HKUST 5143: Introduction to Lie Algebras

Lectures by Eric Marberg, Notes by Gary Hu

In Spring 2022, Eric Marberg taught HKUST 5143: Introduction to Lie Algebras. It consisted of 13 lectures.

This an unofficial set of notes scribed by Gary Hu, who is responsible for all mistakes. If you do find any errors, please report them to: gh7@williams.edu

Contents

1 Foundations of Lie Algebras

1.1 Basic Definitions, Solvable and Nilpotent Lie Algebras

1.2 Overview

The main reference is Humphrey's *Introduction to Lie Algebras and Representation Theory*. Some other books are listed on the website.

Today, we'll see the basic definitions, some important examples, and frame the guiding classification problems.

For some very brief motivation: the most interesting groups in physics, geometry, etc. are Lie groups, which are groups that are also manifolds in a compatible way. For example, the general linear group, the special linear group, the orthogonal group, the symplectic group, etc. The most important features in geometry/representation theory of Lie groups are controlled by the tangent space at the identity element. This tangent space has more structure than just being a vector space - namely, it is what we call a Lie algebra.

1.3 What Is A Lie Algebra

Let's start with a constructive definition. Let $\mathbb{F}$ be some field (e.g., $\mathbb{R}, \mathbb{C}, \mathbb{Q}, \mathbb{F}_p$, etc.). Let n be a positive integer and define

$$gl_n(\mathbb{F}) = \{n \times n \text{ matrices over } \mathbb{F}\}$$

For $X, Y \in gl_n(\mathbb{F})$ let $[X, Y] = XY - YX$.

Definition 1.1. *A (finite-dimensional)* ***Lie algebra*** *is a subspace $L \subseteq gl_n(\mathbb{F})$ such that $[X, Y] \in L \forall X, Y \in L$.*

Example 1.2.

1. *$L = gl_n(\mathbb{F})$, the* ***general linear Lie algebra****.*
2. *$L = \{$diagonal matrices in $gl_n(\mathbb{F})\}$, called $d_n(\mathbb{F})$.*
3. *$L = \{$upper triangular matrices in $gl_n(\mathbb{F})\}$, called $t_n(\mathbb{F})$.*
4. *$L = \{$strictly upper triangular matrices in $gl_n(\mathbb{F})\}$, called $N_n(\mathbb{F})$.*

In fact, any subalgebra $L \subseteq gl_n(\mathbb{F})$ (a subspace of matrices closed under multiplication) is a Lie algebra. Since if $X, Y \in L$ then $[X, Y] = XY - YX \in L$.

But there are more interesting examples.

Example 1.3. *Here are the classical Lie algebras.*

1. *A_ℓ: $\mathfrak{sl}_n(\mathbb{F}) = \{x \in gl_n(\mathbb{F}) | tr(x) = 0\}$, the* ***special linear Lie algebra****. Note that this is not an algebra.*
2. *C_ℓ: Suppose $n = 2m$ is even. We define*

$$\mathfrak{sp}_n(\mathbb{F}) = \{X = \begin{bmatrix} M & N \\ P & -M^T \end{bmatrix} | M, N, P \in gl_m(\mathbb{F}), N = N^T, P = P^T\}$$

to be the ***symplectic Lie algebra****.*

1. *D_ℓ: Suppose $n = 2m$ is even. We define*

$$\mathfrak{O}_n(\mathbb{F}) = \{X = \begin{bmatrix} M & N \\ P & -M^T \end{bmatrix} | M, N, P \in gl_m(\mathbb{F}), N^T = -N, P^T = -P\}$$

to be the ***even orthogonal Lie algebra****.*

1. *B_ℓ: Suppose $n = 2m + 1$ is odd. We define*

$$\mathfrak{O}_n(\mathbb{F}) = \{X = \begin{bmatrix} 0 & A & B \\ -A^T & M & N \\ -B^T & P & -M^T \end{bmatrix} | A, B \text{ are } 1 \times m \text{ vectors}, M, N, P \in gl_m(\mathbb{F}), N^T = -N, P^T = -P\}$$

to be the ***odd orthogonal Lie algebra****.*

It's not entirely clear what the motivation is yet, but we will see this soon. Later today, we will give "basis independent" definitions.

We call $[\cdot, \cdot]$ the **Lie bracket**. Here are it's properties:

1. The Lie bracket is **bilinear**: $[a_1X_1 + a_2X_2, b_1Y_1 + b_2Y_2] = \sum_{i,j=1}^2 a_i b_j [X_i, Y_j] \forall a_i, b_i \in \mathbb{F}, X_i, Y_i \in gl_n(\mathbb{F})$.
2. The Lie bracket is **alternating**: $[X, X] = 0 \forall X$.
3. Properties 1 and 2 imply that the bracket is **skew-symmetric**: $[X, Y] = -[Y, X]$.
4. Let $\text{ad}X$ denote the map $gl_n(\mathbb{F}) \to gl_n(\mathbb{F}), (\text{ad}X)(Y) = [X, Y]$. Then, it holds that $\text{ad}_{[X,Y]} = [\text{ad}X, \text{ad}Y] \forall X, Y$.

How do we know that (4) is true? For any vector space V, let $gl(V)$ be the space of linear maps $V \to V$. For any $f, g \in gl(V)$, define $[f, g] = f \circ g - g \circ f = fg - gf$. Then

$$\begin{aligned}(\text{ad}[X, Y])(Z) &= [[X, Y], Z] \\ &= [XY - YX, Z] \\ &= XYZ - YXZ - ZXY + ZYX\end{aligned}$$

and

$$
\begin{aligned}
[\mathrm{ad}X, \mathrm{ad}Y](Z) &= \mathrm{ad}X(\mathrm{ad}Y(Z)) - \mathrm{ad}Y(\mathrm{ad}X(Z)) \\
&= X(YZ - ZY) - (YZ - ZY)X - Y(XZ - ZX) + (XZ - ZX)Y
\end{aligned}
$$

which are equal after cancellation. Thus, the two are equal as linear maps $gl_n(\mathbb{F}) \to gl_n(\mathbb{F})$. In words: ad commutes with the Lie bracket, or ad is a homomorphism $gl_n(\mathbb{F}) \to gl(gl_n(\mathbb{F}))$. But this is circular without an abstract definition of Lie algebras (and homomorphisms), which brings us to our next topic: defining Lie algebras abstractly.

Suppose L is an $\mathbb{F}$-vector space with a map $[\cdot,\cdot] : L \times L \to L$, called the Lie bracket.

Definition 1.4. *L is a* ***Lie algebra*** *with respect to $[\cdot,\cdot]$ if the following conditions hold:*

1. *The bracket is bilinear.*
2. *The bracket is alternating*
3. *$ad_[X, Y] = ad_X ad_Y - ad_Y ad_X := [ad_X, ad_Y]$ for all $X, Y \in L$.*

Remark 1.5.
1. *L may be infinite dimensional but we will rarely conside this case, the theory is much more involved. Unless stated explicitly, all Lie algebraw L are assumed to have $\dim(L) < \infty$.*
2. *Axioms 1 and 2 imply that the Lie bracket is always skew-symmetric. It might seem more natural to replace the second xiom with skew-symmetry, but this leads to problems when $char(\mathbb{F}) = 2$.*
3. *Axioms 3 is called the* ***Jacobi identity*** *and is equivalent to*

$$
[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \forall X, Y, Z \in L
$$

Usually, given a vector space has only one natural Lie algebra structure and so we reuse the symbol $[\cdot,\cdot]$ to denote the Lie bracket for any Lie algebra.

1.4 Lie Subalgebras and Morphisms

Definition 1.6. *A* ***Lie subalgebra*** *of a Lie algebra L is a subspace $K \subseteq L$ with $[X, Y] \in K \forall X, Y \in K$ a linear map $\phi : L_1 \to L_2$ between Lie algebras is a (Lie algebra)* ***morphism*** *if $\phi([X, Y]) = [\phi(X), \phi(Y)] \forall X, Y$.*

Example 1.7. *Let V be an $\mathbb{F}$-vector space. Then $gl(V)$ is a Lie algebra for the bracket $[f, g] = fg - gf$. The Jacobi identity says that $ad : L \to gl(L)$ is a morphism.*

Definition 1.8. *A Lie algebra* ***isomorphism*** *if a morphism that is a bijection.*

If $\dim(V) = n < \infty$ then choosing a basis for V defines a isomorphism $gl(V) \xrightarrow{\sim} gl_n(\mathbb{F})$.

Here are some more abstract examples.

Example 1.9.

1. *The trace of* $X \in gl(V)$ *is well0defined whenever* $\dim(V) < \infty$*, independent of the choice of basis. So we can define* $\mathfrak{sl}(V) = \{X \in gl(V) | tr(X) = 0\}$ *that is a Lie subalgebra of* $gl(V)$ *by the same argument as earlier.*

2. *Suppose* $B : V \times V \to V$ *is some bilinear form. Example: if* $V = \mathbb{F}^n$ *then every* B *has formula* $B(x, y) = x^T My$ *for some fixed* $n \times n$ *matrix* M*. Then the subspace*

$$L = \{X \in gl(V) | B(Xu, v) = -B(u, Xv) \forall u, v \in V\}$$

is a Lie subalgebra of $gl(V)$*. This is because if* $X, Y \in L$ *then* $B([X, Y]u, v) = B(XYu, v) - B(YXu, v) = -B(Yu, Xv) + B(Xu, Yv) = B(u, YXv) - B(u, XYv) = B(u, [X, Y]v) \forall u, v \in V$ *so* $[X, Y] \in L$*.*

Assume $\dim(V) < \infty$*. If* B *is symmetric and nondegenerate then* $L \cong \mathfrak{O}_n(\mathbb{F})$*. The explicit construction of* $\mathfrak{O}_n(\mathbb{F})$ *earlier corresponds to taking* $B(u, v) = u^T Mv$ *for the matrices* $M = \begin{bmatrix} 0 & I_m \\ I_m & 0 \end{bmatrix}$ *if* $n = 2m$ *is even and* $M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & I_m \\ 0 & I_m & 0 \end{bmatrix}$ *if* $n = 2m + 1$ *is odd.*

If B *is skew-symmetric and nondegenerate, then* $n = 2m$ *has to be even and* $L \cong sl_n(\mathbb{F})$*. The explicit construction earlier had* $B(u, v) = u^T Mv$ *for* $M = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}$.

1. *Suppose* A *is an* $\mathbb{F}$*-algebra, not necessarily associative. (*A *is just a vector space with a bilinear multiplication). We have seen that* $gl(A)$ *is a Lie algebra with bracket* $[X, Y] = XY - YX$*. Let* $Der(A) = \{$ *linear maps*$\delta : A \to A$ *with*$\delta(ab) = a\delta(b) + \delta(a)b \forall a, b\}$*. Call elements* $\delta \in Der(A)$ ***derivations****. It's easy to check that Der is a subalgebra of* $gl(V)$*.*

Should also mention: If A Is an associative algebra, then A can be viewed as a Lie algebra for the bracket $[X, Y] = XY - YX$. The fact that this satisfies the Jacobi identity does require associativity of the algebra.

Next: a laundry list of analogies with group theory.

1.5 Lie Algebras vs. Groups

Notation: $\mathrm{ad}X : Y \mapsto [X, Y]$ for X, Y in a Lie algebra L, $\mathrm{Ad}g : h \mapsto g^{-1}hg$ for g, h in a group G.

Lie Algebras L — **Groups** G — — — An **ideal** of L, $I \subseteq L$, satisfies $(\text{ad}X)(I) \subseteq I$ for all $X \in L$. — A **normal subgroup** of G, $H \subseteq G$, satisfies $(\text{ad}g)(H) \subseteq H$ for all $g \in G$. The **center** of L, $Z(L) = \{Y \in L \mid (\text{ad}X)(Y) = 0$ for all $X \in L\}$. — The **center** of G, $Z(G) = \{h \in G \mid (\text{ad}g)(h) = h$ for all $g \in G\}$. The **quotient Lie algebra** $L/I = \{X+I \mid X \in L\}$ with bracket $[X+I, Y+I] = [X,Y]+I$ for $X, Y \in L$. — The **quotient group** $G/N = \{gN \mid g \in G\}$ is formed with the usual set product. The **derived Lie algebra** $[L,L] = \text{span}(\{[X,Y] \mid X, Y \in L\})$. — The **derived subgroup** $[G,G]$ is generated by $\{ghg^{-1}h^{-1} \mid g, h \in G\}$. L is **abelian** if $L = Z(L)$ or $[L,L] = 0$. — G is **abelian** if $G = Z(G)$ or $[G,G] = \{1\}$. L is **simple** if L is non-abelian and has no proper, nonzero ideals. — G is **simple** if G has no proper, nontrivial normal subgroups.

Some other terminology:

- The **normalizer** of a Lie subalgebra $K \subseteq L$ is $N_L(K) = \{x \in L | (\text{ad}X)(K) \subseteq K\}$. This is a Lie subalgebra, the largest one such that $K \subseteq N_L(K)$ is an ideal.
- The **centralizer** of a subspace $K \subseteq L$ is $C_L(k) = \{X \in L | (\text{ad}X)(K) = 0\}$. This is another Lie subalgebra.

Example 1.10. *Suppose that* $L = \mathfrak{sl}_2(\mathbb{F})$. *Assume char*$(\mathbb{F}) \neq 2$. *A basis is* $X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, Y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. *The Lie brackets are:*

- $[X,X] = [Y,Y] = [H,H] = 0$
- $[X,Y] = -[Y,X] = H$
- $[H,X] = -[X,H] = 2X$
- $[H,Y] = -[Y,H] = -2Y$

Note: adH : $Z \mapsto [H,Z]$ *has eigenvalues* $-2, 0, 2$ *with eigenvectors* Y, H, X. *So adH is diagonalizable.*

The claim is that $L = \mathfrak{sl}_2(\mathbb{F})$ is simple when $\text{char}(\mathbb{F}) = neq2$.

Suppose $0 \neq g := aX + bY + cH$ for $a, b, c \in \mathbb{F}$ belongs to an ideal $I \subseteq L$. Then

$$\begin{aligned} [X,[X,g]] &= [X, bH - 2cX] = -2bX \in I \\ [Y,[Y,g]] &= [Y, aH - 2cY] = -2aY \in I \end{aligned}$$

If $a \neq 0$ then $Y \in I$, but then $H \in I$, so $S \in I = L$. If $b \neq 0$ then $X \in I$, then $H \in I$, so $Y \in I = L$. If $a = b = 0$ then $H \in I$, so $X, Y \in I = L$. Thus, $I = L$.

Now, we'll review some basic facts about quotients.

1. If $\phi : L \to K$ is a surjective Lie algebra morphism, then the kernel is an ideal of L and $L/\text{Ker}(\phi) \cong K$. via the map $X + \text{Ker}(\phi) \mapsto \phi(X)$ for $X \in L$.

2. If $I, J \subseteq L$ are ideals and $I \subseteq J$, then J/I is an ideal of L/I and $(L/I)/(J/I) \stackrel{\cong}{\to} L/J, (X+I)+J/I \mapsto X+J$ as Lie algebras.

3. If $I, J \subseteq L$ are ideals then $(I+J)/J \cong I/(I \cap J), (i+j)+J \mapsto i+I \cap J$.

Some more terminology:

Definition 1.11. *A **representation** of a Lie algebra L is a Lie algebra morphism $\phi : L \to gl(V)$ for some (not necessarily finite dimensional vector space V).*

Example 1.12. *The **adjoint representation** $ad : L \to gl(V)$ is a representation.*

The most interesting Lie algberas arise as subalgebras of $gl(V)$.

Proposition 1.13. *Any simple Lie algebra is isomorphic to a subalgebra of a general linear Lie algebra.*

Proof. More generally, if $Z(L) := \{X \in L | [X, Y] = 0 \forall Y\}$ then $Z(L) = \mathrm{Ker}(\mathrm{ad})$ so $L/Z(L) = L/\mathrm{Ker}(\mathrm{ad}) \cong \mathrm{ad}(L) \subseteq gl(L)$. Therefore $L \cong$ subalgebra of $gl(V)$ whenever $Z(L) = 0$. The center is an ideal so it must be zero if L is simple.

□

1.6 Solvable and Nilpotent Lie Algebras

Derived series of a Lie algebra L: $L^{(0)} = L, L^{(n+1)} = [L^{(n)}, L^{(n)}]$. Recall if $I, J \subseteq L$ then $[I, J]$ is the span of $\{[X, Y] | X \in I, Y \in J\}$.

Definition 1.14. *L is **solvable** if $L^{(n)} = 0$ for some $n > 0$.*

Example 1.15. *One can check that $t_n(\mathbb{F})^{(1)} = N_n(\mathbb{F}), t_n(\mathbb{F})^{(k)} \subseteq span(E_{ij} | j - i \geq 2^{k-1})$ so $t_n(\mathbb{F})^(k) = 0$ if $2^{k-1} > n-1$ so $t_n(\mathbb{F})$ is solvable.*

Proposition 1.16. *L is a Lie algebra. If L is solvable then so are all subalgebras and homomorphic images of L.*

Proof. If $K \subseteq L$ then $K^{(n)} \subseteq L^{(n)}$ and $\phi(L)^{(n)} = \phi(L^{(n)})$ if ϕ is a morphism.

□

Proposition 1.17. *If $I \subseteq L$ is a solvable ideal and L/I is solvable then L is solvable.*

Proof. In this case $L^{(n)} \subseteq I$ for some $n > 0$ and $I^{(m)} = 0$ for some $m > 0$ so $L^{(m+n)} = 0$.

□

Proposition 1.18. *If $I, J \subseteq L$ are both solvable ideals then so is $I + J$.*

Proof. $(I+J)/J \cong I/I \cap J$ is solvable, as is J.

□

Corollary 1.19. *Any Lie algebra L has a unique maximal solvable ideal (which is equal to L if and only if L is solvable).*

Proof. If S is a maximal solvable ideal of L and $I \subseteq L$ is any solvalbe ideal then $S+I$ is solvable and contains S, so it must be equal to S. Thus if I is maximal then $S = S+I = I$.

□

We denote the unique maximal solvable ideal of a Lie algebra L by $\mathrm{Rad}(L)$, call it the **radical**.

Definition 1.20. *L is semisimple if $Rad(L) = 0$. That is, if L has no nonzero solvable ideals.*

Later we will see that semisimple $\iff$ "direct sum of simple".

Proposition 1.21. *$L/Rad(K)$ is semisimple.*

Proof. Preimage of any nonzero ideal in $L/\mathrm{Rad}(K)$ is an ideal $I \subseteq L$ containing $\mathrm{Rad}(L)$ so is not solvable. So by propositions, $I/\mathrm{Rad}(L)$ is not solvable.

□

Proposition 1.22. *Any simple L is semisimple.*

Proof. If L is simmple then $0 \neq [L,L] = L$ so L is not solvable, so $\mathrm{Rad}(L)$ is a proper ideal so it must be zero.

□

Lower/descending central series: $L^0 = L, L^{n+1} = [L, L^n]$.

Definition 1.23. *The lower/descending central series is defined recursively as $L^0 = L, L^{n+1} = [L, L^n]$. L is **nilpotent** if $L^n = 0$ for some $n > 0$.*

Note that nilpotent (strictly upper triangular matrces) $\rightarrow$ solvable (upper triangular matrices) but the reverse does not hold.

We can show that $t_n(\mathbb{F})$ has $t_n(\mathbb{F})^k = N_n(\mathbb{F})$ for all $k \geq 1$. So $t_n(\mathbb{F})$ is solvalbe but not nilpotent. But $N_n(\mathbb{F})$ is nilpotent.

Next time: a little more discussion of nilpotent Lie algebras and Engel's theorem. Then we will discuss the problem of classifying semisimple Lie algebras.

1.7 Engel's and Lie's Theorem, Jordan Decomposition

1.8 Engel's Theorem

Proposition 1.24. *If L is nilpotent, then so are all of its subalgebras and homomorphic images.*

Proof. If $K \subseteq L$ then $K^n \subseteq L^n$ and if $\varphi : L \to K$ is a morphism then $\phi(L)^n = \phi(L^n)$.

□

Proposition 1.25. *If $L/Z(L)$ is nilpotent then L is nilpotent.*

Proof. In this case, $L^n \subseteq Z(L)$ for some $n > 0$ and then $L^{n+1} \subseteq [L, Z(L)] = 0$.

□

Proposition 1.26. *If L is nilpotent and $L \neq 0$ then $Z(L) \neq 0$.*

Proof. If $L^n \neq 0$ and $L^{n+1} = 0$ then $0 \neq L^n \subseteq Z(L)$.

□

Proposition 1.27. *L is nilpotent if and only if there is some $n > 0$ such that $adX_1 adX_2 ... adX_n = 0$ (as a map $L \to L$) for all $X_1, X_2, ..., X_n \in L$.*

Proof. L^n is spanned by elements of $(\mathrm{ad}X_1 \mathrm{ad}X_2 ... \mathrm{ad}X_n)(Y) = [X_1, [X_2, [X_3, ..., [X_n, Y]...]]]$ for $X_i, Y \in L$

□

We say that $X \in L$ is **ad-nilpotent** if $\mathrm{ad}X$ is a nilpotent linear transformation $L \to L$, i.e. $(\mathrm{ad}X)^n = 0$ for some n.

Corollary 1.28. *If L is nilpotent then every $X \in L$ is ad-nilpotent.*

Our first big theorem is Engel's theorem, which is the converse to this corollary.

Theorem 1.29 (Engel's Theorem)**.** *L is nilpotent if and only if every element $X \in L$ is ad-nilpotent.*

In other words, L is nilpotent if and only if the image $\mathrm{ad}L \subseteq gl(L)$ is a set of nilpotent transformations.

Now let's try to prove Engel's Theorem.

Lemma 1.30. *If $X \in gl(V)$ is nilpotent then $ad(x)$ is nilpotent (as an element of $gl(gl(V))$).*

Proof. Let $\lambda_X(Y) = XY$ and $\rho_X(Y) = YX$. Then λ_X and ρ_X are commuting nilpotent elements of $gl(gl(V))$. Since $\lambda_X\rho_X(Y) = \rho_X\lambda_X(Y) = XYX$. If $X^n = 0$ then $\rho_X^n = \lambda_X^n = 0$ so $(\text{ad}X)^{2n} = (\lambda_X - \rho_X)^2 n = \sum_k \binom{2n}{k} \lambda_X^k \rho_X^{2n-k} = 0$.

□

Theorem 1.31. *Suppose $L \subseteq gl(V)$ is a Lie subalgebra and $0 \neq \dim(V) < \infty$. Assume that every $X \in L$ is nilpotent (so $X^n = 0$ for some $n > 0$ depending on X). Then, there exists $0 \neq v \in V$ with $Xv = 0$ for all $X \in L$.*

Proof. Any nilpotent linear transformation has zero as an eigenvector with eigenvalue zero: take any nonzero column of $X^n \neq 0$ if $X^{n+1} = 0$. If $\dim(L) \leq 1$ then we can just take $v \in V$ to be any 0-eigenvector of some $0 \neq X \in L$. Suppose $\dim(L) > 1$ and let K be a maximal proper Lie subalgebra. By induction (with K and L/K replacing L and V), there is a vector $X \in L - K$ with $[Y, X] \in K$ for all $Y \in K$. This means that $K \subsetneq N_L(K)$ because $N_L(K) \ni X \notin K$. Since $K \subseteq L$ is a maximal proper subalgebra, we must have $L = N_L(K)$ so $K \subseteq L$ is actually an ideal. Since K is an ideal, the direct sum $K \oplus \mathbb{F}Z$ is a Lie subalgebra of L for any $Z \in L - K$. Therefore, we must have $L = K \oplus \mathbb{F}Z$ for any $Z \in L - K$ and $\dim(L) = \dim(K) + 1$. By induction on $\dim(L)$, the subspace $W = v \in V | Yv = 0 \forall Y \in K$ is nonzero and we have $LW \subseteq W$ since if $X \in L, Y \in K, w \in W$ then $YX_w = XY_w - [X, Y]w = 0$. Any $Z \in L - K$ acts as a nilpotent linear map $W \to W$ so have a 0-eigenvector $0 \neq v \in W$ with $Zv = 0$. This vector is then a 0-eigenvector for every element $X \in K \oplus \mathbb{F}Z = L$.

□

Now, we have a pretty simple of Engel's theorem.

Proof. Assume every $X \in L$ is ad-nilpotent. Then $\text{ad}L \subseteq gl(L)$ satisfies conditions of the previous theorem. So exists $0 \neq X \in L$ with $(\text{ad}Y)(X) = [Y, X] = 0 \forall Y \in L$ which means that $Z(L) \neq 0$. But now $L/Z(L)$ has smaller dimension with all elements still ad-nilpotent, so by induction $L/Z(L)$ is nilpotent. Hence, by the earlier lemma, L is also nilpotent.

□

Here are some interesting corollaries.

Corollary 1.32. *If $\dim(V) = n < \infty$ and $L \subseteq gl(V)$ is nilpotent then there exists a **flag** of vector spaces $0 = V_0 \subseteq V_1 \subseteq V_2 \subseteq ... \subseteq V_n = V$ such that $XV_i \subseteq V_{i-1}$ for all i and all $X \in L$. Equivalently, there exists a basis of V relative to which the matrices of all elements $X \in L$ are strictly upper-triangular.*

Proof. Set $V_1 = \mathbb{F}v$ where $0 \neq v \in V$ has $Lv = 0$. Then apply induction to image of L in $gl(V/V_1)$.

□

Corollary 1.33. *If L is nilpotent and $K \subseteq L$ is a nonzero ideal then $Z(L) \cap K \neq 0$.*

Proof. L acts on K by adjoint representation so theorem above implies that there exists $0 \neq X \in K$ with $(\text{ad}Y)(X) = [Y, X] = 0 \forall Y \in L$, i.e. X is an nonzero element of $Z(L) \cap K$.

□

1.9 Semisimple Lie Algebras

From now on: $\mathbb{F}$ is an algebraically closed field of characteristic zero.

L is always a Lie algebra over $\mathbb{F}$ of finite-dimension.

Theorem 1.34. *Assume $L \subseteq gl(V)$ is solvable, where $0 < \dim(V) < \infty$. Then there exists $0 \neq v \in V$ that is an eigenvector of every $X \in L$. That is, such that $Xv = \lambda(X)v \forall X \in L$ for some linear map $\lambda : X \to \mathbb{F}$.*

Proof. Mimic the proof of the theorem above. Here is the proof outline:

1. Find an ideal $K \subseteq L$ with $\dim(L) = \dim(K) + 1$. As L is solvable, $[L, L] \subsetneq L$, so the quotient $L/[L, L]$ is nonzero and abelian, so any subspace of this quotient is an ideal. Define K to be the preimage in L of any codim-1 subspace of $L/[L, L]$. Then K is a solvable ideal of L with $\dim(L) = \dim(K) + 1$.

2. By induction, V has a common eigenvector for K. Because $\mathbb{F}$ is algebraically closed, the theorem definitely holds if $\dim(L) \leq 1$. If we assume $\dim(L) \geq 2$, then by induction, there is a linear map $\lambda : K \to \mathbb{F}$ and $0 \neq v \in V$ with $Xv = \lambda(X)v$ for all $X \in K$. Define $W = \{w \in V | Xw = \lambda(X)w \text{ for all } X \in K\} \neq 0$.

3. Check that L stabilizes this common K-eigenspace $\mathbb{F}v$. We need to show L preserves W. Equivalently, we want to check that if $X \in L, Y \in K, w \in W$, then $YXw = \lambda(Y)Xw$. But all we know now is that $YXw = XYw - [X, Y]w = \lambda(Y)Xw - \lambda([X, Y])w$, so it suffices to check that $\lambda([X, Y]) = 0$ for all $X \in L, Y \in K$. Fix $w \in W$. Let $n > 0$ be minimal with $w, Xw, X^2w, ..., X^nw$ linearly dependent. Define $w_i = \mathbb{F}\text{span}\{w, Xw, X^2w, ..., X^{i-1}w\}$ so that $w_0 = 0$ and $\dim(W_i) = \min(i, n)$. Claim that $YX^iw \in \lambda(Y)X^iw + W_i$. This is clear if $i = 0$, and if $i > 0$, then $YX^iw = YXX^{i-1}w = XYX^{i-1}w - [X, Y]X^{i-1}w \in \lambda(Y)X^ix + W_i$. Conclude that relative to the basis $w, Xw, X^2w, ..., X^{i-1}w$, the matrix of Y is

$$\begin{bmatrix} \lambda(Y) & & * \\ & \ddots & \\ 0 & & \lambda(Y) \end{bmatrix}$$

and thus $\mathrm{tr}_{W_n}(Y) = n\lambda(Y)$ but also $\mathrm{tr}_{W_n}([X,Y]) = n\lambda([X,Y])$. Since X and Y preserve W_n, the products XY and YX also preserve W_n, so $\mathrm{tr}_{W_n}([X,Y]) = 0$.

