Structure and Representation Theory of Infinite-dimensional Lie Algebras

David Mehrle

Abstract

Kac-Moody algebras are a generalization of the finite-dimensional semisimple Lie algebras that have many characteristics similar to the finite-dimensional ones. These possibly infinite-dimensional Lie algebras have found applications everywhere from modular forms to conformal field theory in physics. In this thesis we give two main results of the theory of Kac-Moody algebras. First, we present the classification of affine Kac-Moody algebras by Dynkin diagrams, which extends the Cartan-Killing classification of finite-dimensional semisimple Lie algebras. Second, we prove the Kac Character formula, which a generalization of the Weyl character formula for Kac-Moody algebras. In the course of presenting these results, we develop some theory as well.

Contents

Acknowledgements

First, I'd like to thank Bogdan Ion for being such a great advisor. I'm glad that you agreed to supervise this project, and I'm grateful for your patience and many helpful suggestions. Our meetings each week gave me something to look forward to, and I always left excited to get back to work.

Thanks to James Cummings, who introduced me to the wonderful world of abstract algebra, and thanks to Chris Bolognese, who first showed me that math was fun all those years ago. Further thanks should go to Richard Holman for introducing me to Lie algebras, although he was really trying to teach quantum physics at the time.

Chapter 1

Structure

Unlike the finite dimensional simple complex Lie algebras, which are described in terms of the classical matrix algebras and then classified by the Cartan matrices (or equivalently, Dynkin diagrams), Kac-Moody algebras are first constructed *from* **generalized Cartan matrices**, and a comparatively simple description is only recovered later.

Infinite dimensional Lie algebras arise by relaxing the condition that the off-diagonal entries of a Cartan matrix be integers between -3 and 0, and instead allowing them to be any nonnegative integer. The Lie algebras constructed from these matrices are known as **Kac-Moody algebras**, after Victor Kac and Richard Moody who independently discovered them.

Kac-Moody algebras include and subsume the finite-dimensional simple Lie algebras, and also include another class of Lie algebras called the **affine Lie algebras**. These are the first class of Kac-Moody algebra, and are constructed from the finite-dimensional semismile complex Lie algebras by adding a node to the Dynkin diagram. Like the finite-dimensional semisimple complex Lie algebras, the affine Lie algebras admit a classification in terms of Dynkin diagrams. For the remainder of the Kac-Moody algebras, those of so-called **indeterminate** type, little is known. See, for example, [Moo79].

In this chapter, we describe the structure of Kac-Moody algebras first in general, and then focusing on those of affine type. In section 1.1, we give an example that may be helpful to keep in mind while reading the remainder of the sections. In section 1.2, we define generalized Cartan matrices and describe how to create a Kac-Moody algebra. In section 1.3, we delineate the three types of generalized Cartan matrices. Section 1.4 discusses the symmetrizable generalized Cartan matrices and provides another nice characterization of the three types. We introduce Dynkin diagrams for generalized Cartan matrices in section 1.5. Following that, we classify those Dynkin diagrams of affine type in section 1.6. Finally, we briefly describe in section 1.7 other structural results about Kac-Moody algebras as are needed in chapter 2.

1.1 First example: affine $\mathfrak{sl}_2(\mathbb{C})$

In this section, we introduce the simplest example of a Kac-Moody algebra, called affine $\mathfrak{sl}_2$ because it is constructed from the classical Lie algebra $\mathfrak{sl}_2(\mathbb{C})$. Recall that the Lie algebra $\mathfrak{sl}_2$ is the Lie algebra of 2×2 complex matrices with vanishing trace. T by the three matrices

$$h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad e = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad f = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

span $\mathfrak{sl}_2$. The Lie bracket for this Lie algebra is commutator $[x, y] = xy - yx$; the generators satisfy the relations

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

The single root of $\mathfrak{sl}_2$ is just the integer 2, which is also it's Cartan matrix. It has the simplest Dynkin diagram. $\mathfrak{sl}_2$ is the Lie algebra of type A_1. We have a Cartan decomposition $\mathfrak{sl}_2 = \mathbb{C}f \oplus \mathbb{C}h \oplus \mathbb{C}e$. This is a decomposition of $\mathfrak{sl}_2$ into eigenspaces of $\mathrm{ad}(h) = [h, \cdot]$, with eigenvalues -2 for $\mathbb{C}f$, 2 for $\mathbb{C}e$, and 0 for $\mathbb{C}h$. The Killing form on $\mathfrak{sl}_2$ is given by $\langle x \mid y \rangle = \mathrm{tr}(xy)$.

To get the affine form of $\mathfrak{sl}_2$, we first form the **loop algebra** $\widetilde{\mathfrak{sl}_2}$ of $\mathfrak{sl}_2$, which is the Lie algebra of the group of all loops in SL_2, i.e., maps $S^1 \to SL_2$. Algebraically, this corresponds to extending the scalars to all Laurent polynomials over the complex numbers:

$$L(\mathfrak{sl}_2) \; = \; \mathfrak{sl}_2 \otimes_{\mathbb{C}} \mathbb{C}[z, z^{-1}].$$

The Lie bracket on $\mathfrak{sl}_2$ extends to one on its loop algebra via

$$[x \otimes f, y \otimes g] \; = \; [x, y] \otimes fg$$

for $x, y \in \mathfrak{sl}_2$ and $f, g \in \mathbb{C}[z, z^{-1}]$, or alternatively we can define the bracket on only those elements $x \otimes z^n$ for $x \in \mathfrak{sl}_2$ and integers n, as

$$[x \otimes z^n, y \otimes z^m] \; = \; [x, y] \otimes z^{m+n}$$

and extended by bilinearity. In fact, this second approach will be much more useful later.

As a vector space, $L(\mathfrak{sl}_n)$ is the space of 2×2 traceless matrices over the Laurent polynomials. Like $\mathfrak{sl}_2$, it has a decomposition into eigenspaces of $\mathrm{ad}(h)$ as

$$L(\mathfrak{sl}_2) \; = \; \left(\mathbb{C}f \otimes \mathbb{C}[z, z^{-1}]\right) \oplus \left(\mathbb{C}h \otimes \mathbb{C}[z, z^{-1}]\right) \oplus \left(\mathbb{C}e \otimes \mathbb{C}[z, z^{-1}]\right),$$

although here each eigenspace is infinite-dimensional over $\mathbb{C}$.

However, this is not yet a Kac-Moody algebra. From here, we introduce a 1-dimensional **central extension** of the loop algebra by forming a direct sum

$$\widetilde{\mathfrak{sl}_2} \; = \; L(\mathfrak{sl}_2) \oplus \mathbb{C}c,$$

and extend the Lie bracket on $L(\mathfrak{sl}_2)$ to one on this central extension by

$$[x \otimes z^n + \lambda c, y \otimes z^m + \mu c] = [x, y] \otimes z^{m+n} + n\,\mathrm{tr}(xy)\,\delta_{m,-n}\,c,$$

where $\delta_{i,j}$ is the Kronecker delta function. That this defines a Lie bracket is straightforward to verify. Note that c is in the center of this Lie algebra: for any $A \in \widetilde{\mathfrak{sl}_2}$, $[A, c] = 0$. Hence the name central extension.

The last step in making the Kac-Moody algebra associated with $\mathfrak{sl}_2$ is to add a **derivation –** that is, a function $\Delta\colon \widetilde{\mathfrak{sl}_2} \to \widetilde{\mathfrak{sl}_2}$ which obeys the **Leibniz rule**:

$$\Delta[x, y] = [\Delta x, y] + [x, \Delta y].$$

In our case, the derivation will be defined by $\Delta(x \otimes z^n + \lambda c) = z\frac{d}{dz}(x \otimes z^n)$. We form the **Kac-Moody algebra** $\widehat{\mathfrak{sl}_2}$ by a semi direct product $\widetilde{\mathfrak{sl}_2} \rtimes_\Delta \mathbb{C}d$, which is as a vector space $\widetilde{\mathfrak{sl}_2} \oplus \mathbb{C}d$ with bracket extended by the rule that $[d, A] = \Delta(A)$ for all $A \in \widetilde{\mathfrak{sl}_2}$ and linearity.

To summarize, the Kac-Moody algebra $\widehat{\mathfrak{sl}_2}$ is given as a vector space by

$$\widehat{\mathfrak{sl}_2} \;=\; \mathfrak{sl}_2 \otimes_{\mathbb{C}} \mathbb{C}[z, z^{-1}] \oplus \mathbb{C}c \oplus \mathbb{C}d,$$

with Lie bracket

$$\begin{aligned}&[x \otimes z^n + \lambda\, c + \alpha\, d\;,\; y \otimes z^m + \mu\, c + \beta\, d] \;=\\ &\qquad\qquad [x, y] \otimes z^{n+m} + \alpha m(y \otimes z^m) - \beta n(x \otimes z^n) + \delta_{n,-m}\, n\,\mathrm{tr}(xy)\, c.\end{aligned}$$

This is the Kac-Moody algebra of type $A_1^{(1)}$.

So all of that was rather complicated, but it did give us a concrete description of this algebra in terms of matrices with coefficients in Laurent polynomials, with two additional parameters: the central element c and derivation d. There is a simpler, if less concrete, description of the Kac-Moody algebra $A_1^{(1)}$. The Cartan matrix of $\mathfrak{sl}_2$ is the 1×1 matrix $A = (2)$. We add a row and column to form the matrix

$$A' = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}.$$

Remark 1.1.1. In general, given a Lie algebra $\mathfrak{g}$ of rank n with Cartan matrix $A = (a_{ij})_{1 \le i,j \le n}$, we add a row and a column to form a new matrix $A' = (a'_{ij})_{0 \le i,j \le n}$ according to the rules

$$\begin{aligned} a'_{ij} &= a_{ij} && \text{for } 1 \le i, j \le n \\ a'_{i,0} &= -\sum_{j=1}^{n} a_j a_{ij} && \text{for } 1 \le i \le n \\ a'_{0,j} &= -\sum_{i=1}^{n} c_i a_{ij} && \text{for } 1 \le j \le n \\ a'_{00} &= 2, \end{aligned}$$

where $\theta = \sum_{i=1}^n a_i\alpha_i$ is the highest root of $\mathfrak{g}$ and $\theta^\vee = \sum_{i=1}^n c_i\alpha_i^\vee$ is the corresponding coroot. A' will be the generalized Cartan matrix of a Kac-Moody algebra constructed from the Lie algebra $\mathfrak{g}$.

In the next section, we introduce how to do this procedure in general. Given a Generalized Cartan matrix A, or in fact any complex matrix, we will construct a Lie algebra $\mathfrak{g}(A)$.

1.2 Building Lie algebras from complex matrices

Just as the finite-dimensional Lie algebras can be constructed from the Cartan matrices, the Kac-Moody algebras are constructed from generalized Cartan matrices. We begin by defining a generalized Cartan matrix.

Definition 1.2.1. *An $n \times n$ matrix $A = (a_{ij})_{1\le i,j\le n}$ is a* **generalized Cartan matrix** *or* **(GCM)** *if the following four conditions hold:*

- *the entries of A are integers;*
- *the off-diagonal entries are negative;*
- *the diagonal entries are all equal to 2;*
- *$a_{ij} = 0$ if and only if $a_{ji} = 0$.*

In fact, we can construct an infinite-dimensional Lie algebra from any $n \times n$ complex matrix. Beginning in this manner makes the theory significantly easier, so we will work in this more general setting. Later, we will return to the generalized Cartan matrices.

Definition 1.2.2. *A* **realization** *of an $n \times n$ complex matrix $A = (a_{ij})_{1\le i,j\le n}$ and with rank ℓ is a triple $(\mathfrak{h}, \Pi, \Pi^\vee)$, where*

- *$\mathfrak{h}$ is a finite-dimensional $\mathbb{C}$-vector space of dimension $2n - \ell$;*
- *$\Pi = \{\alpha_1, \dots, \alpha_n\}$ and $\Pi^\vee = \{\alpha_1^\vee, \dots, \alpha_n^\vee\}$ are linearly independent subsets of $\mathfrak{h}^*$ and $\mathfrak{h}$, respectively;*
- *α_j and $\alpha_i^\vee$ satisfy $\alpha_j(\alpha_i^\vee) = a_{ij}$;*

Remark 1.2.3. The first thing that we should consider is whether or not such realizations exist for any complex $n \times n$ matrix A of rank ℓ. (They do!) We can construct one as follows. Without loss of generality, assume that the top ℓ rows of A are linearly independent (or permute rows if necessary) and form the $n \times (2n - \ell)$ matrix

$$A' = \begin{pmatrix} \mathrm{A} & \begin{matrix} 0 \\ I_{n-\ell} \end{matrix} \end{pmatrix} \tag{1.1}$$

To get a realization, take $\mathfrak{h} = \mathbb{C}^{2n-\ell}$ and let $\alpha_1, \dots, \alpha_n$ be the first n coordinate functions on $\mathfrak{h}$. Take $\alpha_1^\vee, \dots, \alpha_n^\vee$ to be the rows of A' gives a realization of A.

Definition 1.2.4. *From a matrix A and a realization $(\mathfrak{h}, \Pi, \Pi^\vee)$, we may construct a Lie algebra denoted $\tilde{\mathfrak{g}}(A)$. The Lie algebra $\tilde{\mathfrak{g}}(A)$ is generated by $e_1, \dots, e_n$, $f_1, \dots, f_n$, and x for all $x \in \mathfrak{h}$, subject to the relations*

$$\begin{array}{rcll} [x, y] & = & 0 & \text{for all } x, y \in \mathfrak{h} \\ [e_i, f_j] & = & 0 & \text{if } i \neq j \\ [e_i, f_i] & = & \alpha_i^\vee & \text{for all } i \\ [x, e_i] & = & \alpha_i(x) e_i & \text{for all } x \in \mathfrak{h} \text{ and all } i \\ [x, f_i] & = & -\alpha_i(x) f_i & \text{for all } x \in \mathfrak{h} \text{ and all } i \end{array} \tag{1.2}$$

It is not immediately obvious that this lie algebra $\tilde{\mathfrak{g}}(A)$ depends only on A and not the realization chosen. The next proposition shows that the notation $\tilde{\mathfrak{g}}(A)$ is sensible, insofar as the Lie algebra doesn't depend on the choice of realization but only the matrix A. But first, a notion of isomorphic realizations is in order.

We consider two realizations

$$(\mathfrak{h}, \{\alpha_1, \dots, \alpha_n\}, \{\alpha_1^\vee, \dots, \alpha_n^\vee\}) \quad \text{and} \quad (\mathfrak{g}, \{\beta_1, \dots, \beta_n\}, \{\beta_1^\vee, \dots, \beta_n^\vee\})$$

isomorphic if there is a vector space isomorphism $\phi\colon \mathfrak{h} \to \mathfrak{g}$ such that $\phi^*(\alpha_i) = \beta_i$ and $\phi(\alpha_i^\vee) = \beta_i^\vee$, where ϕ^* is the map of dual spaces induced by ϕ.

Proposition 1.2.5. *The Lie algebra $\tilde{\mathfrak{g}}(A)$ is independent of the realization of A.*

Proof. To show this, it is sufficient to show that any two realizations of A are isomorphic. This is accomplished by showing that they are all isomorphic to the same realization, which was constructed earlier in remark 1.2.3.

Without loss of generality, assume that the top ℓ rows of A are linearly independent. Let $(\mathfrak{h}, \Pi, \Pi^\vee)$ be a realization of A. Complete Π to a basis for $\mathfrak{h}^*$ by adding elements $\alpha_{n+1}, \dots, \alpha_{2n-\ell}$, and consider the $n \times (2n - \ell)$ matrix A' with ij entry $a'_{ij} = \alpha_j(\alpha_i^\vee)$.

Because $\alpha_j(\alpha_i^\vee) = a_{ij}$, A' has the form

$$A' = \begin{pmatrix} \text{A} & \begin{matrix} B \\ D \end{matrix} \end{pmatrix}$$

for an $\ell \times (n - \ell)$ matrix B and an $(n - \ell) \times (n - \ell)$ matrix D. By row-reducing, in particular by adding linear combinations of $\alpha_1, \dots, \alpha_\ell$ to $\alpha_{n+1}, \dots, \alpha_{2n-\ell}$, we may assume that B is 0. Similarly, row reduction can change A' into the form $D = I$. Then A' is the same matrix as in equation (1.1). Hence, the map $\mathfrak{h}^\star \to (\mathbb{C}^{2n-\ell})^\star$ that replaces $\alpha_1, \dots, \alpha_n$ by the relevant linear combination is an isomorphism of realizations. Thus, each realization is isomorphic to this one, and so the Lie algebra $\tilde{\mathfrak{g}}(A)$ doesn't depend on the realization. ☐

This Lie algebra $\tilde{\mathfrak{g}}(A)$ is not quite a Kac-Moody algebra yet – hence the tilde. But $\tilde{\mathfrak{g}}(A)$ does have some important structural properties that will be carried over into the Kac-Moody algebra.

For any realization $(\mathfrak{h}, \Pi, \Pi^\vee)$, we can make sense of $\alpha_1, \ldots, \alpha_n$ as the roots of the Lie algebra, in analogy with the finite-dimensional case and considering the defining relations on the generators.

Definition 1.2.6. *For a matrix A with realization $(\mathfrak{h}, \Pi, \Pi^\vee)$, Π is the set of* **simple roots**, *$\Pi^\vee$ the set of* **simple coroots**, *and*

$$Q = \mathbb{Z}\alpha_1 + \mathbb{Z}\alpha_2 + \ldots + \mathbb{Z}\alpha_n$$

is the **root lattice**. *Further set $Q^+ = \mathbb{Z}_+\alpha_1 + \ldots + \mathbb{Z}_+\alpha_n$.*

Let $\tilde{\mathfrak{n}}^+$ be the sub-algebra of $\tilde{\mathfrak{g}}(A)$ generated by $e_1, \ldots, e_n$ and $\tilde{\mathfrak{n}}^-$ be the sub-algebra generated by $f_1, \ldots, f_n$. Note that there is an involution $\tilde{\omega}$ of $\tilde{\mathfrak{g}}(A)$ defined by

$$\tilde{\omega}(e_i) = -f_i, \quad \tilde{\omega}(f_i) = -e_i, \quad \tilde{\omega}(x) = -x$$

for all $x \in \mathfrak{h}$. This important automorphism is called the **Chevalley involution**.

The following two lemmas provide a structural result about the Lie algebra $\tilde{\mathfrak{g}}(A)$ that shows there is a triangular decomposition for these Lie algebras.

Lemma 1.2.7 ([Kac94, Theorem 1.2]). *$\tilde{\mathfrak{n}}^+$ is freely generated by $e_1, \ldots, e_n$, and $\tilde{\mathfrak{n}}^-$ is freely generated by $f_1, \ldots, f_n$;*

Proof. Let V be an n-dimensional complex vector space with basis $v_1, \ldots, v_n$. Let $\lambda \in \mathfrak{h}^*$. The proof proceeds by defining for each such λ a representation of $\tilde{\mathfrak{g}}(A)$ on the tensor algebra $T(V)$ over an n-dimensional vector space V. It suffices to define the action of the generators of $\tilde{\mathfrak{g}}(A)$ on the generators of $T(V)$.

$$\begin{array}{rcll} f_i \cdot a & = & v_i \otimes a & \text{for all } a \in T(V) \\ x \cdot 1 & = & \lambda(x) & \text{for } x \in \mathfrak{h} \subseteq \tilde{\mathfrak{g}}(A), \text{ and inductively,} \\ x \cdot (v_i \otimes a) & = & -\alpha_i(x) v_i \otimes a + v_i \otimes (x \cdot a) & \text{for } x \in \mathfrak{h}, a \in T(V) \\ e_i \cdot 1 & = & 0 & \text{and inductively,} \\ e_i \cdot (v_j \otimes a) & = & \delta_{ij}\alpha_i^\vee a + v_j \otimes (e_i \cdot a) & \text{for } a \in T(V) \end{array}$$

That this is a bona-fide representation of $\tilde{\mathfrak{g}}(A)$ satisfying the relations in (1.2) can be easily yet tediously checked (if you're a masochist – otherwise, see [Kac94, Theorem 1.2] or [Car05, §14.2]).

