimension greater than 1.

I should make some comments on Lie groups in the finite dimensional cases; infinite dimensional groups in the infinite dimensional cases.

One more lemma. These relations are the *Serre relations*. We will show eventually that these form a complete set of generators for the ideal  $\mathfrak{r}$  we defined abstractly when constructing  $\mathfrak{g}$ , but that needs quite a lot more work about representation theory.

**Lemma 3.7.** For  $i \neq j$ , the following relations hold in  $\mathfrak{g}$ :

$$(\text{ad } e_i)^{1+n_{i,j}} e_j = 0, \quad (\text{ad } f_i)^{1+n_{i,j}} f_j = 0.$$

*Proof.* In view of  $\omega$ , it suffices to prove the second one. Let  $\mathfrak{g}(i)$  be the subalgebra of  $\mathfrak{g}$  generated by  $e_i, f_i, \epsilon_i$ . This is just a copy of  $\mathfrak{sl}_2$ , so we know something about its representation theory already. Consider the adjoint representation of  $\mathfrak{g}(i)$  on  $\mathfrak{g}$ . Let  $v_0 = f_j$ . Note that

$$\operatorname{ad} e_i v_0 = [e_i, f_j] = 0, \quad \operatorname{ad} \epsilon_i v_0 = [\epsilon_i, f_j] = -(\epsilon_i, \epsilon_j) f_j = n_{i,j} f_j.$$

In other words,  $v_0$  is a “primitive vector of weight  $n_{i,j}$  for the action of  $\mathfrak{g}(i)$ ”. Hence we let  $v_n = (\operatorname{ad} f_i)^n v_0$  then we know by induction that

$$e_i v_{n+1} = (n_{i,j} - n)(n + 1) v_n.$$

We are trying to prove that  $\theta := v_{n_{i,j}+1} = 0$ . The identity just proves shows that  $e_i \theta = 0$ . Therefore we just remains to show that  $\operatorname{ad} e_j \theta = 0$  and  $\operatorname{ad} e_k \theta = 0$  for all  $k \neq i, j$ . The latter is obvious since  $[e_k f_i] = [e_k f_j] = 0$ . For the former, we have that

$$\operatorname{ad} e_j \theta = (\operatorname{ad} f_i)^{n_{i,j}+1} [e_j, f_j] = (\operatorname{ad} f_i)^{n_{i,j}+1} \epsilon_j = (\operatorname{ad} f_i)^{n_{i,j}} (\epsilon_i, \epsilon_j) f_i$$

which is obviously zero if  $n_{i,j} > 0$  as  $[f_i, f_i] = 0$  and it is zero if  $n_{i,j} = 0$  as  $(\epsilon_i, \epsilon_j) = -n_{i,j}$ .  $\square$

The goal next: to identify the Lie algebras  $\mathfrak{g}(\Gamma)$  explicitly in the most important Dynkin and Euclidean cases. Here there are alternative constructions which reveal explicitly what the structure is. Things like identifying the roots, seeing the root spaces are all one dimensional, etc... is then easy in these cases. In general it is more tricky but possible to describe the roots abstractly. As far as I know, noone knows a good way to compute the dimensions of the root spaces in general.

**3.3. The Dynkin case.** In the case  $\Gamma$  is a Dynkin diagram, we have that  $\mathfrak{c} = 0, \mathfrak{h}' = \mathfrak{h}$ . I want to give another more concrete definition of  $\mathfrak{g}$  in these cases, which will allow us to identify  $\Delta$  defined abstractly in the previous section with the set  $\Delta$  of roots defined in the previous chapter, and also to see that all the root spaces are one dimensional. Note in particular this obviously identifies our Kac-Moody algebras in the case that  $\Gamma$  is a Dynkin diagram with the corresponding finite dimensional simple Lie algebra, for those of you who know that classification.

So let  $\Gamma$  be a Dynkin diagram. Also let  $\Delta \subset R$  be the set of roots defined as in the previous chapter, i.e. the  $\alpha \in R$  such that  $(\alpha, \alpha) = 2$ . (Sorry, it is a little confusing to keep jumping back and forth, but now I am going to follow the language of the previous chapter for a little while we make the identification!) We know from the previous chapter that  $\Delta$  is a finite set, indeed,  $\Delta$  is just the orbit of one of the simple roots  $\epsilon_i$  under the action of the finite group  $W = \langle s_i \rangle$ .

**Lemma 3.8.** *Assume  $\Gamma$  is Dynkin. If  $\alpha, \beta \in \Delta^+$ , then  $\alpha + \beta \in \Delta^+$  if and only if  $(\alpha, \beta) = -1$ .*

*Proof.* ( $\Leftarrow$ ) In this case,  $s_\alpha(\beta) = \beta + \alpha$  so it is a root.

( $\Rightarrow$ ) Suppose  $\alpha, \beta, \alpha + \beta \in \Delta$ . Then we have that

$$2 = (\alpha + \beta, \alpha + \beta) = (\alpha, \alpha) + (\beta, \beta) + 2(\alpha, \beta) = 4 + 2(\alpha, \beta).$$

Hence  $(\alpha, \beta) = -1$ .  $\square$

Pick an orientation for the edges of  $\Gamma$  and let  $Q$  be the corresponding quiver. Using this choice, we construct an “asymmetry function” on the root lattice

$$\nu : R \times R \rightarrow \{\pm 1\} :$$

define

$$\nu(\epsilon_i, \epsilon_j) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are not connected,} \\ 1 & \text{if } i \rightarrow j, \\ -1 & \text{if } i \leftarrow j, \\ -1 & \text{if } i = j. \end{cases}$$

Extend  $\nu$  to all of  $R \times R$  by “bilinearity”, i.e. by

$$\nu(\alpha + \alpha', \beta) = \nu(\alpha, \beta)\nu(\alpha', \beta)$$

for all  $\alpha, \alpha', \beta, \beta' \in Q$ .

**Lemma 3.9.**  $\nu(\alpha, \alpha) = (-1)^{(\alpha, \alpha)/2}$  for all  $\alpha \in R$ . In particular,

$$\nu(\alpha, \beta)\nu(\beta, \alpha) = (-1)^{(\alpha, \beta)}.$$

for all  $\alpha, \beta \in R$ .

*Proof.* Say  $\alpha = \sum_{i=1}^n \alpha_i \epsilon_i$ . Then,

$$\nu(\alpha, \alpha) = \prod_{i,j=1}^n \nu(\epsilon_i, \epsilon_j)^{\alpha_i \alpha_j} = \prod_{i \leftarrow j} (-1)^{\alpha_i \alpha_j} \prod_i (-1)^{\alpha_i \alpha_i}.$$

This is equal to  $(-1)^{(\alpha, \alpha)/2}$  by definition of the bilinear form. The second statement follows by applying the first to  $\alpha + \beta$  and using bilinearity.  $\square$

Now define a Lie algebra  $\mathfrak{g} = \mathfrak{g}(Q)$  as follows. Let  $\mathfrak{h} = \mathbb{C} \otimes_{\mathbb{Z}} R$  (the same as in the definition of  $\mathfrak{g}(\Gamma)$ ), on basis  $\epsilon_1, \dots, \epsilon_n$ . Let

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathbb{C} e_\alpha$$

as a vector space. Define a multiplication by the formulae

$$\begin{aligned} [\epsilon_i, \epsilon_j] &= 0, \\ [\epsilon_i, e_\alpha] &= (\alpha, \epsilon_i) e_\alpha, \\ [e_\alpha, e_{-\alpha}] &= -\alpha, \\ [e_\alpha, e_\beta] &= 0 \text{ if } \alpha + \beta \notin \Delta \cup \{0\}, \\ [e_\alpha, e_\beta] &= \nu(\alpha, \beta) e_{\alpha+\beta} \text{ if } \alpha + \beta \in \Delta. \end{aligned}$$

You of course have to check that  $\mathfrak{g}$  as just defined is a Lie algebra, i.e. satisfies the axioms. The difficult one is the Jacobi identity, which is checked by a little case analysis. For example, suppose that  $\alpha, \beta, \gamma \in \Delta$  with

$\alpha + \beta, \beta + \gamma, \alpha + \gamma$  all also lie in  $\Delta$  (the hardest situation, actually the only one that is not obvious to analyse). We need to check that

$$[e_\alpha, [e_\beta, e_\gamma]] + [e_\beta, [e_\gamma, e_\alpha]] + [e_\gamma, [e_\alpha, e_\beta]] = 0.$$

Note that by the above lemma,  $(\alpha, \beta) = (\alpha, \gamma) = (\beta, \gamma) = -1$ . Hence  $(\alpha + \beta + \gamma, \alpha + \beta + \gamma) = 0$ , so  $\alpha + \beta + \gamma = 0$  (BECAUSE THE FORM IS POSITIVE DEFINITE!!!) So we can compute to get

$$\nu(\beta, \gamma)(\beta + \gamma) + \nu(\gamma, \alpha)(\gamma + \alpha) + \nu(\alpha, \beta)(\alpha + \beta).$$

Now use the properties of  $\nu$  to show this is zero: it is

$$\nu(\beta, \gamma)(\beta + \gamma) - \nu(\gamma, -\beta - \gamma)(\beta) - \nu(-\beta - \gamma, \beta)(\gamma) = 0$$

since  $\nu(\gamma, -\beta - \gamma) = \nu(-\beta - \gamma, \beta) = \nu(\beta, \gamma)$  by the above properties.

**Theorem 3.10.** *There is a unique isomorphism  $\mathfrak{g}(\Gamma) \rightarrow \mathfrak{g}(Q)$  that is the identity on  $\mathfrak{h}$  and maps  $e_i$  resp.  $f_i$  to  $e_{\epsilon_i}$  resp.  $-e_{-\epsilon_i}$ .*

*Proof.* First you check the given elements satisfy the relations defining  $\tilde{\mathfrak{g}}(\Gamma)$ . Hence there is such an homomorphism  $\tilde{\mathfrak{g}}(\Gamma) \rightarrow \mathfrak{g}(Q)$ .

To show it is onto, we need to show that the  $e_{\pm\epsilon_i}$  generate all of  $\mathfrak{g}(Q)$ . To see this, it suffices to see that given any positive but not simple root  $\alpha$ , we can pick  $i$  such that  $\alpha - \epsilon_i$  is also a root – then you can build  $e_\alpha$  by induction on height. But we can certainly pick  $i$  such that  $(\alpha, \epsilon_i) > 0$  as  $(\alpha, \alpha) > 0$ . Then using the representation theory of  $\mathfrak{sl}_2$  we see that the  $\alpha - \epsilon_i$  root space must be non-zero.

Finally, we check that  $\mathfrak{g}(Q)$  has no non-trivial ideal intersecting  $\mathfrak{h}$  trivially. Well by the little lemma, we can certainly find some  $e_\alpha$  lying in such an ideal. But then bracketing with  $e_{-\alpha}$  gives an element of  $\mathfrak{h}$ . So we get induced a homomorphism  $\mathfrak{g}(\Gamma) \rightarrow \mathfrak{g}$ . Since  $\mathfrak{g}(\Gamma)$  is simple, it must be an isomorphism.  $\square$

**Corollary 3.11.** *In the Dynkin case, the set  $\Delta$  of roots is the set  $\alpha \in R$  with  $(\alpha, \alpha) = 2$ , and  $\dim \mathfrak{g}_\alpha = 1$  for each such  $\alpha$ .*

Since we know the number  $|\Delta^+|$  of positive roots for each of the Dynkin diagrams (see table in previous chapter), we get  $\dim \mathfrak{g} = n + 2|\Delta^+|$  directly from this explicit construction:

**Corollary 3.12.** *In the Dynkin case,  $\mathfrak{g}(\Gamma)$  is a finite dimensional Lie algebra of dimension  $n^2 + 2n$  (type  $A_n$ ),  $2n^2 - n$  (type  $D_n$ ), 78 (type  $E_6$ ), 133 (type  $E_7$ ), 248 (type  $E_8$ ).*

In fact these are the ONLY graphs for which  $\mathfrak{g}(\Gamma)$  is finite dimensional, as we should see in a while.

**Exercise 1.** In the case  $\Gamma = A_n$ , we have now given two explicit constructions of the Kac-Moody algebra  $\mathfrak{g}(\Gamma)$ : as the Lie algebra  $\mathfrak{sl}_{n+1}(\mathbb{C})$  or as the Lie algebra  $\mathfrak{g}(Q)$  for the quiver

$$Q = 1 \rightarrow 2 \rightarrow \cdots \rightarrow n.$$

The goal of this exercise is to identify the two. Let  $\mathfrak{h} = \mathbb{C}\epsilon_1 + \cdots + \mathbb{C}\epsilon_n$  as usual. Identify this with the subalgebra of  $\mathfrak{sl}_{n+1}(\mathbb{C})$  consisting of the diagonal trace zero matrices, so that  $\epsilon_i \leftrightarrow e_{i,i} - e_{i+1,i+1}$  (matrix units). Recall that  $\Delta = \{\epsilon_i - \epsilon_j \mid 1 \leq i \neq j \leq n\}$ . Prove explicitly that the map

$$h \mapsto h \quad (h \in \mathfrak{h}), \quad e_{\epsilon_i - \epsilon_j} \mapsto e_{i,j} \quad (i < j), \quad e_{\epsilon_i - \epsilon_j} \mapsto -e_{i,j} \quad (i > j)$$

is an isomorphism between  $\mathfrak{g}(Q) = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathbb{C}e_\alpha$  and  $\mathfrak{sl}_{n+1}(\mathbb{C}) = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathbb{C}e_{i,j}$ .

**Exercise 2** (Triality). Consider the quiver  $Q = D_4$  consisting of three vertices 1, 2, 3 around the edge, one vertex 4 in the middle with all arrows pointing inwards. Recall that the positive roots are  $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_1 + \epsilon_4, \epsilon_2 + \epsilon_4, \epsilon_3 + \epsilon_4, \epsilon_1 + \epsilon_2 + \epsilon_4, \epsilon_2 + \epsilon_3 + \epsilon_4, \epsilon_1 + \epsilon_3 + \epsilon_4, \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4, \epsilon_1 + \epsilon_2 + \epsilon_3 + 2\epsilon_4$ . Let  $\nu$  be the asymmetry function as defined above, and  $\mathfrak{g} = \mathfrak{h} + \bigoplus_{\alpha \in \Delta} \mathbb{C}e_\alpha$  be the Lie algebra  $\mathfrak{g}(Q)$  ( $\cong \mathfrak{so}_8(\mathbb{C})$ ) as defined above. Let  $\tau : \mathfrak{h} \rightarrow \mathfrak{h}$  be the automorphism sending  $\epsilon_1 \rightarrow \epsilon_2 \rightarrow \epsilon_3 \rightarrow \epsilon_1$  and fixing  $\epsilon_4$ . (This corresponds to the natural symmetry of the quiver  $Q$  so it really is an isometry).

- Prove that  $\nu(\alpha, \beta) = \nu(\tau(\alpha), \tau(\beta))$ .
- Deduce that setting  $\tau(e_\alpha) = e_{\tau(\alpha)}$  for each  $\alpha \in \Delta$  defines an automorphism of  $\mathfrak{g}$  of order 3.
- Construct this automorphism of  $\mathfrak{g}$  in another way using our original definition of  $\mathfrak{g}$  as the quotient of  $\tilde{\mathfrak{g}}$  by the ideal  $\mathfrak{r}$ .

**3.4. The invariant bilinear form.** A quick recap:

- For each graph  $\Gamma$  with no loops, we have defined a vector space  $\mathfrak{h}$  equipped with a non-degenerate bilinear form. The subspace  $\mathfrak{h}'$  of  $\mathfrak{h}$  has basis  $\epsilon_1, \dots, \epsilon_n$  and the form satisfies  $(\epsilon_i, \epsilon_i) = 2, (\epsilon_i, \epsilon_j) = -n_{i,j}$ .
- We have associated a Lie algebra  $\mathfrak{g}$  generated by  $\mathfrak{h}$  and elements  $e_i, f_i$  for  $i = 1, \dots, n$ . We know that  $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$ .
- We have the root space decomposition  $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha$ , from which we defined the set  $\Delta$  of roots. We know  $\Delta = \Delta^+ \sqcup \Delta^-$  and  $\Delta^- = -\Delta^+$ .
- If  $\Gamma$  is connected, then  $\mathfrak{g}$  is “almost” simple.
- If  $\Gamma$  is Dynkin, we have given an alternative more explicit construction of  $\mathfrak{g}$  which proved in particular in that case that the set  $\Delta$  of roots is precisely  $\alpha \in R$  with  $(\alpha, \alpha) = 2$ . These are the finite dimensional simple Lie algebras of type ADE.

In this section, I want to extend the non-degenerate bilinear form on  $\mathfrak{h}$  to a nondegenerate *invariant* bilinear form on all of  $\mathfrak{g}$ . Invariant means:  $([x, y], z) = (x, [y, z])$  for all  $x, y, z \in \mathfrak{g}$ . In the case that  $\mathfrak{g}$  is a finite dimensional simple Lie algebra,  $(\cdot, \cdot)$  will be exactly the usual Killing form (up to a scalar).

**Theorem 3.13.** *There exists a unique nondegenerate symmetric bilinear form  $(\cdot, \cdot)$  on  $\mathfrak{g}$  such that*

- $(\cdot, \cdot)$  is invariant, i.e.  $([x, y], z) = (x, [y, z])$  for all  $x, y, z \in \mathfrak{g}$ ;

- (b) the restriction of  $(.,.)$  to  $\mathfrak{h}$  is the form already defined on  $\mathfrak{h}$ ;
- (c)  $(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$  if  $\alpha \neq -\beta$ ;
- (d) the restriction of  $(.,.)$  to  $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$  is non-degenerate for each  $\alpha \in \Delta$  and moreover  $[x, y] = (x|y)\alpha$  for  $x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_{-\alpha}$ .

*Proof.* For  $j \in \mathbb{Z}$ , let  $\mathfrak{g}_j$  be the sum of all  $\mathfrak{g}_\alpha$  for  $ht(\alpha) = j$ . Note that  $[\mathfrak{g}_j, \mathfrak{g}_k] \subseteq \mathfrak{g}_{j+k}$ . In fancy language, this makes  $\mathfrak{g}$  into a *graded Lie algebra*. For  $N = 0, 1, \dots$ , let  $\mathfrak{g}(N)$  be the sum of  $\mathfrak{g}_j$  with  $|j| \leq N$ . We are going to extend the form  $(.,.)$  on  $\mathfrak{h}$  to  $\mathfrak{g}(N)$  inductively.

To start with, extend it to a symmetric bilinear form on  $\mathfrak{g}(1) = \mathfrak{h} \oplus \bigoplus_{i=1}^n \mathbb{C}e_i \oplus \mathbb{C}f_i$  by setting

$$(e_i, f_j) = \delta_{i,j}, (\mathfrak{h}, e_i) = (\mathfrak{h}, f_i) = (e_i, f_j) = 0.$$

Observe the form  $(.,.)$  on  $\mathfrak{g}(1)$  satisfies (a) as long as  $x, [x, y], z$  and  $[y, z]$  all belong to  $\mathfrak{g}(1)$ . To see this, a little case analysis reduces easily just to having to check

$$([e_i, f_j], h) = (e_i, [f_j, h])$$

(and a similar equality with  $e, f$  swapped). But for this, the left hand side is  $\delta_{i,j}(\epsilon_i, h)$  and the right hand side is  $(\epsilon_i, h)(e_i, f_j) = \delta_{i,j}(\epsilon_i, h)$  so checked.

Now suppose we have defined  $(.,.)$  on  $\mathfrak{g}(N-1)$  so that  $(\mathfrak{g}_i, \mathfrak{g}_j) = 0$  if  $|i|, |j| \leq N-1$  and  $i+j \neq 0$  and also so that (a) is satisfied when it makes sense. To extend it to a form on  $\mathfrak{g}(N)$ , it suffices to consider  $x \in \mathfrak{g}_{\pm N}$  and  $y \in \mathfrak{g}_{\mp N}$ . Write  $y = \sum_i [u_i, v_i]$  where  $u_i, v_i$  are homogeneous of non-zero degree lying in  $\mathfrak{g}(N-1)$ . Then,  $[x, u_i] \in \mathfrak{g}(N-1)$  and we set

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

Note we are forced to do this if we want the form to be invariant!

Now the main thing is to check well-definedness. We first show that if  $i, j, s, t \in \mathbb{Z}$  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

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

To see this, we have by the inductively known invariance of the form on  $\mathfrak{g}(N-1)$  that

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

Now using this we can check well-definedness. Write  $x = \sum_j [u'_j, v'_j]$ . Then,

$$\begin{aligned} (x, y) &= \sum_i ([x, u_i], v_i) = \sum_{i,j} ([[u'_j, v'_j], u_i], v_i) = \sum_{i,j} (u'_j, [v'_j, [u_i, v_i]]) \\ &= \sum_j (u'_j, [v'_j, y]). \end{aligned}$$

This shows  $(x, y)$  is independent of the given choice of expression for  $y$ !

Now we've extended our form to  $\mathfrak{g}(N)$ , and it is automatically invariant there whenever it makes sense because that is how we have defined it. Hence we have constructed by induction an invariant bilinear form on  $\mathfrak{g}$  such that  $(\mathfrak{g}_i, \mathfrak{g}_j) = 0$  if  $i \neq -j$ . We still need to verify (c) and (d). Note once we have done that it will follow at once that our form is symmetric – we don't claim that yet!

