MATH 223A NOTES 2011
LIE ALGEBRAS

Contents

Date: December 22, 2011.

Introduction

These are lecture notes from a graduate course (Math 223a) taught in Fall 2011 at Brandeis University. I mostly followed Humphrey's book on Lie Algebras [5] since it covered the material systematically. However, I used several other books when I felt that more details were needed or a more geometric approach was easier at least for me to understand and present.

I used notation and examples from Fulton and Harris [4] who give a much more geometric approach with lots of pictures. Erdmann and Wildon [2] was also an excellent elementary reference which my students appreciated. I used it to give more details about each type of root system. Procesi's book [6] was also very helpful and I used it for one lecture. Tauvel and Yu [7] give a very nice geometric proof of the fact that Cartan subalgebras are all conjugate. But perhaps this was too heavy for the students. Finally, I used Fulton's book on Young Tableaux [3] (together with [4]) to explain Schur polynomials and their relation to the Weyl character formula. Finally, I referred to Carter's book [1] for the characters of Verma modules (not in the notes).

1. Basic Concepts

1.1. **Definition of Lie algebra.**

Definition 1.1.1. By an *(nonassociative) algebra* over a field F we mean a vector space A together with an F-bilinear operation $A \times A \to A$ which is usually written $(x, y) \mapsto xy$.

The adjective "nonassociative" means "not necessarily associative". An *associative algebra* is an algebra A whose multiplication rule is associative: $x(yz) = (xy)z$ for all $x, y, z \in A$. The existence of a unit 1 is not assumed.

Definition 1.1.2. Let L be a vector space over a field F. Then a bilinear operation $[\cdot\cdot] : L \times L \to L$ sending (x, y) to $[xy]$ is called a *bracket* if it satisfies the following two conditions.

[L1] $[xx] = 0$ for all $x \in L$.
[L2] (*Jacobi identity*) $[x[yz]] + [y[zx]] + [z[xy]] = 0$ for all $x, y, z \in L$.

A vector space L with a bracket $[\cdot\cdot]$ is called a *Lie algebra*. This is an example of a nonassociative algebra.

Let us analyze the two conditions. Condition [L1] implies:

[L1′] $[xy] = -[yx]$ for all $x, y \in L$. (Bracket is *skew commutative*.)

Proof: $[(x + y)(x + y)] = 0 = [xx] + [xy] + [yx] + [yy] = [xy] + [yx]$.

Conversely, if the characteristic of the field F is not equal to 2 then [L1′] implies that $2[xx] = 0$ implies [L1]. So, [L1] and [L1′] are equivalent when *char* $F \neq 2$.

The second condition [L2] can be rewritten as follows:

$$[x[yz]] = [[xy]z] + [[zx]y].$$

The term $[[zx]y]$ prevents L from being associative. Since z, x, y are arbitrary we obtain:

Proposition 1.1.3. *A Lie algebra is associative if and only if* $[[LL]L] = 0$.

The notation $[[LL]L]$ indicates the vector subspace of L generated by all expressions $[[xy]z]$.

Definition 1.1.4. A *(Lie) subalgebra* of a Lie algebra L is defined to be a vector subspace K so that $[KK] \subseteq K$.

For example, $[LL]$ is always a Lie subalgebra of L.

Definition 1.1.5. A *homomorphism* of Lie algebras is a linear map $\varphi : L \to L'$ so that $\varphi([xy]) = [\varphi(x)\varphi(y)]$ for all $x, y \in L$.

1.2. **Examples.**

Example 1.2.1. The simplest example of a Lie algebra is given by letting $[xy] = 0$ for all $x, y \in L$ where L is any vector space over F. All conditions are clearly satisfied. A Lie algebra satisfying this condition (usually written as $[LL] = 0$) is called *abelian*.

The word "abelian" comes from one standard interpretation of the bracket. Suppose that A is an associative algebra over F. Then the *commutator* $[xy]$ is defined by $[xy] = xy - yx$. This is easily seen to be a bracket and is also called the *Lie bracket* of the associative algebra.

Example 1.2.2. Suppose that V is any vector space over F. We define $\mathfrak{gl}(V)$ to be the Lie algebra of all F-linear endomorphisms of V under the Lie bracket operation. A Lie subalgebra of $\mathfrak{gl}(V)$ is called a *linear Lie algebra*.

Definition 1.2.3. A *representation* of the Lie algebra L is defined to be a Lie algebra homomorphism $L \to \mathfrak{gl}(V)$ for some vector space V. The representation is called *faithful* if this homomorphism is injective: $L \hookrightarrow \mathfrak{gl}(V)$.

1.2.1. *linear Lie algebras.* There is a well-known theorem (due to Ado in characteristic 0 and Iwasawa in characteristic p) what every finite dimensional Lie algebra has a faithful finite dimensional representation. I.e., it is isomorphic to a linear Lie algebra. So, our finite dimensional examples are all linear. What are the finite dimensional linear Lie algebras?

If $V = F^n$ then $\mathfrak{gl}(V)$ is denoted $\mathfrak{gl}(n, F)$. This is the vector space of all $n \times n$ matrices with coefficients in F with Lie bracket given by commutator: $[xy] = xy - yx$. A subalgebra is given by a subset of $\mathfrak{gl}(n, F)$ which is closed under this bracket and under addition and scalar multiplication.

Example 1.2.4. Let $\mathfrak{sl}(n, F) \subseteq \mathfrak{gl}(n, F)$ denote the set of all $n \times n$ matrices with trace equal to zero.

(1) $\mathrm{Tr}([xy]) = \mathrm{Tr}(xy) - \mathrm{Tr}(yx) = 0$. So, $\mathfrak{sl}(n, F)$ is closed under $[\,]$.
(2) $\mathrm{Tr}(x + y) = \mathrm{Tr}(x) + \mathrm{Tr}(y) = 0$.
(3) $\mathrm{Tr}(ax) = a\,\mathrm{Tr}(x) = 0$

Therefore, $\mathfrak{sl}(n, F)$ is a linear Lie algebra.

Proposition 1.2.5. *Suppose that $f : V \times V \to F$ is a bilinear form. Then the set of all $x \in \mathfrak{gl}(V)$ so that*

$$f(x(v), w) + f(v, x(w)) = 0$$

for all $v, w \in V$ is a Lie subalgebra of $\mathfrak{gl}(V)$ which we denote $\mathfrak{o}(V, f)$

Proof. It is clear that $\mathfrak{o}(V, f)$ is a vector subspace since the defining equation is linear in x. The following calculation shows that it is closed under Lie bracket.

$$f(xy(v), w) + f(y(v), x(w)) = 0$$
$$f(yx(v), w) + f(x(v), y(w)) = 0$$
$$f(v, xy(w)) + f(x(v), y(w)) = 0$$
$$f(v, yx(w)) + f(y(v), x(w)) = 0$$

If we take the alternating sum (+ - + -) of these equations we see that

$$f([xy](v), w) + f(v, [xy](w)) = 0$$

□

Example 1.2.6. Particular examples of the above definition are as follows.

(1) Suppose that f is a nondegenerate symmetric bilinear form on V. Then $\mathfrak{o}(V, f)$ is called the *orthogonal Lie algebra* relative to f.

(2) Suppose that f is a nondegenerate skew symmetric form on V: $f(v, v) = 0$ for all $v \in V$. (If *char* $F \neq 2$ this is equivalent to the condition that $f(v, w) = -f(w, v)$ for all v, w.) In this case $\dim V = 2n$ (even) and $\mathfrak{o}(V, f)$ is called the *symplectic Lie algebra* relative to f.

We will look at these examples in more detail later.

Example 1.2.7. Other easy examples of linear Lie algebras are:

(1) $\mathfrak{t}(n, F) \subseteq \mathfrak{gl}(n, F)$, the set of upper triangular $n \times n$ matrices over F

(2) $\mathfrak{n}(n, F) \subseteq \mathfrak{t}(n, F)$, the set of strictly upper triangular matrices (with 0 on the diagonal).

(3) $\mathfrak{d}(n, F) \subseteq \mathfrak{t}(n, F)$, the set of diagonal $n \times n$ matrices with coefficients in F.

1.3. Derivations.

Definition 1.3.1. Suppose that A is a nonassociative algebra over F. Then a *derivation* on A is a linear function $\delta : A \to A$ so that

$$\delta(xy) = \delta(x)y + x\delta(y)$$

for all $x, y \in A$. The set of all derivations on A is denoted $Der(A)$.

Proposition 1.3.2. *$Der(A)$ is a Lie subalgebra of $\mathfrak{gl}(A)$.*

Proof. □

Go back to the definition of a Lie algebra. Using the skew symmetry condition [L1′], Condition [L2] can be rephrased as:

$$[z[xy]] = [[zx]y] + [x[zy]$$

In other words, the bracket by z operation $\mathrm{ad}_z(\cdot) = [z(\cdot)]$ satisfies:

$$\mathrm{ad}_z[xy] = [\mathrm{ad}_z(x)y] + [x\mathrm{ad}_z(y)]$$

So any Lie algebra *acts on itself by derivations.* This gives a homomorphism:

$$\mathrm{ad} : L \to Der(A)$$

called the *adjoint representation.*

1.4. Abstract Lie algebras.

We could simply start with the definition and try to construct all possible Lie algebras. Take $L = F^n$.

$n = 1$: Show that all one dimensional Lie algebras are abelian.

$n = 2$: If $L = F^2$ there are, up to isomorphism, exactly two examples.

$n = 3$: Example: Take $L = \mathbb{R}^3$ and take the Lie bracket to be the *cross product.*

$$[xy] = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}$$

Verify the Jacobi identiy. Which Lie algebra is this?

2. Ideals and homomorphisms

2.1. Ideals.

Definition 2.1.1. An *ideal* in a Lie algebra L is a vector subspace I so that $[LI] \subseteq I$. In other words, $[ax] \in I$ for all $a \in L, x \in I$.

Example 2.1.2. (1) 0 and L are always ideals in L.
(2) If L is abelian then every vector subspace is an ideal.
(3) $[LL]$ is an ideal in L called the *derived algebra* of L.

Proposition 2.1.3. *Every 2-dimensional Lie algebra contains a 1-dimensional ideal.*

Proof. As we saw last time, there are only two examples of a 2-dimensional Lie algebra: Either the basis elements commute, in which case L is abelian, or they don't commute, in which case $[LL]$ is 1-dimensional. In both cases, L has a 1-dimensional ideal. □

Proposition 2.1.4. *The kernel of an homomorphism of Lie algebras $\varphi : L \to L'$ is an ideal in L. (The image is a subalgebra of L'.) Conversely, for any ideal $I \subseteq L$, L/I is a Lie algebra and I is the kernel of the quotient map $L \to L/I$.*

Proof. If $x \in \ker \varphi$ and $a \in L$ then $\varphi[ax] = [\varphi(a)\varphi(x)] = [\varphi(a)0] = 0$. So, $[ax] \in \ker \varphi$. Conversely, if $I \subset L$ is an ideal, $a \in L, x \in I$ then

$$[(a + I)(x + I)] \subseteq [ax] + [aI] + [Ix] + [II] \subseteq [ax] + I$$

So, the bracket is well-defined in L/I, $[\varphi(a)\varphi(x)] = \varphi[ax]$ and $I = \ker \varphi$. □

Proposition 2.1.5. *Any epimorphism (surjective homomorphism) of Lie algebras $\varphi : L \to L'$ gives a 1-1 correspondence between ideals I' in L' and ideals I of L containing $\ker \varphi$.*

Proof. The correspondence is given by $I' = \varphi(I)$. This is an ideal since $[L', I'] = [\varphi(L), \varphi(I)] = \varphi[LI] \subseteq \varphi(I) = I'$ and $I = \varphi^{-1}(I')$ which is an ideal since it is the kernel of $L \to L' \to L'/I'$. □

Example 2.1.6. The *center* of a Lie algebra L is defined to be

$$Z(L) = \{x \in L \mid [xL] = 0\}.$$

Then $Z(L)$ is an ideal in L. $Z(L)$ also the kernel of the adjoint representation $\mathrm{ad} : L \to \mathfrak{gl}(L)$. If $I \subseteq L$ is an ideal, then the adjoint representation restricts to a representation on I: $\mathrm{ad} : L \to \mathfrak{gl}(I)$ and the kernel of this is the *centralizer* $Z_L(I)$ of I in L. This is an ideal in L since it is a kernel.

Example 2.1.7. Consider the Lie algebra $\mathfrak{n}(3, F)$. This is 3-dimensional with basis x_{12}, x_{23}, x_{13} and the only nontrivial bracket is $[x_{12}x_{23}] = x_{13}$. Then $L = \mathfrak{n}(3, F)$ has the property that $[LL] = Z(L)$ is one-dimensional. Conversely, it is easy to see that any 3-dimensional Lie algebra with $[LL] = Z(L)$ must be isomorphic to $\mathfrak{n}(3, F)$.

Example 2.1.8. $\mathrm{Tr} : \mathfrak{gl}(n, F) \to F$ is a homomorphism of Lie algebras (with F the abelian Lie algebra) since $\mathrm{Tr}[xy] = \mathrm{Tr}(xy) - \mathrm{Tr}(yx) = 0 = [\mathrm{Tr}(x), \mathrm{Tr}(y)]$. Therefore, $\mathfrak{sl}(n, F)$ is an ideal in $\mathfrak{gl}(n, F)$.

Exercise 2.1.9. What is the center of $\mathfrak{sl}(2,F)$?

A Lie algebra is called *simple* if it is nonabelian and has no nontrivial proper ideals.

Theorem 2.1.10. *$\mathfrak{sl}(2,F)$ is simple if char $F \neq 2$.*

$\mathfrak{sl}(2,F)$ is a key example of a Lie algebra which you need to understand very thoroughly.

Proof. Since $L = \mathfrak{sl}(2,F)$ is 3-dimensional we just need to show that it has no ideals of dimension 1 and no ideals of dimension 2.

The standard basis for $\mathfrak{sl}(2,F)$ is given by

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and computation gives:

$$[hx] = 2x, \quad [hy] = -2y, \quad [xy] = h$$

This proves that $[LL] = L$. This implies that L has no 2-dimensional ideal I since then L/I would be 1-dimensional and thus abelian which would imply that $[LL] \subseteq I$.

This also implies that $L = \mathfrak{sl}(2,F)$ has no 1-dimensional ideal I since, in that case, L/I would be 2-dimensional and thus would have a 1-dimensional ideal by Proposition 2.1.3 and this would correspond to a 2-dimensional ideal in L by Proposition 2.1.5. Therefore, $\mathfrak{sl}(n,F)$ is simple. □

Proposition 2.1.11. *Every simple Lie algebra is linear.*

Proof. The center of L is trivial. So, the adjoint representation gives an embedding $\text{ad} : L \hookrightarrow \mathfrak{gl}(L)$. □

2.2. **nilpotent elements and automorphisms.** First, suppose that *char* $F = 0$ and δ is a nilpotent derivation of any nonassociative algebra A. I.e., $\delta^k = 0$ for some k. Then we claim that

$$\exp\delta = \sum_{i=0}^{k-1} \frac{\delta^i}{i!}$$

is an automorphism of A. It is easy to see that this is a linear automorphism of A since it has the form $1 + \eta$ where η is nilpotent. So, the inverse is $1 - \eta + \eta^2 - \cdots$ which is a finite sum. The following lemma shows that $\exp(-\delta)$ is the inverse of $\exp\delta$.

Lemma 2.2.1. *Suppose that char $F = 0$ and f, g are commuting nilpotent endomorphism of some vector space V (or, more generally, commuting elements of any associative algebra with unity) then $\exp(f+g) = \exp(f)\exp(g)$. In particular, $\exp(f)$ is a linear automorphism of V with inverse $\exp(-f)$.*

Proof. Since f, g commute we have:

$$(f+g)^n = \sum_{i=0}^{n} \binom{n}{i} f^i g^{n-i}$$

Since *char* $F = 0$ we can divide both sides by $n!$ to get

$$\frac{(f+g)^n}{n!} = \sum_{i+j=n} \frac{f^i}{i!}\frac{g^j}{j!}$$

Note that for sufficiently large n, all terms are zero since f, g are nilpotent. Thus we can sum over all $n \geq 0$ to get

$$\exp(f+g) = \exp(f)\exp(g)$$

as claimed. □

To see that $\exp \delta$ is a homomorphism of algebras, we can use the following trick:

$$\delta(xy) = \delta \circ \mu(x, y)$$

where $\mu : A \times A \to A$ is multiplication. Then $\delta \circ \mu = \mu \circ (\delta_1 + \delta_2)$ where

$$\delta_1(x, y) = (\delta x, y), \quad \delta_2(x, y) = (x, \delta y)$$

Since δ_1, δ_2 commute, we can use the lemma to get:

$$\exp \delta(xy) = \exp \delta \circ \mu(x, y) = \mu \circ \exp(\delta_1 + \delta_2)(x, y) = \mu \circ \exp(\delta_1)\exp(\delta_2)(x, y)$$

$$= \exp \delta(x) \exp \delta(y)$$

Definition 2.2.2. Suppose that x is an element of a Lie algebra L so that ad_x is nilpotent. Then $\exp \mathrm{ad}_x$ is called an *inner automorphism* of L.

Proposition 2.2.3. *Suppose that L is the Lie algebra of an associative algebra A with unity* 1. *Let $x \in A$ be a nilpotent element. Then:*

(1) $\exp x = \sum x^k/i!$ *is a unit in A.*
(2) ad_x *is a nilpotent endomorphism of L.*
(3) $\exp \mathrm{ad}_x$ *is conjugation by* $\exp x$.

Proof. By the Lemma, $\exp(-x)$ is the inverse of $\exp x$. The other statement follow from the following trick: Write $\mathrm{ad}_x = \lambda_x + \rho_{-x}$ where λ_x is "left multiplication by x" and ρ_{-x} is "right multiplication by $-x$": $\lambda_x(y) = xy$ and $\rho_{-x}(y) = -yx$. Since left and right multiplication are commuting operations,

$$\exp \mathrm{ad}_x = \exp(\lambda_x + \rho_{-x}) = \exp \lambda_x \exp \rho_{-x}$$

But clearly, $\exp \lambda_x = \lambda_{\exp x}$ and $\exp \rho_{-x} = \rho_{\exp(-x)}$. So,

$$\exp \mathrm{ad}_x(y) = (\exp x)y(\exp(-x)) = (\exp x)y(\exp(x)^{-1})$$

□

2.3. **Exercises.** What about derivations in characteristic p?

(1) Show that δ^p is a derivation.
(2) For the Lie algebra of an associative algebra over a field of characteristic p show that $\mathrm{ad}_x^p = \mathrm{ad}_{x^p}$.
(3) (in any characteristic) Let $\varphi : V \times V \to F$ be a nondegenerate skew-symmetric bilinear pairing. Then the *Heisenberg algebra* of f is given by $L = V \oplus F$ with $[(v, a)(w, b)] = (0, f(v, w))$. Show that $[LL] = Z(L)$ is 1-dimensional and that all Lie algebras with this property are Heisenberg algebras.
(4) If I, J are ideals in L then show that $I \cap J$ and $[IJ]$ are ideals. How are these related? Give an example where these are different.
(5) The *normalizer* $N_L(K)$ of a subalgebra $K \subseteq L$ is the set of all $x \in L$ so that $[xK] \subseteq K$. Note that K is an ideal In $N_L(K)$. Show that $\delta(n, F)$ and $\mathfrak{t}(n, F)$ are self-normalizing in $\mathfrak{gl}(n, F)$.

3. Nilpotent and solvable Lie algebras

I can't find my book. The following is from Fulton and Harris [4].

Definition 3.0.1. A Lie algebra is *solvable* if its iterated derived algebra is zero. In other words, $D^kL = 0$ where $DL = [LL], D^2L = [(DL)(DL)] = [[LL][LL]]$, etc. This is a recursive definition: The *k-th derived algebra* of L is the $k-1$st derived algebra of $[LL]$.

Definition 3.0.2. A Lie algebra is *nilpotent* of class k if

$$\mathrm{ad}_L^k(L) = \underbrace{[L \cdots [L[L[L}_{k} L]]] \cdots] = 0$$

for some k.

Note that if L is nilpotent of class k then $\mathrm{ad}_x^k = 0$ for all $x \in L$ since $\mathrm{ad}_x^k(y) = [x[x[x \cdots [xy] \cdots]]] \in [L[L[L \cdots [LL] \cdots]]] = 0$. Every element of L is ad-nilpotent.

3.1. Engel's Theorem.

This is converse of the above statement.

Theorem 3.1.1 (Engel)**.** *If L is a finite dimensional Lie algebra in which every element is ad-nilpotent then L is nilpotent.*

We prove this theorem in a sequence of lemmas. The first lemma allows us to assume that L is a linear Lie algebra.

Lemma 3.1.2. *Suppose the image of the adjoint representation* ad : $L \to \mathfrak{gl}(L)$ *is a nilpotent subalgebra of $\mathfrak{gl}(L)$ of class k then L is nilpotent of class $\leq k+1$.*

Proof. The kernel of the adjoint representation is the center $Z = Z(L)$. If L/Z is nilpotent of class k then $\mathrm{ad}_{L/Z}^k(L/Z) =$ is the same as $\mathrm{ad}_L^k(L) \subseteq Z$. But $[LZ] = 0$ so $\mathrm{ad}_L^{k+1}(L) = [L\mathrm{ad}_L^k(L)] \subseteq [LZ] = 0$. □

If every element of L is ad-nilpotent then the image of the adjoint representation consists of nilpotent endomorphism of L considered as a vector space.

Lemma 3.1.3. *Suppose that L is a subalgebra of $\mathfrak{gl}(V)$ where V is a nonzero finite dimensional vector space over F. Suppose that every element of L is a nilpotent endomorphism of V. Then there exists a nonzero element $v \in V$ so that $x(v) = 0$ for all $x \in L$.*

Some people call this Engle's Theorem since it is the key step in the proof of the theorem. To see that the lemma implies the theorem, let K be the subspace of V spanned by the vector v. Then the action of L on V induces an action on V/K which is nilpotent. So, the image of L in $\mathfrak{gl}(V/K)$ is a nilpotent Lie algebra of class, say k. This implies that $L^k(V) \subseteq K$. So, $L^{k+1}(V) = 0$ making L nilpotent of class $k+1$.

Proof of key lemma. The proof is by induction on the dimension of L. If L is one-dimensional, the lemma is clear. So, suppose $\dim L \geq 2$. Let J be a maximal proper subalgebra of L.

<u>Claim 1</u>: J is an ideal of codimension 1 in L.

Pf: J acts, by the adjoint action on L and this action leaves J invariant. Therefore, J act on the quotient L/J. By induction on dimension, there is a nonzero vector $y+J \in L/J$ so that ad_J kills this vector. In other words, $[Jy] \subseteq J$. This implies that J and y span a subalgebra of L with dimension one more than the dimension of J. By maximality of J, this subalgebra is equal to L. So, J has codimension 1. Also, J is an ideal since $[LJ] \subseteq [JJ] + [yJ] \subseteq J$.

Since J is smaller than L, there is a nonzero vector $v \in V$ so that $J(v) = 0$. Let W be the set of all $v \in V$ so that $J(v) = 0$. Then W is a nonzero vector subspace of V. Take $y \in L, y \notin J$.

<u>Claim 2</u>: $y(W) \subseteq W$.

Pf: We need to show that, for any $x \in J$ and $w \in W$, $xy(w) = 0$. This follows from the following calculation:

$$xy(w) = [xy](w) + yx(w) = 0 + 0 = 0$$

since $[xy] \in J$ and $J(w) = 0$.

Since y is nilpotent and sends W into W, there is some nonzero $w \in W$ so that $y(w) = 0$. Since $J(w) = 0$, we conclude that $L(w) = 0$ proving the lemma. □

Exercise 3.1.4.
(1) Show that $\mathfrak{n}(n, F)$ is nilpotent of class $n-1$. Hint: there is a filtration of $V = F^n$ by vector subspaces

$$0 = V_0 \subset V_1 \subset V_2 \subset \cdots V_n$$

so that $x(V_i) \subseteq V_{i-1}$ for all $x \in \mathfrak{n}(n, F)$. Such a sequence of subspaces of V is called a *flag*.

(2) Show that, for any linear Lie algebra $L \subseteq \mathfrak{gl}(V)$, the existence of a flag in V with the property that $x(V_i) \subseteq V_{i-1}$ implies that L is a subalgebra of $\mathfrak{n}(n, F)$ up to isomorphism. (Assume $\dim V = n$ is finite.)

(3) Show that any nilpotent subalgebra of $\mathfrak{gl}(V)$ is isomorphic to a subalgebra of $\mathfrak{n}(n, F)$ where $n = \dim V$.

3.2. **Lie's theorem.** When we go to solvable Lie algebras we need the ground field F to be algebraically closed of characteristic 0. So, we might as well assume that $F = \mathbb{C}$.

Then we have the following theorem whose statement and proof are similar to the statement and proof of the key lemma for Engel's Theorem.

Theorem 3.2.1 (Lie)**.** *Suppose that $L \subseteq \mathfrak{gl}(V)$ is a solvable linear Lie algebra over $\mathbb{C}$. Then there exists a nonzero vector $v \in V$ which is a simultaneous eigenvector of every element of L, i.e., $x(v) = \lambda(x)v$ for every $x \in L$ where $\lambda(x) \in \mathbb{C}$.*

Remark 3.2.2. Before proving this we note that, the function $\lambda : L \to \mathbb{C}$ is a linear map.

(1) $ax(v) = a\lambda(x)v$. So, $\lambda(ax) = a\lambda(x)$.
(2) $(x + y)(v) = x(v) + y(v) = \lambda(x)v + \lambda(y)v = (\lambda(x) + \lambda(y))(v)$. So, $\lambda(x + y) = \lambda(x) + \lambda(y)$.

Furthermore, note that if $\lambda : L \to \mathbb{C}$ is a linear map then the equation $x(v) = \lambda(x)(v)$ is a linear equation in x. Therefore, if this equation holds for all x in a spanning set for L then it holds for all x in L.

Proof. The proof is by induction on the dimension of L. If L is one dimensional then we are dealing with one endomorphism x of V. Since $\mathbb{C}$ is algebraically closed, the characteristic polynomial of x has a root $\lambda \in \mathbb{C}$ and a corresponding eigenvector v. ($x(v) = \lambda v$. So, $\lambda_x = \lambda$.) So, suppose that $\dim L \geq 2$.

The next step is to find a codimension one ideal J in L. Since $[LL] \subsetneq L$, this is easy. Take any codimension one vector subspace of $L/[LL]$ and let J be the inverse image of this in L.

By induction there is a nonzero vector $v \in V$ and a linear map $\lambda : J \to \mathbb{C}$ so that $x(v) = \lambda(x)v$ for all $x \in J$. Let W be the set of all $v \in V$ with this property (for the same linear function λ). Let $y \in L, y \notin J$.

<u>Claim</u>: $y(W) \subseteq W$.

Suppose for a moment that this is true. Then, we can find an eigenvalue $\lambda(y)$ and eigenvector $w \in W$ so that $y(w) = \lambda(y)w$. By the remark, this extended linear map λ satisfies the desired equation. The Claim is proved more generally in the following lemma. □

Lemma 3.2.3. *Suppose that J is an ideal in L, V is a representation of L and $\lambda : J \to F$ is a linear map. Assume char $F = 0$. Let*

$$W = \{v \in V \mid x(v) = \lambda(x)v \ \forall x \in J\}$$

Then $y(W) \subseteq W$ for all $y \in L$.

Proof. Define W_i recursively as follows. $W_0 = 0$ and

$$W_{k+1} = \{v \in V \mid x(v) - \lambda(x)(v) \in W_k \ \forall x \in J\}$$

Then $W = W_1 \subseteq W_2 \subseteq \cdots$ since V is finite dimensional we must have $W_k = W_{k+1}$ for some k. This means that $(x - \lambda(x))^k = 0$ on W_k. So, for any $x \in J$, the matrix of x as an endomorphism of W_k is upper triangular with $\lambda(x)$ on the diagonal and we have

$$\mathrm{Tr}(x|W_k) = \lambda(x) \dim W_k$$

<u>Claim</u> For any $y \in L, y(W_k) \subseteq W_{k+1} = W_k$.

Suppose for a moment that this is true. Then, for any $x \in J, y \in L$ we have $\mathrm{Tr}([xy]|W_k) = 0 = \lambda[xy] \dim W_k$. Since *char* $F = 0$ this implies $\lambda[xy] = 0$. We can now show that $y(w) \in W$ for all $y \in L, w \in W$:

$$x(y(w)) = yx(w) + [xy](w) = \lambda(x)y(w) + \lambda[xy](w) = \lambda(x)y(w)$$

Thus it suffices to prove the claim.