Now consider the action of $\tilde{\mathfrak{n}}^-$ on $T(V)$, given by the first equation above. The action of a generator f_i is left-multiplication by a basis element v_i, so we get a Lie algebra homomorphism $\phi\colon \tilde{\mathfrak{n}}^- \to \text{Lie}(T(V))$, where $\text{Lie}(T(V))$ is the Lie algebra structure on $T(V)$ with Lie bracket the commutator. Because $T(V)$ is the free associative algebra on $\{v_1, \ldots, v_n\}$, the free Lie algebra on $\{v_1, \ldots, v_n\}$ lies in $\mathfrak{L}(T(V))$, and lies in the image of ϕ. Furthermore, there is an inverse to ϕ defined by $\psi(v_i) = f_i$, and ψ has domain the free Lie algebra on $\{v_1, \ldots, v_n\}$. So $\tilde{\mathfrak{n}}^-$ is a free Lie algebra with generators $f_1, \ldots, f_n$. Applying the involution $\tilde{\omega}$, which sends $\tilde{\mathfrak{n}}^-$ to $\tilde{\mathfrak{n}}^+$, we see that $\mathfrak{n}^+$ is also freely generated by $e_1, \ldots, e_n$. □

Lemma 1.2.8 ([Kac94, Theorem 1.2]). *There is a triangular decomposition for $\tilde{\mathfrak{g}}(A)$. In particular, $\tilde{\mathfrak{g}}(A) = \tilde{\mathfrak{n}}^+ \oplus \mathfrak{h} \oplus \tilde{\mathfrak{n}}^-$.*

Proof. We keep the same notation as in the proof of lemma 1.2.7. Observe that $I = \tilde{\mathfrak{n}}^- + \mathfrak{h} + \tilde{\mathfrak{n}}^+$ is an ideal of $\tilde{\mathfrak{g}}(A)$ containing all the generators, and so $I = \tilde{\mathfrak{g}}(A)$. It just remains to show that the sum is direct. Let $w \in \tilde{\mathfrak{n}}^-$, $x \in \mathfrak{h}$, and $u \in \tilde{\mathfrak{n}}^+$ such that $w + x + u = 0$. Then this element acts on $T(V)$ by the zero endomorphism, and in particular, we have that $w \cdot 1 + \lambda(x) = 0$. But because $w \cdot 1$ and $\lambda(x)$ are in different graded components of the tensor algebra, it must be that both $w \cdot 1 = 0$ and $\lambda(x) = 0$. Because this holds for all λ, we have that $x = 0$. So $w + x + u = 0$ only if $x = w = 0$, so it must be the case that $u = 0$ as well. Thus, the sum is direct. □

The decomposition as $\tilde{\mathfrak{g}}(A) = \tilde{\mathfrak{n}}^- \oplus \mathfrak{h} \oplus \tilde{\mathfrak{n}}^+$ can actually be refined even further, into what is called the **Cartan decomposition** of the Lie algebra $\tilde{\mathfrak{g}}(A)$. Essentially, this is a decomposition of $\tilde{\mathfrak{g}}(A)$ into the eigenspaces of the adjoint operators for $\mathfrak{h}$, which share eigenspaces because elements of $\mathfrak{h}$ commute.

Proposition 1.2.9.

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

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

Proof. First, it is evident that $\mathfrak{h} = \tilde{\mathfrak{g}}_0$. Further, claim that

$$\tilde{\mathfrak{n}}^+ \subseteq \sum_{\alpha \in Q^+} \tilde{\mathfrak{g}}_\alpha.$$

To establish this claim, observe that for any $x \in \mathfrak{h}$, $[x, e_i] = \alpha_i(x)e_i$ for some $\alpha_i \in Q^+$, and if $w, u \in \tilde{\mathfrak{n}}^+$ with $[x, w] = \beta(x)w$ and $[x, u] = \gamma(x)u$, then

$$\begin{aligned} [x, [w, u]] &= -[w, [u, x]] - [u, [x, w]] \\ &= [w, \gamma(x)u] - [u, \beta(x)w] \\ &= (\beta(x) + \gamma(x))[w, u] \quad \in \quad \tilde{\mathfrak{g}}_{\beta+\gamma}. \end{aligned}$$

Similarly, $\tilde{\mathfrak{n}}^- \subseteq \sum_{\alpha \in Q^+} \tilde{\mathfrak{g}}_{-\alpha}$. These two inclusions show that $\tilde{\mathfrak{g}}(A) = \sum_{\alpha \in Q} \tilde{\mathfrak{g}}_\alpha$.

To show that the sum is direct, let $v_i \in \tilde{\mathfrak{g}}_{\beta_i}$ for some distinct roots β_i, for $i = 1, \ldots, k$. Suppose that $v_1 + v_2 + \ldots + v_k = 0$, yet not all of the v_i are zero. Choose t minimal such that $v_1 + \ldots + v_t = 0$ yet not all v_j are zero for $j = 1, \ldots, t$. Then, for all $x \in \mathfrak{h} = \tilde{\mathfrak{g}}_0$,

$$0 = [x, 0] = [x, v_1 + \ldots + v_t] = [x, v_1] + \ldots + [x, v_t] = \beta_1(x)v_1 + \ldots + \beta_t(x)v_t \quad (1.3)$$

Additionally,

$$\beta_t(x)v_1 + \ldots + \beta_t(x)v_t = 0 \quad (1.4)$$

And by subtracting (1.3) from (1.4), we see that

$$(\beta_1(x) - \beta_t(x))v_1 + \ldots + (\beta_{t-1}(x) - \beta_t(x))v_t = 0$$

But then by the minimality of t, this shows that for each i, $(\beta_i(x) - \beta_t(x))v_i = 0$. This is true for all $x \in \mathfrak{h}$, and because we assumed the β_i were distinct, there is $x \in \mathfrak{h}$ such that $\beta_i(x) - \beta_t(x) \neq 0$. Hence, it must be the case that $v_i = 0$ for all $i = 1, \ldots, t-1$. The equation $v_1 + \ldots + v_t = 0$ shows that $v_t = 0$ as well, contradicting the assumption. Therefore, the sum is direct. □

Fortunately, these root spaces are fairly tame, in that each is finite dimensional. So if $\tilde{\mathfrak{g}}(A)$ is infinite-dimensional, it must be that the direct sum is infinite, and therefore the root lattice must be as well.

Proposition 1.2.10. *Furthermore, each $\mathfrak{g}_\alpha$ is finite dimensional for all $\alpha \in Q$.*

Proof. To show that each $\tilde{\mathfrak{g}}_\alpha$ is finite dimensional, suppose that $\alpha \neq 0$ (if $\alpha = 0$, then $\tilde{\mathfrak{g}}_0 = \mathfrak{h}$ with dimension $2n - \ell$), and further assume that $\alpha \in Q^+$. By the definition of Q, α is a $\mathbb{Z}$-linear combination of $\alpha_1, \ldots, \alpha_n$, say $\alpha = a_1\alpha_1 + \ldots + a_n\alpha_n$ with $a_i \in \mathbb{Z}$. Let $a = a_1 + \ldots + a_n$. Further, $\tilde{\mathfrak{g}}_\alpha \subseteq \tilde{\mathfrak{n}}^+$, which is spanned by Lie monomials in $e_1, \ldots, e_n$. One of these terms lies in the root space $\tilde{\mathfrak{g}}_\alpha$ exactly when it contains each e_i precisely a_i times, and there are no more than a of these terms. Hence, $\dim \tilde{\mathfrak{g}}_\alpha \leq a$.

Finally this is a grading of $\tilde{\mathfrak{g}}(A)$ because $[\tilde{\mathfrak{g}}_\alpha, \tilde{\mathfrak{g}}_\beta] \subseteq \tilde{\mathfrak{g}}_{\alpha+\beta}$ by the Jacobi identity. □

Proposition 1.2.11. *Among the ideals of $\tilde{\mathfrak{g}}(A)$ that meet $\mathfrak{h}$ trivially, there is a unique maximal ideal $\mathfrak{j}$.*

Proof. Let J be any ideal of $\tilde{\mathfrak{g}}(A)$ such that $\mathfrak{h} \cap J = 0$. By proposition 1.2.9, we have a decomposition of $\tilde{\mathfrak{g}}(A)$, and so

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

Because each $\tilde{\mathfrak{g}}_\alpha$ falls entirely inside either $\tilde{\mathfrak{n}}^+$ or $\tilde{\mathfrak{n}}^-$, it follows that $J = (\tilde{\mathfrak{n}}^- \cap J) \oplus (\tilde{\mathfrak{n}}^+ \cap J)$.

Now consider the ideal $\mathfrak{j}$ generated by all other ideals J such that $J \cap \mathfrak{h} = 0$. Because each such ideal lies entirely inside $\tilde{\mathfrak{n}}^+ \oplus \tilde{\mathfrak{n}}^-$, it must be that $\mathfrak{j}$ avoids $\mathfrak{h}$ as well. Clearly $\mathfrak{j}$ is maximal among ideals intersecting $\mathfrak{h}$ trivially. □

Following this proposition, it is now possible to state the definition of the Kac-Moody algebra associated to a generalized Cartan matrix A.

Definition 1.2.12. *Let A be a generalized Cartan matrix and let $\tilde{\mathfrak{g}}(A)$ be the Lie algebra associated with A defined according to the relations in equation* (1.2). *Let $\mathfrak{j}$ be the unique maximal ideal that meets $\mathfrak{h}$ trivially. The* Kac-Moody algebra *with GCM A is defined by*

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

Why do we quotient out by the unique maximal ideal $\mathfrak{j}$? Well, we are seeking an infinite-dimensional analogue of the simple Lie algebras, and by quotienting by the maximal ideal, we ensure that our Kac-Moody algebra is simple.

The Kac-Moody algebra $\mathfrak{g}(A)$ inherits many properties from its precursor $\tilde{\mathfrak{g}}(A)$. Because the intersection between $\mathfrak{h}$ and the maximal ideal is trivial, it is reasonable to identify $\mathfrak{h}$ with its image in $\mathfrak{g}(A)$. Moreover, $\mathfrak{g}(A)$ has a Cartan decomposition into root spaces as

$$\mathfrak{g}(A) = \bigoplus_{\alpha \in Q^+} \mathfrak{g}_{-\alpha} \oplus \mathfrak{h} \oplus \bigoplus_{\alpha \in Q^+} \mathfrak{g}_{\alpha}, \tag{1.5}$$

where $\mathfrak{g}_\alpha$ is the eigenspace of the root α with respect to the adjoint operator of the Cartan subalgebra. If A is a Cartan matrix as opposed to a GCM, then $\mathfrak{g}(A)$ is a semisimple finite dimensional Lie algebra with Cartan matrix A.

A presentation of $\mathfrak{g}(A)$ with generators and relations will be given later in remark 1.4.8, once we define symmetrizable GCMs.

1.3 Three types of generalized Cartan matrices

The structure of a Kac-Moody algebra depends on the generalized Cartan matrix in a crucial way, and the classification of Kac-Moody algebras can be accomplished by classifying the GCM's. In this section, generalized Cartan matrices are placed into three categories: finite, affine, and indeterminate type. Those matrices of finite type correspond to the finite-dimensional semisimple Lie algebras, and the matrices of affine type are called the affine Lie algebras, and like the finite dimensional simple Lie algebras, they admit a classification by (extended) Dynkin diagrams. The big theorem here is Theorem 1.3.10, although its corollary 1.3.11 is generally more useful.

Definition 1.3.1. *Following this proposition, it makes sense to consider GCMs only up to equivalence under permutation of rows and columns. Two such GCMs are called* **equivalent***. Additionally a GCM A is called* **indecomposable** *if it is not equivalent to a block matrix $A_1 \oplus A_2$, where A_1, A_2 are smaller GCMs.*

Proposition 1.3.2. *If two GCM's A and A' are equivalent up to permutation of the rows and columns, then the Kac-Moody algebras $\mathfrak{g}(A)$ and $\mathfrak{g}(A')$ are isomorphic.*

Proof. This follows from proposition 1.2.5. See [Kac94, Proposition 1.1]. □

The definition of the three types of GCM rely on the following convention.

Definition 1.3.3. *Let $v \in \mathbb{R}^n$. Write $v \geq 0$ if all components of v are nonnegative, and $v > 0$ if all components are positive. Using these conventions, define a partial order on vectors by $x \geq y$ if and only if $x - y \geq 0$, and similarly $x > y$ if and only if $x - y > 0$.*

Definition 1.3.4. *An $n \times n$ GCM A has*

- **finite type** *if* $\det A \neq 0$*, and there exists* $v > 0$ *such that* $Av > 0$*, and* $Av \geq 0$ *implies* $v > 0$ *or* $v = 0$*;*
- **affine type** *if* $\operatorname{rank} A = n - 1$*, and there exists* $v > 0$ *such that* $Av = 0$*, and* $Av \geq 0$ *implies* $Av = 0$*;*
- **indefinite type** *if there exists* $v > 0$ *such that* $Av < 0$*, and* $Av \geq 0$ *and* $v \geq 0$ *implies* $v = 0$*.*

Remark 1.3.5. The first thing to observe is that the three categories are disjoint. Clearly affine and finite type are distinct, because the determinant of A is nonzero if and only if A has full rank. If A is of finite type, then there is some $v > 0$ such that $Av > 0$, so A cannot also be of indefinite type, because $Av > 0$ and $v > 0$ implies $v = 0$, a contradiction. Similarly, if A is of affine type, then there is $v > 0$ such that $Av = 0$, but to be of indefinite type, $Av \geq 0$ and $v \geq 0$ implies that $v = 0$, a contradiction. Hence, all three types are genuinely distinct.

There is something very important to notice about the finite type GCMs – namely, these are equivalent to the classical Cartan matrices, although the definition may appear different. The next theorem makes precise the fact that those Cartan matrices coming from the theory of finite dimensional semisimple complex Lie algebras are precisely the generalized Cartan matrices of finite type, which links the finite dimensional theory of Lie algebras to that of Kac-Moody algebras. Hence, the term *generalized* Cartan matrix is an apt description.

Theorem 1.3.6. *An indecomposable GCM A has finite type if and only if A is a Cartan matrix. In particular, A has finite type if and only if*

- $a_{ij} \in \{0, -1, -2, -3\}$ *for* $i \neq j$*;*
- $a_{ij} \in \{-2, -3\} \implies A_{ji} = -1$*;*
- *and the quadratic form*

$$2\sum_{i=1}^{n} x_i^2 - \sum_{i \neq j} a_{ij}a_{ji}x_ix_j$$

is positive definite.

Proof. See [Car05, Theorem 15.9]. □

The next task is to show that that any GCM is one of the above three types. As with many things in Lie theory, this just reduces to a difficult linear algebra problem. The following boring, linear-algebraic fact is essential to this trichotomy. Nevertheless, I at least had never it before, so I will include a proof.

Lemma 1.3.7. *A system of real inequalities $u^i(x) = \sum_{j=1}^n u^i_j x_j > 0$ for $i = 1, \ldots, m$ has a solution if and only if there is no nontrivial dependence $\sum_{i=1}^m \lambda_i u^i = 0$ with $\lambda_i \leq 0$ for all $i = 1, \ldots, m$.*

Proof. Let $x \in \mathbb{R}^n$ be a solution to the system of inequalities: $u^i(x) > 0$ for all i. Suppose $\lambda_1 u^1 + \ldots + \lambda_m u^m = 0$ for some $\lambda_i \leq 0$, $i = 1, \ldots, n$. Then

$$\lambda_1 u^1(x) + \lambda_2 u^2(x) + \ldots + \lambda_m u^m(x) = 0,$$

but $\lambda_i \leq 0$ and $u^i(x) > 0$, so each individual term is negative. Hence, we must have $\lambda_i = 0$ for all $i = 1, \ldots, n$.

Conversely, suppose that there is no nontrivial linear dependence among the dual vectors u^i with $\lambda_i \leq 0$ for all i. Set

$$S = \left\{ \sum_{i=1}^m \lambda_i u^i \;\middle|\; \lambda_i \leq 0, \;\; \sum_{i=1}^m \lambda_i = -1 \right\}.$$

Note that $S \subseteq \mathbb{R}^n$ is compact, and so there is $x \in S$ with $\|x\| \leq \|y\|$ for all other $y \in S$. Furthermore, $x \neq 0$ because S doesn't contain the zero vector; otherwise there would be a nontrivial dependence among the u^i. I claim that this x is the desired solution to the system of inequalities $u^i(x) > 0$.

Let $y \in S$ be distinct from x. Then because S is convex, $ty + (1-t)x \in S$ for all $t \in [0,1]$. So $\|ty + (1-t)x\| \geq \|x\|$, or reinterpreting in terms of inner products,

$$\langle ty + (1-t)x, ty + (1-t)x \rangle \geq \langle x, x \rangle.$$

Rearranging, we see that

$$t^2 \langle y - x, y - x \rangle + 2t \langle y - x, x \rangle \geq 0,$$

for all $t \in [0,1]$. In particular, for $t > 0$, this shows that

$$t \langle y - x, y - x \rangle + 2 \langle y - x, x \rangle \geq 0$$

for all $0 < t \leq 1$. Therefore, $\langle y - x, x \rangle \geq 0$, or $\langle y, x \rangle \geq \langle x, x \rangle > 0$. In particular, taking y to be u^i for any i, we have shown that $u^i(x) > 0$. □

Following this lemma, we have a sequence of other lemmas that lead to the proof of Theorem 1.3.10, which shows that any GCM is one of finite, affine, or indefinite type.

Lemma 1.3.8. *If $A = (a_{ij})$ is any $n \times m$ real matrix such that there is no $u \geq 0$, $u \neq 0$ with $A^T u \geq 0$, then there exists $v > 0$ such that $Av < 0$.*

Proof. We want to find a vector v such that $v_j > 0$ for all $j = 1, \ldots, m$ and $\lambda_i = (Av)_i = \sum_{j=1}^m a_{ij} v_j < 0$. This is equivalent to solving the following system in the variables v_j:

$$\begin{cases} -\lambda_i > 0 & i = 1, \ldots, n \\ v_j > 0 & j = 1, \ldots, m. \end{cases} \tag{1.6}$$

By the previous lemma, this system has a solution if and only if there is no nontrivial dependence among the equations with nonpositive coefficients.

Assume for contradiction that such a nontrivial relation exists, say

$$\sum_{j=1}^{m} b_j v_j + \sum_{i=1}^{n} u_i(-\lambda_i) = 0,$$

where $b_j \leq 0$, $u_i \leq 0$. Rearranging, we see that $v^T b = v^T A^T u$ with vectors $b = (b_j)$ and $u = (u_i)$ nonpositive. This forces $A^T u = a$. Therefore, we have a vector $-u \geq 0$, $-u \neq 0$ such that $A^T(-u) = -b \geq 0$, which contradicts the assumption.

Therefore, the system (1.6) has a solution in the variables v_i, as desired. $\square$

Lemma 1.3.9. *Let A be an $n \times n$ indecomposable GCM. Then $Au \geq 0$ and $u \geq 0$ imply that either $u > 0$ or $u = 0$. In particular, u cannot both have some coordinates zero and others strictly positive.*

Proof. Suppose that $u \neq 0$ satisfies $Au \geq 0$ and $u \geq 0$. Permute the indices such that $u_i = 0$ for $i \leq s$ and $u_i > 0$ for $i > s$ for some s. Then because A is a GCM, we know that $a_{ij} \leq 0$ for $i \neq j$ and $a_{ij} = 0$ if and only if $a_{ji} = 0$. And because $Au \geq 0$, this means that $a_{ij} = a_{ji} = 0$ for $i \leq s$ and for $j > s$ – in particular, A is decomposable, in contradiction to the assumptions. $\square$

The following theorem is due to Vinberg [Vin71].