For (c), take  $h \in \mathfrak{h}$ ,  $x \in \mathfrak{g}_\alpha$  and  $y \in \mathfrak{g}_\beta$ . Then,

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

If  $\alpha + \beta \neq 0$ , we can choose  $h$  so  $(\alpha + \beta, h) \neq 0$  then we get that  $(x, y) = 0$ , which is (c).

Next take  $x \in \mathfrak{g}_\alpha$  and  $y \in \mathfrak{g}_{-\alpha}$  and  $h \in \mathfrak{h}$ . Then,

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

So by the fact that  $(\cdot, \cdot)$  is non-degenerate on  $\mathfrak{h}$ , we must have that  $[x, y] = (x, y)\alpha$  which is the last part of (d).

It now just remains to show that  $(\cdot, \cdot)$  is non-degenerate to complete the proof. Well, let  $I$  be its radical. Since the form is invariant, this is an ideal, and since the form is non-degenerate on  $\mathfrak{h}$ , we have that  $I \cap \mathfrak{h} = 0$ . Hence  $I = 0$ .  $\square$

**Example 3.14.** Suppose  $\Gamma$  is Dynkin. Pick an orientation  $Q$  and let  $\mathfrak{g}(Q) = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathbb{C}e_\alpha$  be our explicit construction for  $\mathfrak{g}(\Gamma)$  from the previous section. Then, the invariant bilinear form satisfies

$$(e_\alpha, e_\beta) = -\delta_{\alpha, -\beta}, (h, e_\alpha) = 0$$

for  $\alpha, \beta \in \Delta$  and  $h \in \mathfrak{h}$ , and it is the form we started with on  $\mathfrak{h}$ . This follows at once from property (d) of the theorem. (In this case you can take these formulae as the definition and verify directly that it is a non-degenerate symmetric invariant bilinear form).

We will always make use of the invariant bilinear form from now on. Note it encodes part of the multiplication in the Lie algebra  $\mathfrak{g}$ .

**3.5. Derivations and central extensions.** To prepare for the construction of the affine Lie algebras in the next section, let us discuss extensions a little bit. Let us first explain how to extend a Lie algebra  $\mathfrak{g}$  by the one

dimensional trivial Lie algebra  $\mathbb{C}$  *on the top*, i.e. how to form an exact sequence of Lie algebras

$$0 \rightarrow \mathfrak{g} \rightarrow \mathfrak{e} \rightarrow \mathbb{C} \rightarrow 0.$$

As a vector space,  $\mathfrak{e} = \mathfrak{g} \oplus \mathbb{C}d$ , and to extend the Lie bracket from  $\mathfrak{g}$  to  $\mathfrak{e}$  we just need to specify in addition what  $[d, x]$  is for all  $x \in \mathfrak{g}$ , i.e. a linear map  $D : \mathfrak{g} \rightarrow \mathfrak{g}$ . For this linear map to make  $\mathfrak{e}$  into a Lie algebra, it is easy to see that  $d$  must be a *derivation*. Given a different derivation  $D' : \mathfrak{g} \rightarrow \mathfrak{g}$  one can form a different extension  $\mathfrak{e}'$ . This is equivalent to the first one (in the usual sense of extensions) if and only if  $D - D'$  is an *inner derivation* (exercise!), i.e. of the form  $\text{ad } x$  for  $x \in \mathfrak{g}$ . So the thing that really matters is  $\text{Der } \mathfrak{g} / \text{ad } \mathfrak{g}$ . This is just  $H^1(\mathfrak{g}, \mathfrak{g})$ .

Next we explain how to extend a Lie algebra  $\mathfrak{g}$  by the one dimensional trivial  $\mathfrak{g}$ -module viewed as an abelian Lie algebra *at the bottom*, i.e.

$$0 \rightarrow \mathbb{C} \rightarrow \mathfrak{e} \rightarrow \mathfrak{g} \rightarrow 0.$$

In other words, we want to form a 1 dimensional *central extension* of  $\mathfrak{g}$ . As a vector space, we must have that  $\mathfrak{e} = \mathfrak{g} \oplus \mathbb{C}\delta$ , and  $\delta$  must be central. To specify the Lie bracket in  $\mathfrak{e}$ , we have for  $x, y \in \mathfrak{g}$  that

$$[x, y]_{\text{new}} = [x, y] + \psi(x, y)\delta$$

for some bilinear function  $\psi : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{C}$ . In order for such a function to make  $\mathfrak{e}$  into a well-defined Lie algebra, it must satisfy the properties:

- (C1)  $\psi(x, y) = -\psi(y, x)$ ;
- (C2)  $\psi([x, y], z) + (\text{cyclic}) = 0$ .

Such a map  $\psi : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{C}$  is called a *2-cocycle with values in  $\mathbb{C}$* . If  $\psi$  and  $\psi'$  are two such 2-cocycles, the corresponding central extensions  $\mathfrak{e}$  and  $\mathfrak{e}'$  are equivalent in the usual sense if and only if  $\psi - \psi'$  is a *2-coboundary*, i.e. a 2-cocycle of the form  $\psi(x, y) = f([x, y])$  for some linear map  $f : \mathfrak{g} \rightarrow \mathbb{C}$ . The 2-cocycles modulo 2-coboundaries is the space  $H^2(\mathfrak{g}, \mathbb{C})$ .

It is known that  $H^2(\mathfrak{g}, \mathbb{C}) = 0$  in the case that  $\mathfrak{g}$  is a finite dimensional simple Lie algebra, so nothing interesting in that case. Also all derivations are inner . . .

Here is a method to construct 2-cocycles.

**Lemma 3.15.** *Suppose that  $\mathfrak{g}$  is a Lie algebra equipped with a symmetric invariant bilinear form  $(\cdot, \cdot)$ , and that  $D : \mathfrak{g} \rightarrow \mathfrak{g}$  is a derivation with the property that  $(Dx, y) = -(x, Dy)$ . Then*

- (i) *The function  $\psi(x, y) = (Dx, y)$  is a 2-cocycle.*
- (ii) *Letting  $\mathfrak{e} = \mathfrak{g} \oplus \mathbb{C}\delta$  be the corresponding one dimensional central extension of  $\mathfrak{g}$ , the derivation  $D$  extends to a derivation  $D$  of  $\mathfrak{e}$  with  $D(\delta) = 0$ .*

*Proof.* The axiom (C1) is immediate. For (C2), we have that

$$\psi([x, y], z) = ((Dx, y) + [x, Dy], z) = (Dx, [yz]) + (Dy, [zx]).$$

Now the Jacobi identity follows since

$$\begin{aligned}
& (Dx, [yz]) + (Dy, [zx]) + (Dz, [xy]) \\
&= -(x, D([yz])) + (Dy, [zx]) + (Dz, [xy]) \\
&= -(x, [Dy, z]) - (x, [y, Dz]) + (Dy, [zx]) + (Dz, [xy]) \\
&= -([zx], Dy) - ([xy], Dz) + (Dy, [zx]) + (Dz, [xy]) = 0.
\end{aligned}$$

So we get  $\mathfrak{e} = \mathfrak{g} \oplus \mathbb{C}\delta$  with

$$[x, y]_{new} = [x, y] + (Dx, y)\delta.$$

To check that  $D$  extends to a derivation, just note that

$$D([x, y]_{new}) = D([x, y] + (Dx, y)\delta) = [Dx, y] + [x, Dy]$$

while

$$[Dx, y]_{new} + [x, Dy]_{new} = [Dx, y] + [x, Dy] + (D(Dx), y)\delta + (Dx, Dy)\delta.$$

Since

$$(D(Dx), y) = -(Dx, Dy)$$

this shows that

$$D([x, y]_{new}) = [Dx, y]_{new} + [x, Dy]_{new}$$

hence it is a derivation.  $\square$

We will apply this lemma at the crucial moment in the next section.

To end the section, let us discuss the Lie algebra  $\mathfrak{d}$  of derivations of the algebra  $\mathcal{L} = \mathbb{C}[t, t^{-1}]$ . Equivalently (since  $\mathbb{C}[t, t^{-1}]$  is the coordinate ring of the complex torus  $\mathbb{C}^\times$ ),  $\mathfrak{d}$  is the Lie algebra of regular vector fields on  $\mathbb{C}^\times$ .

**Exercise 3.** Let  $\mathcal{L} = \mathbb{C}[t, t^{-1}]$ .

- (i) For  $j \in \mathbb{Z}$ , show that  $d_j = -t^{j+1} \frac{d}{dt}$ , i.e. the function  $p \mapsto -t^{j+1} \frac{dp}{dt}$ , is a derivation of  $\mathcal{L}$ .
- (ii) Show that  $\mathfrak{d} = \text{Der } \mathcal{L} = \bigoplus_{j \in \mathbb{Z}} \mathbb{C}d_j$ .
- (iii) Show that  $\psi : \mathfrak{d} \times \mathfrak{d} \rightarrow \mathbb{C}$ ,  $\psi(d_i, d_j) = \frac{1}{12}(i^3 - i)\delta_{i, -j}$  is a 2-cocycle on  $\mathfrak{d}$ .

The corresponding one dimensional central of  $d$  by a 1 dimensional center  $\mathbb{C}\delta$  is called the *Virasoro algebra*. It is an important infinite dimensional Lie algebra beloved by mathematical physicists!

**3.6. Affine Lie algebras.** In this section, we construct the Lie algebra  $\mathfrak{g}(\Gamma)$  in case  $\Gamma$  is a Euclidean diagram. These are known as the *affine Lie algebras*. Recall our convention from §2.10 for numbering the  $\Gamma$  in the Euclidean case by  $0, 1, \dots, n$  unlike the usual  $1, \dots, n$ . It is maybe worth recalling how we defined  $\mathfrak{g}(\Gamma)$  itself: it was generated by  $\mathfrak{h}$  and  $e_0, e_1, \dots, e_n, f_0, f_1, \dots, f_n$ . How did we define  $\mathfrak{h}$ ? Well, we started from  $\mathfrak{h}' = \mathbb{C}\epsilon_0 \oplus \mathbb{C}\epsilon_1 \oplus \dots \oplus \mathbb{C}\epsilon_n$  with the given bilinear form. Its radical was spanned by the element  $\delta = \epsilon_0 + \theta$ , where  $\theta = \sum_{i=1}^n \delta_i \epsilon_i$  (the highest root in the underlying finite root system). We picked a complement to  $\mathfrak{c}$  in  $\mathfrak{h}'$ . Let's always make the following choice:

$\mathfrak{h}'' = \mathbb{C}\epsilon_1 \oplus \dots \oplus \mathbb{C}\epsilon_n$ . Finally let  $d$  be the basis element of  $\mathfrak{e}^*$  such that  $d(\delta) = 1$ . Then,  $\mathfrak{h} = \mathbb{C}\epsilon_0 \oplus \mathbb{C}\epsilon_1 \oplus \dots \oplus \mathbb{C}\epsilon_n \oplus \mathbb{C}d$  and for this basis the bilinear form satisfies

$$\begin{aligned}(\epsilon_i, \epsilon_i) &= 2, \\(\epsilon_i, \epsilon_j) &= -n_{i,j}, \\(d, \epsilon_i) &= \delta_{i,0}, \\(d, \Lambda_0) &= 0\end{aligned}$$

for  $0 \leq i \neq j \leq n$ . It will be really important to remember this important notation!!!

We also defined in chapter 2 the set  $\Delta$  of roots associated to the Euclidean diagram  $\Gamma$ . We showed moreover that

$$\Delta = \mathbb{Z}\delta + \overset{\circ}{\Delta}$$

where  $\overset{\circ}{\Delta}$  are the roots for the underlying finite Dynkin diagram. Maybe this is a hint as to how to construct  $\mathfrak{g}(\Gamma)$ : we need to start with  $\overset{\circ}{\mathfrak{g}}$  and then do something periodic to get the  $\mathbb{Z}\delta$  added on...

Okay, here goes. Let  $\overset{\circ}{\Gamma}$  be the underlying finite Dynkin graph with vertices  $1, \dots, n$  and let  $\overset{\circ}{\mathfrak{g}} = \mathfrak{g}(\overset{\circ}{\Gamma})$  be the associated finite dimensional simple Lie algebra,  $\overset{\circ}{R} = \mathbb{Z}\epsilon_1 \oplus \dots \oplus \mathbb{Z}\epsilon_n$  its root lattice,  $\overset{\circ}{\Delta}$  is set of roots, etc... We have constructed  $\overset{\circ}{\mathfrak{g}}$  absolutely explicitly above as

$$\overset{\circ}{\mathfrak{g}} = \overset{\circ}{\mathfrak{h}} \oplus \bigoplus_{\alpha \in \overset{\circ}{\Delta}} \mathbb{C}e_{\alpha}.$$

Let  $\mathcal{L} = \mathbb{C}[t, t^{-1}]$ . For a Laurent polynomial  $p$ , we write

$$\text{Res } p$$

for its residue, i.e. the  $t^{-1}$ -coefficient. Thus, Res is the linear function characterized by the properties

$$\text{Res } t^{-1} = 1, \quad \text{Res}\left(\frac{dp}{dt}\right) = 0.$$

By the product rule, we have that

$$\text{Res}\left(\frac{dp}{dt}q\right) + \text{Res}\left(p\frac{dq}{dt}\right) = 0.$$

Form the *loop algebra*

$$\mathcal{L}(\overset{\circ}{\mathfrak{g}}) = \mathcal{L} \otimes \overset{\circ}{\mathfrak{g}}.$$

Make this into a Lie algebra by defining  $[p \otimes x, q \otimes y] = pq \otimes [x, y]$ . Any derivation  $D$  of  $\mathcal{L}$  extends to a derivation  $D \otimes 1$  of  $\mathcal{L}(\overset{\circ}{\mathfrak{g}})$ . In particular there is a derivation of  $\mathcal{L}(\overset{\circ}{\mathfrak{g}})$  defined by

$$D(p \otimes x) = t \frac{dp}{dt} \otimes x.$$

Let  $(., .)$  be the invariant bilinear form on  $\mathfrak{g}$ . We lift it to  $\mathcal{L}(\mathring{\mathfrak{g}})$  by setting

$$(t^n \otimes x, t^m \otimes y) = \delta_{n,-m}(x, y).$$

It is obvious that this is a symmetric invariant bilinear form. Moreover,

$$(D(p \otimes x), q \otimes y) = \text{Res}\left(\frac{dp}{dt}q\right)(x, y) = -\text{Res}\left(p\frac{dq}{dt}\right)(x, y) = -(p \otimes x, D(q \otimes y)).$$

Hence the conditions in Lemma 3.15 are satisfied, so the function

$$\psi(p \otimes x, q \otimes y) = (D(p \otimes x), q \otimes y)$$

is a 2-cocycle. Now invoking Lemma 3.15, we get the Lie algebra

$$\mathfrak{g} = \mathbb{C}\delta \oplus \mathcal{L}(\mathring{\mathfrak{g}}) \oplus \mathbb{C}d$$

with Lie bracket defined by

$$\begin{aligned} [\lambda\delta + \mu t^m \otimes x + \nu d, \lambda'\delta + \mu' t^n \otimes y + \nu' d] = \\ \nu\mu' n t^n \otimes y - \nu'\mu m t^m \otimes x + \mu\mu' t^{m+n} \otimes [x, y] + m\delta_{m,-n}(x, y)\delta. \end{aligned}$$

We will show that  $\mathfrak{g}$  is the Kac-Moody algebra  $\mathfrak{g}(\Gamma)$ . Let

$$\mathfrak{h} = \mathbb{C}\delta \oplus \mathring{\mathfrak{h}} \oplus \mathbb{C}d.$$

This is an  $n+2$  dimensional abelian subalgebra of  $\mathfrak{g}$  with basis  $\delta, \epsilon_1, \dots, \epsilon_n, d$ . Let  $\theta = \sum_{i=1}^n \delta_i \epsilon_i$  be the highest root in  $\mathring{\Delta}$ , and let  $\epsilon_0 = \delta - \theta$ . Then we have identified  $\mathfrak{h}$  with the Cartan subalgebra in  $\mathfrak{g}(\Gamma)$  that we are after. Moreover, there is a root space decomposition. Let

$$\Delta = \{j\delta + \gamma \mid j \in \mathbb{Z}, \gamma \in \mathring{\Delta}\} \cup \{j\delta \mid j \in \mathbb{Z} - \{0\}\}.$$

Then, we have the root space decomposition with respect to  $\mathfrak{h}$ :

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

where

$$\mathfrak{g}_{j\delta+\gamma} = t^j \otimes \mathring{\mathfrak{g}}_\gamma, \mathfrak{g}_{j\delta} = t^j \otimes \mathring{\mathfrak{h}}.$$

Note this is exactly the set  $\Delta$  we were expecting in chapter 2...

To identify the generators  $e_i, f_i$ , let  $E_1, \dots, E_n, F_1, \dots, F_n$  be the Chevalley generators of  $\mathring{\mathfrak{g}}$ . In the explicit construction of  $\mathring{\mathfrak{g}}$ , we had that  $E_i = e_{\epsilon_i}$  and  $F_i = -e_{-\epsilon_i}$ . Also let  $F_0 = e_\theta, E_0 = -e_{-\theta}$ . Note  $[E_0, F_0] = -\theta$  and  $(E_0, F_0) = 1$ . Let

$$e_0 = t \otimes E_0, f_0 = t^{-1} \otimes F_0, e_i = 1 \otimes E_i, f_i = 1 \otimes F_i$$

for  $i = 1, \dots, n$ .

**Theorem 3.16.**  $\mathfrak{g}$  is the Kac-Moody algebra  $\mathfrak{g}(\Gamma)$ , with Chevalley generators  $e_0, e_1, \dots, e_n, f_0, f_1, \dots, f_n$ .

*Proof.* We first need to check the relations of  $\tilde{\mathfrak{g}}(\Gamma)$ . Note that  $e_1, \dots, e_n, f_1, \dots, f_n$  generate the copy  $1 \otimes \mathring{\mathfrak{g}}$  inside of  $\mathfrak{g}$ . So the elements  $\epsilon_1, \dots, \epsilon_n, e_1, \dots, e_n, f_1, \dots, f_n$  all satisfy the right sort of relations. Also,

$$[e_0, f_0] = \delta - \theta = \epsilon_0,$$

while  $[e_0, f_i] = 0$  for  $i = 1, \dots, n$  since  $E_0$  lies in the lowest root space. Similarly,  $[e_i, f_0] = 0$ . Next,

$$[\epsilon_0, e_i] = (\epsilon_0, \epsilon_i)e_i, [\epsilon_i, e_0] = (\epsilon_i, -\theta)e_0 = (\epsilon_i, \epsilon_0)e_0.$$

That is all the relations.

Next we show that  $\mathfrak{g}$  has no ideals intersecting  $\mathfrak{g}$  trivially. Let  $I$  be a non-zero ideal of  $\mathfrak{g}$  intersecting  $\mathfrak{h}$  trivially. By the little lemma,  $I$  must intersect some  $\mathfrak{g}_\alpha$  for  $\alpha \in \Delta$ . Hence,  $t^j \otimes x \in I$  for some  $j \in \mathbb{Z}$  and some  $x \in \mathring{\mathfrak{g}}_\gamma$ . Take  $y \in \mathring{\mathfrak{g}}_{-\gamma}$  such that  $(x, y) \neq 0$ . Then,  $[t^j \otimes x, t^{-j} \otimes y] = j(x, y)\delta + [x, y] \in \mathfrak{h} \cap I$ . This implies that  $j = 0$ , whence  $\gamma \neq 0$  and hence  $[x, y] = (x, y)\gamma \neq 0$  which is a contradiction.

Finally we have to show that  $e_0, e_1, \dots, e_n, f_0, f_1, \dots, f_n$  and  $\mathfrak{h}$  generate all of  $\mathfrak{g}$ . Let  $\hat{\mathfrak{g}}$  be the subalgebra that they generate. Since  $E_i, F_i$  for  $i = 1, \dots, n$  generate  $\mathring{\mathfrak{g}}$ , we obtain that  $1 \otimes \mathring{\mathfrak{g}} \subseteq \hat{\mathfrak{g}}$ . Using  $t \otimes E_0 \in \hat{\mathfrak{g}}$  and the simplicity of  $\mathring{\mathfrak{g}}$ , we obtain that  $t \otimes \mathring{\mathfrak{g}} \subseteq \hat{\mathfrak{g}}$ . Since  $[t \otimes x, t^k \otimes y] = t^{k+1} \otimes [x, y]$  we get by induction on  $k$  that  $t^k \otimes \mathring{\mathfrak{g}} \subseteq \hat{\mathfrak{g}}$ . Similarly for the negative part.  $\square$

**Corollary 3.17.** *The roots  $\Delta$  are  $\{\nu + j\delta \mid \nu \in \mathring{\Delta}, j \in \mathbb{Z}\} \cup \{j\delta \mid j \in \mathbb{Z} - \{0\}\}$ . Recall we called the former the real roots and the latter the imaginary roots. The dimension of  $\mathfrak{g}_\alpha$  is 1 if  $\alpha$  is real,  $n$  if  $\alpha$  is imaginary.*

A couple of other things we can identify now we have this explicit realization: For example, the triangular decomposition

$$\mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$$

has

$$\mathfrak{n}_- = (t^{-1}\mathbb{C}[t^{-1}] \otimes (\mathring{\mathfrak{n}}_+ + \mathring{\mathfrak{h}}) + \mathbb{C}[t^{-1}] \otimes (\mathring{\mathfrak{n}}_-)), \mathfrak{n}_+ = (t\mathbb{C}[t] \otimes (\mathring{\mathfrak{n}}_- + \mathring{\mathfrak{h}}) + \mathbb{C}[t] \otimes (\mathring{\mathfrak{n}}_+)).$$

The invariant bilinear form on  $\mathfrak{g}$  is defined by

$$\begin{aligned} (t^m \otimes x, t^n \otimes y) &= \delta_{m, -n}(x, y), (\delta, t^m \otimes x) = (d, t^m \otimes x) = 0, \\ (\delta, \delta) &= (d, d) = 0, (\delta, d) = 1. \end{aligned}$$

To prove this, it is obviously non-degenerate and symmetric, so one has to check that it is invariant and that its restriction to  $\mathfrak{h}$  is the right thing there.