Proof of Claim: $y(W_i) \subseteq W_{i+1}$. To prove this we must show that, for any $w \in W_i, x \in J$ we have $(x - \lambda(x))y(w) \in W_i$. This is a calculation similar to the one above:

$$x(y(w)) - \lambda(x)y(w) = y(x - \lambda(x))(w) + [xy](w) \in y(W_{i-1}) + W_i \subseteq W_i$$

proving the claim by induction. (It is clear when $i = 0$.) □

Exercise 3.2.4. (1) Show that $\mathfrak{t}(n, F)$ is solvable.

(2) Show that any solvable subalgebra of $\mathfrak{gl}(n, \mathbb{C})$ is, up to isomorphism, isomorphic to a subalgebra of $\mathfrak{t}(n, \mathbb{C})$.

(3) If J is an ideal in L then show that L is solvable if and only if J and L/J are solvable.

(4) Prove that the following are equivalent.
   (a) L is solvable.
   (b) L has a sequence of ideals $L \supset J_1 \supset J_2 \cdots \supset J_n = 0$ so that J_k/J_{k+1} is abelian for each k.
   (c) L has a sequence of subalgebras $L \supset L_1 \supset L_2 \cdots \supset L_n = 0$ so that L_{k+1} is an ideal in L_k and L_k/L_{k+1} is abelian for all k.

4. Jordan decomposition and Cartan's criterion

Today I will explain Cartan's criterion which implies that a Lie algebra is solvable. It uses Engel's Theorem (a Lie algebra is nilpotent iff every element is ad-nilpotent) and the Jordan decomposition. The proof requires the ground field F to be the complex numbers.

4.1. **Jordan decomposition.** For this we need the ground field F to be algebraically closed. Suppose that $x \in \mathfrak{gl}(V)$ where V is a finite dimensional vector space over F. Then V has a basis with respect to which x is in *Jordan canonical form.* Here is an example to remind you how it looks.

$$x = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} J_3(\lambda_1) & 0 \\ 0 & J_2(\lambda_2) \end{bmatrix} = J_3(\lambda_1) \oplus J_2(\lambda_2)$$

The *Jordan decomposition* of x is given by writing x as a sum of two matrices: $x = x_s + x_n$ where x_s is "semisimple" (explained below) and x_n is nilpotent:

$$x_s = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}, \quad x_n = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Furthermore, x_s and x_n commute. This is obvious by looking at the matrices. What is not obvious is that this decomposition is unique.

4.1.1. *generalized eigenspaces.* In this example there are two *Jordan blocks* $J_3(\lambda_1), J_2(\lambda_2)$ where $J_m(\lambda)$ is an $m \times m$ matrix with λ along the diagonal, 1 on the super-diagonal and 0 elsewhere. Note that the minimal polynomial of x in this example is

$$p(T) = (T - \lambda_1)^3 (T - \lambda_2)^2.$$

This corresponds to a decomposition of V as a direct sum $V = V_1 \oplus V_2$ where V_i are *generalized eigenspaces* of x defined as follows.

$$V_i = \{v \in V \mid (x - \lambda_i)^m = 0 \text{ for some } m\}.$$

The exponents $3, 2$ correspond to the size of the Jordan blocks in this case. In general we have the following proposition which follows from the fact that $F[T]$ is a principal ideal domain.

Proposition 4.1.1. *Let $x \in \mathfrak{gl}(V)$ and let $p(T) = \prod (T - \lambda_i)^{m_i}$ be the minimal polynomial of x. Then $V = \bigoplus V_i$ where V_i is a the generalized eigenspace of x corresponding to the eigenvalue λ_i. Furthermore we have the following for each i:*

1. $x(V_i) \subseteq V_i$.
2. *V_i is the kernel of $(x - \lambda_i)^{m_i} : V \to V$.* □

4.1.2. *semisimple endomorphisms.*

Definition 4.1.2. $x \in \mathfrak{gl}(V)$ is *semisimple* if any of the following equivalent conditions is satisfied.

(1) The roots of the minimal polynomial of x are distinct. (In other words, the exponents m_i are equal to 1.)
(2) Each generalized eigenspace is equal to the actual eigenspace $V_{\lambda_i} = \{v \in V \mid x(v) = \lambda_i v\}$.
(3) x is *diagonalizable*, i.e., there is a basis for V (consisting of eigenvectors) with respect to which the matrix of x is diagonal.

Lemma 4.1.3. *If x, y are two commuting semisimple endomorphisms of V then $x + y$ is also semisimple.*

Proof. Let V_i be the λ_i-eigenspace of x.
Claim $y(V_i) \subseteq V_i$.
Pf: For any $v \in V_i$ we have: $x(y(v)) = y(x(v)) = y(\lambda_i v) = \lambda_i y(v)$. So, $y(v) \in V_i$.

The minimal polynomial of $y|V_i$ divides the minimal polynomial of y on V. Therefore, $y|V_i$ is semisimple and thus diagonal wrt some basis for each V_i. But $x|V_i$, being multiplication by a scalar λ_i, is diagonal wrt any basis on V_i. So, $x + y$ is diagonal with respect to this basis. So, $x + y$ is semisimple. □

Theorem 4.1.4. *If F is algebraically closed and V is finite dimensional then any $x \in \mathfrak{gl}(V)$ can be written uniquely as a sum $x = x_s + x_n$ where x_s is semisimple, x_n is nilpotent and x_s, x_n commute.*

We have the following basis-independent description of x_s, the semisimple part of x as given by the Jordan canonical form. x_s is the unique endomorphism of V which sends each V_i to V_i by multiplication by λ_i and $x_n = x - x_s$.

Lemma 4.1.5. *Let $V_1, \cdots, V_k$ be the generalized eigenspaces of x and let $\mu_i \in F$. Then there is a polynomial $q(T) \in F[T]$ so that $q(x)$ sends each V_i to V_i by multiplication by μ_i.*

Proof that Lemma implies Theorem. Let $\mu_i = \lambda_i$. Then we see that $x_s = q(x)$ is a polynomial in x. This implies that x_s commutes with x and therefore with $x_n = x - x_s$.

Suppose that $x = x_s' + x_n'$ is another decomposition of x into a semisimple part x_s' and nilpotent part x_n' so that x_s', x_n' commute. Then x_s' commutes with x and therefore with $x_s = q(x)$. Therefore, by the previous lemma, $x_s' - x_s$ is semisimple. But

$$x_s' - x_s = x_n - x_n'$$

is also nilpotent since x_n, x_n' are commuting nilpotent endomorphisms of V. But the only nilpotent diagonal matrix is the zero matrix. So, $x_s = x_s'$ and $x_n = x_n'$. □

Proof of Lemma. We use the Chinese remainder theorem which says that, since $(T-\lambda_i)^{m_i}$ are relatively prime, the projection map

$$F[T] \to \prod F[T]/((T-\lambda_i)^{m_i})$$

is surjective. So, there exists a $q(T) \in F[T]$ so that $q(T) - \mu_i$ is divisible by $(T - \lambda_i)^{m_i}$ for each i. But then $q(x) - \mu_i$ is equal to 0 on V_i. □

4.2. **Cartan's criterion.** This requires $F = \mathbb{C}$. The idea is to find some condition which implies that every element of $[LL]$ is ad-nilpotent. By Engel's Theorem this will imply that $[LL]$ is a nilpotent algebra and thus a solvable algebra. And this will imply that L is solvable.

Lemma 4.2.1. *Suppose that $A \subseteq B \subseteq \mathfrak{gl}(V)$, $V \cong \mathbb{C}^n$, and*

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

Suppose that $x \in M$ so that $\mathrm{Tr}(xy) = 0$ for all $y \in M$. Then x is nilpotent.

Proof. We want to show that all the eigenvalues λ_i of x are equal to zero. Let V_i be the generalized eigenspace of x corresponding to λ_i. Suppose V_i has dimension d_i. By Lemma 4.1.5 there exists $q(T) \in \mathbb{C}[T]$ so that $y = q(x)$ sends V_i to V_i by multiplication by the complex conjugate $\overline{\lambda}_i$ of λ_i for each i. From the fact that $\overline{\lambda}_i$ is a $\mathbb{Q}$-linear function of λ_i it follows that $y \in M$. (See explanation below.) Therefore, $\mathrm{Tr}(xy) = 0$ by assumption. But V_i is the generalized eigenspace of xy with nonnegative real eigenvalue $\lambda_i\overline{\lambda}_i$. So,

$$\mathrm{Tr}(xy) = \sum d_i \lambda_i \overline{\lambda}_i = 0$$

implies that $\lambda_i = 0$ for all i. Therefore x is nilpotent. □

Lemma 4.2.2. *If $x, y, z \in \mathfrak{gl}(V)$ then $\mathrm{Tr}(x[yz]) = \mathrm{Tr}([xy]z)$.*

Proof. $x[yz] - [xy]z = xyz - xzy - xyz + yxz = yxz - xzy = [y, xz]$ has trace zero. □

Theorem 4.2.3 (Cartan's criterion)**.** *Suppose that $L \subseteq \mathfrak{gl}(V)$ where $V \cong \mathbb{C}^m$. Suppose $\mathrm{Tr}(xy) = 0$ for all $x \in [LL], y \in L$. Then L is solvable.*

Proof. We want to show that every $x \in [LL]$ is nilpotent using the lemma. So, let $A = [LL], B = L$ in the lemma. Then

$$M = \{z \in \mathfrak{gl}(V) \,|\, [zL] \subseteq [LL]\}.$$

To apply the lemma we need to show that $\mathrm{Tr}(xz) = 0$ for all $z \in M$. But $x \in [LL]$ is a linear combination of commutators $[ab]$ and $\mathrm{Tr}([ab]z) = \mathrm{Tr}(a[bz]) = 0$ since $a \in L$ and $[bz] \in [LL]$. Therefore, $\mathrm{Tr}(xz) = 0$ for all $z \in M, x \in [LL]$. The Lemma implies that every element of $[LL]$ is nilpotent. So, $[LL]$ is nilpotent and L is solvable. □

Corollary 4.2.4. *Suppose that L is a finite dimensional complex Lie algebra with the property $\mathrm{Tr}(\mathrm{ad}\, x\, \mathrm{ad}\, y) = 0$ for all $x \in [LL], y \in L$. Then L is solvable.*

Proof. By Cartan's criterion, the image of the adjoint representation $\mathrm{ad} : L \to \mathfrak{gl}(L)$ is solvable. Since the kernel $Z(L)$ is abelian, we conclude that L is solvable. □

Exercise 4.2.5.

(1) Prove that $\mathrm{ad}\, x_s = (\mathrm{ad}\, x)_s$ and $\mathrm{ad}\, x_n = (\mathrm{ad}\, x)_n$. We did this at the end of the class. But this was needed in the proof that $y \in M$ in Lemma 4.2.1
(2) Use Cartan's criterion to show that any subalgebra of $\mathfrak{t}(n, \mathbb{C})$ is solvable.
(3) Modify the proof of Cartan's criterion so that it works over any algebraically closed field of characteristic zero.

4.2.1. *Proof that $y \in M$.*

Lemma 4.2.6. *Suppose that $V \cong F^n$ with F algebraically closed. Then, for any $x \in \mathfrak{gl}(V)$ we have $\operatorname{ad} x_s = (\operatorname{ad} x)_s$ and $\operatorname{ad} x_n = (\operatorname{ad} x)_n$.*

Proof. (As we did in class.) The Jordan decomposition of $\operatorname{ad} x$ is

$$\operatorname{ad} x = (\operatorname{ad} x)_s + (\operatorname{ad} x)_n$$

it is characterized by the fact that $(\operatorname{ad} x)_s$ is semisimple, $(\operatorname{ad} x)_n$ is nilpotent and they commute. So, the lemma is equivalent to the following statements.

(1) $\operatorname{ad} x_s$ is semisimple.
(2) $\operatorname{ad} x_n$ is nilpotent.
(3) $\operatorname{ad} x_s$ and $\operatorname{ad} x_n$ commute.

The second and third conditions are obvious, so we just need to prove the first condition.

Take a basis for V so that x is a diagonal matrix with diagonal entries $\lambda_1, \cdots, \lambda_n$. Then we get a basis e_{ij} for $\mathfrak{gl}(V)$ where e_{ij} is the matrix with 1 in the (i, j) position and 0 elsewhere. The endomorphism $\operatorname{ad} x_s$ sends e_{ij} to $(\lambda_i - \lambda_j)e_{ij}$. So, $\operatorname{ad} x_s$ is diagonalizable and thus semisimple. □

Next we need the following corollary of Lemma 4.1.5 which I was explaining in class. (Prove it directly using the Chinese remainder theorem!)

Proposition 4.2.7. *Suppose that $x \in \mathfrak{gl}(V)$ is a diagonal matrix with diagonal entries $\lambda_1, \cdots, \lambda_n$, suppose that $\mu_1, \cdots, \mu_n \in F$ are such that, whenever $\lambda_i = \lambda_j$ then $\mu_i = \mu_j$ (i.e., μ_i is a function of λ_i). Then there is a polynomial $q(T) \in F[T]$ so that $q(x)$ is a diagonal matrix with entries $\mu_1, \cdots, \mu_n$. (Equivalently, $q(\lambda_i) = \mu_i$ for all i.)*

Corollary 4.2.8. *Take $F = \mathbb{C}$. Let $x \in \mathfrak{gl}(V)$. Let $y = q(x)$ be the endomorphism of V which sends each V_i to V_i by multiplication by $\overline{\lambda}_i$ (i.e., y is the complex conjugate of x_s). Then $\operatorname{ad} y$ is a polynomial in $\operatorname{ad} x$.*

This implies that $y \in M$ in the proof of Lemma 4.2.1 since the defining condition for M is that

$$M = \{x \in \mathfrak{gl}(V) \mid \operatorname{ad} x(B) \subseteq A\}$$

If $\operatorname{ad} x$ sends B into $A \subseteq B$ then any polynomial in $\operatorname{ad} x$ also sends B into A. So, $\operatorname{ad} y$ sends B into A making y an element of M as claimed.

Proof. Lemma 4.2.6 implies that $\operatorname{ad} x_s = (\operatorname{ad} x)_s$ is a polynomial in $\operatorname{ad} x$. As we discussed, its eigenvalues are $\lambda_i - \lambda_j$. But $\overline{\lambda}_i - \overline{\lambda}_j$ is a function of $\lambda_i - \lambda_j$. So, by the proposition above, the complex conjugate of $\operatorname{ad} x_s$ (which is $\operatorname{ad} y$) is a polynomial in $\operatorname{ad} x_s$ and therefore also a polynomial in $\operatorname{ad} x$. This proves the corollary. □

5. Semisimple Lie algebras and the Killing form

This section follows Procesi's book [6]. We will define semisimple Lie algebras and the Killing form and prove the following.

Theorem 5.0.9. *The following are equivalent for L a finite dimensional Lie algebra over any subfield $F \subseteq \mathbb{C}$.*

1. *L is semisimple.*
2. *L has no nonzero abelian ideals.*
3. *The Killing form of L is nondegenerate.*
4. *L is a direct sum of simple ideals.*

5.1. **Definition.** First we observe that the sum of two solvable ideals I, J in L is solvable. This follows from the fact that I and $(I+J)/I = J/(I \cap J)$ are both solvable.

Definition 5.1.1. The *solvable radical* Rad L of L is defined to be the sum of all solvable ideals. A Lie algebra is *semisimple* if its solvable radical is zero, i.e., if it has no nonzero solvable ideal.

Proposition 5.1.2. *L is semisimple iff L has no nonzero abelian ideals.*

Proof. If L is semisimple then it has no abelian ideals. Conversely, if L is not semisimple, then L has a solvable ideal J. Then $DJ = [JJ]$ is also an ideal in L since

$$[x[JJ]] \subseteq [[xJ]J] + [J[xJ]] \subseteq [JJ]$$

We have $D^k J = 0$. So $D^{k-1}J$ is a nonzero abelian ideal in L. □

5.2. **Killing form.** The *Killing form* $\kappa : L \times L \to F$ is defined by

$$\kappa(x, y) = \operatorname{Tr}(\operatorname{ad} x \operatorname{ad} y)$$

The Killing form is clearly symmetric: $\kappa(x, y) = \kappa(y, x)$. The Killing form is also "associative":

$$\kappa([xy], z) = \kappa(x, [yz])$$

Proof. Since $\operatorname{ad}[xy] = [\operatorname{ad} x, \operatorname{ad} y]$, we have:

$$\kappa([xy], z) = \operatorname{Tr}(\operatorname{ad}[xy] \operatorname{ad} z) = \operatorname{Tr}([\operatorname{ad} x, \operatorname{ad} y]\operatorname{ad} z) = \operatorname{Tr}(\operatorname{ad} x[\operatorname{ad} y, \operatorname{ad} z]) = \kappa(x, [yz])$$

Proposition 5.2.1. *The Killing form is invariant under any automorphism ρ of L.*

Proof. The equation $\rho[xy] = [\rho(x)\rho(y)]$ for $z = \rho(y)$ is $\rho[x, \rho^{-1}(z)] = [\rho(x)z]$ which can be rewritten as: $\operatorname{ad}\rho(x) = \rho \circ \operatorname{ad} x \circ \rho^{-1}$. So,

$$\kappa(\rho(x), \rho(y)) = \operatorname{Tr}(\operatorname{ad}\rho(x) \operatorname{ad}\rho(y)) = \operatorname{Tr}(\rho \circ \operatorname{ad} x \operatorname{ad} y \circ \rho^{-1}) = \operatorname{Tr}(\operatorname{ad} x \operatorname{ad} y) = \kappa(x, y).$$ □

Lemma 5.2.2. *The kernel (also called the radical) of the Killing form of L is an ideal.*

The definition of the kernel of κ is:

$$S = \{x \in L \mid \kappa(x, z) = 0 \text{ for all } z \in L\}$$

Proof. Suppose $x \in S$ and $y \in L$. Then $\kappa([xy], z) = \kappa(x, [yz]) = 0$. □

Lemma 5.2.3. *Every abelian ideal in L is contained in the kernel of its Killing form.*

Proof. Suppose $J \subseteq L$ is an abelian ideal, $x \in J$ and $y \in L$. Then $\operatorname{ad} x \operatorname{ad} y$ sends L into J and $\operatorname{ad} x \operatorname{ad} y(J) \subseteq \operatorname{ad} x(J) = 0$. So, $(\operatorname{ad} x \operatorname{ad} y)^2 = 0$. Since nilpotent endomorphisms have zero trace, $\kappa(x, y) = \operatorname{Tr}(\operatorname{ad} x \operatorname{ad} y) = 0$ showing that $J \subseteq S$. □

Theorem 5.2.4 (Cartan)**.** *Suppose that L is a finite dimensional Lie algebra over $\mathbb{C}$. Then L is semisimple iff its Killing form is nondegenerate (its kernel $S = 0$).*

Proof. We will show that $S \neq 0$ iff L is not semisimple. If L is not semisimple then it has a nonzero abelian ideal. Any such ideal lies in the kernel S. So, $S \neq 0$.

Conversely, suppose that $S \neq 0$. Then Cartan's criterion shows that the image $\operatorname{ad}_L S$ of S under the adjoint representation $\operatorname{ad}_L : L \to \mathfrak{gl}(L)$ is solvable since

$$\kappa(x, y) = \operatorname{Tr}(\operatorname{ad} x \operatorname{ad} y) = 0$$

for all $x, y \in S$. Since $\operatorname{ad}_L S = S/Z(L)$, this implies that S is solvable. Therefore, L is not semisimple. □

Corollary 5.2.5. *Suppose that L is a finite dimensional Lie algebra over a subfield F of $\mathbb{C}$. Then L is semisimple iff its Killing form is nondegenerate.*

Proof. Suppose that the Killing form of L is nondegenerate. Then L must be semisimple since any abelian ideal is contained in the kernel of κ which is zero. Conversely, suppose that the Killing form of L has a nonzero kernel S.

Let $L_{\mathbb{C}} = L \otimes_F \mathbb{C}$ be the complexification of L. Since $L \subseteq L_{\mathbb{C}}$, it is clear that L is abelian iff $L_{\mathbb{C}}$ is abelian. This implies the L is solvable iff $L_{\mathbb{C}}$ is solvable since $D(L_{\mathbb{C}}) = (DL)_{\mathbb{C}}$. One can also see that the Killing form $\kappa_{\mathbb{C}}$ of $L_{\mathbb{C}}$ is the complexification of the Killing form κ of L. So, the kernel of $\kappa_{\mathbb{C}}$ is $S_{\mathbb{C}}$. We know that $S_{\mathbb{C}}$ is solvable from the proof of Cartan's Theorem. Therefore S is solvable and L is not semisimple. □

Example 5.2.6. Let $L = \mathfrak{sl}(2, \mathbb{R})$. This has basis

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

with respect to this basis we have

$$\operatorname{ad} x = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \operatorname{ad} y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 0 \end{pmatrix}, \quad \operatorname{ad} h = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Then

$$\begin{array}{lll} \kappa(x,x) = 0 & \kappa(x,y) = 4 & \kappa(x,h) = 0 \\ & \kappa(y,y) = 0 & \kappa(y,h) = 0 \\ & & \kappa(h,h) = 8 \end{array}$$

The matrix of the form κ is therefore

$$\begin{bmatrix} 0 & 4 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

Since this matrix is invertible, κ is nondegenerate and $L = \mathfrak{sl}(2,\mathbb{R})$ is semisimple. Since this matrix has negative determinant and positive trace its *signature* (#+ eigenvalues − #− eigenvalues) is 1.

Exercise 5.2.7. Compute the Killing form of the real cross product algebra. Conclude that this algebra is semisimple but not isomorphic to $\mathfrak{sl}(2,\mathbb{R})$.

5.3. **Product of simple ideals.**

Theorem 5.3.1. *Suppose that L is a finite dimensional semisimple Lie algebra over any subfield $F \subseteq \mathbb{C}$. Then L can be expressed uniquely as a product of simple ideals.*

Proof. First we should point out that if L is a direct sum of two ideals $L = J_1 \oplus J_2$ then $L \cong J_1 \times J_2$ with the isomorphism given by the projection maps $L \to L/J_2, L \to L/J_1$.

If L is simple, the statement trivially holds. Otherwise, let $J \subseteq L$ be a minimal ideal. Define $J^\perp$ to be the set of all $x \in L$ so that $\kappa(x, J) = 0$. Then $J^\perp$ is an ideal since $\kappa([xy], J) = \kappa(x, [yJ]) \subseteq \kappa(x, J) = 0$ for all $x \in J^\perp, y \in L$. Therefore, $J \cap J^\perp$ is also an ideal. By minimality of J we have either $J \cap J^\perp = J$ or $J \cap J^\perp = 0$. The first case is not possible since $\kappa = 0$ on $J \cap J^\perp$ which, by Cartan's criterion, would imply that $J_\mathbb{C}$ is solvable, so J would be solvable.

Since κ is nondegenerate, we have $\dim J + \dim J^\perp = \dim L$. Therefore, $L = J \oplus J^\perp$. By induction, $J^\perp$ is a product of simple ideals. So, $L = J_1 \oplus J_2 \oplus \cdots \oplus J_n \cong \prod J_i$. To prove uniqueness of this decomposition, suppose that I is another minimal ideal. Then

$$I = [IL] = [IJ_1] \oplus [IJ_2] \oplus \cdots \oplus [IJ_n]$$

One of these summands must be nonzero. Say $[IJ_i] \subseteq I \cap J_i \neq 0$. Then $I = J_i$. □

Corollary 5.3.2. *Let L be a finite dimensional Lie algebra over $F \subseteq \mathbb{C}$. Then L is semisimple iff it is a product of simple ideals.*

Exercise 5.3.3.
(1) Show that L is nilpotent iff its Killing form is identically zero.
(2) Show that the Killing form of a nonabelian 2-dimensional Lie algebra is nontrivial.
(3) For $F \subseteq \mathbb{C}$ show that L is solvable iff $[LL]$ is contained in the kernel of κ.
(4) Show that κ nondegenerate implies L semisimple over any field.
(5) For *char* $F = 3$ show that $\mathfrak{sl}(3, F)$ modulo its center is semisimple but its Killing form is degenerate.

6. Completely reducible representations

6.1. Modules.

Definition 6.1.1. A *representation* of L is a homomorphism $L \to \mathfrak{gl}(V)$. Then V is called a *module* over L. The action of $x \in L$ on $v \in V$ is denoted $x.v$.

One can describe an action abstractly by saying that $(x, v) \mapsto v$ is F-bilinear in x, v and satisfies the property:

$$[xy].v = x.y.v - y.x.v$$

for all $x, y \in L, v \in V$.

Humphreys points out that V is a module over the associative algebra A_V generated by the image of L in $\mathfrak{gl}(V) = \mathrm{End}_F(V)$. Therefore, any submodule or quotient module of V (or any tensor power) is an L-module. Since $V^* = \mathrm{Hom}(V, F)$ is a right A_V-module, it becomes a *right L-modules*. However, L is isomorphic to its opposite L^{op} since it has an anti-automorphism given by $x \mapsto -x$. So, this explains why V^* is also a (left) L-module.

Definition 6.1.2. If V, W are L-modules then $V \times W = V \oplus W$ is the L-module with action of L given by

$$x.(v, w) = (x.v, x.w)$$

The action of L on $V \otimes W$ is defined by

$$x.(v \otimes w) = x.v \otimes w + v \otimes x.w$$

The action on $\mathrm{Hom}_F(V, W)$ is

$$(x.f)(v) = x.(f(v)) - f(x.v)$$

$$\mathrm{Hom}_L(V, W) = \{f \in \mathrm{Hom}_F(V, W) \,|\, x.f = 0\}$$

The action of L on the dual $V^* = \mathrm{Hom}(V, F)$ is given by

$$(x.f)(v) = -f(x.v)$$

since we take the trivial action: $x.a = 0$ of L on F.

Proposition 6.1.3. *If L is semisimple then any one-dimensional representation is trivial.*

This follows from the more general:

Lemma 6.1.4. *If L is semisimple then any homomorphism $\varphi : L \to \mathfrak{gl}(V)$ has image in $\mathfrak{sl}(V)$.*

Proof. $L = [LL]$. So, $\varphi(L) = \varphi[LL] \subseteq [\mathfrak{gl}(V)\mathfrak{gl}(V)] = \mathfrak{sl}(V)$. □

6.2. Schur's Lemma.

A module V is called *irreducible* if it has no nonzero proper submodules.

Proposition 6.2.1 (Schur's Lemma)**.** *Suppose that F is algebraically closed. Then any endomorphism of an irreducible representation V is a scalar multiple of the identity map.*

Proof. Let $f : V \to V$ be any morphism. Let λ be any eigenvalue of f. Then $f - \lambda id_V$ is an endomorphism of V with nonzero kernel. But the kernel of any morphism is a submodule of V. Since V is irreducible, it must be all of V. So, $f = \lambda id_V$. □

Definition 6.2.2. A module is *completely reducible* if it is a finite direct sum of irreducible representations.

The main theorem is Weyl's Theorem which says that any finite dimensional representation of a semisimple Lie algebra is completely reducible. This requires the Casimir operator

6.3. **Casimir operator.** Given a nongenerate symmetric bilinear form $\beta : L \times L \to F$ which is associative ($\beta([xy], z) = \beta(x, [yz])$) and any V modules we will construct an L-homomorphism $c_\beta : V \to V$. Thus c_β is an operator on every L-module. (It is a central element of the universal enveloping algebra.) The definition uses a basis for L. We need to show it is independent of the choice of basis.

Definition 6.3.1. Choose a basis $\{x_i\}$ for L and choose a dual basis $\{y_i\}$ for L with respect to the form β. Thus y_i are uniquely determined by the property:

$$\beta(x_i, y_j) = \delta_{ij}$$

Given any representation $\varphi : L \to \mathfrak{gl}(V)$, the *Casimir operator* c_β is the endomorphism of V given by

$$c_\beta(v) = \sum_i \varphi(x_i)\varphi(y_i)(v)$$

Note that $\sum_i \varphi(x_i)\varphi(y_i) \in \mathfrak{gl}(V)$ may not be an element of $\varphi(L)$.