Theorem 1.3.10. *An indecomposable generalized Cartan matrix A is either finite type, affine type, or indefinite type (see Definition 1.3.4). Moreover, the type of A^T is the same as that of A.*

Proof. Consider the set of vectors $K_A = \{u \mid Au \geq 0\}$; it forms a convex cone. By the previous lemma, we cannot have any vectors $u \in K_A$ such that u has coordinates both strictly positive and zero. Hence,

$$K_A \cap \{u \mid u \geq 0\} \subsetneq \{u \mid u > 0\} \cup \{0\}.$$

Assume that $K_A \cap \{u \geq 0\}$ is nontrivial, containing more than just the zero vector. Then the intersection contains some vector $u > 0$. If there is no element $v \in K_A$ with $v \notin \{u > 0\}$, then K_A is entirely contained within $\{u > 0\} \cap \{0\}$. This is case (A) below. Otherwise, say $v \in K_A \setminus \{u > 0\}$. Then by the convexity of K_A, all vectors between u and v are contained with in K_A, say this is some line segment $L = \{tu + (1-t)v \mid t \in [0,1]\}$. Because $u > 0$ yet $v \leq 0$, L must meet the boundary of $\{u \geq 0\}$ at some point. However, this is only possible if that point is 0, again by Lemma 1.3.9. Hence, v lies on the line between u and $-u$, and L is a segment of this line. Let U be the subspace generated by u – the line through u and the origin. Because $v \leq 0$, it must be that v is between 0 and $-u$ no U. Therefore, $U \subseteq K_A$. Now assume that there is some $w \in K_A$

yet w is not in the linear subspace U. Again by the convexity of K_A, this means that the plane P generated by U and w is contained in K_A. However P must meet the boundary of the cone $\{u \geq 0\}$ at some point other than the origin, which is forbidden by Lemma 1.3.9. Therefore, $K_A = U$. Moreover, because both $u \in K_A$ and $-u \in K_A$, it must be that $Au = 0$, and so $K_A = \{u \mid Au = 0\}$. This is case (B).

So we have two cases in the occasion that $K_A \cap \{u \mid u \geq 0\}$ is nontrivial, each stated and proved below.

Claim 1.3.10.1 (Case 1). *K_A is entirely contained within $\{u > 0\} \cup \{0\}$ and contains no one-dimensional subspace.*

Proof. The first case 1.3.10.1 here is equivalent to A being of finite type (see Definition 1.3.4). Indeed, if $Av \geq 0$ for some v, then either $v > 0$ or $v = 0$. And moreover, because K_A doesn't contain any linear subspace, A cannot be singular. So A is surjective as a linear transformation, and thus there exists u with $Au > 0$; u is not itself zero because A has full rank. So $u > 0$ because $u \in K_A \subseteq \{u > 0\} \cup \{0\}$. □

Claim 1.3.10.2 (Case 2). *$K_A = \{u \mid Au = 0\}$ and is a one-dimensional subspace.*

Proof. The second case 1.3.10.2 is equivalent to A being of affine type. Here, we know that there is $u > 0$ such that $Au = 0$, and the kernel of A is exactly the one-dimensional subspace K_A, so the rank of A is $n - 1$. □

Moreover, if A is of finite or affine type, there is no $v > 0$ such that $Av < 0$. By the contrapositive of Lemma 1.3.8, this shows that there is some $u \geq 0$, $u \neq 0$ such that $u \in K_{A^T}$. Hence, we know that K_{A^T} falls into either case (A) or case (B). If A is of finite type, then A is nonsingular and so too is A^T, so A^T must also be of finite type. Similarly, if A is of affine type, then A^T also has rank $n-1$, so A^T is also affine type. Finally, if A is indefinite type, then A^T is also of indefinite type.

Throughout we have assumed that the intersection $K_A \cap \{u \mid u \geq 0\}$ is nontrivial. The final option is when the intersection $K_A \cap \{u \mid u \geq 0\}$ is trivial, containing only the origin, then $Au \geq 0$ and $u \geq 0$ implies that $u = 0$. Then we are in the case of Lemma 1.3.8. Applying the lemma, we see that both A and A^T are of indefinite type.

This concludes the proof of Theorem 1.3.10. □

The proof of the above theorem hints at a nice way to characterize the three types of generalized Cartan matrices.

Corollary 1.3.11. *There exists $x > 0$ such that*

- *$Ax > 0$ if and only if A is of finite type;*
- *$Ax = 0$ if and only if A is of affine type;*

- *$Ax < 0$ if and only if A is of indefinite type.*

The above corollary is incredibly useful in the classification of these matrices. To demonstrate its utility, we return to the construction from remark 1.1.1 to show that the matrix constructed from a Cartan matrix of finite type by adding a row and a column is indeed a GCM.

Proposition 1.3.12. *Let A be a Cartan matrix of finite type corresponding to a finite type Lie algebra g. Let A' be the matrix constructed from A as in remark 1.1.1. Then A' is a generalized Cartan matrix of affine type.*

Proof. Recall that $\theta = \sum_{i=1}^{n} a_i \alpha_i$ is the highest root of $\mathfrak{g}$, and $\mathfrak{g}$ has rank n. Then by construction of A', one can check that

$$A' \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = 0.$$

So A' is of affine type by corollary 1.3.11. $\square$

1.4 Symmetrizable GCMs

In the affine case, more can be said about the generalized Cartan matrices, and by extension, the affine Kac-Moody algebras. Affine and finite type GCMs are symmetrizable, meaning that they can be written as $A = DB$ for a nonsingular diagonal matrix D and symmetric matrix B. This allows a nice characterization of the finite and affine type GCMs (Theorem 1.4.9) that will lead to the classification of all finite and affine type GCMs via their Dynkin diagrams in the next section.

Definition 1.4.1. *Let $A = (a_{ij})$ for $i, j = 1, \ldots, n$ be any matrix. A* **principal submatrix** *of A is a matrix (a_{ij}) for $i, j \in I \subseteq \{1, \ldots, n\}$. The determinant of a principal sub matrix is called a* **principal minor***.*

Affine type GCMs are built from those of finite type. In fact, *any* principal submatrix of a GCM A of affine type is a direct sum of those of finite type, which more or less says that the ones we built in remark 1.1.1 are pretty much all there is as far as affine GCMs go. This isn't strictly true, but it is for all the affine Lie algebras on table 1.2!

Lemma 1.4.2. *Let A be an indecomposable GCM of finite or affine type. Then any proper principal submatrix of A is equivalent to a direct sum of matrices of finite type.*

Proof. Let $S \subseteq \{1, \ldots, n\}$ and let A_S be the principal submatrix obtained from A by taking only those rows and columns corresponding to elements of S.

Because A is of finite or affine type, there is a vector $u > 0$ such that $Au \geq 0$. For any vector x, let x_S is the vector built from x by eliminating rows i for $i \notin S$. Since A is a GCM, then a_{ij} for $i \neq j$ are nonpositive. Let $s \in S$, then

$$(Au)s = \sum_{k=1}^{n} a_{sk}u_k \leq \sum_{k \in S} a_{sk}u_k = (A_S u_S)_s,$$

so $(Au)_s \leq A_S u_S$. Because $Au \geq 0$,

$$A_S u_S \geq (Au)_S \geq 0,$$

with equality if and only if $a_{ij} = 0$ for all $i \in S$ and $j \notin S$. So the principal submatrix A_S is of finite or affine type by Corollary 1.3.11. If A_S is of affine type, then $A_S u_S \geq 0$ implies that $A_S u_S = 0$ and in particular $A_S u_S = (Au)_S$. But then $a_{ij} = 0$ for all $i \in S$ and $j \notin S$, which is impossible because A is indecomposable. $\square$

For symmetric matrices A, the question of finite or affine type is simple, as the next lemma shows.

Lemma 1.4.3. *An $n \times n$ symmetric matrix A is*

- *of finite type if and only if A is positive definite;*
- *of affine type if and only if A is positive semidefinite of rank $n - 1$.*

Proof. Let A be positive semidefinite. If there is $u > 0$ such that $Au < 0$, then $u^T Au < 0$, which contradicts the positive semidefinite condition on A. Therefore, if $u > 0$, $Au \geq 0$, so A is of finite or affine type by Corollary 1.3.11. If A has full rank, then A is of finite type. If A has rank $n - 1$, then A is of affine type.

Conversely, if A is a symmetric matrix of finite or affine type, then there is a vector $u > 0$ such that $Au \geq 0$. For a scalar $\lambda < 0$, $(A - \lambda I)u > 0$. So $A - \lambda I$ is of finite type by Corollary 1.3.11. Consequently, $\det(A - \lambda I) \neq 0$ for any $\lambda < 0$, so the eigenvalues of A must be nonnegative. Again, the finite and affine cases can be distinguished by the rank of the matrix. $\square$

Since answering the question of whether or not a matrix is of finite or affine type is so easy for the symmetric matrices, why don't we try to decompose all of our generalized Cartan matrices into symmetric ones and hope for a similar result? We can't in general do eigendecomposition on GCMs, but symmetrizability turns out to be the next best thing.

Definition 1.4.4. *A generalized Cartan matrix A is* **symmetrizable** *if there exists a non-singular diagonal matrix D and a symmetric matrix B such that $A = DB$.*

There are other characterizations of symmetrizable matrices that are sometimes easier to work with.

Lemma 1.4.5. *Let $A = (a_{ij})$ be an $n \times n$ GCM. A is symmetrizable if and only if for all sequences of indices $i_1, i_2, \ldots, i_k$ in $\{1, \ldots, n\}$, we have*

$$a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_k i_1} \;=\; a_{i_2 i_1} a_{i_3 i_2} \cdots a_{i_1 i_k}. \tag{1.7}$$

Proof. If A is symmetrizable, let $A = DB$ with D a diagonal matrix having $d_1, \ldots, d_n$ on the diagonal and $B = (b_{ij})$ an $n \times n$ symmetric matrix. Note that $a_{ij} = d_i b_{ij}$, so

$$\begin{aligned} a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_k i_1} &= d_{i_1} b_{i_1 i_2} d_{i_2} b_{i_2 i_3} \cdots d_{i_k} b_{i_k i_1} \\ &= d_{i_1} b_{i_2 i_1} d_{i_2} b_{i_2 i_3} \cdots d_{i_k} b_{i_1 i_k} \;=\; a_{i_2 i_1} a_{i_3 i_2} \cdots a_{i_1 i_k}, \end{aligned}$$

because B is symmetric.

Conversely, suppose that A satisfies the condition in (1.7) above, for all sequences of indices $i_1, i_2, \ldots, i_k$ in $\{1, \ldots, n\}$. We may further assume that A is indecomposable, because a decomposable matrix is symmetrizable if each of its indecomposable blocks are. Indecomposability implies that for each $i = 1, \ldots, n$, there is a sequence

$$1 = j_1, j_2, \ldots, j_t = i \tag{1.8}$$

such that $a_{j_s j_{s+1}} \neq 0$ for all $s = 1, \ldots, t-1$.

Now choose any nonzero $d_1 \in \mathbb{Q}$, and define

$$d_i \;=\; \frac{a_{j_t j_{t-1}} a_{j_{t-1} j_{t-2}} \cdots a_{j_2 j_1}}{a_{j_1 j_2} a_{j_2 j_3} \cdots a_{j_{t-1} j_t}} d_1.$$

The fact that this definition depends only on i and not on the sequence (1.8) follows from the condition (1.7). Moreover, because $a_{j_s j_{s+1}} \neq 0$ for all s, $d_i \neq 0$. Let D be the diagonal matrix with diagonal entries $d_1, \ldots, d_n$.

Define the matrix B by $b_{ij} = a_{ij}/d_i$. We show that B is indeed symmetric, or that $a_{ij}/d_i = a_{ji}/d_j$. If $a_{ij} = 0$, then by the definition of a GCM $a_{ji} = 0$ as well, so $b_{ij} = b_{ji}$. If $a_{ij} \neq 0$, then $1 = j_1, \ldots, j_t = i$ is a sequence from 1 to i that defines d_i. Furthermore, $1 = j_1, \ldots, j_t, j$ is a sequence from 1 to j that defines d_j, and so

$$d_j = \frac{a_{ji}}{a_{ij}} d_i,$$

so $b_{ij} = b_{ji}$. Therefore, B is symmetric, and by construction $A = DB$ for a diagonal matrix D and symmetric matrix B. □

Remark 1.4.6. In the above proof, we constructed D with nonzero diagonal elements $d_i \in \mathbb{Q}$. By clearing denominators and suitably adjusting B, we may assume that D is actually an integer matrix. Moreover, choosing $d_1 > 0$, we have that each d_i is also positive. Hence, $A = DB$ where D is a diagonal matrix of positive integers. Finally, because $a_{ij} \in \mathbb{Z}$ and $d_i \in \mathbb{Z}$, $b_{ij} \in \mathbb{Q}$ for all i, j.

Theorem 1.4.7. *Let A be an indecomposable GCM of finite or affine type. Then A is symmetrizable.*

Proof. First, consider a sequence of distinct integers $i_1, i_2, \ldots, i_k$ with $k \geq 3$ and $a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_{k-1} i_k} a_{i_k i_1} \neq 0$. If there is no such sequence, then all products $a_{j_1 j_2} a_{j_2 j_3} \cdots a_{j_t j_1}$ vanish, and the conditions of lemma 1.4.5 is satisfied, so A is symmetrizable.

Now, we may assume that such a sequence $i_1, \ldots, i_k$ exists, with

$$a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_{k-1} i_k} a_{i_k i_1} \neq 0.$$

Let $I = \{i_1, i_2, \ldots, i_k\}$. By reordering indices if necessary, we may assume that $i_j = j$ and so we have a principal submatrix A_I of A of the form

$$A_I = \begin{pmatrix} 2 & -r_1 & * & & * & -r_k \\ -s_1 & 2 & \ddots & & * & * \\ * & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ * & * & & \ddots & 2 & -r_{k-1} \\ -s_l & * & \cdots & * & -s_{k-1} & 2 \end{pmatrix},$$

where r_i, s_i are positive integers and the $*$ represent non-positive elements. Observing its structure, we see that A_I is indecomposable, so by lemma 1.4.2 we know that A_I is also of affine or finite type. Thus, there is some vector $u > 0$ such that $A_I u \geq 0$. Let U be the matrix with elements of u on the diagonal. Replacing A_I with $U^{-1} A_I U$, which preserves the structure of A_I, we may assume that u is the vector of all ones. The condition $A_I u \geq 0$ shows that

$$\sum_{j=1}^{k} a_{ij} u_j \geq 0,$$

and from this we conclude that $\sum_{i,j} a_{ij} \geq 0$. Because each of the $*$ elements in A_I is nonpositive, we may conclude in particular that

$$\sum_{i,j} a_{ij} = 2k - \sum_{i=1}^{k} (r_i + s_i) \geq 0. \tag{1.9}$$

Now $\frac{1}{2}(r_i + s_i) \geq \sqrt{r_i s_i} \geq 1$, so $r_i + s_i \geq 2$. But in order to maintain the inequality (1.9), it must be that $r_i + s_i = 2$ and $r_i s_i = 1$. However, r_i, s_i are positive integers, so we have $r_i = s_i = 1$ and all of the $*$ in A_I are zero.

$$A_I \;=\; \begin{pmatrix} 2 & -1 & 0 & & 0 & -1 \\ -1 & 2 & \ddots & & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & & \ddots & 2 & -1 \\ -1 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}, \tag{1.10}$$

Now for the vector $u = (1, 1, \ldots, 1)$, $u > 0$ yet $A_I u = 0$. Hence, by corollary 1.3.11, A_I has affine type. By lemma 1.4.2, we must have $A_I = A$, and A is symmetric and in particular, symmetrizable. □

Remark 1.4.8. There is another presentation of the Kac-Moody algebra $\mathfrak{g}(A)$ in terms of generators and relations, at least when A is a symmetrizable GCM. This actually only involves adding two more relations to the defining relations for $\tilde{\mathfrak{g}}(A)$ in (1.2). For completeness, we will state all of the relations here.

Let $A = (a_{ij})$ be an $n \times n$ symmetrizable GCM, with realization $(\mathfrak{h}, \Pi, \Pi^\vee)$ of The Lie algebra $\mathfrak{g}(A)$ is generated by $\mathfrak{h}$, $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$, subject to the relations

$$\begin{array}{rcll} [x, y] & = & 0 & \text{for all } x, y \in \mathfrak{h} \\ [e_i, f_j] & = & 0 & \text{if } i \neq j \\ [e_i, f_i] & = & \alpha_i^\vee & \text{for all } i \\ [x, e_i] & = & \alpha_i(x) e_i & \text{for all } x \in \mathfrak{h} \text{ and all } i \\ [x, f_i] & = & -\alpha_i(x) f_i & \text{for all } x \in \mathfrak{h} \text{ and all } i \\ (\text{ad}(e_i))^{1-a_{ij}}(e_j) & = & 0 & \text{for } i \neq j \\ (\text{ad}(f_i))^{1-a_{ij}} & = & 0 & \text{for } i \neq j \end{array}$$

By the theorem 1.4.7 above, this gives generators and relations for all Lie algebras of finite or affine type.

This is proved as a consequence Kac Character formula in [Car05, §19.4], and we will not present the proof here. But after reading this thesis and the proof of the Kac character formula in chapter 2, you will hopefully be able to follow the proof presented there.

Theorem 1.4.9. *Let A be an $n \times n$ indecomposable GCM. Then*

(i) *A has finite type if and only if all of its principal minors have positive determinant;*

(ii) *A has affine type if and only if* $\det A = 0$ *and all proper principal minors have positive determinant and* $\det(A) = 0$*;*

(iii) *if neither of the above conditions holds, then A is of indefinite type.*

Proof. Let's prove (i) first. Let A' be a principal submatrix of A. By lemma 1.4.2, A' is of finite type, and by theorem 1.4.7, A' is symmetrizable. So there is a diagonal matrix D with positive integer entries and symmetric matrix B such that $A' = DB$, and B has the same type as A'. In particular, B is of finite type, and so by lemma 1.4.3, B is positive definite. Hence, $\det(B) > 0$. Because all entries of the diagonal matrix D are positive, $\det(D) > 0$ as well, and thus $\det(A') = \det(D)\det(B) > 0$ as well.

Conversely, assume all principal minors of A are positive. Suppose for contradiction that there is a vector $u > 0$ with $Au < 0$, that is, A is of indefinite type. This gives a system of inequalities, one for for each i of the form

$$\sum_{j=1}^{n} a_{ij}u_j < 0.$$

Consider the equation for $i = 1$. By eliminating variables u_j for $j > 1$ from this equation (performing row reduction on the system), we end up with an equation of the form $\lambda u_1 < 0$ (note that these operations do not reverse the inequality because $a_{1,i} \leq 0$ for $i > 1$). Furthermore, we have that $\det(A) = \lambda \det(A')$, where A' is the submatrix consisting of rows and columns 2 through n. Because the principal minors of A are positive, we have that $\lambda > 0$, which makes the equation $\lambda u_1 < 0$ impossible because $u > 0$. Hence, A is of finite or affine type. Yet the principal minor A of A has $\det(A) > 0$, so A cannot be of affine type.

The proof of (ii) is nearly identical to the proof of (i). The statement (iii) follows from statements (i) and (ii). □

1.5 Dynkin diagrams

The purpose of this short section is to introduce Dynkin diagrams.

Definition 1.5.1. *Given an $n \times n$ generalized Cartan matrix $A = (a_{ij})$, the* **Dynkin Diagram** *$D(A)$ of A is a graph on n vertices, labelled $1, \ldots, n$. Each vertex corresponds to a column of A, and vertices i and j are connected by edges as follows:*

- *if $a_{ij}a_{ji} = 0$, there are no edges between i and j;*
- *if $a_{ij}a_{ji} \leq 4$ and $|a_{ij}| \geq |a_{ij}|$, then vertices i and j are connected by $|a_{ij}|$ edges, with an arrow pointing towards i if $|a_{ij}| > 1$;*
- *if $a_{ij}a_{ji} > 4$, then vertices i and j are connected by a single line labelled with the pair $|a_{ij}|, |a_{ji}|$.*

A Dynkin diagram is of **finite, affine**, *or* **indeterminate type** *if the corresponding GCM is of finite, affine, or indeterminate type, respectively.*

Example 1.5.2. *For example, the Generalized Cartan Matrix*

$$B_5^{(1)} = \begin{pmatrix} 2 & 0 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -2 & 2 \end{pmatrix}$$

corresponds to the affine Dynkin diagram

$$B_5^{(1)} \qquad \begin{array}{ccccccc} & & \circ & & & & \\ & & | & & & & \\ \circ & \text{---} & \circ & \text{---} & \circ & \Rightarrow & \circ \end{array}$$

1.6 Classification of affine Lie algebras

The main result of this section is a complete classification of all generalized Cartan matrices of finite and affine type based on their Dynkin diagrams. This classification extends the neat classification of finite dimensional semisimple Lie algebras to the Lie algebras of affine type. This is made precise in the theorem below. For comparison, a classification of finite type Dynkin diagrams (which is very similar to the the classification done here) is done in [Hum72] or [FH91].