The element  $\delta$  by the way is called the *canonical central element*, the element  $d$  is called the *scaling element*.

In class I will then give an explicit example in the case the graph is of type  $\tilde{A}_{n-1}$ . The Lie algebra  $\mathfrak{g}$  in this case is also known as  $\widehat{\mathfrak{sl}}_n$ . Particularly important is  $\widehat{\mathfrak{sl}}_2 \dots$

**Exercise 4.** Recall the Chevalley antiautomorphism  $\omega : \mathfrak{g} \rightarrow \mathfrak{g}$  swapping  $e_i, f_i$ 's and equal to the identity on  $\mathfrak{h}$ . What does  $\omega$  do to  $\delta, d$  and  $t^m \otimes x$  for  $x \in \mathfrak{g}$ ?

**3.7. Integrable representations.** Now we are more or less done with studying the Kac-Moody Lie algebras themselves, and can start studying representations. We will try to develop the general theory for an arbitrary Kac-Moody algebra, then say what is going on in the finite dimensional and affine cases.

So let  $\mathfrak{g} = \mathfrak{g}(\Gamma)$ . Recall that a  $\mathfrak{g}$ -module means the same thing as a  $U(\mathfrak{g})$ -module. The most important example is the adjoint representation of  $\mathfrak{g}$  on itself. Since this is usually infinite dimensional, we need to think about infinite dimensional modules usually! In fact most of the difficulties that we'll face will come from this infinite dimensionality... Some definitions:

- A  $\mathfrak{g}$ -module  $M$  is called *diagonalizable* if

$$M = \bigoplus_{\lambda \in \mathfrak{h}} M_\lambda$$

where

$$M_\lambda = \{m \in M \mid hm = (\lambda, h)m \text{ for all } h \in \mathfrak{h}\}.$$

We call  $M_\lambda$  the  $\lambda$ -*weight space* of  $M$ , and we call vectors in  $M_\lambda$  *weight vectors* of weight  $\lambda$ . For example, the adjoint representation is diagonalizable. The weight spaces in this case are the root spaces together with  $\mathfrak{h}$  ...

- A  $\mathfrak{g}$ -module  $M$  is called *integrable* if it is diagonalizable and all of the  $e_i$ 's and  $f_i$ 's act locally nilpotently, i.e. for each vector  $v \in M$  we have that  $e_i^N v = f_i^N v = 0$  for every  $i$  and  $N \gg 0$ .

If  $\mathfrak{g}$  is finite dimensional, then every *finite dimensional* representation is diagonalizable (not obvious but you can easily prove it starting from the  $\mathfrak{sl}_2$  case), integrable (obvious). However even in this case there are many infinite dimensional representations that are not diagonalizable, not integrable or not restricted. From now on ALL representations that we study here will be assumed diagonalizable.

The following Lemma implies that the adjoint representation  $\mathfrak{g}$  is an integrable module.

**Lemma 3.18.** *Let  $y_1, y_2, \dots$  be a system of generators for a Lie algebra  $\mathfrak{g}$ , and let  $x \in \mathfrak{g}$  be an element such that  $(\text{ad } x)^{N_i} y_i = 0$  for some positive integers  $N_i$ . Then,  $\text{ad } x$  is locally nilpotent on  $\mathfrak{g}$ .*

*Proof.* We have that

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

So if  $\text{ad } x$  is locally nilpotent on both  $y$  and  $z$ , it is also locally nilpotent on the commutator  $[y, z]$ . Now we get the lemma by induction on lengths of commutators...  $\square$

Let us now explain why integrable representations are nice. One of our basic tools as usual will be the use of the representation theory of  $\mathfrak{sl}_2$ , which you should know well by now. For each  $i$ , we have

$$\mathfrak{g}_{(i)} = \mathbb{C}e_i \oplus \mathbb{C}\epsilon_i \oplus \mathbb{C}f_i$$

which is a copy of  $\mathfrak{sl}_2$ . The good thing about integrable modules is the following:

**Lemma 3.19.** *Let  $M$  be an integrable  $\mathfrak{g}$ -module. Then, as a  $\mathfrak{g}_{(i)}$ -module,  $M$  decomposes as a direct sum of finite dimensional, irreducible  $\mathfrak{h}$ -invariant submodules.*

*Proof.* By the relations, we know that for  $m \in M_\lambda$ ,

$$e_i f_i^k(m) = k(1 - k - (\lambda, \epsilon_i)) f_i^{k-1} m + f_i^k e_i(m).$$

Hence, the subspace

$$U = \sum_{k,l \geq 0} f_i^k e_i^l m$$

is invariant under the action of  $\mathfrak{g}_{(i)} + \mathfrak{h}$ . As  $e_i, f_i$  are locally nilpotent,  $U$  is finite dimensional. Now we know that every finite dimensional representation of  $\mathfrak{g}_{(i)}$  is completely reducible (Exercise 10), hence the same is true for  $\mathfrak{g}_{(i)} + \mathfrak{h}$ . So we can decompose  $U$  as a direct sum of finite dimensional irreducible  $\mathfrak{h}$ -invariant  $\mathfrak{g}_{(i)}$ -modules. Since  $M$  is a sum of such  $U$ 's the same is true for  $M$ .  $\square$

Now remember the finite dimensional representation theory of  $\mathfrak{g}_{(i)}$ : there is a unique irreducible  $L(n)$  for each integer  $n$ , the eigenvalues of  $\epsilon_i$  on  $L(n)$  form a chain  $n, n-2, n-4, \dots, -n$ . In other words, the weights of  $L(n)$  are  $\frac{n\epsilon_i}{2}, \frac{n\epsilon_i}{2} - \epsilon_i, \dots, \frac{n\epsilon_i}{2} - n\epsilon_i$ . Say you are in the  $\lambda$  weight space of this chain. So  $\lambda$  is  $q$  steps from the top, or equivalently,  $p$  steps from the bottom, where  $p + q = n$ . Then,

$$(\lambda, \epsilon_i) = \left(\frac{n\epsilon_i}{2} - q\epsilon_i, \epsilon_i\right) = n - 2q = p - q.$$

Observe this is independent of  $n$ !!!! This is very important: it says if you have a vector in a  $\lambda$  weight space inside a finite dimensional irreducible representation of  $\mathfrak{sl}_2$  and you can apply  $e_i$  to the vector  $q$  times before you get zero,  $f_i$   $p$  times before you get zero, then  $p - q = (\lambda, \epsilon_i)$  *no matter what the vector is!*

This has profound implications on the structure of the weights of an integrable  $\mathfrak{g}$ -module  $M$ . Indeed, let  $\lambda \in \mathfrak{h}$  be any weight of  $M$ . Let

$$U = \sum_{k \in \mathbb{Z}} M_{\lambda + k\epsilon_i}.$$

This is invariant under the action of  $\mathfrak{g}_{(i)} + \mathfrak{h}$ , so it is a direct sum of finite dimensional  $\mathfrak{h}$ -invariant irreducible  $\mathfrak{g}_{(i)}$  submodules. So: the weights appearing in  $U$  are a union of strings of the form  $\lambda - p\epsilon_i, \lambda - (p-1)\epsilon_i, \dots, \lambda + (q-1)\epsilon_i, \lambda + q\epsilon_i$  where  $p, q$  are non-negative integers with  $p - q = (\lambda, \epsilon_i)$ . In particular, the weights in  $U$  are “symmetric” about the weight  $\lambda - (\lambda, \epsilon_i)/2$ . This shows:

**Lemma 3.20.** *Let  $M$  be an integrable  $\mathfrak{g}$ -module. If  $\lambda$  and  $\lambda + \epsilon_i$  are both weights of  $M$ , then,  $(\lambda, \epsilon_i) < 0$ . If  $\lambda$  and  $\lambda - \epsilon_i$  are both weights of  $M$ , then,  $(\lambda, \epsilon_i) > 0$ . Moreover,*

$$\dim M_\lambda = \dim M_{\lambda - (\lambda, \epsilon_i)\epsilon_i}$$

for every  $\lambda \in \mathfrak{h}$ .

This motivates the introduction of the Weyl group  $W$  in this setting: by definition,  $W$  is the subgroup of  $GL(\mathfrak{h})$  generated by the *fundamental reflections*  $s_i$  ( $i = 1, \dots, n$ ) defined from

$$s_i(\lambda) = \lambda - (\lambda, \epsilon_i)\epsilon_i$$

for each  $\lambda \in \mathfrak{h}$ . Note that  $W$  preserves the bilinear form on  $\mathfrak{h}$ , because this formula shows  $s_i$  is the reflection in the hyperplane orthogonal to  $\epsilon_i$  with respect to the form.

In view of the lemma, we have for any  $w \in W$  and an integrable module  $M$  that

$$\dim M_\lambda = \dim M_{w\lambda}.$$

Hence the Weyl group permutes the set of weights of an integrable  $\mathfrak{g}$ -module.

There is another way to think of this action of  $W$ . Let  $M$  be an integrable representation. Then it makes sense to consider the *automorphisms*

$$x_i(t) := \exp(te_i), y_i(t) := \exp(tf_i)$$

for any  $t \in \mathbb{C}$  because of the local finiteness. Note  $(x_i(t), x_i(t')) = x_i(t + t')$  and similarly for the  $y$ 's, so this generate an image of the additive from  $(\mathbb{C}, +)$ . The elements  $x_i(t), y_i(t)$  generate a quotient of the group  $SL_2(\mathbb{C})$  acting on  $M$ . To prove this, it suffices by the complete reducibility of  $M$  to understand that these things act in the same way as  $SL_2(\mathbb{C})$  on symmetric powers of the natural two dimensional representation... A consequence of this discussion is that we can consider the automorphism

$$S_i = \exp f_i \exp -e_i \exp f_i \in GL(M)$$

which induces the action of the  $s_i$  on the weight spaces. In class I'll calculate exactly which matrix it is in  $SL_2$ . This really motivates the language “integrable”: one can even consider the group of automorphisms of  $M$  generated by the  $x_i(t), y_i(t)$  for all  $i$ : this would be the *Kac-Moody group* corresponding to  $\Gamma$  in the representation  $M$ .

Now let us apply this discussion to the adjoint representation  $\mathfrak{g}$ . We've already seen it is integrable. So we deduce that  $W$  permutes the set  $\Delta$  of roots of  $\mathfrak{g}$ , and root spaces of roots in the same orbit have the same

dimension. In particular, if  $\alpha$  lies in the same  $W$ -orbit as one of the simple roots  $\epsilon_1, \dots, \epsilon_n$ , then we get at once that  $\dim \mathfrak{g}_\alpha = 1$  since we know that is the case already for the simple root spaces. This is the correct definition of a *real root* for arbitrary  $\Gamma$ : any root lying in the same  $W$ -orbit as a simple root. Thus for real roots,  $(\alpha, \alpha) = 2$  and the corresponding root space is always one dimensional. All other roots in  $\Delta$  are called *imaginary roots*. One can show for an imaginary root that  $(\alpha, \alpha) \leq 0$ . The following lemma summarizes all the familiar properties of real roots:

**Lemma 3.21.** *Let  $\alpha$  be a real root of  $\mathfrak{g}$ . Then,*

- (1)  $\dim \mathfrak{g}_\alpha = 1$ ,  $(\alpha, \alpha) = 2$ ;
- (2)  $k\alpha$  is a root if and only if  $k = \pm 1$ ;
- (3) if  $\beta \in \Delta$  then there exist non-negative integers  $p, q$  such that

$$(p - q)\alpha = (\beta, \alpha)$$

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

*Proof.* We can conjugate by  $W$  to assume that  $\alpha = \epsilon_i$ . Then we know (1) already, (2) follows since if  $k\epsilon_i$  is a root,  $k > 0$ , then it is a bracket of  $e_j$ 's since they generate  $\mathfrak{n}_+$ . But the only such bracket that has weight  $k\epsilon_i$  is  $\epsilon_i$  itself... For (3), think about the longest  $\mathfrak{g}_{(i)}$ -string passing through  $\beta$  – there is such a thing because the  $\beta$ -root space is finite dimensional so there are only finitely many  $\mathfrak{sl}_2$ -strings passing through  $\beta$ .  $\square$

**3.8. The Weyl group.** Now we study the properties of the Weyl group  $W$  in a little more detail. We already looked at this in the finite or Euclidean cases in the previous chapter, and it is exactly the *same* group in those cases. But there we viewed  $W$  only as acting on  $\mathfrak{h}'$ , so we need to extend some of our formulas in the Euclidean case to take care of the action on  $\mathfrak{d}$  too... At the same time we prove some general facts about the Weyl group *in the general case* – including all the facts we stated without proof in the previous chapter in the Dynkin and Euclidean cases.

Recall by definition that  $W$  is the subgroup of  $GL(\mathfrak{h})$  generated by the *fundamental reflections*  $s_i$  ( $i = 1, \dots, n$ ) defined from

$$s_i(\lambda) = \lambda - (\lambda, \epsilon_i)\epsilon_i$$

for each  $\lambda \in \mathfrak{h}$ . Note we really can think of  $s_i$  as a reflection: it fixes  $\epsilon_i^\perp$  which is a hyperplane in  $\mathfrak{h}$  since the form is non-degenerate and it maps  $\epsilon_i$  to  $-\epsilon_i$  – unlike in the previous chapter when I was calling  $s_i$  a reflection even though the form was degenerate. In particular,  $s_i^2 = 1$ !

**Lemma 3.22.** *The set  $\Delta_+ - \{\epsilon_i\}$  is invariant under the simple reflection  $s_i$ .*

*Proof.* Just note that if  $\beta \in \Delta_+ - \{\epsilon_i\}$ , then  $(\beta + \mathbb{Z}\epsilon_i) \cap \Delta \subseteq \Delta_+$ . This follows because  $\beta$  is not a multiple of  $\epsilon_i$ , so it must involve some other  $\epsilon_j$ . When you apply  $s_i$  you leave this  $\epsilon_j$  coordinate fixed, so still keep a positive root.  $\square$

Now we prepare to prove the main facts about the structure of the Weyl group in general.

**Lemma 3.23.** *Let  $\epsilon_i$  be a simple root and suppose that  $s_{i_1} \dots s_{i_t}(\epsilon_i) < 0$ . Then there exists  $r$  such that*

$$s_{i_1} \dots s_{i_r} \dots s_{i_t} s_i = s_{i_1} \dots s_{i_{r-1}} s_{i_{r+1}} \dots s_{i_t}.$$

*Proof.* Let  $\beta_k = s_{i_{k+1}} \dots s_{i_t}(\epsilon_i)$  for  $k < t$  and  $\beta_t = \epsilon_i$ . So  $\beta_0 < 0$  while  $\beta_t > 0$ , hence for some  $r$  we have that  $\beta_{r-1} < 0, \beta_r > 0$ . But  $\beta_{r-1} = s_{i_r}(\beta_r)$  so by the previous lemma we must have that  $\beta_r = \epsilon_{i_r}$ . Therefore,

$$\epsilon_{i_r} = w\epsilon_i$$

where  $w = s_{i_{r+1}} \dots s_{i_t}$ . This implies that

$$s_{i_r} = w s_i w^{-1}.$$

Multiplying both sides by  $s_{i_1} \dots s_{i_{r-1}}$  on the left and  $s_{i_{r+1}} \dots s_{i_t} s_i$  on the right gives that

$$s_{i_1} \dots s_{i_t} s_i = s_{i_1} \dots s_{i_{r-1}} s_{i_{r+1}} \dots s_{i_t} s_i$$

completing the proof.  $\square$

An expression  $w = s_{i_1} \dots s_{i_t}$  is called *reduced* if  $t$  is minimal amongst all such words representing  $w$ . Then  $t$  is called the *length* of  $w$ , denoted  $\ell(w)$ . The determinant of  $w$  acting on  $\mathfrak{h}$  is then  $(-1)^{\ell(w)}$  – which is very familiar in the type  $A$  case since it is just the signature of a permutation.

**Lemma 3.24.** *Let  $w = s_{i_1} \dots s_{i_t}$  be a reduced expression for  $w \in W$ . Let  $\epsilon_i$  be a simple root. Then,*

- (a)  $\ell(ws_i) < \ell(w)$  if and only if  $w(\epsilon_i) < 0$ .
- (b)  $w(\epsilon_{i_t}) < 0$ .
- (c) [Exchange condition] If  $\ell(ws_i) < \ell(w)$  then there exists  $r$  such that

$$s_{i_r} s_{i_{r+1}} \dots s_{i_t} = s_{i_{r+1}} \dots s_{i_t} s_i.$$

*Proof.* By the previous lemma,  $w(\epsilon_i) < 0$  implies that  $\ell(ws_i) < \ell(w)$ . Conversely, if  $w(\epsilon_i) > 0$  then  $ws_i(\epsilon_i) < 0$  hence  $\ell(w) = \ell(ws_i^2) < \ell(ws_i)$ . This proves (a). (b) is immediate from (a). For (c), if  $\ell(ws_i) < \ell(w)$  then we have by (a) that  $w(\epsilon_i) < 0$ , hence by the previous lemma there exists  $r$  such that

$$s_{i_1} \dots s_{i_t} s_i = s_{i_1} \dots s_{i_{r-1}} s_{i_{r+1}} \dots s_{i_t}.$$

Moving things around gives the resulting formula.  $\square$

Now we obtain a geometric way of thinking about the Weyl group. We need to descend to do this to the real numbers: let

$$E' = \mathbb{R}\epsilon_1 \oplus \dots \oplus \mathbb{R}\epsilon_n$$

equipped with the usual bilinear form, and then extend to a real vector space  $E$  with a non-degenerate symmetric bilinear form extending that on  $E'$  so that  $\mathfrak{h} = \mathbb{C} \otimes_{\mathbb{R}} E$ . I think of  $E$  as a real form the vector space  $\mathfrak{h}$ . The Weyl

group  $W$  still acts on  $E$  by the same formula, except now it is acting by *real* reflections!! So it really is correct now to call  $W$  a reflection group. Let

$$C = \{v \in E \mid (\epsilon_i, v) \geq 0 \text{ for all } i = 1, \dots, n\}.$$

This is the fundamental chamber. The sets  $wC$  for  $w \in W$  are called the *chambers*, and their union

$$X = \bigcup_{w \in W} wC$$

is called the *Tits cone*.

- Theorem 3.25.** (a) For  $v \in E$ ,  $\text{Stab}_W(v)$  is generated by the fundamental reflections that it contains.  
 (b) The fundamental chamber  $C$  is a set of orbit representatives for the action of  $W$  on  $X$ .  
 (c)  $C = \{v \in E \mid \text{for every } w \in W, v - wv = \sum_i c_i \epsilon_i \text{ with } c_i \geq 0\}$ .

*Proof.* Let  $w \in W$  and  $w = s_{i_1} \dots s_{i_t}$  be a reduced expression. Take  $v \in C$  and suppose that  $v' = wv \in C$  too. Since  $(\epsilon_{i_t}, v) \geq 0$ , we also have that  $(w(\epsilon_{i_t}), v') \geq 0$ . But since it is a reduced expression,  $w(\epsilon_{i_t}) < 0$  hence  $(w(\epsilon_{i_t}), v') \leq 0$ . So we get that  $(\epsilon_{i_t}, v) = 0$ . Hence,  $s_{i_t}(v) = v$ . Replacing  $w$  by  $s_{i_1} \dots s_{i_{t-1}}$  and repeating, we get that  $s_{i_j}v = v$  for each  $j$ . Hence all  $s_{i_1}, \dots, s_{i_t}$  fix  $v$  and  $v = v'$ . Parts (a) and (b) follow at once.

For (c), the inclusion  $\supseteq$  is obvious: if  $v$  is in the RHS,  $v - s_i(v) = (v, \epsilon_i)\epsilon_i$  is a positive sum of  $\epsilon_j$ 's hence  $(v, \epsilon_i) \geq 0$ . For the reverse inclusion, we prove that  $v - wv$  is a positive sum of  $\epsilon_j$ 's for each  $v \in C$  by induction on  $\ell(w)$ . If  $\ell(w) = 1$ , this is the definition of  $C$ . For  $\ell(w) > 1$ , let  $w = s_{i_1} \dots s_{i_t}$  be a reduced expression. Then,

$$v - w(v) = (v - s_{i_1} \dots s_{i_{t-1}}(v)) + s_{i_1} \dots s_{i_{t-1}}(v - s_{i_t}v).$$

The first sum is a positive sum of  $\epsilon_j$ 's by induction. The second sum is

$$(v, \epsilon_{i_t})s_{i_1} \dots s_{i_{t-1}}(\epsilon_{i_t}) = -(v, \epsilon_{i_t})w(\epsilon_{i_t})$$

and we know that is  $> 0$  by (b) of the previous lemma.  $\square$

The most important thing here is (b): observe it proves something we used a lot in the previous chapter in the finite case.

**Corollary 3.26.** The map  $w \mapsto wC$  is a bijection between  $W$  and the set of chambers.

*Proof.* The stabilizer of a point from the interior of a chamber is trivial by (a) and (b). So the map is injective. It is surjective by definition of chambers.  $\square$

**Corollary 3.27.** Let  $\alpha \in \Delta_+$  be a real root. Then there exists  $i$  such that  $s_i(\alpha)$  is of strictly smaller height than  $\alpha$  itself.