4. Show that some $Z \in L - K$ has an eigenvector in $\mathbb{F}v$ (noting that $L = K \oplus \mathbb{F}Z$) as then we can conclude that v is an eigenvector of every $X \in L$. Write $L = K \oplus \mathbb{F}Z$ for some $Z \in L - K$. Since $\mathbb{F}$ is algebraically closed, Z has an eigenvector $0 \neq v_0 \in W$. But v_0 is then an eigenvector for every $X \in L = K \oplus \mathbb{F}Z$.

□

Now we state Lie's theorem.

Theorem 1.35 (Lie's Theorem)**.** *Suppose $L \subseteq gl(V)$ is a solvable Lie algebra where $\dim(V) = n < \infty$. Then there is some basis of V relative to which the matrices of all elements of L are upper triangular.*

Proof. Choose $0 \neq v_1 \in V$ with $Xv_1 = \lambda(X)v_1 \forall X \in L$ for some linear map $\lambda : L \to \mathbb{F}$. Set $V_1 = \mathbb{F}v_1$ and apply the theorem by induction to V/V_1. This gives a basis $v_2 + V_1, v_3 + V_1, ..., v_n + V_1$ for V/V_1 and the desired basis is then $v_1, v_2, ..., v_n$.

□

Two corollaries:

Corollary 1.36. *Suppose L is solvable with $n = \dim(L) (< \infty)$. There there exists a chain of ideals*

$$0 = L_0 \subset L_1 \subset L_2 \subset ... \subset L_n = L$$

with $\dim(L_i) = i$.

Proof. Apply Lie's theorem to Lie algebra $\mathrm{ad}(L) \subseteq gl(L)$ to get a basis $v_1, v_2, ..., v_n \in L$ such that $(\mathrm{ad}X)v_i \in \mathbb{F} - \mathrm{span}[v_1, ..., v_i] \forall X \in L \forall i$ which means that $L_i := \mathbb{F} - \mathrm{span}[v_1, ..., v_i]$ is an ideal.

□

Corollary 1.37. *Suppose L is solvable. Then $[L,L]$ is nilpotent.*

Proof. Choose a basis of L such that the matrices of $\mathrm{ad}X \in gl(L)$ for every $X \in L$ are upper-triangular. The matrix of $\mathrm{ad}[X,Y] = [\mathrm{ad}X, \mathrm{ad}Y]$ is then strictly upper triangular $\forall X, Y \in L$, so $\mathrm{ad}Z$ is nilpotent for all $Z \in [L,L]$. Engel's theorem then implies that $[L,L]$ is nilpotent.

□

1.10 Jordan-Chevalley Decomposition

Let V be a finite-dimensional vector space. We say that $X \in gl(V)$ is **semisimple** if X is diagonalizable (meaning V has a basis of eigenvectors for X). When $\mathbb{F}$ is not algebraically closed, semieimple means that the roots of the minimal polynomial of X are distinct.

Some quick facts that are quick to check.

Proposition 1.38. *1. If X and Y commute and are both semisimple, then all linear combinations $aX + bY$ $(a, b \in \mathbb{F})$ are semisimple.*

2. If X is semisimple and X preserves $W \subseteq V$, then $X|W$ is semisimple.

Proposition 1.39. *Consider some element $X \in gl(V)$.*

1. *There are unique elements $X_s, X_n \in gl(V)$ with X_s semisimple, X_n nilpotent, $X_sX_n = X_nX_s$ and $X_s + X_n = X$. Call the last part the Jordan-Chevalley decomposition of X.*
2. *There are polynomials $p(T), q(T)$ in a variable T with no constant term (so $p(T), q(T) \in T\mathbb{F}[T]$) such that $X_s = p(X)$ and $X_n = q(X) \implies X_s$ and X_n commute with any $Y \in gl(V)$ that has $XY = YX$.*
3. *If $A \subseteq B \subseteq V$ are subspaces with $XB \subseteq A$ then $X_sB \subseteq A$ and $X_nB \subseteq A$*

We will omit the proof since it's just standard linear algebra. The idea for the first part is that if $V = \mathbb{F}^n$ and the Jordan canonical form of the matrix of X has blocks

$$\begin{bmatrix} a & 1 & 0 & 0 \\ 0 & a & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & a \end{bmatrix}$$

then the Jordan canonical forms of X_s and X_n are

$$\begin{bmatrix} a & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & a \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

respectively.

One proposition that uses this fact is the following:

Proposition 1.40. *Suppose V has a $\dim(V) < \infty$ and $X \in gl(V)$.*

1. *If X is nilpotent then so is $adX \in gl(gl(V))$.*

2. *If X is semisimple then so is adX.*

Proof. If $v_1, v_2, ..., v_n$ is a basis of V consisting of eigenvectors for X, so $Xv_i = a_i v_i$ for some $a_i \in \mathbb{F}$, and e_{ij} is the corresponding basis of $gl(V)$ so that e_{ij} is the linear map with

$$v_k \mapsto \begin{cases} v_i & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

then $(\text{ad}X)(e_{ij}) = (a_i - a_j)e_{ij}$ since $(\text{ad}X)(e_{ij})(v_k) = [X, e_{ij}](v_k) = Xe_{ij}(v_k) - e_{ij}Xv_k = (a_i - a_j)e_{ij}v_k$.

□

One more lemma to mention before we finish.

Lemma 1.41. *Let $X \in gl(V)$, $\dim(V) < \infty$ and the Jordan decomposition of X is $X = X_s + X_n$, then the Jordan decomposition of adX is*

$$adX = adX_s + adX_n$$

Proof. $[\text{ad}X_s, \text{ad}X_n] = \text{ad}[X_s, X_n] = 0$ and we already saw that $\text{ad}X_s$ is semisimple and $\text{ad}X_n$ is nilpotent.

□

Next time: criteria for solvability and semisimplicity.

1.11 Cartan criterion, Killing form, Semisimplicity

1.12 Cartan's Criterion

Lemma 1.42. *Let $A \subseteq B$ be two subspaces of $gl(V)$, where $\dim(V) < \infty$. Define*

$$M = \{X \in gl(V) | [X, B] \subseteq A\}$$

and suppose $X \in M$ has $tr(XY) = 0 \forall Y \in M$. Then X is nilpotent.

Proof. Write the Jordan decomposition of X as $X = X_s + X_n$. Let $v_1, ..., v_n$ be a basis for V such that $X_s v_i = a_i v_i$ for some $a_i \in \mathbb{F}$. Define $E = \mathbb{Q}$-span$\{a_1, ..., a_n\} \subset \mathbb{F}$ to be a rational vector space. We want to show that $X_s = 0$ as then $X = X_n$ is nilpotent. This holds if and only if $a_1 = a_2 = ... = a_n = 0 \iff E = 0$. Let $E^* = \{\mathbb{Q}\text{-linear maps } E \to \mathbb{Q}\}$. Since $\dim_{\mathbb{Q}}(E) \leq n \leq \infty$, it holds that $\dim_{\mathbb{Q}}(E^*) = \dim_{\mathbb{Q}}(E)$ so it suffices to show that $E^* = 0$.

Suppose that $f \in E^*$. Let $Y \in gl(V)$ have the matrix

$$\begin{bmatrix} f(a_1) & & & \\ & f(a_2) & & \\ & & \ddots & \\ & & & f(a_n) \end{bmatrix}$$

in basis $v_1, \ldots, v_n$ for V. Then $Yv_i = f(a_i)v_i$ for all i. If $e_{ij} \in gl(V)$: $v_k \mapsto \begin{cases} v_i & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}$, then $(\text{ad}(X_s))e_{ij} = (a_i - a_j)e_{ij}$ and $(\text{ad}Y)e_{ij} = (f(a_i) - f(a_j))e_{ij}$.

Claim: We can find a polynomial $r(T) \in \mathbb{Q}[T]$ such that $r(a_i - a_j) = f(a_i) - f(a_j) = f(a_i - a_j)$ for all i, j. (The polynomial $r(T)$ passes through the points $(x, y) = (a_i - a_j, f(a_i - a_j))$ for all i, j. This is clear by polynomial interpolation.)

We then have $r(0) = f(0) = f(a_i - a_i) = 0$, so $r(T)$ has no constant term, $r(T) \in T\mathbb{Q}[T]$. Also, we have $\text{ad}(Y) = r(\text{ad}(X_s))$ since both sides give the same result applied to each e_{ij}. Finally, by the earlier factors, $\text{ad}(X_s)$ is the semisimple part of $\text{ad}X$, so $\text{ad}(X_s) = p(\text{ad}X)$ for some polynomial without a constant term.

We assume $(\text{ad}(X))(B) \subseteq A$, so $(\text{ad}(Y))(B) = r(\text{ad}(X_s))(B) = r(p(\text{ad}X))(B) \subseteq A$. This implies that $Y \in M = \{z \in gl(V) | [Z, B] \subseteq A\}$. Hence, by assumption, $\text{tr}(XY) = 0$. But $\text{tr}(XY) = \sum_{i=1}^{\infty} a_i f(a_i) \in E = \mathbb{Q}\text{-span}[a_1, \ldots, a_n]$. Thus, $0 = f(0) = f(\text{tr}(XY)) = \sum_{i=1}^{n} f(a_i)^2$, which can only hold if $f = 0$. Thus, $E^* = 0$.

□

Proposition 1.43. *If $X, Y, Z \in gl(V)$, $\dim(V) < \infty$, then*

$$tr([X, Y]Z) = tr(X[Y, Z])$$

Proof.

$$\text{tr}(XYZ) - \text{tr}(YXZ) = \text{tr}(XYZ) - \text{tr}(XZY)$$

□

Proposition 1.44 (Cartan's Criterion)**.** *Let $L \subseteq gl(V)$ where $\dim(V) < \infty$. If $tr(XY) = 0$ for all $X \in [L, L], Y \in L$, then L is solvable.*

Proof. To show that L is solvable, it is enough to check that $[L, L]$ is nilpotent, and for this, it suffices to check that every $X \in [L, L]$ is nilpotent as an element of $gl(V)$. This will imply that $\text{ad}X$ is nilpotent, so we can use Engel's theorem.

Let $A = [L, L] \subseteq B = L$ and $M = \{X \in gl(V) | [X, L] \subseteq [L, L]\}$. Suppose $X \in [L, L]$ and note that $[L, L] \subseteq M$. If we can show that $\mathrm{tr}[XY] = 0$ for all $Y \in M$, then the lemma implies that it is nilpotent. We assume $\mathrm{tr}(XY) = 0$ for all $Y \in L$. Note that $Y \subseteq M$. But X is a linear combination of elements $[X_1, X_2]$ for $X_i \in L$, so if $Y \in M$ then

$$\mathrm{tr}([X_1, X_2]Y) = \mathrm{tr}(X, [X_2, Y]) = \mathrm{tr}([X_2, Y]X_1) = 0.$$

□

Corollary 1.45. *If L is a Lie algebra such that $tr(adX\, adY) = 0$ for all $X \in [L, L], Y \in L$, then L is solvable.*

Proof. Apply Cartan's criterion to $\mathrm{ad}L \subseteq gl(L)$. As $\mathrm{ad}[L, L] = [\mathrm{ad}L, \mathrm{ad}L]$, we find that $\mathrm{ad}L$ is solvable. But $\ker \mathrm{ad} = Z(L)$ is solvable, so L is solvable since $L/Z(L) \cong \mathrm{ad}L$.

□

1.13 The Killing Form

Let L be a Lie algebra.

Definition 1.46. *The bilinear form $\kappa : L \times L \to L$ defined by*

$$\kappa(X, Y) := tr(adX\, adY)$$

*for $X, Y \in L$ is called the **Killing form** of L.*

Proposition 1.47. *κ is symmetric and associative.*

To compute $\kappa(X, Y)$ need to pick a basis of L and write down the matrices of $\mathrm{ad}X$ and $\mathrm{ad}Y$, doesn't matter which basis you choose.

Lemma 1.48. *Let $I \subseteq L$ be an ideal. Then the Killing form κ_I of I is the restriction of the Killing form $\kappa = \kappa_L$ of L. Thus $\kappa_I(X, Y) = \kappa_L(X, Y) \forall X, Y \in I$.*

Proof. If $\phi : V \to W \subseteq V$ is a linear map then $\mathrm{tr}_V(\phi) = \mathrm{tr}_W(\phi|_W)$ because if $w_1, ..., w_k$ is a basis of W and $w_{k+1}, ..., w_n$ extends this to a basis of V then the matrix of ϕ is 0 for the bottom $n-k$ rows and the $k \times k$ entries in the top left is a matrix of $\phi|_W$. To prove lemma, apply this observation with $V = L, W = I$.

□

Definition 1.49. *The **radical** of any symmetric billinear form $\kappa : L \times L \to L$ is*

$$S = \{X \in L | \kappa(X,Y) = 0 \forall Y \in L\} = \{Y \in L | \kappa(X,Y) = 0 \forall X \in L\}$$

which is $\supset Z(L)$ *since* $adX = 0 \forall X \in Z(L)$.

Definition 1.50. *The form* κ *is* ***non-degenerate*** *if* $S = 0$. *This happens iff*

1. $\kappa(X, \cdot) : L \to L$ *is zero map if and only if* $X = 0$, *or*
2. *The* $n \times n$ *matrix* $[\kappa(X_i, X_j)]_{1 \le i,j \le n}$ *is invertible for same/any basis* $x_1 x_2 ... x_n \in L$.

Example 1.51. *Suppose* $L = \mathfrak{sl}_2(\mathbb{F})$. *Then*

$$adE = \begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, adH = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}, adE = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}$$

so we have $[\kappa(x_i, x_j)]_{1 \le i,j \le 3} = \begin{bmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{bmatrix}$ *where* $x_1 = E, x_2 = H, x_3 = F$, *which is invertible assuming* $char(\mathbb{F}) \neq 2$.

Theorem 1.52. *L is semisimple if and only if the killing form* κ *is nondegenerate.*

Thus, to check if L is semisimple, just need to pick a basis $x_1, ..., x_n$ for L and check if the matrix $[\kappa(x_i, x_j)]_{1 \le i,j \le n}$ nas nonzero determinant.

Proof. Here is the proof that $\mathrm{Rad}L = 0$ if and only if radical S of κ is zero. The radical S of Killing form is an ideal of L. In fact, $\mathrm{ad}_L S$ is a solvable ideal of $\mathrm{ad}L$ by Cartan's criterion:

$$\mathrm{tr}(\mathrm{ad}X \mathrm{ad}Y) = \kappa(X,Y) = 0 \forall X \in S \supset [S,S], \forall Y \in L \supset S$$

The center $Z(S)$ is abelian, hence solvable. As $\mathrm{ad}_L S \cong S/Z(S)$ we conclude that S is solvable. Thus if $\mathrm{rad}L = 0$ then $S = 0$ as $S \subset \mathrm{Rad}L$.

Suppose conversely that $S = 0$. Want now to show that $S = 0$ implies that $\mathrm{Rad}L = 0$. It suffices to check that $I \subset L$ is any abelian ideal then $I \subset S$. This is because if I a nonzero solvable ideal, then $I^{(n)}$ is abelian for n such that $I^{(n)} \neq 0 = I^{(n+1)}$ (so if there are no nonzero abelian ideals, there are also no nonzero solvable ideals). Assume I Is an abelian ideal. If $X \in I, Y \in L$ then $\mathrm{ad}X \mathrm{ad}Y$ is a map $L \to L \to I$ so $(\mathrm{ad}X \mathrm{ad}Y)^2$ is a map $L \to [I, I] = 0$. Thus $\mathrm{ad}X \mathrm{ad}Y$ is nilpotent so it must have zero trace, meaning that $\kappa(X,Y) = 0 \implies I$ abelian $\subset S$.

□

1.14 Simple and Semisimple Ideals

A Lie algebra L is a **direct sum** of ideals $I_1, ..., I_n$ if $L = I_1 \oplus ... \oplus I_n$. This would mean that each $X \in L$ has a unique expression as $x = x_1 + ... + x_n$ with $x_j \in I_j$. Uniqueness here implies that $I_j \cap I_k = 0 \forall j \neq k$. Since each I_j is an ideal, we must also have $[I_j, I_k] = 0 \forall j \neq k$.

Theorem 1.53. *Suppose L is a semisimple Lie algebra. There there exists ideals $L_1, ..., L_n \subseteq L$ such that*

1. *Each L_i is simple.*
2. *$L = L_1 \oplus ... \oplus L_n$.*
3. *Any simple ideal of L is equal to some L_i.*
4. *The Killing form of L_i is just the restriction of the Killing form of L.*

Proof. Let I be any ideal of L and define $I^\perp = \{x \in L | \kappa(X, Y) = 0 \forall Y \in I\}$.

First claim is to show is that $I^\perp$ is an ideal and $L = I \oplus I^\perp$.

1. To see that $I^\perp$ is an ideal, let $X \in L, Y \in I^\perp, Z \in I$. Then $\kappa([X, Y], Z) = -\kappa([Y, X], Z) = -\kappa(Y, [X, Z]) = 0$. So we conclude that $[X, Y] \in I^\perp$.
2. Because L is semisimple, its center $Z(L)$ is zero, so ad is injective. Cartan's criterion applied to $I \cap I^\perp \cong \mathrm{ad} I \cap I^\perp \subseteq gl(L)$ implies that $I \cap I^\perp$ is solvable: $\forall X \in [I \cap I^\perp, I \cap I^\perp], \forall Y \in I \cap I^perp$, we have $\mathrm{tr}(\mathrm{ad}X \mathrm{ad}Y) = \kappa(X, Y) = 0$. Thus $I \cap I^\perp = 0$ as $\mathrm{Rad}L = 0$, and $L = I \oplus I^\perp$.

So our claims both hold. Now proceed by induction on $\dim L$. If L has no nonzero proper ideals then L is simple. Otherwise we can find a minimal proper nonzero ideal $L_1 \subset L$. Any ideal of L_1 is an ideal of $L = L_1 \oplus L_1^\perp$ so L_1 must be simple itself. Likewise, $L_1^\perp$ must be asemisimple since any of its solvable ideals are solvable ideals of L. Thus by induction we can write $L_1^\perp = L_2 \oplus ... \oplus L_n$ for simple ideals L_i and then $L = L_1 \oplus ... \oplus L_n$. This proves parts 1 and 2, 4 is already known by the lemma.

We still have to prove that if I is any simple ideal of L then $I = L_i$ for some $i \in \{1, 2, ..., n\}$. To prove this, we observe that $[I, L] = \mathrm{span}([X, Y] | X \in I, Y \in L)$ is also an ideal of L since if $X \in L, Y \in I, Z \in L$ then

$$[X, [Y, Z]] = \mathrm{ad}X[Y, Z] = [[X, Y], [X, Z]] \in [I, L]$$

If $[I, L] = 0$ then $I \subseteq Z(L) = 0$. As $I \neq 0$ is simple, we must have $I = [I, L]$. But $[I, L] = \bigoplus_j [I, L_j] = I$ means that $[I, L_j]$ must be zero for all but one j and $I = L_j$.

∎

Our original definition of semisimple was the property of having no nonzero sovlable ideals. Now we have a more intuitive characterization.

Corollary 1.54. *L is semisimple if and only if L is a direct sum of simple Lie algebras.*

Proof. Only if direction: Previous theorem.

If direction: If $L = L_1 \oplus ... \oplus L_n$ with L_i simple then the radical of the Killing form κ of L is $\bigoplus_{i=1}^{n} \mathrm{Rad}(\kappa|_{L_i \times L_i})$ since $L_i^{\perp} = \bigoplus_{j \neq i} L_j$. But each simple L_i is semisimple so $\mathrm{Rad}(\kappa|_{L_i \times K_i}) = 0$.

□

Corollary 1.55. *If L is semisimple then $L = [L, L]$ and all ideals and homomorphic images of L are also semisimple.*

Proof. If $L = \bigoplus_i L_i$, each L_i simple, then $[L_i, L_i] = L_i \forall i$ and $[L_i, L_j] = 0 \forall i \neq j$ so $[L, L] = \bigoplus_{i,j} [L_i, L_j] = \bigoplus_i L_i = L$. If $I \subseteq L$ is an ideal then I is semisimple as any of its solvable ideals are also ideals of L. Final claim about homomorphic images is left as an exercise.

□

1.15 Modules, Casimir Element, Weyl's Theorem, and Abstract Jordan Decomposition

1.16 Goals

1. Basic concepts in the representation theory of Lie algebras
2. Discuss Casimir element as a tool for proving Weyl's theorem: finite dimensional representations of a semisimple L are "completely reducible"
3. Some consequences for Jordan decomposition

1.17 Representation Theory

Terminology: Throughout, L is a semisimple Lie algebra.

Definition 1.56. *An L-**representation** is a Lie algebra morphism $\phi : L \to gl(V)$ for some vector space V.*

Definition 1.57. *An L-**module** is a vector space V with a bilinear operation $L \times V \to L$ such that $(X, v) \mapsto X \cdot v$ with $[X, Y] \cdot v = X \cdot (Y \cdot v) - y \cdot (X \cdot v) \forall X, Y \in L, v \in V$.*

L-representations and L-modules are "equivalent" notions, just different syntax. Equivalently in this sense:

Proposition 1.58. *If $\phi : L \to gl(V)$ is an L-rep then V is an L-module for the action $X \cdot v := \phi(x)(v)$ for $X \in L, v \in V$.*

Proof. The action is bilinear and we have

$$X \cdot (Y \cdot v) - Y \cdot (X \cdot v) = \phi(X)(\phi(Y)(v)) - \phi(Y)(\phi(X)(v))$$
$$= (\phi(X)\phi(Y) - \phi(Y)\phi(X))(v) \qquad = [\phi(X), \phi(Y)](v) = \phi([X, Y])(v) = [X, Y] \cdot v$$

$\forall X, Y \in L, v \in V$.

□

Proposition 1.59. *If v is an L-module then the map $\phi : L \to gl(V)$ is defined by $\phi(x) : v \mapsto X \cdot v$ for $v \in V$ is an L-rep.*

Proof. Similar straightforward algebra.

□

Suppose V is an L-module.

Definition 1.60. *A **submodule** of V is a subspace $U \subseteq V$ such that $X \cdot u \in U \forall X \in L, \forall u \in U$.*

Definition 1.61. *A **morphism** of two L-modules V and W is a linear map $f : V \to W$ such that*

$$f(X \cdot v) = X \cdot f(v) \forall v \in V, X \in L.$$

Definition 1.62. *The **kernel** of a morphism $f : V \to W$ is the submodule $Ker(f) := \{v \in V | f(v) = 0\}$.*

Definition 1.63. *If an L-module morphism $f : V \to W$ is a bijection then f is a **isomorphism**.*

Definition 1.64. *An L-module V is **irreducible** if its only submodules are 0 and $V \neq 0$ (meaning V has exactly two submodules).*

Zero modules are not considered irreducible because we want a unique direct sum decomposition into irreducible submodules.

Definition 1.65. *V is **completely reducible** if there are irreducible submodules $V_i \subseteq V$ such that $V = \bigoplus_i V_i$.*

Here $\oplus$ refers to the obvious notion of direct sum for L-modules.

The fundamental result (stale without proof) is Schur's lemma.

Lemma 1.66 (Schur's Lemma)**.** *Suppose $\phi : L \to \text{gl}(V)$ is an irreducible L-repn (meaning that the associated L-module structure on V is irreducible). Then the only linear maps $f : V \to V$ with $f \circ \phi(X) = \phi(X) \circ f$ for all $X \in L$ are the scalar maps $f_c : V \to V$, $v \mapsto cv$ for fixed $c \in \mathbb{F}$.*

Remark 1.67. *We require $\mathbb{F}$ to be algebraically closed with characteristic zero.*

Now let's talk about the dual/contragradient of an L-module. Suppose V is an L-module. Define $V^* = \{\text{linear maps } V \to \mathbb{F}\}$.

Proposition 1.68. *V^* is an L-module for the action*

$$X \cdot f = \{\textit{the linear map } V \to \mathbb{F} \textit{ sending } v \mapsto f(X \cdot v)\}$$

for $f \in V^*$.

Proof. For $X, Y \in L$, $f \in V^*$, $v \in V$, we have

$$\begin{aligned}
([X,Y] \cdot f)(v) &= -f([X,Y] \cdot v) \\
&= -f(X \cdot Y \cdot v - Y \cdot X \cdot v) \\
&= -f(X \cdot Y \cdot v) + f(Y \cdot X \cdot v) \\
&= (X \cdot f)(Y \cdot v) - (Y \cdot f)(X \cdot v) \\
&= -(Y \cdot X \cdot f)(v) + (X \cdot Y \cdot f)(v) \\
&= (X \cdot Y \cdot f - Y \cdot X \cdot f)(v)
\end{aligned}$$

□

Now, let's talk about tensor products of L-modules. Suppose V and W are L-modules, say with bases $[v_i]_{i \in I}$ and $[w_j]_{j \in J}$.

Definition 1.69. *The **tensor product** $V \otimes W$ is the vector space spanned by all tensors $v \otimes w$ ($v \in V, w \in W$), where $(v + v') \otimes w = v \otimes w + v' \otimes w$ and $v \otimes (w + w') = v \otimes w + v \otimes w'$ and $(av) \otimes w = v \otimes (aw)$ for $a \in \mathbb{F}$, $v, v' \in V$, $w, w' \in W$, with basis $\{v_i \otimes w_j\}_{(i,j) \in I \times J}$.*

Proposition 1.70. *$V \otimes W$ is an L-module for the action*

$$X \cdot (v \otimes w) = (X \cdot v) \otimes w + v \otimes (X \cdot w)$$

for $X \in L$, $v \in V$, $w \in W$.

Proof.

$$\begin{aligned}
[X,Y] \cdot (v \otimes w) &= ([X,Y] \cdot v) \otimes w + v \otimes ([X,Y] \cdot w) \\
&= X \cdot (Y \cdot v) \otimes w - Y \cdot (X \cdot v) \otimes w + v \otimes X \cdot (Y \cdot w) - v \otimes Y \cdot (X \cdot w)
\end{aligned}$$

But we also have

$$X \cdot Y \cdot (v \otimes w) - Y \cdot (X \cdot (v \otimes w)) = X \cdot (Yv \otimes w + v \otimes Yw) - Y \cdot (Xv \otimes w + v \otimes Xw)$$

which are equal on the right-hand side, and after some cancellation, this is the same as above.

□

Proposition 1.71. *The linear map $V^* \otimes V \to \mathrm{gl}(V)$, $f \otimes w \mapsto \{$the linear map $V \to V$ sending $v \to f(v)w\}$, is an isomorphism of vector spaces if $\dim(V) < \infty$. The L-module structure on $\mathrm{gl}(V)$ making this map a module isomorphism is $X \cdot f = \{$the linear map $V \to V$ sending $v \to X \cdot f(v) - f(X \cdot v)\}$.*

1.18 Casimir Element

Definition 1.72. *An L-rep $\phi : L \to gl(V)$ is* ***faithful*** *if $Ker(\phi) = 0$, meaning ϕ is injective.*

Assume $\phi : L \to gl(V)$ is a faithful L-rep. Define $\beta : L \times L \to \mathbb{F}, (X, Y) \mapsto \beta(X, Y) := \mathrm{tr}_V(\phi(X)\phi(Y))$. This bilinear form is symmetric and associative. The Killing form of L is β for $\phi = \mathrm{ad} : L \to gl(L)$.

Definition 1.73. *The* ***radical*** *of β is the* ***ideal*** *$S = \{X \in L | \beta(X, Y) = 0 \forall Y \in L\}$.*

In fact, S is a solvable ideal of L, since $S \cong \phi(S)$ (by faithfulness of ϕ) and Cartan's criterion for $\phi(S)$ ($\mathrm{tr}(XY) = 0 \forall X \in \phi([S, S] \forall Y \in \phi(S)))$.