Proposition 6.3.2. *c_β is independent of the choice of x_i.*

Proof. Any other basis is given by $x_i' = \sum a_{ij} x_j$ where $(a_{ij}) \in GL(n, F)$. The dual basis is given by $y_j' = \sum b_{jk} y_k$ where $(b_{jk}) = (a_{ij})^{-t}$ (inverse transpose). This gives the same operator c_β. □

To prove the key property (next proposition) we need the following observation. We can express any $z \in L$ in terms of both bases: $z = \sum a_i x_i = \sum b_j y_j$ and the formula for a_i, b_j is

$$a_i = \beta(z, y_i), \quad b_i = \beta(x_i, z)$$

Thus

$$z = \sum \beta(z, y_i) x_i = \beta(x_i, z) y_i$$

Proposition 6.3.3. *$c_\beta : V \to V$ is a homomorphism of L-modules. (I.e., $[z.c_\beta] = 0$ for all $z \in L$.)*

Proof. $z.c_\beta = \varphi(z) \sum \varphi(x_i)\varphi(y_i) - \sum \varphi(x_i)\varphi(y_i)\varphi(z) =$

$$\sum \varphi[zx_i]\varphi(y_i) + \sum \varphi(x_i)\varphi[zy_i]$$

$$= \sum \beta([zx_i], y_i)\varphi(x_i)\varphi(y_i) + \beta(x_i, [zy_i])\varphi(x_i)\varphi(y_i)$$

This $= 0$ since each summand is zero by associativity of β:

$$\beta([zx_i], y_i) = -\beta([x_i z], y_i) = -\beta(x_i, [zy_i])$$

□

6.4. **Weyl's Theorem.**

Theorem 6.4.1 (Weyl)**.** *Any finite dimension representation $\varphi : L \to \mathfrak{gl}(V)$ of a semisimple Lie algebra L is completely reducible. Assume F is algebraically closed with characteristic zero or prime characteristic greater than* $\dim L$

Lemma 6.4.2. *A representation V is completely reducible iff every submodule W has a complementary module X so that $V = W \oplus X$.* □

The idea of the proof is to use the Casimir operator c_β corresponding to the bilinear form

$$\beta(x, y) = \mathrm{Tr}(\varphi(x)\varphi(y))$$

Since L is semisimple, this is nondegenerate (and symmetric and associative). Then

$$\mathrm{Tr}(c_\beta) = \mathrm{Tr}(\sum \varphi(x_i)\varphi(y_i)) = \sum \beta(x_i, y_i) = \dim L \neq 0$$

By induction on $\dim L$ we may assume that φ is a faithful representation.

Proof of Weyl's Theorem. Given any submodule W of V we will show that there is a complementary submodule Z.

Claim 1 We may assume that W is irreducible. (I.e., if W is not irreducible then it has a complement.)

Pf: We know by induction that W is completely reducible: $W = \sum W_i$ where W_i are irreducible. If there is more than one component then we get a nontrivial decomposition $W = X \oplus Y$. Then $W/X \subseteq V/X$. So, by induction there is a complementary module C/X. Then $W + C = V$ and $W \cap C = X$ since C is smaller than V, the submodule $X \subseteq C$ has a complement Z making

$$V = Y \oplus X \oplus Z = W \oplus Z$$

Proving that Z is a complement for W. Therefore, we may assume W is irreducible.

Case 1 Suppose that V/W is 1-dimensional. Then we proved that L acts trivially on V/W. This implies that the Casimir operator c_β also acts trivially on V/W. In other words $c_\beta(V) \subseteq W$. Since W is irreducible, c_β is multiplication by a scalar on W. Since $\mathrm{Tr}(c_\beta) = \dim L \neq 0$, this scalar is nonzero. Then we get $V = W \oplus \ker c_\beta$ as desired.

Case 2 In the general case, we consider

$$\mathcal{V} = \{f \in \mathrm{Hom}_F(V, W) \mid f|W \text{ is multiplication by a scalar}\}$$

$$\mathcal{W} = \{f \in \mathrm{Hom}_F(V, W) \mid f|W = 0\} \subseteq \mathcal{V}$$

The following calculation shows that $x.f \in \mathcal{W}$ for all $f \in cV$. So$\mathcal{V}$ is an L-submodule of $\mathrm{Hom}_F(V, W)$ and $\mathcal{W}$ is a codimension 1 submodule.

$$x.f(w) = x(f(w)) - f(x.w) = \lambda x.w - \lambda x.w = 0$$

Therefore, by Case 1, there is a complementary one dimensional submodule for $\mathcal{W}$ in $\mathcal{V}$. It is generated by one element $f \in \mathcal{V}$. Since 1 dim reps of L are trivial, $x.f = 0$ for all $x \in L$. This means $f \in \mathrm{Hom}_L(V, W)$. So, f is a multiple of a retraction of V to W and $\ker f$ is a complement for W in V. □

6.5. **Preservation of Jordan decomposition.** *From now on, we will assume that F is algebraically closed of characteristic zero.*

The following theorem is crucial to the next section.

Theorem 6.5.1. *Suppose $L \subseteq \mathfrak{gl}(V)$ is semisimple and V is finite dimensional. Then L contains the semisimple and nilpotent parts of its Jordan decomposition: $x_s, x_n \in L$ for all $x \in L$.*

Remark 6.5.2. The decomposition $x = x_s + x_n$ is unique and therefore does not depend on the representation. The proof is that the endomorphism $\mathrm{ad}_L x$ of L decomposes uniquely as $\mathrm{ad}_L x = (\mathrm{ad}_L x)_s + (\mathrm{ad}_L x)_n = \mathrm{ad}_L x_s + \mathrm{ad}_L x_n$. So, $\mathrm{ad}_L x_s$ is the semisimple part of $\mathrm{ad}_L x$. But the adjoint representation of L is faithful since L is semisimple. Therefore, $\mathrm{ad}_L x_s$ determines x_s uniquely.

Proof. We know that $x_s, x_n \in \mathfrak{gl}(V)$ are polynomials in x. Thus, x_s, x_n leave invariant any L-submodule W of V. Furthermore, we claim:

$$x_s|W \in \mathfrak{sl}(W)$$

This follows from the calculation:

$$\mathrm{Tr}(x_s|W) = \mathrm{Tr}(x|W) - \mathrm{Tr}(x_n|W) = 0$$

since x_n is nilpotent and $L \to \mathfrak{gl}(W)$ has image in $\mathfrak{sl}(W)$ since L is semisimple.

Also, $\mathrm{ad}\, x_s = (\mathrm{ad}\, x)_s$ and $\mathrm{ad}\, x_n = (\mathrm{ad}\, x)_n$ are polynomials in $\mathrm{ad}\, x$. So, $[x_s, L] \subseteq L$ and $[x_n, L] \subseteq L$.

Let L^* be the set of all $y \in \mathfrak{sl}(V)$ so that $[y, L] \subseteq L$ and $y|W \in \mathfrak{sl}(W)$ for all L-submodules W of V. We have that $L \subseteq L^*$ and $x_s, x_n \in L^*$ for all $x \in L$. So, it suffices to show that $L^* = L$.

But, L^* is a module over L and the condition $[y, L] \subseteq L$ implies that the action of L sends L^* into L. By Weyl's Theorem, L has a complement M in L^* and the action of L on M is trivial. Let $y \in M$. Then $[y, L] = 0$. This means that $y : V \to V$ is a homomorphism of L-modules. By definition, y preserves each submodule W of V. Take a decomposition of V into irreducible submodules W. By Schur's Lemma, y acts by multiplication by a scalar λ on each such W. But $\mathrm{Tr}(y|W) = \lambda \dim W = 0$ implies that $\lambda = 0$. Since this holds on each component of V, $y = 0$. So, $M = 0$ and $L^* = L$. So, $x_s, x_n \in L^* = L$. □

7. Representations of $\mathfrak{sl}(2, F)$

Let $L = \mathfrak{sl}(2, F)$. Recall that the standard basis is:

$$x = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

and $[hx] = 2x, [hy] = -2y, [xy] = h$.

Since $\mathrm{ad}_L h$ is semisimple, it follows from Theorem 6.5.1 and the Remark after the theorem that $\varphi(h)$ is semisimple for any finite dimensional representation $\varphi : L \to \mathfrak{gl}(V)$. Therefore, V decomposes into a direct sum of eigenspaces

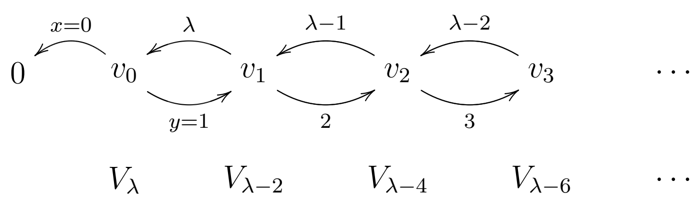
$$V_\lambda = \{v \in V \mid h.v = \lambda v\}$$

of the action of h on V. If $V_\lambda \neq 0$ then V_λ is called a *weight space* and the decomposition $V = \oplus V_\lambda$ is called the *weight space decomposition* of V.

Lemma 7.0.3. *If $v \in V_\lambda$ then $x.v \in V_{\lambda+2}$ and $y.v \in V_{\lambda-2}$.*

If V is finite dimensional, it follows that $x^n.v = 0$ for some positive n. We say that v_0 is a *maximal vector* if $v_0 \in V_\lambda$ for some λ and $x.v_0 = 0$. If $x^n.v \neq 0$ and $x^{n+1}.v = 0$ then $x^n.v$ is a maximal vector.

Lemma 7.0.4. *Suppose that $v_0 \in V_\lambda$ is a maximal vector for the action of L on V. For all $i \geq 0$ let $v_i = \frac{1}{i!}y^i.v_0$ and let $v_{-1} = 0$. Then, for all $i \geq 0$ we have:*
(1) $h.v_i = (\lambda - 2i)v_i$ *i.e.,* $v_i \in V_{\lambda-2i}$
(2) $y.v_i = (i+1)v_{i+1}$
(3) $x.v_i = (\lambda - i + 1)v_{i-1}$



Proof. (1) follows from the previous lemma. (2) follows directly from the definition of v_i. (3) is an easy induction. If $i = 0$ it is true since $x.v_0 = 0$ by assumption. Assuming (3) holds for i we get

$$(i+1)x.v_{i+1} = x.y.v_i = (h+yx).v_i = (\lambda-2i)v_i + y.(\lambda-i+1)v_{i-1} = (\lambda-2i+i(\lambda-i+1))v_i$$

But

$$\lambda - 2i + i(\lambda - i + 1) = i\lambda - i^2 + \lambda - i = (i+1)(\lambda - i)$$

Therefore,

$$x.v_{i+1} = (\lambda - i)v_i$$

showing that (3) holds for all $i \geq 0$. □

Since V is finite dimensional, and the v_i are linearly independent by (1), we have that $v_i = 0$ for sufficiently large i. Let m be maximal so that $v_m \neq 0$. Then $v_{m+1} = 0$. So:

$$x.v_{m+1} = 0 = (\lambda - m)v_m$$

This implies that $\lambda = m$.

Example 7.0.5. Take $V = F^2$ with the standard action of $L - \mathfrak{sl}(2, F)$. The basis vectors are $e_1 = \binom{1}{0} \in V_1$ and $e_2 = \binom{0}{1} \in V_{-1}$. Since $x.e_1 = 0$ we have $v_0 = e_1$. Furthermore, $y.e_1 = e_2$. Therefore, $v_1 = e_2$. Since $y.v_1 = 0$ we have $m = 1$.

Now consider the m-fold tensor product

$$W = V^{\otimes m} = \underbrace{V \otimes V \otimes \cdots \otimes V}_{m}$$

Take the element $e_1 \otimes e_1 \otimes \cdots \otimes e_1$. Then

$$h(e_1 \otimes \cdots \otimes e_1) = h.e_1 \otimes e_1 \otimes \cdots + e_1 \otimes h.e_1 \otimes \cdots + \cdots = m(e_1 \otimes \cdots \otimes e_1)$$

Therefore, $e_1 \otimes \cdots \otimes e_1 \in W_m$. Since $x.(e_1 \otimes \cdots \otimes e_1) = 0$, this is a maximal vector.

In general, the lemma implies that $v_0, \cdots, v_m$ span an L-submodule of V. If V is irreducible then it must be all of V. Conversely, any irreducible L-module must be of this kind. Reexamination of the lemma shows that V is uniquely determined by the number $\lambda = m$ which is a nonnegative integer. Furthermore, every nonnegative integer m occurs by the example above.

Theorem 7.0.6. *Up to isomorphism the irreducible finite dimensional modules over* $\mathfrak{sl}(2, F)$ *are* $V(m)$ *for* $m \geq 0$ *where* $V(m)$ *is an* $m+1$ *dimensional representation generated by a maximal vector of weight* m *and*

$$V(m) = V_m \oplus V_{m-2} \oplus V_{m-4} \oplus \cdots \oplus V_{-m}$$

Corollary 7.0.7. *For any representation* V *of* $\mathfrak{sl}(2, F)$*, the weights are all integers, and* V *is uniquely determined up to isomorphism by the dimensions of the weight space* V_i*. Furthermore,* $\dim V_i = \dim V_{-i}$ *and* $\dim V_0 + \dim V_1$ *is equal to the number of irreducible components of* V*.*

Example 7.0.8. The weight space decomposition of L under the adjoint action is

$$L = L_{-2} \oplus L_0 \oplus L_2$$

This confirms what we already know: L is a simple module over itself. This weight space decomposition is called the *root space decomposition* of L.

Exercise 7.0.9. Given $\dim V_i$ for all i, find the decomposition of V as a direct sum of the irreducible modules $V(m)$.

Exercise 7.0.10. This is Exercise 7 on p.34. Let $\lambda \in F$ be any scalar and let $Z(\lambda)$ be the infinite dimensional vector space with basis $v_0, v_1, v_2, v_3, \cdots$.

(1) Show that the formulas in Lemma 7.0.4 define an action of $L = \mathfrak{sl}(2, F)$ on $Z(\lambda)$
(2) Show that every L-submodule of $Z(\lambda)$ contains a maximal vector.
(3) Show that $Z(\lambda)$ is irreducible if $\lambda + 1$ is not a nonnegative integer.
(4) If λ is a nonnegative integer then show that $V(\lambda)$ is a quotient module of $Z(\lambda)$ with kernel isomorphic to $Z(-\lambda - 2)$.

Exercise 7.0.11. Find the Casimir element for the adjoint representation of $L = \mathfrak{sl}(2, F)$.

Exercise 7.0.12. A Lie algebra L is called *reductive* if $Rad\, L = Z(L)$. For example, $L = \mathfrak{gl}(n, F)$. Show that L is reductive if and only if L is completely reducible as an L-module under the adjoint representation.

7.0.1. *avoiding computations.* In order to show that the formulas in Lemma 7.0.4 define an action of $L = \mathfrak{sl}(2, F)$ on $Z(\lambda)$, we need to show that three equations hold.

Lemma 7.0.13. *Suppose that $f(x), g(x) \in F[x]$ are polynomials of degree $\leq n$ with coefficients in a field F so that $f(a) = g(a)$ for at least $n+1$ values of a. Then $f(x) = g(x)$.*

Proof. If $f(x) \neq g(x)$ then $f(x) - g(x)$ can have at most n roots, i.e. there are at most n elements of F at which $f = g$. So, $f(a) = g(a)$ for $n + 1$ values of a implies that $f(x) = g(x)$ as polynomials in x. □

To prove that $Z(\lambda)$ is a module over L we need to verify three polynomial identities for each i. But we know that these identities hold for all integers λ greater than i. Therefore, by the Lemma, the identities hold for all $\lambda \in F$. So, $Z(\lambda)$ is a module as claimed.

Here is an example of an identity. We need to verify that $[xy].v_i = h.v_i$. Both sides are defined by the equations which give the action of x, y, h as multiplication by linear polynomials in λ:

$$[xy].v_i = x.y.v_i - y.x.v_i = x.(i+1)v_{i+1} - y.(\lambda - i + 1)v_{i-1}$$
$$= (\lambda - i)(i+1)v_i - i(\lambda - i + 1)v_i = ((\lambda - i)(i+1) - i(\lambda - i + 1))v_i$$
$$h.v_i = (\lambda - 2i)v_i$$

Therefore $[xy].v_i = h.v_i$ iff $(\lambda - i)(i+1) - i(\lambda - i + 1) = \lambda - 2i$ for all integers $i \geq 0$ and for all $\lambda \in F$. This is an easy calculation but the point is that we already know it is true for all $\lambda \in F$ by the lemma since it is true for infinitely many integers λ (by the existence of the modules $V(m)$ for all $m \geq 0$). There are two other similar computations to verify that $[hx].v_i = 2x.v_i$ and $[hy].v_i = -2y.v_i$. Each of these is a similar calculation which we do not need to do by the formal argument given above.

8. Root space decomposition

Now we come to root spaces and the classification of semisimple Lie algebras using Dynkin diagrams. My aim is to gloss over the combinatorics and emphasize the algebraic foundations.

First a review of the key definitions and theorems that we need.

8.0. **Review.** We assume L is a *semisimple* Lie algebra. This means $L = [LL]$ and L has no solvable ideals. So, the adjoint representation

$$L \hookrightarrow \mathfrak{gl}(L)$$

is faithful (a monomorphism).

The *Killing form* is the nondegenerate, symmetric, associative form on L given by

$$\kappa(x, y) = \mathrm{Tr}(\mathrm{ad}\, x \ \mathrm{ad}\, y)$$

For any $x \in L$ we have the *abstract Jordan decomposition* $x = x_s + x_n$ where x_s, x_n are uniquely determined by the formula $\mathrm{ad}\, x_s = (\mathrm{ad}\, x)_s$, $\mathrm{ad}\, x_n = (\mathrm{ad}\, x)_n$. For any representation $\varphi : L \to \mathfrak{gl}(V)$, $\varphi(x)_s = \varphi(x_s)$ and $\varphi(x)_n = \varphi(x_n)$. (Here is another proof of this: Take $\psi = (\mathrm{ad}, \varphi) : L \hookrightarrow \mathfrak{gl}(L \oplus V)$. Then by Thm 6.5.1, $\forall x, \exists y \in L$ so that $\psi(x)_s = \psi(y) = ((\mathrm{ad}\, x)_s, \varphi(x)_s) = (\mathrm{ad}\, y, \varphi(y))$. This implies $y = x_s$.) Since x_s, x_n are polynomials in x, anything that commutes with x also commutes with x_s and x_n. In particular $[xy] = 0$ implies $[x_s y] = [x_n y] = 0$.

Engel's Theorem If every element of a Lie algebra L is ad-nilpotent, then L is nilpotent.

Exercise 8.0.1.
(1) (crucial observation) If $[xy] = 0$ then show that $\kappa(x_n, y) = 0$.
(2) If C is a nilpotent Lie algebra which is not abelian then show that $Z(C) \cap [CC] \neq 0$.
(3) If $S \subseteq \mathfrak{gl}(V)$ is solvable then $[SS]$ is nilpotent.

8.1. **Cartan subalgebra H.** By Engel's Theorem, the semisimple Lie algebra L has at least one element x which is not nilpotent. Then $x_s \neq 0$. The subspace spanned by x_s is an abelian subalgebra of L all of whose elements are semisimple.

Definition 8.1.1. The *Cartan subalgebra* of a semisimple Lie algebra L is defined to be a maximal abelian subalgebra of L consisting only of semisimple elements.

8.1.1. *preview.* We choose any Cartan subalgebra H of L. Before proceeding, we will list the important properties that we want to prove about H.

(1) $\mathrm{ad}_L H$ is simultaneously diagonalizable. This gives us the *root space decomposition* of L:
$$L = L_0 \oplus \bigoplus_{\alpha \in H^*} L_\alpha$$
where
$$L_\alpha = \{x \in L \,|\, [hx] = \alpha(h)x\}$$
(2) $L_0 = H = C_L(H)$, i.e., H is *self-centralizing.*

Definition 8.1.2. $\alpha \in H^*$ is called a *root* if $\alpha \neq 0$ and $L_\alpha \neq 0$. The set of roots is denoted Φ.

$\Rightarrow$ The root space decomposition is

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

(3) Each $L_\alpha, \alpha \in \Phi$ is 1-dimensional.
(4) $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$ for all $\alpha, \beta \in H^*$.
(5) The restriction κ_H of the Killing form κ to $L_0 = H$ is nondegenerate. The dual of κ_H gives an inner product $(\cdot,\cdot)$ on H^*.
(6) H^* with inner product $(\cdot,\cdot)$ and the set of roots $\Phi \subset H^*$ completely determines the multiplication table for L.

8.1.2. *basic properties.* We prove the easy properties first.

Proposition 8.1.3. $\mathrm{ad}_L H$ *is simultaneously diagonalizable.*

Proof. Chose any basis $h_1, \cdots, h_n$ for H. Starting with h_1, choose a diagonalization of $\mathrm{ad}\, h_1$. Then $L \cong \bigoplus L_\lambda$ where L_λ is the λ eigenspace of h_1. Since H is abelian, $h_2, \cdots, h_n$ act as commuting semisimple operators on each L_λ. By induction on n, $\mathrm{ad}\, h_2, \cdots, \mathrm{ad}\, h_n$ are simultaneously diagonalizable. This gives a diagonalization of $\mathrm{ad}_L H$. □

Remark 8.1.4. This proof shows that any linear combination of commuting semisimple elements of L is semisimple.

As pointed out in the preview, this gives the root space decomposition

$$L = L_0 \oplus \bigoplus_{\alpha\in H^*} L_\alpha$$

Proposition 8.1.5. (1) $[L_\alpha L_\beta] \subseteq L_{\alpha+\beta}$.
(2) *If* $\alpha \neq 0$ *then any element of* L_α *is* ad-*nilpotent.*
(3) $\kappa(L_\alpha, L_\beta) = 0$ *if* $\alpha + \beta \neq 0$.

Proof. (1) Suppose $x \in L_\alpha, y \in L_\beta$ and $h \in H$. Then

$$\mathrm{ad}\, h[xy] = [\mathrm{ad}\, h(x), y] + [x, \mathrm{ad}\, h(y)] = [\alpha x, y] + [x, \beta y] = (\alpha + \beta)[xy]$$

So, $[xy] \in L_{\alpha+\beta}$.

(2) $\mathrm{ad}\, x$ is nilpotent since it sends L_β to $L_{\beta+\alpha}$.

(3) Since $\alpha + \beta \neq 0$ there is some $h \in H$ so that $\alpha(h) + \beta(h) \neq 0$. Then for any $x \in L_\alpha, y \in L_\beta$ we have:

$$\alpha(h)\kappa(x, y) = \kappa([hx], y) = -\kappa(x, [hy]) = -\beta(h)\kappa(x, y)$$

So, $(\alpha(h) + \beta(h))\kappa(x, y) = 0$ making $\kappa(x, y) = 0$. □

Corollary 8.1.6. *The restriction of* κ *to* $L_0 = C_L(H)$ *is nondegenerate.*

Proof. We know that κ is nondegenerate on L. So, $\kappa(h, L) = 0$ iff $h = 0$. If the restriction of κ to L_0 is degenerate then there is an $h \in L_0$ so that $\kappa(h, L_0) = 0$. But $\kappa(h, L_\alpha) = 0$ for all $\alpha \in \Phi$ by (3) above. So, $\kappa(h, L) = 0$ making $h = 0$. □

8.2. **Centralizer of H.** Let $C = L_0 = C_L(H)$. We want to show that $C = H$. Since H is abelian, we know that $H \subseteq C$.

Theorem 8.2.1. *Any Cartan subalgebra of L is self-centralizing: $H = C_L(H)$.*

Proof. The outline in Humphreys works even though we used a different definition of H. Let $C = C_L(H)$. Then $H \subseteq Z(C)$.

Claim 1. If $x \in C$ then $x_s, x_n \in C$.

Pf: $x \in C$ iff $[xH] = 0$. This implies $[x_sH] = 0$ and $[x_nH] = 0$ as I pointed out in the review. So, $x_s, x_n \in C$.

Claim 2. All semisimple elements of C lie in H.

Pf: If $x = x_s \in C$ then the span of x, H is an abelian subalgebra of L all of whose elements are semisimple by Remark 8.1.4. By maximality of H, $x \in H$.

Claim 3. The restriction of κ to H is nondegenerate.

Pf: Suppose that $h \in H$ so that $\kappa(h, H) = 0$. To show that $h = 0$ it suffices to show that $\kappa(h, C) = 0$. So, let $x \in C$. Then $x = x_s + x_n$ where $x_s \in H$ by Claim 2. So, $\kappa(h, x_s) = 0$. But $[hx] = 0$ implies $\kappa(h, x_n) = 0$ by Exercise 8.0.1. So, $\kappa(h, C) = 0$ making $h = 0$ since κ is nondegenerate on C.

Claim 4. C is nilpotent.

Pf: Take any element $x \in C$. Then $x = x_s + x_n$. But $x_s \in H \subseteq Z(C)$ by Claim 2. So, $\mathrm{ad}_C x = \mathrm{ad}_C x_n$ is nilpotent. So, C is nilpotent by Engel's Theorem.

Claim 5. $H \cap [CC] = 0$.

Pf: Since $[HC] = 0$, $\kappa([HC], C) = 0 = \kappa(H, [CC])$. Since κ is nondegenerate on H by Claim 3, $H \cap [CC]$ must be 0.

Claim 6. C is abelian.

Pf: We use (1) and (2) of Exercise 8.0.1. If C is not abelian then $Z(C) \cap [CC] \neq 0$ by part (2). Let $x \neq 0 \in Z(C) \cap [CC]$. Then $[xC] = 0$. So, $\kappa(x_n, C) = 0$ by part (1). Therefore, $x_n = 0$ since $\kappa|C$ is nondegenerate. So, $x = x_s \in H$ contradicting Claim 5.

Claim 7. $C = H$.

Pf: For any $x \in C$, $[xC] = 0$ implies $\kappa(x_n, C) = 0$. So, $x_n = 0$. Thus $x = x_s \in H$. □

Corollary 8.2.2. (1) *The restriction of κ to H is nondegenerate.*
(2) *There is a linear isomorphism $H^* \cong H$ sending $\varphi \in H^*$ to $t_\varphi \in H$ so that*

$$\varphi(h) = \kappa(t_\varphi, h)$$

for all $h \in H$. □

8.3. **Embedded $\mathfrak{sl}(2, F)$.** For any $\alpha \in \Phi$ we will find a subalgebra of L isomorphic to $\mathfrak{sl}(2, F)$ with generators $x_\alpha \in L_\alpha, y_\alpha \in L_{-\alpha}$ and $h_\alpha \in L_0 = H$.

Lemma 8.3.1. *If $x \in L_\alpha, y \in L_{-\alpha}$ then $[xy] = \kappa(x, y)t_\alpha$.*

Proof. For any $h \in H$ we have

$$\kappa(h, [xy]) = \kappa([hx], y) = \alpha(h)\kappa(x, y)$$

Since $\kappa(h, t_\alpha) = \alpha(h)$, the Lemma follows. □

Take $x \neq 0 \in L_\alpha$. Then $\kappa(x, -) = 0$ on all L_β for $\beta \neq -\alpha$. Since κ is nondegenerate, there must be some $y \in L_{-a}$ so that $\kappa(x, y) \neq 0$. This gives:

Proposition 8.3.2. *If $\alpha \in \Phi$ then $-\alpha \in \Phi$ and $[L_\alpha L_{-\alpha}] = Ft_\alpha$.* □

Next, we need to find an element $h_\alpha \in [L_\alpha L_{-\alpha}]$ so that $\alpha(h) = 2$. This will imply that

$$(\forall x \in L_\alpha)\ [hx] = \alpha(h)x = 2x, \quad (\forall y \in L_{-\alpha})\ [hy] = -2y$$

We can then take the appropriate scalar multiples of x, y to get $[xy] = h$. Then x, y, h will span a subalgebra of L isomorphic to $\mathfrak{sl}(2, F)$ and we will use the notation $x_\alpha, y_\alpha, h_\alpha$ for this choice of x, y, h. By the proposition, it suffices to prove the following.

Lemma 8.3.3. *$\alpha(t_\alpha) \neq 0$.*

Proof. Suppose $\alpha(t_\alpha) = 0$. Then $[t_\alpha x] = 0 = [t_\alpha y]$ for all $x \in L_\alpha, y \in L_{-\alpha}$. By the proposition above, we can choose x, y so that $[xy] = t_\alpha$. Then the span S of x, y, t_α is a solvable subalgebra of $L \subseteq \mathfrak{gl}(L)$ with $[SS] = Ft_\alpha$. This implies that $\operatorname{ad} t_\alpha$ is nilpotent as we reviewed in Exercise 8.0.1. But the Jordan decomposition is unique and the only element which is both semisimple and nilpotent is 0. So $t_\alpha = 0$ which contradicts it definition. □

By the discussion preceding the Lemma, this proves the following.

Theorem 8.3.4. *For all $\alpha \in \Phi$ there are elements $x_\alpha \in L_\alpha$, $y_\alpha \in L_{-\alpha}$ and $h_\alpha \in H$ so that $x_\alpha, y_\alpha, h_\alpha$ span a subalgebra of L isomorphic to $\mathfrak{sl}(2, F)$ and they correspond to the standard generators x, y, h. Furthermore, $\alpha(h_\alpha) = 2$.* □

Although we have some freedom in choosing x_α, y_α we must have:

$$h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$$

This follows from the fact that h_α must be some multiple of t_α and

$$\alpha(h_\alpha) = \kappa(t_\alpha, h_\alpha) = 2$$

Also, note that $t_{-\alpha} = -t_\alpha$. So, we must have $h_{-\alpha} = -h_\alpha$. Also, we may choose $x_{-\alpha} = y_\alpha$ and $y_{-\alpha} = x_\alpha$.