*Proof.* Suppose not. Then,  $s_i(\alpha) = \alpha - (\alpha, \epsilon_i)\epsilon_i$  is of height  $\geq ht(\alpha)$  for all  $i$ , hence  $(\alpha, \epsilon_i) \leq 0$  for all  $i$ . This means that  $-\alpha \in C$ . Hence by (c) of the theorem,  $-\alpha + w\alpha \geq 0$  for all  $w \in W$ . But we can pick  $w$  such that  $w\alpha$  is a simple root to get a contradiction.  $\square$

One can show with only a tiny bit more work that  $X = E$  if and only if  $W$  is finite, which is if and only if  $\Gamma$  is a Dynkin diagram. But I think we've now proved all of the things we are going to need about  $W$  later on in the general case. But it would be remiss if I didn't at least state the following theorem, which doesn't take that much more work to prove now that we've established the deletion condition:

**Theorem 3.28.** *The group  $W$  is generated by the elements  $s_1, \dots, s_n$  subject ONLY to the following relations:*

$$s_i^2 = 1, (s_i s_j)^{m_{i,j}} = 1$$

where

$$m_{i,j} = 2, 3, \infty$$

according to whether

$$n_{i,j} = 0, 1, \geq 2.$$

The Theorem shows that  $W$  is a *Coxeter group*, which is a class of groups defined by such generators and relations.

**Exercise 5.** Suppose that  $\Gamma$  is the Dynkin diagram  $A_{n-1}$ . We know in this case that the Weyl group  $W$  is the symmetric group  $S_n$  on generators  $s_1, \dots, s_{n-1}$  where  $s_i = (i, i+1)$  and the Euclidean space  $E$  is the subspace of the standard Euclidean space  $\mathbb{R}^n = \mathbb{R}v_1 \oplus \dots \oplus \mathbb{R}v_n$  spanned by  $\epsilon_i = v_i - v_{i+1}$  for  $i = 1, \dots, n-1$ . Use Lemma 3.24 above to show that for  $w \in S_n$ ,

$$\ell(w)$$

is the number of pairs  $(i, j)$  with  $1 \leq i < j \leq n$  and  $w(i) > w(j)$ .

In the Dynkin cases,  $W$  is exactly the finite reflection group we discussed in chapter 2. Let me end by recalling what we know about the Euclidean case. We showed in chapter 2 in these cases that  $W$  was the semidirect product of the translation group  $T$  by  $\mathring{W} = \langle s_1, \dots, s_n \rangle$ , where  $T$  was the group  $\{t_\alpha \mid \alpha \in \mathring{R}\}$  and  $t_\alpha(\lambda) = \lambda - (\lambda, \alpha)\delta$  as a translation  $t_\alpha : E' \rightarrow E'$ . I mention this because in chapter 2 we only discussed  $W$  acting on  $E' = \bigoplus \mathbb{R}\epsilon_0 \oplus \dots \oplus \mathbb{R}\epsilon_n$ . Now we have one extra dimension tacked on, because  $E = E' \oplus \mathbb{R}d$ . The Weyl group  $W$  is the same here as before – but to understand its action on the bigger space I need to write down the formula for how  $t_\alpha : E \rightarrow E$  acts on  $d$  as well. Here is the formula:

$$t_\alpha(d) = d + \alpha - \frac{(\alpha, \alpha)}{2}\delta$$

for  $\alpha \in \mathring{R}$ . It looks mysterious, but you discover it in the same way as we worked out the structure of  $W$  back in chapter 2.

**3.9. Highest weight theory.** After that digression about the Weyl group, lets get back to representation theory. Let  $\mathfrak{g} = \mathfrak{g}(\Gamma)$ . Recall the triangular decomposition:

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

Now consider the universal enveloping algebras. Pick a basis for  $\mathfrak{n}_-$ , for  $\mathfrak{h}$  and for  $\mathfrak{n}_+$ : together they form a basis for  $\mathfrak{g}$ . By the PBW theorem,  $\mathfrak{g}$  has a basis consisting of all monomials in this ordered basis, as does  $\mathfrak{n}_-, \mathfrak{h}, \mathfrak{n}_+$ . Hence multiplication induces a vector space isomorphism

$$U(\mathfrak{n}_-) \otimes U(\mathfrak{h}) \otimes U(\mathfrak{n}_+) \rightarrow U(\mathfrak{g}).$$

We pick these bases so that they are compatible always with the root space decomposition.

Remember also that  $\mathfrak{h}$  is abelian, so  $U(\mathfrak{h})$  is commutative  $\cong S(\mathfrak{h})$ , and (by the Nullstellensatz) the irreducible  $U(\mathfrak{h})$ -modules (a.k.a. the maximal ideals) are parametrized by the points in the dual space  $\mathfrak{h}^*$  which we have identified by  $\mathfrak{h}$ , i.e. they are the modules  $\mathbb{C}_\lambda = \mathbb{C}v_\lambda$  for each  $\lambda \in \mathfrak{h}$ , where  $h.v_\lambda = (\lambda, h)v_\lambda$ . We are only considering *diagonalizable* representations, meaning the ones that decompose over  $\mathfrak{h}$  as a direct sum

$$M = \bigoplus_{\lambda \in \mathfrak{h}} M_\lambda$$

of weight spaces: so  $M_\lambda$  here is a direct sum of copies of the irreducible  $\mathbb{C}_\lambda$ . In other words, the diagonalizable representations are the ones that are completely reducible as  $\mathfrak{h}$ -modules.

Let  $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}^+$ , and call it the *Borel subalgebra*. Note that  $\mathfrak{n}^+$  is an ideal in  $\mathfrak{b}$  by the relations. So we can extend  $\mathbb{C}_\lambda$  to a  $\mathfrak{b}$ -module by declaring that  $\mathfrak{n}^+$  acts as zero. Now form the *Verma module*

$$M(\lambda) := U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} \mathbb{C}_\lambda,$$

for each  $\lambda \in \mathfrak{h}$ . We also write  $v_\lambda$  for the vector  $1 \otimes v_\lambda \in M(\lambda)$ . This is a weight vector of weight  $\lambda$  by definition. By the triangular decomposition above,  $U(\mathfrak{g})$  is a free right  $U(\mathfrak{b})$ -module on basis given by the set of all monomials in the ordered basis we picked above for  $\mathfrak{n}_-$ . So these same monomials applied to  $v_\lambda$  give a basis for  $M(\lambda)$ . In other words,  $M(\lambda)$  is a free  $U(\mathfrak{n}_-)$ -module of rank 1, on basis  $v_\lambda$ . Consider for example the monomial  $b_1 \dots b_N$  where  $b_i \in \mathfrak{g}_{-\alpha_i}$  for  $\alpha_i \in \Delta^+$ . Then,

$$\begin{aligned} h(b_1 \dots b_N v_\lambda) &= (b_1 \dots b_N) h v_\lambda + [h, b_1 \dots b_N] v_\lambda \\ &= (h, \lambda) b_1 \dots b_N v_\lambda + \sum_{i=1}^N b_1 \dots [h, b_i] \dots b_N v_\lambda \\ &= ((h, \lambda) - \sum_{i=1}^N (h, \alpha_i)) b_1 \dots b_N v_\lambda \end{aligned}$$

So it is of weight  $\lambda - \alpha_1 - \cdots - \alpha_N$ . This shows that  $M(\lambda)$  is *diagonalizable* and that its weight spaces are all *finite dimensional*. Moreover, the  $\lambda$ -weight space  $M(\lambda)_\lambda$  is one dimensional, spanned by  $v_\lambda$ , and all the other weights appearing are  $< \lambda$  in the dominance ordering.

We call a  $\mathfrak{g}$ -module a *highest weight module* of *highest weight*  $\lambda$  if it is generated by a *highest weight vector* of weight  $\lambda$ , i.e.  $v_\lambda$  of weight  $\lambda$  and is annihilated by each  $e_i$ . We have just seen that  $M(\lambda)$  is such a module. In fact it is the universal such module:

**Lemma 3.29.** *Suppose that  $M$  is a highest weight module of highest weight  $\lambda$ . Then there is a unique up to scalars surjective homomorphism from  $M(\lambda)$  to  $M$ .*

*Proof.* By adjointness of tensor and hom,

$$\mathrm{Hom}_{\mathfrak{g}}(U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} \mathbb{C}_\lambda, M) \cong \mathrm{Hom}_{\mathfrak{b}}(\mathbb{C}_\lambda, M).$$

Since  $M$  is generated by some highest weight vector  $m$  of weight  $\lambda$ , we have that  $M = U(\mathfrak{n}_-)m$ , hence  $M_\lambda = \mathbb{C}\lambda$  and all other weight spaces are lower. So this space of homomorphisms is exactly one dimensional. Moreover, the image of any such homomorphism sends  $v_\lambda$  to a non-zero multiple of  $m$ , which generates  $M$  so it is onto.  $\square$

**Lemma 3.30.**  *$M(\lambda)$  has a unique maximal submodule  $\mathrm{rad} M(\lambda)$ , and the quotient  $M(\lambda)/\mathrm{rad} M(\lambda)$  is an irreducible module  $L(\lambda)$  of highest weight  $\lambda$ . Any other irreducible module of highest weight  $\lambda$  is isomorphic to  $L(\lambda)$ , by an isomorphism that is unique up to scalars. In particular,*

$$\dim \mathrm{End}_{\mathfrak{g}}(L(\lambda)) = 1.$$

*Proof.* If  $M$  is a proper submodule of  $M(\lambda)$ , we must have that  $M_\lambda = 0$ , since any non-zero vector in the  $\lambda$ -weight space of  $M(\lambda)$  is a scalar multiple of  $v_\lambda$  so already generates all of  $M(\lambda)$ . Hence, the sum of all the proper submodules of  $M(\lambda)$  also has  $\lambda$ -weight space equal to 0, so is still proper. This proves the existence of a unique maximal submodule. Now suppose that  $L$  is an irreducible module of highest weight  $\lambda$ . By the previous lemma,

$$\mathrm{Hom}_{\mathfrak{g}}(M(\lambda), L)$$

is one dimensional. Pick any non-zero  $f : M(\lambda) \rightarrow L$ . Since  $L$  is irreducible, its kernel is a maximal submodule of  $M(\lambda)$ , hence equals  $\mathrm{rad} M(\lambda)$ , so  $f$  factors to induce an isomorphism  $L(\lambda) \rightarrow L$ .  $\square$

Thus we have constructed a nice family of irreducible modules  $L(\lambda)$  parametrized by the space  $\mathfrak{h}$ . We know  $L(\lambda)_\lambda$  is one dimensional and  $L(\lambda)_\mu = 0$  unless  $\mu \leq \lambda$ . Define

$$D(\lambda) = \{\mu \in \mathfrak{h} \mid \mu \leq \lambda\} :$$

the weights of  $M(\lambda)$  or  $L(\lambda)$  are subsets of  $D(\lambda)$ . We now introduce an appropriate category of  $\mathfrak{g}$ -modules for which all the irreducible modules are highest weight modules, i.e.  $L(\lambda)$ 's.

The category  $\mathcal{O}$  is the category consisting of all  $\mathfrak{g}$ -modules  $M$  that are diagonalizable, have finite dimensional weight spaces, and for which there exist elements  $\lambda_1, \dots, \lambda_N \in \mathfrak{h}$  such that the weights of  $M$  belong to  $D(\lambda_1) \cup \dots \cup D(\lambda_N)$ . (“Finitely many roofs”). Morphisms in  $\mathcal{O}$  are all  $\mathfrak{g}$ -module homomorphisms.

Observations: and submodule or quotient module of something in category  $\mathcal{O}$  is in  $\mathcal{O}$ , direct sums or tensor products of finitely many modules in  $\mathcal{O}$  is in  $\mathcal{O}$ .

One more piece of terminology: we call a vector  $m$  in a  $\mathfrak{g}$ -module  $M$  a *primitive vector* of weight  $\lambda$  if there exists a submodule  $M'$  of  $U(\mathfrak{g})m$  such that the image of  $m$  in  $U(\mathfrak{g})m/M'$  is a highest weight vector of weight  $\lambda$ .

**Lemma 3.31.** *Let  $M \in \mathcal{O}$ .*

- (a)  *$M$  contains a non-zero highest weight vector, i.e. a weight vector annihilated by the  $e_i$ 's.*
- (b)  *$M$  is irreducible if and only if  $M \cong L(\lambda)$  for some  $\lambda$ .*
- (c)  *$M$  is generated by its primitive vectors as a  $\mathfrak{g}$ -module.*

*Proof.* (a) By the definition of a module in category  $\mathcal{O}$ , we can pick a maximal weight of  $M$  in the dominance ordering. Then any vector in that weight space will be such.

(b) If  $M$  is irreducible, pick a highest weight vector  $m_\lambda \in M$  of weight  $\lambda$  using (a). By irreducibility,  $M = U(\mathfrak{g})m_\lambda$  so it is an irreducible highest weight of highest weight  $\lambda$ . By the previous lemma,  $M \cong L(\lambda)$ . Conversely,  $L(\lambda)$  is an irreducible in category  $\mathcal{O}$ .

(c) Let  $M'$  be the submodule in  $M$  generated by all the primitive vectors. If  $M' \neq M$ , then  $M/M'$  contains a highest weight vector, any preimage of which in  $M$  is a primitive vector. Contradiction.  $\square$

Note this shows that the  $L(\lambda)$ 's give all the irreducibles in  $\mathcal{O}$ .

**3.10. Formal characters.** Unfortunately, a module  $V$  in  $\mathcal{O}$  need not admit a composition series. So we cannot define things like composition multiplicities  $[M : L(\lambda)]$  in the usual way. The following provides a substitute for this:

**Lemma 3.32.** *Let  $M \in \mathcal{O}$  and  $\lambda \in \mathfrak{h}$ . Then there exists a filtration*

$$M = M_t \supset M_{t-1} \supset \dots \supset M_0 = 0$$

*and a subset  $J \subset \{1, \dots, t\}$  such that*

- (i) *if  $j \in J$  then  $M_j/M_{j-1} \cong L(\lambda_j)$  for some  $\lambda_j \geq \lambda$ ;*
- (ii) *if  $j \notin J$  then  $(M_j/M_{j-1})_\mu = 0$  for every  $\mu \geq \lambda$ .*

*Proof.* Let

$$a(M, \lambda) = \sum_{\mu \geq \lambda} \dim M_\mu.$$

This is well-defined by the definition of category  $\mathcal{O}$ . We prove the lemma by induction on  $a(M, \lambda)$ . If  $a(M, \lambda) = 0$  then  $0 = M_0 \subset M_1 = M$  is the

required filtration! If  $a(M, \lambda) > 0$ , let  $\mu$  be a maximal weight of  $M$  such that  $\mu \geq \lambda$ . Choose a non-zero vector  $m \in M_\mu$  and let  $U = U(\mathfrak{g})m$ . Clearly  $U$  is a highest weight module. Hence it has a unique maximal submodule  $\text{rad } U$ . Now we have

$$0 \subset \text{rad } U \subset U \subset M$$

with  $U/\text{rad } U \cong L(\mu)$  and  $\mu \geq \lambda$ . Since  $a(\text{rad } U, \lambda) < a(M, \lambda)$  and also  $a(M/U, \lambda) < a(M, \lambda)$  we now can proceed by induction.  $\square$

Now fix  $\mu \geq \lambda$  and construct a filtration like in the lemma. Denote by  $[M : L(\mu)]$  the number of times  $\mu$  appears among the  $\{\lambda_j \mid j \in J\}$ . This number is independent of the choice of filtration: for given two such filtrations you can form a common refinement like in the proof of the Jordan-Holder theorem. Moreover, this number is independent of the particular choice of  $\lambda$ : given two such  $\lambda$ 's you can find a third lower than them both and now you see that you can always lower  $\lambda$  to preserve these multiplicities. We call  $[M : L(\mu)]$  the *multiplicity* of  $L(\mu)$  in  $M$ . In case  $M$  DOES have a composition series it is the same as the usual composition multiplicity.

Now we are ready to introduce a crucial numerical tool: the notion of character of a module in category  $\mathcal{O}$ . The way we're going to treat these, they are just some numerical invariant associated to a module  $M$  which, although not fine enough to characterize  $M$  as a module, is good enough to show that if  $M$  and  $N$  have the same characters, they have the same composition multiplicities.

Given a module  $M$  in category  $\mathcal{O}$ , we have by definition that all its weight spaces are finite dimensional. The idea of character is to record the dimensions of each of the weight spaces. Since there may be infinitely many weight spaces, we are going to have to work with certain formal infinite sums recording all this information. But it is no big deal . . .

So let  $\mathcal{E}$  be the  $\mathbb{C}$ -algebra whose elements are series of the form

$$\sum_{\lambda \in \mathfrak{h}} c_\lambda e^\lambda$$

where  $c_\lambda \in \mathbb{C}$  and  $c_\lambda = 0$  for  $\lambda$  outside the union of a finite number of sets of the form  $D(\mu)$ . The  $e^\lambda$ 's here should just be treated as formal symbols (though there is another way to view characters through which the  $e^\lambda$ 's really are exponential functions converging absolutely on some convex subset of the space  $\mathfrak{h}$ ). The sum of two formal characters is the obvious coordinatewise thing, the product of two formal characters also makes sense starting from  $e^\lambda e^\mu = e^{\lambda+\mu}$  (the point is to calculate the coefficient of a given  $e^\nu$  in the product of two elements if  $\mathcal{E}$  involves only calculating a finite sum).

Given a module  $M \in \mathcal{O}$ , let

$$\text{ch } M = \sum_{\lambda \in \mathfrak{h}} (\dim M_\lambda) e^\lambda.$$

By the definition of  $\mathcal{O}$  this is a well-defined element of  $\mathcal{O}$ . Also note

$$\text{ch}(M \oplus N) = \text{ch } M + \text{ch } N, \text{ch}(M \otimes N) = \text{ch } M \cdot \text{ch } N.$$

The basic result showing that characters really do characterize modules is:

**Lemma 3.33.** *Let  $M \in \mathcal{O}$ . Then,*

$$\text{ch } M = \sum_{\lambda \in \mathfrak{h}} [M : L(\lambda)] \text{ch } L(\lambda).$$

*Proof.* Let  $\phi$  be the map which sends  $M \in \mathcal{O}$  to the difference  $\phi(M)$  between the LHS and the RHS of the formula we're trying to prove. So  $\phi(L(\lambda)) = 0$ . Moreover, given a short exact sequence of modules

$$0 \rightarrow M_1 \rightarrow M_2 \rightarrow M_3 \rightarrow 0$$

we have by the definition of composition multiplicities that  $\phi(M_2) = \phi(M_1) + \phi(M_3)$ . Now let us focus on a particular  $\lambda$  and  $M$ . There exists a filtration

$$M = M_t \supset \cdots \supset M_0 = 0$$

such that, setting  $N_i = M_i/M_{i-1}$ , we have that either  $N_i \cong L(\lambda_i)$  for some  $\lambda_i \geq \lambda$  or that  $(N_i)_\lambda = 0$ . In the former case,  $\phi(N_i) = 0$ . In the latter case,  $\phi(N_i)$  has  $e^\lambda$ -coefficient equal to zero. Hence,  $\phi(M) = \sum \phi(N_i)$  has  $\lambda$ -coefficient equal to zero. This was for all  $\lambda$ , so it is actually zero.  $\square$

In the rest of the section, I am going to work out a load of examples.

**Example 3.34.** (1) Consider the adjoint representation  $\mathfrak{g}$ . We already observed that it was integrable, so it was a good representation from our old point of view. However it *may not be* in category  $\mathcal{O}$ : it fails the finitely many roots condition even in the affine case. In the Dynkin case it is in category  $\mathcal{O}$  however, as all finite dimensional representations are, for  $\Gamma$  Dynkin. and its character is

$$\text{ch } \mathfrak{g} = n + \sum_{\alpha \in \Delta} e^\alpha.$$

Since  $\mathfrak{g}$  is simple, this is an irreducible representation in category  $\mathcal{O}$ , so it should be one of the  $L(\lambda)$ 's. To work out which, we identify the highest root: it is  $\theta = \sum_{i=1}^n \delta_i \epsilon_i$ . Thus,  $\mathfrak{g} = L(\theta)$  in the Dynkin case. In particular, observe that all other elements of  $\Delta$  are  $< \theta$  in the dominance ordering – giving another sense in which  $\theta$  is the maximal root.

(2) Consider next the graph  $\Gamma = A_{n-1}$ , so  $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ . We know that  $\mathfrak{h}$  is the space  $\mathbb{C}\epsilon_1 \oplus \cdots \oplus \mathbb{C}\epsilon_{n-1}$  with  $(\epsilon_i, \epsilon_j) = 0$  if  $|i - j| > 1$ ,  $-1$  if  $|i - j| = 1$  and  $2$  if  $|i - j| = 0$ . As usual, it is useful to work with  $\mathfrak{h}$  by working in a bigger space  $\mathbb{C}\nu_1 \oplus \cdots \oplus \mathbb{C}\nu_n$  with  $(\nu_i, \nu_i) = \delta_{i,j}$  then  $\epsilon_i = \nu_i - \nu_{i+1}$ . Note the  $\nu_i$ 's themselves do not belong to  $\mathfrak{h}$ , so let  $\bar{\nu}_i = \nu_i - \frac{1}{n}(\nu_1 + \cdots + \nu_n)$  which does. The point is that  $(\nu_i, h) = (\bar{\nu}_i, h)$  for all  $h \in \mathfrak{h}$  since such  $h$  have trace zero. If we let

$$\Lambda_i = \bar{\nu}_1 + \cdots + \bar{\nu}_i$$

then we have that  $(\Lambda_i, \epsilon_j) = \delta_{i,j}$ . These are called the *fundamental dominant weights*, and are convenient for writing down weights. I am going to compute

the characters of some finite dimensional representations – recall any finite dimensional representation in the Dynkin case is necessarily in category  $\mathcal{O}$ .