Theorem 1.6.1. *The finite Dynkin diagrams are exactly those in table 1.1. The affine Dynkin diagrams are exactly those in tables 1.2, 1.3, and 1.4.*

There are several parts to this theorem, and we will prove each of them separately. First, by theorem 1.3.6, the finite type Dynkin diagrams are exactly those corresponding to the Cartan matrices. The classification of Cartan matrices is a standard part of the theory of finite dimensional complex Lie algebras, and can be found in [Car05] or [FH91], among other places, and we will not repeat it here.

In table 1.1, the numbers to the right of the diagrams are the determinants of the corresponding Cartan matrices. These will be used to show that each of the diagrams in tables 1.2, 1.3 and 1.4 actually do come from affine GCMs. To see this, first we prove a lemma.

Lemma 1.6.2. *Let A be an $n \times n$ indecomposable GCM. Then A is of affine type if and only if there exists a vector $\delta > 0$ such that $A\delta = 0$. Furthermore, such a vector δ is unique up to a constant scale factor.*

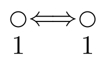
Proof. If A is of affine type, then there is a vector $u > 0$ such that $Au = 0$. Conversely, if there exists such a vector $\delta > 0$ such that $A\delta = 0$, then A is of affine type by 1.3.11.

Table 1.1: Dynkin Diagrams of Finite Type.

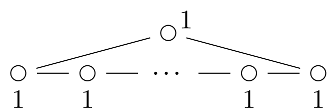
$A_n\ (n \geq 1)$	$\circ - \circ - \cdots - \circ - \circ$	$(n+1)$
$B_n(n \geq 2)$	$\circ - \circ - \cdots - \circ \Rightarrow \circ$	(2)
$C_n\ (n \geq 3)$	$\circ - \circ - \cdots - \circ \Leftarrow \circ$	(2)
$D_n(n \geq 4)$	$\begin{array}{ccccccccc} & & & & & & \circ & & \\ & & & & & & \vert & & \\ \circ & - & \circ & - & \cdots & - & \circ & - & \circ \end{array}$	(4)
E_6	$\begin{array}{ccccccccc} & & & & \circ & & & & \\ & & & & \vert & & & & \\ \circ & - & \circ & - & \circ & - & \circ & - & \circ \end{array}$	(3)
E_7	$\begin{array}{ccccccccccc} & & & & \circ & & & & & & \\ & & & & \vert & & & & & & \\ \circ & - & \circ & - & \circ & - & \circ & - & \circ & - & \circ \end{array}$	(2)
E_8	$\begin{array}{ccccccccccccc} & & & & & & & & \circ & & & & \\ & & & & & & & & \vert & & & & \\ \circ & - & \circ & - & \circ & - & \circ & - & \circ & - & \circ & - & \circ \end{array}$	(1)
F_4	$\circ - \circ \Rightarrow \circ - \circ$	(1)
G_2	$\circ \Rrightarrow \circ$	(1)

Table 1.2: Dynkin Diagrams of Affine Type (1).

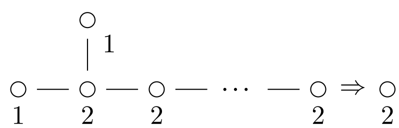
$A_1^{(1)}$



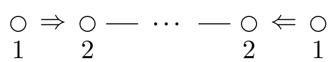
$A_n^{(1)}$ $(n \geq 2)$



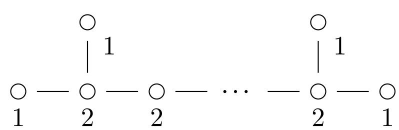
$B_n^{(1)}$ $(n \geq 3)$



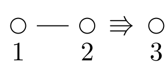
$C_n^{(1)}$ $(n \geq 2)$



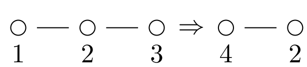
$D_n^{(1)}$ $(n \geq 4)$



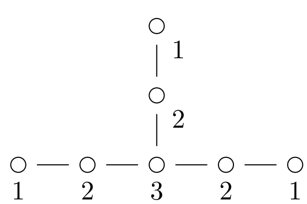
$G_2^{(1)}$



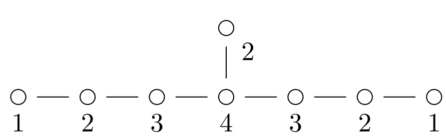
$F_4^{(1)}$



$E_6^{(1)}$



$E_7^{(1)}$



$E_8^{(1)}$

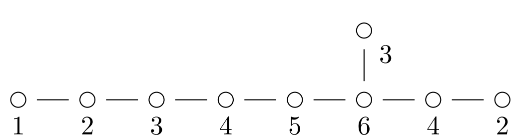


Table 1.3: Dynkin Diagrams of Affine Type (2).

$A_2^{(2)}$	$\overset{2}{\circ} \Lleftarrow \overset{1}{\circ}$
$A_{2n}^{(2)}$ $(n \geq 2)$	$\overset{2}{\circ} \Leftarrow \overset{2}{\circ} — \cdots — \overset{2}{\circ} — \overset{1}{\circ}$
$A_{2n-1}^{(2)}$ $(n \geq 3)$	$\overset{1}{\circ} — \overset{\overset{\circ^{1}}{\mid}}{\circ}{}^{2} — \overset{2}{\circ} — \cdots — \overset{2}{\circ} \Leftarrow \overset{1}{\circ}$
$D_{n+1}^{(2)}$ $(n \geq 2)$	$\overset{1}{\circ} \Leftarrow \overset{1}{\circ} — \cdots — \overset{1}{\circ} \Rightarrow \overset{1}{\circ}$
$E_6^{(2)}$	$\overset{1}{\circ} — \overset{2}{\circ} — \overset{3}{\circ} \Leftarrow \overset{2}{\circ} — \overset{1}{\circ}$

Table 1.4: Dynkin Diagrams of Affine Type (3).

$D_4^{(3)}$	$\overset{1}{\circ} — \overset{2}{\circ} \Lleftarrow \overset{1}{\circ}$

To show uniqueness, recall that because A is affine type, then $\operatorname{rank}(A) = n-1$, so the nullity of A is 1. Hence, if v is any other vector such that $Av = 0$, then v is a scalar multiple of u. Thus, the vector δ is unique up to scalar multiple. $\square$

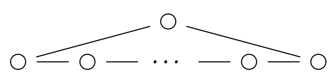
This lemma is key in showing that all of the Dynkin diagrams in tables 1.2, 1.3 and 1.4 are of affine type.

Lemma 1.6.3. *The Dynkin diagrams in tables 1.2, 1.3 and 1.4 are Dynkin diagrams of affine type.*

Proof. The nodes of the Dynkin diagrams in these tables are labelled with the coefficients of a linear dependence among the columns of the corresponding generalized Cartan matrices, that is, a vector δ such that $A\delta = 0$, where A is the GCM corresponding to the Dynkin diagram. This can easily be checked in each individual case. By lemma 1.6.2, each of these diagrams is of affine type. $\square$

Remark 1.6.4. From the above and lemma 1.4.2, we may conclude that all of the Dynkin diagrams appearing in table 1.1 are of finite type, as they each appear as a subdiagram of a diagram of affine type, and therefore correspond to a principal submatrix of a GCM of affine type.

Lemma 1.6.5. *If A is an indecomposable GCM of finite or affine type and if the corresponding Dynkin diagram $D(A)$ of contains a cycle of length at least three, then $D(A)$ is the cycle*



Note that this is the Dynkin diagram $A_\ell^{(1)}$ from table 1.2.

Proof. If $D(A)$ contains a cycle of length at least three, then there is a sequence $i_1, i_2, \ldots, i_k$ of indices such that $a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_k i_1} \neq 0$. So by the proof of theorem 1.4.7, the matrix A has the form as in equation (1.10). The associated Dynkin diagram is exactly the cycle pictured above. $\square$

Now, using the above lemmas, we show that if any Dynkin diagram of finite or affine type appears in either table 1.1 or one of the tables 1.2, 1.3 or 1.4, respectively. This is done by successively eliminating possibilities from the allowable connected Dynkin diagrams according to rules previously established in lemmas above. Those connected Diagrams remaining will necessarily be exactly those in the tables 1.1, 1.2, 1.3 and 1.4.

Lemma 1.6.6. *If A is an indecomposable GCM of finite type, then its Dynkin diagram $D(A)$ appears in table 1.1. If A is an indecomposable GCM of affine type, then its Dynkin diagram $D(A)$ appears in one of the tables 1.2, 1.3 or 1.4.*

Because this is such a long endeavor, we split the proof of this lemma into numerous easily digestible claims. And moreover, as stated before, the classification of Cartan matrices, which are exactly those GCMs of finite type by theorem 1.3.6, is a standard part of finite dimensional Lie theory, and will hence be assumed. For a full proof, see either [Car05] or [FH91]. So we know that the Dynkin diagrams of finite type are exactly those in table 1.1.

Proof of lemma 1.6.6. The proof proceeds by induction on the number of nodes n.

Claim 1.6.6.1. *There is only one possibility when $D(A)$ has a single vertex.*

Proof. If $D(A)$ has only one vertex, then $D(A)$ is A_1 and of finite type. $\square$

Claim 1.6.6.2. *If $D(A)$ has only two vertices, and A is of affine type, then $D(A)$ is $A_1^{(1)}$ or $A_2^{(2)}$.*

Proof. If $D(A)$ has two vertices, then

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

with a, b positive integers. For A to be of affine type, $\det(A) = 0$, so we must have that $ab = 4$. The possibilities are $\{a, b\} = \{1, 4\}$ or $a = b = 2$, corresponding to the Dynkin diagrams $A_1^{(1)}$ and $A_2^{(2)}$. $\square$

Now we begin the inductive step, which is the meat of this proof. Suppose that $D(A)$ has at least three vertices.

Claim 1.6.6.3. *We may assume that $D(A)$ contains no cycle.*

Proof. If $D(A)$ contains a cycle, then by lemma 1.6.5 $D(A)$ is the diagram A_ℓ for some $\ell > 2$. So we assume that $D(A)$ contains no cycle. $\square$

Claim 1.6.6.4. *The allowable edge types of $D(A)$ are the following:*

$$\circ - \circ \qquad\qquad \circ \Rightarrow \circ \qquad\qquad \circ \Rrightarrow \circ$$

Proof. By lemma 1.4.2, we know that every connected subdiagram of $D(A)$ with two vertices lies on the finite list. In particular, every connected subdiagram of $D(A)$ with two vertices is A_2, B_2 or G_2, so every edge of $D(A)$ looks like one of the above. $\square$

Claim 1.6.6.5. *If $D(A)$ has a triple edge and more than two vertices, then it must have exactly three vertices.*

Proof. Otherwise, it would have a proper connected subdiagram with a triple edge that is not G_2, which must be of finite type by lemma 1.4.2. But no such diagram appears on the affine list in table 1.1. Hence, we must have that

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -3 & 2 & -a \\ 0 & -b & 2 \end{pmatrix}$$

or the transpose of this matrix, where a, b are positive integers. From this, observe that

$$0 = \det A = 2(1 - ab).$$

The only solution to this equation in the positive integers is $a = b = 1$, so the diagram is $D_4^{(3)}$. □

So we may assume that $D(A)$ has no triple edge.

Claim 1.6.6.6. *We know that $D(A)$ has at most two double edges.*

Proof. Every proper connected subdiagram must appear on the finite list, and finite Dynkin diagrams have at most one double edge. □

Claim 1.6.6.7. *If $D(A)$ has two double edges, then $D(A)$ is $C_\ell^{(1)}$, $D_{\ell+1}^{(2)}$, or $A_{2\ell}^{(2)}$ for $\ell \geq 2$.*

Proof. Assume that $D(A)$ has two double edges; it also has no triple edge and no other type of edges. Each connected subdiagram is of finite type, so the second double edge must be on the end of a diagram of finite type that has a double edge already. The only such possible diagrams are $C_\ell^{(1)}$, $D_{\ell+1}^{(2)}$, or $A_{2\ell}^{(2)}$ for $\ell \geq 2$. □

Claim 1.6.6.8. *If $D(A)$ has only one double edge, then it is $B_\ell^{(1)}$ or $A_{2\ell-1}^{(2)}$ for $\ell \geq 3$.*

Proof. Assume that $D(A)$ has only one double edge. If $D(A)$ has a node of degree greater than two, it has a branch point. No proper subdiagram contains both the branch point and the double edge, because the subdiagram must be of finite type. Hence, $D(A)$ is either $B_\ell^{(1)}$ or $A_{2\ell-1}^{(2)}$ for $\ell \geq 3$. □

So far, we have eliminated all diagrams which have more than one double edge or a double edge and a branch point, so the only diagrams with a double edge not yet considered are those with a single double edge and no branches.

Claim 1.6.6.9. *If $D(A)$ has exactly one double edge and non branch points, then $D(A)$ is either $F_4^{(1)}$ of $E_6^{(2)}$*

Proof. In this case, $D(A)$ has the form

$$\circ - \cdots - \circ - \circ \Rightarrow \circ - \circ - \cdots - \circ$$

with l many nodes to the left of the double arrow, and r many nodes to the right of the double arrow, and $l + r + 2$ total vertices. Both $r \le 2$ and $l \le 2$, because otherwise $F_4^{(1)}$ and its transpose would appear as proper sub diagrams. Hence, we have that $\{r, l\} = \{1, 1\}, \{2, 1\}$, or $\{2, 2\}$. The case where $l = r = 1$ is the finite diagram F_4, and the case $r = l = 2$ would have $F_4^{(1)}$ as a subdiagram, and hence not be affine. Thus, $\{r, l\} = \{1, 2\}$ and the diagram is $F_4^{(1)}$ of $E_6^{(2)}$. □

Now we may assume that $D(A)$ has no edges other than single edges, and no cycles.

Claim 1.6.6.10. *$D(A)$ cannot have a node of degree greater than four. If it has a node of degree four, then $D(A)$ is $D_4^{(1)}$.*

Proof. If a node has degree greater than four, then there is a proper subdiagram of $D_4^{(1)}$, which is not finite type. So $D(A)$ cannot have any nodes with degree greater than four. If there is a node with degree four, then $D(A)$ is the diagram $D_4^{(1)}$, because otherwise it would again contain $D_4^{(1)}$ as a proper subdiagram. □

Therefore, we have reduced to the case where $D(A)$ has nodes of degree at most three.

Claim 1.6.6.11. *$D(A)$ has no more than two nodes of degree three.*

Proof. There cannot be more than two nodes of degree three, because otherwise there would be a proper connected subdiagram with two nodes of degree three, but such diagrams appear nowhere on the finite list in table 1.1. □

Claim 1.6.6.12. *If $D(A)$ has two nodes of degree three then $D(A)$ is $D_\ell^{(1)}$ for $\ell \ge 5$.*

Proof. It must be the case that every proper subdiagram of $D(A)$ had only one branch point in order for the proper subdiagram to be affine, so the only option is that $D(A)$ is $D_\ell^{(1)}$ for $\ell \ge 5$. □

Hence, we have reduced to the case where $D(A)$ has no multiple edges, no cycles, and at most one node of degree three.

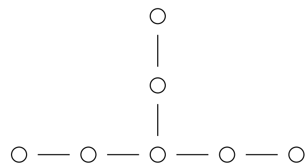
Claim 1.6.6.13. *If $D(A)$ has nodes with degree at most two, then $D(A)$ is finite type.*

Proof. If $D(A)$ has no node of degree three, then the diagram is A_ℓ on the finite list, and so not affine. □

Thus, we have a diagram $D(A)$ with no multiple edges, no cycles, and exactly one node of degree three.

Claim 1.6.6.14. *If $D(A)$ is a tree with one vertex of degree three, let the lengths of the three branches be a, b, c, respectively, with $a + b + c + 1$ total vertices. Then each of a, b, c must be at most two. If $a = b = c = 2$, then the diagram is $E_6^{(1)}$.*

Proof. Otherwise there would be a subdiagram of $E_6^{(1)}$, pictured here:



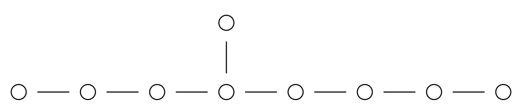
This diagram is not of finite type, so this eliminates the case in which $a = b = 2$ and $c > 2$. By observing the diagrams on table 1.2, we see that if $a = b = c = 2$, then the diagram is $E_6^{(1)}$. $\square$

Claim 1.6.6.15. *Keep the notation of the previous claim 1.6.6.14. If two of the branches are of length* 1*, then* $D(A)$ *is of finite type.*

Proof. Look at table 1.1. The diagram would be of finite type: either D_4, E_6, E_7 or E_8. $\square$

Claim 1.6.6.16. *Keep the notation of claim 1.6.6.14. If* $a = 1$ *but the diagram is of affine type, either* $b \leq 3$ *or* $c \leq 3$*.*

Proof. By the previous claim, we must have both $b > 1$ and $c > 1$. But in addition, $b \leq 3$ or $c \leq 3$, as otherwise there would be a proper subdiagram that is not of affine type as below:

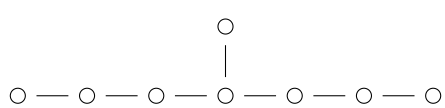


Hence, we have either $b \leq 3$ or $c \leq 3$. $\square$

Without loss of generality, suppose that $b \leq 3$. Consider first the case when $b = 2$.

Claim 1.6.6.17. *Keep the notation of 1.6.6.14. If* $a = 1$ *and* $b = 2$*, then we must have* $c = 3$*, and the diagram* $D(A)$ *is* $E_7^{(1)}$*.*

Proof. Otherwise there would be a proper subdiagram that looks like the following:



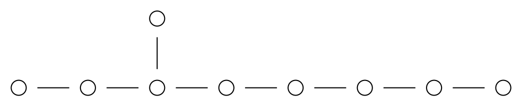
As before, this diagram is not of finite type . So if $a = 1$ and $b = 2$, then $c = 3$ and the diagram is $E_7^{(1)}$. $\square$

Claim 1.6.6.18. *Keep the notation of 1.6.6.14. Assume that* $a = 1$ *and* $b = 2$*. If* $D(A)$ *is finite type, then it is one of* E_6*,* E_7 *or* E_8*.*

Proof. Look at table 1.1. The diagrams with $c = 2, 3, 4$ are the finite diagrams E_6, E_7 and E_8. $\square$

Claim 1.6.6.19. *Keep the notation of 1.6.6.14. Assume that $a = 1$ and $b = 2$. If $D(A)$ is affine type, then $D(A)$ is $E_8^{(1)}$.*

Proof. So if this diagram is to be affine, then $c \geq 5$. Moreover, $c \leq 5$ because otherwise there would be a proper subdiagram that is not of affine type, as pictured below:



Therefore, this leaves only the diagram $E_8^{(1)}$. $\square$

This completes the proof of lemma 1.6.6. $\square$

Finally, we combine results to prove the central result of this section, which completes the classification of affine and finite type GCMs.

Proof of Theorem 1.6.1. The fact that each of the diagrams in table 1.1 are of finite type is shown in remark 1.6.4. The fact that each of the diagrams in tables 1.2, 1.3 and 1.4 is of affine type is shown in lemma 1.6.3. That these are all of the diagrams of finite and affine type is shown in lemma 1.6.6. $\square$

Based on theorem 1.6.1, we can also determine what the Dynkin diagrams of indefinite type are. Let A be an indecomposable GCM. Then A has indefinite type if and only if its Dynkin diagram appears on none of the tables 1.1, 1.2, 1.3, and 1.4.