We are assuming that L is semisimple, so we can conclude that

Proposition 1.74. *The form*

$$\beta(X, Y) = tr(\phi(X)\phi(Y))$$

is nondegenerate (ie. $S = 0$)

Conversely, assume $\beta : L \times L \to \mathbb{F}$ is any symmetric, associative, nondegenerate billinear form. Choose a basis $X_1, ..., X_n$ for L and define $Y_1, ..., Y_n$ to the (unique) dual basis with

$$\beta(X_i, Y_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Fix $Z \in L$ and define $a_{ij}, b_{ij} \in \mathbb{F}$ such that $[Z, X_i] = \sum_j a_{ij} X_j$ and $[Z, Y_i] = \sum_j b_{ij} Y_j$.

Lemma 1.75.

$$a_{ik} = -b_{ki} \forall i, k$$

Proof.

$$\begin{aligned} a_{ik} &= \sum_j \beta(X_j, X_k) a_{ij} \\ &= \beta([Z, X_i], Y_k) \\ &= \beta(X_i, -[Z, Y_k]) \\ &= \sum_j \beta(X_i, Y_j) b_{kj} \\ &= -b_{ki} \end{aligned}$$

□

Suppose $\phi : L \to gl(V)$ is an L-rep. Define $c_\phi(\beta) := \phi(X_i)\phi(Y_i) \in gl(V)$ where $\phi(X_i), \phi(Y_i)$ are dual bases defined with respect to β. This definition looks like it depends on the choice of these bases, but this is in fact not true.

Proposition 1.76. *Then* $[\phi(Z), c_\phi(\beta)] = 0 \forall Z \in L$. *So (multiplication by)* $c_\phi(\beta)$ *is a linear map* $V \to V$ *that commutes with the action of* L *via* ϕ.

Proof.

$$[\phi(Z), c_\phi(\beta)] = \sum_i [\phi(Z), \phi(X_i)]\phi(Y_i) + Z_i \phi(X_i)[\phi(Z), \phi(Y_i)] = \sum_{i,j} (a_{ij} + b_{ij})\phi(X_i)\phi(Y_j) = 0$$

□

Definition 1.77. *When* $\phi : L \to gl(V)$ *is a faithful* L*-rep we define the* ***Casimir element*** *to be* $c_\phi = c_\phi(\beta) \in gl(V)$ *for the form* $\beta(X, Y) = tr(\phi(X)\phi(Y))$.

This makes sense because we already checked that this form is nondegenerate and associative.

Two key facts:

Proposition 1.78.

$$tr(c_\phi) = \dim(L)$$

Proof.

$$\mathrm{tr}(c_\phi) = \sum_i \mathrm{tr}(\phi(X_i)\phi(Y_i)) = Z_i \beta(X_i, Y_i) \dim(L)$$

□

Proposition 1.79. *If* ϕ *is irreducible then* $c_\phi = \frac{\dim(L)}{\dim(V)} \in \mathbb{F} \in gl(V)$

Proof. If ϕ is irreducible then Schur's lemma implies that c_ϕ is a scalar since $[\phi(X), c_\phi] = 0 \forall x \in L$.

□

Example 1.80. *Let* $L = \mathfrak{sl}_2$ *with basis* $X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, Y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. *Suppose* $V = \mathbb{F}^2$ *and* $\phi : L \to gl(V)$ *is the identity map. The basis dual to* X, H, Y *in trace form is* $Y, \frac{1}{2}H, X$ *(*$\beta(A,B) = tr(AB)$ *since* $\phi = id$*), so* $XY + \frac{1}{2}H^2 + YX = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{3}{2} \end{bmatrix}$

Definition 1.81. *When* L *is semisimple but* $\phi : L \to gl(V)$ *is not faithful, we define* $c_\phi \in gl(V)$ *to be the* ***Casimir element*** *of the faithful rep* $\phi : L/Ker(\phi) \to gl(V)$.

Lemma 1.82. *Let* $\phi : L \to gl(V)$ *be an* L*-rep with* L *semisimple. Then* $\phi(L) \subseteq sl(V) \subseteq gl(V)$. *Thus if* $\dim(V) = 1$, *then* $\phi(L) = 0$ *as* $sl(V) = 0$.

Proof. We have $L = [L, L]$ by semisimplicity so

$$\phi(L) = \phi([L, L]) = [\phi(L), \phi(L)] \subseteq [gl(V), gl(V)] = sl(V).$$

□

Now, we can finally state Weyl's theorem.

1.19 Weyl's Theorem

Theorem 1.83. *Suppose* $\phi : L \to gl(V)$ *is an* L*-rep. As usual, we assume* L *is semisimple and* $\dim(V) < \infty$. *Then* ϕ *is* ***completely reducible****, meaning that there are irreducible* L*-submodules* $V_1, V_2, ..., V_n \subseteq V$ *such that*

$$V = V_1 \oplus V_2 \oplus ... \oplus V_n.$$

1.20 Proof of Weyl's Theorem

By replacing L to $L/\text{Ker}(\phi)$, we may assume that ϕ is faithful. Suppose $W \subseteq V$ is a proper L-submodule. By induction on dimension, we just need to show that there is a complementary L-submodule U with $V = W \oplus U$. (We can find a subspace such that direct sum holds as vector spaces, but the hard part is to find a submodule.)

Three steps for the proof.

Part 1. Assume W is irreducible and $\dim(V) = \dim(W) + 1$ so $V = W \oplus \mathbb{F}$ as vector spaces. Let $c = c_\phi$ be the Casimir element. Then $v \mapsto cv$ is an L-module endomorphism as c commutes with $\phi(L)$. Thus $cW \subseteq W$ (because W is a submodule and $c \in \phi(L)$) and $\text{Ker}(c) := \{v \in V | cv = 0\}$ is an L-submodule (because c commutes with $\phi(L)$). L acts trivially on $V/W = \mathbb{F}$ because all 1-dim representations of semisimple Lie algebras are trivial. This means $c(V/W) = 0 \implies cV \subseteq W$. Therefore $\dim(\text{Ker}(c)) \geq 1$, since otherwise $cV = V$. But c acts on W as a scalar by Schur's lemma, and this scalar cannot be zero since $\text{tr}_W(c) = \text{tr}_V(c) = \dim(L) \neq 0$. Therefore $\text{Ker}(c) \cap W = 0$ since c acts as a nonzero scalar on W. As $\dim(\text{Ker}(c)) + \dim(W) \geq \dim(V)$ we conclude that $V = W \oplus \text{Ker}(c)$.

Part 2. Suppose $V = W \oplus \mathbb{F}$ as vector spaces (so $\dim(V) = \dim(W+1)$) but W is not irreducible as an L-module. Then there is a nonzero submodule $W' \subset W$ and by induction $V/W' = W/W' \oplus \widetilde{U}$ for some L-submodule $\widetilde{U} \subset V/W'$. Define U to be the preimage of $\widetilde{U}$ under the quotient map $V \to V/W'$. Then U is an L-submodule containing W' (so $\widetilde{U} = U/W'$). By induction, $U = W' \oplus W''$ for some L-module W'' (as L-modules) and then $V = W \oplus W''$ (as L-modules) since $\dim(W) + \dim(W'') = \dim(V)$ and $W \cap W'' = 0$ (as $W \cap W'' \subseteq W'$ but $W' \cap W'' = 0$).

Part 3. Finally, suppose W is arbitrary and $\dim(V/W) \geq 1$. Let $\text{Hom}(V, W)$ be the L-module of linear maps $f : V \to W$ with L acting as $X \cdot f : v \mapsto X \cdot f(v) - f(X \cdot v)$. Let $A = \{f \in \text{Hom}(V, W) | f|_w \text{ is a scalar map}\}$ and $B = \{f \in \text{Hom}(V, W) | f|_w = 0\}$. Then A and B are L-submodules with $L \cdot A \subseteq B$ since if $f(w) = aW \forall w \in W$ where $a \in \mathbb{F}$ then

$$(X \cdot f)(w) = X \cdot f(w) - f(X \cdot w) = aX \cdot a - aX \cdot w = 0.$$

But $\dim(A) = \dim(B) + 1$ since if $f, g \in A$ then $af + bg \in B$ for some $a, b \in \mathbb{F}$. Therefore $A = B \oplus C$ for some L-submodule C with $\dim(C) = 1$. Suppose $C = \mathbb{F}\text{-span}\{h\}$ for some $h : V \to W$. We may assume $h|_W = \text{id}$ (after rescaling). The main claim is as follows:

Lemma 1.84. *$Ker(h) = \{v \in V | h(v) = 0\}$ is an L-submodule and $V = W \oplus Ker(h)$ as L-modules.*

Proof. If $h(v) = 0$ and $X \in L$ then $h(X \cdot v) = -(X \cdot h)(v) + X \cdot h(v) = 0$. Now observe $V = \text{Im }(h) \oplus \text{Ker}(h) = W \oplus \text{Ker}(h)$.

□

This finishes the proof.

1.21 Abstract Jordan Decomposition

Assume L is semisimple. Then $Z(L) = 0$ so the adjoint rep $\mathrm{ad} : L \to gl(L)$ is faithful.

Definition 1.85. *The* ***abstract Jordan decomposition*** *of $X \in L$ to be $X = X_s + X_n$ where $X_s, X_n \in L$ are the unique elements with $ad(X_s) = (ad(X))_s$ and $ad(X_n) = (ad(X))_n$.*

This definition is ambiguous if it already holds $L \subseteq gl(V)$ for some V.

Theorem 1.86. *If $L \subseteq gl(V)$ then the components X_s and X_n of the usual Jordan decomposition of $X \in L$ are both contained in L, and they coincide with the components of the abstract Jordan decomposition of X.*

Remark 1.87. *The second claim is a consequence of the first, via the properties defining both decompositions. The first claim is nontrivial because although we know X_s and X_n are polynomials in X, L is not a subalgebra of $gl(V)$.*

Proof. V is an L-submodule since $L \subseteq gl(V)$. For each L-submodule $W \subseteq V$ define $L_W = \{Y \in gl(V) | YW \subseteq W \text{ and } \mathrm{tr}_W(Y) = 0\}$. Since $L = [L, L]$ we have $L \subseteq L_W$. Define $L' := \bigcap_{W \text{ submodule } V} L_W \cap N_{gl(V)}(L)$ where $N_{gl(V)}(L) := \{Y \in gl(V) | [Y, L] \subseteq L\}$.

Fix $X \in L$, and let X_s, X_n be the Jordan decomposition of X which are in $gl(V)$. Since X_s and X_n are polynomials in L, and as $(\mathrm{ad}(X))(L) \subseteq L$, we have $X_s, X_n \in N_{gl(V)}(L)$. Also $X_s, X_n \in L_W$ for all W. So it suffices to show $L = L'$, which is a consequence of Weyl's theorem.

□

Corollary 1.88. *If L is semisimple and $\phi : L \to gl(V)$ is an L-rep with $\dim(V) < \infty$, then for any $X \in L$ with abstract Jordan decomposition $X = X_s + X_n$ the usual Jordan decomposition of $\phi(X)$ is $\phi(X) = \phi(X_s) + \phi(X_n)$. We say this earlier for $\phi = ad$.*

We'll continue looking at this next class.

2 Root Spaces

2.1 Special Linear Representations and Root Space Decompositions

2.2 Special Linear Representations

Today we'll look at the representations of $\mathfrak{sl}_2(\mathbb{F})$.

Consider an arbitrary $\mathfrak{sl}_2(\mathbb{F})$-module of finite dimension. Since H is semisimple = (diagonalizable in adjoint rep), the theorem just proved says that V must have a basis of eigenvectors for H.

This property relies on $\mathbb{F}$ being algebraically closed, so that all eigenvalues for H are present.

Key takeaway: we may decompose

$$V = \bigoplus_{\text{eigenvalues } \lambda \text{ for } H} V_\lambda$$

where $V_\lambda = \{v \in V | Hv = \lambda v\}$.

Note that our definition of V_λ makes sense even when λ is not an eigenvalue for H, but in that case $V_\lambda = 0$.

Definition 2.1. *We refer to the eigenvalues of H as **weights** and the nonzero subspaces V_λ as **weight spaces**.*

Lemma 2.2. *If $v \in V_\lambda$ and $Xv \in V_{\lambda+2}$ and $Yv = V_{\lambda-2}$.*

Proof. $HXv = [H, x]v + XHv = 2Xv + X\lambda v = (\lambda + 2)Xv$. Argument show $HYx = (\lambda - 2)Y_V$ is similar.

□

Assume our $\text{sl}_2(\mathbb{F})$-module V has $0 < \dim(V) < \infty$. There must exist at least one $\lambda \in \mathbb{F}$ with $V_\lambda \neq 0 = V_{\lambda+2}$. For this λ, we have $Xv = 0 \forall v \in V_\lambda$.

Definition 2.3. *We call the nonzero elements of this V_λ **maximal weight vectors** of V with weight λ.*

Lemma 2.4. *Assume V is irreducible $\mathfrak{sl}_2(\mathbb{F})$-module. Choose a maximal weight vector $v_0 \in V_\lambda$. Define $v_{-1} = 0$ and $v_k = \frac{1}{k!}Y^k v_0$. Then:*

1. $Hv_i = (\lambda - 2i)v_i$.
2. $Yv_i = (i + 1)v_{i+1}$.
3. $Xv = (\lambda - i + 1)v_{i-1}$.

Proof.

1. Apply previous lemma since $v_i \in V_{\lambda-2i}$.
2. By definition.
3. By induction using formulas for Lie brackets and parts a and b.

□

Continue to assume V is irreducible $\dim(V) < \infty$. Since the nonzero v_k's are H-eigenvectors with distinct eigenvalues, they are linearly independent.

There exists a smallest m with $v_m \neq 0 = v_{m+1} = v_{m+2} = ...$ Then umst have $V = \mathbb{F} - span\{v_0, v_1, ..., v_m\}$. In the basis $v_0, v_1, ..., v_m$ for V the matrices of H, X, Y are diagonal, strictly upper triangular and strictly lower triangular.

Moreover: $0 = Xv_{m+1} = (\lambda - m)v_m$.

Corollary 2.5. *Thus $\lambda = m \in \mathbb{Z}_{\geq 0}$ and the weight of any highest weight vector in an irreducible finite dimension $\mathfrak{sl}_2(\mathbb{F})$-module is a nonnegative integers, called the* ***highest weight****.*

Theorem 2.6. *Let V be an irreducible $\mathfrak{sl}_2(\mathbb{F})$-rep with $\dim(V) = m + 1 < \infty$.*

1. *Then $V = V_{-m} \oplus V_{-m+2} \oplus V_{-m+4} \oplus ... \oplus V_{m-2} \oplus V_m$ where each $V_i = \{v \in V | Hv = iv\} \neq 0$.*
2. *V has a unique highest weight space of weight m.*
3. *For each $m \geq 0$, there exists exactly one irreducible $\mathfrak{sl}_2(\mathbb{F})$-module of dimension $m + 1$ up to isomorphism.*

Proof. Check that the formulas for the action of X, Y, H in the previous lemma define an $\mathfrak{sl}_2(\mathbb{F})$-module.

□

Note that if $m = \dim(V) - 1$ is odd then V looks like

$$V_{-m} \to V_{-m+2} \to ... \to V_{-2} \to V_0 \to V_2 \to ... \to V_{m-2} \to V_m$$

while if m is even, V looks like

$$V_{-m} \to V_{-m+2} \to ... \to V_{-3} \to V_{-1} \to V_1 \to V_3 \to ... \to V_{m-2} \to V_m$$

so exactly one of V_0 or V_1 is nonzero when V is irreducible.

Corollary 2.7. *If V is any finite dimensional $\mathfrak{sl}_2(\mathbb{F})$-module then the eigenvalues for $H \in \mathfrak{sl}_2(\mathbb{F})$ acting on V are integers and if λ is one of these eigenvalues then so is $-\lambda$. Also, if $V_i = \{v \in V | Hv = iv\}$ then the number of summands is any irreducible of V is $\dim(V_0) + \dim(V_1)$.*

Proof. $\mathfrak{sl}_2(\mathbb{F})$ is semisimple so just apply Weyl's theorem.

□

2.3 Root Space Decomposition

This will generalize the weight space decomposition we just saw for $\mathfrak{sl}_2$-modules.

Suppose L is any finite-dimensional nonzero semisimple Lie algebra.

Definition 2.8. *A subalgebra of L is* ***toral*** *if all of its elements are semisimple (that is, if $X \in T$ has abstract Jordan decomposition $X = X_s + X_n$ then $X = X_s, X_n = 0$).*

Equivalently: $X \in T$ if and only if L has a basis consisting of eigenvectors for ad$X$.

Lemma 2.9. *Any total subalgebra $T \subseteq L$ is abelian: $[X, Y] = 0 \forall X, Y \in T$.*

Proof. Assume $T \subseteq L$ is toral. Suffices to show $\text{ad}_T(X) := (\text{ad}X_)|_T$ is zero $\forall X \in T$. Since $\text{ad}(X)$ is diagonalizable, preserves T, and $\mathbb{F}$ is algebraically closed, we can prove that $\text{ad}_T(X)$ has no nonzero eigenvalues (it is an easy exercise to show that T is spanned by eigenvectors for $\text{ad}_T(X)$).

Argue by contradiction. Assyme $[X, Y] = aY$ for some $Y \in T, 0 \neq a \in \mathbb{F}$. Then $(\text{ad}_T Y)(X) = [Y, X] = -[X, Y] = -aY \neq 0$ is an eigenvector for $\text{ad}_T(Y)$ with eigenvalue zero (since $\text{ad}_T(Y)(-aY) = -a[Y, Y] = 0$). But X is a linear combination of eigenvectors for $\text{ad}(Y)$ (since $Y \in T$) and also for $\text{ad}_T(Y)$ (since $X, Y \in T$). If we write $X = \sum_{\lambda \in \mathbb{F}} c_\lambda X_\lambda$ where $[Y, X_\lambda] = \lambda X_\lambda$ then $[Y, X] = \sum_{\lambda \neq 0} c_\lambda \lambda X_\lambda$ and $[Y, [Y, X]] = \sum_{\lambda \neq 0} c_\lambda \lambda^2 X_\lambda \neq 0$ contradicting $[Y, [Y, X]] = -a[Y, Y] = 0$. Thus $[X, Y] = 0 \forall X, Y \in T$.

□

Choose a maximal toral subalgebra $H \subseteq L$.

Example 2.10. *If $L = \mathfrak{sl}_n(\mathbb{F})$ then one choice for H is the subalgebra of (traceless) diagonal matrices.*

L is never abelian (since L is semisimple, $Z(L) = 0$) so any toral subalgebra $T \subseteq L$ has $T \neq L$ and is never an ideal.

H is abelian so $\text{ad}_L(H)$ is a family of commuting diagonalizable/semisimple linear maps $L \to L$. Therefore $\text{ad}_L(H)$ is simultaneously diagonalizable, meaning there is a decomposition

$$L = \bigoplus_{\alpha \in H^*} L_\alpha$$

where H^* = linear maps $H \to \mathbb{F}$ and

$$L_\alpha = \{X \in L | [h, x] = \alpha(h) X \forall h \in H\}.$$

1. If $L_\alpha \neq 0$ and $0 \neq \alpha \in H^*$ then α is called a **root**. Let Φ be the set of all roots: this is a finite subset of $H^* \backslash 0$.
2. We call $L = C_L(H) \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$ a Cartan/root space decomposition of L.

We have $L = C_L(H) \oplus \bigoplus_{\alpha \in \Phi} L_\alpha, L_\alpha = \{X \in L | [h, X] = \alpha(h)X \forall h \in H\}$.

Proposition 2.11. *For all $\alpha, \beta \in H^*$, $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$.*

Proof. Jacobi identity for $h \in H, X \in L_\alpha, Y \in L_\beta \implies$

$$\begin{aligned}
[h, [X, Y]] &= -[X, [Y, h]] - [Y, [h, X]] \\
&= [X, [h, Y]] + [[h, X], Y] \\
&= (\alpha(h) + \beta(h))[X, Y]
\end{aligned}$$

□

Proposition 2.12. *If $X \in L_\alpha$ and $0 \neq \alpha \in H^*$ then ad(X) is nilpotent.*

Proof. $(\text{ad}(X))^n(Y) \in L_{n\alpha+\beta}$ for any $Y \in L_\beta$, and if $n >> 0$ then $L_{n\alpha+\beta} = 0$ since $\dim(L) < \infty$. Since $L = \bigoplus_{\beta \in H^*} L_\beta$ it follows that $(\text{ad}(X))^n = 0$ for $n >> 0$.

□

Proposition 2.13. *If $\alpha, \beta \in H^*$ with $\alpha+\beta \neq 0$ and $\kappa(X, Y) = 0 \forall X \in L_\alpha, Y \in L_\beta$. Thus L_α and L_β are orthogonal with respect to κ if $\alpha + \beta \neq 0$.*

Proof. Since $\alpha+\beta \neq 0$, there is $h \in H$ with $(\alpha+\beta)(h) \neq 0$. Let $X \in L_\alpha, Y \in L_\beta$. Then

$$\kappa([h, X], Y) = -\kappa([X, h], Y) = -\kappa(X, [h, Y])$$

which implies $(\alpha + \beta)(h)\kappa(X, Y) = 0 \forall \kappa(X, Y) = 0$.

□

Corollary 2.14. *Killing form κ of L restricts to a nondegenerate form on $L_0 = C_L(H) = \{x \in L | [X, h] = 0 \forall h \in H\}$.*

Proof. Let $0 \neq X \in L_0$. Since $\kappa(X, Y) = 0$ for all $Y \in \bigoplus_{\alpha \in \Phi} L_\alpha$ by previous property, we must have $\kappa(X, Y) \neq 0$ for some $Y \in L_0$ since otherwise $\kappa : L \times L \to \mathbb{F}$ would be degenerate with $\kappa(X, \cdot) = 0 \in L^*$.

□

Easy fact from linear algebra: If $X, Y \in gl(V)$ with $\dim(V) < \infty$ and $XY = YX$ and Y is nilpotent, then XY is also nilpotent and $\text{tr}(XY) = \text{tr}(Y) = 0$.

Theorem 2.15. *Suppose H is a maximal toral subalgebra of a semisimple Lie algebra L with $\dim(L) < \infty$. Then*

$$H = C_L(H) = \{x \in L | [X, h] = 0 \forall h \in H\}.$$

So the Cartan decomposition of L with respect to H is just

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

Proof. Let $C = C_L(H)$. We proceed with a series of claims.

Claim 1: If $X \in C$ then $X_s \in C$ and $X_n \in C$.

If $X \in C$ then $\text{ad}(X)$ maps $H \to 0$. Since $(\text{ad}X)_s$ and $(\text{ad}X)_n$ are polynomials in $\text{ad}X$ with zero constant term, they also map $H \to 0$. But $\text{ad}(X_s)$ and $(\text{ad}X)_n = \text{ad}(X_n)$ so this means that $X_s, X_n \in C$.

Claim 2: If $X = X_s \in C$ then $X \in H$.

Suppose $X = X_s \in C$. Then $H + \mathbb{F}x$ is a toral subalgebra so must be equal to H, so $X \in H$.

Claim 3: $\kappa|_{H \times H}$ is nondegenerate.

Suppose $\kappa(h, H) = 0$ for some $h \in H$. Want to show that $h = 0$. Consider some $X \in C$. By claims 1 and 2, we have $X_n \in C$ and $X_s \in H \subseteq C$. So $\kappa(h, X) = \kappa(h, X_n) = \text{tr}(\text{ad}h \, \text{ad}X_n) = 0$ therefore $\kappa(h, C) = 0$. But this contradicts since we already shown that $\kappa|_{C \times C}$ is nondegenerate, as $H \subseteq C$ unless $h = 0$ as desired.

Claim 4: C is nilpotent, ie. $\text{ad}_C(X)$ is nilpotent $\forall X \in C$.

If $X = X_s \in C$ then $X \in H$ so $\text{ad}_C(X) = 0$ (which is clearly nilpotent). If $X = X_n \in C$ then $\text{ad}_C X_n$ is nilpotent by definition. For general, $X = X_S + X_n \in C$, we have $X_n, X_s \in C$ and $\text{ad}_C X_s$ commutes with $\text{ad}_C X_n$, so $\text{ad}(X) = \text{ad}(X_s) + \text{ad}(X_n)$ is nilpotent.

Claim 5: $H \cap [C, C] = 0$.

$$\kappa(H, [C, C]) = \kappa([H, C], C) = \kappa(0, C) = 0.$$

Since $\kappa|_{H \times H}$ is nondegenerate, this means no nonzero $X \in H$ is in $[C, C]$.

Claim 6: C is abelian, meaniing $[C, C] = 0$.

Suppose $[C, C] \neq 0$. This is a nonzero ideal of C, which is nilpotent by claim 4. So, by theorem proved to show Engel's theorem, $\mathrm{ad}(C)$ acts on $[C, C]$ as nilpotent linear transforms, which in the same basis are all strictly upper-triangular matrices. In other words, there is some $0 \neq Z \in [C, C]$ with $[X, Z] = 0$ for all $X \in C$. This element is evidently in $[C, C] \cap Z(C)$. It cannot be semisimple as then we would have $0 \neq Z = Z_s \in H \cap [C, C] = 0$. Thus we must have $0 \neq Z_n \in C$. But $\mathrm{ad}(Z_n)$ is polynomial in $\mathrm{ad}(Z)$ without constant term, so $Z_n \in Z(C)$. But then $\kappa(Z_n, C) = 0$ contradicting that $\kappa|_{C \times C}$ is nondegenerate.

Claim 7: $C = H$.

If $C \neq H$ then there exists a nilpotent nonzero element $0 \neq X = X_n \in C$. But at $Z(C) = C$ by claim 6, the argument just given implies that $\kappa(X, C) = 0$ contradicting $\kappa|_{C \times C}$ is nondegenerate.

□

2.4 Properties of Root Spaces, Abstract Root Systems

2.5 Root Space Decomposition (Continued)

Example 2.16 ($\mathfrak{sl}_3(\mathbb{F})$). *Take H to be the maximal toral subalgebra. Let $\epsilon_i : H \to \mathbb{F}$ by $\epsilon_i(M) = M_{ii}$. Each $\epsilon_i \in H^*$, but $\epsilon_1, \epsilon_2, \epsilon_3$ are not a basis as $\dim(H^*) = 2$. Then we have*

$$L = H \oplus \mathbb{F}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \oplus \mathbb{F}\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \oplus \mathbb{F}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \oplus \mathbb{F}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \oplus \mathbb{F}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \oplus \mathbb{F}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

We have $\Phi = \{\epsilon_i - \epsilon_j | 1 \leq i, j \leq 3, i \neq j\}$. Recall we have a root space decomposition

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

Because $\kappa|_{H \times H}$ is nondegenerate, for each $\alpha \in H^*$, there is a unique element $t_a \in H$ such that $\alpha(h) = \kappa(t_\alpha, h) \forall h \in H$.

Theorem 2.17 (Orthogonality properties).

1. *$H^* = \mathbb{F}$-span $\{\alpha \in \Phi\} = \mathbb{F}\Phi$.*
2. *If $\alpha \in \Phi$ then $-\alpha \in \Phi$.*
3. *If $\alpha \in \Phi, x \in L_\alpha, Y \in L_{-\alpha}$ then $[X, Y] = \kappa(X, Y) t_\alpha$.*

Proof.

1. Otherwise there is some $0 \neq h \in H$ with $\alpha(h) = 0 \forall a \in \Phi$ and then $[h, L_\alpha] = \alpha(h)L_\alpha = 0 \forall \alpha \in \Phi$ and also $[h, H] = 0 \implies [h, L] = 0 \implies 0 \neq h \in Z(L) = 0$, contradiction.

2. By a prop last time, $\kappa(L_\alpha, L_\beta) = 0$ if $\alpha, \beta \in \Phi$ and $\alpha + \beta = 0$ and $\kappa(L_\alpha, H) = 0$. So if $\alpha \in \Phi, -\alpha \in \Phi$ then it would follow that $\kappa(L_\alpha, L) = 0$, contradicting nondegeneracy of κ.

3. If $h \in H$ then $\kappa(h, [X, Y]) = \kappa([h, X], Y) = \alpha(h)\kappa(X, Y) = \kappa(t_a, h)\kappa(X, Y) = \kappa(h, \kappa(X, Y)t_\alpha) \implies \kappa(h, [X, Y] - \kappa(X, Y)t_\alpha) = 0 \forall h \in H \implies [X, Y] = \kappa(X, Y)t_\alpha$ by non degeneracy of $\kappa_{H \times H}$.