Consider first the natural  $n$ -dimensional module  $V$  on basis  $v_1, \dots, v_n$ . We have that  $(h, v_i) = (h, \nu_i)v_i$ . Hence  $v_i$  is of weight  $\bar{\nu}_i$ . So

$$\text{ch } V = e^{\bar{\nu}_1} + \dots + e^{\bar{\nu}_n}.$$

It is tidier to set  $x_i = e^{\bar{\nu}_i}$  for short. Then the character of any finite dimensional module lies in the polynomial ring  $\mathbb{C}[x_1, \dots, x_n] \subset \mathcal{E}$ . (Of course in this ring  $x_1 \dots x_n = 1$ .) Since such modules are integrable, the symmetric group  $S_n$  (the Weyl group) permutes the weight spaces around, so such characters are invariant under  $S_n$  permuting  $x_1, \dots, x_n$ , i.e. they are symmetric functions belonging to  $\mathbb{C}[x_1, \dots, x_n]^{S_n}$ .

What about the  $d$ th exterior power of  $V$ ? Adopting the above conventions, its character is

$$\text{ch } \bigwedge^d V = \sum_{1 \leq i_1 < \dots < i_d \leq n} x_{i_1} \dots x_{i_d},$$

usually called the  $d$ th elementary symmetric function. By the way,  $\bigwedge^d V$  is an irreducible representation of highest weight  $\Lambda_d$ . To see it is irreducible, note that the Weyl group in this case acts *transitively* on the set of weights, and all the weight spaces are irreducible.

What about the  $d$ th symmetric power of  $V$ ? Its character is

$$\text{ch } S^d V = \sum_{1 \leq i_1 \leq \dots \leq i_d \leq n} x_{i_1} \dots x_{i_d}.$$

This is the  $d$ th complete symmetric function. You can also show that  $S^d V$  is irreducible, but there isn't a cheap trick like before. But for instance you can simply check that any vector can produce any other. So this is  $L(d\Lambda_1)$ .

Let's now work out the composition multiplicities of  $V \otimes V$ . Its character is  $(x_1 + \dots + x_n)^2 = \sum_{i \leq j} x_i x_j + \sum_{i < j} x_i x_j = \text{ch } L(\Lambda_2) + \text{ch } L(2\Lambda_1)$ . Hence by the above lemma its composition factors are  $S^2 V$  and  $\bigwedge^2 V$ . Of course this is rather obvious in this case, but it illustrates how characters are a good tool for calculating.

(3) Return to the general case. Let us compute the character of the Verma module  $M(\lambda)$  for any  $\lambda \in \mathfrak{h}$ . Recall that  $M(\lambda)$  is a free  $U(\mathfrak{n}_-)$ -module on basis  $v_\lambda$ . Let  $\beta_1, \beta_2, \dots$  be all the positive roots of  $\mathfrak{g}$ . Let  $m_\alpha = \dim \mathfrak{g}_\alpha$  denote the dimension of the root space. We can pick a basis  $e_{-\beta_s, 1}, \dots, e_{-\beta_s, m(\beta_s)}$  for  $\mathfrak{g}_{-\beta_s}$ . Then the vectors

$$e_{-\beta_1, 1}^{n_{1,1}} \dots e_{-\beta_1, m(\beta_1)}^{n_{1, m(\beta_1)}} e_{-\beta_2, 1}^{n_{2,1}} \dots e_{-\beta_2, m(\beta_2)}^{n_{2, m(\beta_2)}} \dots v_\lambda$$

such that  $(n_{1,1} + n_{1,2} + \dots)\beta_1 + (n_{2,1} + n_{2,2} + \dots)\beta_2 + \dots = \lambda - \mu$  form a basis for  $M(\lambda)_\mu$ . Therefore, rewriting the sum as a product,

$$\text{ch } M(\lambda) = e^\lambda \prod_{\alpha \in \Delta_+} (1 + e^{-\alpha} + e^{-2\alpha} + \dots)^{m(\alpha)}.$$

It makes sense to write instead

$$\text{ch } M(\lambda) = \frac{e^\lambda}{\prod_{\alpha \in \Delta_+} (1 - e^{-\alpha})^{m(\alpha)}}$$

since the elements  $1 - e^{-\alpha}$  are invertible in the ring  $\mathcal{E}$  for each  $\alpha \in \Delta_+$ .

**Exercise 6.** Continue with the above example of  $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ . Compute the character of  $V^*$ , the dual of the natural module. Hence show that  $V^* \cong \bigwedge^{n-1} V$  and prove that the tensor product  $V \otimes V^*$  has exactly two composition factors. What are their highest weights?

**3.11. The Casimir operator.** Let  $\rho$  is some fixed element of  $\mathfrak{h}$  such that  $(\rho, \epsilon_i) = 1$  for each  $i = 1, \dots, n$ . Note  $\rho$  is not necessarily uniquely determined unless the  $\epsilon_i$  form a basis for  $\mathfrak{h}$ . In the Dynkin case there is a very nice formula for  $\rho$  which motivates its introduction as we'll see:

**Lemma 3.35.** *Suppose  $\Gamma$  is Dynkin. Then  $\rho = \frac{1}{2} \sum_{\alpha \in \Delta_+} \alpha$ .*

*Proof.* Let  $\rho = \frac{1}{2} \sum_{\alpha \in \Delta_+} \alpha$ . We need to show that  $(\rho, \epsilon_i) = 1$  for each  $i = 1, \dots, n$ . Equivalently, we need to show that  $s_i(\rho) = \rho - \epsilon_i$  for each  $i$ . But this is obvious since  $s_i$  sends  $\epsilon_i$  to  $-\epsilon_i$  and it permutes all the other positive roots amongst themselves.  $\square$

Of course we couldn't hope for such a formula in any other case, because there are infinitely many positive roots!!!

The goal in this section is to prove the following fundamental theorem:

**Theorem 3.36.** *Let  $V$  be a  $\mathfrak{g}$ -module with highest weight  $\lambda$ . Then,*

$$\text{ch } V = \sum_{\mu} c_{\mu} \text{ch } M(\mu)$$

where  $c_{\mu} \in \mathbb{Z}$ ,  $c_{\lambda} = 1$ , and the sum is over  $\mu \leq \lambda$  with  $(\lambda + \rho, \lambda + \rho) = (\mu + \rho, \mu + \rho)$ .

To prove this we need to introduce the Casimir operator  $\Omega$ . This will be some operator that acts on any integrable  $\mathfrak{g}$ -module and commutes with the action of  $\mathfrak{g}$ . Moreover, we will show that if  $v$  is a highest weight vector of weight  $\lambda$ , then

$$\Omega v = ((\lambda + \rho, \lambda + \rho) - (\rho, \rho))v$$

Before I explain the construction of the Casimir operator, let us see how the theorem follows from this. First of all it suffices by Lemma 3.33 to consider the case that  $V$  is the irreducible representation  $L(\lambda)$ . On that  $\Omega$  must act as a scalar by Schur's lemma, namely, the scalar  $(\lambda + \rho, \lambda + \rho) - (\rho, \rho)$ . We know that

$$\text{ch } M(\mu) = \sum_{\nu} [M(\mu) : L(\nu)] \text{ch } L(\nu),$$

applying Lemma 3.33, with  $c_{\mu, \mu} = 1$ . Since  $\Omega$  acts as the scalar  $(\mu + \rho, \mu + \rho) - (\rho, \rho)$  on  $M(\mu)$  we get that  $[M(\mu) : L(\nu)] = 0$  unless  $(\mu + \rho, \mu + \rho) =$

$(\nu + \rho, \nu + \rho)$ . Let  $B = \{\mu \leq \lambda \mid (\mu + \rho, \mu + \rho) = (\lambda + \rho, \lambda + \rho)\}$ . Order the elements of  $B$   $\lambda_1, \lambda_2, \dots$  so that  $\lambda_i \geq \lambda_j$  implies  $i \leq j$ . Then,

$$\text{ch } M(\lambda_i) = \sum_j c_{i,j} \text{ch } L(\lambda_j).$$

Moreover,  $c_{i,i} = 1$  and  $c_{i,j} = 0$  for  $i > j$ . So we can solve this system of linear equations to complete the proof of the theorem.

So in the remainder of the section, we need to construct the Casimir operator  $\Omega$ . Before we do this, some motivation. Suppose first that  $\mathfrak{g}$  is a finite dimensional Lie algebra with non-degenerate symmetric bilinear form  $(\cdot, \cdot)$ . Let  $x_1, \dots, x_n$  be a basis, let  $y_1, \dots, y_n$  be the dual basis. The classical Casimir operator is the element

$$\Omega = \sum x_i y_i \in U(\mathfrak{g}).$$

(One can easily check that its definition is independent of the particular choice of bases). Claim:  $\Omega$  is central, hence its action on any module commutes with the action of  $\mathfrak{g}$ .

Let us prove the claim. Pick any  $z \in \mathfrak{g}$ . Write

$$[z, x_i] = \sum_j a_{i,j} x_j, \quad [y_j, z] = \sum_i b_{i,j} y_i.$$

Now compute

$$[z, \sum_i x_i y_i] = \sum_i [z, x_i] y_i + \sum_j x_j [z, y_j] = \sum_{i,j} a_{i,j} x_j y_i - \sum_{i,j} b_{i,j} x_j y_i.$$

Therefore we need to prove that  $a_{i,j} = b_{i,j}$  and we're done. But we have that

$$a_{i,j} = (y_j, [z, x_i]) = ([y_j, z], x_i) = b_{i,j}.$$

QED.

This would work just great in the Dynkin case. We have the basis  $e_\alpha$  for  $\alpha \in \Delta_+$  for  $\mathfrak{n}_+$ . Pick a dual basis  $f_\alpha$  for  $\mathfrak{n}_-$ . Also let  $u_i$  be a basis for  $\mathfrak{h}$  and let  $v_i$  be the dual basis. Then the Casimir operator is

$$\sum u_i v_i + \sum_{\alpha \in \Delta_+} f_\alpha e_\alpha + \sum_{\alpha \in \Delta_+} e_\alpha f_\alpha.$$

This is central. Now it acts as a scalar on each  $L(\lambda)$  by Schur's lemma. To find what scalar, we need to act on a h/w vector. For that it is more convenient to rewrite:

$$\sum u_i v_i + 2 \sum_{\alpha \in \Delta_+} f_\alpha e_\alpha + \sum_{\alpha \in \Delta_+} m(\alpha) \alpha$$

(remembering that  $[e_\alpha, f_\alpha] = (e_\alpha, f_\alpha) \alpha$  by the properties of the invariant bilinear form). So we can work out the scalar by which it acts on  $L(\lambda)$  since the  $f_\alpha e_\alpha$  part acts as zero... The final term here contributes

$$\left( \sum \alpha, \lambda \right) = 2(\rho, \lambda)$$

to this scalar in view of the first lemma in the section. The first term contributes  $(\lambda, \lambda)$  (we'll prove this shortly). So overall we have  $(\lambda, \lambda) + 2(\rho, \lambda) = (\lambda + \rho, \lambda + \rho) - (\lambda, \lambda)$ , and we are done – in the Dynkin case.

We want to try to do the same in our infinite dimensional case. We do still have the invariant bilinear form! So, let  $u_1, u_2, \dots$  and  $v_1, v_2, \dots$  be dual bases of  $\mathfrak{h}$  with respect to this form. Let  $\{e_\alpha^{(i)}\}$  be a basis of the space  $\mathfrak{g}_\alpha$  for each  $\alpha \in \Delta_+$ , let  $\{f_\alpha^{(i)}\}$  be the dual basis for the space  $\mathfrak{g}_{-\alpha}$ . We may choose this basis so that  $e_{\epsilon_i} = e_i$  and  $f_{\epsilon_i} = f_i$ . Now the classical Casimir operator would be the expression

$$\sum_i u_i v_i + \sum_{\alpha \in \Delta_+} \sum_i f_\alpha^{(i)} e_\alpha^{(i)} + \sum_{\alpha \in \Delta_+} \sum_i e_\alpha^{(i)} f_\alpha^{(i)}.$$

Ignoring for the moment that this doesn't make sense as an element of  $U(\mathfrak{g})$  (its an infinite sum) we would then hopefully have a central element. To compute its action on a h/w vector we can rewrite it as

$$\sum_i u_i v_i + 2 \sum_{\alpha \in \Delta_+} \sum_i f_\alpha^{(i)} e_\alpha^{(i)} + \sum_{\alpha \in \Delta_+} \alpha.$$

Still that is not going to make sense because of the last sum. But the last sum acts on a h/w vector as something that does make sense, so we replace it by  $2\rho$  instead to get something finite and sensible. Moreover this expression – though still an infinite sum – does still make sense now as an operator on any module in category  $\mathcal{O}$  because only finitely many of the  $e$ 's act as non-zero on any given vector in such a module. So this is where we should begin!!!! So, with the motivation out of the way, let

$$\Omega = 2\rho + \sum_i u_i v_i + 2 \sum_{\alpha \in \Delta_+} \sum_i f_\alpha^{(i)} e_\alpha^{(i)}.$$

If  $\Delta_+$  is finite this makes sense as an element of  $U(\mathfrak{g})$ , hence as an operator on *any* representation. BUT usually  $\Delta_+$  is infinite, so this is meaningless as yet. We need to fix now a module  $M \in \mathcal{O}$ . On any particular vector  $m \in M$ , all but finitely many of the  $e_\alpha^{(i)}$  act as zero. So  $\Omega$  does make sense as an element of  $\text{End}(M)$ . Now we finish the section with the brutal calculation (but amazingly the only brutal calculation!!!)

**Lemma 3.37.**  *$\Omega$  commutes with the action of  $\mathfrak{g}$  on  $M$  Moreover, if  $v \in M$  is a highest weight vector of highest weight  $\lambda$ , then  $\Omega v = ((\lambda + \rho, \lambda + \rho) - (\rho, \rho))v$ .*

*Proof.* It is convenient to split  $\Omega$  into two pieces

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

and

$$\Omega_1 = 2\rho + \sum_i u_i v_i.$$

Let us first analyse the easy bit,  $\Omega_1$ : for  $\lambda, \mu \in \mathfrak{h}$ ,

$$\lambda = \sum_i (\lambda, u_i) v_i, \quad \mu = \sum_i (\mu, v_i) u_i$$

hence

$$(\lambda, \mu) = \sum_i (\lambda, u_i) (\mu, v_i).$$

If  $v$  is a h/w vector of h/w  $\lambda$ , then  $\Omega_0$  acts on  $v$  as zero and

$$\begin{aligned} \Omega_1 v &= ((2\rho, \lambda) + \sum_i (u_i, \lambda) (v_i, \lambda)) v \\ &= (2(\lambda, \rho) + (\lambda, \lambda)) v = ((\lambda + \rho, \lambda + \rho) - (\rho, \rho)) v \end{aligned}$$

which proves the moreover. Now take  $x \in \mathfrak{g}_\alpha$ . Then,

$$\begin{aligned} [\sum_i u_i v_i, x] &= \sum_i (\alpha, u_i) x v_i + \sum_i u_i (\alpha, v_i) x \\ &= \sum_i (\alpha, u_i) (\alpha, v_i) x + x (\sum_i u_i (\alpha, v_i) + v_i (\alpha, u_i)) \\ &= x((\alpha, \alpha) + 2\alpha). \end{aligned}$$

Hence,

$$[\Omega_1, x] = x(2(\rho, \alpha) + (\alpha, \alpha) + 2\alpha)$$

for  $x \in \mathfrak{g}_\alpha$ . To complete the proof of the lemma, it suffices to show that

$$[\Omega_0, x] = -x(2(\rho, \alpha) + (\alpha, \alpha) + 2\alpha)$$

as operators on  $M$  for any  $x \in \mathfrak{g}_\alpha$ .

So now we've got to analyse  $\Omega_0$ . We'll prove slightly more than we need here. Let  $U = U(\mathfrak{g})$  and for  $\alpha \in R$ , let

$$U_\alpha = \{u \in U \mid [h, u] = (\alpha, h)u \text{ for all } h \in \mathfrak{h}\}.$$

We'll show that

$$[\Omega_0, u] = -u(2(\rho, \alpha) + (\alpha, \alpha) + 2\alpha)$$

for each  $u \in U_\alpha$  (as operators on  $M$  of course). This is obvious for  $u \in \mathfrak{h}$ . Observe first that if this is true for  $u \in U_\alpha$  and for  $u' \in U_\beta$  then it is true for  $uu' \in U_{\alpha+\beta}$ :

$$\begin{aligned} [\Omega_0, uu'] &= [\Omega_0, u]u' + u[\Omega_0, u'] \\ &= -u(2(\rho, \alpha) + (\alpha, \alpha) + 2\alpha)u' - uu'(2(\rho, \beta) + (\beta, \beta) + 2\beta) \\ &= -uu'((2(\rho, \alpha) + (\alpha, \alpha) + 2\alpha) + 2(\alpha, \beta) + 2(\rho, \beta) + (\beta, \beta) + 2\beta) \\ &= -uu'(2(\rho, \alpha + \beta) + (\alpha + \beta, \alpha + \beta) + 2(\alpha + \beta)). \end{aligned}$$

Since  $U$  is generated by  $\mathfrak{h}$  and by the  $e_i, f_i$ , it now suffices to check our statement for  $u = e_i, f_i$ .

Now let us just check:

$$[\Omega_0, e_i] = -4e_i - 2e_i \epsilon_i = -2\epsilon_i e_i$$

as operators on  $M$ . (You also need to prove the claim for  $f_i$  but that works out by a similar but easier calculation which I'm going to omit.) This is what we want because  $2(\rho, \epsilon_i) + (\epsilon_i, \epsilon_i) = 4$ . To prove this, we are going to need to use the *main identity*: for  $\alpha, \beta \in \Delta$  and  $z \in \mathfrak{g}_{\beta-\alpha}$ ,

$$\sum_i f_\alpha^{(i)}[e_\alpha^{(i)}, z] = - \sum_j [f_\beta^{(j)}, z] e_\beta^{(j)} \quad (1)$$

written in  $U(\mathfrak{g})$  hence in any representation. This follows from the lemma below on applying the map  $x \otimes y \mapsto xy$ . Given the main identity, we have that

$$\begin{aligned} [\Omega_0, e_i] &= 2 \sum_{\alpha \in \Delta_+} [f_\alpha^{(i)} e_\alpha^{(i)}, e_i] \\ &= 2 \sum_{\alpha \in \Delta_+} ([f_\alpha^{(i)}, e_i] e_\alpha^{(i)} + f_\alpha^{(i)} [e_\alpha^{(i)}, e_i]) \\ &= 2[f_i, e_i] e_i + 2 \sum_{\alpha \in \Delta_+ - \{\epsilon_i\}} ([f_\alpha^{(i)}, e_i] e_\alpha^{(i)} + f_\alpha^{(i)} [e_\alpha^{(i)}, e_i]) \\ &= -2\epsilon_i e_i + 2 \sum_{\alpha \in \Delta_+ - \{\epsilon_i\}} \left( \sum_i [f_\alpha^{(i)}, e_i] e_\alpha^{(i)} + \sum_j f_{\alpha + \epsilon_i}^{(j)} [e_{\alpha - \epsilon_i}^{(j)}, e_i] \right) \\ &= -2\epsilon_i e_i \end{aligned}$$

(For the middle step, set  $\beta = \alpha + \epsilon_i$  and note the sum is also over  $\beta \in \Delta_+ - \{\epsilon_i\}$ , then replace  $\beta$  back with  $\alpha$ ).  $\square$

**Lemma 3.38.** *If  $\alpha, \beta \in \Delta$  and  $z \in \mathfrak{g}_{\beta-\alpha}$ , then we have that*

$$\sum_i f_\alpha^{(i)} \otimes [z, e_\alpha^{(i)}] = \sum_j [f_\beta^{(j)}, z] \otimes e_\beta^{(j)}$$

in  $\mathfrak{g} \otimes \mathfrak{g}$ .

*Proof.* Extend the form  $(\cdot, \cdot)$  on  $\mathfrak{g}$  to  $\mathfrak{g} \otimes \mathfrak{g}$  in the usual way. Pick  $e \in \mathfrak{g}_\alpha$  and  $f \in \mathfrak{g}_{-\beta}$ . It suffices by non-degeneracy of the form to check that pairing both sides with  $e \otimes f$  gives the same result. LHS:

$$\sum_i (f_\alpha^{(i)}, e) ([z, e_\alpha^{(i)}], f) = \sum_i (f_\alpha^{(i)}, e) (e_\alpha^{(i)}, [f, z]) = ([f, z], e) = (z, [e, f])$$

since they are dual bases! RHS:

$$\sum_j ([f_\beta^{(j)}, z], e) (e_\beta^{(j)}, f) = \sum_j (f_\beta^{(j)}, [z, e]) (e_\beta^{(j)}, f) = ([z, e], f) = (z, [e, f])$$

We're done.  $\square$

**Exercise 7.** Suppose that  $\mathfrak{g}$  is a finite dimensional Lie algebra with a non-degenerate invariant symmetric bilinear form  $(\cdot, \cdot)$ . Let  $x_i$  and  $y_i$  be a pair

of dual basis. Let

$$\Omega_2 = \sum_{i=1}^n x_i \otimes y_i.$$

Show that for any  $\mathfrak{g}$ -module  $V$ , the action of the operator  $\Omega_2$  on  $V \otimes V$  commutes with the usual tensor product action of  $\mathfrak{g}$ .