Let $S_\alpha \cong \mathfrak{sl}(2, F)$ be the span of $x_\alpha, y_\alpha, h_\alpha$. Then L is a module over S_α. Using our knowledge of all representations of $\mathfrak{sl}(2, F)$, we will be able to give a very complete description of the structure of the semisimple Lie algebra L.

8.4. **Root strings.** This subsection is based on Erdmann and Wildon "Introduction to Lie Algebras" [2] an undergraduate textbook which is the place to look if you don't understand something. We first review what we have so far using an example.

8.4.1. *example:* $\mathfrak{sl}(3,F)$. Let $L = \mathfrak{sl}(3,F)$. This is 8 dimensional with H being the 2-dimensional subalgebra of diagonal matrices with trace zero.

$$H = \left\{ \begin{bmatrix} h_1 & 0 & 0 \\ 0 & h_2 & 0 \\ 0 & 0 & h_3 \end{bmatrix} : h_1 + h_2 + h_3 = 0 \right\}$$

The off-diagonal entries have an obvious basis given by x_{ij}, the matrix with 1 in the ij position and 0 elsewhere:

$$x_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad x_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad x_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and x_{21}, x_{32}, x_{31}. Note that each x_{ij} is an eigenvector: For example,

$$[hx_{12}] = hx_{12} - x_{12}h = (h_1 - h_2)x_{12}$$

So,

(1) $x_{12} \in L_\alpha$ where $\alpha(h) = h_1 - h_2$,
(2) $x_{23} \in L_\beta$ where $\beta(h) = h_2 - h_3$ and
(3) $x_{13} \in L_{\alpha+\beta}$ with $(\alpha+\beta)(h) = h_1 - h_3$.
We also have
(4) $x_{21} \in L_{-\alpha}$
(5) $x_{32} \in L_{-\beta}$
(6) $x_{31} \in L_{-\alpha-\beta}$.

This gives the root space decomposition:

$$L = H \oplus L_\alpha \oplus L_\beta \oplus L_{\alpha+\beta} \oplus L_{-\alpha} \oplus L_{-\beta} \oplus L_{-\alpha-\beta}$$

What is h_α?

$$[x_{12}x_{21}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = h_\alpha$$

This matrix is h_α since $\alpha(h_\alpha) = h_1 - h_2 = 2$. $x_a = x_{12}, y_\alpha = x_{21}$ and

$$S_\alpha = span(x_{12}, x_{21}, h_\alpha)$$

Find the decomposition of L into irreducible S_α-modules.

One component is $V(2) = L_\alpha \oplus Fh_\alpha \oplus L_{-\alpha}$. To find the others we draw the root spaces in the following pattern:

$$\begin{array}{ccccccc} & L_{\alpha+\beta} & \oplus & L_\beta & & & \cong V(1) \\ L_\alpha & \oplus & H & \oplus & L_{-\alpha} & & \cong V(2)\oplus V(0) \\ & L_{-\beta} & \oplus & L_{-\alpha-\beta} & & & \cong V(1) \end{array}$$

Claim: $L_{\alpha+\beta} \oplus V_\beta \cong V(1)$.

Pf: Since $x_\alpha \in L_\alpha$, $\operatorname{ad} x_\alpha$ sends $L_{\alpha+\beta}$ to $L_{2\alpha+\beta} = 0$. Therefore, $x_{13} = x_{\alpha+\beta}$ is a maximal vector. But $h_\alpha(x_{13}) = (h_1 - h_3)x_{13} = x_{13}$. So, the weight is 1. So, x_{13} generates a submodule isomorphic to $V(1)$. (Recall that $V(\lambda)$ is generated by a maximal vector with maximal weight λ and that λ is a nonnegative integer.)

Similarly, $L_{-\beta} \oplus L_{-\alpha-\beta} \cong V(1)$. This leaves only a one dimensional submodule $V(0)$ contained in H.

8.4.2. *α-root string.* We now return to the general case. Recall that we have a root space decomposition:

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

Definition 8.4.1. If $\alpha \in \Phi$ and $\beta \in \Phi \cup \{0\}$, the *$\alpha$-root string through β* is the S_α-module:

$$M = \bigoplus_{c\in F} L_{\beta+c\alpha}$$

(We will prove shortly that only integer values of c can occur in any root string.)

Proposition 8.4.2. *For any $\alpha \in \Phi$, L is a direct sum of α-root strings.*

From the example we know that the α-root string though 0 is not irreducible in general.

Proposition 8.4.3. *If $\alpha \in \Phi$ then L_α is 1-dimensional. Furthermore, the only multiples of α which are roots are $\pm\alpha$.*

Proof. Consider the α-root string though 0:

$$M = H \oplus \bigoplus_{c\alpha\in\Phi} L_\alpha$$

Let $K = \ker(\alpha : H \to F)$. Then $H = K \oplus Fh_\alpha$.

Claim: K is an S_α-module.

Pf: For any $k \in K$ we have $[x_\alpha k] = -\alpha(k)x_\alpha = 0$. Similarly, $[y_\alpha k] = 0$ and $[hk] = 0$ since $h, k \in H$. So, every nonzero element of K generated an S_α-module isomorphic to $V(0)$. This implies that

$$M/K \cong W = Fh_\alpha \oplus \bigoplus_{c\alpha\in\Phi} L_{c\alpha}$$

is an S_α-module. But the weight zero part of W is $W_0 = Fh_\alpha$. So, W contains only one irreducible $V(even)$. Since we know that W contains $S_\alpha \cong V(2)$, there cannot be any other $V(even)$. So, $L_{2\alpha} = 0$. In other words, twice a root cannot be a root. But then $\frac{1}{2}\alpha$ also cannot be a root since twice of it is a root. And this implies that $W_1 = L_{\alpha/2} = 0$. So, $W = V(2)$ is irreducible. This also implies that L_α is one dimensional. □

Corollary 8.4.4. *If $\alpha, \beta \in \Phi$ and β is not $\pm\alpha$ then the α-root string through β is irreducible and has the form:*

$$V(m) \cong L_{\beta+q\alpha} \oplus L_{\beta+(q-1)\alpha} \oplus \cdots \oplus L_{\beta-r\alpha}$$

and $m = q + r$. And $\beta(h_\alpha) = r - q$ is an integer. ($\beta(h_\alpha)$ are the Cartan integers*.)*

Proof. Let M be the α-root string through β. Then $M_0 = 0$ since the root string does not go through 0. (The proof of the last proposition showed that $S_\alpha = L_\alpha \oplus Fh_\alpha \oplus L_{-\alpha}$ is the only α-root string through 0.) Therefore, M is a direct sum of $V(odd)$s. So, the α-root string has only M_{odd} but a difference of 2 in h_α-weights corresponds to a difference of α in roots since $\alpha(h_\alpha) = 2$. So, the root string contains only $L_{\beta+k\alpha}$ for integer k. One of these is M_1 and thus is 1-dimensional. So, M is irreducible.

If $M \cong V(m)$ then $M_m = L_{\beta+q\alpha}$ and $M_{-m} = L_{\beta-r\alpha}$. The dimension of M is $q+r+1 = m+1$. This implies that

$$q + r = m = (\beta + q\alpha)(h_\alpha) = \beta(h_\alpha) + 2q$$

So, $\beta(h_\alpha) = r - q \in \mathbb{Z}$. □

8.5. **Inner product on H^*.** We just proved that $\beta(h_\alpha) \in \mathbb{Z}$. We will rephrase this in terms of the inner product on H^*.

Recall that $h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$. And, by definition of t_β we have:

$$\beta(h_\alpha) = \kappa(t_\beta, h_\alpha) = \frac{2\kappa(t_\beta, t_\alpha)}{\kappa(t_\alpha, t_\alpha)} = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z}$$

where we used the definition of the inner product on H^*:

$$(\alpha, \beta) := \kappa(t_\alpha, t_\beta).$$

Proposition 8.5.1. *The set of roots Φ spans H^*.*

Proof. Suppose not. Then there is some $h \in H$ so that $\alpha(h) = 0$ for all $\alpha \in \Phi$. But then $[hx] = \alpha(h)x = 0$ for all $x \in L_\alpha$ and $[hx] = 0$ for all $x \in H$ since H is abelian. So, h commutes with all the generators of L and is therefore in $Z(L)$. But $Z(L) = 0$ since L is semisimple. □

This means we can choose a basis for H^* consisting of roots: $\alpha_1, \cdots, \alpha_n$. If $\beta \in \Phi$ then $\beta = \sum c_i\alpha_i$ where $c_i \in F$. Recall that we are assuming F is algebraically closed with characteristic 0. Thus F contains $\mathbb{Q}$, the rational numbers.

Claim 1: $c_i \in \mathbb{Q}$.

Pf: $(\beta, \alpha_j) = \sum c_i(\alpha_i, \alpha_j)$. Therefore,

$$\frac{2(\beta, \alpha_j)}{(\alpha_j, \alpha_j)} = \sum c_i \frac{2(\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)}$$

These fractions are Cartan integers. Also the matrix (α_i, α_j) is nonsingular since the form is nondegenerate. Therefore, c_i is the ratio of two determinants of integer matrices and therefore a rational number.

This implies that the roots lie in the $\mathbb{Q}$-span of $\alpha_1, \cdots, \alpha_n$. Let $E_{\mathbb{Q}}$ denote this rational vector space.

Claim 2: The inner product $(\cdot, \cdot)$ on $E_{\mathbb{Q}}$ has rational values and is positive definite.

Pf: Take any $\lambda \in H^*$. Then

$$(\lambda, \lambda) = \kappa(t_\lambda, t_\lambda) = \operatorname{Tr}(\operatorname{ad} t_\lambda \operatorname{ad} t_\lambda) =_{(1)} \sum_{\beta \in \Phi} \beta(t_\lambda)^2 =_{(2)} \sum_{\beta \in \Phi} (\beta, \lambda)^2$$

where (1) follows from the fact that $\operatorname{ad} t_\lambda$ acts on L_β by multiplication by $\beta(t_\lambda)$ and (2) follows from the observation that $\beta(t_\lambda) = \kappa(t_\beta, t_\lambda) = (\beta, \lambda)$. Thus $(\cdot, \cdot)$ is positive semi-definite. Since it is nondegenerate, it must be positive definite.

Let $E = E_{\mathbb{Q}} \otimes \mathbb{R}$.

Theorem 8.5.2. (1) *E is a real vector space with a positive definite inner product, i.e., E is Euclidean space.*

(2) *Φ spans E and $0 \notin \Phi$.*
(3) *If $\alpha \in \Phi$ then $-\alpha \in \Phi$ and not other real multiple of α is in Φ.*
(4) *If $\alpha, \beta \in \Phi$ then $\frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$.*
(5) *If $\alpha, \beta \in \Phi$ then*

$$\beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha \in \Phi$$

Proof. We proved (1)-(4). The class figured out the proof of (5):

$$\frac{2(\beta, \alpha)}{(\alpha, \alpha)} = \beta(h_\alpha) = r - q$$

in the notation of the α-root string through β. So,

$$\beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha = \beta + (q - r)\alpha$$

and $q \geq q - r \geq -r$. So, this is one of the roots that occur in the α string through β. $\square$

9. Abstract root systems

We now attempt to reconstruct the Lie algebra based only on the information given by the set of roots Φ which is embedded in Euclidean space E.

9.1. **Definition.** Any finite subset Φ of Euclidean space E satisfying the conditions of Theorem 8.5.2 will be called a root system. To repeat:

Definition 9.1.1. A *root system* is defined to be a subset Φ of standard Euclidean space E (i.e., finite dimensional real vector space with a positive definite inner product $(\cdot,\cdot)$) satisfying the following.

(1) Φ is finite, spans E and $0 \notin \Phi$.
(2) If $\alpha \in \Phi$ then $-\alpha \in \Phi$ and not other real multiple of α is in Φ.
(3) If $\alpha, \beta \in \Phi$ then
$$\langle \beta, \alpha \rangle := \frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$$
(4) If $\alpha, \beta \in \Phi$ then
$$\beta - \langle \beta, \alpha \rangle \alpha \in \Phi$$

Note that $\langle \alpha, \alpha \rangle = 2$.

Proposition 9.1.2. *Let $\sigma_\alpha : E \to E$ be the linear mapping given by*
$$\sigma_\alpha(x) = x - \langle x, \alpha \rangle \alpha$$
Then σ_α is reflection through the hyperplane perpendicular to α.

Proof. σ_α sends α to $-\alpha$ and leaves fixed every vector perpendicular to α. □

Condition (4) says that $\sigma_\alpha(\Phi) = \Phi$ for all $\alpha \in \Phi$.

Definition 9.1.3. The *Weyl group* W is defined to be the subgroup of $GL(E)$ generated by the reflections σ_α for all $\alpha \in \Phi$.

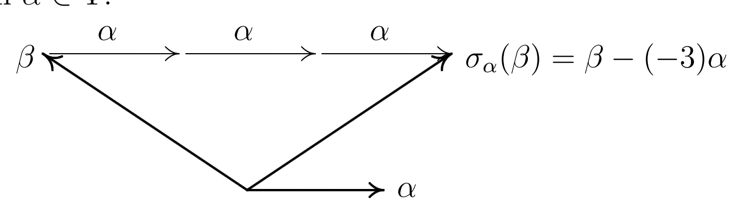


Figure 9.1.1. In this example, $\langle \beta, \alpha \rangle = -3$ and $\sigma_\alpha(\beta) = \beta + 3\alpha \in \Phi$. We will prove later that the vectors in the middle: $\beta + \alpha, \beta + 2\alpha$ are also in Φ.

Example 9.1.4. (1)
$$A_1^n = \{\pm\epsilon_i \,|\, 1 \leq i \leq n\} \quad \subseteq E = \mathbb{R}^n$$
is a root system. σ_{ϵ_i} sends ϵ_i to $-\epsilon_i$ and leaves the other unit vectors ϵ_j fixed. The matrix of this reflection map is the diagonal matrix with -1 in the ith position with other diagonal entries equal to 1. Therefore, the Weyl group is the group of diagonal matrices with ± 1 on the diagonal: $W = \mathbb{Z}_2^n = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \cdots \times \mathbb{Z}_2$. $(\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z})$

(2)
$$A_{n-1} = \{\pm(\epsilon_i - \epsilon_j) \mid 1 \le i < j \le n\}$$
is a root system in $E = \{x \in \mathbb{R}^n \mid \sum x_i = 0\}$. For $\alpha = \epsilon_i - \epsilon_j$, the reflection σ_α switches ϵ_i and ϵ_j and leaves the other ϵ_k fixed. This clearly sends the set A_{n-1} to itself. The reflections are transpositions. So, the Weyl group is the symmetric group S_n.

(3)
$$D_n = \{\pm(\epsilon_i \pm \epsilon_j) \mid 1 \le i < j \le n\}$$
is a root system. For $\alpha = \epsilon_i + \epsilon_j$, the reflection σ_α switches ϵ_i and $-\epsilon_j$ (and $-\epsilon_i \leftrightarrow \epsilon_j$) and leaves the other ϵ_k fixed. This clearly sends the set D_n to itself. One can show that the Weyl group is the group of signed permutations with an even number of negative signs.

Exercise 9.1.5. Find an explicit isomorphism $D_3 \cong A_3$.

9.2. **Two roots.**

Lemma 9.2.1. *Suppose that $\alpha, \beta \in \Phi$ and $\beta \neq \pm\alpha$. Then*
$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 0, 1, 2 \text{ or } 3.$$

Proof. This follows from the Schwartz inequality:
$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = \frac{4(\alpha,\beta)(\beta,\alpha)}{(\beta,\beta)(\alpha,\alpha)} = \frac{4(\alpha,\beta)^2}{||\beta||^2 ||\alpha||^2} \le 4$$
where equality holds iff α, β are collinear. □

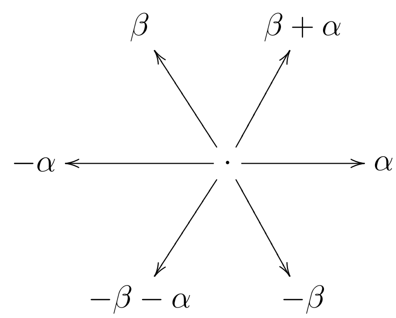
This is also equal to $4\cos^2\theta$ where θ is the angle between α, β.
$$\begin{array}{rcllll} \cos^2\theta &=& 0 & \frac14 & \frac12 & \frac34 \\ \pm\cos\theta &=& 0 & \frac12 & \frac{\sqrt2}{2} & \frac{\sqrt3}{2} \\ \theta \text{ or } \pi-\theta &=& \frac{\pi}{2} & \frac{\pi}{3} & \frac{\pi}{4} & \frac{\pi}{6} \end{array}$$

If the product $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle$ is 0 then α, β are perpendicular and $\sigma_\alpha(\beta) = \beta$. This is illustrated in the first example above.

If the product of $\langle \alpha, \beta \rangle, \langle \beta, \alpha \rangle$ is nonzero, one of them must be ±1. By symmetry assume $\langle \alpha, \beta \rangle = \pm1$. Then $\langle \beta, \alpha \rangle = \pm1, \pm2, \pm3$.
$$|\langle \beta, \alpha \rangle| = \frac{\langle \beta, \alpha \rangle}{\langle \alpha, \beta \rangle} = \frac{2(\beta,\alpha)/||\alpha||^2}{2(\beta,\alpha)/||\beta||^2} = \frac{||\beta||^2}{||\alpha||^2} = 1, 2, 3$$
This looks like 6 cases, but this reduces to 3 cases with the following observation. Let
$$\gamma = \sigma_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha$$
Then
$$\langle \gamma, \alpha \rangle = \langle \beta, \alpha \rangle - 2\langle \beta, \alpha \rangle = -\langle \beta, \alpha \rangle$$
So, by replacing β with γ if necessary, we may assume that $\langle \beta, \alpha \rangle$ is *negative*.

9.2.1. *root system A_2.* Suppose that $\langle \beta, \alpha \rangle = -1$. Then $||\alpha|| = ||\beta||$ and $\theta = 2\pi/3$. Also

$$\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha = \beta + \alpha$$



This is the root system of $\mathfrak{sl}(3, F)$. Also, it is a special case of example 2 above with $n = 3$. The correspondence is

$$\alpha = \epsilon_1 - \epsilon_2, \quad \beta = \epsilon_2 - \epsilon_3 \qquad (\alpha + \beta = \epsilon_1 - \epsilon_3)$$

with $||\alpha|| = ||\beta|| = \sqrt{2}$ and

$$\langle \beta, \alpha \rangle = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} = \frac{-2}{2} = -1$$

In terms of H, the Cartan subalgebra of $L = \mathfrak{sl}(3, F)$, $\epsilon_i \in H^*$ is defined by $\epsilon_i(h) = h_i$. So,

$$\alpha(h) = \epsilon_1(h) - \epsilon_2(h) = h_1 - h_2$$

which agrees with the earlier terminology.

Exercise 9.2.2. Generalize this correspondenc to show that Example 2 is the root system for $\mathfrak{sl}(n, F)$.

An animation of $A_3 = D_3$, the root system of $\mathfrak{sl}(4, F)$ is on the webpage. In that rotating figure, the green arrow is $\alpha = \epsilon_1 - \epsilon_2$, the red arrow is $\beta = \epsilon_2 - \epsilon_3$ and the blue arrow is $\gamma = \epsilon_3 - \epsilon_4$. The angles are:

$$\langle \alpha, \beta \rangle = 2, \quad \theta = 2\pi/3$$

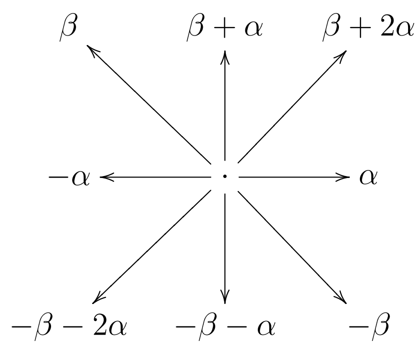
$$\langle \beta, \gamma \rangle = 2, \quad \theta = 2\pi/3$$

$$\langle \alpha, \gamma \rangle = 0, \quad \theta = \pi/2$$

The white arrow is the sum of these three roots: $\alpha + \beta + \gamma = \epsilon_1 - \epsilon_4$. Note that all roots have the same length: $\sqrt{2}$.

9.2.2. *root system* $B_2 = C_2$. Suppose that $\langle \beta, \alpha \rangle = -2$. Then $||\beta|| = \sqrt{2}||\alpha||$ and $\theta = 3\pi/4$. Also

$$\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha = \beta + 2\alpha$$



This root system has *short roots* of length 1 and *long roots* of length $\sqrt{2}$.

Exercise 9.2.3. Show that the 26 vectors in $\mathbb{R}^3$ given by

$$\pm\epsilon_i(\text{ 6 vectors}), \pm\epsilon_i \pm \epsilon_j(\text{ 12 vectors}), \pm\epsilon_1 \pm \epsilon_2 \pm \epsilon_3(\text{ 8 vectors})$$

do not form a root system. (These vectors form a $2 \times 2 \times 2$ cube just as the figure above forms a 2×2 square.)

Definition 9.2.4. The root system B_n in $\mathbb{R}^n$ is defined to be the union of the set of $2n$ short roots $\pm\epsilon_i$ and the $4\binom{n}{2}$ long roots $\pm\epsilon_i \pm \epsilon_j$.

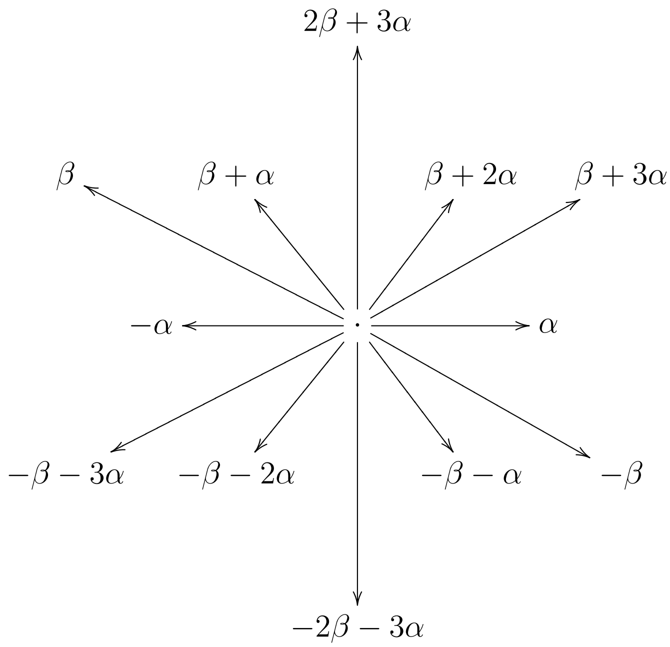
The long roots of B_n form a subsystem D_n as in Example 9.1.4 (3) and the short roots form A_1^n as in Example 9.1.4 (1). For example, for $n = 3$, there are 6 short roots and 12 long roots and these long roots form $D_3 = A_3$.

Definition 9.2.5. The root system C_n in $\mathbb{R}^n$ is defined to be the union of the set of $2n$ long roots $\pm 2\epsilon_i$ and the $4\binom{n}{2}$ short roots $\pm\epsilon_i \pm \epsilon_j$.

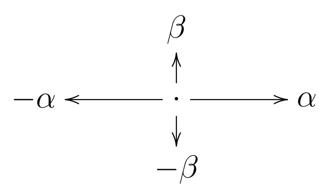
The short roots of C_n form a subsystem D_n. For $n = 3$ there are 12 short roots and 6 long roots. The case B_3, C_3 can be seen in these animations: B_3, C_3. The short roots are red and the long roots are blue in both.

9.2.3. *root system* G_2. Suppose that $\langle \beta, \alpha \rangle = -3$. Then $||\beta|| = \sqrt{3}||\alpha||$ and $\theta = 5\pi/6$. Also

$$\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha = \beta + 3\alpha$$



9.2.4. *root system* $A_1 \times A_1$. When $\langle \beta, \alpha \rangle = 0$ the two roots are perpendicular. They could have any length.



This is the root system of $\mathfrak{sl}(2, F) \times \mathfrak{sl}(2, F)$.

9.3. **Irreducible root systems.**

Definition 9.3.1. We say that a root system Φ *decomposes* if it is a disjoint union $\Phi_1 \cup \Phi_2$ of two nonempty subsets so that every root in Φ_1 is perpendicular to every root of Φ_2. We say that Φ is *irreducible* if there is no such decomposition.

If Φ decomposes, then E also decomposes as an orthogonal direct sum $E = E_1 \oplus E_2$ where E_i is the span of Φ_i. Each $\Phi_i \subset E_i$ is a root system.

Exercise 9.3.2. Show that the root system of a product $L = L_1 \times L_2$ of two semisimple Lie algebras decomposes as the union of the root systems of L_1, L_2. Conversely, any decomposition of the root system of L comes from such a factorization of L.

Proposition 9.3.3. *The decomposition of a root system Φ into irreducible components is unique.*

Proof. If $\Phi = \bigcup \Phi_i$ and $\Phi = \bigcup \Psi_j$ are any two decompositions then so is $\Phi = \bigcup \Phi_i \cap \Psi_j$. $\square$

9.4. **Bases for root systems.**

Definition 9.4.1. A subset $\Delta = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ of Φ is called a *base* for the root system if

1. Δ is a basis for E
2. Every element $\beta \in \Phi$ can be written as $\beta = \sum k_i \alpha_i$ where all k_i are integers with the same sign (or zero).

Given a choice of base Δ, the root in Δ are called *simple roots*, roots which are positive linear combinations of simple roots are called *positive roots* the set of positive roots is denoted Φ_+. The set of *negative roots* (negatives of positive roots) is Φ_-. Thus

$$\Phi = \Phi_+ \cup \Phi_-$$

Example 9.4.2. In the root systems $A_2, B_2, G_2, A_1 \times A_1$ above, the roots α, β were chosen to form a base for the root system.

Example 9.4.3. In the root system

$$A_{n-1} = \{\epsilon_i - \epsilon_j \in \mathbb{R}^n\}$$

the roots $\alpha_i := \epsilon_i - \epsilon_{i+1}$, $i = 1, \cdots, n-1$ form a base. In the A_3-animation, the simple roots are green, red and blue.

Example 9.4.4. In the second set of animations for the root systems B_3 and C_3 the simple roots have green tips, and the positive roots are in the foreground (with negative roots behind a semitransparent plane).

Lemma 9.4.5. *If $\alpha, \beta \in \Phi$ and $(\alpha, \beta) < 0$ then $\alpha + \beta$ is a root. Similarly, if $(\alpha, \beta) > 0$ then $\alpha - \beta$ is a root.*

Proof. By symmetry we may assume $\langle \alpha, \beta \rangle = \pm 1$ and this has the same sign as (α, β). Then

$$\sigma_\beta(\alpha) = \begin{cases} \alpha + \beta & \text{if } (\alpha, \beta) < 0 \\ \alpha - \beta & \text{if } (\alpha, \beta) > 0 \end{cases}$$

and $\sigma_\beta(\alpha) \in \Phi$ by definition of root system. □

Theorem 9.4.6. *Every root system has a base.*

Proof. First choose a vector x which is not perpendicular to any root. Since there are only finitely many roots and their perpendicular hyperplane have zero measure, their union also has zero measure. Take any point in the complement.

Define a root $\beta \in \Phi$ to be *positive* if $(\beta, x) > 0$. Let S be the set of all real numbers $(\beta, x) > 0$ for $\beta \in \Phi$. Then S is finite. Define a positive root to be *indecomposable* if it cannot be written as a sum of other positive roots.

Claim 1 Every positive root β can be written as a sum of indecomposable positive roots.

Pf: This holds by induction on the number of elements in the set S which are less than (β, x). If this number is zero then β is indecomposable. If β decomposes into a sum $\beta = \beta_1 + \beta_2$, we have $(\beta_i, x) < (\beta, x)$. So, by induction, each β_i is a sum of indecomposable positive roots.

<u>Claim 2</u> If α, β are indecomposable positive roots then $(\alpha, \beta) \leq 0$.

Pf: If $(\alpha, \beta) > 0$ then $\gamma = \alpha - \beta$ is a root and therefore either γ or $-\gamma$ is a positive room. In the first case, $\alpha = \beta + \gamma$ is not indecomposable. In the second case, $\beta = \alpha + (-\gamma)$ is not indecomposable.

<u>Claim 3</u> Let Δ be the set of indecomposable positive roots. Then Δ is a base for Φ.

Pf: Since the second condition in the definition of a base is satisfied by construction, it suffices to show that Δ is a basis for E. Since Δ spans Φ it also spans E. So, it suffices to show Δ is linearly independent.