□

Example 2.18. *If $L = \sim\lessdot_3(\mathbb{F})$ where every root has form $\alpha = \epsilon_i - \epsilon_j (i \neq j)$ and every root space is $L_{\epsilon_i - \epsilon_j} = \mathbb{F}E_{ij}$ it follows that*

$$t_(\epsilon_i - \epsilon_j) = \frac{1}{\kappa(E_{ij}, E_{ji})}[E_{ij}, E_{ji}] = \frac{1}{4}(E_{ii} - E_{jj}).$$

This works if $L = \sim\lessdot_n(\mathbb{F})$ for any n.

Here are two more properties of root space decomposition:

Theorem 2.19.

1. *If $\alpha \in \Phi$ then $[L_\alpha, L_{-\alpha}] = \mathbb{F}$-span $\{t_\alpha\} \neq 0$.*

2. *$\alpha(t_\alpha)$ which by definition is $\kappa(t_\alpha, t_\alpha)$ is nonzero for all $\alpha \in \Phi$.*

3. *If $\alpha \in \Phi$ and $x_\alpha \in L_\alpha$ is nonzero then there is some $Y_\alpha \in L_{-\alpha}$ such that $\mathbb{F}$-span $\{X_\alpha, Y_\alpha, H_\alpha\} \cong \sim\lessdot_2(\mathbb{F})$ (where $H_\alpha := [X_\alpha, Y_\alpha]$) via the obvious map $X_\alpha \mapsto \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, Y_\alpha \mapsto \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, H_\alpha \mapsto \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$*

4. *In the setup of the previous statement, we have $H_\alpha = \frac{2+\alpha}{\kappa(t_\alpha, t_\alpha)} = -H_{-\alpha}$.*

Proof.

1. Just need to show $[L_\alpha, L_{-\alpha}] \neq 0$ given part 3 from the previous theorem. If $0 \neq X \in L_\alpha$ and $\kappa(X, L_{-\alpha}) = 0$ then $\kappa(X, L) = 0$ which is impossible as κ is nondegenerate.

2. We can find $X \in L_\alpha, Y \in L_{-\alpha}$ which $[X, Y] = t_\alpha$ so by part 1 if $\alpha(t_\alpha) = 0$ then $[t_\alpha, X] = [t_\alpha, Y] = 0$. In this case $\mathrm{ad}t_\alpha$ is nilpotent and semisimple so $\mathrm{ad}t_\alpha = 0 \implies 0 \neq t_\alpha \in Z(L) = 0$, contradiction.

3. Define Y_α such that $\kappa(X_\alpha, Y_\alpha) = \frac{2}{\kappa(t_\alpha, t_\alpha)}$ and then do some checking.

4. Straight forward.

□

Theorem 2.20 (Integrality Properties of Φ). *1. $\dim(L_\alpha) = 1 \forall \alpha \in \Phi$.*

2. If $\alpha \in \Phi$ then $\mathbb{F}\alpha \cap \Phi = \{-\alpha, \alpha\}$

Proof. Let $\alpha \in \Phi$. Let $S_\alpha = \mathbb{F}\text{-span}\ \{X_\alpha, Y_\alpha, H_\alpha = [X_\alpha, Y_\alpha]\} \cong \sim\lessdot_2(\mathbb{F})$ where $0 \neq X_\alpha \in L_\alpha, 0 \neq Y_\alpha \in L{-}\alpha$. Let $M = \bigoplus_c L_{c\alpha} \oplus H = H \oplus L_\alpha \oplus L_{-\alpha} \oplus$ (possibly other root spaces, but we will prove that this is zero). M is an S_α-module with weights 0 and $2c$ since $c\alpha(H_\alpha) = c \cdot \alpha\left(\frac{2+\alpha}{\alpha(t_\alpha)}\right) = 2c$ by previous properties. This implies we must have $c \in \frac{1}{2}\mathbb{Z}$ since all $\mathfrak{sl}_2$-weights are integers. Every irreducible S_α-submodule of M of even highest weight contributes one dimension (just H) to the zero weigths pace of M. But $S_\alpha \subseteq M$ is irreducible and

$$H = \text{Ker}(\alpha) \oplus \mathbb{F}H_\alpha,$$

where S_α acts as zero on the subspace $\text{Ker}(\alpha)\oplus$, which has dimension $\dim(H)-1$, and $\mathbb{F}H_\alpha$ is the 0-weight space in S_α. Since we already have $L_\alpha, L_{-\alpha} \subseteq S_\alpha$, it must hold that $L_{c\alpha} = 0$ if c is an even integer with $c \neq -2, 0, 2$.

We can conclude that $\alpha \in \Phi$ then $2a \notin \Phi$. Hence we cannot have $\alpha, \frac{1}{2}\alpha \in \Phi$ so if $\alpha \in \Phi$ then $\frac{1}{2}\alpha \notin \Phi$. This means that 1 cannot occur as a weight for $M \implies M = H + S_\alpha = \text{Ker}(\alpha) \oplus \mathbb{F}H_\alpha \oplus \mathbb{F}X_\alpha \oplus \mathbb{F}Y_\alpha$ so $\dim(L_\alpha) = 1$.

□

Proposition 2.21. *1. If $\alpha, \beta \in \Phi$ then $\beta(H_\alpha) \in Z$ (call this a **Cartan integer**) and $\beta - \beta(H_\alpha)\alpha \in \Phi$.*

2. If $\alpha, \beta \in \Phi$ and $\alpha + \beta \in \Phi$ then $[L_\alpha, L_\beta] = L_{\alpha+\beta}$.

3. If $\alpha, \beta \in \Phi$ and $\alpha + \beta \neq 0$ then there are integers $r, q \geq 0$ such that $(\beta + \mathbb{Z}\alpha) \cap \Phi = \{\beta + i\alpha | i \in \mathbb{Z}, -r \leq i \leq q\}$ ("no gaps in the α-root string through β"). Also, it holds that $\beta(H_\alpha) = r - q$.

4. L is generated by the root spaces L_α as a Lie algebra.

Proof. We will show that (3) holds. The other are straightforward.

Let $K = \sum_{i\in\mathbb{Z}} L_{\beta+i\alpha}$ where $\alpha, \beta \in \Phi$ with $\alpha + \beta = 0$. No multiple of α except $\pm\alpha$ is a root. So we have $\beta + i\alpha \neq 0 \forall i \in \mathbb{Z}$. K is a submodule of $S_\alpha \cong \mathfrak{sl}_2(\mathbb{F})$ and each $L_{\beta+i\alpha}$ is either zero if $\beta + i\alpha \notin \Phi$ or 1-dimensional if $\beta + i\alpha \in \Phi$ in which case $(\beta + i\alpha)(H_\alpha) = \beta(H_\alpha) + 2i$. In the latter case, $\beta(H_\alpha) + 2i$ is the weight of H_α on $L_{\beta+i\alpha}$.

Because all of these weights differ by an even integer, exactly one of the numbers 0 or 1 can occur as a weight, so κ is an irreducible S_α-module. Thus if r, q are maximal with $\beta - r\alpha \in \Phi, \beta + q\alpha \in \Phi$ then the corresponding weights $\beta(H_\alpha) - 2r$ and $\beta(H_\alpha) + 2q$ sum to zero, and (3) follows.

□

We have our root space decomposition $= H \oplus \bigoplus_{\alpha\in\Phi\subseteq H^*\setminus 0} L_\alpha$ and $\kappa|_{H\times H}$ is nondegenerate, and we defined $t_\alpha \in H$ for $\alpha \in H^*$ to have $\kappa(t_\alpha, h) = \alpha(h)\forall h \in H$. We now further define $(\alpha, \beta) := \kappa(t_\alpha, t_\beta)$ for $\alpha, \beta \in H^*$.

Let $E_{\mathbb{Q}} = \mathbb{Q}$-span $\{\alpha \in \Phi\}$ and $E = \mathbb{R} \otimes_{\mathbb{Q}} \in \mathbb{Q}$. One can show that

Theorem 2.22. $(\cdot,\cdot)$ *restricts to a positive definite form on* E *with* $(\alpha, \beta) \in \mathbb{Q}\forall\alpha, \beta \in \Phi$. *Additionally,*

1. Φ *spans* $\mathbb{R}$.
2. *If* $\alpha \in \Phi$ *then* $R\alpha \cap \Phi = \{-\alpha, \alpha\}$
3. *If* $\alpha, \beta \in \Phi$ *then* $\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha \in \Phi$.
4. $\frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}\forall\alpha, \beta \in \Phi$.

If $L = \sim\lessdot_n(\mathbb{F})$ and $H =$ diagonal matrices in L, then $\Phi = \{\epsilon_i - \epsilon_j | 1 \leq i, j \leq n, i = j\}$ where $\epsilon_i : H \mapsto \mathbb{F}, D \mapsto D_{ii}$. As noted earlier, we have $t_(\epsilon_i - \epsilon_j) = \frac{1}{4}(E_{ii} - E_{jj})$ and $(\epsilon_i - \epsilon_j, \epsilon_k - \epsilon_l) := \kappa(t_{\epsilon_i-\epsilon_j}, t_{\epsilon_k-\epsilon_l}) = \frac{1}{4}\langle\epsilon_i - \epsilon_j, \epsilon_k - \epsilon_l\rangle$ where

$$\langle\epsilon_i - \epsilon_j\rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

so

$$2\frac{\epsilon_i - \epsilon_j, \epsilon_k - \epsilon_l}{\epsilon_k - \epsilon_l, \epsilon_k - \epsilon_l} = \langle\epsilon_i - \epsilon_j, \epsilon_k - \epsilon_k\rangle.$$

2.6 Root Systems

Properties of the root space decomposition of L motivate the axiomatic definition of a **root system**.

Let E be a finite dimensional real vector space with a symmetric, positive definite bilinear form $(\cdot,\cdot)$.

For $0 \neq \alpha \in E$, define $r_\alpha : E \to E$ by $r_\alpha(\beta) :=$ the vector obtained by reflecting β across the hyperplane $H_\alpha = \{v \in E | (\alpha, v) = 0\} = (\mathbb{F}\alpha)^\perp$.

If $c \in \mathbb{R}$ such that $\beta - c\alpha \in H_\alpha$, then $r_\alpha(\beta) = \beta - 2c\alpha$. But $\beta - c\alpha \in H_\alpha \implies (\beta - c\alpha, \alpha) = 0 \implies (\beta, \alpha) = c(\alpha, \alpha) \implies c = \frac{(\beta,\alpha)}{(\alpha,\alpha)}$.

Thus the reflection $r_\alpha : E \to E$ belongs to $GL(E)$ and has the formula $r_\alpha(\beta) = \beta - \langle\beta, \alpha\rangle\alpha$ where $\langle\beta, \alpha\rangle := \frac{2(\beta,\alpha)}{(\alpha,\alpha)}$.

Thus,

$$r_\alpha^{-1} = r_\alpha$$
$$r_{c\alpha} = r_\alpha \text{ if } 0 \neq c \in \mathbb{R}$$
$$(r_\alpha(\beta), r_\alpha(\gamma)) = (\beta, \gamma)$$

Definition 2.23. *A subset $\Phi \in E$ is a* ***root system*** *if*

1. $|\Phi| < \infty$ *and* $0 \neq \Phi$ *and* Φ *spans* E.
2. *If* $\alpha \in \Phi$ *then* $R\alpha \cap \Phi = \{\pm\alpha\}$.
3. *If* $\alpha \in \Phi$ *then* $r_\alpha(\beta) \in \Phi \forall \beta \in \Phi$.
4. *If* $\alpha \in \Phi$ *then* $\langle \beta, \alpha \rangle \in \mathbb{Z}$.

The ***Weyl group*** *of* Φ *is* $W = \langle r_\alpha | \alpha \in \Phi \rangle \subseteq GL(E)$.

Since Φ is finite and spans E, and each r_α defines a permutation of Φ, it follows that W is isomorphic to a subgroup of the symmetric group of all permutations of Φ. Thus, the Weyl group has $|W| < \infty$.

Quick intuitive idea for root system: Suppose W is any finite subgroup of E generated by reflections r_α. Consider the set lines $\mathbb{R}_\alpha$ for $\alpha \neq 0$ with $r_\alpha \in W$¿ Replace each of these lines by a pair of vectors α and $-\alpha$. Morally, the result is a root system with Weyl group W, and any root system arises like this.

Example 2.24 ($\Phi_{A_1 \times A_1}$)**.** *4 roots,* $(\alpha, \beta) = \langle \alpha, \beta \rangle = 0$. $r_\alpha : \alpha \mapsto \alpha, -\alpha \mapsto \alpha, -\beta \mapsto \beta, -\beta \mapsto -\beta$. r_β *fixes* $\pm\alpha$, *negates* $\pm\beta$. *We have* $W = \langle r_\alpha, r_\beta \rangle \cong S_2 \times S_2 \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

Example 2.25 (Φ_{A_2})**.** *6 roots (diagonal of hexagon),* $(\alpha, \beta) = ||\alpha||\ ||\beta|| \cos\left(\frac{2\pi}{3}\right) = \frac{-||\beta||^2}{2} \implies \langle \alpha, \beta \rangle = \frac{2(\alpha,\beta)}{(\beta,\beta)} = -1$, *similarly for* $\langle \alpha, \gamma \rangle, \langle \beta, \gamma \rangle$.

$r_\alpha : \beta \iff \gamma, \alpha \iff -\alpha, -\beta \iff -\gamma$. $r_\beta : \alpha \iff \gamma, \beta \iff -\beta, -\alpha \iff -\gamma$. $r_\gamma : \alpha \iff -\beta, \beta \iff -\alpha, \gamma \iff -\gamma$.

Can check that $W \cong S_3$.

Example 2.26 (Φ_{B_2})**.** *8 roots,* $||\beta|| = \sqrt{2}||\alpha||, ||\alpha + \beta|| = ||\alpha||$. $\langle \alpha, \beta \rangle = \frac{2||\alpha||\ ||\beta|| \cos\left(\frac{3\pi}{2}\right)}{||\beta||^2} = -1$, *likewise for other inner products.*

$r_\alpha : \pm\beta \iff \pm(2\alpha + \beta), \alpha \iff -\alpha, \pm(\alpha + \beta)$ *is fixed.* $r_\beta : \pm\alpha \iff \pm(\alpha + \beta), \beta \iff -\beta, \pm(2\alpha + \beta)$ *is fixed.*

Can check that $W \cong \langle \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \rangle$.

Example 2.27 (Φ_{G_2})**.** *12 roots (6 short and 6 long like* Φ_{A_2}*), can show that* $W \cong D_{12}$.

Pairs of roots: The **rank** of Φ is $\dim(E)$. Examples above are rank 2.

Suppose $\alpha, \beta \in \Phi$ and $\beta \neq \pm\alpha$. Then

$$\langle \beta, \alpha \rangle = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} = 2\frac{||\beta||}{||\alpha||}\cos\theta \in \mathbb{Z}$$

and

$$\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle = 4\cos^2(\theta) \in \mathbb{Z}$$

as $\cos^2(\phi) \in [0, 1]$, the only possitbilities for $\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle, \theta, \frac{||\beta||^2}{||\alpha||^2}, \Phi$ are as follows:

— $\langle \alpha, \beta \rangle$ — $\langle \beta, \alpha \rangle$ — θ — $\frac{||\beta||^2}{||\alpha||^2}$ — Φ — ———————————————————————————— — 0
— 0 — $\frac{\pi}{2}$ — ? — $A_1 \times A_1$ — — 1 — 1 — $\frac{\pi}{3}$ — 1 — A_2 — — -1 — -1 — $\frac{2\pi}{3}$
— 1 — A_2 — — 1 — 2 — $\frac{\pi}{4}$ — 2 — B_2 — — -1 — -2 — $\frac{3\pi}{4}$ — 2 — B_2 — —
1 — 3 — $\frac{\pi}{6}$ — 3 — G_2 — — -1 — -3 — $\frac{5\pi}{6}$ — 3 — G_2 —

So in fact the four examples given account for all possible rank two root systems (up to isomorphism).

2.7 Properties of Simple Roots and the Weyl group

2.8 Abstract Root Systems (Continued)

Set $\langle \beta, \alpha \rangle = \frac{2(\beta,\alpha)}{(\alpha,\alpha)}$ for $\alpha, \beta \in \Phi$.

Definition 2.28. *If $\Phi \in E$ and $\Phi' \in E'$ are root systems, then an **isomorphism** $\Phi \to \Phi$ is a linear bijection $f : E \to E'$ such that $\langle f(\beta), f(\alpha) \rangle = \langle \beta, \alpha \rangle \forall \alpha, \beta \in \Phi$.*

Motivation: Suppose L is a semisimple Lie algebra over $\mathbb{C}$, finite dimensional and nonzero. Choose a maximal toral subalgebra $H \subseteq L$ and let $H^* = \{\text{linera maps } H \to \mathbb{C}\}$. For each $\alpha \in H^*$ define $L_\alpha := \{X \in L | [h, x] = \alpha(h)X \forall h \in H\}$. Set $\Phi = \{\alpha \in H^* \backslash 0 | L_\alpha \neq 0\}$. We showed $H = L_0$ is abelian. So we have a decomposition $L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$. Here, Φ is a root system in $E = \mathbb{R} - \text{span}\{\alpha \in \Phi\}$, where the relevant form $(\cdot, \cdot)$ is the Killing form of L, restricted H, and then transferred to H^* by nondegeneracy.

Also: $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta} \forall \alpha, \beta \in \Phi$.

Up to isomorphism, there are 4 root systems in $\mathbb{R}^2$: $\Phi_{A_1 \times A_1}, \Phi_{A_2}, \Phi_{B_2}$, and Φ_{G_2}.

Proposition 2.29. *Let Φ be a root system with Weyl group W. If $\sigma \in GL(E)$ has $\sigma(\Phi) = \Phi$ then $\sigma r_\alpha \sigma^{-1} = r_{\sigma(\alpha)}$ and $\langle \beta, \alpha \rangle = \langle \sigma(\beta), \sigma(\alpha) \rangle \forall \alpha, \beta \in \Phi$.*

Proof. Compute $\sigma r_\alpha \sigma^{-1}(\sigma(\beta)) = \sigma r_\alpha(\beta) = \sigma(\beta) - \langle \beta, \alpha \rangle \sigma(\alpha)$. Clearly, $\sigma r_\alpha \sigma^{-1}$ preserves Φ and sends $\sigma(\alpha) \mapsto -\sigma(\alpha)$. Also, $\sigma r_\alpha \sigma$ fixes the hyperplane $\sigma(H_\alpha)$ where $H_\alpha = \{v \in E | (v, \alpha) = 0\}$.

A priori, we don't know that $\sigma(H_\alpha) = H_{\sigma(\alpha)}$. If we knew this then it would be clear by comparing formulas that $\sigma r_\alpha \sigma^{-1} = r_{\sigma(\alpha)}$ and also $\langle \beta, \alpha \rangle = \langle \sigma(\beta), \sigma(\alpha) \rangle \forall \alpha, \beta \in \Phi$. So just need to show:

Step 1: If $\sigma \in GL(E)$ has $\sigma(\Phi) = \Phi$ and σ fixes a hyperplane $H \subseteq E$ while sending some $0 \neq \alpha \in E$ to $-\alpha$, then $H = H_\alpha$ and $\sigma = \sigma_\alpha$.

Proof idea: Define $\tau = \sigma r_\alpha$. Then $\tau(\alpha) = \alpha, \tau(\Phi) = \Phi, \tau$ fixes H point-wise. Choose a basis $v_1, v_2, ..., v_{n-1}$ for H. Set $v_n = \alpha$. Since $\alpha \notin H, v_1, v_2, ..., v_n$ is a basis for E. But the matrix of τ in this basis is the identity matrix, so $\tau = 1$.

Step 2: Let $\alpha, \beta \in \Phi$ be non-proportional (so $\alpha \neq \pm\beta$).

1. If $(\alpha, \beta) > 0$ then $\alpha - \beta \in \Phi$.
2. Of $(\alpha, \beta) < 0$ then $\alpha + \beta \in \Phi$.

Proof: Part 2 follows from part 1, swapping β and $-\beta$. For part 1: $(\alpha, \beta) > 0 \implies \langle \alpha, \beta \rangle > 0$. The acute angle between α and β must be $\frac{\pi}{3}, \frac{\pi}{4}$, or $\frac{\pi}{6}$ (since α, β not orthogonal) and must have $\langle \alpha, \beta \rangle = 1$ or $\langle \beta, \alpha \rangle = 1$. If $\langle \alpha, \beta \rangle = 1$ then $\alpha - \beta = \sigma_\beta(\alpha) \in \Phi$. If $\langle \beta, \alpha \rangle = 1$ then $\alpha - \beta = -sigma_\alpha(\beta) \in \Phi$.

□

Definition 2.30. *For $\alpha, \beta \in \Phi$ with $\beta \neq \pm\alpha$, the α-**string through** β is the set of roots $\{\beta + i\alpha | i \in \mathbb{Z}\} \cap \Phi$.*

Proposition 2.31. *There are integers $q, r \geq 0$ such that the α-string through β is exactly $\{\beta + i\alpha | -r \leq i \leq q\}$.*

Proof. If there were any gaps in the string, then we could find $p, s \in \mathbb{Z}$ with $-r \leq p < s \leq q$ where $\beta + p\alpha, \beta + s\alpha \in \Phi$ but $\beta + (p+1)\alpha, b + (s-1)\alpha \notin \Phi$.

Previous lemma implies $(\beta + \beta\alpha, \alpha) \geq 0 \geq (\beta + s\alpha, \alpha) \implies ((s-p)\alpha, \alpha) = |s - p|(\alpha, \alpha) \leq 0$, impossible as $(\cdot, \cdot)$ is positive definite.

□

Corollary 2.32. *The integers $r, q \geq 0$ such that the α-string through β is $\{\beta + i\alpha | -r \leq i \leq q\}$ satisfy $r - q = \langle \beta, \alpha \rangle = \{0, \pm 1, \pm 2, \pm 3\}$. So every α-string has at most 4 elements.*

Proof. The reflection r_α preserve the α-string through β since $r_\alpha(\beta + i\alpha) = \beta - (\langle \beta, \alpha \rangle + i)\alpha$. Therefore, must have $r_\alpha(\beta + q\alpha) = \beta - \langle \beta, \alpha \rangle \alpha - q\alpha$ so $\langle \beta, \alpha \rangle = r - q$.

□

2.9 Simple Roots and the Weyl Group

Definition 2.33. *A **base** or a **simple system** for Φ is a basis Δ for E such that each $\alpha \in \Phi$ can be written as $\alpha = \sum_{\beta \in \Delta} k_{\alpha\beta}\beta$ where coefficients $k_{\alpha\beta}$ are either all nonnegative integers or all nonpositive integers.*

Necessarily $|\Delta| = \dim(E)$. Not clear a priori that any base exists.

Example 2.34. *In each root system in $\mathbb{R}^2$, the roots labeled $\{\alpha, \beta\}$ form a base.*

Lemma 2.35. *If Δ is a base of Φ and $\alpha, \beta \in \Delta$ have $\alpha \neq \beta$, then $\alpha - \beta \in \Phi$ so $(\alpha, \beta) \leq 0$.*

Proof. If $(\alpha, \beta) > 0$ then our earlier lemma says $\alpha - \beta \in \Phi$. Since if $\alpha \neq \beta$ then also $\alpha \neq -\beta$ (since elements of Δ are linearly independent). But if $\alpha - \beta \in \Phi$ then Δ would not be a base.

□

Definition 2.36. *Given a system Δ for Φ, define the **height** of a root $\alpha = \sum_{\beta \in \Delta} k_{\alpha\beta}\beta$ to be the sum*

$$ht(\alpha) = \sum_{\beta \in \Delta} k_{\alpha\beta} \in \mathbb{Z}\backslash 0.$$

Definition 2.37. *We define $\Phi^+ = \{\alpha \in \Phi | ht(\alpha) > 0\}$ and $\Phi^- = -\Phi^+$ so that $\Phi = \Phi^+ \sqcup \Phi^-$. We call Φ^+ the set of **positive roots**, and Φ^- the set of **negative roots**.*

Theorem 2.38. *Φ does have a base/simple system.*

For each $\gamma \in E$ define $\Phi^+(\gamma) = \{\alpha \in \Phi | (\gamma, \alpha) > 0\}$. One can always choose $\gamma \in E \backslash \bigcup_{\alpha \in \Phi} H_\alpha$ and we call such γ **regular**.

If γ is regular then $\Phi = \Phi^+(\gamma) \sqcup \Phi^-(\gamma)$ where $\Phi^-(\gamma) = -\Phi^+(\gamma)$. Call $\alpha \in \Phi^+(\gamma)$ **indecomposable** if we cannot write $\alpha = \beta_1 + \beta_2$ where $\beta_i \in \Phi^+(\gamma)$.

Theorem 2.39. *If $\gamma \in E$ is regular, then the set $\Delta(\gamma)$ of indecomposable roots in Φ is a base, and every base arises in this way.*

Proof. We make a series of claims:

1. Each $\alpha \in \Phi^+(\gamma)$ is in $\mathbb{Z}_{\geq 0} - \text{span}\{\beta \in \Delta(\gamma)\}$.

Otherwise, choose $\alpha \in \Phi^+(\gamma)$ not in $\text{span}\{\beta \in \Delta(\gamma)\}$ with (α, γ) minimal. Then $\alpha = \beta_1 + \beta_2$ for some $\beta_1, \beta_2 \in \Phi^+(\gamma)$ (α cannot be indecomposable). Thus $(\alpha, \gamma) = (\beta_1, \gamma) + (\beta_2, \gamma)$ so by minimality of (α, γ) it must hold that $\beta_1, \beta_2 \in \mathbb{Z}_{\geq 0} - \text{span}\{\beta \in \Delta(\gamma)\}$, a contradiction.

1. If $\alpha, \beta \in \Delta(\gamma)$ and $\alpha \neq \beta$, then $(\alpha, \beta) \leq 0$.

Otherwise $\alpha - \beta \in \Phi, \beta = \pm\alpha$, so $\alpha - \beta$ or $\beta - \alpha$ is in $\Phi^+(\gamma)$. But then $\alpha = \beta + (\alpha - \beta)$ or $\beta = \alpha + (\beta - \alpha)$ would be decomposable.

1. $\Delta(\gamma)$ is linearly independent.

Suppose we can write $0 = \sum_\alpha c_\alpha \alpha - \sum_\beta d_\beta \beta$ where α, β range over distinct subsets of $\Delta(\gamma)$ and $c_\alpha, d|beta \geq 0$. Then

$$
\begin{aligned}
0 &\leq \left(\sum_\alpha c_\alpha \alpha, \sum_\alpha c_\alpha \alpha)\right) \\
&= \left(\sum_\alpha c_\alpha \alpha, \sum_\beta d_\beta \beta\right) \\
&= \sum_{\alpha,\beta} c_\alpha d_\beta (\alpha, \beta) & \leq 0
\end{aligned}
$$

so all $c_\alpha = 0$. Similarly derive that all $d_\beta = 0$.

1. $\Delta(\gamma)$ is a base of Φ.

Proof is clear from previous parts.

1. Every base of Φ arises as $\Delta(\gamma)$ for some regular $\gamma \in E$.

Given some base Δ for Φ, we need to find γ with $\Delta = \Delta(\gamma)$. Choose a regular γ with $(\gamma, \alpha) > 0$ for all $\alpha \in \Delta$. Then $\Phi^{+/-} = \Phi^{+/-}(\gamma)$ so every $\alpha \in \Delta$ must be indecomposable with respect to γ. This means $\Delta \subseteq \Delta(\gamma)$. As $|\Delta| = |\Delta(\gamma)| = \dim(E)$, must have $\Delta = \Delta(\gamma)$.

□

Call elements of Δ **simple roots**.

The hyperplanes H_α for $\alpha \in \Phi$ divide E into finitely many regions. We call the connected component of

$$E \setminus \bigcup_{\alpha \in \Phi} H_\alpha$$

the **Weyl chambers** of E.

2.10 Properties of Simple Roots

Fix a base Δ of Φ and define $\Phi^{+/-}$ relative to Δ. Elements of Φ^+ are **positive roots**, elements of Φ^- are **negative roots**.

Lemma 2.40. *If $\alpha \in \Phi^+$ but $\alpha \notin \Delta$ then $\alpha - \beta \in \Phi^+$ for some $\beta \in \Delta$.*

Proof. If $(\alpha, \beta) \leq 0 \forall \beta \in \Delta$ then the argument presented earlier in the 3rd part of the previous proof would show that $\Delta \sqcup \{\alpha\}$ is linearly independent. As this is impossible, must have $(\alpha, \beta) > 0$ for some $\beta \in \Delta$ and then $\alpha - \beta \in \Phi$. Since α, β cannot be proportional, $\alpha - \beta$ must be in Φ^+, since at least one coefficient in $\alpha - \beta = \sum_{\delta \in \Delta} c_\delta \delta$ must have $c_\delta > 0$.