**3.12. The category  $\mathcal{O}_{int}$ .** Let  $\mathcal{O}_{int}$  denote the category of all integrable modules in category  $\mathcal{O}$ . In the case  $\Gamma$  is Dynkin, this is exactly the category of all finite dimensional  $\mathfrak{g}$ -modules (though I am not going to explain why) – in general you should think of  $\mathcal{O}_{int}$  as being the right analog of finite dimensional for general  $\Gamma$ .

Our first question: which  $L(\lambda)$ 's belong to  $\mathcal{O}_{int}$ ? i.e. what are the irreducibles in this category. Let

$$P = \{\lambda \in \mathfrak{h} \mid (\lambda, \epsilon_i) \in \mathbb{Z} \text{ for all } i = 1, \dots, n\},$$

$$P^+ = \{\lambda \in P \mid (\lambda, \epsilon_i) \geq 0 \text{ for all } i = 1, \dots, n\}.$$

We call elements of  $P$  *integral weights* and elements of  $P^+$  *dominant integral weights*. Note  $P^+$  is a sublattice of the fundamental chamber  $C$ . For example, in the Dynkin case,  $P$  is the free  $\mathbb{Z}$ -module on basis  $\Lambda_1, \dots, \Lambda_n$  where  $(\Lambda_i, \epsilon_j) = \delta_{i,j}$ , and then  $P^+$  are the non-negative linear combinations of these *fundamental dominant weights*.

**Lemma 3.39.** *Let  $\lambda \in \mathfrak{h}$ . Then,  $L(\lambda)$  is integrable if and only if  $\lambda \in P^+$ .*

*Proof.* By the representation theory of  $\mathfrak{sl}_2$ , we must have that  $(\lambda, \epsilon_i)$  is a non-negative integer for each  $i = 1, \dots, n$ . Conversely, suppose each  $(\lambda, \epsilon_i)$  is a non-negative integer. Then,

$$e_i f_i^{(\lambda, \epsilon_i)+1} v_\lambda = 0$$

for each  $i = 1, \dots, n$ . Suppose that  $f_i^{(\lambda, \epsilon_i)+1} v_\lambda$  is non-zero for some  $i$ . Then it is a highest weight vector of weight  $< \lambda$ , which contradicts the irreducibility of  $L(\lambda)$ . Hence each  $f_i^{(\lambda, \epsilon_i)+1} v_\lambda = 0$ . Hence the  $e_i$  and  $f_i$  act locally finitely on the vector  $v_\lambda$ , and since  $v_\lambda$  generates all of  $L(\lambda)$  it follows that they act locally finitely on the whole thing.  $\square$

It follows in particular that for  $\lambda \in P^+$ ,

$$\dim L(\lambda)_\mu = \dim L(\lambda)_{w\mu}$$

for each  $w \in W$ .

It would be nice if we could make the Weyl group  $W$  act on the character ring  $\mathcal{E}$ , then we could rephrase the above statement as saying that characters of  $L(\lambda)$ 's in  $\mathcal{O}_{int}$  are  $W$ -invariant. But that doesn't quite make sense, since  $s_i$  sends  $\epsilon_i$  to  $-\epsilon_i$  so it gives things that fail the finitely many roots condition. So we need the bigger *vector space*  $\tilde{\mathcal{E}}$  consisting of *all* expressions of the form

$$\sum_{\lambda \in \mathfrak{h}} c_\lambda e^\lambda$$

for  $c_\lambda \in \mathbb{Z}$  (no finiteness assumptions at all). We let  $w \in W$  act on this by

$$w \sum_{\lambda \in \mathfrak{h}} c_\lambda e^\lambda = \sum_{\lambda \in \mathfrak{h}} c_\lambda e^{w\lambda}.$$

Note products do not often make sense in  $\tilde{\mathcal{E}}$ , but when they do it will certainly be the case that

$$w(P_1 P_2) = (wP_1)(wP_2).$$

Now we can restate the observation in the previous paragraph as saying that

$$w \operatorname{ch} L(\lambda) = \operatorname{ch} L(\lambda)$$

for all  $w \in W$  and  $\lambda \in P^+$ .

Recall now the character of the Verma module

$$\operatorname{ch} M(\lambda) = \frac{e^\lambda}{R}$$

where

$$R = \prod_{\alpha \in \Delta_+} (1 - e^{-\alpha})^{m(\alpha)}.$$

Also let  $\rho \in \mathfrak{h}$  be the element chosen before with  $(\rho, \epsilon_i) = 1$  for each  $i = 1, \dots, n$ .

**Lemma 3.40.** *For  $w \in W$ ,  $w(e^\rho R) = (-1)^{\ell(w)} e^\rho R$ .*

*Proof.* Since  $(-1)^{\ell(w)} = \det_E w$ , the map  $w \mapsto (-1)^{\ell(w)}$  is a group homomorphism. So it suffices to check the lemma for  $w = s_i$ . Recall that  $\Delta_+ - \{\epsilon_i\}$  is  $s_i$ -invariant, and  $m(\alpha) = m(s_i(\alpha))$ . Hence, since  $s_i(\rho) = \rho - \epsilon_i$ ,

$$\begin{aligned} s_i(e^\rho R) &= e^{\rho - \epsilon_i} s_i(1 - e^{-\epsilon_i}) s_i \prod_{\alpha \in \Delta_+ - \{\epsilon_i\}} (1 - e^{-\alpha})^{m(\alpha)} \\ &= e^\rho e^{-\epsilon_i} (1 - e^{\epsilon_i}) \prod_{\alpha \in \Delta_+ - \{\epsilon_i\}} (1 - e^{-\alpha})^{m(\alpha)} \\ &= -e^\rho (1 - e^{-\epsilon_i}) \prod_{\alpha \in \Delta_+ - \{\epsilon_i\}} (1 - e^{-\alpha})^{m(\alpha)} \\ &= -e^\rho \prod_{\alpha \in \Delta_+} (1 - e^{-\alpha})^{m(\alpha)} \end{aligned}$$

As required! □

In the next lemma I'll use two more things. First for  $\lambda \in \mathfrak{h}' = \mathbb{C}\epsilon_1 \oplus \dots \oplus \mathbb{C}\epsilon_n$ , let  $\operatorname{supp}(\lambda)$  denote the set of  $i$  such that the  $\epsilon_i$ -coefficient of  $\lambda$  is non-zero. Second, note that if  $\lambda - \rho \in P^+$ , this means that

$$(\lambda - \rho, \epsilon_i) \in \mathbb{Z}_{\geq 0},$$

hence

$$(\lambda, \epsilon_i) \in \mathbb{Z}_{> 0},$$

i.e.  $\lambda$  is “very” dominant.

**Lemma 3.41.** *Let  $\mu, \lambda \in P$  be such that  $\mu \leq \lambda$  and  $\mu + \lambda \in P^+$ . Let  $\text{supp}(\mu - \lambda)$  be the set of all  $i$  such that the  $\epsilon_i$ -coefficient of  $\mu - \lambda$  is non-zero. Then either  $(\mu + \lambda, \epsilon_i) = 0$  for  $i \in \text{supp}(\mu - \lambda)$  or  $(\lambda, \lambda) > (\mu, \mu)$ .*

*In particular, if  $\lambda - \rho \in P^+$ ,  $\mu \in P^+$  and  $\mu < \lambda$  then  $(\lambda, \lambda) > (\mu, \mu)$ .*

*Proof.* Note the second part follows from the first, for under that extra circumstance  $(\mu + \lambda, \epsilon_i)$  is never zero.

For the first, let  $\mu = \lambda - \beta$  with  $\beta = \sum k_i \epsilon_i$ ,  $k_i \geq 0$ . Note the  $i$  with  $k_i \neq 0$  is  $\text{supp}(\mu - \lambda)$ . Then,

$$(\lambda, \lambda) - (\mu, \mu) = 2(\lambda, \beta) - (\beta, \beta) = (\lambda + \mu, \beta) = \sum_i k_i (\lambda + \mu, \epsilon_i).$$

Since  $\lambda + \mu \in P^+$  this is  $\geq 0$ . If it is  $> 0$  then  $(\lambda, \lambda) > (\mu, \mu)$  and we're done. If it is  $= 0$  then  $(\lambda + \mu, \epsilon_i) = 0$  whenever  $k_i \neq 0$ .  $\square$

Now we can prove the Weyl-Kac character formula.

**Theorem 3.42.** *Let  $\lambda \in P^+$ . Then,*

$$\text{ch } L(\lambda) = \frac{\sum_{w \in W} (-1)^{\ell(w)} e^{w(\lambda + \rho) - \rho}}{\prod_{\alpha \in \Delta_+} (1 - e^{-\alpha})^{m(\alpha)}}.$$

*(RHS computed in  $\tilde{\mathcal{E}}$ ).*

*Proof.* We know that

$$\text{ch } L(\lambda) = \sum_{\mu} c_{\mu} \text{ch } M(\mu)$$

for coefficients  $c_{\mu} \in \mathbb{Z}$  with  $c_{\lambda} = 1$ , where the sum is over  $\mu \leq \lambda$  with  $(\lambda + \rho, \lambda + \rho) = (\mu + \rho, \mu + \rho)$ . Multiplying both sides by  $e^{\rho} R$  and expanding the formula for  $\text{ch } M(\mu)$  gives

$$e^{\rho} R \text{ch } L(\lambda) = \sum_{\mu} c_{\mu} e^{\mu + \rho}$$

summing over the same  $\mu$ . The left hand side is  $W$ -skew-invariant, i.e. multiplying by  $w$  scales it by  $(-1)^{\ell(w)}$ . Hence the RHS is too:

$$c_{\mu} = (-1)^{\ell(w)} c_{\nu}$$

if  $\nu = w(\mu + \rho) - \rho$  for some  $w \in W$ .

Now fix  $\mu \leq \lambda$  with  $c_{\mu} \neq 0$ , so certainly  $\mu \leq \lambda$  and  $(\lambda + \rho, \lambda + \rho) = (\mu + \rho, \mu + \rho)$ . Choose  $\nu \in \{w(\mu + \rho) - \rho \mid w \in W\}$  so that  $ht(\lambda - \nu)$  is minimal. By the previous paragraph, we have that  $c_{\nu} \neq 0$ , and clearly  $\nu + \rho \in P^+$  (else we could apply some  $s_i$  to get something with smaller height). Since  $c_{\nu} \neq 0$ ,  $\nu \leq \lambda$  and  $(\nu + \rho, \nu + \rho) = (\lambda + \rho, \lambda + \rho)$ . Applying the preceding lemma to  $\lambda + \rho$  and  $\nu + \rho$  we deduce that  $\nu = \lambda$ .

Thus  $c_{\mu} \neq 0$  implies that  $\mu = w(\lambda + \rho) - \rho$  for some  $w \in W$  and hence that  $c_{\mu} = (-1)^{\ell(w)}$ . Finally note that the stabilizer in  $W$  of  $\lambda + \rho$  is (since it lies in the fundamental chamber) generated by the  $s_i$  that fix it, which

is none of them since it is very dominant. We've just calculated the RHS! Hence,

$$e^\rho \text{ch } L(\lambda) = \sum_{w \in W} (-1)^{\ell(w)} e^{w(\lambda+\rho)}.$$

The theorem follows.  $\square$

**Remarks 3.43.** (0) Let's practise in type  $A_1$ ! Here,  $P^+$  is the set of all weights  $n\epsilon_1/2$  for  $n \in \mathbb{N}$  – because those guys have inner product with  $\epsilon_1$  equal to  $n \in \mathbb{N}$ ! Let  $L(n)$  be the irreducible representation of  $\mathfrak{sl}_2$  of highest weight  $n\epsilon_1/2$  – i.e. the ones generated by a vector  $v$  with  $ev = 0, hv = nv$ . Of course we already know this is the  $(n+1)$ -dimensional rep whose character is

$$e^n + e^{n-2} + \dots + e^{-n}$$

if  $n \in \mathbb{N}$ , writing  $e^n$  not  $e^{n\epsilon_1/2}$ . What does WCF say? Here,  $\rho = \epsilon_1/2$ . So it is

$$\frac{\sum_{w \in S_2} (-1)^{\ell(w)} e^{w(n+1)-1}}{1 - e^{-2}}$$

which is

$$(e^n - e^{-n-2})(1 + e^{-2} + e^{-4} + \dots)$$

which is

$$e^n + e^{n-2} + e^{n-4} + \dots + e^{-n} + e^{-n-2} + e^{-n-4} + \dots - e^{-n-2} - e^{-n-4} - \dots$$

which is correct!

(1) Take  $\lambda = 0$  in the Weyl-Kac character formula. Obviously  $L(0)$  is the trivial representation so  $\text{ch } L(0) = 1$ . Hence we deduce that

$$\sum_{w \in W} (-1)^{\ell(w)} e^{w(\lambda+\rho)-\rho} = \prod_{\alpha \in \Delta_+} (1 - e^{-\alpha})^{m(\alpha)}.$$

This is the very non-trivial *denominator identity*. Of course its pretty stupid for  $\mathfrak{sl}_2$ : it says that  $(1 - e^{-2})^{-1} = 1 + e^{-2} + e^{-4} + \dots$ . But already for  $\mathfrak{sl}_3$  it is not quite so trivial. In the case of  $\widehat{\mathfrak{sl}}_2$  it has been used by Kac, Lepowsky and Wilson prove an impressive combinatorial identity: the Rogers-Ramanujan identity.

(2) In the case  $\Gamma$  is a Dynkin diagram, the theorem is the classical *Weyl character formula*. It was generalized to arbitrary symmetrizable Kac-Moody Lie algebras in 1974 (we have explained it here only in the symmetric case). In 1987 Kumar and Mathieu extended it to *arbitrary* (not even symmetrizable) Kac-Moody Lie algebras.

(3) The proof of the Weyl-Kac character formula never used the fact that  $L(\lambda)$  was irreducible: just that it was an integrable highest weight module. You could start instead with any integrable h/w module  $L'(\lambda)$  (where  $\lambda \in P^+$  necessarily by the proof of Lemma 3.39). The proof goes through to show that  $\text{ch } L'(\lambda) = \text{ch } L(\lambda)$ . Hence, *any integrable h/w module of h/w  $\lambda$  is isomorphic to  $L(\lambda)$ .*

(4) Take in particular  $L'(\lambda)$  to be the quotient of  $M(\lambda)$  by the submodule generated by the vectors

$$f_i^{(\lambda, \epsilon_i)+1} v_\lambda$$

for each  $i = 1, \dots, n$  (for  $\lambda \in P^+$  still). The proof of Lemma 3.39 shows that these vectors are h/w vectors in  $M(\lambda)$  so they belong to its unique maximal submodule, hence  $L'(\lambda)$  is non-zero. Moreover, the proof of Lemma 3.39 shows already that  $L'(\lambda)$  is integrable. Hence  $L'(\lambda) = L(\lambda)$ . In other words, we can define  $L(\lambda)$  for  $\lambda \in P^+$  in a completely different way as the quotient of  $U(\mathfrak{g})$  by the left ideal generated by the elements  $e_i, f_i^{(\lambda, \epsilon_i)+1}$  and  $\epsilon_i - (\lambda, \epsilon_i)$ .

Let me mention one other consequence of the Weyl-Kac character formula.

**Lemma 3.44.** *For  $\lambda$  with  $\lambda - \rho \in P^+$ ,*

$$\sum_{w \in W} (-1)^{\ell(w)} q^{(\lambda, \rho - w\rho)} = \prod_{\alpha \in \Delta_+} (1 - q^{(\lambda, \alpha)})^{m(\alpha)}.$$

*Proof.* Let  $s_i = (\lambda, \epsilon_i)$  for each  $i = 1, \dots, n$ . So  $s_i > 0$  for each  $i$ . Define a homomorphism  $\mathbb{C}[[e^{-\epsilon_1}, \dots, e^{-\epsilon_n}]] \rightarrow \mathbb{C}[[q]]$  by

$$e^{-\epsilon_i} \mapsto q^{s_i}.$$

Say  $\alpha = \sum k_i \epsilon_i$ . Then  $e^{-\alpha}$  maps to  $q^{\sum k_i s_i}$ . But  $\sum k_i s_i = (\lambda, \alpha)$ . So our map sends  $e^{-\alpha}$  to  $q^{(\lambda, \alpha)}$ . In particular, it sends

$$e^{w\rho - \rho} \mapsto q^{(\lambda, \rho - w\rho)}$$

and

$$1 - e^{-\alpha} \mapsto 1 - q^{(\lambda, \alpha)}.$$

Now take the identity

$$\sum_{w \in W} (-1)^{\ell(w)} e^{w\rho - \rho} = \prod_{\alpha \in \Delta_+} (1 - e^{-\alpha})^{m(\alpha)}$$

from Remark (1) and apply the homomorphism to both sides.  $\square$

Now fix  $\lambda \in P^+$ . For  $d \geq 0$ , let  $L(\lambda)_d$  be the sum of all  $L(\lambda)_\mu$  with  $ht(\lambda - \mu) = d$ , i.e. the sum of all the weight spaces of  $L(\lambda)$  that are  $d$  simple roots down from the highest weight. Let

$$\dim_q L(\lambda) = \sum_{d \geq 0} (\dim L(\lambda)_d) q^d \in \mathbb{N}[[q]].$$

This is the  $q$ -dimension of  $L(\lambda)$ .

**Theorem 3.45.** *For  $\lambda \in P^+$ ,*

$$\dim_q L(\lambda) = \prod_{\alpha \in \Delta_+} \left( \frac{1 - q^{(\lambda + \rho, \alpha)}}{1 - q^{(\rho, \alpha)}} \right)^{m(\alpha)}.$$

*Proof.* Note that  $ht(\lambda - \mu) = (\lambda - \mu, \rho)$  and that

$$(\lambda - (w(\lambda + \rho) - \rho), \rho) = (\lambda + \rho, \rho - w\rho).$$

So by WCF and Remark (1) above, we need to compute

$$\frac{\sum_{w \in W} (-1)^{\ell(w)} q^{(\lambda + \rho, \rho - w\rho)}}{\sum_{w \in W} (-1)^{\ell(w)} q^{(\rho, \rho - w\rho)}}.$$

By the preceding lemma, that is what we want!  $\square$

As a corollary we get the famous Weyl dimension formula.

**Corollary 3.46.** *Suppose  $\Gamma$  is Dynkin and  $\lambda \in P^+$ . Then,*

$$\dim L(\lambda) = \prod_{\alpha \in \Delta^+} \frac{(\lambda + \rho, \alpha)}{(\rho, \alpha)}.$$

*In particular,  $L(\lambda)$  is finite dimensional!*

*Proof.* Let  $q \rightarrow 1$  and apply l'Hopital.  $\square$

**Exercise 8.** Suppose  $\mathfrak{g} = \mathfrak{sl}_3(\mathbb{C})$ . Let  $\lambda \in P^+$  be a weight with  $(\lambda, \epsilon_1) = a$ ,  $(\lambda, \epsilon_2) = b$ . Prove that

$$\dim L(\lambda) = \frac{1}{2}(a+1)(b+1)(a+b+2).$$

What well known representation is the one with  $a = b = 1$ ?

**3.13. Complete reducibility.** The goal in this section is to prove:

**Theorem 3.47.** *Every  $M \in \mathcal{O}_{int}$  is isomorphic to a direct sum of  $L(\lambda)$ 's for  $\lambda \in P^+$ .*

**Remark 3.48.** In case  $\Gamma$  is Dynkin, we have already observed that any finite dimensional  $\mathfrak{g}$ -module belongs to  $\mathcal{O}_{int}$ . So the theorem implies *Weyl's theorem*: any finite dimensional  $\mathfrak{g}$ -module is completely reducible. Moreover, we have proved that each  $L(\lambda)$  belonging to  $\mathcal{O}_{int}$  is finite dimensional. So any object  $M$  in  $\mathcal{O}_{int}$  in the Dynkin case is a direct sum of various  $L(\lambda)$ 's and – since there are only finitely many dominant weights that are  $\leq$  any given  $\lambda$  – you can see moreover from the finitely many roots assumption in the definition of  $\mathcal{O}$  that there must be just finitely many summands, i.e.  $M$  is finite dimensional. Hence,  $\mathcal{O}_{int}$  is *precisely* the category of finite dimensional  $\mathfrak{g}$ -modules in the Dynkin case.

To prove the complete reducibility theorem, we need a couple of preparatory lemmas. Recall that we called a vector  $m \in M$  primitive of weight  $\lambda$  if there was a submodule  $m \notin U \subset M$  such that  $m + U$  was a h/w vector of h/w  $\lambda$  in  $M/U$ . We say  $\lambda$  is a primitive weight of  $M$  if there exists a primitive vector of weight  $\lambda$ .

**Lemma 3.49.** *Let  $M \in \mathcal{O}$ . If for any two primitive weights  $\lambda, \mu$  of  $M$  the inequality  $\lambda \geq \mu$  implies  $\lambda = \mu$ . Then the module  $M$  is completely reducible, i.e. a direct sum of  $L(\lambda)$ 's.*

*Proof.* Let  $M^0 = \{v \in M \mid \mathfrak{n}_+ v = 0\}$ . This is  $\mathfrak{h}$ -invariant, hence it is equal to the sum of its weight spaces. Let  $0 \neq v \in M_\lambda^0$  for some  $\lambda \in \mathfrak{h}$ , so  $v$  is a h/w vector of weight  $\lambda$ . Then the  $\mathfrak{g}$ -module  $U(\mathfrak{g})v$  is irreducible,  $\cong L(\lambda)$ . Indeed, otherwise it would have a proper submodule which necessarily contains a h/w vector of some weight  $\mu < \lambda$  contrary to the assumptions. Therefore, the  $\mathfrak{g}$ -submodule  $M'$  of  $M$  generated by  $M^0$  is completely reducible.