Suppose not. Then there is a linear dependence $\sum k_i \alpha_i = 0$ where $\alpha_i \in \Delta$. Separate the positive and negative coefficients κ_i to obtain a relation:

$$\sum s_i \alpha_i = \sum t_j \alpha_j$$

where s_i, t_j are all ≥ 0 and $\alpha_i \neq \alpha_j \in \Delta$. Let z denote this sum. Then

$$(z, z) = \sum s_i t_j (\alpha_i, \alpha_j) \leq 0$$

by Claim 2. So, $z = 0$ proving Claim 3 and the theorem. □

10. Weyl group and Weyl chambers

We will use the Weyl group and the geometry of Weyl chambers to prove basic properties of root systems, such as the uniqueness of the base up to isomorphism.

Recall that the Weyl group is the subgroup W of $GL(E)$ generated by the reflections σ_β through the roots $\beta \in \Phi$. Since $\sigma_{-\beta} = \sigma_\beta$ we can restrict to positive roots β. It is important to notice that reflections and thus all elements of W are orthogonal transformations, i.e., that they preserve the inner product:

$$(w(x), w(y)) = (x, y)$$

The *Weyl chambers* are defined to be the components of the complement in E of the union of all hyperplanes perpendicular to the roots. The elements of W are orthogonal and permute the roots. Therefore, W permutes the Weyl chambers.

10.1. Simple reflections.

Choose a base Δ for Φ. Then the *simple reflections* are defined to be σ_{α_i} where α_i are the simple roots (elements of Δ).

Lemma 10.1.1. *If $\alpha \in \Delta$, the simple reflection σ_α sends α to $-\alpha$, $-\alpha$ to α and permutes all of the other positive roots.*

Proof. Suppose that β is a positive root not equal to α. Then β is not equal to a scalar multiple of α. So, in the expansion of β as a positive linear combination of simple roots: $\beta = \sum k_i\alpha_i$ where, say, $\alpha = \alpha_1$, one of the other coefficients, say $k_2 > 0$. Then

$$\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha\rangle \alpha = (k_1 - \langle \beta, \alpha\rangle)\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$$

is a positive root since $k_2 > 0$. □

Proposition 10.1.2. *For every $\beta \in \Phi$ there is a simple root α and a sequence of simple reflections whose composition carries α to β.*

Proof. We may assume that β is a positive root: If we know that a sequence of simple reflections carries α to β then, if we first do σ_α (sending α to $-\alpha$) and then do these same simple reflections, then we will carry α to $-\beta$.

Define the *height* of a positive root β as the sum of the coefficients k_i in the expansion $\beta = \sum k_i\alpha_i$. If the height of β is 1 then β is simple and there is nothing to prove. Then we proceed by induction.

Suppose that β is a positive root with height ≥ 2. Then at least two of the coefficients, say k_1, k_2 are positive.

Claim There exists a simple root α so that $(\alpha_i, \alpha) > 0$ for some α_i so that $k_i > 0$.

Pf: If not then $(\beta, \alpha) \leq 0$ for all simple roots α and therefore, $(\beta, \beta) = \sum k_i(\beta, \alpha_i) \leq 0$ which is not possible since $\beta \neq 0$.

Choose α as in the Claim. Then

$$\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha\rangle \alpha$$

is a positive root by the lemma and it has smaller height than β since $\langle \beta, \alpha\rangle$ is positive. Therefore, by induction on height, we can do a few more simple reflections to send $\sigma_\alpha(\beta)$ to a simple root. □

Lemma 10.1.3. *If $w \in W$ and β is any root then $w\sigma_\beta w^{-1} = \sigma_{w(\beta)}$.* □

Corollary 10.1.4. *The Weyl group W is generated by simple reflections.*

Proof. For any $\beta \in \Phi$ there is a simple root α and $w \in W$, a product of simple reflections, so that $w(\alpha) = \beta$. But then the generator σ_β of W is a product of simple reflections: $\sigma_\beta = w\sigma_a w^{-1}$, proving the corollary. □

Definition 10.1.5. A *reduced expression* for w is a factorization of w as a product of k simple reflections $w = \sigma_1\sigma_2 \ldots \sigma_k$ where k is minimal. This minimal number is called the *length* of w.

Lemma 10.1.6. *For any $w \in W$, the following two numbers are equal:*

(1) *$n(w)$ = the number of positive roots β which are sent to negative roots by w.*
(2) *$\ell(w)$ = the smallest number of simple reflections whose product is w.*

Proof. Every simple reflection sends exactly on positive root to a negative root and the other positive roots to positive roots. Therefore, a product of k simple reflections can send at most k positive roots to negative roots. Therefore, $n(w) \leq \ell(w)$.

To prove the lemma it suffices to show that, if w is a product of $k+1$ simple reflections $w = \sigma_0\sigma_1 \cdots \sigma_k$ and $n(w) < k+1$ then the expression is not reduced. So, suppose the expression is reduced. Then, the subword $w_1 = \sigma_1\sigma_2 \cdots \sigma_k$ is also reduced with length k. By induction on k we have $\ell(w_1) = n(w_1) = k$. So, w_1 sends k positive roots to negative roots. Since $n(w) < k+1$, the last simple reflection $\sigma_0 = \sigma_\alpha$ sends one of these negative roots back to a positive root. This root must be $-\alpha$ which is equal to some $w(\beta), \beta \in \Phi_+$. But this means the simple reflections σ_i keep β positive for a while, switches the sign then keeps it negative. So

$$w_1 = \sigma_1 \cdots \sigma_i \cdots \sigma_k = g\sigma_i h$$

where $h(\beta) = \alpha_j \in \Delta$ and $\sigma_i = \sigma_{\alpha_j}$ and $g(\alpha_j) = \alpha$ so that:

$$w_1(\beta) = g\sigma_{\alpha_j} h(\beta) = g\sigma_{\alpha_j}(\alpha_j) = g(-\alpha_j) = -\alpha$$

But this last step $g(\alpha_j) = \alpha$ implies that $g\sigma_i g^{-1} = \sigma_{g(\alpha_j)} = \sigma_\alpha$. Therefore,

$$w = \sigma_\alpha w_1 = \sigma_\alpha g\sigma_{\alpha_j} h = gh$$

showing that the original expression for w was not reduced. □

Proposition 10.1.7. *The only element of W which sends all positive roots to positive roots is the identity.* □

Corollary 10.1.8. *The only element $w \in W$ which sends Δ to Δ is $w = 1$.* □

10.2. **Weyl chambers.** We will show that there is a 1-1 correspondence between bases Δ, Weyl chambers C and elements w of the Weyl group. This is equivalent to saying that the action of W on both of these sets is simply transitive (transitive and effective).

Lemma 10.2.1. *The Weyl group acts transitively on the set of Weyl chambers.*

Proof. Whereas hyperplanes cut E up into disjoint regions, the intersection of any two hyperplanes has codimension two and the complement of any finite number of codimension 2 subspaces is connected.

If we take any two Weyl chambers C_0, C_1 then there is a path connecting C_0 to C_1 which passes through a finite number of the hyperplanes separating the chambers but does not pass through intersections of hyperplanes. If this finite number is 0 then $C_0 = C_1$ and there is nothing to prove. If this number is positive, then take the first hyperplane, say $\beta^\perp$ which crosses the path from C_0 to C_1. Then σ_β takes C_0 to the chamber C_2 on the other side of the hyperplane $\beta^\perp$. This is closer to C_1 by construction and therefore $w(C_2) = C_1$ for some $w \in W$ by induction. Then $C_1 = w\sigma_\beta(C_0)$. □

Lemma 10.2.2. *Let C be the set of all $x \in E$ with the property that $(x, \beta) > 0$ for all positive roots β. Then C is a Weyl chamber.*

We call C the *fundamental chamber.*

Proof. Clearly C is convex and therefore connected. Also C is disjoint from all hyperplanes $\beta^\perp$. Therefore, C is contained in some Weyl chamber C_0. Suppose that $y \in C_0$ then, since C_0 is connected, there is a path $\gamma(t)$ in C_0 connecting $x \in C$ to y. This path does not cross any of the hyperplanes. Therefore, by the intermediate value theorem, the sign of $(\gamma(t), \beta)$ remains unchanged. Since it starts as positive, it remains positive. So, $y \in C$ proving that $C = C_0$ is a Weyl chamber. □

The above proof has a gap: it assumes that C is nonempty. To fix this I added the following in class.

Lemma 10.2.3. *Let C' be the set of all $x \in E$ so that $(x, \alpha) > 0$ for all positive roots α. Then C' is nonempty and $C' = C$.*

Proof. It is easy to see that $C \subseteq C'$ because the condition (x, β) for all positive β implies that $(x, \alpha) > 0$ for simple α. Conversely, if $(x, \alpha) > 0$ for simple α then $(x, \beta) = \sum k_i(x, \alpha_i) > 0$ for all positive β.

To see that C' is nonempty, note that the mapping $\psi : E \to E^*$ given by $\alpha \mapsto (-, \alpha)$ is an isomorphism. Therefore it sends basis to basis. So, $\{(-, \alpha_i)\}$ is a basis for E^*. Equivalently,

$$x \mapsto [(x, \alpha_1), (x, \alpha_2), \cdots, (x, \alpha_n)] \in \mathbb{R}^n$$

gives a linear isomorphism $E \cong \mathbb{R}^n$. So, there is some $x \in E$ which goes to $[1, 1, \cdots, 1] \in \mathbb{R}^n$. Then $x \in C'$ showing that C' is nonempty. □

Theorem 10.2.4. *There is a one-to-one correspondence between bases Δ, Weyl chambers C and elements of the Weyl group W.*

Proof. We know that W acts transitively on Weyl chambers and effectively on bases. Therefore, it suffices to show that there is a 1-1 corresondence between bases and chambers which respects the action of W.

For each Δ we have the set Φ_+ of all roots which are positive linear combinations of elements of Δ. The lemma associates a corresponding fundamental chamber C. Conversely, given any chamber C, the positive roots are those with $(x, \beta) > 0$ for all $x \in C$

and the simple roots are the indecomposable positive roots. It is easy to see that this correspondence is W-equivariant, i.e., $w\Delta$ corresponds to wC. $\square$

Exercise 10.2.5. (1) Show that there is a unique element of the Weyl group $w_0 \in W$ of maximal length. (w_0 corresponds to $-C_0$ where C_0 is the fundamental chamber.)

(2) Suppose that $\beta = \sum k_i \alpha_i$ is a positive root with maximal height. ($\sum k_i$ is maximal).
   (a) Show that β is in the closure of C_0.
   (b) Show that any α_j with $k_j = 0$ is perpendicular to the α_i with $k_i > 0$.

(3) Suppose that Φ is an irreducible root system. Then show that there exists a unique root of maximal height. (If there are two then their difference would be a root.)

11. Classification of semisimple Lie algebras

I will explain how the Cartan matrix and Dynkin diagrams describe root systems. Then I will go through the classification using classical examples of Lie algebras following Erdmann and Wildon [2]. I will skip the combinatorial proof. (See Humphreys [5]. This part of the book requires no knowledge of Lie algebras!)

11.1. **Cartan matrix and Dynkin diagram.**

Definition 11.1.1. Suppose that Φ is a root system with base Δ. Let $\alpha_1, \cdots, \alpha_n$ be the list of simple roots. Then the *Cartan matrix* is defined to be the $n \times n$ matrix C with entries $\langle \alpha_i, \alpha_j \rangle$. We recall that $\langle \alpha_i, \alpha_j \rangle = \alpha_i(h_{\alpha_j})$ where h_α is the unique element of $[L_\alpha, L_{-\alpha}]$ so that $\alpha(h_\alpha) = 2$.

Example 11.1.2. If we let α_1 be the short root, the Cartan matrices for the root systems of rank 2 are:

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & -1 \\ -2 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$$

$$A_1 \times A_1 \qquad A_2 \qquad B_2 = C_2 \qquad G_2$$

The diagonal entries are always equal to 2. The off diagonal entries are negative or zero.

Proposition 11.1.3. *The Cartan matrix of a semisimple Lie algebra is nonsingular. In fact, its determinant is positive and all of its diagonal minors are positive.*

Proof. By multiplying the j column of C by the positive number $(\alpha_j, \alpha_j)/2$, it becomes the matrix (α_i, α_j) which has positive determinant since it is the matrix of a positive definite symmetric form on $E \cong \mathbb{R}^n$. Any diagonal submatrix of C is proportional to the matrix of the same form on a vector subspace. This is positive definite since it is the restriction of a positive definite form. □

Definition 11.1.4. The *Dynkin diagram* of a root system of rank n is defined to be a graph with n vertices labelled with the simple roots α_i and with edges given as follows.

(1) No edge connecting roots α_i, α_j if they are orthogonal (equivalently, if $c_{ij} = 0$.)

$$\circ \qquad \circ$$

(2) A single edge connecting α_i, α_j if α_i, α_j are roots of the same length which are not orthogonal (equivalently, $c_{ij} = c_{ji} = -1$.)

$$\circ \!-\!\!-\!\!-\! \circ$$

(3) A double edge pointing from α_i to α_j if α_i, α_j are not perpendicular and $||\alpha_i||^2 = 2||\alpha_j||^2$

$$\circ_i \Longrightarrow \circ_j$$

(4) A triple edge pointing from α_i to α_j if α_i, α_j are not perpendicular and $||\alpha_i||^2 = 3||\alpha_j||^2$

$$\circ_i \Rrightarrow \circ_j$$

The *Coxeter graph* is the same graph without the orientation on the multiple edges.

Corollary 11.1.5. *The Coxeter graph of a root system contains no cycle of single edges.*

Proof. If the diagram contains such a cycle then the vector v which is sum of the simple roots corresponding to the vertices of the cycle have the property that $v^tCv = 0$ since the only positive number in each column is a 2 and the each simple edge incident to α_i puts a -1 in the ith column. But v^tCv is proportional to (v, v) since all simple roots in the cycle have the same length. This contradicts the fact that the inner product is positive definite. □

In a similar way, we can list more "forbidden subgraphs" of the Coxeter graph. The classification is given by listing all possible graphs which contain no forbidden subgraphs. This gives a list of all possible Dynkin diagrams. We need to construct semisimple Lie algebras for each diagram and show that the diagram determines the Lie algebra uniquely.

11.2. **Orthogonal algebra $\mathfrak{so}(2n, F)$.** Recall that any bilinear form $f : V \times V \to F$ defines a subalgebra of $\mathfrak{gl}(V)$ consisting of all x so that

$$f(x(v), w) + f(v, x(w)) = 0$$

We will take $V = F^n$ and take only $x \in \mathfrak{sl}(n, F)$ satisfying the above. When $V = F^n$ there is a unique $n \times n$ matrix S so that

$$f(v, w) = v^tSw$$

So, the condition $f(x(v), w) + f(v, x(w)) = 0$ becomes:

$$x^tS + Sx = 0$$

For the even dimensional *orthogonal algebra* $\mathfrak{so}(2n, F)$ we take the nondegenerate symmetric bilinear form given by the matrix

$$S = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}$$

Thus $\mathfrak{so}(2n, F)$ is the set of all $2n \times 2n$ matrices

$$x = \begin{bmatrix} m & p \\ q & -m^t \end{bmatrix}$$

where $p = -p^t$ and $q = -q^t$. The Cartan subalgebra H is the set of diagonal matrices and H^* is generated by the n functions ϵ_i where $\epsilon_i(h) = h_i$ is the ith diagonal entry of h. The off-diagonal entries of the matrix m are coefficients of the matrix x_{ij} and

$$[h, x_{ij}] = (h_i - h_j)x_{ij} = (\epsilon_i - \epsilon_j)(h)x_{ij}$$

So, x_{ij} is in L_α where $\alpha = \epsilon_i - \epsilon_j$ and L_α is 1-dimensional. Also, $x_{ji} \in L_{-\alpha}$ and $[x_{ij}, x_{ji}] = h_\alpha$ has $\alpha(h_\alpha) = 2$.

The ij entry of p is $p_{ij} = -p_{ji}$. Let $y_{ij}, 1 \le i < j \le n$ be the basis element corresponding to this entry. Then

$$[h, y_{ij}] = (h_i + h_j)y_{ij} = (\epsilon_i + \epsilon_j)(h)y_{ij}$$

Similarly, the ij entry of q is $q_{ij} = -q_{ji}$. If z_{ij} is the negative of the corresponding basis element (with $q_{ij} = -1$) then

$$[h, z_{ij}] = (-h_i - h_j)z_{ij} = (-\epsilon_i - \epsilon_j)(h)z_{ij}$$

So, $y_{ij} \in L_\beta$ and $z_{ij} \in L_{-\beta}$ where $\beta = \epsilon_i + \epsilon_j$. Also, $[y_{ij}, z_{ij}] = h_\beta$ where $\beta(h_\beta) = 2$.

We see that the set of roots seems to be of type D_n (but we did not yet show that the inner product is the same). A base for this is the set of roots:

$$\Delta = \{\epsilon_1 - \epsilon_2, \epsilon_2 - \epsilon_3, \cdots, \epsilon_{n-1} - \epsilon_n, \epsilon_{n-1} + \epsilon_n\}$$

(Proof that this is a base) Adding consecutive simple roots from the first $n-1$ gives all roots of the form $\epsilon_i - \epsilon_j$. To get $\epsilon_i + \epsilon_j$ take

$$\epsilon_i - \epsilon_j + 2(\epsilon_j - \epsilon_{n-1}) + (\epsilon_{n-1} - \epsilon_n) + (\epsilon_{n-1} + \epsilon_n)$$

All other roots are negatives of these. So, Δ is a base.

We also have a root space decomposition

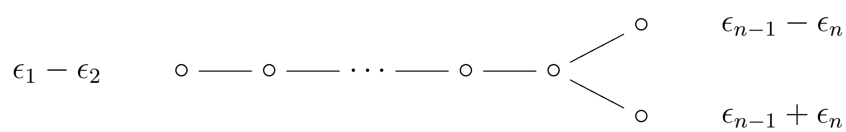
$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$$

And, each L_α is one-dimensional and the intersection of the kernels of $\alpha : H \to F$ is zero. These facts together imply the following.

Theorem 11.2.1. *$\mathfrak{so}(2n, F)$ is semisimple with root system of type D_n.*

Proof. To show that $\mathfrak{so}(2n, F)$ is semisimple we need to show that there is no nontrivial abelian ideal. So, suppose there is an abelian ideal J. Then H acts on J and, therefore, J decomposes into weight spaces for this action. But the weight space decomposition of J must be compatible with the decomposition of L. So J is a sum of $J_0 = J \cap H$ with $J_\beta = L_\beta$ for certain roots β. If any $J_\beta = L_\beta$ is nonzero, then J also contains $h_\beta \in [L_{-\beta} L_\beta] \subseteq [LJ] \subseteq J$. But h_β and L_β do not commute, contradicting the assumption that J is abelian. So, $J = J_0 \subseteq H$. If $h \in J \neq 0$ there is some root α which is nonzero on h. Then $[h, x_\alpha] = \alpha(h)x_\alpha$ is a nonzero element of L_α which is a contradiction. So, $\mathfrak{so}(2n, F)$ is semisimple.

The Cartan matrix has entries $\alpha(h_\beta)$. But the calculation shows that each h_β is the dual of β with respect to the basis ϵ_i. So, this is equal to the dot product of α, β. Apply this to the base to get the usual Dynkin diagram for D_n since we have two copies of A_{n-1} differing only in the last simple roots which are perpendicular.



□

11.3. **Symplectic algebra $\mathfrak{sp}(2n, F)$.** For the next example we take f to be nondegenerate and skew symmetric:

$$f(v, v) = 0$$

This is called a *symplectic form* on V.

Proposition 11.3.1. *Given a symplectic form f on V, there is a basis $v_1, \cdots, v_n, w_1, \cdots, w_n$ for V so that $f(v_i, v_j) = 0 = f(w_i, w_j)$ for all i, j and $f(v_i, w_j) = \delta_{ij}$. (This is called a* symplectic basis *for V.)*

Proof. This is by induction on the dimension of V. If the dimension is 0 there is nothing to prove. If $\dim V > 0$ then choose any nonzero vector $v_1 \in V$. SInce f is nondegenerate, there is a $w_1 \in V$ so that $f(v_1, w_1) = 1$. Since $f(v_1, v_1) = 0$, v_1, w_1 are linearly independent and their span W is 2-dimensional. Let

$$W^\perp = \{x \in V \mid f(x, v_1) = f(x, w_1) = 0\}$$

Then $W \cap W^\perp = 0$ and $\dim W^\perp = \dim V - 2$. So, $V = W \oplus W^\perp$. Since $W^\perp$ is perpendicular to W, the restriction of f to $W^\perp$ is symplectic. So, $W^\perp$ has a basis $v_2, \cdots, v_n, w_2, \cdots, w_n$ with the desired properties. Together with v_1 and w_1 we get a symplectic basis for W. □

Definition 11.3.2. The *symplectic Lie algebra* $\mathfrak{sp}(2n, F)$ is defined to be the algebra of all $x \in \mathfrak{sl}(2n, F)$ so that

$$x^t \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} = - \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} x$$

In other words, $\mathfrak{sp}(2n, F)$ is the set of all $2n \times 2n$ matrices

$$x = \begin{bmatrix} m & p \\ q & -m^t \end{bmatrix}$$

where $p = p^t$ and $q = q^t$. Again we have H the diagonal matrices. H^* is n-dimensional with basis ϵ_i, off-diagonal entries of m give x_{ij} with

$$[h, x_{ij}] = \alpha(h) x_{ij}$$

where $\alpha = \epsilon_i - \epsilon_j$ and $[x_{ij}, x_{ji}] = h_\alpha$. The entries of p give y_{ij} with

$$[h, y_{ij}] = \beta(h) y_{ij}$$

where $\beta = \epsilon_i + \epsilon_j$ where now we include the case $i = j$. The entries of q give basis elements z_{ij} with

$$[h, z_{ij}] = -\beta(h) z_{ij}$$

and

$$[y_{ij}, z_{ij}] = h_\beta$$

Thus we have a root system of type C_n. Note that, when $i = j$, the element h_β is not exactly the dual of β since $\beta = 2\epsilon_i$ whereas h_b is dual to ϵ_i. So, the Cartan matrix $\alpha_i(h_{\alpha_j})$ is not symmetric in this case.

Theorem 11.3.3. *$\mathfrak{sp}(2n, F)$ is semisimple with root system C_n.*

Proof. This follows from the root space decomposition just as in the proof of Theorem 11.2.1. The base for this root system is

$$\Delta = \{\epsilon_1 - \epsilon_2, \cdots, \epsilon_{n-1} - \epsilon_n, 2\epsilon_n\}$$

with Dykin diagram

$$\epsilon_1 - \epsilon_2 \qquad \circ - \circ - \circ - \cdots - \circ - \circ \Leftarrow \circ \qquad 2\epsilon_n$$

(The double arrow points to the shorter root.) □

11.4. **Orthogonal algebra $\mathfrak{so}(2n+1,F)$.** The odd dimensional *orthogonal algebra* $\mathfrak{so}(2n+1,F)$ is the Lie algebra given by the nondegenerate symmetric bilinear form corresponding to the matrix

$$S=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & I_n\\ 0 & I_n & 0\end{bmatrix}$$

Thus $\mathfrak{so}(2n+1,F)$ is the set of all $2n+1\times 2n+1$ matrices x so that $x^tS+Sx=0$. In other words

$$x=\begin{bmatrix}0 & c^t & -b^t\\ b & m & p\\ -c & q & -m^t\end{bmatrix}$$

where $p=-p^t$ and $q=-q^t$. We number the rows and columns $0,1,2,\cdots,2n$.

The Cartan subalgebra H is the set of diagonal matrices and H^* is generated by the n functions ϵ_i where $\epsilon_i(h)=h_i$ is the ith diagonal entry of h. As in the case of $\mathfrak{so}(2n,F)$, the off-diagonal entries of the matrix m are coefficients of the matrix x_{ij} and

$$[h,x_{ij}]=(h_i-h_j)x_{ij}=\alpha(h)x_{ij}$$

where $\alpha=\epsilon_i-\epsilon_j$. So, x_{ij} is in L_α and $x_{ji}\in L_{-\alpha}$ and $[x_{ij},x_{ji}]=h_\alpha$ has $\alpha(h_\alpha)=2$.

We also have basic matrices $y_{ij}, 1\le i<j\le n$ for p and z_{ij} for q so that

$$[h,y_{ij}]=(h_i+h_j)y_{ij}=\beta(h)y_{ij}$$

where $\beta=\epsilon_i+\epsilon_j$ and

$$[h,z_{ij}]=-\beta(h)z_{ij}$$

Also, $[y_{ij},z_{ij}]=h_\beta$ where $\beta(h_\beta)=2$.

What is new is the basic matrices b_i,c_i for b,c with

$$[h,b_i]=h_ib_i=\gamma(h)b_i$$

where $\gamma=\epsilon_i$ and

$$[h,c_i]=-h_ic_i=-\gamma(h)c_i$$

But $h_\gamma=2[b_i,c_i]$ since

$$\gamma([b_i,c_i])=1$$

The root system consists of $\pm\epsilon_i\pm\epsilon_j$ and $\pm\epsilon_i$. The basic roots are given by

$$\epsilon_1-\epsilon_2,\cdots,\epsilon_{n-1}-\epsilon_n,\epsilon_n$$

To see that this forms a base for the root system, note that adding consecutive elements gives any $\epsilon_i-\epsilon_j$ and adding the terms from $\epsilon_j-\epsilon_{j+1}$ to ϵ_n gives ϵ_j. And $\epsilon_i-\epsilon_j+2\epsilon_j=\epsilon_i+\epsilon_j$. So, this is a base. Since ϵ_n is a short root, the Dynkin diagram is:

$\epsilon_1-\epsilon_2$ ○ —— ○ —— ○ —— ⋯ —— ○ —— ○ ⟹ ○ ϵ_n

and this is a root system of type B_n.

Exercise 11.4.1. Show that $\mathfrak{sl}(n+1,F)$ is a semisimple Lie algebra with root system A_n:

$\epsilon_1-\epsilon_2$ ○ —— ○ —— ○ —— ⋯ —— ○ —— ○ $\epsilon_n-\epsilon_{n+1}$

12. Exceptional Lie algebras and automorphisms

We have constructed four infinite families of semisimple algebras:

(1) $\mathfrak{sl}(n+1, F)$ has type A_n
(2) $\mathfrak{so}(2n, F)$ has type D_n
(3) $\mathfrak{so}(2n+1, F)$ has type B_n
(4) $sp(2n, F)$ has type C_n.

Theorem 12.0.2. *If L is a simple Lie algebra over an algebraically closed field of characteristic zero then L either belongs to one of the four infinite families listed above or it is isomorphic to one of the five exceptional Lie algebras of type E_6, E_7, E_8, F_4, G_2.*

We will construct a Lie algebra of type G_2 later. For now we will just construct the root systems of type E_8 and F_4.

12.1. E_6, E_7, E_8. The root systems E_6, E_7 can be constructed from the root system (E, Φ) of type E_8 as follows. Let Δ' be the subset of the base Δ corresponding to the subdiagram of E_8 corresponding to E_6 or E_7. Let E' be the span of Δ' in E. Let $\Phi' = \Phi \cap E'$. Then (E/Φ') will be a root system of type E_6 or E_7. Therefore, it suffices to construct a root system of type E_8.

Let $E = \mathbb{R}^8$. The root system is:

$$\Phi = \{\pm\epsilon_i \pm \epsilon_j\} \cup \left\{ \frac{1}{2} \sum_{i=1}^{8} \pm\epsilon_i \text{ with even number of } - \text{ signs} \right\}$$

Note that the first part $\{\pm\epsilon_i \pm \epsilon_j\}$ is the root system of type D_8. Also, all of these roots have the same length. A base for this root system is

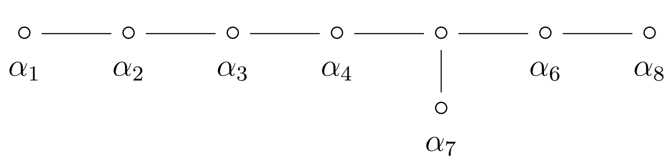
$$\epsilon_2 - \epsilon_3, \cdots, \epsilon_7 - \epsilon_8, \epsilon_7 + \epsilon_8, \frac{1}{2}\left(\epsilon_1 + \epsilon_8 - \sum_{i=2}^{7} \epsilon_i\right)$$

We show that this is a base:

Since the first 7 roots form a base for D_7 we know that any root of the form $\epsilon_i \pm \epsilon_j$ for $2 \le i < j \le n$ is a positive linear combination of the first 7 roots. By adding twice the last simple root α_8, we can get $\epsilon_1 - \epsilon_2$:

$$\epsilon_1 - \epsilon_2 = 2\alpha_8 + (\epsilon_3 + \epsilon_4) + (\epsilon_5 + \epsilon_6) + (\epsilon_7 - \epsilon_8)$$

with a single α_8 plus roots of the form $\epsilon_i \pm \epsilon_j$ we get any sum $\frac{1}{2}\sum \pm\epsilon_i$ with the sign of ϵ_1 being positive (with an even number of - signs). Note that α_8 is perpendicular to every other simple root except for $\epsilon_7 - \epsilon_8 = \alpha_6$. So, the Dynkin diagram is:



12.2. F_4. This root system is given by $E = \mathbb{R}^4$ and

$$\Phi = \{\pm\epsilon_i\} \cup \{\pm\epsilon_i \pm \epsilon_j\} \cup \{\tfrac{1}{2}(\pm\epsilon_1 \pm \epsilon_2 \pm \epsilon_3 \pm \epsilon_4)\}$$

where these roots have lengths $1, \sqrt{2}, 1$ resp. The base is given by:

$$\epsilon_1 - \epsilon_2, \epsilon_2 - \epsilon_3, \epsilon_3, \tfrac{1}{2}(-\epsilon_1 - \epsilon_2 - \epsilon_3 + \epsilon_4)$$

The last two roots are smaller so the Dynkin diagram is:

$$\begin{array}{ccccccc} \circ & \text{——} & \circ & \Rightarrow & \circ & \text{——} & \circ \\ \alpha_1 & & \alpha_2 & & \alpha_3 & & \alpha_4 \end{array}$$

12.3. **Automorphisms of Φ.**

Definition 12.3.1.

$$\mathrm{Aut}(\Phi) := \{f \in O(E) \mid f(\Phi) = \Phi\}$$

The *orthogonal group* $O(E)$ is the group of all linear isometries of E, i.e., f preserves the inner product, equivalently, f preserves lengths and angles.