□

By induction:

Corollary 2.41. *Each $\alpha \in \Phi^+$ can be written $\alpha = \alpha_1 + ... + \alpha_k$ where $\alpha_i \in \Delta \forall i$ and where each partial sum $\alpha_1 + ... + \alpha_j \in \Phi^+$ for $1 \leq j \leq k$.*

Lemma 2.42. *If $\alpha \in \Delta$ then $r_\alpha(\alpha) = -\alpha$ and $r_\alpha(\Phi^+ \backslash \{\alpha\}) = \Phi^+ \backslash \{\alpha\}$.*

Proof. Suppose $\beta \in \Phi^+ \backslash \{\alpha\}$. Write $\beta = \sum_{\gamma \in \Delta} k_\gamma \gamma$ where $k_\gamma \in \mathbb{Z}_{\geq 0}$. Note: β is not proportional to α. Thus $k_\gamma \neq 0$ for some $\gamma \neq \alpha$. Then the coefficient of γ in $r_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha$ is also $k_\gamma > 0$, so $r_\alpha(\beta)$ must still be in Φ^+ since it is a valid root. The lemma follows as $r_\alpha : E \to E$ is a bijection.

□

Corollary 2.43. *Set $\delta = \frac{1}{2} \sum_{\beta \in \Phi^+} \beta$ then $r_\alpha(\delta) = \delta - \alpha \forall \alpha \in \Delta$.*

Lemma 2.44. *Suppose we have a sequence $\alpha_1, ..., \alpha_m \in \Delta$. Write $r_i = r_{\alpha_i}$. Suppose $r_1 r_2 ... r_{m-1}(\alpha_m) \in \Phi^-$. Then $r_1 r_2 ... r_m = r_1 ... r_{s-1} r_{s+1} ... r_{m-1}$ for some index $1 \leq s \leq m - 1$.*

Remark 2.45. *The roots $\alpha_1, \alpha_2, ..., \alpha_m$ don't need to be all distinct.*

Proof. Set $\beta_i := r_{i+1} r_{i+2} ... r_{m-1}(\alpha_m)$, with $\beta_{m-1} := \alpha_m$. Then $\beta_0 \in \Phi^-$ and $\beta_{m-1} \in \Delta \subset \Phi^+$ so there is some smallest index s with $\beta_s \in \Phi^+$. Then $r_s(\beta_{s-1}) = \beta_s$ since $r_s^2 = 1 \implies r_s(\beta_s) = \beta_{s-1} \in \Phi^- \implies \beta_s = \alpha_s$ by previous lemma. This implies $r_s := r_{\alpha_s} = r_{\beta_s} = r_{r_{s+1} r_{s+2} ... r_{m-1}(\alpha_m)} = (r_{s+1} ... r_{m-1}) r_m (r_{m-1} ... r_{s+1})$. The result follows by substituting this expression (into $r_1 ... r_s ... r_m$) for r_s, noting that $r_i^2 = 1$.

□

Corollary 2.46. *If $\sigma = r_{\alpha_1} r_{\alpha_2} ... r_{\alpha_m}$ is an expression for $\sigma \in W$ with m as small as possible and if $\alpha_i \in \Delta$, then $\sigma(\alpha_m) \in \Phi^-$.*

Recall: Φ is a root system with Weyl group W.

Proposition 2.47. *Any given $\alpha \in \Phi$ belongs to some base of Φ.*

Proof. The hyperplanes H_β for $\beta \in \Phi \backslash \{\pm\alpha\}$ are distinct from H_α, so if we choose $\gamma \in H_\alpha$ with $\gamma \notin H_\beta \forall \beta \in \Phi \backslash \{\pm\alpha\}$, and then choose some regular γ' close to γ with $(\gamma', \alpha) = \epsilon > 0$ and $(\gamma', \beta) > \epsilon \forall \beta \in \Phi \backslash \{\pm\alpha\}$ then we'll have $\alpha \in \Delta(\gamma')$.

□

Fix a base Δ for Φ.

Theorem 2.48. *If Δ' is any base for Φ then there exists a unique element $\sigma \in W$ with $\sigma(\Delta') = \Delta$. Moreover, it holds that $W = \langle r_\alpha | \alpha \in \Delta \rangle$.*

Proof. Let $\widetilde{W} = \langle r_\alpha | \alpha \in \Delta \rangle \subseteq W$. We'll show that $\widetilde{W} = W$. Let $\delta = \frac{1}{2}\sum_{\alpha \in \Phi^+} \alpha$ and choose a regular $\gamma \in E$ along with $\sigma \in \widetilde{W}$ such that $(\sigma(\gamma), \delta)$ is maximal. If α is a simple root then $r_\alpha \sigma \in \widetilde{W}$ so our maximality assumption implies $(\sigma(\gamma), \delta) \geq (r_\alpha \sigma(\gamma), \delta) = (\sigma(\gamma), r_\alpha(\delta)) = (\sigma(\gamma), \delta - \alpha) = (\sigma(\gamma), \delta) - (\sigma(\gamma), \alpha) \forall \alpha \in \Delta$.

Thus $(\sigma(\gamma), \alpha) \geq 0 \forall \alpha \in \Delta$. Equality never holds since γ is regular and $0 \neq (\gamma, \sigma^{-1}(\alpha)) = (\sigma(\gamma), \alpha)$.

Thus we have $(\sigma(\gamma), \alpha) > 0 \forall \alpha \in \Delta$. If Δ' is any base then $\Delta' = \Delta(\gamma)$ for some reular $\gamma \in E$ and if we choose $\sigma \in \widetilde{W}$ as above then evidently $\Delta = \Delta(\sigma(\gamma)) = \sigma^{-1}(\Delta(\gamma)) = \sigma^{-1}(\Delta')$.

So for any base Δ' there is at least some $\sigma \in \widetilde{W} \subset W$ with $\sigma(\Delta') = \Delta$.

To show that $\widetilde{W} = W$, it suffices to check that $r_\alpha \in \widetilde{W} \forall \alpha \in \Phi$. Given $\alpha \in \Phi$, choose a base Δ' with $\alpha \in \Delta'$ and then choose $\sigma \in \widetilde{W}$ with $\sigma(\Delta') = \Delta$. Set $\beta = \sigma(\alpha) \in \Delta$ and then we have $r_\beta = r_{\sigma(\alpha)} = \sigma r_\alpha \sigma^{-1} \in \widetilde{W}$ as well.

Finally, we need to show that the element $\sigma \in \widetilde{W} = W$ with $\sigma(\Delta') = \Delta$ is unique for a given base Δ' of Φ.

We appeal to the technical lemma above: it's enough to show that if $\sigma \in W$ has $\sigma(\Delta) = \Delta$ then $\sigma = 1$. Assume $\sigma(\Delta) = \Delta$ and write $\sigma = r_1 r_2 ... r_m$ where $r_i = r_{\alpha_i}$ for some simple roots $\alpha_1, ..., \alpha_m \in \Delta$, and assume m is minimal. If $\sigma \neq 1$ then $m > 0$. So by the corollary above $\sigma(\alpha_m) \in \Phi^- \implies \sigma(\Delta) \neq \Delta \subseteq \Phi^+$. Thus the only way to have $\sigma(\Delta) = \Delta$ is if $m = 0$ and then $\sigma = 1$.

□

Fix an ordering $\alpha_1 \alpha_2 ... \alpha_n$ of the roots in Δ.

Definition 2.49. *We call any minimal length expression*

$$\sigma = r_{i_1} r_{i_2} ... r_{i_\ell}$$

*where $r_j := r_{\alpha_j}$ a **reduced expression** for $\sigma \in W$.*

Set $\ell(w) = \ell$, and call this the **length** of W.

Proposition 2.50. *If $\sigma \in W$ then $\ell(\sigma) = \#\{\alpha \in \Phi^+ | \sigma(\alpha) \in \Phi^-\}$.*

Remark 2.51. *This gives $\ell(r_\alpha) = 1 \forall \alpha \in \Delta$.*

Proof. Use induction and earlier lemmas.

□

2.11 Irreducible Root Systems

Definition 2.52. *A root system Φ is **irreducible** if it cannot be partitioned as a disjoint union $\Phi = \Phi_1 \sqcup \Phi_2$ where Φ_1 and Φ_2 are both nonempty and $(\alpha, \beta) = 0$ for all $\alpha \in \Phi_1, \beta \in \Phi_2$. If Φ can be partitioned in this way then Φ is **reducible**.*

Next time: there is a natural notion of **root subsystem** and **direct sum** for root systems, and any Φ is isomorphic to the direct sum of its irreducible subsystems.

Example 2.53. *The root system $\Phi_{A_1 \times A_1} = \{\pm\alpha\} \sqcup \{\pm\beta\}$ is reducible, but $\Phi_{A_2}, \Phi_{B_2}, \Phi_{G_2}$ are all irreducible.*

Proposition 2.54. *Suppose Φ has a base Δ. Then Φ is reducible if and only if there is a portion $\Delta = \Delta_1 \sqcup \Delta_2$ where $\Delta_1, \Delta_2 \neq \emptyset$ and $(\alpha, \beta) = 0 \forall \alpha \in \Delta_1, \beta \in \Delta_2$.*

3 More on Root Systems and Enveloping Algebras

3.1 Classification of Irreducible Root Sytems, Isomorphism and Conjugacy Theorems

3.2 More on Bases of Root Systems

Fix a base Δ for Φ from now on. Some facts:

Theorem 3.1.

1. *$W := \langle r_\alpha | \alpha \in \Phi \rangle = \langle r_\alpha | \alpha \in \Phi^+ \rangle = \langle r_\alpha | \alpha \in \Delta \rangle$.*
2. *If $\beta \in \Phi$ then there is some base of Φ containing β and there is some $w \in W$ with $w(\beta) \in \Delta$.*

Theorem 3.2. *For a root system Φ with base Δ, the following are equivalent:*

1. *We can write $\Phi = \Phi_1 \sqcup \Phi_2$ for some nonempty disjoint subsets Φ_i with $(\alpha, \beta) = 0 \forall \alpha \in \Phi_1, \beta \in \Phi_2$.*
2. *We can write $\Delta = \Delta_1 \sqcup \Delta_2$ for some nonempty disjoint sets Δ_i with $(\alpha, \beta) = 0 \forall \alpha \in \Delta_1, \beta \in \Delta_2$.*
3. *Φ is reducible $\equiv$ not irreducible.*

Assume these properties hold. Let $E_i = \mathbb{R}\text{-span}\{\alpha \in \Delta_i\}$. Then $[\cdot, \cdot]$ restricts to a nondegenerate form on each E_i and $E = E_1 \oplus E_2$ and each Φ_i is a root system in E_i with Δ_i as a base.

All of this extends from two to k factors as follows:

Proposition 3.3. *There is a maximal partition $\Delta = \Delta_1 \sqcup \Delta_2 \sqcup ... \sqcup \Delta_k$ into nonempty pairwise disjoint and orthogonal subsets, which is unique up to permutation of indices, and if $E_i := \mathbb{R}$-span$\{\alpha \in \Delta_i\}$ and $\Phi_i := \Phi \cap E_i$ then $E = E_1 \oplus E_2 \oplus ... \oplus E_k$ and each Φ_i is a root system in E_i with base Δ_i and $\Phi = \Phi_1 \sqcup \Phi_2 \sqcup ... \sqcup \Phi_k$.*

We call the root systems Φ_i the **irreducible components** of Φ. The proposition shows that Φ is determined up to isomorphism by these components.

Note: Φ is irreducible if and only if $k = 1$ in the prop.

Proof. The only part that is not clear is the claim that $\Phi = \Phi_1 \sqcup \Phi_2 \sqcup ... \sqcup \Phi_k$. To show this, consider some $\gamma \in \Phi$. Then there is $w \in W$ with $w(\gamma) \in \Delta$, so γ is an W-orbit of an element of some Δ_i. But orthogonality $+W = \langle r_\alpha | \alpha \in \Delta \rangle$ which means that W preserves the subspace E_i so $\gamma \in \Phi_i$. □

3.3 Invariants of Root Systems

Fix an ordering $\alpha_1, \alpha_2, ..., \alpha_\ell$ of the simple roots in our fixed base $\Delta \subseteq \Phi$.

Definition 3.4. *The **Cartan matrix** of Φ is the $\ell \times \ell$ matrix $[\langle \alpha_i, \alpha_j \rangle]_{1 \le i,j \le \ell}$, where $\langle \alpha, \beta \rangle := 2\frac{(\alpha,\beta)}{\beta,\beta} \in \mathbb{Z}$.*

Example 3.5.

- *Cartan matrix for $\Phi_{A_1 \times A_1}$ is $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ as $(\alpha_1, \alpha_2) = 0$.*
- *Cartan matrix for Φ_{A_2} is $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ because $(\alpha_1, \alpha_2) = ||\alpha_1|| \; ||\alpha_2|| \cos \frac{2\pi}{3}$ and $||\alpha_1|| = ||\alpha_2||$ so we have $\langle \alpha_1, \alpha_2 \rangle = \langle \alpha_2, \alpha_1 \rangle = 2\cos(\frac{2\pi}{3}) = -1$.*
- *Cartan matrix for Φ_{B_2} is $\begin{bmatrix} 2 & -1 \\ -2 & 2 \end{bmatrix}$ because $(\alpha_1, \alpha_1) = 1, (\alpha_2, \alpha_2) = 2, (\alpha_1, \alpha_2) = \sqrt{1}\sqrt{2} \cdot \cos \frac{3\pi}{2}$.*
- *Cartan matrix for Φ_{G_2} is $\begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$.*

Proposition 3.6. *The Cartan matrix (up to reordering of rows/columns) determines Φ (up to isomorphism). More precisely, if there is another root system $\Phi' \subseteq E'$ with ordered base Δ' and there is a bijection $f : \Delta \to \Delta'$ such that*

$$\langle a, b \rangle = \langle f(\alpha), f(\beta) \rangle \forall \alpha, \beta \in \Delta$$

then the unique linear map $E \to E'$ extending f is a root system isomorphism $\Phi \xrightarrow{\sim} \Phi'$. In particular, the linear extension of f has $\langle \alpha, \beta \rangle = \langle f(\alpha), f(\beta) \rangle \forall \alpha, \beta \in \Phi$

Proof. The linear extension $f : E \to E'$ is invertible since Δ, Δ' are bases. For $\alpha \in \Delta$, it holds $r_{f(\alpha)} = f \circ r_\alpha \circ f^{-1}$. Hence the Weyl group W' of Φ of exactly

$$\{f \circ w\ circ f^{-1} | w \in W\}.$$

Each $\beta \in \Phi$ has $\beta = w(\alpha)$ for some $w \in W, \alpha \in \Delta$. So $f(\beta) = f \circ w(\alpha) = f \circ w \circ f^{-1}(f(\alpha)) \in \Delta'$. Similar argument shows that $f^{-1}(b) \in \Phi \forall \beta \in \Phi'$ so we can conclude that f is a bijection $\Phi \to \Phi'$. Finally observe for $\alpha, \beta \in \Phi$ that

$$\begin{aligned} r_{f(\alpha)}(f(\beta)) &= f \circ r_\alpha \circ f^{-1}(f(\beta)) \\ &= f(r_\alpha(\beta)) \\ &= f(\beta) - \langle \beta, \alpha \rangle f(\alpha) \end{aligned}$$

so we must have $\langle \beta, \alpha \rangle = \langle f(\beta), f(\alpha) \rangle$.

Checking that $r_{f(\alpha)} = f \circ r_\alpha \circ f^{-1}$ for any $\alpha \in \Phi$ follows from acse when $\alpha \in \Delta$. ☐

The Coxeter graph of a root system Φ with base Δ: this is the undirected graph with vertices labeled by the elements of Δ and with exactly $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = \frac{4(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)}$ edges between vertices α and β.

Example 3.7. *$\Phi_{A_1 \times A_1}, \Phi_{A_2}, \Phi_{B_2}, \Phi_{G_2}$ have 0, 1, 2, 3 edges, respectively.*

The number of edges between α_i, α_j is the product of entries (i, j) and (j, i) of Cartan matrix.

If all roots have the same length (eg. for Φ_{A_2}) then $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle$. If roots have different lengths then we need a little extra information to recover the Cartan matrix from the Coxeter graph.

Define the **Dynkin diagram** of Φ by taking the Coxeter diagram and addig an arrow from larger root to shorter root to each double or triple edge.

Example 3.8. *For $\Phi_{A_1 \times A_1}, \Phi_{A_2}$, the Coxeter graph is the Dynkin diagram. For Φ_{B_2}, Φ_{G_2}, the arrow points towards the α_1 vertex because $||\alpha_2|| > ||\alpha_1||$.*

The Dynkin diagram determines the Cartan matrix.

Corollary 3.9. *The Dynkin diagram of Φ determines Φ up to $\cong$. Moreover, the irreducible components of Φ correspond to the connected components of the Dynkin diagram, and so Φ is irreducible if and only if the Dynkin diagram is connected.*

3.4 Classification Results and Constructions

We'll mostly skip the proofs.

Proposition 3.10. *Let* $E = \{v \in \mathbb{R}^{n+1} | v_1 + ... + v_{n+1} = 0\} \equiv \mathbb{R}^n$. *Write* $e_1, ..., e_{n+1}$ *for standard basis of* $\mathbb{R}^{n+1}$. *Define* $\Phi_{A_n} = \{e_i - e_j | 1 \leq i, j \leq n, i \neq j\}$. *Then* Φ_{A_n} *is a root system with base* $\Delta_{A_n} = \{e_i - e_{i+1} | i = 1, 2, ..., n\}$ *and Dynkin diagram of Type A. Also,* $r_{e_i - e_{i+1}}\left(\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}\right) = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} - (v_i - v_{i+1})(e_i - e_{i+1}) = \begin{bmatrix} v_1 \\ \vdots \\ v_{i+1} \\ v_i \\ \vdots \\ v_n \end{bmatrix}$ *so it follows Weyl Group* $W_{A_n} \equiv S_{n+1}$.

Proof. Straightforward calculation. □

Proposition 3.11. *Let* $\Phi_{B_n} \subseteq \mathbb{R}^n$ *be a set of* $2n + 4\binom{n}{2}$ *vectors* $\{\pm e_i | i = 1, 2, ..., n\} \sqcup [\pm e_i \pm e_j | 1 \leq i < j \leq n]$. *Then* Φ_{B_n} *is a root system with base*

$$\Delta_{B_n} = \{e_1 - e_2, e_2 - e_3, ..., e_{n-1} - e_n, e_n\}$$

and Dynkin diagram of type B. The Weyl group $W_{B_n} \equiv$ *(signed* $n \times n$ *permutation matrices).*

Proposition 3.12. *Let* $\Phi_{C_n} = \{\pm 2e_i | i = 1, 2, ..., n\} \sqcup \{\pm e_i \pm e_j | 1 \leq i < j \leq n\}$. *Then* Φ_{C_n} *is a root system with base* $\Delta_{C_n} = \{e_1 - e_2, ..., e_{n-1} - e_n, 2e_n\}$ *and Dynkin diagram of type C. Also,* $W_{C_n} = W_{B_n}$.

Proposition 3.13. *Finally, let* $\Phi_{D_n} \subseteq \mathbb{R}^n$ *be the set of* $4\binom{n}{2}$ *vectors* $\{\pm e_i \pm e_j | 1 \leq i < j \leq n\}$. *Then* Φ_{D_n} *is a root system with base* $\Delta_{D_n} = \{e_1 - e_2, e_2 - e_3, ..., e_{n-1} - e_n, e_{n-1} + e_n\}$ *and Dynkin diagram of type E. The Weyl group* W_{D_n} *is an index two normal subgroups of* $W_{B_n} = W_{C_n}$.

Φ_{A_n} is irreducible $\forall n \geq 1$ (Dynkin diagram is connected). $\Phi_{B_1} \cong \Phi_{A_1}$ as Dynkin diagram is just isolate vertex. So we only consider Φ_{B_n} for $n \geq 2$. $\Phi_{C_1} \cong \Phi_{B_1}$ and $\Phi_{C_2} \cong \Phi_{B_2}$ since Dynkin diagrams are isomorphic. So we only consider Φ_{C_n} for $n \geq 3$. $\Phi_{D_1} \cong \Phi_{A_1}, \Phi_{D_2} \cong \Phi_{A_1 \times A_1}$ (not irreducible), $\Phi_{D_3} \cong \Phi_{A_3}$ so we only consider Φ_{D_n} for $n \geq 4$.

Theorem 3.14. *Suppose* Φ *is an irreducible root system. Then the Dynkin diagram of* Φ *is either isomorphic to the Dynkin diagram of* Φ_{A_n} *(some* $n \geq 1$*),* Φ_{B_n} *(some* $n \geq 2$*),* Φ_{C_n} *(come* $n \geq 3$*),* Φ_{D_n} *(some* $n \geq 4$*), or to one of 5 exceptional diagrams:* E_6, E_7, E_8, F_4*, and* G_2. *Moreover, each of these* *exceptional diagrams does arise as the Dynkin diagram of an (irreducible) root system.*

Now let's construct these.

We've already seen Φ_{G_2}.

We have

$$\Phi_{F_4} := \{\pm e_i \pm e_j | 1 \leq i < j \leq 4\} \sqcup \{\frac{1}{2}(a_1 e_1 + a_2 e_2 + a_3 e_3 + a_4 e_4) | a_1, a_2, a_3, a_4 \in \{\pm 1\}\} \sqcup \{\pm e_i | 1 \leq i \leq 4\} \subseteq \mathbb{R}^4.$$

Φ_{F_4} has 48 elements and it's Weyl group has size 1152. It's base is $\{e_2 - e_3, e_3 - e_4, e_4, \frac{1}{2}(e_1 - e_2 - e_3 - e_n)\}$.

It suffices to construct Φ_{E_8} as Φ_{E_6}, Φ_{E_7} can be realized as subsystems. We can construct $\Phi_{E_8} \subseteq \mathbb{R}^8$ as the set of 240 vectors of the form

$$\{\pm e_i \pm e_j | 1 \leq i < j \leq 8\} \sqcup \{\frac{1}{2}(a_1 e_1 + a_8 e_8) | a_1, a_2, ..., a_8 \in \{\pm 1\} \text{ with } a_1 a_2 ... a_8 = 1\}.$$

This is a root system with base

$$\Delta_{E_8} = \{\frac{1}{2}(e_1 - e_2 - e_3 - ... - e_7 + e_8), e_1 + e_2, e_2 - e_3, e_3 - e_4, ..., e_7 - e_8\}.$$

3.5 Isomorphism and Conjugacy Theorems

Recall that if L is a semisimple Lie algebra (over an algebraically closed, characteristic zero field $\mathbb{F}$), and $H \subseteq L$ is a maximal toral subalgebra then there is a finite set $\Phi \subseteq H^*$ [0] with $L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$ where $L_\alpha := \{X \in L | [h, x] = \alpha(h) X \forall h \in H\} \neq 0$ for $\alpha \in \Phi$.

The set Φ is a root system in $E = \mathbb{R} \otimes_{\mathbb{Q}} \mathbb{Q}\text{-span}\{\alpha \in \Phi\}$ with the bilinear form on H^* dual to the Killing form of L restricted to H.

Proposition 3.15. *If L is simple then Φ is irreducible.*

Proof. If $\Phi = \Phi_1 \sqcup \Phi_2$ were reducible (with Φ_1, Φ_2 nonempty and orthogonal) and $\alpha \in \Phi_1, \beta \in \Phi_2$, then $\alpha + \beta$ is neither in Φ_1 (since $(\beta, \alpha + \beta) = (\beta, \beta) \neq 0$) nor in Φ_2 (since $(\alpha, \alpha + \beta) = (\alpha, \alpha) \neq 0$) so $\alpha + \beta \notin \Phi$ and it follows that the subalgebra of L generated by L_α for $\alpha \in \Phi_1$ is a proper nonzero ideal (since $[L_\alpha, L_\beta] = 0 \forall \alpha \in \Phi_1, \beta \in \Phi_2$). ∎

Proposition 3.16. *If $L = L_1 \oplus L_2 \oplus ... \oplus L_n$ is the decomposition of L into simple ideals then $H_i := H \cap L_i$ is a maximal toral subalgebra of L_i and the* *irreducible root system Φ_i determined by $H_i \subseteq L_i$ may be viewed as a subsystem of Φ such that $\Phi = \Phi_1 \sqcup \Phi_2 \sqcup ... \sqcup \Phi_n$ is the decomposition of Φ into irreducible components.*

Theorem 3.17. *Suppose L' is another semisimple Lie algebra with a maximal toral subalgebra H' and root system Φ'. Suppose there exists a root system isomorphism $f : \Phi \to \Phi'$. Extend f to a vector space isomorphism $f : H \xrightarrow{\sim} H'$ by setting $f(t_\alpha) = t'_{f(\alpha)}$ where for $\alpha \in \Phi, \alpha' \in \Phi'$, $t_\alpha \in H$ and $t'_{\alpha'} \in H'$ are the elements with $\kappa(t_\alpha, h) = \alpha(h), \kappa(t'_\alpha, h') = \alpha'(h')$. Choose a base $\Delta \subseteq \Phi$ along with isomorphisms between the 1-dim root spaces $L_\alpha \xrightarrow{\sim} L'_{f(\alpha)}$ for $\alpha \in \Delta$. Then there is a unique Lie algebra isomorphism $L \xrightarrow{\sim} L'$ extending $f : H \to H'$ and these chosen isomorphisms.*

3.6 Cartan Subalgebras

Definition 3.18. *A **Cartan subalgebra** of a Lie algebra L is a nilpotent subalgebra $H \subseteq L$ with $H = N_L(H)$, where $N_L(H) := \{X \in L | [X, h] \in H \forall h \in H\}$.*

Theorem 3.19. *If L is semisimple and defined over an algebraically closed field $\mathbb{F}$ with characteristic zero, then a subalgebra $H \subseteq L$ is a **maximal toral subalgebra** if and only if H is a Cartan subalgebra.*

Remark 3.20. *If the characteristic is not zero, this does not hold.*

Definition 3.21. *A **Borel subalgebra** of a Lie algebra L is a maximal solvable subalgebra.*

Theorem 3.22. *If B_1, B_2 are two Borel subalgebras of a Lie algebra L, then there is an automorphism $f \in Aut(L)$ with $f(B_1) = B_2$. Moreover, the same fact holds if B_1 and B_2 are two Cartan subalgebras.*

Example 3.23. *If $L = sl_2(\mathbb{F})$ then two Borel subalgebras are B_1 = upper triangular matrices and B_2 = lower triangular matrices. We have $f(B_1) = B_2$ for $f(X) = -X^T$.*

The textbook proves the stronger fact that $f \in \text{Aut}(L)$ can be chosen in a subgroup $E(L) \subseteq \text{Aut}(L)$ generated by $\exp(\text{ad}(X))$ for $X \in L$ that are strongly ad-nilpotent in a certain sense.

Theorem 3.24. *In a nice setting (ie. when L is semisimple, defined over $\mathbb{C}$ or a similar field), Cartan subalgebras are the same thing as maximal toral subalgebras.*

For semisimple theory over more general fields $\to$ one works with Cartan subalgebras instead of maximal toral subalgebras.

One last related important fact:

Theorem 3.25. *In a general Lie algebra, all Cartan subalgebras H are isomorphic.*

3.7 Universal Enveloping Algebras

3.8 Introduction

The following constructions pertian to arbitrary Lie algebras over any field $\mathbb{F}$. The main idea is to construct from a Lie algebra L an associative unital algebra $U(L)$ ”as freely as possible” subject to the commutation relations of L. That is, we want to build the ”most general possible algebra” $L \subseteq U(L)$ that has $X \cdot Y - Y \cdot X = [X, Y] \forall X, Y \in L$.