It remains to show that  $M' = M$ . Suppose not. Considering a h/w vector  $M/M'$ , there exists a weight vector  $v \in M$  of weight  $\mu$  such that  $v \notin M'$  but  $e_i v \in M'$  for each  $i$ . Since  $v \notin M'$  we must have that  $e_i v \neq 0$  for some  $i$ . But then since  $M$  is in  $\mathcal{O}$  and  $e_i v$  is of weight  $\mu + \epsilon_i$ , there must be some maximal weight  $\lambda$  of  $M$  with  $\lambda \geq \mu + \epsilon_i > \mu$ . This gives primitive weights  $\lambda, \mu$  of  $M$  contradicting the assumptions of the lemma.  $\square$

**Lemma 3.50.** *Let  $M \in \mathcal{O}$  such that for any two primitive weights  $\lambda, \mu$  of  $M$  with  $\lambda > \mu$  one has  $(\lambda + \rho, \lambda + \rho) \neq (\mu + \rho, \mu + \rho)$ , then  $M$  is completely reducible.*

*Proof.* Let  $\Omega$  be the Casimir operator on  $M$

$$\Omega = \sum_i h_i k_i + 2 \sum_{\alpha \in \Delta_+} \sum_j f_\alpha^{(j)} e_\alpha^{(j)} + 2\rho.$$

For  $a \in \mathbb{C}$ , let

$$M_a = \bigcup_{n \geq 0} \ker(\Omega - a)^n.$$

Since  $\Omega$  commutes with the action of  $\mathfrak{g}$  on  $M$ , each  $M_a$  is  $\mathfrak{g}$ -invariant. Think of a weight space  $M_\lambda$ . Since  $\Omega$  leaves  $M_\lambda$  invariant and  $M_\lambda$  is finite dimensional,  $M_\lambda$  is a direct sum of Jordan blocks over  $\Omega$  – these are the  $M_\lambda \cap M_a$ 's. This shows that

$$M = \bigoplus M_a.$$

Note for each primitive weight  $\lambda$  of  $M_a$ , we must have that  $(\lambda + \rho, \lambda + \rho) - (\rho, \rho) = a$  since that is the scalar that  $\Omega$  acts on a h/w vector of weight  $\lambda$ .

Now we show that  $M_a$  is completely reducible to prove the lemma. Take two primitive vectors  $\lambda \geq \mu$  of  $M_a$ . We must have that  $(\lambda + \rho, \lambda + \rho) = (\mu + \rho, \mu + \rho)$ , hence by the assumption  $\lambda = \mu$ . Hence by the preceding lemma,  $M_a$  is completely reducible.  $\square$

Now we can prove the theorem. Let  $M \in \mathcal{O}_{int}$ . By the preceding lemma, it suffices to show that if  $\lambda, \mu$  are primitive weights with  $\lambda > \mu$  then  $(\lambda + \rho, \lambda + \rho) \neq (\mu + \rho, \mu + \rho)$ . Since  $M$  is integrable,  $(\lambda, \epsilon_i), (\mu, \epsilon_i) \in \mathbb{N}$  for each  $i$ . Hence, letting  $\beta = \lambda - \mu > 0$ ,

$$\begin{aligned} (\lambda + \rho, \lambda + \rho) - (\mu + \rho, \mu + \rho) &= (\lambda + \rho, \lambda + \rho) - (\lambda - \beta + \rho, \lambda - \beta + \rho) \\ &= 2(\lambda + \rho, \beta) - (\beta, \beta) \\ &= (\lambda + \mu + 2\rho, \beta) > 0. \end{aligned}$$

We're done.

**3.14. Generators and relations.** The next thing to polish off: the Serre relations. Recall how we defined  $\mathfrak{g}$ . We started with  $\tilde{\mathfrak{g}}$  defined by generators  $\mathfrak{h}, e_i, f_i$  and the easy relations. We had

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

Then we factored out the unique maximal ideal  $\mathfrak{r} = \mathfrak{r} \cap \tilde{\mathfrak{n}}_- \oplus \mathfrak{r} \cap \tilde{\mathfrak{n}}_+$  (direct sum of ideals!) that intersects  $\mathfrak{h}$ -trivially.

We showed that the following elements lay in  $\mathfrak{r}$ :

$$(\text{ad } e_i)^{1+n_{i,j}} e_j, (\text{ad } f_i)^{1+n_{i,j}} f_j$$

for each  $i \neq j$ . These are the *Serre relations*. Our goal is to show that these relations are actually *sufficient* to generate all of  $\mathfrak{r}$ . This implies a simpler definition of  $\mathfrak{g}$  from the outset by generators and relations – one can simply impose the Serre relations in addition to the easy relations we imposed for  $\tilde{\mathfrak{g}}$  and that is the Kac-Moody Lie algebra directly.

**Lemma 3.51.** *Let  $\mathfrak{n}$  be a Lie algebra and  $\mathfrak{r}$  be a subalgebra. Then*

$$\mathfrak{r} \cap \mathfrak{r}U_0(\mathfrak{n}) \subseteq [\mathfrak{r}, \mathfrak{r}],$$

where  $U_0(\mathfrak{n})$  denotes the ideal in  $U(\mathfrak{n})$  generated by  $\mathfrak{n}$ .

*Proof.* Pick a basis  $x_i$  for  $\mathfrak{r}$  and extend to a basis  $x_i, y_j$  for  $\mathfrak{n}$ . Then  $\mathfrak{r}U_0(\mathfrak{n})$  is spanned by the ordered monomials that start from some  $x_i$  and involve at least one more basis element. Since the resulting linear combination has to lie back in  $\mathfrak{r}$ , one deduces from the PBW theorem that  $\mathfrak{r} \cap \mathfrak{r}U_0(\mathfrak{n}) = \mathfrak{r} \cap U_0(\mathfrak{r})^2$ . Now we show that

$$\mathfrak{r} \cap U_0(\mathfrak{r})^2 \subseteq [\mathfrak{r}, \mathfrak{r}].$$

Pasing to  $U(\mathfrak{r}/[\mathfrak{r}, \mathfrak{r}])$ , we may assume that  $\mathfrak{r}$  is commutative, and then need to see that

$$\mathfrak{r} \cap U_0(\mathfrak{r})^2 = 0$$

which is clear since  $U(\mathfrak{r})$  is a polynomial ring.  $\square$

The main step in the proof is the following:

**Lemma 3.52.** *The ideal  $\mathfrak{r}_- = \mathfrak{r} \cap \tilde{\mathfrak{n}}_-$  is generated as an ideal in  $\tilde{\mathfrak{n}}_-$  by those  $\mathfrak{r}_{-\alpha}$  for which  $\alpha \in R^+ - \{\epsilon_1, \dots, \epsilon_n\}$  with  $2(\rho, \alpha) = (\alpha, \alpha)$ . Similarly for  $\mathfrak{r}_+$ .*

*Proof.* For  $\lambda \in \mathfrak{h}$ , define the Verma module for  $\tilde{\mathfrak{g}}$  in exactly the same way as for  $\mathfrak{g}$ :

$$\tilde{M}(\lambda) = U(\tilde{\mathfrak{g}}) \otimes_{U(\mathfrak{h} \oplus \tilde{\mathfrak{n}}_+)} \mathbb{C}\lambda.$$

Again it has a unique maximal submodule  $\text{rad } \tilde{M}(\lambda)$ . In particular,

$$\tilde{M}(0)/\text{rad } \tilde{M}(0) \cong \mathbb{C}$$

so

$$\text{rad } \tilde{M}(0) \cong \bigoplus_{i=1}^n \tilde{M}(-\epsilon_i).$$

This is because  $U(\tilde{\mathfrak{n}}_-)$  is free on  $f_1, \dots, f_n$ . Hence,

$$U(\mathfrak{g}) \otimes_{U(\tilde{\mathfrak{g}})} \text{rad } \tilde{M}(0) \cong U(\mathfrak{g}) \otimes_{U(\tilde{\mathfrak{g}})} \bigoplus_{i=1}^n \tilde{M}(-\epsilon_i) \cong \bigoplus_{i=1}^n M(-\epsilon_i).$$

Let  $\pi : \tilde{\mathfrak{g}} \rightarrow \mathfrak{g}$  be the canonical map. Define a map

$$\psi : \mathfrak{r}_- \rightarrow U(\mathfrak{g}) \otimes_{U(\tilde{\mathfrak{g}})} \text{rad } \tilde{M}(0)$$

by  $\psi(a) = 1 \otimes a(\tilde{v}_+)$  where  $\tilde{v}_+$  is a h/w vector of  $\tilde{M}(0)$ . Viewing the ideal  $\mathfrak{r}_-$  as a  $\tilde{\mathfrak{g}}$ -module by the adjoint action (it IS an ideal!), this is a  $\tilde{\mathfrak{g}}$ -module homomorphism; indeed for  $x \in \tilde{\mathfrak{g}}$  and  $a \in \mathfrak{r}_-$  we have

$$\begin{aligned} \psi([x, a]) &= 1 \otimes (xa - ax)\tilde{v}_+ = \pi(x) \otimes a(\tilde{v}_+) - \pi(a) \otimes x(\tilde{v}_+) \\ &= \pi(x) \otimes a(\tilde{v}_+) = x(\psi(a)) \end{aligned}$$

since  $\pi(a) = 0$ . A similar calculation gives that  $\pi([\mathfrak{r}_-, \mathfrak{r}_-]) = 0$ . So  $\psi$  factors to a  $\tilde{\mathfrak{g}}$ -module homomorphism

$$\bar{\psi} : \mathfrak{r}_- / [\mathfrak{r}_-, \mathfrak{r}_-] \rightarrow \bigoplus_{i=1}^n M(-\epsilon_i).$$

Recalling that  $\mathfrak{r} = \mathfrak{r}_- \oplus \mathfrak{r}_+$  (direct sum of ideals),  $[\mathfrak{r}, \mathfrak{r}_-] = [\mathfrak{r}_-, \mathfrak{r}_-] + [\mathfrak{r}_+, \mathfrak{r}_-] = [\mathfrak{r}_-, \mathfrak{r}_-]$ . So  $\mathfrak{r}_- / [\mathfrak{r}_-, \mathfrak{r}_-]$  is actually a  $\mathfrak{g}$ -module and we've constructed a homomorphism  $\mathfrak{r}_- / [\mathfrak{r}_-, \mathfrak{r}_-] \hookrightarrow \bigoplus M(-\epsilon_i)$  of  $\mathfrak{g}$ -modules.

We claim that  $\bar{\psi}$  is injective. This implies that  $\mathfrak{r}_- / [\mathfrak{r}_-, \mathfrak{r}_-]$  belongs to the category  $\mathcal{O}$ , hence it is generated – even as an  $\mathfrak{n}_-$  – by its primitive vectors. Let  $-\alpha$  be a primitive weight, where clearly  $\alpha \in R^+ - \{\epsilon_1, \dots, \epsilon_n\}$  since none of the  $f_i$ 's are in  $\mathfrak{r}$ . To prove the lemma for  $\mathfrak{r}_-$ , we just need to show that  $2(\rho, \alpha) = (\alpha, \alpha)$  – then it follows for  $\mathfrak{r}_+$  by twisting by  $\tilde{\omega}$ . But  $-\alpha$  is then also a primitive weight in some  $M(-\epsilon_i)$ , hence Casimir must act in the same way:

$$(-\alpha + \rho, -\alpha + \rho) = (-\epsilon_i + \rho, -\epsilon_i + \rho)$$

for some  $i$ . That implies that  $(\alpha, \alpha) - 2(\rho, \alpha) = (\epsilon_i, \epsilon_i) - 2(\epsilon_i, \rho) = 0$ . We're done.

Still we have to show that  $\bar{\psi}$  is injective. It may be described explicitly as follows: write  $a \in \mathfrak{r}_-$  in the form  $a = \sum_i u_i f_i$  with  $u_i \in U_0(\tilde{\mathfrak{n}}_-)$  (we can do this since the  $f_i$  generate  $U(\tilde{\mathfrak{n}}_-)$ ). Let  $v_i$  be the h/w vector in  $M(-\epsilon_i)$ . Then,

$$\bar{\psi}(a + [\mathfrak{r}_-, \mathfrak{r}_-]) = \sum \pi(u_i) v_i.$$

Suppose this is zero. Then  $\pi(u_i) = 0$  for each  $i$ . Hence,  $a \in \mathfrak{r}_- U_0(\tilde{\mathfrak{n}}_-) \cap \mathfrak{r}_-$ . Now we're done since the right hand side is  $[\mathfrak{r}_-, \mathfrak{r}_-]$  by the opening lemma.  $\square$

**Theorem 3.53.** *The elements*

$$(\text{ad } e_i)^{1+n_{i,j}} e_j, (\text{ad } f_i)^{1+n_{i,j}} f_j$$

for  $i \neq j$  generate  $\mathfrak{r}_+$  resp.  $\mathfrak{r}_-$ .

*Proof.* Let  $\mathfrak{g}'$  be the quotient of  $\tilde{\mathfrak{g}}$  by these relations. Let  $\mathfrak{r}'_{\pm}$  be the image of  $\mathfrak{r}_{\pm}$ . Note  $\mathfrak{g}'$  is still a direct sum of root spaces  $\bigoplus_{\alpha \in R} \mathfrak{g}'_{\alpha}$ . We have already shown that the Serre relations hold in  $\mathfrak{g}$ . So it remains to show that  $\mathfrak{r}'_{\pm} = 0$ . Suppose for a contradiction that  $\mathfrak{r}'_{+} \neq 0$  (the case of  $\mathfrak{r}'_{-}$  being similar). Choose a root  $\alpha$  of minimal height among all  $\alpha \in R^{+} - \{0\}$  such that  $\mathfrak{r}'_{\alpha} \neq 0$ . By the preceding lemma,  $2(\rho, \alpha) = (\alpha, \alpha)$ .

The relations we have factored out are all we needed when we were proving that each  $e_i$  and  $f_i$  acted locally finitely, i.e. that  $\mathfrak{g}$  was an integrable representation. So we can do exactly the same things with  $\mathfrak{g}'$  as we did for  $\mathfrak{g}$  regarding the Weyl group. So we let  $s_i$  be the automorphism of  $\mathfrak{g}'$

$$s_i = \exp(\text{ad } f_i) \exp(-\text{ad } e_i) \exp(\text{ad } f_i).$$

This maps  $\mathfrak{g}'_{\alpha}$  to  $\mathfrak{g}'_{s_i(\alpha)}$ . Since  $\alpha \neq k\epsilon_i$ ,  $s_i(\alpha)$  must still be a positive root. Moreover, since  $\mathfrak{r}'$  is the unique maximal ideal of  $\mathfrak{g}'$  intersecting  $\mathfrak{h}$  trivially  $s_i$  leaves  $\mathfrak{r}'$  invariant. Hence,  $s_i$  maps  $\mathfrak{r}'_{\alpha}$  to  $\mathfrak{r}'_{s_i(\alpha)}$  and by the minimality of the choice of  $\alpha$  we must have  $ht(s_i(\alpha)) \geq ht(\alpha)$ , hence

$$(\alpha, \epsilon_i) \leq 0.$$

This is true for all  $i$ , hence  $(\alpha, \alpha) \leq 0$ . But  $2(\rho, \alpha) > 0$ . This is the contradiction.  $\square$

Let me stress this result by stating a couple of examples. First, we've derived the generators and relations for the finite dimensional simple Lie algebras discovered originally by Serre: generators  $e_1, \dots, e_n, f_1, \dots, f_n, \epsilon_1, \dots, \epsilon_n$ ; relations

$$\begin{aligned} [\epsilon_i, \epsilon_j] &= 0, \\ [e_i, f_j] &= \delta_{i,j} \epsilon_i, \\ [\epsilon_i, e_j] &= (\epsilon_i, \epsilon_j) e_j, \\ [\epsilon_i, f_j] &= -(\epsilon_i, \epsilon_j) f_j, \\ (\text{ad } e_i)^{1+n_{i,j}} e_j &= 0, \\ (\text{ad } f_i)^{1+n_{i,j}} f_j &= 0 \end{aligned}$$

where  $n_{i,j}$  is the number of edges from  $i$  to  $j$  and  $(\epsilon_i, \epsilon_j) = 2$  if  $i = j$ ,  $-n_{i,j}$  otherwise.

Second we've derived generators and relations for the affine Lie algebras: numbering the affine Dynkin diagram according to our usual convention, the

generators are  $e_0, e_1, \dots, e_n, f_0, f_1, \dots, f_n, \epsilon_0, \epsilon_1, \dots, \epsilon_n, d$ . The relations are

$$\begin{aligned} [\epsilon_i, \epsilon_j] &= 0, \\ [d, \epsilon_i] &= 0, \\ [d, e_i] &= \delta_{i,0} e_i, \\ [d, f_i] &= -\delta_{i,0} f_i, \\ [\epsilon_i, e_j] &= (\epsilon_i, \epsilon_j) e_j, \\ [\epsilon_i, f_j] &= -(\epsilon_i, \epsilon_j) f_j, \\ [e_i, f_j] &= \delta_{i,j} \epsilon_i, \\ (\text{ad } e_i)^{1+n_{i,j}} e_j &= 0, \\ (\text{ad } f_i)^{1+n_{i,j}} f_j &= 0. \end{aligned}$$

I find it amusing to write the generators for  $\widehat{\mathfrak{sl}}'_2$  out one more time: generators  $e_0, f_0, e_1, f_1, \epsilon_0, \epsilon_1$ , relations

$$\begin{aligned} [e_0, f_0] &= \epsilon_0, [e_1, f_1] = \epsilon_1, [e_0, f_1] = [e_1, f_0] = 0, \\ [\epsilon_0, e_0] &= 2e_0, [\epsilon_0, f_0] = -2f_0, \\ [\epsilon_1, e_1] &= 2e_1, [\epsilon_1, f_1] = -2f_1, \\ [\epsilon_0, e_1] &= -2e_1, [\epsilon_0, f_1] = 2f_1, \\ [\epsilon_1, e_0] &= -2e_0, [\epsilon_1, f_0] = 2f_0, \\ [e_0, [e_0, [e_0, e_1]]] &= [e_1, [e_1, [e_1, e_0]]] = 0, \\ [f_0, [f_0, [f_0, f_1]]] &= [f_1, [f_1, [f_1, f_0]]] = 0. \end{aligned}$$

Or something like that.

To end the section, I want to do one more baby example: the Heisenberg algebra  $\mathfrak{h}$ . This is the Lie algebra with generators  $p_i, q_i$  ( $i \geq 1$ ) and  $c$  with relations

$$[p_i, q_i] = ic$$

and all other brackets are zero. In the next section we're going to meet a slightly more complicated Heisenberg algebra, but let's just explain this one and the general one we'll meet later goes similarly.

Note  $\mathfrak{h}$  looks very like a Kac-Moody algebra: it has a triangular decomposition with the  $p_i$ 's generating  $\mathfrak{h}_+$ ,  $c$  generating the Cartan part and the  $q_i$ 's generating  $\mathfrak{h}_-$ . So you can talk about highest weight modules over  $\mathfrak{h}$ , meaning modules generated by a vector killed by the  $p$ 's on which  $c$  acts as a scalar  $\lambda$ . I want to convince you that  $\mathfrak{h}$  has a unique irreducible module generated by a vector killed by the  $p$ 's on which  $c$  acts as 1. This is called the *canonical representation of the Heisenberg Lie algebra*.

To construct it is easy: take

$$U(\mathfrak{h}) \otimes_{U(\mathfrak{h}_+ \oplus \mathbb{C}c)} \mathbb{C}$$

with  $c$  acting on  $\mathbb{C}$  as 1. This should be the "universal" highest weight module of weight 1. It has a basis by the PBW theorem given by all ordered

monomials in the  $q$ 's. Since the  $q$ 's commute with each other, we may as well simply identify this module with the polynomial algebra

$$B := \mathbb{C}[q_1, q_2, \dots] = S(\mathfrak{s}_-).$$

Then  $q_i$  is acting as the operator of multiplication by  $q_i$ , and  $c$  is acting everywhere as 1. How does  $p_i$  act? Well  $[p_i, q_j] = 0$  for  $j \neq i$  and  $[p_i, q_i] = ic$ . So whenever you hit  $q_i^r$  in  $B$  by  $p_i$  it comes past like a derivation and you find that in  $B$ ,

$$p_i q_i^r = ir q_i^{r-1}.$$

Hence  $p_i$  acts like the operator  $i \frac{d}{dq_i}$ . Now it is easy to prove that  $B$  is actually irreducible: take any element of the symmetric algebra and differentiate it to get back to 1.

**3.15. Basic representations of affine Lie algebras.** . Finally in the chapter I want to show you a beautiful thing involving the representation theory of the affine Lie algebras.

So assume  $\mathfrak{g} = \mathfrak{g}(\Gamma)$  for a Euclidean diagram  $\Gamma$ . We'll now switch back to affine Lie algebra mode, labelling vertices  $0, 1, \dots, n$ . Recall also how we've realized the form on  $\mathfrak{h}$ : it is spanned by  $\epsilon_0, \epsilon_1, \dots, \epsilon_n$  and  $d$  with  $(\epsilon_i, \epsilon_j) = 2$  if  $i = j$ ,  $-n_{i,j}$  if  $i \neq j$ , and  $(d, \epsilon_i) = \delta_{i,0}$ ,  $(d, d) = 0$ . Recall the dominant integral weights

$$P^+ = \{\lambda \in \mathfrak{h} \mid (\lambda, \epsilon_i) \in \mathbb{N}\}$$

that parametrize the irreducibles in  $\mathcal{O}_{int}$ . Note by the way that the canonical central element  $\delta = \sum_{i=0}^n \delta_i \epsilon_i$  must act as a scalar on any  $L(\lambda)$ . In the case that  $\lambda \in P^+$ , this scalar is a natural number, and is called the *level* of the module  $L(\lambda)$ .