Lemma 12.3.2. *The Weyl group W is a normal subgroup of* $\mathrm{Aut}(\Phi)$*.*

Proof. Clearly, W is a subgroup of $\mathrm{Aut}(\Phi)$ since it sends roots to roots and is generated by reflections. It remains to show that W is a normal subgroup.

W is generated by the reflections σ_β where $\beta \in \Phi$. For any $\tau \in \mathrm{Aut}(\Phi)$ we have:

$$\tau\sigma_\beta\tau^{-1} = \sigma_{\tau(\beta)} \in W$$

So, W is normal in $\mathrm{Aut}(\Phi)$. The equation follows from the fact that $\tau\sigma_\beta\tau^{-1}$ sends $\tau(\beta)$ to $-\tau(\beta)$ and also fixes every $x \perp \tau(\beta)$. □

Theorem 12.3.3. $\mathrm{Aut}(\Phi) \cong W \rtimes \Gamma$ *where Γ is the group of all $\tau \in \mathrm{Aut}(\Phi)$ so that $\tau(\Delta) = \Delta$.*

Proof. To prove that $\mathrm{Aut}(\Phi)$ is this semi-direct product it suffices to show that $W \cap \Gamma = \{1\}$ and that $W\Gamma = \mathrm{Aut}(\Phi)$.

To prove the first statement we recall that W acts simply transitively on the set of all bases. So, the only element of W which fixes Δ is the identity. To prove the second statement, take any $\tau \in \mathrm{Aut}(\Phi)$. Then $\tau(\Delta) = \Delta'$ is another base for Φ. So, there is some $w \in W$ so that $w(\Delta) = \Delta'$. Then $w^{-1}\tau \in \Gamma$ and $\tau = w(w^{-1}\tau) \in W\Gamma$. So, $\mathrm{Aut}(\Phi) = W \rtimes \Gamma$. □

Theorem 12.3.4. Γ *is the group of automorphisms of the Dynkin diagram.*

Before proving this, I pointed out that the automorphisms of the diagram are very easy to compute.

(1) For $A_n, n \geq 2$, $\Gamma = \mathbb{Z}/2$.
(2) For $D_n, n \geq 5$, $\Gamma = \mathbb{Z}/2$.
(3) For D_4, $\Gamma = S_3$, the symmetric group on 3 letters.
(4) For E_6, $\Gamma = \mathbb{Z}/2$.
(5) For all other Dynkin diagrams, Γ is trivial.

Proof. Take any $\tau \in \Gamma$. Then $\tau(\Delta) = \Delta$ means that τ permutes the elements of Δ and preserves lengths and angles. Therefore, τ gives an automorphism of the Dynkin diagram.

Conversely, suppose that τ is an automorphism of the Dynkin diagram. Then τ sends basic roots to basic roots. Since the basic roots form a basis for E, τ gives a linear isometry of E. The Dynkin diagram contains the information of which simple roots are long and which are short. So, τ sends long/short simple roots to long/short simple roots and also preserves the angles. So, τ is an isometry of E.

It remains to show that τ sends roots to roots. (I will finish this later.) □

14. Isomorphism Theorem

This section contain the important theorem that two simple Lie algebras with the same Dynkin diagram are isomorphic. The proof uses the existence of a unique maximal root (Exercise 10.2.5). It also uses another lemma which we skipped:

Lemma 14.0.5. *Suppose that β is a positive root which is not simple. Then there exists a simple root α so that $\beta-\alpha$ is a positive root.*

Proof. Let $\beta=\sum k_i\alpha_i$. Then $(\beta,\beta)=\sum k_i(\beta,\alpha_i)>0$ implies that $(\beta,\alpha_i)>0$ for some α_i. But this implies that $\beta-\alpha_i$ is a root (Lemma 9.4.5). And it must be a positive root since one of the other coefficients $k_j>0$. □

14.1. **Proof of the isomorphism theorem.**

Proposition 14.1.1. *A semisimple Lie algebra L is generated by $L_\alpha, L_{-\alpha}$ for all $\alpha\in\Delta$.*

Proof. Consider the subalgebra L^+ generated by L_α for all $\alpha\in\Delta$. We claim that L^+ contains all L_β for $\beta\in\Phi_+$. If not then take the smallest positive root $\beta=\sum k_i\alpha_i$ so that L_β is not contained in L^+. Then there is some simple root α so that $\gamma=\beta-\alpha$ is a positive root (which is smaller than β). But then $L_\gamma\subseteq L^+$ and $[L_\alpha,L_\gamma]=L_\beta\subseteq L^+$.

Similarly, the subalgebra L^- generated by all $L_{-\alpha}$ for simple α contains all L_β for $\beta\in\Phi_-$. So the subalgebra L' of L generated by all $L_\alpha, L_{-\alpha}$ contains L_β for all $\beta\in\Phi$. On the other hand $h_\alpha\in[L_\alpha,L_{-\alpha}]$. So, L' also contains all h_α. But the h_α generate H since the α are a basis for H^*. So, $L'=L$. □

Theorem 14.1.2. *Let L,L' be simple Lie algebras over F. Let H,H' be Cartan subalgebras for L,L'. Let Φ,Φ' be the corresponding root systems. Suppose that $\Phi\cong\Phi'$. Choose any isomorphism $\Phi\cong\Phi'$. Let Δ,Δ' be corresponding bases. Let $\pi:H\to H'$ be the corresponding isomorphism (whose dual $\pi^*:H'^*\to H^*$ sends Δ' to Δ). Choose any nonzero $x_\alpha\in L_\alpha, x'_\alpha\in L'_\alpha$. Then there is a unique isomorphism $\overline{\pi}:L\to L'$ sending x_α to x'_α so that $\overline{\pi}|H=\pi$.*

In short: if $\Phi\cong\Phi'$ then $L\cong L'$.

Proof. (Uniqueness) First note there are unique elements $y_\alpha\in L_{-\alpha}$ and $y'_\alpha\in L'_{-\alpha}$ so that $[x_\alpha,y_\alpha]=h_\alpha$ and $[x'_\alpha,y'_\alpha]=h'_\alpha$. Since $\pi:H\to H'$ sends h_α to h'_α, $\overline{\pi}$ must send y_α to y'_α. But the elements x_α,y_α generate L by the proposition above. So, $\overline{\pi}$ is uniquely determined.

Next we consider the algebra $L\oplus L'$. This is a semisimple Lie algebra with exactly two nonzero proper ideal: L,L'. Take the "diagonal" D which is the subalgebra of $L\oplus L'$ generated by the elements $\overline{x}_\alpha=(x_\alpha,x'_\alpha)$ and $\overline{y}_\alpha=(y_\alpha,y'_\alpha)$. Then the projection map $L\oplus L'\to L$ sends D onto L and similarly projection to the second factor sends D onto L'. We will show that both of these projection maps are isomorphisms giving $L\cong D\cong L'$.

Let D^- be the subalgebra of $L\oplus L'$ generated by the elements $\overline{y}_\alpha$ for $\alpha\in\Delta$. Let $\beta\in\Phi_+$ be the unique maximal root and choose nonzero elements $z\in L_\beta, z'\in L'_\beta$. Let M be the D^- submodule of $L\otimes L'$ generated by the element $\overline{z}=(z,z')\in L_\beta\oplus L'_\beta$. In other words, M is the span of $\overline{z}$ and all elements of the form

$$[\overline{y}_{\alpha_1}[\overline{y}_{\alpha_2}\cdots[\overline{y}_{\alpha_k}\overline{z}]\cdots]] \tag{14.1}$$

Claim 1 $M \cap (L_\beta \oplus L'_\beta)$ is one-dimensional.

Pf: The generators $\overline{y}_\alpha$ of D^- sends $L_\beta \oplus L'_\beta$ into $L_{\beta-\alpha} \oplus L'_{\beta-\alpha}$ and $L_\gamma \oplus L'_\gamma$ into $L_{\gamma-\alpha} \oplus L'_{\gamma-\alpha}$. So, we never come back to the maximal root β.

Claim 2 $[D, M] \subseteq M$.

Pf: We show by induction on k that ad$\overline{x}_\alpha$ sends the expression 14.1 above into M.

The generators $\overline{x}_\alpha$ of D act trivially on $(z, z') \in L_\beta \oplus L'_\beta$. So, this statement is true for $k = 0$. Since the difference of two simple roots is never a root, $\overline{x}_\alpha$ commutes with $\overline{y}_{\alpha'}$ if $\alpha' \neq \alpha$. So, if $\alpha_1 \neq \alpha$ then

$$[\overline{x}_\alpha[\overline{y}_{\alpha_1}[\overline{y}_{\alpha_2} \cdots [\overline{y}_{\alpha_k}\overline{z}] \cdots]]] = [\overline{y}_{\alpha_1}[\overline{x}_\alpha[\overline{y}_{\alpha_2} \cdots [\overline{y}_{\alpha_k}\overline{z}] \cdots]]]$$

lies in $[\overline{y}_{\alpha_1} M] \subseteq M$ by induction on k. Therefore, we may assume that $\alpha_1 = \alpha$. Then

$$[\overline{x}_\alpha[\overline{y}_\alpha[\overline{y}_{\alpha_2} \cdots [\overline{y}_{\alpha_k}\overline{z}] \cdots]]] = [\overline{y}_\alpha[\overline{x}_\alpha[\overline{y}_{\alpha_2} \cdots [\overline{y}_{\alpha_k}\overline{z}] \cdots]]] + [\overline{h}_\alpha[\overline{y}_{\alpha_2} \cdots [\overline{y}_{\alpha_k}\overline{z}] \cdots]]]$$

where $\overline{h}_\alpha = (h_\alpha, h'_\alpha) = [\overline{x}_\alpha, \overline{y}_\alpha]$. The first summand lies in M by induction. The second summand also lies in M since ad$\,\overline{h}_\alpha$ acts by multiplication by the same scalar in both coordinates of $L_\gamma \oplus L\gamma'$.

Claim 3 $D \neq L \oplus L'$.

Pf: Otherwise, M is an ideal by Claim 2 and proper by Claim 1. This is a contradiction since the only nontrivial proper ideals are $L \oplus 0$ and $0 \oplus L'$.

Claim 4 $D \cap L = 0 = D \cap L'$.

Pf: Suppose that $D \cap L$ is nonzero. Then it contains some $(w, 0)$ where $w \in L$ is nonzero. But $\overline{x}_\alpha \in D$ acts on $(w, 0)$ by $[\overline{x}_\alpha, (w, 0)] = ([x_\alpha w], 0)$ and similarly for $\overline{y}_\alpha$. By the Proposition, x_α, y_α generate L. So, $[D, (w, 0)] = ([L, w], 0) = (L, 0)$ which would imply that $D = L \oplus L'$ contradicting Claim 3. Similarly $D \cap L' = 0$.

This implies that the projection maps $D \to L, D \to L'$ are isomorphisms as claimed. These isomorphisms send $\overline{x}_\alpha$ to x_α, x'_α and similarly for $\overline{y}_\alpha$. But then the composition $L \to D \to L'$ sends x_α to x'_α as claimed. □

14.2. **Automorphisms of L.** The isomorphism theorem can also be used when $L = L'$. In that case it says that any automorphism of Φ extends to an automorphism of the Lie algebra L.

Corollary 14.2.1. *We have an epimorphism of groups:* $\mathrm{Aut}(L) \twoheadrightarrow \mathrm{Aut}(\Phi)$.

There is one automorphism in particular that we can write down:

Corollary 14.2.2. *If L is a simple Lie algebra and $x_\alpha \in L_\alpha, y_\alpha \in L_{-\alpha}$ so that $[x_\alpha, y_\alpha] = h_\alpha$ then there is an automorphism σ of L so that $\sigma(x_\alpha) = -y_\alpha$, $\sigma(y_\alpha) = -x_\alpha$ and $\sigma(h_\alpha) = -h_\alpha$. (Consequently, $\sigma(h) = -h$ for all $h \in H$ and $\sigma^2 = id_L$.)*

Exercise 14.2.3.
(1) Show that $\mathrm{Aut}\, L$ contains a subgroup isomorphic to Γ.
(2) Prove that Φ has a unique maximal root. (See Exercise 10.2.5.)
(3) Complete the proof of Theorem 12.3.4.

15. Cartan subalgebras

The next project is to show that the Cartan subalgebra H of a semisimple Lie algebra L is unique up to an automorphism of L. However, it turns out to be easier to generalize the notion of Cartan subalgebra to a not necessarily semisimple Lie algebra and then show that this more general notion is unique up to an inner automorphism of L.

There are two equivalent definitions for a Cartan subalgebra H of a general Lie algebra L:

(1) H is a self-normalizing nilpotent subalgebra, i.e., a nilpotent subalgebra of L so that
$$H = N_L(H) := \{x \in L \mid [xH] \subseteq H\}$$
For L semisimple, this is equivalent to the theorem that $H = L_0$.

(2) H is a minimal Engel subalgebra (defined below).

Humphreys points out that the ground field F can be of arbitrary characteristic. But, when F is finite, we need to assume that F has more than $2 \dim L$ elements.

15.1. **Engel subalgebras.**

For any endomorphism x of $V \cong F^n$ let
$$V_0(x) := \ker(x^n) = \ker(x^m) \quad \forall m \geq n.$$
Note that x acts nilpotently on $V_0(x)$ and x acts as an automorphism on $V/V_0(x)$.

Lemma 15.1.1. *Suppose that $z, x \in \mathfrak{gl}(V)$ and $V \cong F^n$.*

(1) *The set of all $c \in F$ so that $z + cx$ is nilpotent is either finite with $\leq n$ elements or all of F.*

(2) *The set of all $c \in F$ so that $z + cx$ is not an automorphism of V is either finite with $\leq n$ elements or all of F.*

Proof. The characteristic polynomial of $z + cx$,
$$p(T) = \det(z + cx - T)$$
is a polynomial in T of degree n whose coefficients are polynomials in the entries of the matrix for $z+cx$. Considering z, x to be constant, these coefficients are polynomials in c:
$$p(T) = (-1)^n T^n + g_1(c)T^{n-1} + \cdots + g_n(c)$$
where $\deg g_i(c) \leq i \leq n$.

(2) $z + cx$ is a not an automorphism iff its determinant is zero: $p(0) = g_n(c) = 0$. But g_n has at most n roots unless it is the zero polynomial in which case $g_n(c) = 0$ for all $c \in F$.

(1) $z + cx$ is nilpotent iff $p(T) = (-1)^n T^n$. This happens iff $g_i(c) = 0$ for all i. This is an intersection of sets each of which is either finite with $\leq n$ elements or all of F. □

Exercise 15.1.2. Show that $L_0(\operatorname{ad} x)$ is a subalgebra of L and $x \in L_0(\operatorname{ad} x)$. Show that a subalgebra K of L is nilpotent iff $K \subseteq L_0(\operatorname{ad} x)$ for all $x \in K$.

Definition 15.1.3. An *Engel subalgebra* of L is any subalgebra of the form $E = L_0(\operatorname{ad} x)$ for some $x \in L$. We call x an *annihilator* for E.

Lemma 15.1.4. *Let K any subalgebra of L and let $z \in K$ be so that the Engel subalgebra $E = L_0(\operatorname{ad} z)$ is minimal among all those with annihilator in K. Suppose that $K \subseteq E$. Then $E \subseteq L_0(\operatorname{ad} x)$ for all $x \in K$. In particular, K is nilpotent.*

Recall that we are assuming F has at least $2n+1$ elements where $n = \dim_F L$.

Proof. Take any $x \in K$. Then, for any $c \in F$, $z + cx \in K \subseteq E$. So, $\operatorname{ad}(z+cx)$ stabilizes E and induces an endomorphism of both E and L/E. Since $\operatorname{ad} z$ is an isomorphism on L/E, $\operatorname{ad}(z+cx)$ will be an isomorphism on L/E for all but a finite number ($\leq n$) of $c \in F$. Since F has at least $2n+1$ elements, $\operatorname{ad}(z+cx)$ is an isomorphism on L/E for at least $n+1$ values of c. For each of these values of c we must have

$$L_0(\operatorname{ad}(z+cx)) \subseteq E$$

But, the minimality of E implies that these must be equal. Therefore, $\operatorname{ad}_E(z+cx)$ is nilpotent for $n+1$ values of c. By the previous lemma, $\operatorname{ad}_E(z+cx)$ must be nilpotent for all values of $c \in F$. This implies that $L_0(\operatorname{ad}(z+x)) \supseteq E$ for all $x \in K$. Replacing x with $x - z$ gives the result. □

Lemma 15.1.5. *Any subalgebra K of L which contains an Engel subalgebra is self-normalizing, i.e., $K = N_L(K)$.*

Proof. Suppose that K contains $E = L_0(\operatorname{ad} x)$. Then $x \in E \subseteq K$. So, $[x, K] \subseteq K$ and $[x, N_L(K)] \subseteq K$. So, $\operatorname{ad} x$ stabilizes E, K and $N_L(K)$. But we know that $\operatorname{ad} x$ acts as an automorphism on L/E and therefore on L/K. And $\operatorname{ad} x$ annihilates $N_L(K)/K$. Therefore, $N_L(K)/K = 0$. So, $N_L(K) = K$. □

15.2. Cartan subalgebras.

Theorem 15.2.1. *Let H be a subalgebra of L. Then tfae.*

(1) *H is a minimal Engel subalgebra of L.*
(2) *H is nilpotent and self-normalizing.*

Proof. (1) ⇒ (2) If $H = L_0(\operatorname{ad} x)$ is minimal then $x \in H$ and H is nilpotent by Lemma 15.1.4. By Lemma 15.1.5, any Engel subalgebra is self-normalizing.

(2) ⇒ (1) Since H is nilpotent, $H \subseteq L_0(\operatorname{ad} x)$ for all $x \in H$. In particular, H cannot properly contain any Engel subalgebra. Therefore, it suffices to show that $H = L_0(\operatorname{ad} x)$ for some $x \in H$.

Suppose not. Let $E = L_0(\operatorname{ad} z)$ be minimal. Then, $E \subseteq L_0(\operatorname{ad} x)$ for all $x \in H$ by Lemma 15.1.4. This implies that every element of H acts nilpotently on E/H. By Engel's Lemma 3.1.3, there is some nonzero element $x + H \in E/H$ which is annihilated by every element of H. In other words $[H, x+H] \subseteq H$. So, $[H, x] \subseteq x$ which means $x \in N_L(H) = H$. This is a contradiction. □

Corollary 15.2.2. *If H is a subalgebra of a semisimple Lie algebra L then tfae.*

(2) *H is nilpotent and self-normalizing.*
(3) *H is a Cartan subalgebra of L, i.e., a maximal subalgebra which is abelian in which all the elements are semisimple.*

Proof. We already observed that (3) $\Rightarrow$ (2) since abelian implies nilpotent and $H = L_0$ implies $H = N_L(H)$. To show the converse, suppose that H is a minimal Engel subalgebra of L. Then $H = L_0(\operatorname{ad} x)$. But $x = x_s + x_n$ and x_n is nilpotent on all of L. So, $L_0(\operatorname{ad} x) = L_0(\operatorname{ad} x_s)$. But, for semisimple elements, generalized eigenspaces are the same as eigenspaces. So,

$$H = L_0(\operatorname{ad} x_s) = \ker(\operatorname{ad} x_s) = C_L(x_s)$$

Since Fx_s is abelian and its elements are all semisimple, Fx_s is contained in some Cartan subalgebra C of L. Since C is abelian it is contained in $H = C_L(x_s)$. But we just proved that Cartan subalgebras are Engel subalgebras. Therefore C is an Engel subalgebra contained in the minimal Engel subalgebra H. Therefore $C = H$, making H a Cartan subalgebra. □

Definition 15.2.3. A *Cartan subalgebra* (CSA) of a Lie algebra L is defined to be any nilpotent self-normalizing subalgebra of L.

By the corollary above, this definition agrees with the previous definition when L is semisimple. The proof of the corollary also gives the following.

Corollary 15.2.4. *For any Cartan subalgebra H of a semisimple Lie algebra L, there exists an element $x \in H$ so that $H = C_L(x)$.*

We call x a *regular semisimple element* of L.

15.3. **Functorial properties of CSA's.**

Lemma 15.3.1. *Any epimorphism of Lie algebras $\varphi : L \to L'$ takes CSA's of L onto CSA's of L'.*

Proof. Let H be a CSA of L. Then $\varphi(H)$ is nilpotent. So, it suffices to show that $\varphi(H)$ is self-normalizing. But $\varphi(x)$ normalizes $\varphi(H)$ iff x normalizes $H + \ker\varphi$ in L. But $H + \ker\varphi$ contains the Engel subalgebra H and is thus self-normalizing by Lemma 15.1.5. Therefore, $x \in H + \ker\varphi$ making $\varphi(x) \in \varphi(H)$. So, $\varphi(H)$ is self-normalizing and nilpotent. □

Proposition 15.3.2. *For any epimorphism $\varphi : L \to L'$ and any CSA H' of L', let $K = \varphi^{-1}(H')$. Then any CSA of K is a CSA of L.*

Proof. Let H be a CSA of K. Then H is nilpotent and $H = N_K(H)$. So, it suffices to show that $H = N_L(H)$. Since $\varphi(K) = H'$, the lemma above implies that $\varphi(H) = H'$ (since H' is the only CSA of H'). Let $x \in N_L(H)$. Then $\varphi(x)$ normalizes $\varphi(H) = H'$. So, $\varphi(x) \in H'$ making $x \in K$. But $H = N_K(H)$. So, $x \in H$, proving that $N_L(H) = H$ as required. □

Exercise 15.3.3.
1. If H is a CSA of L then prove that H is a maximal nilpotent subalgebra of L.
2. Prove Lemma 15.1.4 using the assumption that F has at least $n+1$ elements where $n = \dim L$.

16. Conjugacy theorems

We can now prove that the Cartan subalgebra H of any Lie algebra L is unique up to an inner automorphism of L. However, the proof in the book is too complicated. I will present an alternate proof which is based on Tauvel and Yu [7]. They prove directly that CSA's are all conjugate and derive as a fairly easy consequence the fact that Borel subalgebras are all conjugate.

The proof of Tauvel and Yu uses algebraic geometry. In order not to go too far into that area I will assume that the ground field is the complex number $\mathbb{C}$. I will use the elementary fact that any Zariski open subset of $\mathbb{C}^n$ is path connected. I will also use the inverse function theorem from multivariable calculus.

Recall that, for any field F, a subset of F^n is *Zariski closed* if it is the common set of zeros of a set of polynomials. When $n = 1$, a Zariski closed set is either finite or all of F.

Proposition 16.0.4. *Any nonempty Zariski open subset U of $\mathbb{C}^n$ is path connected and dense in the usual Euclidean topology on $\mathbb{C}^n$.*

Proof. Take any two points x, y in U. Let L be the affine line connecting x, y. Topologically, L is a plane since $L \cong \mathbb{C} \cong \mathbb{R}^2$. The intersection with U is Zariski open in L and therefore the complement of a finite subset of L. But the complement of a finite subset of a plane is path connected.

Doing the same thing with $x \in U$ and $y \notin U$ we see that y is in the closure of U. So, U is dense in $\mathbb{C}^n$. □

16.1. Generic elements.

Now take any Lie algebra L. As a vector space $L \cong F^n$. For each element $x \in L$, take the characteristic polynomial $p_x(T)$:

$$p_x(T) = \det(\operatorname{ad} x - TI_n) = (-1)^n T^n + g_1(x)T^{n-1} + \cdots + g_n(x)$$

Each coefficient $g_i(x)$ is a polynomial in x. The last term $g_n(x) = 0$ since $\operatorname{ad} x$ is never an automorphism of L. (Why?) The first term $g_0(x) = (-1)^n$ is never zero.

Definition 16.1.1. The *rank* of L is defined to be the smallest integer r so that $g_{n-r}(x) \neq 0$ for some $x \in L$. The elements $x \in L$ for which $g_{n-r} \neq 0$ are called *generic* or *regular*.

- The generic elements clearly form a nonempty Zariski open subset of L.
- x is generic iff the multiplicity of 0 as an eigenvalue of $\operatorname{ad} x$ is r (the rank of L).
- For $y \in L$ not generic, the multiplicity of 0 as eigenvalue of $\operatorname{ad} y$ is greater than r.
- Any automorphism ψ of L sends generic elements to generic elements.

Lemma 16.1.2. *Suppose that $x \in L$ is generic. Then $L_0(\operatorname{ad} x)$ is the unique CSA of L which contains x and it dimension is equal to the rank r of L.*

Proof. By definition, $\dim L_0(\operatorname{ad} x) = r$ iff x is regular. If y is not regular then $\dim L_0(\operatorname{ad} y) > r$. Therefore, $H = L_0(\operatorname{ad} x)$ is minimal Engel. So, it is a CSA. If C is any other CSA containing x then C is nilpotent. So, $\operatorname{ad} x$ is nilpotent on C which means $C \subseteq H$. But H is minimal. So, $C = H$. □

Next we will show that every CSA $H \subseteq L$ contains a generic element. The idea is to deform H by an automorphism of L close to the identity map. Since generic elements form an open dense subset of L, a general deformation of H will contain a generic element. So, H must contain a generic element. To do this rigorously, we first need to define the automorphism groups that we want to use.

16.2. **The group $\mathcal{E}(L)$.** Assume that F is algebraically closed of characteristic 0. Then for any $x \in L$, L decomposes into a direct sum of the generalized eigenspaces of $\operatorname{ad} x$:

$$L = \bigoplus_{\lambda} L_\lambda(\operatorname{ad} x)$$

where

$$L_\lambda(\operatorname{ad} x) = \ker(\operatorname{ad} x - \lambda)^n$$

Definition 16.2.1. $x \in L$ is called *strongly ad-nilpotent* if $x \in L_\lambda(\operatorname{ad} y)$ for some $y \in L$ and $\lambda \in F$. The set of strongly ad-nilpotent elements of L is denoted $\mathcal{N}(L)$. The subgroup of $\operatorname{Aut}(L)$ generated by $\exp \operatorname{ad} x$ for all $x \in \mathcal{N}(L)$ is denoted $\mathcal{E}(L)$.

Exercise 16.2.2. If K is a subalgebra of L then show that $\mathcal{N}(K) \subseteq \mathcal{N}(L)$ and that, for any $x \in \mathcal{N}(K)$, $\operatorname{ad}_K(x)$ is the restriction to K of $\operatorname{ad}_L(x)$. (Let $\mathcal{E}(L, K)$ denote the subgroup of $\mathcal{E}(L)$ generated by $\exp \operatorname{ad} x$ for all $x \in \mathcal{N}(K)$.)

If $\varphi : L \to L'$ is an epimorphism and $y \in L$ then show that $\varphi(L_\lambda(\operatorname{ad} y)) = L'_\lambda(\operatorname{ad} \varphi(y))$. Conclude that $\varphi(\mathcal{N}(L)) = \mathcal{N}(L')$.