1.7 Odds and ends: the root system, the standard invariant form and Weyl group

In this section, we will speed through some structural results of Kac-Moody algebras in order to get to the good stuff (by which I mean representation theory) in the next chapter. Mostly, I'm just going to write down some definitions and references. These will be the facts important for later proving the Kac Character formula.

Let A be an $n \times n$ symmetrizable matrix, and let

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

be the Cartan decomposition of $\mathfrak{g}(A)$, as in equation (1.5).

Definition 1.7.1. *The* **root system** *of* $\mathfrak{g}(A)$ *is*

$$\Phi = \{\alpha \in \mathbb{Q} \mid \alpha \neq 0, \mathfrak{g}_\alpha \neq 0\}.$$

The **positive roots** *are* $\Phi^+ = \Phi \cap Q^+$*, where* $Q^+ = \sum_{i=1}^n \mathbb{Z}_+ \alpha_i$ *as in definition 1.2.6. The* **negative roots** *are* $\Phi^- = \Phi \setminus \Phi^+$*.*

Next up is the generalization of the Killing form to Kac-Moody algebras. Here, the name is much less morbid – it's called the **standard invariant form**, which is both much more descriptive and much less fun.

Recall that because A is a symmetrizable matrix, we have $A = DB$ for D a diagonal matrix with diagonal elements $d_1, \dots, d_n$ and B a symmetric matrix $B = (b_{ij})_{1 \leq i,j \leq n}$. Let $(\mathfrak{h}, \Pi, \Pi^\vee)$ be a realization of A, with $\Pi = \{\alpha_1, \dots, \alpha_n\}$. Let $\mathfrak{h}'$ be the $\mathbb{C}$-span of $\alpha_i^\vee$ for $i = 1, \dots, n$, and let $\mathfrak{h}''$ be a complimentary subspace to $\mathfrak{h}'$ in $\mathfrak{h}$.

To define the standard invariant form, we first define a symmetric, bilinear form $\langle \cdot , \cdot \rangle$ on $\mathfrak{h}$ and then extend it to all of the Kac-Moody algebra $\mathfrak{g}(A)$. On $\mathfrak{h}$, this form is given by three equations

$$\begin{array}{rcll} \langle \alpha_i^\vee, \alpha_j^\vee \rangle & = & d_i d_j b_{ij} & \\ \langle \alpha_i^\vee, h \rangle & = & \alpha_i(h) d_i & \text{for } h \in \mathfrak{h} \\ \langle h_1'', h_2'' \rangle & = & 0 & \text{for } h_1'', h_2'' \in \mathfrak{h}''. \end{array}$$

Now, this form $\langle \cdot , \cdot \rangle$ is non-degenerate on $\mathfrak{h}$, as can be seen in[Car05, Proposition 16.1] or [Kac94, Lemma 2.1b]. So it gives us a bijection between $\mathfrak{h}^*$ and $\mathfrak{h}$, given by $\alpha \mapsto h'_\alpha$, where h'_α is the element of $\mathfrak{h}$ such that $\langle h'_\alpha, h \rangle = \alpha(h)$.

Via this bijection, we can extend the form on $\mathfrak{h}$ to one on $\mathfrak{h}^*$, which we will also notate by $\langle \cdot , \cdot \rangle$. This allows us to recover the generalized Cartan matrix $A = (a_{ij})_{1 \leq i,j \leq n}$ from the roots, as in [Kac94, page 20]:

$$a_{ij} = \frac{2\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle}$$

The form we constructed on $\mathfrak{h}$ extends uniquely to a form on all of $\mathfrak{g}(A)$; for details, see [Car05, Theorem 16.2] or [Kac94, Theorem 2.2].

Definition 1.7.2. *The* **standard invariant form** *on the Kac-Moody algebra* $\mathfrak{g}(A)$ *is the symmetric, bilinear,* $\mathbb{C}$*-valued form satisfying the following rules:*

- *(invariant)* $\langle x, [y, z] \rangle = \langle [x, y], z \rangle$*;*
- $\langle \mathfrak{g}(A)_\alpha, \mathfrak{g}(A)_\beta \rangle = 0$ *unless* $\alpha + \beta = 0$*;*
- *the pairing* $\mathfrak{g}_\alpha \times \mathfrak{g}_{-\alpha} \to \mathbb{C}$ *given by restricting* $\langle \cdot , \cdot \rangle$ *is nondegenerate for* $\alpha \neq 0$*;*
- $[x, y] = \langle x, y \rangle h'_\alpha$ *for* $x \in \mathfrak{g}_\alpha, y \in \mathfrak{g}_{-\alpha}, \alpha \in \Delta$*.*

As with the Killing form for finite-type Lie algebras, any nondegenerate, invariant, symmetric, bilinear form on $\mathfrak{g}(A)$ is a scalar multiple of the standard invariant form. This also means that the standard invariant form need not reduce directly to the Killing form, but only a multiple of it.

Now we'll talk about the **Weyl group** of a Kac-Moody algebra $\mathfrak{g}(A)$, with generalized Cartan matrix $A = (a_{ij})_{1 \leq i,j \leq n}$.

Definition 1.7.3. *The* **Weyl group** W *of a Kac-moody algebra is the Coxeter group generated by* s_i $i = 1, \ldots, n$ *and subject to the following relations:*

- $s_i^2 = 1$ *for all* i*;*
- $(s_i s_j)^2 = 1$ *if* $a_{ij} a_{ji} = 0$*;*
- $(s_i s_j)^3 = 1$ *if* $a_{ij} a_{ji} = 1$*;*
- $(s_i s_j)^4 = 1$ *if* $a_{ij} a_{ji} = 2$*;*
- $(s_i s_j)^6 = 1$ *if* $a_{ij} a_{ji} = 3$*.*

Remark 1.7.4. The Weyl group is easy to read off from a Dynkin diagram. You have one generator for each node, each of order 2. The other relations correspond to the number of edges: s_i and s_j commute if the corresponding nodes are disconnected, $(s_i s_j)$ has order 3 if there is a single edge between nodes i and j, order 4 if there is a double edge, order 6 if there is a triple edge, and infinite order otherwise.

For example, given the generalized Cartan matrix associated to affine $\mathfrak{sl}_2$ (type $A_1^{(1)}$),

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

the Weyl group is generated by s and t such that $s^2 = t^2 = 1$, with no other relations. This is the infinite Dihedral group.

Via the standard invariant form, we can define an action of the Weyl group W on the root system of $\mathfrak{g}(A)$. It suffices to define how the generators for W act on Φ, which is

$$s_i(\lambda) = \lambda - \lambda(\alpha_i^\vee)\alpha_i$$

for $\lambda \in \Phi$. This action is well-defined ([Car05, Proposition 16.15]) and faithful ([Car05, Theorem 16.16]).

Chapter 2

Representation Theory

The representation theory of Kac-Moody algebras is in many ways quite similar to the representation theory of Lie algebras of finite type, although there are some difficulties in accounting for the possibility of infinite dimensions. Nevertheless, one can make sense of **Verma modules**, **characters**, and even the **Casimir operator**. As in the finite dimensional case, we restrict attention mostly to the **category** $\mathcal{O}$ of those modules diagonalizable over the Cartan subalgebra. Within this framework, we are able to prove an analogue of the Weyl character formula, called the Kac character formula.

In this chapter, we begin in section 2.1 by recalling some basics of the representation theory of Lie algebras and define the **universal enveloping algebra**. We then introduce the category $\mathcal{O}$ of representations in section 2.2. Characters are introduced in section 2.3 and the appropriate generalization of the Casimir operator is given in 2.4. Finally, we combine these ingredients to prove the Kac Character formula in section 2.5.

2.1 The Universal Enveloping Algebra

Given any complex Lie algebra $\mathfrak{g}$, there is a unique associative algebra $U(\mathfrak{g})$ in such a way that the representation theory of $\mathfrak{g}$ is the same as the representation theory of $U(\mathfrak{g})$. This is the universal enveloping algebra of $\mathfrak{g}$. Because the representation theory of associative algebras is in many ways more familiar than the representation theory of Lie algebras, the universal enveloping algebra is commonly employed in the study of Lie algebra representations. As the name might suggest, it has a universal property among associative algebras that contain $\mathfrak{g}$.

First, a bit of notation. For any associative algebra A we can define a Lie algebra Lie(A) with bracket $[a, b] = ab - ba$ for all $a, b \in A$.

Definition 2.1.1. *The* **Universal Enveloping Algebra** *of a Lie algebra* $\mathfrak{g}$ *is an asso-*

ciative algebra $U(\mathfrak{g})$ together with a $\mathbb{C}$-linear map $i\colon \mathfrak{g} \to \mathrm{Lie}(U(\mathfrak{g}))$ such that for any associative algebra A and Lie algebra map $\theta\colon \mathfrak{g} \to \mathrm{Lie}(A)$, there is a unique associative algebra homomorphism $\phi\colon U(\mathfrak{g}) \to A$ such that $\phi \circ i = \theta$.

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "i"] \arrow[dr, "\theta"'] & \mathrm{Lie}(U(\mathfrak{g})) \arrow[d, dashed, "\phi"] \\ & \mathrm{Lie}(A) \end{tikzcd} \qquad \begin{tikzcd} U(\mathfrak{g}) \arrow[d, dashed, "\exists!\,\phi"] \\ A \end{tikzcd}$$

Remark 2.1.2. Although universal properties are quite useful, it is often helpful to have a concrete construction of these objects. These can be constructed as a quotient of the tensor algebra $T(\mathfrak{g})$ of $\mathfrak{g}$, and the universal property is then inherited from that of the tensor algebra.

This construction proceeds as follows. Given the direct sum of vector spaces

$$T(\mathfrak{g}) = \bigoplus_{n=0}^{\infty} \mathfrak{g}^{\otimes n}, \tag{2.1}$$

where $\mathfrak{g}^{\otimes 0} = \mathbb{C}$, we have a product structure given on simple tensors by

$$(x_1 \otimes \cdots \otimes x_i)(y_1 \otimes \cdots \otimes y_j) = x_1 \otimes \cdots \otimes x_i \otimes y_1 \otimes \cdots \otimes y_j,$$

which extends to a product on all of $T(\mathfrak{g})$ by linearity. The unit of $T(\mathfrak{g})$ is $1 \in g^{\otimes 0} = \mathbb{C}$. This is the tensor algebra of $\mathfrak{g}$.

Let J be the two-sided ideal of $T(\mathfrak{g})$ generated by all elements of the form $(x \otimes y - y \otimes x) - [x, y]$ for all $x, y \in \mathfrak{g}$. Then the universal enveloping algebra is the quotient

$$U(\mathfrak{g}) := T(\mathfrak{g})/J.$$

Furthermore, the map $i\colon \mathfrak{g} \to U(\mathfrak{g})$ is given by the composition of the maps $\mathfrak{g} \to T(\mathfrak{g}) \to U(\mathfrak{g})$, where the first is the inclusion into the direct sum (2.1) as $\mathfrak{g} = \mathfrak{g}^{\otimes 1}$, and the second is the natural quotient homomorphism.

From this construction, it seems to be the case that $U(\mathfrak{g})$ is always infinite dimensional as a $\mathbb{C}$-vector space. This is indeed the case, but the basis of $U(\mathfrak{g})$ nevertheless has a nice description in relation to the basis of $\mathfrak{g}$. This is the Poincaré-Birkhoff-Witt basis theorem (often abbreviated as the PBW theorem).

Theorem 2.1.3 (Poincaré-Birkhoff-Witt Basis Theorem)**.** *Let $\mathfrak{g}$ be a Lie algebra with basis $\{x_j : j \in I\}$, with a total order $<$ on the set I. Then if $i\colon \mathfrak{g} \to U(\mathfrak{g})$ is the natural linear map, and $y_j = i(x_j)$, $U(\mathfrak{g})$ has a basis all elements of the form*

$$y_{j_1}^{k_1} \cdots y_{j_n}^{k_n},$$

for all $n \geq 0$, $k_i \geq 0$, and $j_1, \ldots, j_n \in I$ with $j_1 < j_2 < \ldots < j_n$. Note that if $n = 0$, this is the empty product, or just the unit 1 of $T(\mathfrak{g})$.

Remark 2.1.4. This theorem is useful for understanding the structure of $U(\mathfrak{g})$, but it's proof doesn't really yield any new ideas in the representation theory, so we move on. A detailed proof can be found in [Car05, Theorem 9.4].

Amusingly, it is unclear if Poincaré actually successfully proved this theorem when he introduced it. Although some believe that the proof that Poincaré gave in the paper [Poi99] was incorrect, Ton-That and Tran claim in [TTT99] that this is a misunderstanding of his work.

This theorem also has many interesting corollaries:

Corollary 2.1.5 (Corollaries of the PBW Theorem)**.**

1. *The map* $i\colon \mathfrak{g} \to U(\mathfrak{g})$ *is injective.*
2. *Therefore,* i *identifies* $\mathfrak{g}$ *with a Lie subalgebra of* $\mathrm{Lie}(U(\mathfrak{g}))$*. In particular, every Lie algebra is a Lie subalgebra of an associative algebra.*
3. $U(\mathfrak{g})$ *has no zerodivisors.*

Remark 2.1.6. The PBW theorem has an important consequence that will be used freely. Namely, if $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$ is the triangular decomposition of $\mathfrak{g}$, then

$$U(\mathfrak{g}) \cong U(\mathfrak{n}^-) \otimes U(\mathfrak{h}) \otimes U(\mathfrak{n}^+).$$

These are direct sums and tensor products as vector spaces.

Finally, the following proposition describes the reason for studying the universal enveloping algebra: representation theory. This proposition is extremely useful in obtaining information about $\mathfrak{g}$-modules by studying $U(\mathfrak{g})$-modules. This development parallels the study of representations of a finite group G by instead considering modules over the group algebra $\mathbb{C}[G]$.

Proposition 2.1.7. *There is a bijection between representations of the Lie algebra* $\mathfrak{g}$ *and representations of* $U(\mathfrak{g})$*. In particular, if* V *is a complex vector space, then maps* $\phi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ *are in bijection with maps* $\psi\colon U(\mathfrak{g}) \to \mathrm{End}(V)$ *by the condition that* $\psi \circ i = \phi$*, where* $i\colon \mathfrak{g} \to U(\mathfrak{g})$ *is the inclusion map.*

We will use this proposition freely and often without explicit reference.

Proof. Let $\phi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation of $\mathfrak{g}$. By the universal property of $U(\mathfrak{g})$, there is a unique associative algebra homomorphism $\psi\colon U(\mathfrak{g}) \to \mathrm{End}(V)$ such that $\psi \circ i = \phi$ because $\mathfrak{gl}(V)$ is the Lie algebra formed from the associative algebra $\mathrm{End}(V)$.

Conversely, given a representation $\psi\colon U(\mathfrak{g}) \to \mathrm{End}(V)$, we define $\phi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ by $\phi = \psi \circ i$; we may consider this as a Lie algebra map.

It is clear that these two operations are inverse, so the representations of $\mathfrak{g}$ are in bijection with representations of $U(\mathfrak{g})$. ☐

Here is one more useful fact about universal enveloping algebras.

Proposition 2.1.8. *Let $\mathfrak{g}_1$, $\mathfrak{g}_2$ be two Lie algebras with maps $i_1\colon \mathfrak{g}_1 \to U(\mathfrak{g}_1)$ and $i_2\colon g_2 \to U(\mathfrak{g}_2)$. Then $U(\mathfrak{g}_1 \oplus \mathfrak{g}_2) \cong U(\mathfrak{g}_1) \otimes U(\mathfrak{g}_2)$. This is an isomorphism as associative algebras.*

Proof. Consider the homomorphism of Lie algebras $\mathfrak{g}_1 \oplus \mathfrak{g}_2 \to U(\mathfrak{g}_1) \otimes U(\mathfrak{g}_2)$ given by

$$x_1 + x_2 \mapsto i_1(x_1) \otimes 1 + 1 \otimes i_2(x_2).$$

This is a homomorphism because $x_1 \in \mathfrak{g}_1$ and $x_2 \in \mathfrak{g}_2$ commute in $\mathfrak{g}_1 \oplus \mathfrak{g}_2$. By the universal property of the enveloping algebra $U(\mathfrak{g}_1 \oplus \mathfrak{g}_2)$, this extends to an associative algebra morphism $\psi\colon U(\mathfrak{g}_1 \oplus \mathfrak{g}_2) \to U(\mathfrak{g}_1) \otimes U(\mathfrak{g}_2)$.

Furthermore, we have maps $\mathfrak{g}_1 \to U(\mathfrak{g}_1 \oplus \mathfrak{g}_2)$ and $\mathfrak{g}_2 \to U(\mathfrak{g}_1 \oplus \mathfrak{g}_2)$ given by $x \mapsto i_1(x) + 0$ and $y \mapsto 0 + i_2(y)$, respectively. These extend to associative algebra maps by the universal property, which we will call $\phi_1\colon U(\mathfrak{g}_1) \to U(\mathfrak{g}_1 \oplus \mathfrak{g}_2)$ and $\phi_2\colon U(\mathfrak{g}_2) \to U(\mathfrak{g}_1 \oplus \mathfrak{g}_2)$. Now define $\phi\colon U(\mathfrak{g}_1) \otimes U(\mathfrak{g}_2)$ by

$$\phi(x_1 \otimes x_2) = \phi_1(x_1)\phi_2(x_2).$$

To check that ϕ is well-defined, note that because $[x_1, x_2] = 0$ for all $x_1 \in \mathfrak{g}_1$ and $x_2 \in \mathfrak{g}_2$, we have that $\phi_1(x_1)\phi_2(x_2) = \phi_2(x_2)\phi_1(x_1)$. Because the images of element in $\mathfrak{g}$ generate $U(\mathfrak{g})$ as an algebra, for any Lie algebra $\mathfrak{g}$, this is exactly the condition needed for ϕ to be well-defined.

Now, I claim that ϕ and ψ are inverses. We need only check this on the generators, that is, elements of the form $x_1 + x_2 \in U(\mathfrak{g}_1 \oplus \mathfrak{g}_2)$ or $x_1 \otimes 1 + 1 \otimes x_2 \in U(\mathfrak{g}_1) \otimes U(\mathfrak{g}_2)$, for $x_1 \in \mathfrak{g}_1, x_2 \in \mathfrak{g}_2$. So compute:

$$(\phi \circ \psi)(x_1 + x_2) = \phi(x_1 \otimes 1) + \phi(1 \otimes x_2) = x_1 + x_2$$

$$(\psi \circ \phi)(x_1 \otimes 1 + 1 \otimes x_2) = \psi(x_1 + x_2) = x_1 \otimes 1 + 1 \otimes x_2$$

Hence, $U(\mathfrak{g}_1 \otimes \mathfrak{g}_2) \cong U(\mathfrak{g}_1) \oplus U(\mathfrak{g}_2)$. □

2.2 The Category $\mathcal{O}$

Let $\mathfrak{g}$ be a Kac-Moody algebra with triangular decomposition $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$. The category $\mathcal{O}$ is a particularly nice category of $\mathfrak{g}$-modules introduced in [BGG76] by Bernstein, Gelfand and Gelfand. This category is the full subcategory of all $\mathfrak{g}$-modules that satisfies some finiteness conditions, and is Abelian as a category, and closed under quotients, direct sums, and tensor products. Each finite-dimensional $\mathfrak{g}$-module is in $\mathcal{O}$, as well as an important class of $\mathfrak{g}$-modules known as the Verma modules.

Definition 2.2.1. *Given a Kac-Moody algebra $\mathfrak{g}$, the category $\mathcal{O}$ consists of those $\mathfrak{g}$-modules V such that:*

1. *V has a weight-space decomposition: $V = \bigoplus_{\lambda \in \mathfrak{h}^*} V_\lambda$, where V_λ is the eigenspace of the action of $\mathfrak{h}$ with eigenvalue λ: $V_\lambda = \{v \in V \mid x \cdot v = \lambda(x)v \text{ for all } x \in \mathfrak{h}\}$;*

2. $\dim V_\lambda$ *is finite for each* $\lambda \in \mathfrak{h}^*$;

3. *There is a finite set* $\{\lambda_1, \dots, \lambda_s\} \subset \mathfrak{h}^*$ *such that for each* λ *with* $V_\lambda \neq 0$, $\lambda \leq \lambda_i$ *for some* i, $1 \leq i \leq s$.