In particular,  $d$  is a dominant integral weight since

$$(d, \epsilon_0) = 1, (d, \epsilon_1) = \dots = (d, \epsilon_n) = 0.$$

The integrable module  $L(d)$  is called the *basic representation of level one* – it is the “simplest” of all of the integrable representations. Note it really is of level one since  $(c, d) = 1$ . The goal is to give an explicit construction of the basic representation of  $\mathfrak{g}$ , discovered by Frenkel and Kac in 1980. I probably won't give the full proof, but I want to give you the gist.

Along the way I'll review our explicit construction of  $\mathfrak{g}$  from scratch.

- (1) Let  $\overset{\circ}{R} = \mathbb{Z}\epsilon_1 \oplus \dots \oplus \mathbb{Z}\epsilon_n$  be the root lattice associated to the underlying finite diagram, with inner product  $(\cdot, \cdot)$ . Let  $\overset{\circ}{\Delta}$  be the roots, i.e. the  $\alpha \in \overset{\circ}{R}$  with  $(\alpha, \alpha) = 2$ . Let  $\epsilon_1, \dots, \epsilon_n$  be the simple roots. Pick an orientation on the Dynkin diagram, let

$$\nu : \overset{\circ}{R} \times \overset{\circ}{R} \rightarrow \{\pm 1\}$$

be the associated asymmetry function, so it is bilinear and

$$\nu(\epsilon_i, \epsilon_j) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are not connected,} \\ 1 & \text{if } i \rightarrow j, \\ -1 & \text{if } i \leftarrow j, \\ -1 & \text{if } i = j. \end{cases}$$

Recall then that

$$\mathring{\mathfrak{g}} = \mathring{\mathfrak{h}} \oplus \bigoplus_{\alpha \in \mathring{\Delta}} \mathbb{C}e_\alpha$$

where  $\mathring{\mathfrak{h}} = \mathbb{C} \otimes_{\mathbb{Z}} \mathring{R}$ . The multiplication is defined so that  $\mathring{\mathfrak{h}}$  is abelian and

$$\begin{aligned} [\epsilon_i, e_\alpha] &= (\epsilon_i, \alpha)e_\alpha, \\ [e_\alpha, e_{-\alpha}] &= -\alpha \quad (\alpha \in \mathring{\Delta}), \\ [e_\alpha, e_\beta] &= 0 \quad (\alpha + \beta \notin \mathring{\Delta} \cup \{0\}), \\ [e_\alpha, e_\beta] &= \nu(\alpha, \beta)e_{\alpha+\beta} \quad (\alpha + \beta \in \mathring{\Delta}). \end{aligned}$$

Finally the invariant bilinear form on  $\mathring{\mathfrak{g}}$  extending that on  $\mathring{\mathfrak{h}}$  is defined by  $(e_\alpha, e_\beta) = -\delta_{\alpha, -\beta}$ .

- (2) The affine Lie algebra  $\mathfrak{g}$  is then constructed as

$$\mathfrak{g} = \mathring{\mathfrak{g}} \otimes \mathbb{C}[t, t^{-1}] \oplus \mathbb{C}\delta \oplus \mathbb{C}d$$

with Lie bracket defined by declaring  $\delta$  to be central and

$$\begin{aligned} [x \otimes t^m, y \otimes t^k] &= [x, y] \otimes t^{m+k} + m\delta_{m, -k}(x, y)\delta, \\ [d, x \otimes t^m] &= mx \otimes t^m. \end{aligned}$$

The Cartan subalgebra  $\mathfrak{h} = \mathring{\mathfrak{h}} \oplus \mathbb{C}\delta \oplus \mathbb{C}d$ . The bilinear form on  $\mathfrak{h}$  extends that on  $\mathring{\mathfrak{h}}$  so that  $\delta, d$  are orthogonal to  $\mathring{\mathfrak{h}}$ , and  $(d, \delta) = 1, (d, d) = (\delta, \delta) = 0$ . I think we will sometimes write for  $h \in \mathfrak{h}$   $\bar{h}$  for the projection of  $h$  onto  $\mathring{\mathfrak{h}}$  along this orthogonal direct sum decomposition. The zero-th simple root  $\epsilon_0$  is the element

$$\epsilon_0 = \delta - \theta$$

where  $\theta = \sum_{i=1}^n \delta_i \epsilon_i$  is the highest root of  $\mathring{\Delta}$ . The root system  $\Delta$  is

$$\Delta = \{m\delta + \gamma \mid m \in \mathbb{Z}, \gamma \in \mathring{\Delta}\} \cup \{m\delta \mid m \in \mathbb{Z} - \{0\}\}.$$

The real root spaces are all one dimensional, the imaginary root spaces are the  $\mathring{\mathfrak{h}} \otimes t^m$  so are  $n$  dimensional. The invariant bilinear form on  $\mathfrak{g}$  extending that on  $\mathfrak{h}$  satisfies  $(x \otimes t^m, y \otimes t^k) = \delta_{m, -k}(x, y)$ .

- (3) Introduce the *Heisenberg subalgebra*  $\mathfrak{t}$  of  $\mathfrak{g}$ :

$$\mathfrak{t} = \mathbb{C}\delta \oplus \bigoplus_{m \in \mathbb{Z} - \{0\}} \mathring{\mathfrak{h}} \otimes t^m.$$

Thus,  $\mathfrak{t} = \mathfrak{t}_- \oplus \mathfrak{t}_0 \oplus \mathfrak{t}_+$  where  $\mathfrak{t}_0 = \mathbb{C}\delta$ ,  $\mathfrak{t}_-$  has basis  $\epsilon_i \otimes t^m$  for  $i = 1, \dots, n, m < 0$  and  $\mathfrak{t}_+$  has basis  $\epsilon_i \otimes t^m$  for  $i = 1, \dots, n, m > 0$ . The multiplication is defined by

$$[\epsilon_i \otimes t^m, \epsilon_j \otimes t^{-m}] = m(\epsilon_i, \epsilon_j)$$

all other brackets being zero. We need the *canonical commutation relations module*. This is the unique irreducible highest weight  $\mathfrak{t}$ -module on which  $\delta$  acts as 1. It can be constructed as an induced module – like a Verma module. Here is the alternate description which is more handy: it is the vector space

$$S(\mathfrak{t}_-),$$

with highest weight vector 1. The action of  $\delta$  is as the identity. The action of  $\epsilon_i \otimes t^m \in \mathfrak{t}_-$  for  $m < 0$  is as multiplication by this operator. The action of  $\epsilon_i \otimes t^m \in \mathfrak{t}_+$  for  $m > 0$  is as the derivation that maps

$$\epsilon_j \otimes t^{-m} \mapsto m(\epsilon_i, \epsilon_j)1$$

and all other generators (including the h/w vector 1) are sent to zero. You of course need to check this really is a module – it is easiest just to see it is the Verma module. Then it follows easily that it is an irreducible representation – just argue directly using the non-degeneracy of the form that any non-zero vector can be sent to any other.

(4) Now consider the vector space

$$V = S(\mathfrak{t}_-) \otimes \mathbb{C}[\overset{\circ}{R}]$$

where  $\mathbb{C}[\overset{\circ}{R}]$  denotes the group algebra over  $\mathbb{C}$  of the abelian group  $\overset{\circ}{R}$ . To avoid confusion, we denote the standard basis of this group algebra by  $e^\alpha$  for  $\alpha \in \overset{\circ}{R}$  – then multiplication obeys the exponential laws! Now we are going to make  $V$  into a  $\mathfrak{g}$ -module. Let us start by specifying how  $\mathfrak{h}$  acts: first

$$\epsilon_i(f \otimes e^\alpha) = (\delta_{i,0} + (\epsilon_i, \alpha)).f \otimes e^\alpha$$

for  $i = 0, \dots, n, f \in S(\mathfrak{t}_-)$  and  $\alpha \in \overset{\circ}{R}$ . Second, for  $f \in S(\mathfrak{t}_-)$ ,

$$d(f \otimes e^\alpha) = -(\deg(f) + \frac{1}{2}(\alpha, \alpha)).f \otimes e^\alpha$$

where  $\deg(f)$  is the degree of this homogeneous polynomial in the  $\epsilon_i \otimes t^{-m}$  defined by declaring that  $\epsilon_i \otimes t^{-m}$  is of degree  $m$ . Before we go anywhere, let us understand what this says about the *character* of the module  $V$  that we haven't constructed yet. The vectors of the form  $1 \otimes e^\alpha$  for  $\alpha \in \overset{\circ}{R}$  contribute

$$e^{d+\alpha-\frac{1}{2}(\alpha,\alpha)\delta}$$

to the character. Then for a monomial  $f \in S(\mathfrak{t}_-)$  of degree  $m$ , the vector  $f \otimes e^\alpha$  contributes

$$e^{d+\alpha-(m+\frac{1}{2}(\alpha,\alpha))\delta}$$

to the character, i.e. the same as  $1 \otimes e^\alpha$  but shifted down by  $m$  lots of  $\delta$ . To work out the dimension of this weight space of  $V$ , we need to compute the dimension of the degree  $m$  part of  $S(\mathfrak{t}_-)$ . Recall  $\epsilon_i \otimes t^{-m}$  is of degree  $m$ . So to get a monomial of degree  $m$ , we need to write  $m = m_1 + 2m_2 + \dots + rm_r$  and then choose  $m_i$  generators of degree  $i$  (out of  $n$ ) for each  $i = 1, \dots, r$ . For  $n = 1$  this is just the number of partitions of  $m$  (take the partition  $1^{m_1}2^{m_2} \dots$ ). For general  $n$  it is the number of partitions of  $m$  whose parts are colored with  $n$  colors. Denote this number by  $P_n(m)$ . For example:

$$P_1(4) = 5, P_2(4) = 20.$$

Now the character of  $V$  is

$$\sum_{\alpha \in \mathring{R}} \sum_{m \geq 0} P_n(m) e^{d+\alpha-(m+\frac{1}{2}(\alpha,\alpha))\delta}.$$

(5) Note that

$$d + \alpha - (m + \frac{1}{2}(\alpha, \alpha))\delta = d + \alpha - (m + \frac{1}{2}(\alpha, \alpha))\epsilon_0 - (m + \frac{1}{2}(\alpha, \alpha))\theta.$$

Since  $(\alpha, \alpha) \geq 2$  and  $\theta$  is the highest root,  $\alpha - \frac{1}{2}(\alpha, \alpha)\theta < 0$ . This checks that the unique highest weight of  $V$  in the dominance ordering is  $d$ . When we've finally finished constructing  $V$ , it will be obvious that it is an irreducible h/w module generated by the vector  $1 \otimes e^0$  which has weight  $d$ . So it will be automatic that it is the irreducible module  $L(d)$ .

(6) Recall the Weyl group  $W$  looked like  $T \cdot \mathring{W}$ , where  $T$  was the *translation group*  $\{t_\alpha \mid \alpha \in \mathring{R}\}$  acting on  $\mathfrak{h}$  so that

$$t_\alpha(d) = d + \alpha - \frac{1}{2}(\alpha, \alpha)\delta.$$

The  $\mathring{W}$  part on the other hand fixes  $d$  because  $(d, \epsilon_i) = 0$  for each  $i = 1, \dots, n$ . All of  $W$  fixes  $\delta$ . Hence we can rewrite the character as

$$\sum_{\alpha \in \mathring{R}} t_\alpha \left( \sum_{m \geq 0} P_n(m) e^{d-m\delta} \right)$$

Since  $V$  is going to be an integrable representation  $W$  will act permuting the weight spaces. In particular, the  $W$ -orbit of the highest weight space will give all the contributions here with  $m = 0$ , and so on. This gives a nice picture of the weight spaces.

- (7) Now we have defined the action of  $\mathfrak{h}$ . Let us next define the action of the Heisenberg part  $\mathfrak{t}$ . This is easy: we just make  $\mathfrak{t}$  act on  $V = S(\mathfrak{t}_-) \otimes \mathbb{C}[\mathring{R}]$  on the first tensor in the way it acts on its canonical commutation relations module. Thus as a  $\mathfrak{t}$ -module,  $V$  looks like a direct sum

$$\bigoplus_{\alpha \in \mathring{R}} S(\mathfrak{t}_-) \otimes e^\alpha$$

of copies of its irreducible module. Now we can see why  $V$  is going to be irreducible overall. Start from the vector  $1 \otimes e^0$ . Acting by  $W$ , we get inside the module generated by this vector that all the vectors  $1 \otimes e^\alpha$  along the top lie in the submodule generated by  $1 \otimes e^0$ . Then acting with  $\mathfrak{t}$  we get everything. That's how we'll prove  $V$  is a highest weight module. It is also how we'll prove its irreducible: take a non-zero weight vector. It must lie in one of the irreducible Heisenberg module, so we can raise it up to some  $1 \otimes e^\alpha$  using  $\mathfrak{t}$ . Then using  $W$  we can conjugate it to the h/w vector. Hence it will be irreducible.

- (8) Well we've proved lots of things already before we had a right to. We've still got to define the action of the  $e_\alpha \otimes t^m$ 's for  $\alpha \in \mathring{\Delta}$  and  $m \in \mathbb{Z}$  – then after that we have to verify that the relations are satisfied so that  $V$  really is a  $\mathfrak{g}$ -module. Given those things the arguments sketched above will prove that  $V \cong L(d)$ . This is where the vertex operators come in. I'm now going to start working heavily with generating functions in indeterminates  $z, w, \dots$ . Most of the things I write down should be interpreted as elements of  $\text{End}(V)[[z^{\pm 1}]]$ , i.e. formal power series with coefficients in  $\text{End}(V)$ . For  $\alpha \in \mathring{\Delta}$ , let

$$P_\alpha(z) = \exp\left(\sum_{n \geq 1} \frac{\alpha \otimes t^{-n} z^n}{n}\right)$$

and

$$Q_\alpha(z) = \exp\left(-\sum_{n \geq 1} \frac{\alpha \otimes t^n z^{-n}}{n}\right).$$

These both make sense in  $\text{End}(V)[[z^{\pm 1}]]$ . Let

$$E_\alpha(z) = P_\alpha(z)Q_\alpha(z)e^\alpha z^\alpha s_\alpha$$

(noting it makes sense to compose  $P_\alpha(z)$  with  $Q_\alpha(z^{-1})$  because  $Q_\alpha$  actually only involves a finite number of non-zero operators on any vector of  $V$ ). Here,  $z^\alpha : V[[z^{\pm 1}]] \rightarrow V[[z^{\pm 1}]]$  is the operator  $z^\alpha(f \otimes e^\beta) = z^{(\alpha, \beta)}(f \otimes e^\beta)$ ,  $e^\alpha : V \rightarrow V$  is the operator  $f \otimes e^\beta \mapsto f \otimes e^{\alpha+\beta}$  and  $s_\alpha(f \otimes e^\beta) = \nu(\alpha, \beta)f \otimes e^\beta$ . Again  $E_\alpha(z)$  makes sense in

$\text{End}(V)[[z^{\pm 1}]]$ . Now expand this power series as

$$E_\alpha(z) = \sum_{m \in \mathbb{Z}} E_\alpha(m) z^{-m-1}.$$

This gives us operators  $E_\alpha(m) : V \rightarrow V$ . DEFINE the action of  $e_\alpha \otimes t^m$  on  $V$  to be this linear map  $E_\alpha(m)$ . Given our remarks along the way, we're done if we can prove the following theorem due to Kac and Frenkel:

**Theorem 3.54.** *The above definitions make  $V$  into an integrable  $\mathfrak{g}$ -module.*

*Proof.* We have to check relations! I am not going to do the whole thing – see Theorem 14.8 in Kac's book for that. I just want to check one relation to convince you that the generating functions are things you can work with. We'll need the formal delta function

$$\delta(z - w) = \sum_{r \in \mathbb{Z}} z^r w^{-r-1}.$$

Its basic property is

$$\text{Res}_{z=0} f(z) \delta(z - w) = f(w)$$

where  $\text{Res}_{z=0}$  denotes the  $z^{-1}$ -coefficient of a formal Laurent series in  $z$ . Of course you have to be careful: this only holds when it makes sense to multiply  $f(z)$  by  $\delta(z - w)$  as formal Laurent series. To prove this property, write  $f(z) = \sum f_j z^j$  then the left hand side is the  $z^{-1}$ -coefficient of

$$\sum_{j,k} f_j z^j z^k w^{-k-1}$$

which is  $\sum_k f_{-k-1} w^{-k-1} = f(w)$ . A consequence of this property: for formal Laurent series  $f(z, w)$  in  $z^{\pm 1}, w^{\pm 1}$ ,

$$f(z, w) \delta(z - w) = f(w, w) \delta(z - w)$$

whenever both sides make sense. To see this, multiply both sides by  $z^n$  and check the equality of  $\text{Res}_{z=0}$  using the basic property above: both sides give  $w^n f(w, w)$ .

Now I'm going to check the one relation

$$[e_\alpha \otimes t^r, e_\beta \otimes t^s] = \nu(\alpha, \beta) e_{\alpha+\beta} \otimes t^{r+s}$$

for  $\alpha + \beta \in \Delta$ . This transcribes into power series notation as the relation

$$[E_\alpha(z), E_\beta(w)] = \sum_{r,s \in \mathbb{Z}} \nu(\alpha, \beta) E_{\alpha+\beta}(z) \delta(z - w).$$

Let's really check this. The left hand side is

$$\sum_{r,s} \nu(\alpha, \beta) e_{\alpha+\beta} \otimes t^{r+s} z^{-r-1} w^{-s-1}$$

for sure. Substitute  $q = r + s$  to get

$$\sum_{q,s} \nu(\alpha, \beta) e_{\alpha+\beta} \otimes t^q z^{s-q-1} w^{-s-1} = \nu(\alpha, \beta) E_{\alpha+\beta}(z) \delta(z-w).$$

To prove this identity, we'll have to commute  $E_\alpha(z) = P_\alpha(z) Q_\alpha(z) e^{\alpha z^\alpha s_\alpha}$  with  $E_\beta(w) = P_\beta(w) Q_\beta(w) e^{\beta w^\beta s_\beta}$ . Note the vertex operator part acts only on  $S(\mathfrak{t}_-)$  while the  $e^{\alpha z^\alpha s_\alpha}$  part acts only on  $\mathbb{Z}[\overset{\circ}{R}]$ . Now I claim

$$Q_\alpha(z) P_\beta(w) = P_\beta(w) Q_\alpha(z) \left(1 - \frac{w}{z}\right)^{(\alpha, \beta)}$$

and

$$e^{\alpha z^\alpha s_\alpha} e^{\beta w^\beta s_\beta} = e^{\alpha+\beta z^\alpha w^\beta z^{(\alpha, \beta)} s_\alpha s_\beta} \nu(\alpha, \beta).$$

The second of these identities is easy to check from the definitions. To check the first one, you need to apply the following facts:

- (1)  $e^A e^B = e^B e^A e^{[A, B]}$  for two operators  $A$  and  $B$  such that  $[A, B]$  commutes with  $A$  and  $B$ ;
- (2)  $\exp(-\sum_{j \geq 1} x^j / j) = 1 - x$  (take log's!!).

Okay now using the two identities, we compute

$$E_\alpha(z) E_\beta(w) = \left(1 - \frac{w}{z}\right)^{(\alpha, \beta)} z^{(\alpha, \beta)} \nu(\alpha, \beta) \times \\ P_\alpha(z) P_\beta(w) Q_\alpha(z) Q_\beta(w) \times e^{\alpha+\beta z^\alpha w^\beta s_\alpha s_\beta}.$$

Hence the commutator  $[E_\alpha(z), E_\beta(w)]$  is equal to

$$\left[ z^{(\alpha, \beta)} \left(1 - \frac{w}{z}\right)^{(\alpha, \beta)} - w^{(\alpha, \beta)} \left(1 - \frac{z}{w}\right)^{(\alpha, \beta)} \right] \nu(\alpha, \beta) \times \\ P_\alpha(z) P_\beta(w) Q_\alpha(z) Q_\beta(w) \times e^{\alpha+\beta z^\alpha w^\beta s_\alpha s_\beta}.$$

Since  $\alpha + \beta \in \Delta$ ,  $(\alpha, \beta) = -1$ . Now we observe that

$$\left[ z^{-1} \left(1 - \frac{w}{z}\right)^{-1} - w^{-1} \left(1 - \frac{z}{w}\right)^{-1} \right] = \delta(z-w).$$

So we've got

$$\delta(z-w) \nu(\alpha, \beta) \times \\ P_\alpha(z) P_\beta(w) Q_\alpha(z) Q_\beta(w) \times e^{\alpha+\beta z^\alpha w^\beta s_\alpha s_\beta}.$$

Hence it equals

$$\delta(z-w) \nu(\alpha, \beta) \times \\ P_\alpha(z) P_\beta(z) Q_\alpha(z) Q_\beta(z) \times e^{\alpha+\beta z^{\alpha+\beta} s_{\alpha+\beta}}.$$

That is what we were after!  $\square$

**Exercise 9.** Go through the details of the proof that the module  $S(\mathfrak{t}_-)$  is an irreducible  $\mathfrak{t}$ -module.