Lemma 16.2.3. *If $\varphi : L \to L'$ is an epimorphism and $\sigma' \in \mathcal{E}(L')$ then there is a $\sigma \in \mathcal{E}(L)$ so that $\varphi\sigma = \sigma'\varphi$. I.e., the following diagram commutes.*

$$\begin{array}{ccc} L & \xrightarrow{\varphi} & L' \\ {\scriptstyle\sigma}\downarrow & & \downarrow{\scriptstyle\sigma'} \\ L & \xrightarrow{\varphi} & L' \end{array}$$

Proof. σ' is a product (composition) of automorphism of L' of the form $\exp \operatorname{ad} x'$ for some $x' \in \mathcal{N}$. By the exercise, $x' = \varphi(x)$ for some $x \in \mathcal{N}(L)$. Then

$$\exp \operatorname{ad}(\varphi(x))\varphi(y) = \varphi\left((\exp \operatorname{ad} x)(y)\right)$$

In other words, $\varphi(\exp \operatorname{ad} x) = (\exp \operatorname{ad} x')\varphi$. Let σ be the product of the liftings $\exp \operatorname{ad} x$ for each factor $\exp \operatorname{ad} x'$ of σ'. $\square$

From now on, we assume that $F = \mathbb{C}$.

Lemma 16.2.4. *The derivative of* $\exp \operatorname{ad} ty = \exp(t \operatorname{ad} y)$ *at* $t = 0$ *is* $\operatorname{ad} y$.

Proof. Let $z = \operatorname{ad} y$. Then

$$\exp(tz) = id_L + tz + t^2 z/2 + \cdots$$

$$\frac{d}{dt}\exp(tz) = z + tz + t^2 z/2 + \cdots$$

which is equal to z at $t = 0$. $\square$

The lemma says that, when $t \in \mathbb{C}$ is close to 0 ($|t|$ is small).

$$(\exp \operatorname{ad} ty)(x) \approx x + t[y, x] \qquad (= x - t[x, y])$$

16.3. **Conjugacy of CSA's.**

Lemma 16.3.1. *Every CSA H of L contains a generic element. ($\Rightarrow \dim H = \text{rank } L$)*

Proof. Since H is Engel, $H = L_0(\operatorname{ad} x)$ for some $x \in H$. The generalized eigenspace decomposition of L wrt $\operatorname{ad} x$ is:

$$H \oplus L_{\lambda_1} \oplus L_{\lambda_2} \oplus \cdots \oplus L_{\lambda_k}$$

with $H = L_0$ having $\dim H = k \geq r$ and suppose $\dim L_{\lambda_i} = m_i$. For each i, choose a Jordan canonical form for $\operatorname{ad} x$ acting on L_{λ_i}. This gives a basis of y_{ij} for each L_{λ_i} so that the matrix of $\operatorname{ad} x$ is $\lambda_i I_{m_i}$ plus an upper triangular matrix. Then each $y_{ij} \in \mathcal{N}(L)$ and $\sum m_i = n - k$.

Consider the function $\psi : H \times \mathbb{C}^{n-k} \to L$ given by

$$\psi(h,t) = \sigma_t(h) \text{ where } \sigma_t = \prod_{j=1}^{m} \exp \operatorname{ad} t_{ij} y_{ij} \in \mathcal{E}(L)$$

The function ψ is smooth (C^∞) since, in matrix form, every entry is a polynomial. (ψ is *algebraic.*)

Claim The derivative of ψ at $(x, 0)$ is nonsingular.

Pf: By the lemma, the partial derivative of $\psi(h,t)$ wrt t_{ij} at $(h,t) = (x,0)$ is

$$\operatorname{ad} y_{ij}(x) = -[x, y_{ij}] = -\lambda_i y_{ij} - (*) y_{i,j-1}$$

where $* = 0$ or 1. if $y_j \in L_{\lambda_i}$. Therefore the matrix of the derivative $D\psi_{(x,0)}$ wrt the basis $\{y_{ij}\}$ plus any basis for H is

$$D\psi_{(x,0)} = \begin{bmatrix} I_k & 0 \\ 0 & U \end{bmatrix}$$

where U is an upper triangular matrix with nonzero diagonal entries $-\lambda_i$.

By the inverse function theorem, the image of ψ contains an open (in the Euclidean topology) neighborhood U of the point $x \in L$. Since the set of generic elements is open and dense in L, this open neighborhood contains a generic point x_{gen}. By definition $x_{gen} = \psi(h,t)$ for some $(h,t) \in H \times \mathbb{C}^{n-k}$. This implies that $h \in H$ is generic. Furthermore, $x_{gen} = \sigma(h)$ for $\sigma = \sigma_t \in \mathcal{E}(L)$. □

Note that, at the last step, we also get that $\sigma(H) = L_0(x_{gen})$ since this is the unique CSA containing x_{gen}. This proves the following lemma.

Lemma 16.3.2. *For any generic x in L there is an open neighborhood U of x so that, for any other generic element $x' \in U$ there is a $\sigma \in \mathcal{E}(L)$ so that $\sigma(L_0(\operatorname{ad} x)) = L_0(\operatorname{ad} x')$.*

Theorem 16.3.3. *The group $\mathcal{E}(L)$ acts transitively on the set of CSA's of L.*

Proof. Take any two CSA's $H_0, H_1 \subseteq L$. By the first lemma above, each $H_i = L_0(\operatorname{ad} x_i)$ for x_0, x_1 generic. Choose a continuous path $x_t, 0 \leq t \leq 1$ from x_0 to x_1 in the set of generic elements. Let $H_t = L_0(\operatorname{ad} x_t)$ for each $t \in [0,1]$. By the second lemma above, there is an open neighborhood U_t of each x_t and therefore an open interval J_t around $t \in [0,1]$ so that for all $s \in J_t$, $H_s = \sigma(H_t)$ for some $\sigma \in \mathcal{E}(L)$. Since $[0,1]$ is compact, there is a finite covering by these intervals J_t. We conclude that $H_1 = \sigma H_0$ for some $\sigma \in \mathcal{E}(L)$. □

16.4. **Borel subalgebras.** The book uses Borel subalgebras to prove that CSA's are conjugate. We are doing this backwards. Given that CSA's are conjugate by the action of $\mathcal{E}(L)$, we will show that Borel subalgebras are all conjugate. This is Humphreys proof which we read backwards.

16.4.1. *definition and basic properties.* A *Borel subalgebra* B of L is defined to be a maximal solvable subalgebra. The first two basic properties are clear.

Lemma 16.4.1. *Every Borel subalgebra is self normalizing: $B = N_L(B)$.* □

Recall that $Rad(L)$ is the unique maximal solvable ideal of L and that $L/Rad(L)$ is semisimple.

Lemma 16.4.2. *Every Borel subalgebra B of L contains the solvable radical $Rad(L)$. And $B \leftrightarrow B/Rad(L)$ gives a bijection between the set of Borel subalgebras of L and those of $L/Rad(L)$.* □

This lemma reduces the problem to the case when L is semisimple. We will show the following.

Theorem 16.4.3. *For any two Borel subalgebras B, B' of a semisimple Lie algebra L, there is an automorphism $\sigma \in \mathcal{E}(L)$ so that $B' = \sigma(B)$.*

With the lemma and the functorial properties of $\mathcal{E}(L)$ we get the following.

Corollary 16.4.4. *For any two Borel subalgebras B, B' of a Lie algebra L, there is an automorphism $\sigma' \in \mathcal{E}(L)$ so that $B' = \sigma'(B)$.*

Proof. By the theorem, there is a $\sigma \in \mathcal{E}(L/Rad(L))$ so that $\sigma(B/Rad(L)) = B'/Rad(L)$. By the functorial properties of $\mathcal{E}$, $\sigma \in \mathcal{E}(L/Rad(L))$ lifts to an element $\sigma' \in \mathcal{E}(L)$ so that $\varphi\sigma' = \sigma\varphi$. So,

$$\varphi\sigma'(B) = \sigma'(B)/Rad(L) = \sigma(B/Rad(L)) = B'/Rad(L)$$

and therefore, $\sigma(B) = B'$. □

It remains to prove the theorem.

16.4.2. *standard Borel subalgebras.* The standard example of a Borel subalgebra is given as follows.

Lemma 16.4.5. *Let L be a semisimple Lie algebra H a CSA, with root system Φ and base Δ. Then*

$$B(\Delta) := H \oplus \bigoplus_{\beta\in\Phi_+} L_\beta$$

is a Borel subalgebra of L. These are called the Standard Borel subalgebras *of L. Conversely, any Borel subalgebra of L which contains H is standard.*

Exercise 16.4.6. Show that $N(\Delta) := \bigoplus_{\beta\in\Phi_+} L_\beta$ is a maximal nilpotent subalgebra of L.

Proof. It is clear that $B(\Delta)$ is solvable since $[B(\Delta, B(\Delta)] = N(\Delta)$ is nilpotent.

Conversely, we claim that any solvable subalgebra B of L which contains H is contained in $B(\Delta)$ for some base Δ for Φ. This will prove both statements in our lemma. To prove this, let B be solvable and $H \subseteq B$. Then B has a weight space decomposition wrt H:

$$B = H \oplus \bigoplus_{\beta \in S} L_\beta$$

for some set of roots $S \subseteq \Phi$. But S cannot contain contain both β and $-\beta$ since, otherwise, L would contain the semisimple Lie algebra $S_\beta = L_{-\beta} \oplus Fh_\beta \oplus L_\beta$ which is not possible since all subalgebras of solvable algebras are solvable. But this implies that $S \subseteq \Phi_+$ with respect to some base Δ and thus $B \subseteq B(\Delta)$. □

Lemma 16.4.7. *Given a fixed CSA H of a semisimple Lie algebra L, any two Borel subalgebras B containing H are conjugate by an element of $\mathcal{E}(L)$.*

Proof. By the previous lemma we know that each $B \supseteq H$ has the form $B(\Delta)$. We also know that the Weyl group W acts transitively on the set of bases Δ. But W is generated by reflections σ_α which lift to the elements of $\mathcal{E}(L)$ given by:

$$\tau_\alpha = \exp \operatorname{ad} x_\alpha \cdot \exp \operatorname{ad} (-y_\alpha) \cdot \exp \operatorname{ad} x_\alpha$$

Claim 1 $\tau_\alpha(h_\alpha) = -h_\alpha$

Pf: Since $[x_\alpha, h_\alpha] = -2x_\alpha$, $[y_\alpha, h_\alpha] = 2y_\alpha$ and $[y_\alpha, x_\alpha] = -h_\alpha$ we get:

$$\exp \operatorname{ad} x_\alpha(h_\alpha) = h_\alpha + [x_\alpha, h_\alpha] + \frac{1}{2}[x_\alpha[x_\alpha, h_\alpha]] + \cdots = h_\alpha - 2x_\alpha$$

$$\begin{aligned}\exp \operatorname{ad} (-y_\alpha)(h_\alpha - 2x_\alpha) &= (h_\alpha - 2x_\alpha) - [y_\alpha, (h_\alpha - 2x_\alpha)] + \frac{1}{2}[y_\alpha[y_\alpha, (h_\alpha - 2x_\alpha)]] + \cdots \\ &= (h_\alpha - 2x_\alpha) - (2y_\alpha + 2h_\alpha) + \frac{1}{2}(4y_\alpha) \\ &= -h_\alpha - 2x_\alpha\end{aligned}$$

$$\exp \operatorname{ad} x_\alpha(-h_\alpha - 2x_\alpha) = (-h_\alpha - 2x_\alpha) + 2x_\alpha = -h_\alpha$$

Claim 2 If $h \in K = \ker \alpha$ then $\tau_\alpha(h) = h$.

Pf: This follows from the fact that $[h, x_\alpha] = \alpha(h)x_\alpha = 0$ and similarly $[h, y_\alpha] = 0$.

Together, these two Claims imply that $\tau_\alpha^* : H^* \to H^*$ sends α to $-\alpha$ since $\alpha\tau_\alpha(h) = \alpha(h) = 0$ for $h \in K$ and $\alpha\tau_\alpha(h_\alpha) = -2h_\alpha$. For any $x \in H^*$ orthogonal to α we have $x(h_\alpha) = 0$. So, $x\tau_\alpha(h_\alpha) = 0$ and $x\tau_\alpha(h) = x(h)$ for any $h \in K$. So, $\tau_\alpha^*(x) = x$. In other words, τ_α^* is equal to the reflection σ_α as claimed. Therefore, the subgroup of $\mathcal{E}(L)$ generated by the τ_α acts transitively on the set of Borel subalgebras of L containing H. □

Since $\mathcal{E}(L)$ acts transitively on the set of CSA's of L we get the following.

Theorem 16.4.8. *$\mathcal{E}(L)$ acts transitively on the set of all standard Borel subalgebras of L.*

The only thing left to show is the following.

Theorem 16.4.9. *All Borel subalgebras of a semisimple Lie algebra L are standard.*

Proof. Let B be a Borel subalgebra of L. To show that B is standard, it suffices to show that B contains a CSA H of L. So, we may assume that B does not contain any CSA.

Claim 1 B contains at least one nonzero semisimple element t.

Pf: We know that B is self-normalizing. So, it cannot be nilpotent. Otherwise, B would be a CSA. So, B contains a nonzero element x which is not nilpotent. Then $x_s \neq 0$. But $[x_s, B] \subseteq B$ since x_s normalizes any subalgebra which is normalized by x. So, $x_s \in N_L(B) = B$ as claimed.

Claim 2 We may assume t is not in the center of B.

Since t is semisimple, it is contained in a CSA H of L and H is abelian. If t is central in B then $H, B \subset C = C_L(t) \subsetneq L$ are CSA and Borel respectively in C. By induction on the size of L, there is some $\sigma \in \mathcal{E}(C)$ so that $\sigma(H) \subseteq B$. But $\sigma(H)$ is also a CSA of L and we are done.

Since t is semisimple, B decomposes as eigenspaces of the action of $\operatorname{ad} t$:

$$B = B_0 \oplus \bigoplus B_\lambda$$

Since t is not central in B, one of the B_λ is nonzero, $\lambda \neq 0 \in F$. Take $x \neq 0$ in B_λ. Then x is nilpotent (since $[x, B_a] \subseteq B_{a+\lambda}$) and $[t, x] = \lambda x$. So, $[t_0, x] = x$ where $t_0 = \frac{1}{\lambda} t$.

Let H be a CSA of L which contains t. Then, using the root space decomposition of L wrt H, we get

$$x = x_0 + \sum_{\beta \in S} c_\beta x_\beta$$

where $x_0 \in H, x_\beta \in L_\beta$ and $c_\beta \neq 0 \in F$ for each β is a subset $S \subseteq \Phi$.

Claim 3 $x_0 = 0$ and $S \subseteq \Phi_+$ with respect to some base Δ. Thus $x \in N(\Delta)$.

Pf: The calculation

$$[t_0, x] = x = \sum c_\beta \beta(t_0) x_\beta$$

implies that $x_0 = 0$ and $\beta(t_0) = 1$ for all $\beta \in S$. So, β and $-\beta$ cannot both be in S. Claim 3 follows.

This implies $A = B \cap B(\Delta)$ contains the nonzero nilpotent element $x \in N = B \cap N(\Delta)$. Choose Δ so that A is maximal and $N \neq 0$. (If $A = B$ we are done.) Since $N(\Delta)$ is an ideal in $B(\Delta)$, N is an ideal in A. So, $A \subset K = N_L(N) \subsetneq L$. ($L$ has no ideals.)

Claim 4 A is properly contained in $B \cap K$ and in $B(\Delta) \cap K$.

Pf: Since $N \subseteq B$ is nilpotent, the action of N on B/A is nilpotent and there is a $y \in B$, $y \notin A$ so that $[N, y] \subseteq A$. But $[N, y] \subseteq [B, B]$ is nilpotent. So, we must have $[N, y] \subseteq N$. So, $y \in K \cap B$ but $y \neq A$. The other case is similar.

Let C, C' be Borel subalgebras of K which contain $B \cap K, B(\Delta) \cap K$ respectively. (See Figure 1 on p.85 in the book.) Then, by induction on $\dim L$, there is a $\sigma \in \mathcal{E}(K)$ sending C to C'. So, we may assume that $C = C'$. Let B' be a Borel subalgebra of L containing $C = C'$. Then, $B' \cap B(\Delta) \supseteq K \cap B(\Delta) \supsetneq A = B \cap B(\Delta)$. Therefore, by maximality of A, B' is standard. But we also have $B \cap B' \subseteq B \cap K \supsetneq A = B \cap B(\Delta)$. So, B is also standard by maximality of A. So, every Borel subalgebra of L is standard. □

17. Universal enveloping algebras

Recall that, for an associative algebra A with unity (1), a Lie algebra structure on A is given by the Lie bracket $[ab] = ab - ba$. Let $\mathcal{L}(A)$ denote this Lie algebra. Then $\mathcal{L}$ is a *functor* which converts associative algebras into Lie algebras. Every Lie algebra L has a universal enveloping algebra $\mathcal{U}(L)$ which is an associative algebra with unity. The functor $\mathcal{U}$ is "adjoint" to the functor $\mathcal{L}$. The universal enveloping algebra is defined by category theory. The Poincaré-Birkoff-Witt Theorem gives a concrete description of the elements of the elements of $\mathcal{U}(L)$ and how they are multiplied. There is also a very close relationship with the multiplication rule in the associated Lie group.

17.1. Functors.

I won't go through the general definition of categories and functors since we will be working with specific functors not general functors. I will just use vector spaces over a field F, Lie algebras and associative algebras (always with unity) as the main examples.

Definition 17.1.1. A *functor* from the category of vector spaces to the category of associative algebras both over F is defined to be a rule $\mathcal{F}$ which assigns to each F-vector space V an associative algebra $\mathcal{F}(V)$ over F and to each linear map $f : V \to W$, an F-algebra homomorphism $f_* : \mathcal{F}(V) \to \mathcal{F}(W)$ so that two conditions are satisfied:

(1) $(id_V)_* = id_{\mathcal{F}(V)}$
(2) $(fg)_* = f_* g_*$.

Recall that an *F-algebra* is an algebra which is also a vector space over F so that multiplication if F-bilinear. An *F-algebra homomorphism* is a ring homomorphism which is also F-linear. We say that the homomorphism is *unital* if it takes 1 to 1.

In short: a functor takes objects to objects and morphisms to morphism and satisfies the two conditions listed above.

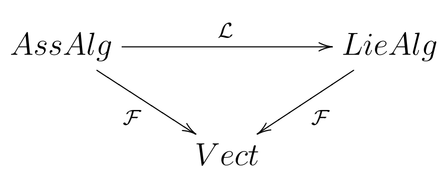
Example 17.1.2. The mapping $A \mapsto \mathcal{L}(A)$ is a functor from associative algebras to Lie algebras. For this functor, $f_* = f$ for all F-algebra homomorphisms $f : A \to B$. The reason that this works is elementary:

$$f[a, b] = f(ab - ba) = f(a)f(b) - f(b)f(a) = [f(a), f(b)]$$

We say that f_* is f considered as a homomorphism of Lie algebras $\mathcal{L}(A) \to \mathcal{L}(B)$. The two conditions are obviously satisfied and this defines a functor.

Example 17.1.3. The *forgetful functor* $\mathcal{F}$ takes an associative algebra (or Lie algebra) A to the underlying vector space. $\mathcal{F}$ is defined on morphisms by $f_* = f$. Since F-algebra homomorphisms are F-linear by definition, this defines a functor.

Exercise 17.1.4. Show that the following diagram commutes.



17.2. **Tensor and symmetric algebras.** These are two important algebras associated to any vector space. They are both graded algebras.

Definition 17.2.1. A *graded algebra* over F is an algebra A together with a direct sum decomposition:
$$A = A^0 \oplus A^1 \oplus A^2 \oplus \cdots$$
so that $A^iA^j \subseteq A^{i+j}$. If A has unity (1) it should be in A^0. Elements in A^n are called *homogeneous of degree n*.

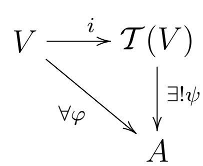
Example 17.2.2. The polynomial ring $P = F[X_1, \cdots, X_n]$ is a graded ring with P^k being generated by degree k monomials. The noncommutative polynomial ring $Q = F\langle X_1, \cdots, X_n\rangle$ is also a graded ring with Q^k being generated by all words of length k in the letters $X_1, \cdots, X_n$. An example of a graded Lie algebra is the standard Borel subalgebra B of any semisimple Lie algebra L. Then $B^0 = H$ is the CSA and B^k is the direct sum of all B_β where β has height k.

Exercise 17.2.3. Show that P^k has dimension $\binom{n+k-1}{k}$. For example, for $n = 2$, $\dim P^k = k+1$ with basis elements x^iy^{k-i} for $i = 0, \cdots, k$. Q^k has dimension n^k.

Definition 17.2.4. Given a vector space V, the *tensor algebra* $\mathcal{T}(V)$ is defined to be the vector space
$$\mathcal{T}(V) = F \oplus V \oplus V \otimes V \oplus V \otimes V \otimes V \oplus \cdots$$
with multiplication defined by tensor product (over F). This is an associative graded algebra with $\mathcal{T}^k(V) = V^{\otimes k}$, the k-fold tensor product of V with itself. Note that, in degree 1, we have $\mathcal{T}^1(V) = V$.

If $V = F^n$ then $\mathcal{T}(F^n) \cong F\langle X_1, \cdots, X_n\rangle$. For example, $V \otimes V$ is n^2 dimensional with basis given by $e_i \otimes e_j$. The tensor algebra has the following universal property.

Proposition 17.2.5. *Any linear map φ from a vector space V to an associative algebra A with unity extends uniquely to a unital algebra homomorphism $\psi : \mathcal{T}(V) \to A$:*



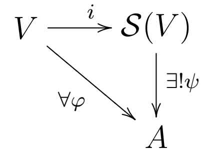
Proof. ψ is must be given by $\psi(1) = 1$, $\psi(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = \varphi(v_1)\varphi(v_2)\cdots\varphi(v_k)$. □

This proposition means that $\mathcal{T}(V)$ is the universal associative algebra with unity generated by V.

Definition 17.2.6. The *symmetric algebra* $\mathcal{S}(V)$ generated by V is defined to be the quotient of $\mathcal{T}(V)$ by the ideal generated by all elements of the form $x \otimes y - y \otimes x$. This makes $\mathcal{S}(V)$ into a commutative graded algebra with unity. Since the relations are in degree 2, the degree 1 part is still the same: $\mathcal{S}^1(V) = \mathcal{T}^1(V) = V$.

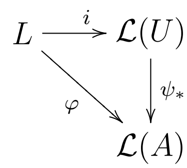
For example, when $V = F^n$ we have $\mathcal{S}(F^n) \cong F[X_1, \cdots, X_n]$.

Proposition 17.2.7. *Any linear map φ from a vector space V to a commutative algebra A with unity extends uniquely to a unital algebra homomorphism $\psi : \mathcal{S}(V) \to A$:*



17.3. **Universal enveloping algebra.** Following tradition, we define this by its desired universal property.

Definition 17.3.1. For any (possibly infinite dimensional) Lie algebra L, the *universal enveloping algebra* of L is defined to be any pair (U, i) where U is an associative algebra with unity and $i : L \to \mathcal{L}(U)$ is a Lie algebra homomorphism with the property that, for any other associative algebra with unity A and any Lie algebra homomorphism $\varphi : L \to \mathcal{L}(A)$ there is a unique unital algebra homomorphism $\psi : U \to A$ so that the following diagram commutes where $\psi_* = \psi$ considered as a homomorphism of Lie algebras.



Example 17.3.2. One important example is the case when $A = \mathrm{End}_F(V)$ is the algebra of F-linear endomorphisms of a vector space V. Then $\mathcal{L}(A) = \mathfrak{gl}(V)$ and $\varphi : L \to \mathcal{L}(A) = \mathfrak{gl}(V)$ is a representation of L making V into an L-module. The algebra homomorphism $\psi : U \to A = \mathrm{End}_F(V)$ makes V into a module over the associative algebra U. Therefore, a module over L is the same as a module over U.

Proposition 17.3.3. *The universal enveloping algebra (U, i) of L is unique up to isomorphism if it exists.*

Proof. If there is another pair (U', i') then, by the universal property, there are algebra homomorphisms $\psi : U \to U'$ and $\psi' : U' \to U$ so that $i' = \psi_* i$ and $i = \psi'_* i'$. But then $i = \psi'_* \psi_* i = (\psi'\psi)_* i$. By uniqueness, we must have $\psi'\psi = id_U$. Similarly $\psi\psi' = id_{U'}$. So, $U \cong U'$ and i, i' correspond under this isomorphism. □

The construction of U is easy when we consider the properties of an arbitrary Lie algebra homomorphism

$$\varphi : L \to \mathcal{L}(A)$$

Since φ is a linear mapping from L to A, it extends uniquely to a unital algebra homomorphism $\overline{\varphi} : \mathcal{T}(L) \to A$. Taking into account that φ is also a Lie algebra homomorphism, we see that, for any two elements $x, y \in L$, we must have

$$\varphi([x, y]) = [\varphi(x), \varphi(y)] = f(x)f(y) - f(y)f(x) = \overline{\varphi}(x \otimes y - y \otimes x)$$

In other words, $\overline{\varphi} : \mathcal{T}(L) \to A$ has the elements

$$x \otimes y - y \otimes x - [x, y]$$

in its kernel. Let J be the two-sided ideal in $\mathcal{T}(L)$ generated by all elements of this form. Then J is in the kernel of $\overline{\varphi}$ and we have an induced unital algebra homomorphism $\psi : \mathcal{T}(L)/J \to A$.

Definition 17.3.4. $\mathcal{U}(L)$ is defined to be the quotient of $\mathcal{T}(L)$ by the ideal generated by all $x \otimes y - y \otimes x - [x, y]$ for all $x, y \in L$. Let $i : L \to \mathcal{U}(L)$ be the inclusion map $i(x) = x$.

Note that the relations imposed on $\mathcal{U}(L)$ are the minimal ones needed to insure that $i : L \to \mathcal{L}(\mathcal{U}(L))$ is a Lie algebra homomorphism. The fact that $(\mathcal{U}(L), i)$ satisfies the definition of a universal enveloping algebra is supposed to be obvious.

Exercise 17.3.5. Show that, for any associative algebra A, there is a canonical unital algebra homomorphism $\mathcal{U}(\mathcal{L}(A)) \to A$. When is this an isomorphism? What happens when A is commutative?

In the special case that L is a graded Lie algebra, such as a standard Borel algebra, $\mathcal{T}(L)$ has another grading given by

$$\mathcal{T}(L)^k = \bigoplus_{\sum j_i = k} L^{j_1} \otimes L^{j_2} \otimes \cdots \otimes L^{j_m}$$

and the ideal J is generated by homogeneous elements. This makes $\mathcal{U}(L)$ into a graded algebra. In general, there is no graded structure on $\mathcal{U}(L)$. However, there is a filtration $\mathcal{U}_k(L)$ induced by the filtration

$$\mathcal{T}_k(L) = \mathcal{T}^0(L) \oplus \mathcal{T}^1(L) \oplus \cdots \oplus \mathcal{T}^k(L)$$

of $\mathcal{T}(L)$.

By a *filtration* of an algebra A we mean a sequence of vector subspaces

$$A_0 \subseteq A_1 \subseteq A_2 \subseteq \cdots$$

so that $A = \bigcup A_k$ and so that $A_j A_k \subseteq A_{j+k}$. The *associated graded algebra* is defined by $G^k(A) = A_k / A_{k-1}$ with multiplication $G^j G^k \to G^{j+k}$ induced by the filtered multiplication on A. We let $\mathcal{G}(L)$ denote the associated graded algebra of $\mathcal{U}(L)$.

Example 17.3.6. Suppose that L is abelian. Then $\mathcal{U}(L)$ is $\mathcal{T}(L)$ modulo the ideal generated by $x \otimes y - y \otimes x$ since $[x, y] = 0$. But this means $\mathcal{U}(L)$ is the symmetric algebra $\mathcal{S}(L)$. If $\dim L = n$ then, in filtration k, we have $\mathcal{U}_k(L) = \mathcal{S}_k(L)$ which is equivalent to the vector space of polynomials of degree $\leq k$ in n variables, or equivalently, homogeneous polynomials of degree equal to k in $n + 1$ variables. Thus

$$\dim \mathcal{U}_k(L) = \dim \mathcal{S}_k(F^n) = \dim \mathcal{S}^k(F^{n+1}) = \binom{n+k}{k}$$

In this example, $\mathcal{U}(L) \cong \mathcal{G}(L) \cong \mathcal{S}(L)$.