Definition 3.26. *An **associative unital algebra** is a vector space A with an associative bilinear multiplication operation and a compatible unit $1 \in A$.*

Definition 3.27. *An **enveloping algebra** of a given Lie algebra L is a pair (A, ϕ) where A is an associative unital algebra and $\phi : L \to A$ is a linear map such that*

$$\phi([X, Y]) = \phi(X)\phi(Y) - \phi(Y)\phi(X) \forall X, Y \in L.$$

Example 3.28. *If $L \subseteq gl(V)$ for a vector space V then $gl(V)$ is an enveloping algebra with respect to to obvious inclusion $\phi : L \hookrightarrow gl(V)$.*

Definition 3.29. *A **morphism** of enveloping algebras $f : (A_1, \phi_1) \to (A_2, \phi_2)$ is an algebra homomorphism $f : A_1 \to A_2$ such that*

¡div className=”text-center”¿ ¡Image alt=”diagram” src=”/static/images/com/liealg/1.png” width=116 height=96 /¿ ¡/Image¿

commutes. ¡/Image¿

Definition 3.30. *A **universal enveloping algebra** of L is an initial object in the category of enveloping algebras for L: that is, an enveloping algebra (U, i) such that if (A, ϕ) is any enveloping algebra for L then there is a unique morphism $(U, i) \to (A, \phi)$.*

Proposition 3.31. *If (U_1, i_1) and (U_2, i_2) are both universal enveloping algebras for L then there is a unique isomorphism $(U_1, i_1) \xrightarrow{\sim} (U_2, i_2)$.*

Proof. By definition, there are unique morphisms $f : (U_1, i_1) \to (U_2, i_2)$ and $g : (U_2, i_2) \to (U_1, i_1)$. But the identity morphism is the only morphism $(U_j, i_j) \to (U_j, i_j)$. So $f \circ g = g \circ f = \text{id}$. □

So there is at most one universal enveloping algebra of L (up to isomorphism). More involved:

Theorem 3.32. *Any Lie algebra L has a universal enveloping algebra.*

Remark 3.33. *This is always infinite-dimensional if $L \neq 0$.*

The proof requires a short digression on tensor algebras.

Let V be a finite-dimensional vector space over a field $\mathbb{F}$. Define $T^0(V) = \mathbb{F}, T^1(V) = V, T^2(V) = V \otimes V, ..., T^n(V) = V \otimes ... \otimes V$ (n factors).

Let $T(V) = \bigoplus_{n \geq 0} T^n(V)$. This is a vector space whose elements are finite linear combinations of tensors $v_1 \otimes ... \otimes v_n$ for any $n \geq 0$, any $v_i \in V$.

We make $T(V)$ into an associative unital algebra with unit $1 \in \mathbb{F} = T^0(V) \subset T(V)$ by setting

$$(v_1 \otimes ... \otimes v_k)(w_1 \otimes ... \otimes w_\ell) := v_1 \otimes ... \otimes v_k \otimes w_1 \otimes ... \otimes w_\ell$$

and extending by linearity for $v_i, w_j \in V$. The resulting structure is the **tensor algebra** of V.

Some properties of the tensor algebra $T(V)$:

- Associative
- Infinite-dimensional
- Graded as $T^m(V) \times T^n(V) \to T^{m+n}(V)$.
- Generated as an algebra by any basis of V.

The tensor algebra of V is characterized by this universal property: for any associative unital algebra A and any linear map $\phi : V \to A$, there is a unique algebra morphism $\psi : T(V) \to A$ such that

¡div className="text-center"¿ ¡Image alt="diagram" src="/static/images/com/liealg/2.png" width=176 height=96 /¿ ¡/Image¿ commutes.

Let I be the two-sided ideal in $T(V)$ generated by the set $\{x \otimes y - y \otimes x | x, y \in V\}$.

Definition 3.34. *The* ***symmetric algebra*** *of V is the quotient $S(V) := T(V)/I$.*

This is a commutative algebra with the same universal property as $T(V)$ but restricted to commutative algebras.

If $X_1, ..., X_n$ is a basis for V then $S(V)$ is isomorphic to the polynomial algebra $\mathbb{F}[X_1, ..., X_n]$ in n commuting variables.

The tensor algebra $T(V)$ is similarly isomorphic to the free associative algebra $F\langle X_1, ..., X_n \rangle$ of polynomials in n noncommuting variables.

Let's look at the proof of existence of universal enveloping algebras.

Proof. Let J be the two-sided ideal in $T(L)$ generated by the set $\{x \otimes y - y \otimes x - [X, Y] | X, Y \in L\}$.

Next we set $U(L) := T(L)/J$. Also define $\Pi : T(L) \to U(L)$ to be the quotient map and define $i : L \to U(L)$ to be the quotient map and define $i : L \to U(L)$ to be the composition $L \hookrightarrow T(L) \xrightarrow{\Pi} U(L)$.

Since $J \subseteq \bigoplus_{n>0} T^n L$ the quotient $U(L)$ is nonzero and contains $T^0 L = \mathbb{F}$.

It is not yet clear whether or not i is injective (this will turn out to be true but is not part of any definitions).

To show that $(U(L), i)$ is a universal enveloping algebra: suppose (A, j) is some enveloping algebra for L. The universal property of $T(L)$ gives us a unique algebra homomorphism $\phi' : T(L) \to A$ such that the diagram

¡div className=”text-center”¿ ¡Image alt=”diagram” src=”/static/images/com/liealg/3.png” width=176 height=96 /¿ ¡/Image¿

commutes.

But all elements $X \otimes Y - Y \otimes X - [X, Y]$ for $X, Y \in L$ are in $\mathrm{Ker}(\phi')$, since $\phi'(X \otimes Y) = \phi'(X)\phi'(Y)$ as ϕ' is an algebra homomorphism. Thus $j \subseteq \mathrm{Ker}(\phi')$ so ϕ' descends to the desired unique morphism $(U(L), i) \to (A, \phi)$. ¡/Image¿

$\square$

Example 3.35. *Suppose L is abelian so that $[X, Y] = 0 \forall X, Y \in L$. Then $J = I$ and $U(L) = S(L)$ is the symmetric algebra of L.*

Next, we'll look at the algebra structure of $U(L)$ and the Poincaré-Birkhoff-Witt theorem describing a basis for $U(L)$.

3.9 Structure of $U(L)$

Let $T_m := T^0 L \oplus T^1 L \oplus ... \oplus T^m L$ and $U_m L = \Pi(T_m)$ where $\Pi : T(L) \to U(L)$ is a quotient map and $U_{-1} := 0$.

Clearly $U_m \cdot U_n \subseteq U_{m+n}$ and $U_m \subseteq U_{m+1}$. So we can define a vector space $G^m = U_m / U_{m-1}$ and set

$$G(L) = \bigoplus_{m \geq 0} G^m \neq U(L).$$

There is a well-defined associative bilinear map $G^m \times G^n \to G^{m+n}$ so we can view $G(L)$ as a grade associative algebra.

There is also a (surjective) map $T(L) \to G(L) = \bigoplus_{m \geq 0} U_m / U_{m-1}$.

Call this map $\phi : T(L) \to G(L)$. This is surjective because $\Pi(T_m - T_{m-1}) = U_m - U_{m-1}$.

Lemma 3.36. *The map $\phi : T(L) \to G(L)$ is an algebra homomorphism with $\phi(I) = 0$, so ϕ descends to a morphism $S(L) \to G(L)$.*

Note: the algebra structure on $G(L)$ encodes multiplication in $U(L)$ modulo lower degree terms.

Proof. Let $X = X_1 \otimes ... \otimes X_m \in T^m L$ and $Y = Y_1 \otimes ... \otimes Y_n \in T^n L$. Then $\phi(XY) = \phi(X)\phi(Y)$ so ϕ is an algebra morphism. For any $X, Y \in L$, we have $\Pi(X \otimes Y - Y \otimes X) \in U_2$ but $\Pi(X \otimes Y - Y \otimes X) = \Pi([X, Y]) \in U_1$, so it follows that $\phi(X \otimes Y - Y \otimes X) \in U_1/U_1 = 0$ and $I \subseteq \mathrm{Ker}(\phi)$. □

This leads to a fundamental result, the PBW theorem.

Theorem 3.37 (The PBW Theorem)**.** *Let $w : S(L) \to G(L)$ be the morphism induced by ϕ in the previous lemma. Then w is an isomorphism of algebras.*

The proof is elementary but very technical, so we'll skip it. This theorem was several useful consequences.

Corollary 3.38. *Let W be a subspace of $T^m L$. Suppose the quotient map $T^m L \to S^m L$ sends W isomorphically onto $S^m L$. Then $U_m = U_{m-1} \oplus \Pi(W)$.*

Remark 3.39. *The symmetric algebra inherits a grading $S(L) = \bigoplus_{m \geq 0} S^m L$ where $S^m L$ is the image of $T^m L$ under $T(L) \to S(L)$.*

Proof. Consider the diagram

¡div className=”text-center”¿ ¡Image alt=”diagram” src=”/static/images/com/liealg/4.png” width=176 height=137 /¿ ¡/Image¿

The lemma and PBW theorem imply that this diagram commutes. Since w is an isomorphism, bottom two maps send w isomorphically onto G^m. Result follows as kernel of $U_m \to G^m$ is U_{m-1}. ¡/Image¿ □

Corollary 3.40. *The associated map $i : L \to U(L)$ is injective.*

Proof. Recall that i is composition $L = T^{-1} L \hookrightarrow T(L) \xrightarrow{\Pi} U(L)$. Thus if we take $W = T^1 L = L$ then quotient map $T(L) \to S(L)$ sends W isomorphically onto $S^1 L = T^1 L$, so previous corollary implies that $\Pi(L) \oplus U_0 = i(L) \oplus \mathbb{F} = U_1$, so $i(L)$ is complementary to U_0 in U_1 and i is injective. □

Corollary 3.41. *If (U, i) is any universal enveloping algebra for L then i is injective.*

Proof. (U, i) is isomorphic as an enveloping algebra to explicit constructions just given. □

Corollary 3.42. *Suppose $X_1, X_2, ...$ is an ordered basis of L (which could be infinite). Then a basis for $U(L)$ is provided by all elements of the form*

$$X_{i_1} ... X_{i_m} := \prod (X_{i_1} \otimes ... \otimes X_{i_m})$$

where $m \geq 0$ is an nonnegative integer and the indices $i_1 \leq i_2 \leq ... \leq i_m$ are weakly increasing.

Remark 3.43. *In this setting the case $m = 0$ contributes the unit element 1. Call this set of elements the PBW basis of L.*

Proof. Let W be the subspace of $T^m L$ spanned by the PBW basis elements of degree m. THen W is mapped isomorphically onto $S^m L$ and so the corollary above implies that $\Pi(W)$ is complementary to U_{m-1} in U_m.

By induction on m, follows that PBW basis spans $U(L)$ and is a basis. □

Corollary 3.44. *Suppose H is a subalgebra of L with an order basis $(h_1, h_2, ...)$ that can be extended to a basis of L by adding $(X_1, X_2, ...)$. Then the inclusion $H \hookrightarrow L$ extends to an injective algebra morphism $U(H) \mapsto U(L)$ and $U(L)$ is a free $U(H)$-module with basis given by the PBW basis elements involving only $X_1, X_2, ...$*

Proof. Clear from the description of the PBW basis. □

3.10 Free Lie Algebras

These are analogous to free groups and will enable us to define Lie algebras by generators and relations, on taking certain quotients.

Let L be a Lie algebra over a field $\mathbb{F}$ generated by a set X.

Definition 3.45. *L is **free** on X if for any map $\phi : X \to M$, where M is a Lie algebra, there exists a unique Lie algbra morphism $\psi : L \to M$ such that the diagram*

¡div className=”text-center”¿ ¡Image alt=”diagram” src=”/static/images/com/liealg/5.png” width=176 height=106 /¿ ¡/Image¿

commutes. ¡/Image¿

The usual universal property arguments show that ay two Lie algebras that are free on a given set X are uniquely isomorphic.

To establish the existence of a free Lie algebra:

- Begin with a vector space V with X as a basis.
- Form the tensor algebra $T(V)$, viewed as a Lie algebra with $[X, Y] = XY - YX$.
- Then let L be the Lie subalgebra of $T(V)$ generated by X.

Proposition 3.46. *This gives a free Lie algebra on X.*

Proof. Suppose $\phi : X \to M$ is a map, with M a Lie algebra. First extend ϕ to a linear map $V \to M \subset U(M)$. Then canonically extend this to an algebra morphism $T(V) \to U(M)$, and restrict this to a Lie algebra morphism. □

Definition 3.47. *If L is free on X and R is the ideal of L generated by some elements $\{f_j | j \in I\}$ then call quotient Lie algebra $L/R = \langle X | f_i = 0\rangle$ the **Lie algebra generated by** X with relations $f_i = 0$.*

3.11 Serre's Theorem, Weight Spaces, Standard Cyclic Modules

3.12 Setup

In this post, we'll do two things:

1. Sketch the statement of Serre's theorem and explain how to go from abstract root system to semisimple Lie algebras.
2. Begin studying highest weight representations of semisimple Lie algebras.

For the entire post, let L be a semisimple Lie algebra defined over an algebraically closed field of characteristic zero. L will always be finite-dimensional.

We fix a Cartan subalgebra $H \subseteq L$. The choice of H gives us a root space decomposition of L with corresponding root system to be denoted by $\Phi \subseteq H^*$.

Choose a set of simple roots $\Delta = \{\alpha_1, ..., \alpha_n\}$. This determines a subset of positive roots Φ^+ that is not a subset of Φ.

Define $h_j \in H$ such that $\alpha_i(h_j) = \dfrac{2(\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)} := \langle \alpha_i, \alpha_j \rangle$. Finally, choose elements $x_i \in L_{\alpha_i}$ and $y_i \in L_{-\alpha_i}$ with $[x_i, y_i] = h_i$.

Example 3.48. *When $L = \mathfrak{sl}_{n+1}(\mathbb{F})$, we can take H to be the subalgebra of diagonal matrices in L, and then one gets*

$$h_i = E_{i,i} - E_{i+1,i+1} (1 \leq i \leq n)$$
$$x_i = E_{i,i+1}$$
$$y_i = E_{i+1,i}$$

Proposition 3.49. *In this notation, we have $L = \langle x_i, y_i, h_i | i = 1, 2, ..., n\rangle$ and these generators satisfy the following relations:*

1. $[h_i, h_j] = 0$.
2. $[x_i, y_i] = h_i$ *and* $[x_i, y_i] = 0$ *if* $i \neq j$.
3. $[h_i, x_j] = \langle \alpha_j, \alpha_i \rangle x_j$ *and* $[h_i, y_j] = -\langle \alpha_j, \alpha_i \rangle y_j$.

4. $(ad(x_i))^{1-\langle\alpha_j,\alpha_i\rangle}(x_j) = 0$ *for* $i \neq j$.
5. $(ad(y_i))^{1-\langle\alpha_j,\alpha_i\rangle}(y_j) = 0$ *for* $i \neq j$.

When $i \neq j, \langle\alpha_j, \alpha_i\rangle = \dfrac{2(\alpha_j, \alpha_i)}{(\alpha_i, \alpha_i)} \in \mathbb{Z}_{\leq 0}$ so these identities make sense.

Proof.

1. Holds since H is abelian.
2. By definition and since $[X_i, Y_j] \in L_{\alpha_i - \alpha_j} = L_0$ but $\alpha_i - \alpha_j \notin \Phi$ when $i \neq j$.
3. Holds by definition: as $[h_i, x_j] = \alpha_j(h_i)x_j$ and $[h_i, y_j] = -\alpha_j(h_i)y_j$.

To check that last two, it suffices to check (S_{ij}^+) by symmetry. We assume $i \neq j$. Then $\alpha_j - \alpha_i$ is not a root, so the α_i-root string through α_j the form

$$\alpha_j, \alpha_j + \alpha_i, \alpha_j + 2\alpha_i, ..., \alpha_j + q\alpha_i$$

where $q = -\langle\alpha_j, \alpha_i\rangle$ by the earlier result.

So why does this imply that $(\mathrm{ad}(x_i))^{1-\langle\alpha_j,\alpha_i\rangle} = 0$?

Just note that $(\mathrm{ad}(x_i))^k$ maps $L_{\alpha_j} \to L_{\alpha_j + k\alpha_i}$ so if $k > -\langle\alpha_j, \alpha_i\rangle$ then $(\mathrm{ad}(x_i))^k(L_{a_j}) = 0$. □

Remark 3.50. *When* $\mathrm{rank}(\Phi) = 1$ *so that* $L \cong \mathfrak{sl}_2(\mathbb{F})$*, then the relations* $(S_{ij}^{+/-})$ *are vacuous since we have too few indices to have* $i \neq j$*, and the first three relations reduce to the usual relations* $[x, y] = h$ *and* $[h, x] = -2x$ *and* $[h, y] = 2y$*.*

Key point about these relations: they only involve constants depending on the root system Φ.

Theorem 3.51 (Serre's Theorem)**.** *Suppose* Φ *is any root system with simple system* $\Delta = \{\alpha_1, \alpha_n\}$*. Let* L *be the Lie algebra generated by the* $3n$ *elements* $\{x_i, y_i, h_i | i = 1, 2, ..., n\}$ *subject to the just the first 5 relations from the previous property. Then* L *is a finite-dimensional semisimple Lie algebra with Cartan subalgebra* $H := \mathbb{F} - span\{h_1, ..., h_n\}$ *and with corresponding root system* Φ *(viewing* $\Phi \subseteq H^*$ *by setting* $\alpha_i(h_j) = \langle\alpha_i, \alpha_j\rangle$ *and extending by linearity).*

We will skip the proof out of time constraints.

Remark 3.52. *If you do this construction but leave out* (S_{ij}^+) *and* (S_{ij}^-) *then the resulting Lie algebra is usually infinite-dimensional, but the proof of Serre's theorem proceeds by first studying this object.*

Corollary 3.53. *For each root system* Φ *there is a semisimple Lie algebra with* Φ *as its root system (relative to some Cartan subalgebra). Moreover, if we have isomorphic root systems* $\Phi \xrightarrow{\sim} \Phi'$ *and with a simple system* $\Delta \subseteq \Phi$ *whose image* *in Φ' is Δ', then the obvious bijection between generating sets of the associated Lie algebras extends to an isomorphism of Lie algebras.*

Proof. The described map on generators extends to a Lie algebra homomorphism because the images of the generators of one Lie algebra satisfy the same defining relations in the other Lie algebra. (This is a general phenomena of morphisms from Lie algebras constructed by generators and relations: a map from the generating set extends uniquely to a morphism if and only if the images of the generators still satisfy all defining relations).

You can build such a morphism in either direction, and there must be inverses because their composition is the identity map on the generating sets. $\square$

3.13 Criteria for Semisimplicity

Recall that L is **reductive** if $\text{Rad}(L) = Z(L)$.

1. Semisimple $\implies$ reductive since $Z(L) \subseteq \text{Rad}(L)$ so if $\text{Rad}(L) = 0$ then $Z(L) = 0 = \text{Rad}(L)$.
2. Abelian $\implies$ reductive since if $Z(L) = L$ then $\text{Rad}(L) = L = Z(L)$.

Proposition 3.54. *If L is reductive then $[L, L]$ is semisimple and $L = [L, L] \oplus Z(L)$.*

Proof. Since $L/Z(L) \cong textadL$ is semisimple, it acts completely reducibly on L, so $L = M \oplus Z(L)$ for some ideal M. But then $[L, L] = [M, M] \subseteq M$ and we must have equality since $[L/Z(L), L/Z(L)] = L/Z(L)$ as this is semisimple. $\square$

Proposition 3.55. *Let $L \subseteq gl(V)$, where V is finite-dimensional, be a nonzero Lie algebra acting irreducibly on V. Then L is reductive with $\dim(Z(L)) \leq 1$ and if $L \subseteq \mathfrak{sl}(V)$ is not a subset of $gl(V)$, then L is semisimple.*

Proof. Let $S = \text{Rad}(L)$. By Lie's theorem, S has a common eigenvector in V, call this $v \in V$, with $s \cdot v = \lambda(s)v \forall s \in S$ where $\lambda \in S^*$. If $x \in L$ then $[S, x] \in S$ so

$$S \cdot (x \cdot v) = x \cdot (S \cdot v) + [S, x] \cdot v = \lambda(S)x \cdot v + \lambda([S, x])v.$$

Since L acts irreducibly on V, all vectors in V are obtained by applying sequences of elements $x \in L$ to $v \in V$ and taking linear combinations. Thus in some basis of V_j, each $s \in S$ acts as a triangular matrix with $\lambda(S)$ on diagonal.

Because the commutators $[S, L] \subset S$ all have trace zero, λ must vanish on $[S, L]$. But this means that $s \in S$ acts exclusively as the scalars $\lambda(s)$. Therebefore $\text{Rad}(L) = S \subseteq Z(L)$ whence $\text{Rad}(L) = Z(L)$ so L is reductive and $\dim(S) \leq 1$. Finally if $L \subseteq \mathfrak{sl}(V)$ then $S = 0$ since $\mathfrak{sl}(V)$ has no scalars except zero, in which case L is semisimple.

□

Recall the classical Lie algebras:

- Type A_n: $\mathfrak{sl}_{n+1}(\mathbb{F}) =$ (traceless matrices)
- Type C_n: $\mathfrak{sp}_{2n}(\mathbb{F}) \cong \{x \in \mathfrak{gl}_{2n}(\mathbb{F}) | Jx + x^T J = 0\}$
- Type B, D: $\mathfrak{D}_n(\mathbb{F}) \cong \{x \in \mathfrak{gl}_n(\mathbb{F}) | x + x^T = 0\}$

Proposition 3.56. *Each classical Lie algebra is semisimple and in fact simple.*

Proof. $\mathfrak{gl}(V) = \mathfrak{sl}(V) +$ (scalar matrices) and $\mathfrak{gl}(V)$ acts irreducibly on V, so $\mathfrak{sl}(V)$ also acts irreducibly. Then $\mathfrak{sl}_n(\mathbb{F}) \cong \mathfrak{sl}_n(\mathbb{V})$ for $V = \mathbb{F}^n$ is semisimple.

We observed that $\mathfrak{sp}_{2n}(\mathbb{F}) \subseteq \mathfrak{sl}_{2n}(\mathbb{F})$ and $\mathfrak{D}_n(\mathbb{F}) \subseteq \mathfrak{sl}_n(\mathbb{F})$ so we just need to show that these Lie algebras act irreducibly on $\mathbb{F}^{2n}$ or $\mathbb{F}^n$. This is straight forward from explicit constructions: just want to find a way to express any $x \in \mathfrak{gl}_n(\mathbb{F})$ as a linear combination of elements in your classical algebra and scalars. In this way, deduce that each algebra is semisimple. Simplicity follows by computing the root systems and their Dynkin diagrams. □

3.14 Representation Theory of Semisimple Lie Algebras

Assumptions:

- L is a semisimple Lie algebra over a field $\mathbb{F}$.
- $\mathbb{F}$ is algebraically closed with characteristic zero.
- $H \subseteq L$ is a fixed Cartan subalgebra.
- $\Phi \subseteq H^*$ is the corresponding root system.
- $\Delta \subseteq \Phi$ is a simple system with elements $\alpha_1, ..., \alpha_n$.
- $W := \langle r_\alpha | \alpha \in \Phi \rangle$ is a Weyl group of Φ.

Our goal is to understand the finite-dimensional L-modules, in particular those which are irreducible.

Suppose V is a finite-dimensional L-module. Then H acts on V as commuting diagonalizable operators, so V can be decomposed into simultaneous eigenspaces for H.

Specifically, we can write $V = \bigoplus_{\lambda \in H^*} V_\lambda$ where $V_\lambda := \{v \in V | h \cdot v = \lambda(h) v \forall h \in H\}$. If $V_\lambda \neq 0$ (which can only happen for finitely-many $\lambda \in H^*$) then we call V_λ a **weight space** and λ a **weight**.

Example 3.57. *If $V = L, L$ acting by the adjoint representation, then the weight spaces are just the root spaces L_α (along with H) and the weights are the roots $\alpha \in \Phi$ (along with 0).*

Example 3.58. *If* $L = \mathfrak{sl}_2(\mathbb{F}) = \langle x = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \rangle$ *and* V *is irreducible, then* V *looks like*

$$V = V_{-m} \oplus V_{-m+2} \oplus ... \oplus V_m$$

for some integer $m \geq 0$*. Each* V_i *is a weight space for the weight* $\lambda : h \mapsto i$*. Everything is easy because* $H = \mathbb{F}$*-span*$\{h\}$*.*

Some pathologies: if $\dim(V) = \infty$ then the sum of the weight spaces $V_\lambda \subseteq V$ may be a proper subspace, though this sum of subspaces is always direct.

However:

Lemma 3.59. *Let* V *be an arbitrary* L*-module. Then:*

1. L_α *maps* V_λ *into* $V_{\lambda+\alpha} \forall \lambda \in H^*$ *and* $\alpha \in \Phi$*.*
2. $U := \sum_{\lambda \in H^*} V_\lambda$ *is equal to* $\bigoplus_{\lambda \in H^*} V_\lambda$ *and is an* L*-submodule of* V*.*
3. *If* $\dim(V) < \infty$ *then* $U = V$*.*

Proof. We'll just prove the first part. Note for $x \in L_\alpha, v \in V_\lambda, h \in H$ that $h \cdot x \cdot v = x \cdot h \cdot v + [h, x] \cdot v = (\lambda(h) + \alpha(h))x \cdot v$ so L_α sends V_λ to $V_{\lambda+a}$. □

3.15 Standard Cyclic Modules

Definition 3.60. *A* ***maximal vector*** *of weight* $\lambda \in H^*$ *in an* L*-module* V *is a nonzero vector* $v^+ \in V$ *with* $x_{v^+} = 0 \forall \alpha \in \Delta, x \in L_\alpha$*.*

Remark 3.61. *This depends implicitly on choice of simple roots* Δ*. If* $\dim(V) = \infty$ *then it could happen that there are no such vectors.*

But if $\dim(V) < \infty$ then the **Borel subalgebra** $B = H \oplus \bigoplus_{\alpha \in \Phi^+} L_\alpha$ is solvable and so has a common eigenvector in V (by Lie's theorem) and this eigenvector provides a maximal vector (because it is killed by all L_α for $\alpha \in \Phi^+$).

So the idea is to first study L-modules generated by a maximal vector.

Note that any L-module structure on V corresponds to a map $L \to \mathfrak{gl}(V)$ which his an algebra, and so extends uniquely to an associative algebra module structure on V related to $\mathcal{U}(L)$.

Definition 3.62. *If* $V = \mathcal{U}(L) \cdot v^+$ *for a maximal vector* v^+ *of weight* λ*, then we say* V *is a* ***standard cyclic*** *of weight* λ*, and we call* v^+ *the* ***highest weight vector*** *of* V*.*

Fix $x_\alpha \in L_\alpha, y_\alpha \in L_{-\alpha}$ with $[x_\alpha, y_\alpha] = h_\alpha$ for each $\alpha \in \Phi^+$. Write $\lambda > \mu$ for λ, μ is a sum of positive roots.

Theorem 3.63 (Structure Theorem for Standard Cyclic Modules)**.** *Let V be a standard cyclic L-module with highest weight vector $v^+ \in V_\lambda$. Write $\Phi^+ = \{\beta_1, ..., \beta_m\}$ and $y_i := y_{\beta_i}$. Then:*

1. *V is spanned by the vectors $y_{i_1}...y_{i_k}v^+$ as $(i_1, ..., i_k)$ ranges over all weakly increasing sequences $1 \leq i_1 \leq ... \leq i_k \leq m$. Also V is the direct sum of its weight spaces.*
2. *All weights μ for V have the form*

$$\mu = \lambda - \sum_{i=1}^{n} k_i \alpha_i$$

where $k_i \in \mathbb{Z}_{\geq 0}$ and therefore $\mu < \lambda$.

1. *For each $\mu \in H^*, \dim(V_\mu) < \infty$ and $\dim(V_\lambda) = 1$.*
2. *Each submodule of V is a direct sum of weight spaces.*
3. *V is an indecomposable L-module with a unique maximal proper submodule whose quotient is irreducible.*
4. *Every nonzero homomorphic image of V is also standard cyclic of weight λ.*

Proof. Let $N^- = \bigoplus_{\alpha\in\Phi^-} L_\alpha$ and $B = H \oplus \bigoplus_{\alpha\in\Phi^+} L_\alpha$ so $L = N^- \oplus B$. PBW theorem implies that $\mathcal{U}(L)v^+ = \mathcal{U}(N^-)\mathcal{U}(B)v^+ = \mathcal{U}(N^-)\mathbb{F}v^+$ since v^+ is a common eigenvectof rof B. Part 1 follows from the PBW theorem for N^-.

Our lemma above implies that $y_{i_1}...y_{i_k}v^+$ has weight $\mu = \lambda - \beta_{i_1} - ... - \beta_{i_k}(\star)$ so part 2 also follows. There are only finitely many vectors in part 1 that can give rise a given weight μ via $(\star)$ so $\dim(V_\mu) < \infty$, and the only such weight vector of weight λ is v^+ so $\dim(V_\lambda) = 1$.