The morphisms in $\mathcal{O}$ *are homomorphisms of* $\mathfrak{g}$*-modules.*

In the above definition, the order on weights is the partial order defined as follows: $\mu \leq \lambda$ if and only if $\lambda - \mu$ is a sum of positive roots. The third condition shows that each modules in $\mathcal{O}$ has finitely many highest weights.

In order for the category $\mathcal{O}$ to be useful in the representation theory of $\mathfrak{g}$, we want to know that all of the usual tools of the theory of $\mathfrak{g}$-modules are available to us. The next lemma shows that $\mathcal{O}$ is closed under taking submodules, quotients, direct sums, and tensor products – all of the usual constructions on $\mathfrak{g}$-modules. The proof is given in section 19.1 of [Car05].

Lemma 2.2.2.

1. *If* $V \in \mathcal{O}$ *and* U *is a submodule of* V, *then* $U \in \mathcal{O}$ *and* $V/U \in \mathcal{O}$;

2. *If* $V_1, V_2 \in \mathcal{O}$, *then* $V_1 \oplus V_2 \in \mathcal{O}$ *and* $V_1 \otimes V_2 \in \mathcal{O}$.

Proof. Kac thinks this is obvious [Kac94, top of page 146]; I do not. In any case, it is proved satisfactorily in [Car05, Lemma 19.1]. □

For any $\lambda \in \mathfrak{h}^*$, we may define the Verma module $M(\lambda)$ which has unique highest weight λ. This is a class of modules important to the representation theory of $\mathfrak{g}$; their quotients by their unique maximal submodules are exactly the irreducible modules of $\mathcal{O}$. First we need to define the highest weight modules, of which the Verma modules are a special case.

Definition 2.2.3. *A* $\mathfrak{g}$*-module* V *is a* **highest-weight module** *with highest weight* $\Lambda \in \mathfrak{h}^*$ *if there is a nonzero vector* $v_\Lambda \in V$ *such that*

- $xv_\Lambda = 0$ *for all* $x \in \mathfrak{n}^+$;
- $hv_\Lambda = \Lambda(h)v_\Lambda$ *for all* $h \in \mathfrak{h}$;
- v_Λ *generates* V *as a* $U(\mathfrak{g})$*-module.*

The vector v_Λ is called the **highest-weight vector** for V. Furthermore, it follows from remark 2.1.6 and the first and third conditions above that v_Λ actually generates V as a $U(\mathfrak{n}^-)$-module.

It is not immediately clear that these highest weight modules are actually in the category $\mathcal{O}$, but this is indeed the case. A highest weight module V splits into weight spaces as

$$V = \bigoplus_{\lambda \leq \Lambda} V_\lambda,$$

with $V_\Lambda = \mathbb{C}v_\Lambda$ and $\dim V_\lambda$ finite for all λ (see [Kac94], section 9.2 for proof). In fact, this is the meaning of the term "highest-weight"– it refers to the fact that these modules necessarily have a single highest weight Λ, i.e. all weight spaces of V have weight $\lambda \leq \Lambda$.

The Verma modules are the highest weight modules that contain all of the others, in the following sense:

Definition 2.2.4. *A $\mathfrak{g}$-module $M(\Lambda)$ with highest weight Λ is a* **Verma module** *if every $\mathfrak{g}$-module with 1-dimensional highest Λ-weight space is a quotient of $M(\Lambda)$.*

From this definition, we may easily conclude that a Verma module with highest weight Λ is unique up to isomorphism; indeed, given two Verma modules M_1 and M_2, we have surjective homomorphisms $M_1 \to M_2$ and $M_2 \to M_1$, and these two maps will be mutually inverse.

Remark 2.2.5. So far we have not given an example of a highest weight module, so we will construct one now. In fact, we will have a Verma module for each element of $\mathfrak{h}^*$. Given any $\Lambda \in \mathfrak{h}^*$, let J be the left-ideal of $U(\mathfrak{g})$ generated by all $x \in \mathfrak{n}^+$ and all elements of the form $h - \Lambda(h)$ for $h \in \mathfrak{h}$, and set

$$M(\Lambda) \;=\; U(\mathfrak{g})/J,$$

which has the structure of a left $U(\mathfrak{g})$-module. The highest weight vector is given by the image of $1 \in U(\mathfrak{g})$, that is, $1 + J \in M(\Lambda)$.

I claim that this module is actually a Verma module. To see this, given any other $\mathfrak{g}$-module V with highest weight Λ, the annihilator $\mathrm{Ann}(V_\Lambda)$ of V_Λ is a left ideal of $U(\mathfrak{g})$ containing J, by the choice of generators for J. We have that $V \cong U(\mathfrak{g})/\mathrm{Ann}(V_\Lambda)$, and thus we get a surjection of $\mathfrak{g}$-modules $M(\Lambda) \to V$, that is, V is a quotient of $M(\Lambda)$.

Remark 2.2.6. Earlier, we noticed that any highest weight module is generated as a $U(\mathfrak{n}^-)$-module by its highest weight vector. In the case that this highest weight module is the Verma module $M(\Lambda)$, we can say quite a bit more: $M(\Lambda)$ is the free $U(\mathfrak{n}^-)$-module generated by its highest weight vector m_Λ. This follows from the Poincaré-Birkhoff-Witt theorem (theorem 2.1.3) and the construction of $M(\Lambda)$ given above. See [Kac94] section 9.2 for details.

Now, the Verma modules give us a tool for getting our hands on the irreducible objects of $\mathcal{O}$, which are central to the representation theory of any object. In fact, as stated before, each Verma module $M(\Lambda)$ has a unique maximal submodule $M'(\Lambda)$, and the quotients $L(\Lambda) = M(\Lambda)/M'(\Lambda)$ for $\Lambda \in \mathfrak{h}^*$ are precisely the irreducible objects of $\mathcal{O}$.

Proposition 2.2.7. *The Verma module $M(\Lambda)$ has a unique maximal submodule $M'(\Lambda)$, and the quotients $L(\Lambda) = M(\Lambda)/M'(\Lambda)$ are the only irreducible modules of $\mathcal{O}$.*

More elegantly stated, the irreducible objects of $\mathcal{O}$ are in bijection with the elements of $\mathfrak{h}^*$, via $\Lambda \mapsto L(\Lambda)$.

Proof. Let V be a proper $U(\mathfrak{g})$-submodule of $M(\Lambda)$. By Lemma 2.2.2, we have that

$$V = \bigoplus_{\mu} V_\mu$$

I claim that the weight space $V_\Lambda = 0$. If $V_\Lambda \neq 0$, then because $\dim M(\Lambda)_\Lambda = 1$ and $V_\Lambda \subseteq M(\Lambda)_\Lambda$, we have $V_\Lambda = M(\Lambda)_\Lambda = \mathbb{C}v_\Lambda$, where v_Λ is the highest weight vector. But $M(\Lambda)$ is generated by the highest weight vector v_Λ as a $U(\mathfrak{g})$-module, and if $v_\Lambda \in V_\Lambda \subseteq V$, then $V = M(\Lambda)$. Yet we assumed V was a proper submodule, so this cannot be the case. Hence, $V_\Lambda = 0$.

From this, we may conclude that every proper submodule of $M(\Lambda)$ lies inside $\sum_{\mu;\, \mu \neq \Lambda} M(\Lambda)_\mu$. Take $M'(\Lambda)$ to be the sum of all proper submodules of $M(\Lambda)$, which must lie in the above sum, and is therefore a proper submodule. Furthermore, $M'(\Lambda)$ is the unique maximal submodule because it contains all proper submodules of $M(\Lambda)$ by construction. $L(\Lambda) = M(\Lambda)/M'(\Lambda)$ has no submodules, and so is irreducible.

To show that these quotients are the only irreducible modules of $\mathcal{O}$, let $V \in \mathcal{O}$ be an irreducible $\mathfrak{g}$-module. By the definition of $\mathcal{O}$, we know that V has a maximal weight, say μ. For a vector $w_\mu \in V$ with weight μ, we have that $xv_\mu = 0$ for $x \in \mathfrak{n}^+$ and $xw_\mu = \mu(x)w_\mu$ for $x \in \mathfrak{h}$.

Define a map ϕ from the Verma module $M(\mu)$ to V as follows. Because $M(\mu)$ is generated as a $U(\mathfrak{n}^-)$-module by a highest-weight vector v_μ, each element of $M(\mu)$ is of the form uv_μ for $u \in U(\mathfrak{n}^-)$. So define $\phi(uv_\mu) = uw_\mu$. It can be checked that this is a homomorphism.

The image of ϕ is a nonempty submodule of V (containing v_μ), so $\operatorname{im}\phi = V$ because V is irreducible. Hence, by the first isomorphism theorem,

$$M(\mu)/\ker\phi \cong V.$$

But because V is irreducible, $M(\mu)/\ker\phi$ must also be irreducible, which is only the case when $\ker\phi$ is a maximal submodule. There is only one such submodule, which is $M'(\mu)$. Hence, V is isomorphic to $M(\mu)/M'(\mu) = L(\mu)$. □

Finally, we have a variant of Schur's lemma, which says that the endomorphisms of an irreducible object are merely multiplication by a scalar. The proof is short.

Lemma 2.2.8. $\operatorname{End}_{U(\mathfrak{g})} L(\Lambda) \cong \mathbb{C}\mathrm{id}_{L(\Lambda)}$

Proof. If α is an endomorphism of $L(\Lambda)$ and v_Λ is a highest-weight vector, then α sends a highest weight vector v_Λ to another highest weight vector. But $\dim L(\Lambda)_\Lambda = 1$ because $L(\Lambda) = M(\Lambda)/M'(\Lambda)$ and $M'(\Lambda)$ avoids $M(\Lambda)_\Lambda$. Thus, $\alpha(v_\Lambda) = av_\Lambda$ for some $a \in \mathbb{C}$. However, $L(\Lambda)$ is generated by v_Λ as a $U(\mathfrak{g})$-module, so any element is of the form uv_Λ for some $u \in U(\mathfrak{g})$. So $\alpha(uv_\Lambda) = u\alpha(v_\Lambda) = auv_\Lambda$. Hence, $\alpha = a\mathrm{id}_{L(\Lambda)}$. □

2.3 Characters of $\mathfrak{g}$-modules

As is usually done, we consider here formal characters of $\mathfrak{g}$-modules in $\mathcal{O}$ rather than the characters arising from representations a Lie group with Lie algebra $\mathfrak{g}$. This does not, however, make the characters we consider any less useful in revealing the structure of representations. In fact, if V is a $\mathfrak{g}$-module that admits a character $\operatorname{ch} V$, then the structure of V as an $\mathfrak{h}$-module is completely determined by $\operatorname{ch} V$.

In working with the characters, we will use the notion of formal exponentials e^λ, with which we create a $\mathbb{C}$-algebra which contains the characters and is analogous to the ring of class functions on a finite group. In fact, there is an alternative description of the characters, and this ring, as functions $\mathfrak{h}^* \to \mathbb{Z}$. We will in this section explicitly determine the character for any Verma module $M(\Lambda)$ in terms of these formal exponentials, and set up the basic theory of characters necessary to approach the Kac character formula, which determines the character of the irreducible representation $L(\Lambda)$ for Λ a dominant integral weight.

Definition 2.3.1. *The* **formal exponentials** *e^λ are symbols in bijection with $\mathfrak{h}^*$ that obey the rule $e^\lambda e^\mu = e^{\lambda+\mu}$. The $\mathbb{C}$-algebra $\mathcal{E}$ is the algebra of series of the form*

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

for complex numbers c_λ that vanish outside the union of a finite number of sets of the form $S(\mu) = \{\nu \in \mathfrak{h}^* \mid \nu \leq \mu\}$. The identity of $\mathcal{E}$ is e^0. The sum and product of two such series are defined in the usual manner.

Definition 2.3.2. *The* **character** *of a $\mathfrak{g}$-module V in $\mathcal{O}$ is*

$$\operatorname{ch} V = \sum_{\lambda \in \mathfrak{h}^*} \dim V_\lambda e^\lambda.$$

Note that by the third condition of Definition 2.2.1, the support of $\operatorname{ch} V$ is limited to a finite union of the sets $S(\mu)$; there are only finitely many maximal weights of V. So $\operatorname{ch} V$ is indeed an element of $\mathcal{E}$.

Remark 2.3.3. There is another interpretation of the algebra $\mathcal{E}$ as the functions $\mathfrak{h}^* \to \mathbb{Z}$, with e^λ the characteristic function $e^\lambda(\lambda) = 1$, $e^\lambda(\mu) = 0$ for $\lambda \neq \mu$. Here, multiplication $e^\lambda e^\mu$ is convolution determined by

$$(e^\mu e^\nu)(\lambda) \;=\; \sum_{\substack{\alpha,\beta \in \mathfrak{h}^* \\ \alpha+\beta=\lambda}} e^\mu(\alpha) e^\mu(\beta)$$

This makes the connection to the class functions on a finite group a bit more transparent. In this interpretation, the character of a $\mathfrak{g}$-module V in $\mathcal{O}$ is the function $\mathfrak{h}^* \to \mathbb{Z}$ defined by

$$(\operatorname{ch} V)(\lambda) = \dim V_\lambda.$$

Many of the typical facts about characters which are familiar from the representation theory of finite groups carry over into this setting. Namely, we have the following proposition for modules in $\mathcal{O}$. This follows easily from lemma 2.2.2 and dimension considerations.

Proposition 2.3.4.

1. *Let V be a $\mathfrak{g}$-module and U a submodule of V. Then* $\operatorname{ch} U + \operatorname{ch}(V/U) = \operatorname{ch} V$.
2. *Suppose V, W are both $\mathfrak{g}$-modules. Then* $\operatorname{ch}(V_1 \oplus V_2) = \operatorname{ch} V_1 + \operatorname{ch} V_2$ *and* $\operatorname{ch}(V_1 \otimes V_2) = \operatorname{ch} V_1 \otimes \operatorname{ch} V_2$.

This next proposition gives an explicit description for the character of the prototypical modules in $\mathcal{O}$, the Verma modules.

Proposition 2.3.5. *Let $M(\lambda)$ be a Verma module for the Kac-Moody algebra $\mathfrak{g}$. Then*

$$\operatorname{ch} M(\Lambda) = \frac{e^{\Lambda}}{\prod_{\alpha \in \Phi^+} (1 - e^{-\alpha})^{\operatorname{mult} \alpha}}.$$

Proof. As in remark 2.2.6, we know that $M(\Lambda)$ is the free $U(\mathfrak{n}^-)$-module generated by its highest weight vector v_Λ. In particular, the map $u \mapsto uv_\Lambda$ is a bijection between $U(\mathfrak{n}^-)$ and $M(\Lambda)$ that sends the weight space $U(\mathfrak{n}^-)_{-\mu}$ to $M(\Lambda)_{\Lambda-\mu}$. We will use this to construct a basis for each weight space of $M(\Lambda)$ and rewrite the character in this basis.

For each positive root $\alpha \in \Phi^+$, the **multiplicity** of α is the dimension of the weight space $\mathfrak{g}_{-\alpha} = (\mathfrak{n}^-)_{-\alpha}$, denoted as $\operatorname{mult} \alpha = \dim(\mathfrak{n}^-)_{-\alpha}$. Choose for each α a basis $e_{-\alpha,i}$ of $\mathfrak{g}_{-\alpha}$ for $1 \leq i \leq \operatorname{mult} \alpha$. By the PBW theorem (theorem 2.1.3), we obtain a basis of $U(\mathfrak{n}^-)$ consisting of all products

$$\prod_{\alpha \in \Phi^+} \prod_{i=1}^{\operatorname{mult} \alpha} e_{-\alpha,i}^{k_{\alpha,i}} \tag{2.2}$$

for nonnegative integers $k_{\alpha,i}$ and an appropriate choice of ordering on the $e_{-\alpha,i}$.

This tells us that the weight space $U(\mathfrak{n}^-)_{-\mu}$ has a basis consisting of elements of the form (2.2) such that

$$\sum_{\alpha \in \Phi^+} \left(\sum_{i=1}^{\operatorname{mult} \alpha} k_{\alpha,i} \right) \alpha = \mu \tag{2.3}$$

because for any $x \in U(\mathfrak{g})$ and basis element $v = \prod_{\alpha \in \Phi^+} \prod_{i=1}^{\operatorname{mult} \alpha} e_{-\alpha,i}^{k_{\alpha,i}} \in U(\mathfrak{n}^-)$, we have

$$x \cdot v = \left(\sum_{\alpha \in \Phi^+} \left(\sum_{i=1}^{\operatorname{mult} \alpha} k_{\alpha,i} \right) \alpha(x) \right) v.$$

That is, v has weight μ, and is in the subspace $U(\mathfrak{n}^-)_\mu$, if and only if (2.3) holds.

We may rewrite the formal character of $U(\mathfrak{n}^-)$ as follows

$$\operatorname{ch} U(\mathfrak{n}^-) \;=\; \sum_{\alpha\in\Phi^+} \dim(U(\mathfrak{n}^-)_{-\alpha})e^{-\alpha} \;=\; \prod_{\alpha\in\Phi^+} \left(1+e^{-\alpha}+(e^{-\alpha})^2+\dots\right)^{\operatorname{mult}\alpha}, \tag{2.4}$$

the right equality holding because the number of times $e^{-\mu}$ appears on the right-hand side is exactly the number of sets of nonnegative integers $k_{\alpha,i}$ such that (2.3) holds.

Because $U(\mathfrak{n}^-)_{-\mu}$ is in bijection with $M(\Lambda)_{\Lambda-\mu}$, these weight spaces have the same dimension and we may relate the characters by $\operatorname{ch} U(\mathfrak{n}^-) = e^{\Lambda}\operatorname{ch} M(\Lambda)$. It follows from (2.4) that

$$\operatorname{ch} M(\Lambda) = e^{\Lambda} \prod_{\alpha\in\Phi^+} \left(1+e^{-\alpha}+(e^{-\alpha})^2+\dots\right)^{\operatorname{mult}\alpha}.$$

But the element $1+e^{-\alpha}+(e^{-\alpha})^2+\dots$ of the algebra $\mathcal{E}$ is invertible, with inverse $1-e^{-\alpha}$; this is a simple computation done in [Car05, Lemma 12.3]. So as a formal series, we write

$$\operatorname{ch} M(\Lambda) = \frac{e_{\Lambda}}{\prod_{\alpha\in\Phi^+}(1-e^{-\alpha})^{\operatorname{mult}\alpha}}. \qquad \square$$

This formula is a precursor to the Kac character formula, which gives an explicit method of computing the character of the irreducible module $L(\Lambda)$ in the case that Λ is what is called a dominant integral weight. In the finite dimensional case, the analogue of the Kac character formula, called the Weyl character formula, is proven using multiplicities $[V\colon L(\Lambda)]$ for $V\in\mathcal{O}$.

These multiplicities are usually found using composition series of modules $V\in\mathcal{O}$: $[V\colon L(\Lambda)]$ should be the number of composition factors isomorphic to $L(\Lambda)$, which would be independent of the choice of series by the Jordan-Hölder Theorem. But in the infinite dimensional case, the definition of these multiplicities is not so clear. One obstruction to this is, for example, the fact that the Verma module $M(0)$ has no irreducible submodule when $\mathfrak{g}$ is infinite dimensional. We can nevertheless find a suitable definition of $[V\colon L(\Lambda)]$ in the infinite dimensional case, following [Kac94].