For part 4, let W be a submodule of V and write $w \in W$ as a sum of vectors $v_i \in V_{\mu_i}$ for distinct weights μ_i. We want to show that each v_i is in W. Suppose otherwise and choose $w = v_1 + ... + v_n$ with n minimal where none of $v_1, ..., v_n$ are in W. (Then $n > 1$). Find $h \in H$ with $\mu_1(h) \neq \mu_2(h)$. Then $h \cdot w = \sum_i \mu_i(h)v_i \in W$ so $(h - \mu_1(h))w \in W$ but $(h - \mu_1(h))w$ has the form $(\mu_2(h) - \mu_1(h))v_2 + ... + (\mu_n(h) - \mu_1(h))v_n \neq 0$ contradicting minimality of n. Thence each $v_i \in W$ and part 4 holds.

We conclude from parts 3 and 4 that each proper submodule of V is in the sum of weight spaces other than V_λ, so the sum W of all proper submodules is proper, so the quotient V/W must be irreducible. This proves part 5, and part 6 holds by definition. □

Corollary 3.64. *If V is as in the theorem and V is irreducible then v^+ is the unique maximal weight vector up to rescaling.*

Proof. If there were another such vector of weight λ' then the theorem implies that $\lambda < \lambda'$ and $\lambda' < \lambda$ so $\lambda = \lambda'$. □

4 Module Theory and Character Formulas in Lie Algebras

4.1 Finite-Dimensional Modules, Multiplicity Formulas, Formal Characters

4.2 More About Standard Cyclic Modules

Theorem 4.1. *If V and W are irreducible standard cyclic L-modules with same highest weight $\lambda \in H^*$, then $V \equiv W$.*

Theorem 4.2. *If $\lambda \in H^*$ then there exists an irreducible standard L-module $V(\lambda)$ of highest weight λ.*

Proof. Let $X = V \oplus W := \{v + w | v \in V, w \in W\}$. This is an L-module and if $v^+ \in V$ and $w^+ \in W$ are highest weight vectors then $x^+ := v^+ + w^+ \in X$ is a maximal vector also of weight λ.

Let Y be the submodule of X generated by x^+. This is standard cyclic by definition. But $v \equiv Y/\text{Ker}(\pi_1)$ and $W \cong Y/\text{Ker}(\phi_2)$ where $\pi_1 : Y \to V$ and $\pi_2 : Y \to W$ are the obvious surjective homomorphisms. This means V and W are both isomorphic to the unique irreducible quotient of Y. □

To prove the second theorem we must explain how to construct standard cyclic modules.

Begin with a 1-dimensional vector space $D_\lambda = \mathbb{F}-\text{span}\{v^+\}$ spanned by some vector v^+. Let $\lambda \in H^*$ and $B = B(\Delta) := H \oplus \bigoplus_{\alpha\in\Phi^+} L_\alpha \subset L$.

The **Borel subalgebra** B acts on D_λ linearly by $h \cdot v^+ := \lambda(h)v^+$ and $x_{v^+} := 0$ for $h \in H, \alpha \in \Phi^+, x \in L_\alpha$.

This makes D_λ into a module for B and for $\mathcal{U}(B)$.

Definition 4.3. *Let $Z(\lambda) := \mathcal{U}(L) \otimes_{\mathcal{U}(B)} D_\lambda$. This is a general construction of a $\mathcal{U}(L)$-module: $\mathcal{U}(L) + D_\lambda$. Concretely, $Z(\lambda)$ is a vector space spanned by the tensors $x \otimes y$ ($x \in \mathcal{U}(L), y \in D_\lambda$) subject to relations*

$$
\begin{aligned}
&(x + x') \otimes y = x \otimes y + x' \otimes y \\
&x \otimes (y + y') = x \otimes y + x \otimes y' \\
&c(x \otimes y) = (cx) \otimes y = x \otimes (cy) \\
&xb \otimes y = x \otimes by
\end{aligned}
$$

for $x \in \mathcal{U}(L), b \in \mathcal{U}(L), y \in D_\lambda$.

THe way L acts on $Z(L)$ is a $A \cdot (x \otimes x) := (Ax) \otimes y$.

Proposition 4.4. *$Z(\lambda)$ is a standard cyclic L-module of weight λ.*

Proof. Every $y \in D_\lambda$ is a scalar multiple of v^+ so every tensor $x \otimes y \in Z(\lambda)$ is equal to $\widetilde{x} \cdot (1 \otimes v^+)$ where $\widetilde{x} \in \mathcal{U}(L)$ is a vector multiple of x. For $x \in L_\alpha$ for $\alpha \in \Delta$ we have $x \cdot (1 \otimes v^+) = x \otimes v^+ = 1 \otimes xv^+ = 1 \otimes 0 = 0$. Also for $h \in H \subset B$ we have $h \cdot (1 \otimes v^+) = h \otimes v^+ = 1 \otimes hv^+ = 1 \otimes \lambda(h) v^+ = \lambda(h)(1 \otimes v^+)$. ☐

Let $N^- = \bigoplus_{\alpha \in -\Phi^-} L_\alpha$. The relation $xb \otimes v^+ = x \otimes bv^+] for all b \in B$, since $b = N^- \oplus B$, implies that if $\Phi^+ = \{\beta_1, \beta_2, ...\}$ and $\{x_i = x_{\beta_i}$ spans $L_{\beta_i}\}, \{y_i = y_{\beta_i}$ spans $L_{-\beta_i}\}$ then

$$\{y_{i_1} y_{i_2} ... y_{i_k} \otimes v^+ | k \geq 0 \text{ and } i_1 \leq ... \leq i_k\}$$

is a basis for $Z(\lambda)$, via the PBW theorem.

Proposition 4.5. *$Z(\lambda) \cong \mathcal{U}(L)/I(\lambda)$ as $\mathcal{U}(L)$-modules, where $I(\lambda)$ is the left ideal generated in $\mathcal{U}(L)$ by the elements*

$$\{x_1, x_2, ...\} \cup \{h_\alpha - \lambda(h_\alpha) \cdot 1 | \alpha \in \Phi\}.$$

Proof. These generators annihilate $1 \otimes v^+$ so there is a surjective morphism $\mathcal{U}(L)/I(\lambda) \to Z(\lambda)$ which is injective using PBW theorem. ☐

Theorem 4.6. *Define $V(\lambda)$ for $\lambda \in H^*$ to be the unique irreducible quotient of the standard cyclic module $Z(\lambda)$. Then $V(\lambda)$ is standard cyclic of weight λ and irreducible.*

Remark 4.7. *$V(\lambda)$ might still be infinite-dimensional.*

Proof. Since $Z(\lambda)$ is standard cyclic, and since $V(\lambda)$ as a quotient is a homomorphic image of $Z(\lambda)$, everythnig follows from structure theorem for standard cyclic modules. ☐

In some sense, the hardest part of the theorem is showing $Z(\lambda) \neq 0$ (but we will not discuss this issue in detail, follows from PBW theorem).

4.3 Two New Goals

Two new goals:

1. Explain when $V(\lambda)$ is finite-dimensional.
2. Determine weight spaces $V(\lambda)_\mu \subseteq V(\lambda)$.

Proposition 4.8. *If V is any irreducible L-module with $\dim(V) < \infty$ then $V \cong V(\lambda)$ for some $\lambda \in H^*$.*

Proof. If $\dim(V) < \infty$ then Lie's theorem applied to B-action on V implies existence of a maximal vector of some weight λ. This vector must generate V by irreducibility, so $V \cong V(\lambda)$ by the first theorem. □

For each simple root $\alpha_i \in \Delta$ let $S_i = S_{\alpha_i} = L_{-\alpha_i} \oplus \mathbb{F}h_{\alpha_i} \oplus L_{\alpha_i} \cong \mathfrak{sl}_2(\mathbb{F})$. Then $V(\lambda)$ is a module for S_i and a maximal vector for L is also maximal for S_i.

Theorem 4.9. *If $V \cong V(\lambda)$ and $\dim(V) < \infty$ then $\lambda(h_{\alpha_i}) \in \mathbb{Z}_{\geq 0} \forall \alpha_i \in \Delta$ and if $\mu \in H^+$ is any weight for V then $\mu(h_{\alpha_i}) \in \mathbb{Z} \forall \alpha_i \in \Delta$.*

Proof. Follows from $\mathfrak{sl}_2$-representation theory, as V decomposes as sum of finite dimensional irreducible S_i-modules. □

Definition 4.10. *Call $\lambda \in H^*$*

- ***dominant** if $\lambda(h_\alpha) > 0 \forall \alpha \in \Delta$ (equivalently $\forall \alpha \in \Phi^+$)*
- ***integral** if $\lambda(h_\alpha) \in \mathbb{Z} \forall \alpha \in \Delta$ (equivalently $\forall \alpha \in \Phi$).*

*Then $\lambda \in H^*$ is **dominant integral** if $\lambda(h_\alpha) \in \mathbb{Z}_{\geq 0} \forall \alpha \in \Delta$.*

Let Λ be an abelian group of integral weights and Λ^+ the subset of dominant integral weights. Note that $\Phi \subset \Lambda$. For an L-module V let $\Pi(v) \subseteq H^*$ be its set of weights and define $\Pi(\lambda) = \Pi(v(\lambda))$. If $\dim(V) < \infty$ then $\Pi(\lambda) \subset \Lambda$.

Now, let's look at the main theorem.

Theorem 4.11. *Suppose $\lambda \in \Lambda^+$. Then $V(\lambda)$ has finite dimension and the Weyl group $W \subseteq GL(H^+)$ permutes $\Pi(\lambda)$ with $\dim(V(\lambda)_\mu) = \dim(V(\lambda)_{\sigma\mu}) \forall \sigma \in W$.*

Corollary 4.12. *The map $\lambda \mapsto V(\lambda)$ is a bijection from Λ^+ to isomorphism classes of irreducible finite dimensional L-modules.*

Proof. Combine the main theorem with the last two theorems and propositions. □

4.4 Proof of Main Theorem

Now let's prove the main theorem.

Proof. Some identities in $\mathcal{U}(L)$: writing $x_i = x_{\alpha_i}, y_i = y_{\alpha_i}, h_i = h_{\alpha_i}$ for $\alpha_i \in \Delta$:

1. $[x_j, y_i^{k+1}] = 0$ when $i \neq j, k \geq 0$.
2. $[h_j, y_i^{k+1}] = -(k+1)\alpha_i(h_j)y_i^{k+1}$ $(k \geq 0)$
3. $[x_i, y_i^{k+1}] = -(k+1)y_i^k(k - h_i)$ $(k \geq 0)$

Straightforward algebra by induction on $k \geq 0$.

Now we derive a series of claims:

Claim 1: $y_i^{m_i+1}v^+ = 0$ where $m_i = \lambda(h_i) \in \mathbb{Z}_{\geq -}$, and $v^+ \in V = V(\lambda)$ is a highest weight vector.

Proof. Otherwise we can use parts (1-3) to show that $y_i^{m_i+1}v^+$ is a second maximal vector of weight $\neq \lambda$ which is impossible.

Claim 2: v contains a nonzero finite dimensional $S_i = S_{\alpha_i} \cong \mathfrak{sl}_2(\mathbb{F})$-module.

Proof. Consider subspace spanned by $v^+, y_iv^+, y_i^2v^+, \ldots$ This is finite-dimensional by claim 1.

Claim 3: V is a sum of finite-dimensional S_i-modules.

Proof. Let V' be the sum of all S_i-submodules of finite dimension in V. Then $V' \neq 0$ by claim 2. Check that V' is an L-module, hence $V' = V$ since V is irreducible.

Claim 4: If $\phi : L \to \mathfrak{gl}(V)$ is a representation corresponding to an L-module structure on V then $\phi(x_i)$ and $\phi(y_i)$ are both **locally nilpotent** (meaning nilpotent when restricted to a finite-dimensional subspace).

Proof. Each $v \in V$ is a finite sum of finite dimensional S_i-modules, on which $\phi(x_i), \phi(y_i)$ act as nilpotent operators, by $\mathfrak{sl}_2$-representation theory.

Claim 5: Define $\sigma_i := \exp(x_i)\exp(-y_i)\exp(x_i)$. This is an automorphism of V (as a vector space).

Proof. Just need to check that σ_i is well-defined, but this follows from previous claim.

Claim 6: if μ is a weight of V then $\sigma_i(V_\mu) = V_\nu$ for $\nu := r_{\alpha_i}(\mu)$ with $r_\alpha \in W$ the usual reflection.

Proof. Follows from $\mathfrak{sl}_2$-representation theory since V_μ is finite-dimensional S_i-submodule.

Claim 7: If $\mu \in \Pi(V) = \Pi(\lambda)$ and $w \in W$ then $w(\mu) \in \Pi(\lambda)$ and $\dim(V_{w(\mu)}) = \dim(V_\mu)$.

Proof. Immediate from claim 6 as $W = \langle r_{\alpha_i} | \alpha_i \in \Delta \rangle$.

Claim 8: $\Pi(\lambda)$ is finite.

Proof. $\Pi(\lambda)$ is a subset of the set of W-conjugates of all dominant integral $\mu \in H^*$ with $\mu < \lambda$ by previous claim and structure theorem of standard cyclic modules. Results in chapter 13 of textbook imply this set is finite.

Claim 9: $\dim(V) < \infty$ since $\Pi(V) = \Pi(\lambda)$ is finite and each $\mu \in \Pi(\lambda)$ has $\dim(V_\mu) < \infty$. □

4.5 Multiplicity Formula

Fix $\lambda \in \Lambda^+$. Then $V(\lambda)$ is finite dimensional and irreducible. For $\mu \in H^*$ let $m_\lambda(\mu) := \dim(V(\lambda)\mu) \in \mathbb{Z}_{\geq 0}$.

This is zero if $\mu \notin \Pi(\lambda)$. The $m_\lambda(\mu)$ the **multiplicity** of μ in $V(\lambda)$. If $\mu \in H^*$ and $\mu \notin \Lambda$ then $\mu \notin \Pi(\lambda)$ so $m_\lambda(\mu) = 0$.

Theorem 4.13 (Freudenthal's Formula)**.** *If $\mu \in \Lambda$ and $\sigma = \frac{1}{2} \sum_{\alpha \in \Phi^+} \alpha$ then*

$$((\lambda + \delta, \lambda + \delta) - (\lambda + \sigma, \lambda + \sigma)) m_\lambda(\mu) = 2 \sum_{\alpha \in \Phi^+} \sum_{i=1}^{\infty} m_\lambda(u + i\alpha)(u + i\alpha, \alpha)$$

and this formula provides an effective algorithm to compute $m_\lambda(\mu)$.

Remark 4.14. *Key point: if $\lambda \neq \mu$ then $||\lambda + \sigma||^2 \neq ||\mu + \sigma||^2$, so we can divide both sides by this number.*

Minor point: $m_\lambda(\lambda) = 1$.

4.6 Formal Characters

We want to assign each finite-dimensional L-module a vector (similar to character of a group representation) that identifies its isomorphism class.

Notation: Let $\mathbb{Z}[\Lambda]$ be the free $\mathbb{Z}$-module with basis given by symbols $\{e^\lambda | \lambda \in \Lambda\}$ and make this additive group into a ring by setting $e^\lambda e^\mu = e^{\lambda + \mu}$, where $\Lambda \subset H^*$ is the infinite set of integral weights, including $0 \in A$.

Definition 4.15. *If $\lambda \in \Lambda^+$ then the **formal character** of $V \cong V(\lambda)$ is $ch_V = ch_\lambda := \sum_{\mu \in \Pi(\lambda)} m_\lambda(\mu) e^\mu \in \mathbb{Z}[\Lambda]$. If V is an arbitrary finite dimensional L-module then V has unique decomposition $V \cong V(\lambda_1) \oplus ... \oplus V(\lambda_k)$ with each $\lambda_i \in \Lambda^+$ and we set $ch_V = \sum_{i=1}^k ch_{\lambda_i}$.*

Example 4.16. *If $L = sl_2(\mathbb{F})$ then $ch_\lambda = e^\lambda + e^{\lambda - \alpha} + ... + e^{\lambda - m\alpha}$ where $m = \langle lambda, \alpha \rangle$. Here, $\alpha = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}, m = \lambda_1 - \lambda_2$.*

Weyl group W acts on $\mathbb{Z}[\Lambda]$ by

$$w \cdot \left(\sum_{\mu \in \Lambda} c_\mu e^\mu \right) = \sum_{\mu \in \Lambda} c_\mu e^{w(\mu)}$$

where $c_\mu \in \mathbb{Z}$.

Corollary 4.17. *ch_V is fixed by every $w \in W$.*

Proof.

$$m_\lambda(\mu) = m_\lambda(w(\mu)) \forall w \in W.$$

□

Proposition 4.18. *If $f \in \mathbb{Z}[\Lambda]$ is fixed by all $w \in W$ then f has a unique expansion as a finite linear combination of formal characters ch_λ for $\lambda \in \Lambda^+$.*

Proof. Write $f = \sum_{\lambda \in \Lambda} c_\lambda e^\lambda$ with $c_\lambda \in \mathbb{Z}$ all but finitely many c_λ's must be zero. Find a maximal $\lambda \in \Lambda^+$ with $c_\lambda \neq 0$, form $g = f - c_\lambda \mathrm{ch}_\lambda$, and argue that you may conclude by induction that g has desired expansion. □

Proposition 4.19. *Suppose V and W are both finite dimensional L-modules. Then $ch_{v \otimes w} = ch_V ch_W$.*

Proof. Straightforward calculation. □

4.7 Harish-Chandra's Theorem, More Formal Characters, Kostant's Formula

4.8 Harish-Chandra's Formula

Some technical proofs will just be outlined.

Let $\mathcal{Z}$ denote the center of the algebra $\mathcal{U}(L)$:

$$\mathcal{Z} := \{x \in \mathcal{U}(L) | xy = yx \forall y \in L\}$$

This is a commutative subalgebra, and each of the L-module is also a $\mathcal{U}(L)$-module and, by restriction, a $\mathcal{Z}$-module.

Consider the standard cyclic L-module

$$\mathcal{Z}(\lambda) = \mathcal{U}(L) \otimes_{\mathcal{U}(B)} D_\lambda$$

for some $\lambda \in H^*$, now viewed as a $\mathcal{Z}$-module.

If v^+ is a maximal vector in $\mathcal{Z}(\lambda)$ and $z \in \mathcal{Z}$, then

$$
\begin{aligned}
h \cdot z \cdot v^+ &= z \cdot h \cdot v^+ = \lambda(h) z \cdot v^+ \forall h \in H \\
x \cdot z \cdot v^+ &= z \cdot x \cdot v^+ = 0 \forall \alpha \in \Delta, \in L_\alpha
\end{aligned}
$$

Thus $z \cdot v^+$ is also a maximal vector of weight λ. Therefore $z \cdot v^+$ is a scalar multiple of v^+.

Define $\chi_\lambda : \mathcal{Z} \to \mathbb{F}$ to be the map with $z \cdot v^+ = \chi_\lambda(z) v^+ \forall z \in \mathcal{Z}$. Does not depend on choice of v^+, as all maximal vectors in $\mathcal{Z}(\lambda)$ are scalar multiples of each other.

Proposition 4.20. *χ_λ is an algebra homomorphism.*

Proof.

$$\chi_\lambda(z_1 z_2) v^+ = z_1 z_2 \cdot v^+ = z_1 \cdot (z_2 \cdot v^+) = \chi_\lambda(z_2) z_1 \cdot v^+ = \chi_\lambda(z_1)\chi(z_2) v^+.$$

□

Definition 4.21. *Call $\chi_\lambda : Z \to \mathbb{F}$ the **central character** of $\lambda \in H^*$ (or of $Z(\lambda)$).*

These central character χ_λ may coincide for different λ's, and Harish-Chandra's theorem will tell us precisely when this happens.

Proposition 4.22. *If $z \in \mathcal{Z}$ and $u \in Z(\lambda)$ is any vector then $z \cdot u = \chi_\lambda(z) u$.*

Proof. Since v^+ generates $Z(\lambda)$ and z commutes with all elements of L, the result follows. □

Corollary 4.23. *The action of $z \in \mathcal{Z}$ on any submodule of homomorphic image of $Z(\lambda)$ is by the scalar $\chi_\lambda(z)$.*

Definition 4.24. *Two elements $\lambda, \mu \in H^*$ are **linked** (by $w \in W$) if $\lambda + \sigma = w \cdot (\mu + S)$ where $\sigma := \frac{1}{2} \sum_{\alpha \in \Phi^+} \alpha$.*

In this situation we write $\mu \sim \lambda$. Given $\alpha \in \Phi^+$ choose $0 \neq x_\alpha \in L_\alpha$. Then there exists a unique $y_\alpha in L_{-\alpha}$ such that if $h_a := [x_\alpha, y_\alpha]$ then $\langle x_\alpha, y_\alpha, h_\alpha \rangle \cong \mathfrak{sl}_2(\mathbb{F})$ via $x_\alpha \mapsto \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, y_\alpha \mapsto \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, h_\alpha \mapsto \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

Proposition 4.25. *Let $\lambda \in \Lambda, \alpha \in \Delta, m = \langle \lambda, \alpha \rangle \in \mathbb{Z}$. If $m \geq 0$ then $y_\alpha^{m+1} \otimes_{\mathcal{U}(B)} v^+ \in Z(\lambda)$ is a maximal vector of weight $\lambda - (m+1)\alpha$. Here, $v^+ \in D_\lambda$ so $1 \otimes_{\mathcal{U}(B)} v^+$ generates $Z(\lambda)$.*

Proof. Formulas last time tell us that:

- For $\alpha \neq \beta$ in Δ: $[x_\beta, y_\alpha^{m+1}] = 0 \implies x_\beta \cdot y_\alpha^{m+1} \otimes v^+ = y_\alpha^{m+1} \otimes x_\beta v^+ = 0$.
- For any $\alpha, \beta \in \Delta$: $[h_\beta, y_\alpha^{m+1}] = -(m+1)\alpha(h_\beta) y_\alpha^{m+1} = \lambda(h_\beta) \implies h_\beta \cdot y_\alpha^{m+1} \otimes v^+ = -(m+1)\alpha(h_\beta) y_\alpha^{m+1} \otimes v^+ + y_\alpha^{m+1} \otimes h_\beta v^+ = (\lambda - (m+1)\alpha)(h_\beta) y_\alpha^{m+1} \otimes v^+$.

□

Corollary 4.26. *If $\lambda \in \Lambda, \alpha \in \Delta, \mu = r_\alpha \cdot (\lambda + \sigma) - \sigma$ where $r_\alpha \in W, x \mapsto x - \langle x, \alpha \rangle \alpha$ then $\chi_\lambda = \chi_\mu$.*

Proof. r_α sends $\alpha \mapsto -\alpha$ and permutes $\Phi^+ \backslash \{\alpha\}$, so $r_\alpha \sigma - \sigma = -\alpha$ and $\mu = r_\alpha \lambda - \alpha = \lambda - (\langle \lambda, \alpha \rangle + 1)\alpha$. We always have $\langle \lambda, \alpha \rangle \in \mathbb{Z}$.

If $\langle \lambda, \alpha \rangle \in \mathbb{Z}_{\geq 0}$ then previous proposition shows that $Z(\lambda)$ has maximal vector of weight μ.

As $z \in \mathcal{Z}$ acts on this vector by the scalar $\chi_\mu(z)$ and also $\chi_\lambda(z)$, we must hvae $\chi_\lambda = \chi_\mu$.

If $\langle \lambda, \alpha \rangle < 0$ then $\langle \mu, \alpha \rangle = \langle \lambda, \alpha \rangle - 2(\langle \lambda, \alpha \rangle + 1) = -\langle \lambda, \alpha \rangle - 2$ is ≥ 0 so we can apply proposition with μ in place of λ to deduce the same conclusion. □

Because $W = \langle r_\alpha | \alpha \in \Delta \rangle$ we can conclude:

Corollary 4.27. *If $\lambda \sim \mu$ where $\lambda \in \Lambda$, then $\chi_\lambda = \chi_\mu$.*

Theorem 4.28 (Harish-Chandra's Theorem)**.** *Let $\lambda, \mu \in H^*$. Then $\chi_\lambda = \chi_\mu$ if and only if $\lambda \sim \mu$.*

4.9 Outline of Proof

First part: already know that $\lambda \sim \mu \implies \chi_\lambda = \chi_\mu$ when $\lambda \in \Lambda$. We want to extend this to a statement allowing any $\lambda \in H^*$.

Construct PBW bases of $\mathcal{U}(L)$ and $\mathcal{U}(H)$ from the basis $\{h_\alpha | \alpha \in \Delta\} \sqcup \{x_\alpha, y_\alpha | \alpha \in \Phi^+\}$ for L, under any order putting all y_α's first, then the h_α's, then the x_α's.

Then we can define a linear map $\zeta : \mathcal{U}(L) \to \mathcal{U}(H)$ sending each PBW basis elements in $\mathcal{U}(A)$ to itself, every other PBW basis elements to 0.

Since $\Pi_{\alpha \in \Phi^+} y_\alpha^{i_\alpha} \Pi_{\alpha \in \Delta} h_\alpha^{k_\alpha} \Pi_{\alpha \in \Phi^+} x_\alpha^{j_\alpha}$ will either kill $v^+ \in Z(\lambda)$ if any $j_\alpha > 0$, or sent v^+ to lower weight space if all $j_\alpha = 0$ and any $i_\alpha > 0$. It follows that $\chi_\lambda(z) = \lambda(\zeta(z)) \forall z \in \mathcal{Z}$.

Now define another Lie algebra homomorphism $\eta : H \to \mathcal{U}(H)$ with $\eta(h_\alpha) = h_\alpha - 1 \forall \alpha \in \Delta$. This extends to an algebra automorphism $\eta : \mathcal{U}(H) \to \mathcal{U}(H)$. Define

$$\psi : \mathcal{Z} \xrightarrow{\zeta} \mathcal{U}(H) \xrightarrow{\eta} \mathcal{U}(H)$$

so $\psi = \eta \circ \zeta$.

We can write $\sigma = \sum_{\alpha\in\Delta} \lambda_\alpha$ as a sum over fundamental weights λ_α which have $\lambda_\alpha(h_\beta) = \begin{cases} 1 & \text{if } \alpha = \beta, \\ 0 & \text{if } \alpha \neq \beta. \end{cases}$ for $\alpha, \beta \in \Delta$. Then we have

$$(\lambda + \delta)(h_\alpha - 1) = (\lambda + \sigma)(h_\alpha) - (\lambda + \sigma)(1) = \lambda(h_\alpha)$$

so $(\lambda + \sigma)(\psi(z)) = \lambda(\zeta(z)) \forall z \in \mathcal{Z}, \gamma \in H^* \implies (\lambda + \sigma)(\psi(z)) = \chi_\lambda(z) \forall z \in \mathcal{Z}, \lambda \in H^*$.

Now check that $\psi(z)$ is W-invariant (using the easy case of theorem and properties of W-orbits in Λ) and use this to conclude that if $\lambda \sim \mu$ then $(\lambda+\sigma)(\psi(z)) = (\mu + \sigma)(\psi(z))$ and hence that $\chi_\lambda = \chi_\mu$ (for any $\lambda, \mu \in H^*$).

The other half of the theorem remains: if $\chi_\lambda = \chi_\mu$, then we need toshow that $\lambda \sim \mu$. This requires a more involved argument, see section 23.3 of textbook.

4.10 Applications of Harish-Chandra Theorem

We want to introduce **formal characters** for $Z(\lambda)$ and similar modules. Let $\mathcal{X}$ be the vector space of all formal $\mathbb{Z}$-linear combinations

$$\sum_{\lambda\in H^*} c_\lambda e^\lambda$$

($c_\lambda \in \mathbb{Z}, e^\lambda$ is a symbol)

which are finitely supported in the sense that there are finitely many $\lambda_1, ..., \lambda_k \in H^*$ such that $c_\lambda \neq 0 \implies \lambda \leq \lambda_i$ for some i. Then the **formal character** $\mathrm{ch}_{Z(\lambda)} := \sum_{\mu\in H^*} \dim(Z(\lambda))_\mu e^\mu$ belongs to $\mathcal{X}$, where $\lambda \leq \mu$ means $\mu - \lambda \in \mathbb{Z}_{\geq 0}$−span$\{\alpha \in \Delta\}$.

Proposition 4.29. *$\mathcal{X}$ is closed under usual mutliplication extending ring structure on $Z\not\geq$.*

We now have a well-defined notion of formal character

$$\mathrm{ch}_V := \sum_{\mu\in H^*} \dim(V_\mu e^\mu) \in \mathcal{X}$$

for any standard cuclic L-module V.

Let $p(\lambda)$ for $\lambda \in H^*$ be the number of functions $k : \Phi^+ \to \mathbb{Z}_{\geq 0}$ such that $\lambda + \sum_{\alpha\in\Phi^+} k(\alpha)\alpha = 0$. Clearly $p(\lambda) = 0$ unless $(-\lambda) \in \mathbb{Z}_{\geq 0}$-span$\{\alpha \in \Delta\}$.

Call p the **Kostant (partition) function** and identify $p \Longleftrightarrow \sum_{\lambda\in H^*} p(\lambda)e^\lambda \in \mathcal{X}$.

Also let $q = \prod_{\alpha \in \Phi^+}(e^{\frac{\alpha}{2}} - e^{-\frac{\alpha}{2}})$ and call this the **Weyl function**.

Finally set $f_\alpha = e^0 + e^{-\alpha} + e^{-2\alpha} + ... \in \mathcal{X}$ for $\alpha \in \Phi^+$.

Lemma 4.30. *1.* $p = \prod_{\alpha \in \Phi^+} f_\alpha$

2. $(e^0 - e^{-\alpha})f_\alpha = e^0$.

3. $q = e^\sigma \prod_{\alpha \in \Phi^+}(e^0 - e^{-\alpha})$ *where* $\sigma = \frac{1}{2}\sum_{\alpha \in \Phi^+} \alpha$.

Proof.

1. Holds by definition
2. Is basic algebra
3. Is clear

□

Lemma 4.31. *For any* $w \in W$ *it holds that* $wq = sgn(w)q$.

Proof. Suffices to show $r_\alpha q = -q$ for any $\alpha \in \Delta$. Easy enough: $r_\alpha q = r_\alpha(e^{\frac{\alpha}{2}} - e^{-\frac{\alpha}{2}})r_\alpha\left(\prod_{\beta \in \Phi^+ \setminus \{\alpha\}}(e^{\frac{\alpha}{2}} - e^{-\frac{\alpha}{2}})\right) = -q$. □

Lemma 4.32.

$$qpe^{-\sigma} = e^0 = 1$$

Proof.

$$\begin{aligned} qpe^{-\sigma} &= \prod_{\alpha \in \Phi^+}(e^0 - e^{-\alpha}) \cdot e^\sigma \cdot p \cdot e^{-\sigma} \\ &= \prod_{\alpha \in \Phi^+}(e^0 - \alpha^{-\alpha})pp \\ &= \prod_{\alpha \in \Phi^+}((e^0 - e^{-\alpha})f_\alpha) \\ &= \prod_{\alpha \in \Phi^+} e^0 \\ &= e^0 \\ &= 1 \end{aligned}$$

□

Lemma 4.33.

$$ch_{Z(\lambda)} = \sum_{\mu \in H^*} p(\mu - \lambda)e^\mu = e^\lambda p$$

Proof. Straightforward from properties of $Z(\lambda)$. □

Lemma 4.34.

$$q ch_{Z(\lambda)} = e^{\lambda+\sigma}$$

Proof. $qpe^{-\sigma} = e^0 = 1$ and $\text{ch}_{Z(\lambda)} = e^{\lambda}p$ so $q\text{ch}_{Z(\lambda)} = e^{\lambda+\sigma}qpe^{-\sigma} = e^{\lambda+\sigma}$. □

Want to express $\text{ch}_\lambda = \text{ch}_{V(\lambda)}$ are linear combination of $\text{ch}_{Z(\mu)}$'s.

Define M_λ (for $\lambda \in H^*$) to be the family of L-modules V such that

1. V is direct sum of its weight spaces.
2. Z-action on V is by scalar $\chi_\lambda(z)$.
3. $\text{ch}_V \in \mathcal{X}$.

M_λ is closed under taking submodules homomorphic images, direct sums, contains each standard cyclic module.

Corollary 4.35. *$M_\lambda = M_\mu$ if and only if $\lambda \sim \mu$.*

Lemma 4.36. *Suppose $0 \neq V \in M_\lambda$. Then V has a maximal vector.*

Proof. Since $\text{ch}_V \in \mathcal{X}$, for each weight μ of V, and each $\alpha \in \Phi^+$ there is a maximal $k \in \mathbb{Z}_{\geq 0}$ with $\mu + k\alpha$ still a weight. So we can find a weight μ for V such that $\mu + \alpha$ is not a weight $\forall \alpha \in \Phi^+$, and then any nonzero vector in the corresponding weight space is maximal. □

For $\lambda \in H^*$ let $\theta(\lambda) = \{\mu \in H^* | \mu < \lambda \text{ and } \mu \sim \lambda\}$.

Proposition 4.37. *Let $\lambda \in H^*$.*

1. *$Z(\lambda)$ has a composition.*
2. *Each composition factor of $Z(\lambda)$ is $\cong V(\mu)$ for some $\mu \in \theta(\lambda)$.*
3. *$V(\lambda)$ occurs as exactly one composition factor.*

Proof.

1. Nothing to prove if $Z(\lambda)$ is irreducible (then $Z(\lambda) = V(\lambda)$). Otherwise $Z(\lambda)$ has a proper nonzero submodule $V \in M_\lambda$. Since $\dim(Z(\lambda)_\lambda) = 1$, λ is not a weight of V. So by lemma, V has maximal vector, of some weight $\mu > \lambda$. V contains homomorphic image W of $Z(\mu)$, so $\chi_\lambda = \chi_\mu \implies \lambda \sim \mu \implies \mu \in \theta(\lambda)$. Continue inductively, repeating same argument applied to W and $Z(\lambda)/W$.
2. Each composite factor is in M_λ so has a maximal vector and is irreducible, so must be standard cyclic, have $\cong V(\mu)$ for some $\mu \in \theta(\lambda)$.
3. Clear since $\dim(Z(\lambda)_\lambda) = 1$.

□

Corollary 4.38. *Let $\lambda \in H^*$. Then $ch_{V(\lambda)} = \sum_{\mu\in\theta(\lambda)=\{v\in H^*|v\le\lambda, v\sim\lambda\}} c_\mu ch_Z(\mu)$ for some coefficients $c_\mu \in \mathbb{Z}$ with $c_\lambda = 1$.*

Proof. Proposition says we can write $\mathrm{ch}_{Z(\lambda)} = \mathrm{ch}_{V(\lambda)} + \sum_{\mu\in\theta(\lambda)} d_\mu \mathrm{ch}_{V(\mu)}$ where $d_\mu \in \mathbb{Z}_{\ge 0}$. Thus $\mathrm{ch}_{V(\lambda)} = \mathrm{ch}_{Z(\lambda)} - \sum_{\mu\in\theta(\lambda)} d_\mu \mathrm{ch}_{V(\mu)}$ and expanding the RHS recursively gives desired formula. □

Theorem 4.39 (Kostant's Formula)**.** *Let $\lambda \in \Lambda^+$ then*

$$m_\lambda(\mu) = \sum_{w\in W} sgn(w)p(\mu + \sigma - w(\lambda + \sigma)).$$

Proof. $\mathrm{ch}_\lambda = \sum_{\mu\in\theta(\lambda)} c_\mu \mathrm{ch}_{Z(\mu)}$ with $c_\lambda = 1$. Earlier lemmas tell us that $q\mathrm{ch}_\lambda = \sum_{\mu\in\theta(\lambda)} c_\mu e^{\mu+\sigma}$ and $w(q\mathrm{ch}_\lambda) = w(q)w(\mathrm{ch}_\lambda) = \mathrm{sgn}(w)q\mathrm{ch}_\lambda \forall w \in W$. But also $w\left(\sum_{\mu\in\theta(\lambda)} c_\mu e^{\mu+\sigma}\right) = \sum_{\mu\in\theta(\mu)} c_\mu e^{w(\mu+\sigma)}$ since $w \in W$. Permutes $\theta(\lambda)$ while $c_\lambda = 1$, deduce that $c_\mu = \mathrm{sgn}(w)$ if $w^{-1}(\mu + \sigma) = \lambda + \sigma$.

So

$$q\mathrm{ch}_\lambda = \sum_{w\in W} \mathrm{sgn}(w)e^{w(\lambda+\sigma)}.$$

By one of the earlier lemmas,

$$\begin{aligned}
\mathrm{ch}_\lambda &= qpe^{-\sigma}\mathrm{ch}_\lambda \\
&= pe^{-\sigma}\left(\sum_{w\in W} \mathrm{sgn}(w)e^{w(\lambda+\sigma)}\right) \\
&= p\sum_{w\in W} \mathrm{sgn}(w)e^{w(\lambda+\sigma)-\lambda} \\
&= \sum_{w\in W} \mathrm{sgn}(w)pe^{w(\lambda+\sigma)-\sigma}.
\end{aligned}$$

□

Corollary 4.40.

$$q = \sum_{w\in W} sgn(w)e^{w\sigma}$$

Proof. Take $\lambda = 0$ for

$$q\mathrm{ch}_\lambda = \sum_{w\in W} \mathrm{sgn}(w)e^{w(\lambda+\sigma)}.$$

□

Next time, we'll look at the Weyl character formula and Chevalley groups.

4.11 Weyl Character Formula, Chevalley Bases

4.12 Weyl Character Formula

Setup throughout: L is a finite-dimensional semisimple Lie algebra, defined over an algebraically closed field $\mathbb{F}$ with $\mathrm{char}(\mathbb{F}) = 0$, choose a Cartan subalgebra $H \subset L$, write $\Phi \subset H^*$ for corresponding root system, choose a set of simple roots Δ, positive roots ϕ^+, write

$$W = \langle r_\alpha | \alpha \in \Phi \rangle \subset \mathrm{GL}(H^*)$$

for Weyl group.

Last time: Let $\mathcal{Z}$ be the center of the universal enveloping algebra $\mathcal{L}$. For each $\lambda \in H^*$ we have a standard cyclic L-module $Z(\lambda) := \mathcal{U}(L) \otimes_{\mathcal{U}(B)} D_\lambda$.

Proposition 4.41. *There exists a unique algebra homomorphism $\chi_\lambda : \mathcal{Z} \to \mathbb{F}$ (called the **central character**) with $a \cdot u = \chi_\lambda(a)u \forall a \in \mathcal{Z}, u \in Z(\lambda)$.*

Harish-Chandra's theorem gives necessary and sufficient condition to have $\chi_\lambda = \chi_\mu$ for $\lambda, \mu \in H^*$. Namely: say that $\lambda, \mu \in H^*$ are **linked** (and write $\lambda \sim \mu$) if $\lambda + \delta$ and $\mu + \delta$ are in the same W-orbit where $\delta = \frac{1}{2}\sum_{\alpha \in \Phi^+} \alpha$.

Theorem 4.42. *For $\lambda, \mu \in H^*$ we have $\chi_\lambda = \chi_\mu$ if and only if $\lambda \sim \mu$.*

Define the **formal character** of any (standard cyclic) L-module V to be the formal expression $\mathrm{ch}_V = \sum_{\mu \in H^*} \dim(V_\mu) e^\mu$. Here, $V_\mu = \{v \in V | h \cdot v = \mu(h) v \forall h \in H\}$, and e^μ is just a formal symbol.

The reason for this notation is that we want to enable adding and multiplying characters like formal power series (or polynomials) under convention that $e^\lambda e^\mu = e^{\lambda+\mu}$.

For this kind of multiplication to be well the set of nonzero coefficients $c_\mu \neq 0$ in a character $\sum_{\mu \in H^*} c_\mu e^\mu$ must be finitely supported in some sense.

Relevant property: If V is standard cyclic then $\mathrm{ch}_V \in \mathcal{X}$ where $\mathcal{X}$ is the set of expressions $\sum_{\mu \in H^*} c_\mu e^\mu$ for which there are finitely many $\lambda_1, ..., \lambda_k \in H^*$ such that $c\mu \neq 0 \implies \mu < \lambda_i$ for some index i, where $\mu < \lambda$ means $\lambda - \mu \in \mathbb{Z}_{\geq 0}\mathrm{-span}\{\alpha \in \Delta\}$.

Recall that $V(\lambda)$ is unique irreducible quotient of $Z(\lambda)$. Now, given $\lambda, \mu \in H^*$ define

$$m_\lambda(\mu) = \dim(V(\lambda)\mu) = (\text{dim of } \mu \text{ weight space in } V(\lambda))$$

so that $\chi_{V(\lambda)} := \mathrm{ch}_\lambda = \sum_{\mu \in H^*} m_\lambda(\mu) e^\mu$.

Let $\mathrm{sgn} : W \to \{\pm 1\}$ be the unique group homomorphism with $\mathrm{sgn}(r_\alpha) = -1$.

Let

$$\begin{aligned} p(\lambda) &= \text{number of ways of writing } -\lambda \text{ as a sum of positive roots)} \\ &= (\text{number of functions } k : \phi^+ \to \mathbb{Z}_{\geq 0} \text{ such that } \lambda + \sum_{\alpha \in \Phi^+} \kappa(\alpha)\alpha = 0) \end{aligned}$$

Explicit formula but still less efficient than recursive, less explicit algorithms for computation.

In the proof of Kostant's formula, we encountered two identities: Let $q := \prod_{\alpha \in \Phi^+} (e^{\frac{\alpha}{2}} - e^{-\frac{\alpha}{2}}) = e^\delta \prod_{\alpha \in \Phi^+} (1 - e^{-\alpha}) \in \mathcal{X}$.

Then

1. $q\mathrm{ch}_\lambda = \sum_{w \in W} \mathrm{sgn}(w) e^{w(\lambda+\delta)}$
2. $q = \sum_{w \in W} \mathrm{sgn}(w) e^{w\delta}$

Substituting (2) into (1) gives the Weyl character formula.

Theorem 4.43 (Weyl Character Formula)**.** *If $\lambda \in \Lambda^+$ then*

$$\left(\sum_{w \in W} sgn(w) e^{w\delta} \right) ch_\lambda = \sum_{w \in W} sgn(w) e^{w(\lambda+\delta)}$$

Thus can compute ch_λ by doing "long division" in ring $\mathcal{X}$, but this is somewhat complicated in practice if $|\Delta|$ is large.

4.13 An Application

We can use this to find an explicit formula for $\deg(\lambda) := \dim(V(\lambda)) = \sum_{\mu \in H^*} m_\lambda(\mu)$ only defined for $\lambda \in \Lambda^+$.

Let $\mathcal{X}_0 \subset \mathcal{X}$ be $\mathbb{Z}$-span$\{e^\lambda | \lambda \in H^*\}$ (so formal characters with finite number of nonzero coefficients c_μ). Then we can define $\mathrm{eval} : \mathcal{X}_0 \to \mathbb{F}, \sum_\mu c_\mu e^\mu \mapsto \sum_\mu c_\mu$. Then $\deg(\lambda) = \mathrm{eval}(\mathrm{ch}_\lambda)$. Also $\mathrm{eval} : \mathcal{X}_0 \to \mathbb{F}$ is a ring homomorphism so $\mathrm{eval}(\mathrm{ch}_1, \mathrm{ch}_2) = \mathrm{eval}(\mathrm{ch}_1)\mathrm{eval}(\mathrm{ch}_2)$.

For $\alpha \in \Phi$ let $D_\alpha : \mathcal{X}_0 \to \mathcal{X}_0$ be the linear map with $D_\alpha e^\lambda = (\lambda, \alpha) e^\lambda$. This is a derivation $D_\alpha(fg) = fD_\alpha(g) + gD_\alpha(f)$ since

$$D_\alpha(e^\lambda e^\mu) = D_\alpha(e^{\lambda+\mu}) = (\lambda + \mu, \alpha)e^{\lambda+\mu} = D_\alpha(e^\lambda)e^\mu + e^\lambda D_\alpha(e^\mu).$$

Let $D = \prod_{\alpha\in\Phi^+} D_\alpha : \mathcal{X}_0 \to \mathcal{X}_0$ (no longer a derivation). Let $Q = \sum_{w\in W} \operatorname{sgn}(w)e^{w\delta}$ and $P = \sum_{w\in W} \operatorname{sgn}(w)e^{w(\delta+\lambda)}$, so the Weyl character formula is just $Q \cdot \mathrm{ch}_\lambda = P$.

We want to apply eval $\circ\, D$ to both sides of this since each D_α is derivation, $Q = e^{-\delta}\prod_{\alpha\in\Phi^+}(e^\alpha - 1)$, and $\operatorname{eval}(e^\alpha - 1) = 0$, one can show that this gives

$$\operatorname{eval}(D(Q))\operatorname{eval}(\mathrm{ch}_\lambda) = \operatorname{eval}(D(P)).$$

This implies that $\deg(\lambda) = \dfrac{\operatorname{eval}(D(P))}{\operatorname{eval}(D(Q))}$. Now observe that $\operatorname{eval}(D(e^\delta)) = \operatorname{eval}\left(\prod_{\alpha\in\Phi^+}(\delta,\alpha)\cdot e^\delta\right) = \prod_{\alpha\in\Phi^+}(\delta,\alpha)$. Similarly,

$$\operatorname{eval}(D(e^{w\delta})) = \prod_{\alpha\in\Phi^+}(w\delta,\alpha) = \prod_{\alpha\in\Phi^+}(\delta, w^{-1}\alpha) = (-1)^{\ell(w)}\prod_{\alpha\in\Phi^+}(\delta,\alpha) = \operatorname{sgn}(w)\prod_{\alpha\in\Phi^+}(\delta,\alpha).$$

Thus $\operatorname{eval}(D(Q)) = \sum_{w\in W}\operatorname{sgn}(w)\operatorname{eval}(D(e^{w\delta})) = \sum_{w\in W}\operatorname{sgn}(w)^2\prod_{\alpha\in\Phi^+}(\delta,\alpha) = |W|\cdot\prod_{\alpha\in\Phi^+}(\delta,\alpha)$.

Similarly, we derive $\operatorname{eval}(D(P)) = \sum_{w\in W}\operatorname{sgn}(w)\operatorname{eval}(D(e^{w(\lambda+\delta)})) = |W|\cdot\prod_{\alpha\in\Phi^+}(\delta+\lambda,\alpha)$. Thus, as $\deg(\lambda) = \dfrac{\operatorname{eval}(D(P))}{\operatorname{eval}(D(Q))}$.

Corollary 4.44 (Weyl Dimension Formula)**.** *If $\lambda\in\Lambda^+$ then*

$$\mathit{deg}(\lambda) := \dim(V(\lambda)) = \prod_{\alpha\in\Phi^+}\frac{(\delta+\lambda,\alpha)}{(\delta,\alpha)} = \prod_{\alpha\in\Phi^+}\frac{\langle\delta+\lambda,\alpha\rangle}{\langle\delta,\alpha\rangle}$$

where $\delta = \frac{1}{2}\sum_{\alpha\in\Phi^+}\alpha$ and $\langle\beta,\alpha\rangle = \dfrac{2(\beta,\alpha)}{(\alpha,\alpha)}$.

Example 4.45. *Consider type A_2. Let $\lambda_1,\lambda_2\in\mathbb{R}^2$ be such that*

$$\langle\lambda_i,\alpha_j\rangle = \begin{cases} 1 & \text{if } i = j,\\ 0 & \text{if } i \neq j.\end{cases}$$

which implies $\langle\delta,\alpha_1\rangle = \langle\delta,\alpha_2\rangle = 1, \langle\delta,\alpha_1+\alpha_2\rangle = 2 \implies \prod_{\alpha\in\Phi^+}\langle\delta,\alpha\rangle = 2$.

Every weight $\lambda\in\Lambda^+$ can be written uniquely as $\lambda = m_1\lambda_1 + m_2\lambda_2$. As $\langle\lambda+\delta,\alpha_1\rangle = m_1+1, \langle\lambda+\delta,\alpha_2\rangle = m_2+1, \langle\lambda+\delta,\alpha_1+\alpha_2\rangle = m_1+m_2+2$ we end up with $\mathit{deg}(\lambda) = \frac{1}{2}(m_1+1)(m_2+1)(m_1+m_2+2)$.

4.14 Chevalley Algebras and Groups

Keep same notation for $L, H, \Phi, \Delta, \Phi^+, W, \mathbb{F}$, etc.

Idea: there is a basis for L whose structure constants are all integers, so we can realize L as a Lie algebra generated by matrices over $\mathbb{Z}$. By extending scalars, one can construct L and its representations over arbitrary fields (rather than just the nic field $\mathbb{F}$).

Recall that $L = H \oplus \bigoplus_{\alpha\in\Phi} L_\alpha$ where

$$L_\alpha := \{X \in L | [h, X] = \alpha(h)X \forall h \in H\}.$$

Key properties:

1. $\dim(L_\alpha) = 1$
2. $\dim(H) = |\Delta|$
3. $[L_\alpha, L_{-\alpha}]$ is a 1-dimensional subspace of H spanned by a certain element h_α.

To be explicit: writing $\kappa(X, Y) = \mathrm{tr}(\mathrm{ad}(X)\mathrm{ad}(Y))$, we can define $t_\alpha = \dfrac{1}{\kappa(X,Y)}[X, Y]$ for any nonzero $X \in L_\alpha, Y \in L_{-\alpha}$. Then t_α is a unique element of H with $\kappa(t_\alpha, h) = \alpha(h)$. Then $h_\alpha = \dfrac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$.

Definition 4.46. *A **Chevalley basis** for L is a basis*

$$\{x_\alpha \in L_\alpha | \alpha \in \Phi\} \sqcup \{h_1, h_2, ..., h_n \in H\}$$

st.

1. $[x_\alpha, x_{-\alpha}] = h_\alpha \forall \alpha \in \Phi$.
2. *If* $\alpha, \beta, \alpha + \beta \in \Phi$ *and* $[x_\alpha, x_\beta] = c_{\alpha\beta}x_{\alpha+\beta}$ *for* $c_{\alpha\beta} \in \mathbb{F}$*, then* $c_{\alpha\beta} = -c_{-\alpha,-\beta}$.
3. Φ *has a simple system* Δ *such that* $\{h_1, ..., h_n\} = \{h_\alpha | \alpha \in \Delta\}$.

Proposition 4.47. *L has a Chevalley basis, and the coefficients $c_{\alpha\beta}$ corresponding to any such basis satisfy*

$$c_{\alpha\beta}^2 = q(r + q)\frac{(\alpha + \beta, \alpha + \beta)}{\beta, \beta}$$

where the α-root string through β is

$$\beta - r_\alpha, \beta - r_\alpha + \alpha, ..., \beta + q\alpha$$

Proof. Slightly technical but elementary algebraic exercise given what we know about the root space decomposition. □

Theorem 4.48. *If $\{x_\alpha \in L_\alpha | \alpha \in \Phi\} \sqcup \{h_1, h_2, ..., h_n \in H\}$ is a Chevalley basis for L the corresponding structure constants are all integers:*

1. $[h_i, h_j] = 0 \forall i, j$
2. $[h_i, x_\alpha] = \langle \alpha, \alpha_i \rangle x_\alpha$.
3. $[x_\alpha, x_{-\alpha}] = h_\alpha \in \mathbb{Z}$*-span*$\{h_1, ..., h_n\}$.
4. *If* $\alpha, \beta \in \Phi$ *are non-proportional then* $[x_\alpha, x_\beta] = \begin{cases} 0 & \text{if } \alpha + \beta \in \phi, \\ \pm(r+1)x_{\alpha+\beta} & \text{if } \alpha + \beta \notin \phi. \end{cases}$.

Proof. We have $[h_i, h_j] = 0$ since H is abelian. We have by definition $[h_i, x_\beta] = \alpha(h_i)x_\beta = \dfrac{2\beta(t_{\alpha_i})}{\kappa(t_{\alpha_i}, t_{\alpha_i})} x_\beta = \frac{2(\beta, \alpha_i)}{\alpha_i, \alpha_i} x = \langle \beta, \alpha_i \rangle x.$

We have $[x_\alpha, x_{-\alpha}] = h_\alpha$ by definition of Chevalley basis. We want to show that $h_\alpha \in \mathbb{Z}$-span$\{h_1, ..., h_n\}$. For this, let $\alpha^\vee = \frac{2\alpha}{(\alpha, \alpha)}$ then $\Phi^\vee \{\alpha^\vee | \alpha \in \Phi\}$ is a root system base $\Delta^\vee = \{\alpha^\vee | \alpha \in \Delta\}$, under the Killing form identification of H with $H^*, t_\alpha \iff \alpha$ and $h_\alpha \iff \alpha^\vee$, and each $\alpha^\vee$ is a $\mathbb{Z}$-linear combination of $\Delta^\vee$ so each h_α is a linear combination of $\{h_1, ..., h_n\}$.

Finally property follows from the proposition and ths lemma:

Lemma: If $\alpha, \beta \in \Phi$ are nonproportional and the α-string through β is $\beta - r_\alpha, ..., \beta + q\alpha$ then

$$r + 1 = \frac{q(\alpha + \beta, \alpha + \beta)}{\beta, \beta}.$$

Proof. Either case by case argument, considering rank two root systems, or see uniform geometric argument in textbook. □

About uniqueness: given H, Φ, etc. $\subset L$, a Chevalley basis is "almost unique": once a simple system $\Delta \subset \Phi$ is chosen, the h_i's are determined, but there is some flexibility in constructing the x_α's.

Fix a Chevalley basis for L. Let

$$\begin{aligned} L(\mathbb{Z}) &:= \mathbb{Z} - \text{span}\{\text{this Chevalley basis}\} \\ &= \mathbb{Z} - \text{span}\{x_\alpha(\alpha \in \Phi), h_1, h_2, ..., h_n\} \end{aligned}$$

This is the "Lie algebra over $\mathbb{Z}$" in the obvious sense, inheriting the Lie bracket from L.

For a prime finite field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ ($p > 0$ prime) can define $L(\mathbb{F}_p) := L(\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{F}_p$. More generally, for any field extension $\mathbb{F}_p \subseteq \mathbb{K}$ can define $L(\mathbb{F}) := L(\mathbb{F}_p) \otimes_{\mathbb{F}_p} \mathbb{K} = L(\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{K}$.

Proposition 4.49. *Then $L(\mathbb{K})$ is a Lie algebra over $\mathbb{K}$. Call this a* ***Chevalley algebra****.*

How do tensors $\otimes_{\mathbb{Z}}$ work?

You have symbols $a \otimes_{\mathbb{Z}} b$ but not for any $k \in \mathbb{Z}$ wehave $ka \otimes_{\mathbb{Z}} b = a \otimes_{\mathbb{Z}} kb$, plus usual tensor formalities.

Proposition 4.50. *Isomorphism class of $L(\mathbb{K})$ depends only on L not on the choice of Chevalley basis.*

Example 4.51. *If $L = \mathfrak{sl}_n(\mathbb{F})$ then $L(\mathbb{K})$ has the same multiplication table relative to usual standard basis, so $L(\mathbb{K}) \cong \mathfrak{sl}_n(\mathbb{K})$. Only change in this case is that $\mathfrak{sl}_n(\mathbb{K})$ may no longer be simple if $char(\mathbb{K})$ divides n.*

Proposition 4.52. *Let $\alpha \in \Phi$ and $m \in \mathbb{Z}_{\geq 0}$. Then $\dfrac{(ad(x_\alpha))^m}{m!}$ preserves $L(\mathbb{Z})$.*

Proof. Straightforward calculation. □

Definition 4.53. *Express $ad(x_\alpha)$ as a (nilpotent) matrix M_α relative to the Chevalley basis. Define $x_\alpha(t) = exp(tM_\alpha) = \sum_{k \geq 0} \dfrac{t^k}{k!} M_\alpha^k$. The* ***Chevalley group (of adjoint type)*** *is then*

$$G(\mathbb{K}) := \langle x_\alpha(t) | \alpha \in \Phi, t \in \mathbb{K} \rangle \subset GL(L(\mathbb{K}))$$

Here are some key facts about Chevalley groups:

Proposition 4.54.

1. *When $\mathbb{K}$ is finite the group $G(\mathbb{K})$ is finite.*
2. *When $\mathbb{K}$ is finite and L is simple, $G(\mathbb{K})$ is a finite simple group outside a short list of exceptional cases.*
3. *There is a uniform way of proving that these Chevalley groups are almost always simple, and this gives most of the families of finite simple groups.*
4. *As usual, isomorphism type of $G(\mathbb{K})$ depends on L but not on the choice of Chevalley basis.*