For an approach which avoids the use of composition series, see [Run12]

Lemma 2.3.6. *Let $V\in\mathcal{O}$ and let $\lambda\in\mathfrak{h}^*$. Then V has a filtration*

$$0 = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_t = V$$

of finite length by a sequence of submodules and a subset $J\subseteq\{1,\ldots,t\}$ such that

- *if $j\in J$, then $V_j/V_{j-1}\cong L(\lambda_j)$ for some $\lambda_j\geq\lambda$;*
- *if $j\notin J$, then $(V_j/V_{j-1})_\mu = 0$ for all $\mu\geq\lambda$.*

Proof. Because V is in $\mathcal{O}$, it has only finitely many weights μ with $\mu \geq \lambda$. Let $a(V, \lambda)$ be the sum of the dimensions of weight spaces V_μ for $\mu \geq \lambda$, that is $a(V, \lambda) = \sum_{\mu \geq \lambda} \dim V_\mu$. The proof proceeds by induction on $a(V, \lambda)$.

In the case that $a(V, \lambda) = 0$, then the filtration is $V = V_1 \supseteq V_0 = 0$, as the quotient $V_1/V_0 = V$ has the property that every weight space V_μ for $\mu \geq \lambda$ is trivial.

So suppose that $a(V, \lambda) > 0$. Then there is some weight μ with $\mu \geq \lambda$; suppose that μ is maximal among such weights. Choose a vector $v \in V$ with weight μ, so $xv = 0$ for $x \in \mathfrak{n}^+$ (because μ is maximal) and $xv = \mu(x)v$ for $x \in \mathfrak{h}$. Let $V' = U(\mathfrak{g})v$ be the submodule of V generated by v. By the definition of the Verma module $M(\mu)$, V' is a quotient of $M(\mu)$, because V' has highest weight μ. So we have a surjective homomorphism of $U(\mathfrak{g})$-modules $\phi\colon M(\mu) \to V'$ defined by $\phi(um_\mu) = uv$ for $u \in U(\mathfrak{n}^-)$, where m_μ is the highest-weight vector of $M(\mu)$. Via ϕ, V' is isomorphic to a quotient of $M(\mu)$, and so has a unique maximal submodule V'' and

$$V'/V'' \cong M(\mu)/M'(\mu) \cong L(\mu).$$

Now we can create a filtration: $V \supseteq V' \supseteq V'' \supseteq 0$. We note that $a(V'', \lambda) < a(V, \lambda)$ and $a(V/V', \lambda) < a(V, \lambda)$ because $\mu \geq \lambda$ is a weight in $V'/V'' \cong L(\mu)$. So by induction there exist filtrations of the required kind for $V'' \in \mathcal{O}$ and $V/V' \in \mathcal{O}$, and we can put these filtrations together to get such a filtration for V. $\square$

Now, we want to use these filtrations to define the multiplicity $[V : L(\Lambda)]$ as the number of times $L(\Lambda)$ appears as a composition factor of such a filtration, which is equal to the number of times Λ appears in the set $\{\lambda_j : j \in J\}$. But it is not *a priori* clear that this number is independent of the choice of filtration or the choice of λ (although Kac claims that it is in [Kac94, middle of page 151]). This is indeed the case, but it takes some work to establish.

Lemma 2.3.7. *Let $V \in \mathcal{O}$ and $\lambda \in \mathfrak{h}^*$. Consider filtrations of V as in lemma 2.3.6 with respect to λ. Suppose $\mu \in \mathfrak{h}^*$ satisfies $\mu \geq \lambda$. Then the number of factors isomorphic to $L(\mu)$ in such a filtration is independent of both the filtration and the choice of λ.*

Proof. Immediately, we notice that a filtration with respect to λ is also a filtration with respect to μ when $\mu \geq \lambda$, and whether regarded as a filtration of μ or λ, the multiplicity of $L(\mu)$ is unchanged. Thus, to prove the lemma we take two filtrations with respect to μ and show that $L(\mu)$ has the same multiplicity in each. We do that by a modification of the proof of the Jordan-Hölder theorem.

Let

$$V \;=\; V_0 \subseteq V_1 \subseteq \ldots \subseteq V_t \;=\; V \tag{2.5}$$

$$0 = W_0 \subseteq W_1 \subseteq \ldots \subseteq W_s = V \tag{2.6}$$

be two such filtrations, as in Lemma 2.3.6. The proof proceeds by induction on $m = \min\{t, s\}$.

If $m = \min\{t, s\} = 1$, then either the two filtrations are identical, or μ is not a weight of V. In either case, the multiplicity of $L(\mu)$ in V is the same.

So suppose that $m > 1$. This is the hard part. We divide this into cases and handle each separately. We prove that the conclusions of 2.3.7 hold:

- when $V_{t-1} = W_{s-1}$ in claim 2.3.7.1;
- when $V_{t-1} \subsetneq W_{s-1}$ in claim 2.3.7.2;
- if neither of V_{t-1}, W_{s-1} contain the other in claim 2.3.7.3.

Claim 2.3.7.1. *If $V_{t-1} = W_{s-1}$, then the number composition factors isomorphic to $L(\mu)$ is the same in* (2.5) *and* (2.6).

Proof. Let $U = V_{t-1} = W_{s-1}$. The two filtrations

$$0 = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_{t-1}$$

$$0 = W_0 \subseteq W_1 \subseteq \ldots \subseteq W_{s-1}$$

of U are still filtrations of the type in Lemma 2.3.6, and by induction contain the same number of factors of $L(\mu)$. The multiplicity of $L(\mu)$ in the original filtrations (2.5) and (2.6) are possibly one greater, depending on the factor $V/V_{t-1} = V/W_{s-1}$, which is the same for both. Hence, the multiplicities of $L(\mu)$ in these two filtrations is the same. □

Claim 2.3.7.2. *If $V_{t-1} \subsetneq W_{s-1}$, then the number composition factors isomorphic to $L(\mu)$ is the same in* (2.5) *and* (2.6).

Proof. We have that V/V_{t-1} is not irreducible, having V/W_{s-1} as a submodule. Thus, μ is not a weight of V/V_{t-1}, and so not a weight of V/W_{s-1}. Neither V/V_{t-1} nor V/W_{s-1} can possibly be isomorphic to $L(\mu)$ without μ as a weight. Now let

$$0 = U_0 \subseteq U_1 \subseteq \ldots \subseteq U_m \subseteq V_{t-1}$$

be any filtration of V_{t-1} with respect to μ satisfying the conditions outlined in Lemma 2.3.6. From this, we construct two more filtrations:

$$0 = U_0 \subseteq \ldots \subseteq U_m \subseteq V_{t-1} \subseteq V \tag{2.7}$$

$$0 = U_0 \subseteq \ldots \subseteq U_m \subseteq V_{t-1} \subseteq W_{s-1} \subseteq V. \tag{2.8}$$

Both of these are filtrations with respect to μ. By applying the induction hypothesis, we can see that $L(\mu)$ has the same multiplicity in (2.7) as it does in (2.5), and the same multiplicity in (2.8) as in (2.6), given that they have the same leading term. And the argument above showed that none of V/V_{t-1}, V/W_{s-1} or W_{s-1}/V_{t-1} are isomorphic to $L(\mu)$, so the multiplicity of $L(\mu)$ in (2.7) and (2.8) must be the same. Hence, we conclude that the multiplicity of $L(\mu)$ in (2.5) and (2.6) are the same. □

Claim 2.3.7.3. *If neither of V_{t-1}, W_{s-1} contain the other, then the number composition factors isomorphic to $L(\mu)$ is the same in* (2.5) *and* (2.6)*.*

Proof. Let $U = V_{t-1} \cap W_{s-1}$ and choose a filtration

$$0 = U_0 \subseteq U_1 \subseteq \ldots \subseteq U_m = U$$

of U of the required kind with respect to μ. From this, again construct two more filtrations:

$$0 = U_0 \subseteq U_1 \subseteq \ldots U_m \subseteq V_{t-1} \subseteq V \tag{2.9}$$

$$0 = U_0 \subseteq U_1 \subseteq \ldots U_m \subseteq W_{s-1} \subseteq V. \tag{2.10}$$

To know that these are filtrations of the right type, observe that by the second isomorphism theorem,

$$V_{t-1}/U \cong (V_{t-1} + W_{s-1})/W_{s-1} \quad \text{and} \quad W_{s-1}/U \cong (V_{t-1} + W_{s-1})/V_{t-1}. \tag{2.11}$$

That these composition factors have the right property follows.

Applying the induction hypothesis, we see that because filtrations (2.9) and (2.5) share leading terms, $L(\mu)$ has the same multiplicity in each. Similarly, the multiplicity of $L(\mu)$ in (2.10) is the same as that in (2.6). So it suffices to show that the multiplicity of $L(\mu)$ is the same among (2.9) and (2.10). Since these filtrations differ only by the first two factors, we check that the number of times $L(\mu)$ appears among the first two factors is the same.

If $V = V_{t-1} + W_{s-1}$, then by (2.11) we have that $V/V_{t-1} \cong W_{s-1}/U$ and $V/W_{s-1} \cong V_{t-1}/U$, so the multiplicities are the same.

If $V_{t-1} + W_{s-1} \neq V$, then V/V_{t-1} and V/W_{s-1} are not irreducible and therefore do not have μ as a weight. In this case, μ is also not a weight of any of V/V_{t-1}, V_{t-1}/U, V/W_{s-1} , or W_{s-1}/U. Hence, none of these modules can be isomorphic to $L(\mu)$. □

This completes the proof of lemma 2.3.7 . □

Finally, given the above lemma, we make formal the definition of the multiplicity $[V : L(\Lambda)]$.

Definition 2.3.8. *The* **multiplicity** *of $L(\Lambda)$ in a filtration of $V \in \mathcal{O}$ of the type considered in lemma 2.3.6 is the number of times $L(\Lambda)$ appears as a composition factor, and is denoted $[V : L(\Lambda)]$.*

Note that $[V, L(\mu)]$ is zero when μ is not a weight of V.

The reason that this multiplicity is useful to us is that it allows us to rewrite the character of any $V \in \mathcal{O}$ in terms of the the characters of $L(\lambda)$ for weights λ of V. The following is an incredibly useful lemma in the theory of characters of modules in $\mathcal{O}$, which also gives some motivation for understanding the characters of irreducible $L(\Lambda)$ via the Kac character formula. Once we know the characters of $L(\Lambda)$ via the Kac character formula, we can write the characters of any $V \in \mathcal{O}$ in terms of the characters of the irreducibles. And moreover, the first step of the proof of the Kac character formula will be to rewrite the character of a Verma module $M(\Lambda)$ in terms of the irreducible modules $L(\mu)$ for $\mu \leq \lambda$.

Lemma 2.3.9. *Let $V \in \mathcal{O}$. Then,*

$$\operatorname{ch} V \;=\; \sum_{\lambda \in \mathfrak{h}^*} [V \colon L(\lambda)] \operatorname{ch} L(\lambda).$$

Proof. Because both sides of the equation above are formal characters, we match coefficients for each $\mu \in \mathfrak{h}^*$. On the left hand side, the coefficient of e^μ in $\operatorname{ch} V$ is $\dim V_\mu$, and the coefficient of e^μ on the right hand side is

$$\sum_{\lambda \in \mathfrak{h}^*} [V \colon L(\lambda)] \dim L(\lambda)_\mu.$$

Now choose a filtration of V with respect to μ as in lemma 2.3.6. Each composition factor is either isomorphic to $L(\lambda)$ for some $\lambda \geq \mu$, or μ is not a weight of the composition factor, and the number of times that $L(\lambda)$ is a composition factor is $[V \colon L(\lambda)]$. From this, we see that

$$\dim V_\mu \;=\; \sum_{\lambda \colon \lambda \geq \mu} [V \colon L(\lambda)] \dim L(\lambda)_\mu$$

In fact, the sum over all λ such that $\lambda \geq \mu$ is the same as the sum over all $\lambda \in \mathfrak{h}^*$, because $\dim L(\lambda)_\mu = 0$ unless $\lambda \geq \mu$. □

2.4 The Generalized Casimir Operator

There is one more ingredient necessary in the proof of the Kac character formula – the generalized Casimir operator. This is an element in the center of the universal enveloping algebra, and acts on the Verma modules by scalar multiplication, which is key in the determination of the character of $L(\mu)$. Calling it an "operator" originates with physics, where an example is the square of the angular momentum operator.

For a finite dimensional Lie algebra $\mathfrak{g}$ with basis $x_1, \ldots, x_n$, choose a dual basis $y_1, \ldots, y_n$ with respect to the Killing form. Then the **Casimir operator** is given by

$$\sum_{i=1}^{n} x_i y_i \;\in\; U(\mathfrak{g}). \tag{2.12}$$

This acts on the Verma module $M(\lambda)$ by scalar multiplication by $\langle \lambda + \rho, \lambda + \rho \rangle - \langle \rho, \rho \rangle$, where $\langle \cdot\,, \cdot \rangle$ is the Killing form on $\mathfrak{g}$ and $\rho \in \mathfrak{h}^*$ is the half-sum of positive roots $\frac{1}{2}\sum_{\alpha\in\Phi^+}\alpha$, which satisfies $\rho(\alpha_i^\vee) = 1$ for $i = 1, \ldots, r$. For details, see [FH91].

We wish to generalize the Casimir operator to a Kac-Moody algebra $\mathfrak{g}(A)$ for a symmetrizable matrix A (definition 1.4.4), but we cannot define a Casimir element as in (2.13) since the sum may be infinite when $\mathfrak{g}(A)$ is infinite-dimensional. Nevertheless, we can define an operator $\Omega\colon V \to V$ for any $\mathfrak{g}(A)$-module V which behaves in the same way as the Casimir for finite dimensional algebras and reduces to (2.13) in the case $\mathfrak{g}(A)$ is finite-dimensional. In place of the Killing form, we use the standard invariant bilinear form on $\mathfrak{g}(A)$, also denoted $\langle \cdot \mid \cdot \rangle$. In the case that $\mathfrak{g}(A)$ is finite-dimensional, this is just a scalar multiple of the Killing form.

Given a root space decomposition

$$\mathfrak{g}(A) = \mathfrak{h} \oplus \bigoplus_{\alpha\in Q} \mathfrak{g}_\alpha,$$

choose a basis $e_\alpha^{(1)}, e_\alpha^{(2)}, \ldots, e_\alpha^{(t)}$ for $\mathfrak{g}_\alpha$, and using the standard invariant bilinear form, choose a corresponding dual basis $f_\alpha^{(1)}, f_\alpha^{(2)}, \ldots, f_\alpha^{(t)}$ for $\mathfrak{g}_{-\alpha}$ such that $\langle e_\alpha^{(i)}, f_\alpha^{(j)}\rangle = \delta_{ij}$. Additionally, choose a basis $h_1', h_2', \ldots, h_s'$ for $\mathfrak{h}$ and let $h_1'', h_2'', \ldots, h_s''$ be the dual basis of $\mathfrak{h}$ with respect to the standard invariant bilinear form, i.e., $\langle h_i', h_j''\rangle = \delta_{ij}$.

With the notation set up, recall that the fundamental coroots $\alpha_1^\vee, \ldots, \alpha_n^\vee \in \mathfrak{h}$ are linearly independent. Thus, there exists $\rho \in \mathfrak{h}^*$ such that $\rho(\alpha_i^\vee) = 1$ for all $i = 1, \ldots, n$, but ρ may not in general be uniquely determined by this condition. So just choose some ρ that satisfies this condition. Corresponding to ρ there is $\rho^\vee$ such that $\rho(x) = \langle \rho^\vee, x\rangle$ for all $x \in \mathfrak{h}$. Then set

$$\Omega \;=\; \sum_{i=1}^{s} h_i' h_i'' + 2\rho^\vee + 2\sum_{\alpha\in\Phi^+}\sum_{i=1}^{t} f_\alpha^{(i)} e_\alpha^{(i)}. \tag{2.13}$$

Note that the expression above doesn't necessarily make sense as an element of $U(\mathfrak{g}(A))$, because there could be infinitely many positive roots $\alpha \in \Phi^+$, and thus the last sum is infinite. But for $\mathfrak{g}(A)$-module $V \in \mathcal{O}$, we know that $\mathfrak{g}(A)_\alpha V \neq 0$ for only finitely many positive roots $\alpha \in \Phi^+$. Therefore, as an operator $V \to V$, Ω is well-defined. Moreover, a straightforward calculation will show that $\Omega\colon V \to V$ does not depend on the choices of h_1', h_2', …, h_s', h_1'', h_2'', …, h_s'', $e_\alpha^{(1)}$, $e_\alpha^{(2)}$, …, $e_\alpha^{(t)}$ or $f_\alpha^{(1)}$, $f_\alpha^{(2)}$, …, $f_\alpha^{(t)}$. However, it may depend on the choice of ρ.

Definition 2.4.1. *The operator $\Omega\colon V \to V$ defined in* (2.13) *is called the* **generalized Casimir operator** *on V with respect to ρ.*

In the finite-dimensional case, the Casimir operator is in the center of the universal enveloping algebra. While Ω may not even make sense as an element of $U(\mathfrak{g}(A))$ in general, we can still show that it behaves as if it were in the center of $U(\mathfrak{g}(A))$ in the following sense.

Theorem 2.4.2. *Let $u \in U(\mathfrak{g}(A))$ and let $V \in \mathcal{O}$. Then the maps $u\colon V \to V$ (multiplication by u) and $\Omega\colon V \to V$ commute.*

The proof of this theorem is a straightforward, if tedious, calculation. Simply put, it suffices to check that Ω commutes with the generators of $U(\mathfrak{g}(A))$, which are e_i, f_i for $i = 1, \ldots, n$ and the elements of $\mathfrak{h}$. These calculations are done in all of their excruciating detail in section 19.2 of [Car05].

Now knowing this theorem, we can use it to show that Ω acts on a Verma module $M(\lambda)$ as a scalar, as in the finite-dimensional case.

Proposition 2.4.3. *The generalized Casimir operator Ω with respect to ρ acts on the Verma module $M(\Lambda)$ as scalar multiplication by $\langle \Lambda + \rho, \Lambda + \rho \rangle - \langle \rho, \rho \rangle$.*

Proof. Let m_Λ be the highest-weight vector of $M(\Lambda)$. Recall that $M(\Lambda)$ is generated by m_Λ as a $U(\mathfrak{n}^-)$-module, as in remark 2.2.6. And moreover, by the previous theorem, for any $u \in U(\mathfrak{n}^-)$, the action of u and Ω commute. Therefore, it suffices to determine the action of Ω on m_Λ. Hence, we calculate

$$\begin{aligned}
\Omega m_\Lambda &= \left(\sum_{i=1}^{s} h_i' h_i'' + 2\rho^\vee + 2 \sum_{\alpha \in \Phi^+} \sum_{i=1}^{t} f_\alpha^{(i)} e_\alpha^{(i)} \right) m_\Lambda \\
&= \left(\sum_{i=1}^{s} \Lambda(h_i')\Lambda(h_i'') + 2\Lambda(\rho^\vee) \right) m_\Lambda \\
&= \left(\langle \Lambda, \Lambda \rangle + 2\langle \Lambda, \rho \rangle \right) m_\Lambda \\
&= \left(\langle \Lambda + \rho, \Lambda + \rho \rangle - \langle \rho, \rho \rangle \right) m_\Lambda.
\end{aligned}$$

□

Corollary 2.4.4. *Ω acts on the irreducible $L(\Lambda)$ by scalar multiplication by $\langle \Lambda + \rho, \Lambda + \rho \rangle - \langle \rho, \rho \rangle$.*

That Ω acts by a scalar on $L(\Lambda)$ would have followed from our version of Schur's Lemma (lemma 2.2.8), but the particular scalar by which it acts is somewhat more difficult to determine by that lemma alone.

2.5 Kac Character Formula

In this section, we finally prove the Kac Character formula, which is the generalization of the Weyl character formula for a Kac-Moody algebra $\mathfrak{g}$. The Kac Character formula has many further applications – from it, one can construct an explicit set of generators and relations for the Kac-Moody algebra $\mathfrak{g}(A)$ of a Generalized Cartan Matrix A (these were given in remark 1.4.8), deduce the Jacobi Triple Product identity (which appears everywhere from physics to number theory) and its generalizations the MacDonald Identities. Of course, the calculation of the formal character of an irreducible module $L(\Lambda)$ from the category $\mathcal{O}$ is an interesting representation-theoretic result in its own right.

Let $\mathfrak{g}$ be a Kac-Moody algebra of rank n and let $\mathfrak{h}$ be the Cartan sub algebra. We write

$$\begin{aligned} P &= \{\lambda \in \mathfrak{h}^* \mid \lambda(\alpha_i^\vee) \in \mathbb{Z} \text{ for all } i\}; \\ P^+ &= \{\lambda \in P \mid \lambda(\alpha_i^\vee) \geq 0 \text{ for all } i\}; \\ P^{++} &= \{\lambda \in P \mid \lambda(\alpha_i^\vee) > 0 \text{ for all } i\}. \end{aligned}$$

for the sets of the **integral weights** P of $\mathfrak{g}$, the **dominant integral weights** P^+ of $\mathfrak{g}$, and the **regular dominant integral weights** P^{++}, respectively. For most of the remainder of this section, we will restrict our attention to the dominant integral weights P^+. The reason for this restriction to only dominant integral weights is that the modules $L(\Lambda)$ are integrable if and only if $\Lambda \in P^+$.

Definition 2.5.1. *A $\mathfrak{g}$-module V is called* **integrable** *if*

1. *V is $\mathfrak{h}$-diagonalizable, that is, $V = \bigoplus_{\lambda \in \mathfrak{h}^*} V_\lambda$;*
2. *the actions of e_i, f_i are locally nilpotent on V, i.e., for any $v \in V$, there is $N \in \mathbb{N}$ such that $e_i^N \cdot v = 0$.*

We immediately can see that the adjoint module $\mathfrak{g}$ is an example of an integrable module, as in the finite-dimensional case – in proving that, one does not need the finite-dimensional assumption.

The following two lemmas are all that we need to know about integrable $\mathfrak{g}$-modules to prove the character formula. In fact, what we really desire is to know that the action of the Weyl group on the character of $L(\Lambda)$ is trivial. The proofs of these two lemmas are rather similar in strategy: reduce to the finite dimensional case, and specifically to the case of $\mathfrak{sl}_2$.

Lemma 2.5.2. *$L(\Lambda)$ is integrable if and only if Λ is a dominant integral weight ($\Lambda \in P^+$).*

Lemma 2.5.3. *Let V be an integrable $\mathfrak{g}$-module. Then $\dim V_\lambda = \dim V_{w(\lambda)}$ for each $\lambda \in \mathfrak{h}^*$ and $w \in W$.*

Because these lemmas can be reduced to the finite dimensional case rather easily, we refer to [Car05, Prop. 19.14, Prop. 19.13] and [Kac94, Lemma 10.1, Prop. 10.1] for proofs.

We now let the Weyl group act on the algebra $\mathcal{E}$ from definition 2.3.1 of linear combinations of formal exponentials by

$$w \cdot \left(\sum_\lambda c_\lambda e^\lambda \right) = \sum_\lambda c_\lambda e^{w(\lambda)}.$$

An important consequence of the previous lemma is that the action of the Weyl group leaves the character of $L(\Lambda)$ invariant when Λ is a dominant integral weight, because the Weyl group action simply interchanges the coefficients in the formula for the character of $L(\Lambda)$.

Corollary 2.5.4. *For $\Lambda \in P^+$ and $w \in W$,*

$$w(\operatorname{ch} L(\Lambda)) \;=\; \operatorname{ch} L(\Lambda).$$

The following lemma regarding equality of dominant and regular dominant weights will be used several times in the course of proving the character formula and other results, so we present it here, independent of its use in the character formula.

Lemma 2.5.5. *Let $\xi \in P^{++}$ and $\eta \in P^+$ such that $\eta \le \xi$ and $\langle \eta, \eta \rangle = \langle \xi, \xi \rangle$. Then $\eta = \xi$.*

Proof. Since $\eta \le \xi$, then $\xi - \eta$ is a sum of simple roots, say

$$\xi - \eta = \sum_{i=1}^{n} k_i \alpha_i,$$

for nonnegative integers k_i. It follows that

$$\langle \xi, \xi \rangle - \langle \eta, \eta \rangle \;=\; \langle \xi + \eta, \xi - \eta \rangle \;=\; \sum_{i=1}^{n} k_i \langle \xi + \eta, \alpha_i \rangle \;=\; \sum_{i=1}^{n} \frac{k_i}{2} \langle \alpha_i, \alpha_i \rangle (\xi + \eta)(h_i).$$

But for simple roots, $\langle \alpha_i, \alpha_i \rangle > 0$, and because $\xi \in P^{++}$, $\eta \in P^+$, we also have that $(\xi + \eta)(h_i) > 0$. And on the other hand, we have that $\langle \xi, \xi \rangle - \langle \eta, \eta \rangle = 0$, so $k_i = 0$ for each i. Thus, $\xi = \eta$. $\square$

Finally, we turn our attention to the Kac Character formula. Fix ρ such that $\rho(\alpha_i^\vee) = 1$ for all i and let Ω be the Casimir operator with respect to ρ.

For elements w of the Weyl group W, define the **length** $\ell(w)$ as the smallest number k such that w is the product of k-many reflections by simple roots. Then let $\epsilon(w) = (-1)^{\ell(w)}$.

Theorem 2.5.6 (Kac Character Formula)**.** *Let $\mathfrak{g}(A)$ be a symmetrizable Kac-Moody algebra and $L(\Lambda)$ an irreducible $\mathfrak{g}(A)$-module with Λ a dominant integral weight. Then we have the following equality in $\mathcal{E}$*

$$\operatorname{ch} L(\Lambda) \;=\; \frac{\displaystyle\sum_{w \in W} \varepsilon(w) e^{w(\Lambda+\rho)-\rho}}{\displaystyle\prod_{\alpha \in \Phi^+} (1 - e^{-\alpha})^{\operatorname{mult}\alpha}}.$$

The proof of this formula proceeds in several steps. As a general outline, we first use Lemma 2.3.9 to write the character for the Verma module $M(\Lambda)$ in terms of the characters for $L(\mu)$, which involves coefficients $[M(\Lambda)\colon L(\mu)]$, as follows

$$\operatorname{ch} M(\Lambda) = \sum_{\mu \in \mathfrak{h}^*} [M(\Lambda)\colon L(\mu)] \operatorname{ch} L(\mu).$$

The action of the Casimir operator on $M(\Lambda)$ and $L(\mu)$ shows us that we may invert the sum to write the character for $L(\Lambda)$ in terms of the characters of $M(\mu)$,

$$\operatorname{ch} L(\Lambda) = \sum_{\mu \in \mathfrak{h}^*} b_{\Lambda\mu} M(\mu).$$

We know what the characters of $M(\mu)$ are by Proposition 2.3.5. To nail down the coefficients on the right hand side, we then consider the action of the Weyl group on the characters. By corollary 2.5.4, the action of W on $\operatorname{ch} L(\Lambda)$ is trivial, but it allows us to exactly determine what the coefficients $b_{\Lambda\mu}$ should be.

So we begin.

Proof of the Kac Character Formula. By lemma 2.3.9, we see that

$$\operatorname{ch} M(\Lambda) = \sum_{\mu \in \mathfrak{h}^*} [M(\Lambda) \colon L(\mu)] \operatorname{ch} L(\mu). \tag{2.14}$$

As $M(\Lambda)$ is a Verma module with highest weight Λ, a filtration of $M(\Lambda)$ can only contain composition factors $L(\mu)$ for $\mu \leq \lambda$. So if $[M(\Lambda) \colon L(\mu)] \neq 0$, then we must have that $\mu \leq \Lambda$.

Now by proposition 2.4.3 Ω acts on $M(\Lambda)$ as scalar multiplication by $\langle \Lambda + \rho, \Lambda + \rho\rangle - \langle \rho, \rho\rangle$. Similarly, by corollary 2.4.4 Ω acts on $L(\mu)$ by scalar multiplication by $\langle \mu + \rho, \mu + \rho\rangle - \langle \rho, \rho\rangle$. Therefore, if $[M(\Lambda) \colon L(\mu)] \neq 0$, we must have

$$\langle \Lambda + \rho, \Lambda + \rho\rangle - \langle \rho, \rho\rangle = \langle \mu + \rho, \mu + \rho\rangle - \langle \rho, \rho\rangle,$$

or equivalently,

$$\langle \Lambda + \rho, \Lambda + \rho\rangle = \langle \mu + \rho, \mu + \rho\rangle. \tag{2.15}$$

Let $S = \{\mu \in \mathfrak{h}^* \mid \mu \leq \Lambda,\ \langle \mu + \rho, \mu + \rho\rangle = \langle \Lambda + \rho, \Lambda + \rho\rangle\}$. We can re-index the sum in (2.14) to run over only those $\mu \in S$, because those are the only terms with nonzero coefficient as per (2.15).

$$\operatorname{ch} M(\Lambda) \ = \sum_{\mu \in S} [M(\Lambda) \colon L(\mu)] \operatorname{ch} L(\mu). \tag{2.16}$$

By extending the partial ordering $\leq$ on weights to a total ordering on those weights in S, we may write

$$\operatorname{ch} M(\Lambda) \ = \sum_{\mu \in S} a_{\Lambda\mu} \operatorname{ch} L(\mu), \tag{2.17}$$

where $a_{\Lambda\mu} = [M(\Lambda) \colon L(\mu)]$.

The next step is to invert the sum in (2.17) to write the character of $L(\Lambda)$ in terms of the characters for $M(\mu)$. To that end, note that the infinite matrix $(a_{\Lambda\mu})$ has nonnegative integer entries, is upper triangular, and has entries $a_{\lambda\lambda} = 1$ along the diagonal. Such a matrix is invertible, so we may write

$$\operatorname{ch} L(\Lambda) \ = \sum_{\mu \in S} b_{\Lambda\mu} \operatorname{ch} M(\mu) \tag{2.18}$$

for some integer coefficients $b_{\Lambda\mu}$ such that the infinite matrix $(b_{\lambda\mu})$ is upper triangular with $b_{\lambda\lambda} = 1$ for all $\lambda \in S$.

We know the character of $M(\mu)$ by proposition 2.3.5, so we substitute in (2.18) to get

$$\operatorname{ch} L(\Lambda) \;=\; \sum_{\mu \in S} b_{\Lambda\mu} \,\frac{e^{\mu}}{\prod_{\alpha\in\Phi^+}(1-e^{-\alpha})^{\operatorname{mult}\alpha}}. \tag{2.19}$$

For ease of notation, let $\Delta = \prod_{\alpha\in\Phi^+}(1-e^{-\alpha})^{\operatorname{mult}\alpha}$, and multiply both sides of (2.19) by $e^{\rho}\Delta$, whence we arrive at

$$e^{\rho}\,\Delta \operatorname{ch} L(\Lambda) \;=\; \sum_{\mu\in S} b_{\Lambda\mu} e^{\mu+\rho}. \tag{2.20}$$

Now we wish to determine the coefficients $b_{\Lambda\mu}$, or at least rewrite the sum on the right-hand side without these unknowns. To do this, we use the action of the Weyl group on both sides of (2.20). It suffices to consider the action of generators of W. The generators of the Weyl group are the reflections s_i, one for each $\alpha_i \in \Phi$. On roots, s_i takes α_i to $-\alpha_i$ and leaves $\Phi^+ \setminus \{\alpha_i\}$ untouched. By lemma 2.5.3, $\operatorname{mult} s_i\alpha = \operatorname{mult}\alpha$.

$$\begin{aligned}
s_i \cdot (e^{\rho}\Delta) &= s_i \cdot \left(e^{\rho}(1-e^{-\alpha_i}) \prod_{\alpha\in\Phi^+\setminus\{\alpha_i\}} (1-e^{-\alpha})^{\operatorname{mult}\alpha} \right) \\
&= e^{\rho-\alpha_i}(1-e^{\alpha_i}) \prod_{\alpha\in\Phi^+\setminus\{\alpha_i\}} (1-e^{-\alpha})^{\operatorname{mult}\alpha} \\
&= e^{\rho} e_{-\alpha_i}(1-e^{\alpha_i}) \prod_{\alpha\in\Phi^+\setminus\{\alpha_i\}} (1-e^{-\alpha})^{\operatorname{mult}\alpha} \\
&= e^{\rho}(e^{-\alpha}-1) \prod_{\alpha\in\Phi^+\setminus\{\alpha_i\}} (1-e^{-\alpha})^{\operatorname{mult}\alpha} \\
&= -e^{\rho}\Delta
\end{aligned}$$

It follows that for any $w \in W$,

$$w(e^{\rho}\Delta) = \varepsilon(w) e^{\rho}\Delta. \tag{2.21}$$

Note also that $\operatorname{ch} L(\Lambda)$ is invariant under the action of W by corollary 2.5.4. So the action of the Weyl group on (2.20) picks up a factor of $\varepsilon(w)$ on the left hand side. Hence, for any $w \in W$,

$$w \cdot \left(\sum_{\mu\in S} b_{\Lambda\mu} e_{\mu+\rho} \right) \;=\; \varepsilon(w) \sum_{\mu\in S} b_{\Lambda\mu} e_{\mu+\rho}$$

But also, by the definition of the action of W on $\mathcal{E}$,

$$w \cdot \left(\sum_{\mu\in S} b_{\Lambda\mu} e^{\mu+\rho} \right) \;=\; \sum_{\mu\in S} b_{\Lambda\mu} e^{w(\mu+\rho)}.$$

Equating coefficients in the previous two expressions, we see that

$$b_{\Lambda\mu} = \varepsilon(w) b_{\Lambda\nu} \tag{2.22}$$

if $w(\mu+\rho) = \nu+\rho$ for some $w \in W$. This is exactly the fact that will allow us to determine the coefficients $b_{\Lambda\mu}$ of (2.20). This is the last step of the proof.

Suppose that μ is a weight for which $b_{\Lambda\mu} \neq 0$, and consider the set T_μ of all weights ν such that $w(\mu+\rho) = \nu+\rho$ for some $w \in W$, that is, $T_\mu = \{\nu \in \mathfrak{h}^* \mid w(\mu+\rho) = \nu+\rho \text{ for some } w \in W\}$. By (2.22), all such weights $\nu \in T_\mu$ satisfy $b_{\Lambda\nu} \neq 0$, and moreover we must have $\nu \leq \lambda$. Choose a weight ν such that the height of $\Lambda - \nu$ is minimal. I claim that $\nu+\rho$ must be a dominant integral weight. If not, then there is some i such that $(\nu+\rho)(h_i) < 0$, so set

$$\xi = s_i w(\mu+\rho) = s_i(\nu+\rho) = \nu+\rho-(\nu+\rho)(h_i)\alpha_i.$$

Note that $\xi \in T_\mu$ by construction. The height of $\Lambda-\xi$ is less than the height of $\Lambda-\nu$, contradicting the minimality of the height of $\Lambda-\nu$. So it must be that $\nu+\rho \in P^+$.

Because $b_{\Lambda\nu} \neq 0$, $b_{\Lambda\mu} \neq 0$, each of the weights μ, ν appears in the sum on the right hand side of (2.20). Hence, $\mu, \nu \in S$, which means that $\langle \mu+\rho, \mu+\rho\rangle = \langle \Lambda+\rho, \Lambda+\rho\rangle$ and $\langle \nu+\rho, \nu+\rho\rangle = \langle \Lambda+\rho, \Lambda+\rho\rangle$. So we satisfy all of the conditions of lemma 2.5.5 for the weights $\nu+\rho$ and $\Lambda+\rho$. Namely, $\nu+\rho \in P^+$, $\Lambda+\rho \in P^{++}$, $\nu+\rho \leq \lambda+\rho$, and $\langle \nu+\rho, \nu+\rho\rangle = \langle \Lambda+\rho, \Lambda+\rho\rangle$. This implies that $\nu = \Lambda$.

Then for any weight μ such that $b_{\Lambda\mu} \neq 0$, there is $w \in W$ such that $\mu+\rho = w(\Lambda+\rho)$, and moreover by (2.22), $b_{\Lambda\mu} = \varepsilon(w) b_{\Lambda\Lambda} = \varepsilon(w)$. Happily, this means that we can reduce in (2.20) the sum over a possibly infinite set S of weights to a finite sum over $w \in W$.

$$\sum_{\mu \in S} b_{\Lambda\mu} e^{\mu+\rho} \;=\; \sum_{w \in W} \varepsilon(w) e^{w(\lambda+\rho)}$$

Finally, we substitute this back into the right-hand side of equation (2.20) and divide both sides by $e_\rho \Delta$ to isolate $\operatorname{ch} L(\Lambda)$. The result is the Kac Character formula

$$\operatorname{ch} L(\Lambda) \;=\; \frac{\displaystyle\sum_{w \in W} \varepsilon(w) e^{w(\lambda+\rho)-\rho}}{\Delta}.$$

Expanding Δ, we get the form promised in the theorem.

$$\operatorname{ch} L(\Lambda) \;=\; \frac{\displaystyle\sum_{w \in W} \varepsilon(w) e^{w(\Lambda+\rho)-\rho}}{\displaystyle\prod_{\alpha \in \Phi^+} (1-e^{-\alpha})^{\operatorname{mult}\alpha}}. \qquad \square$$

We would be amiss to prove the character formula and omit some of the immediate corollaries that it entails. In particular, when $\Lambda = 0$, $L(0)$ is the trivial $\mathfrak{g}$-module. Hence, we know that $\operatorname{ch} L(0) = e^0$, which is the identity element of $\mathcal{E}$. The character formula states that:

$$e^0 = \operatorname{ch} L(0) = \frac{\sum_{w\in W} \varepsilon(w) e^{w(\Lambda+\rho)-\rho}}{\prod_{\alpha\in\Phi^+}(1-e^{-\alpha})^{\operatorname{mult}\alpha}}.$$

But e^0 is the identity element of $\mathcal{E}$, so we may clear denominators and multiply both sides by e^ρ to arrive at the following. This is sometimes known as **Kac's denominator formula**.

Corollary 2.5.7 (Kac's denominator formula)**.**

$$e^\rho \prod_{\alpha\in\Phi^+}(1-e^{-\alpha})^{\operatorname{mult}\alpha} \;=\; \sum_{w\in W} \varepsilon(w) e^{w(\rho)}.$$

We can then use this to obtain an alternative form of Kac's character formula: take the character formula and multiply numerator and denominator of the right hand side by e_ρ. Then apply the denominator formula (corollary 2.5.7) to the denominator, and we get this exciting new form of Kac's character formula.

Corollary 2.5.8.

$$\operatorname{ch} L(\Lambda) \;=\; \frac{\displaystyle\sum_{w\in W} \varepsilon(w) e^{w(\Lambda+\rho)}}{\displaystyle\sum_{w\in W} \varepsilon(w) e^{w(\rho)}}.$$

Bibliography

[BGG76] I.N. Bernstein, I.M. Gelfand, and S.I. Gelfand. A certain category of $\mathfrak{g}$-modules. *Funkt. Analis i ego Prilozh*, 10(2):1–8, 1976. English Translation: Funct. Anal. Appl. 12:4 (1978), 79-80.

[Car05] Roger William Carter. *Lie algebras of finite and affine type*. Number 96 in Cambridge studies in advanced mathematics. Cambridge University Press, 2005.

[FH91] William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer, 1991.

[Hum72] James E Humphreys. *Introduction to Lie algebras and representation theory*, volume 9. Springer Science & Business Media, 1972.

[Kac94] Victor G. Kac. *Infinite-dimensional Lie algebras*, volume 44. Cambridge university press, 1994.

[Moo79] Robert V Moody. Root systems of hyperbolic type. *Advances in Mathematics*, 33(2):144–160, 1979.

[Poi99] Henri Poincare. Sur les groupes continus. *Transactions. Cambridge Philosophical Society*, 18:220, 1899.

[Run12] Ingo Runkel. http://www.math.uni-hamburg.de/home/runkel/Material/SS12/afflie-weylkac.pdf, 2012. Accessed: March 23, 2015.

[TTT99] Tuong Ton-That and Thai-Duong Tran. Poincaré's proof of the so-called birkhoff-witt theorem. *Rev. Histoire Math*, 5:249–284, 1999.

[Vin71] E.B. Vinberg. Discrete linear groups generated by reflections. *Izvestija AN USSR*, 35:1072–1112, 1971. English Translation: Math. USSR-Izvestija 5 (1971), 1083-1119.