17.4. **Poincaré-Birkov-Witt.** The PBW Theorem gives a detailed description of the structure of the universal enveloping algebra $\mathcal{U}(L)$. It gives a formula for a vector space basis and how they are multiplied. One of the amazing features of the theorem is that it says that the dimension of $\mathcal{U}_k(L)$ depends only on the dimension of L. I.e., the dimension is always $\binom{n+k}{k}$. For each $k \geq 0$ consider the composition:

$$\varphi^k : \mathcal{T}^k(L) = L^{\otimes k} \hookrightarrow \mathcal{T}_k(L) \twoheadrightarrow \mathcal{U}_k(L) \twoheadrightarrow \mathcal{G}^k(L) = \mathcal{U}_k(L)/\mathcal{U}_{k-1}(L)$$

Lemma 17.4.1. *$\varphi = \sum \varphi^k : \mathcal{T}(L) \to \mathcal{G}(L)$ is a graded algebra epimorphism which induced a graded algebra epimorphism $\mathcal{S}(L) \twoheadrightarrow \mathcal{G}(L)$.*

Proof. The morphism φ is multiplicative by definition. So, it is an algebra epimorphism. Elements of the form $x \otimes y - y \otimes x \in \mathcal{T}^2(L)$ are sent to $[x, y] \in \mathcal{U}_1(L)$ which is zero in $G^2(L)$. Therefore, there is an induced algebra epimorphism $\mathcal{S}(L) \twoheadrightarrow \mathcal{G}(L)$. □

Theorem 17.4.2 (PBW)**.** *The natural graded epimorphism $\mathcal{S}(L) \to \mathcal{G}(L)$ is always an isomorphism.*

Corollary 17.4.3 (PBW basis)**.** *If $x_1, \cdots, x_n$ is a vector space basis for L then a vector space basis for $\mathcal{U}_k(L)$ is given by all monomials of length $\leq k$ of the form*

$$x_{j_1} x_{j_2} \cdots x_{j_\ell}$$

where $j_1 \leq j_2 \leq j_3 \leq \cdots \leq j_\ell$ plus 1 (given by the empty word). In particular, $i : L \to \mathcal{U}(L)$ is a monomorphism.

Proof. Monomials as above (with nondecreasing indices) of length equal to k form a basis for $\mathcal{S}^k(L)$ and therefore give a basis for $\mathcal{U}_k(L)$ modulo $\mathcal{U}_{k-1}$. □

Corollary 17.4.4. *If H is any subalgebra of L then the inclusion $H \hookrightarrow L$ extends to a monomorphism $\mathcal{U}(H) \hookrightarrow \mathcal{U}(L)$. Furthermore $\mathcal{U}(L)$ is a free $\mathcal{U}(H)$-modules.*

Proof. Extend an ordered basis of H to an ordered basis for L and use PBW bases. □

We will skip the proof of the PBW Theorem in the lectures. So, you need to read the proof. There is a purely algebraic proof in the book. Here I will give a proof using combinatorial group theory.

17.5. **Proof of PBW Theorem.** I will first reduce the proof of PBW to a the proof of a key lemma and then use combinatorial group theory to prove the key lemma.

17.5.1. *proof assuming key lemma.* Choose a basis B for L. Then $\mathcal{T}^n = \mathcal{T}^n(L)$ has as basis the set of all tensors $x_1 \otimes x_2 \otimes \cdots \otimes x_n$ where $x_i \in B$. Let J^n be the span in $\mathcal{T}_n$ of all elements of the form

$$x_1 \otimes \cdots \otimes x_{i-1} \otimes (x_i \otimes x_{i+1} - x_{i+1} \otimes x_i - [x_i, x_{i+1}]) \otimes x_{i+2} \otimes \cdots \otimes x_n$$

($J^n = 0$ if $n < 2$). Modulo $\mathcal{T}_{n-1}$, these elements span the kernel I^n of the epimorphism $\mathcal{T}^n(L) \twoheadrightarrow \mathcal{S}^n(L)$. Therefore $(J^n + \mathcal{T}_{n-1}) \cap \mathcal{T}^n = I^n$.

Let $J_m = \bigoplus_{k \le m} J^k$. Then

$$(J_n + \mathcal{T}_{n-1}) \cap \mathcal{T}^n = I^n.$$

Lemma 17.5.1 (Key Lemma)**.** *For any $n \ge 2$ we have:*

$$J_n \cap \mathcal{T}_{n-1} \subseteq J_{n-1}$$

This implies that $J_m \cap \mathcal{T}_{n-1} \subseteq J_{n-1}$ for all $m \ge n$ since

$$J_m \cap \mathcal{T}_{n-1} = (J_m \cap \mathcal{T}_{m-1}) \cap \mathcal{T}_{n-1} \subseteq J_{m-1} \cap \mathcal{T}_{n-1}$$

which is contained in J_{n-1} by induction on $m - n$.

Suppose for a moment that this is true. Let I^n be the kernel of $\mathcal{T}^n(L) \twoheadrightarrow \mathcal{S}^n(L)$.

Lemma 17.5.2. *For all $m \ge n$ we have*

$$(J_m + \mathcal{T}_{n-1}) \cap \mathcal{T}^n = I^n$$

Proof. We already observed that this statement is true for $m = n$. For $m > n$ we have the following by the Key Lemma.

$$I^n = (J_n + \mathcal{T}_{n-1}) \cap \mathcal{T}^n \subseteq (J_m + \mathcal{T}_{n-1}) \cap \mathcal{T}^n = (J_m + \mathcal{T}_{n-1}) \cap \mathcal{T}_n \cap \mathcal{T}^n \subseteq (J_{m-1} + \mathcal{T}_{n-1}) \cap \mathcal{T}^n$$

which is equal to I^n by induction on $m - n$. □

Proof of PBW. The kernel of the map $\mathcal{T}^n \twoheadrightarrow \mathcal{U}_n(L)/\mathcal{U}_{n-1}(L)$ is the union of all $(J_m + \mathcal{T}_{n-1}) \cap \mathcal{T}^n$. But these are all equal to I^n. So, $\mathcal{T}^n/I^n = \mathcal{S}^n(L) \cong \mathcal{U}_n(L)/\mathcal{U}_{n-1}(L)$ as claimed. □

17.5.2. *braid relations and Jacobi identity.* To prove the Key Lemma, we will first find special elements of $J_n \cap \mathcal{T}_{n-1}$ given by braid relations and show that each one lies in J_{m-1} we will call these *braid elements.* Since the symmetric group on n letters is generated by the simple reflections $s_1, \cdots, s_{n-1}$ modulo the braid relations it will follow that that the braid elements span $J^n \cap \mathcal{T}_{n-1}$ and the Key Lemma will follow.

Define a left action of the group S_n on $\mathcal{T}^n$ in the obvious way by permutation of tensor factors. For any generator $x = x_1 \otimes \cdots \otimes x_n$ in $\mathcal{T}^n$ and any $\sigma \in S_n$ the formula for this action is

$$\sigma x := x_{\rho(1)} \otimes x_{\rho(2)} \otimes \cdots \otimes x_{\rho(n)}$$

where $\rho = \sigma^{-1}$. E.g., if $\sigma = (123)$ then $\sigma x = x_3 \otimes x_1 \otimes x_2 = x_{\sigma^{-1}(1)} \otimes x_{\sigma^{-1}(2)} \otimes x_{\sigma^{-1}(3)}$.

For any $1 \le i \le n-1$, let $\partial_i x \in \mathcal{T}^{n-1}$ be the element given by

$$\partial_i x = x_1 \otimes \cdots \otimes x_{i-1} \otimes [x_i, x_{i+1}] \otimes x_{i+2} \otimes \cdots \otimes x_n$$

Then the generators of J^n can be written as

$$d(i,x) := x - s_i x - \partial_i x \quad \in \mathcal{T}_n$$

Lemma 17.5.3 (First braid relations)**.** (1) $B0(i,x) := d(i,x) = 0$ *if* $x_i = x_{i+1}$.
(2) $B1(i,x) := d(i,x) + d(i, s_i x) = 0$

Definition 17.5.4. For any $i, j \leq n-1$ with $|i-j| \geq 2$ we define the *braid element* $B2(i,j,x)$ by

$$B2(i,j,x) := d(i,x) + d(j, s_i x) + d(i, s_j s_i x) + d(j, s_j x)$$

Lemma 17.5.5 (Second braid relation)**.** $B2(i,j,x) \in J^{n-1}$.

Proof. If we expand $B2(i,j,x)$, the degree n parts cancel and we get:

$$B2(i,j,x) = -\partial_i x - \partial_j s_i x - \partial_i s_j s_i x - \partial_j s_j x \quad \in \mathcal{T}^{n-1}$$

If $i < j$ then $\partial_i \partial_j x = \partial_{j-1} \partial_i x$. So, the difference between the two generators

$$d(i, \partial_j x) = \partial_j x - \partial_j s_i x - \partial_i \partial_j x$$

$$d(j-1, \partial_i x) = \partial_i x - \partial_i s_j x - \partial_{j-1} \partial_i x$$

of J^{n-1} is

$$d(i, \partial_j x) - d(j-1, \partial_i x) = \partial_j x - \partial_j s_i x - \partial_i x + \partial_i s_j x$$

which is equal to $B2(i,j,x)$ by the first braid relation. □

Definition 17.5.6. For all $1 \leq i \leq n-2$ we define the *braid element* $B3(i,x)$ by

$$B3(i,x) := d(i,x) + d(i+1, s_i x) + d(i, s_{i+1} s_i x)$$

$$+ d(i+1, s_i s_{i+1} s_i x) + d(i, s_i s_{i+1} x) + d(i+1, s_{i+1} x)$$

Lemma 17.5.7 (Third braid relation)**.** $B3(i,x) \in J^{n-1}$.

Proof. Expanding $B3(i,x)$ we get

$$B3(i,x) = -\partial_i x - \partial_{i+1} s_i x - \partial_i s_{i+1} s_i x - \partial_{i+1} s_i s_{i+1} s_i x - \partial_i s_i s_{i+1} x - \partial_{i+1} s_{i+1} x$$

Compare this with the Jacobi relation $[[xy]z] + [[yz]x] + [[zx]y]$ with $x, y, z = x_i, x_{i+1}, x_{i+2}$ which gives the following identity in $\mathcal{T}^{n-2}$.

$$\partial_i \partial_i (x + s_i s_{i+1} x + s_{i+1} s_i x) = 0$$

In terms of the generators of J^{n-1}, this identity gives the following.

$$d(i, \partial_i x) + d(i, \partial_i s_i s_{i+1} x) + d(i, \partial_i s_{i+1} s_i x) =$$

$$\partial_i x - s_i \partial_i x + \partial_i s_i s_{i+1} x - s_i \partial_i s_i s_{i+1} x + \partial_i s_{i+1} s_i x - s_i \partial_i s_{i+1} s_i x$$

Using the identity $s_i \partial_i = -\partial_{i+1} s_i s_{i+1} s_i = -\partial_{i+1} s_{i+1} s_i s_{i+1}$, this becomes

$$\partial_i x + \partial_{i+1} s_i s_{i+1} s_i x + \partial_i s_i s_{i+1} x + \partial_{i+1} s_i x + \partial_i s_{i+1} s_i x + \partial_{i+1} s_{i+1} x$$

which is equal to $-B3(i,x)$. □

17.5.3. *general elements of J^n.* We use the following elementary theorem without proof.

Theorem 17.5.8. *The permutation group on n letters $1, 2, \cdots, n$ is generated by the $n-1$ simple reflections $s_i = (i, i+1)$ for $i = 1, \cdots, n-1$ modulo the following relations.*

(1) $s_i^2 = e$
(2) $s_j s_i s_j s_i = e$ *if* $|i-j| \geq 2$
(3) $s_{i+1} s_i s_{i+1} s_i s_{i+1} s_i = e$

Whenever a group G is given by generators and relations, we can form an exact sequence

$$C_2(G) \xrightarrow{d_2} C_1(G) \xrightarrow{d_1} FG \xrightarrow{\epsilon} F \to 0$$

where $C_1(G), C_2(G)$ are right FG-modules freely generated by the set of generators and relations respectively and ϵ, d_1, d_2 are given on F-basis elements as follows.

(1) $\epsilon(g) = 1$ for all $g \in G$.
(2) $d_1([x]g) = g - xg$ for all generators x of G and all $g \in G$.
(3) $d_2([y]g) = [y_1]g + [y_2]y_1 g + \cdots + [y_r]y_{r-1} \cdots y_1 g$ if $y = y_r \cdots y_1$ is a relation and $g \in G$ (with the convention that $[y^{-1}]g = -[y]y^{-1}g$.)

In the case at hand, the generators of S_n are the simple transpositions s_i and the homomorphism $d_1 : C_1(S_n) \to FS_n$ is given by

$$d_1([s_i]\sigma) = \sigma - s_i\sigma$$

Corollary 17.5.9. *The kernel of d_1 is spanned by elements of the following form.*

R1 $[s_i]\sigma + [s_i]s_i\sigma$
R2 $[s_i]\sigma + [s_j]s_i\sigma + [s_i]s_j s_i\sigma + [s_j]s_i s_j s_i\sigma$ *for* $|i-j| \geq 2$
R3 $[s_i]\sigma + [s_{i+1}]s_i\sigma + [s_i]s_{i+1}s_i\sigma + [s_{i+1}]s_i s_{i+1} s_i\sigma + [s_i]s_i s_{i+1}\sigma + [s_{i+1}]s_{i+1}\sigma$

Recall that B is a vector space basis for L. For any $x_1, \cdots, x_n \in B$ consider $x = x_1 \otimes \cdots \otimes x_n \in \mathcal{T}^n(L)$. Let

$$\psi_x : C_1(S_n) \to \mathcal{T}_n$$

be the linear map given by

$$\psi_x([s_i]\sigma) = d(i, \sigma x) = \sigma x - s_i\sigma x - \partial_i \sigma x \in J^n$$

Lemma 17.5.10. *The linear map ψ_x sends the generators $R1, R2, R3$ of $\ker d_1$ to braid elements in J^{n-1}.*

Proof. This follows directly from the definitions. □

The image of ψ_x is the component of J^n corresponding to the monomial x. So, to prove the Key Lemma it suffices to show that the intersection of the image of ψ_x with $\mathcal{T}_{n-1}$ is spanned by braid elements for all x.

However, ψ_x sends $z \in C_1(S_n)$ into $\mathcal{T}_{n-1}$ iff z lies in the kernel of the composite map

$$C_1(S_n) \xrightarrow{\psi_x} \mathcal{T}_n \twoheadrightarrow \mathcal{T}_n/\mathcal{T}_{n-1} = \mathcal{T}^n$$

This composition $\overline{\psi}_x$ is given on generators by

$$\overline{\psi}_x([s_i]\sigma) = \sigma x - s_i\sigma x$$

Thus, it suffices to show that the kernel of $\overline{\psi}_x$ is generated by elements which map to braid elements under ψ_x.

Proposition 17.5.11. *The kernel of* $\overline{\psi}_x$ *is spanned by elements of the form*

R0a $[s_1]\sigma$ *where* $x_{\sigma^{-1}(1)} = x_{\sigma^{-1}(2)}$

R0b $[s_i]\sigma - [s_i]\sigma'$ *if* $\sigma x = \sigma' x$

and those of the form $R1, R2, R3$ *as given in the previous corollary.*

First we point out that this implies the Key Lemma. The reason is that the element in R0a, R0b lie in the kernel of ψ_x and the other elements map to braid elements by the lemma above. Therefore, all elements of $J^n \cap \mathcal{T}_{n-1}$ are linear combinations of braid elements which we have shown to lie in J^{n-1} in the previous subsection.

Proof. The map $\overline{\psi}_x$ factors as a composition of $d_1 : C_1(S_n) \to FS_n$ with the map $\varphi_x : FS_n \to \mathcal{T}^n$ given by $\sigma \mapsto \sigma x$. The kernel of φ_x is spanned by elements of the form $\sigma - \tau\sigma$ where τ is a transposition which switches two equal coordinates of σx. Any element in the kernel of $\overline{\psi}_x$ will map, under d_1, to a linear combination of elements of FS_n of the form $\sigma - \tau\sigma$. Therefore, to prove the proposition, it suffices to show that there is a linear combination of the elements R0a, R0b which maps onto the element $\sigma - \tau\sigma$.

Since any transposition is a conjugate of s_1, $\tau = \rho s_1 \rho^{-1}$. For simplicity of notation, we assume that $\rho = abc$ where a, b, c are simple transpositions. Then $\tau = abcs_1cba$ and

$$\sigma - \tau\sigma = (\sigma - a\sigma) + (a\sigma - ba\sigma) + \cdots + (bcs_1cba\sigma - \tau\sigma)$$

is the image under $\psi_x : C_1(S_n) \to FS_n$ of the sum

$$= [a]\sigma + [b]a\sigma + [c]ba\sigma + [s_1]cba\sigma + [c]s_1cba\sigma + [b]cs_1cba\sigma + [a]bcs_1cba\sigma$$

But $[s_1]cba\sigma$ is an element of type R0a and the other six elements form a sum of three elements of type R0b (and elements of type R1,R2,R3 in the kernel of ψ_x) since $s_1cba\sigma x = cba\sigma x$. For example,

$$[b]cs_1cba\sigma - [b]ba\sigma \in R0b$$
$$[b]ba\sigma + [b]a\sigma \in R1$$

adding up to the two terms with $[b]$. □

We will now skip sections 18 and 19 to concentrate on Representation Theory.

Representation Theory

Throughout the rest of these notes L will be a finite dimensional semisimple Lie algebra over $F = \mathbb{C}$ with CSA H, root system Φ, base Δ and Weyl group W. Although L will be finite dimensional, we need to consider infinite dimensional representations V of L. The main goal will be to explain the Weyl character formula. The proof will come afterwards.

20. Weights and maximal vectors

The statement is: Irreducible representations V of L are uniquely determined up to isomorphism by their highest weight and are generated by any vector of highest weight. This is true when V is finite dimensional and is also true for many infinite dimensional V. The main problem is that an infinite dimensional representation may not have a highest weight.

20.1. **definitions.** Recall that L has a root space decomposition

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

For any representation V of L and any $\lambda : H \to F = \mathbb{C}$ recall that the λ weight space of V is:

$$V_\lambda = \{v \in V \mid h(v) = \lambda(h)v\}$$

Let V' be the sum of all the weight spaces V_λ.

Proposition 20.1.1. (1)

$$V' = \bigoplus_{\lambda} V_\lambda$$

(2) $L_\alpha(V_\lambda) \subseteq V_{\lambda+\alpha}$.
(3) $V' = V$ *if* V *is finite dimensional.*

Definition 20.1.2. A *highest weight* for V is a weight λ so that $V_\lambda \neq 0$ but $V_{\lambda+\alpha} = 0$ for all $\alpha \in \Phi_+$.

It is clear that any (nonzero) finite dimensional representation has a highest weight.

Example 20.1.3. For the adjoint representation $V = L$, the highest weight is equal to the maximal root.

Definition 20.1.4. A *maximal vector* $v^+ \in V$ is a nonzero element with the property that

$$x_\alpha v^+ = 0$$

for all $x_\alpha \in L_\alpha$ where α is a positive root.

It is clear that any nonzero vector of highest weight is a maximal vector. The converse is not true.

It is enough to have $x_\alpha v^+ - 0$ for $\alpha \in \Delta$.

Example 20.1.5. Let $L = \mathfrak{sl}(2,F) = H \oplus L_\alpha \oplus L_{-\alpha}$. Recall that $H = Fh_\alpha, L_\alpha = Fx_\alpha, L_{-\alpha} = Fy_\alpha$. Since there is only one positive root α, a maximal weight in a representation V is any nonzero $v \in V$ so that $x_\alpha(v) = 0$.

Let $V = \mathfrak{sl}(3,F)$ with positive roots $\alpha, \beta, \alpha+\beta$. The weight space decomposition of V is

$$V = V_\alpha \oplus V_{\frac{1}{2}\alpha} \oplus V_0 \oplus V_{-\frac{1}{2}\alpha} \oplus V_{-\alpha}$$

Identifying $\alpha = 2$ since H^* is one dimensional and $\alpha(h_\alpha) = 2$, this can be rewritten:

$$V = V_2 \oplus V_1 \oplus V_0 \oplus V_{-1} \oplus V_{-2}$$

- The vector $x_\alpha \in V_2$ is a maximal vector since it has highest weight.
- The vector $x_{\alpha+\beta} \in V$ is maximal since $[x_\alpha, x_{\alpha+\beta}] = 0$. It also lies in V_1:

  $$h_\alpha(x_{\alpha+\beta}) = (\alpha(h_\alpha) + \beta(h_\alpha))x_{\alpha+\beta} = (2-1)x_{\alpha+\beta} = x_{\alpha+\beta}$$

  so it has highest weight since $V_{\frac{1}{2}\alpha+\alpha} = V_3 = 0$.
- The vector $h_\alpha + 2h_\beta \in V_0$ is also a maximal vector since

  $$x_\alpha(h_\alpha + 2h_\beta) = -[h_\alpha + 2h_\beta, x_\alpha] = -(\alpha(h_\alpha) + 2\beta(h_\alpha))x_\alpha = -(2-2)x_\alpha = 0$$

  but $h_\beta \in V_0$ so it does not have highest weight.

Note that, in this example, V has two highest weights.

20.2. **Standard cyclic modules.**

Definition 20.2.1. A *standard cyclic module* of highest weight λ is a representation V which is generated by a single maximal vector v^+ of weight λ.

This means that V is spanned by elements of the form $a_1 a_2 \cdots a_m v^+$ where $a_i \in L$. I.e., $V = \mathcal{U}(L)v^+$. The fact that the finite dimensional Lie algebra L can have infinite dimensional cyclic modules comes from the fact that $\mathcal{U}(L)$ is infinite dimensional in general.

Lemma 20.2.2. *Let V be a standard cyclic module generated by $v^+ \in V_\lambda$. Then V is spanned by elements of the form*

$$y_{\beta_1} y_{\beta_2} \cdots y_{\beta_k} v^+$$

where β_i are positive roots and $y_\beta \in L_{-\beta}$.

Proof. Use PBW to see that $\mathcal{U}(L)v^+ = \mathcal{U}(N_-(L))\mathcal{U}(B(\Delta))v^+ = \mathcal{U}(N_-(L))v^+$ □

Theorem 20.2.3. *If V is standard cyclic as above then*

1. *λ is a highest weight.*
2. *V_λ is one dimensional.*
3. *V has a weight space decomposition $V = \bigoplus V_\beta$ where β runs over weights of the form $\lambda - \sum k_i \alpha_i$ where $\alpha_i \in \Delta$ and k_i are nonnegative integers.*

In the proof of the corollary below we used the following lemma.

Lemma 20.2.4. *$v^+ \in V$ is a maximal vector iff it satisfies the condition:*

$$Bv^+ = \mathbb{C}v^+$$

In other words, v^+ is a common eigenvector for all elements of the Borel subalgebra $B = B(\Delta)$.

Proof. Let $W = \mathbb{C}v^+$. Then W is a representation of B and therefore also of $H \subseteq B$. So, v^+ is an eigenvector of H and we have a linear map $\lambda : H \to \mathbb{C}$ given by $\lambda(h)v^+ = h(v^+)$. Thus $W = W_\lambda$. For any positive root α we have $x_\alpha \in B$ and $x_a(v^+) \subseteq W_{\lambda+\alpha} = 0$. So, v^+ is a maximal vector of weight λ. The converse is obvious. □

Corollary 20.2.5. *V is indecomposable and all quotient modules are cyclic with highest weight λ. V has a unique maximal proper submodule. If V is irreducible then λ is unique.*

Proof. Suppose that $V = V_1 \oplus V_2$. Then each element of V has two coordinates. So, $v^+ = (v_1^+, v_2^+)$. For every $b \in B$ we have $bv^+ = av^+$ for some $a \in \mathbb{C}$. But $av^+ = (av_1^+, av_2^+)$. So, $Bv_1^+ = \mathbb{C}v_1^+$ and $Bv_2^+ = \mathbb{C}v_2^+$. Therefore, $(v_1^+, 0)$ and $(0, v_2^+)$ are maximal vectors of weight λ. But V_λ is one-dimensional. So, either $v_1^+ = 0$ or $v_2^+ = 0$. Since v^+ generates V, v_i^+ generates V_i. So, either $V_1 = 0$ or $V_2 = 0$ showing that V is indecomposable.

Given any submodule W of V, since W is an H-submodule of V, it must be the sum of weight spaces W_μ. Since $W \neq V$, we must have $W_\lambda = 0$. So, $(V/W)_\lambda = V_\lambda/W_\lambda = V_\lambda \neq 0$. So, $v^+ + W$ is a nonzero maximal vector for V/W of weight λ and it clearly generates V/W. So, V/W is cyclic.

To show that there is a unique maximal proper submodule, note that all proper submodules of V lie in the vector subspace $\bigoplus_{\mu\neq\lambda} V_\mu$. But then the sum of all proper submodules of V is a proper submodule which is unique since it contains all other proper submodules.

Finally, if V is irreducible then λ is uniquely determined since, given any other maximal vector $w^+ \in V_\mu$, the submodule generated by w^+ must be equal to V. But then $\lambda = \mu - \sum k_i\alpha_i$ which implies that $\mu = \lambda + \sum k_i\alpha_i$ which implies that $\lambda = \mu$. □

20.3. **Existence and uniqueness of cyclic modules.** I proved the existence theorem first:

Theorem 20.3.1. *For any $\lambda : H \to \mathbb{C}$, there exists an irreducible standard cyclic module with highest weight λ.*

Proof. Start with a one dimensional representation $D_\lambda = \mathbb{C}v^+$ of B given by taking the action of any $h \in H$ to be multiplication by $\lambda(h)$ and the action of any $x_\alpha \in L_\alpha$ to be zero. Then take:

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

This is the L-module obtained from D_λ by "extension of scalars" which is also called the "induced representation." (Recall that for any homomorphism of rings $R \to S$ and any S-module M we have an R-module given by $R \otimes_S M$.)

The L-module V is generated by the element $1 \otimes v^+$ which is a maximal vector of weight λ since $b(1 \otimes v^+) = 1 \otimes bv^+$ is a scalar multiple of $1 \otimes v^+$ and that scalar is equal to $\lambda(h)$ when $b = h \in H$.

By the corollary, V has a unique maximal proper submodule M and the quotient V/M is the desired irreducible cyclic module with prescribed highest weight λ. □

Theorem 20.3.2. *There is only one irreducible V with highest weight λ (up to isomorphism).*

Proof. Suppose there are two of irreducible standard cyclic modules V^1, V^2 with the same highest weight λ. Then $V^1_\lambda = \mathbb{C}v_1$ and $V^2_\lambda = \mathbb{C}v_2$. Let $V = V^1 \oplus V^2$. Then $V_\lambda = V^1_\lambda \oplus V^2_\lambda$. So, $v^+ = (v_1^+, v_2^+)$ is a maximal vector since, for all $b \in B$ we have $bv^+ = (bv_1^+, bv_2^+) = (av_1^+, av_2^+) = av^+$ for some scalar a. (Since a is uniquely determined by b and λ, it is the same scalar in both coordinates.)

Let W be the cyclic module generated by v^+. Then the projection map $p_1 : W \to V_1$ is onto since it sends v^+ to the generator v_1^+ of V_1. Since V_1 is irreducible, the kernel of p_1 is the unique maximal proper submodule M of W. So, $V_1 \cong W/M$. Similarly, $V_2 \cong W/M$. So, $V_1 \cong V_2$. Furthermore, this isomorphism sends v_1^+ to v_2^+. □

21. Finite dimensional modules

Given any weight $\lambda : H \to \mathbb{C}$ we have the cyclic module

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

This is called the *Verma module* of highest weight λ. This module has a unique maximal proper submodule and the quotient $V(\lambda)$ is the unique irreducible module with highest weight λ. In this section we determine when $V(\lambda)$ is finite dimensional. If we recall Weyl's Theorem (Every finite dimensional representation of a semisimple Lie algebra is a direct sum of irreducible modules.) this will give a complete classification of all finite dimensional representations of semisimple Lie algebras.

The statement is:

Theorem 21.0.3. *$V(\lambda)$ is finite dimensional iff λ is a dominant weight, i.e., $\lambda(h_\alpha) = \langle \lambda, \alpha \rangle$ is a nonnegative integer for all positive roots α.*

The theory of dominant weights was done abstractly in Section 13 which we skipped. Now we need some of the basic concepts from that section.

21.1. **Dominant weights.** Suppose that $\alpha_1, \cdots, \alpha_n$ are the positive simple roots (the elements of the base Δ). For each i we have a copy $S_i = S_{\alpha_i}$ of $\mathfrak{sl}(2, F)$ with basis $x_i, y_i, h_i = h_{\alpha_i} \in H$. Then the h_i form a vector space basis for H.

Definition 21.1.1. An *abstract or integral weight* is a linear function $\lambda : H \to \mathbb{C}$ with the property that $\lambda(h_i) \in \mathbb{Z}$ for all i. An abstract weight is called *dominant* if $\lambda(h_i) \geq 0$ for all i. The *fundamental dominant weights* λ_i are the ones given by:

$$\lambda_i(h_j) = \delta_{ij}$$

I.e., these form the dual basis for the basis of H given by the h_i.

Example 21.1.2. For $L = \mathfrak{sl}(2, \mathbb{C})$ there is only one fundamental weight $\lambda_1 = \frac{1}{2}\alpha$:

$$\lambda_1(h_1) = \frac{1}{2}\alpha(h_\alpha) = \frac{2}{2} = 1$$

Exercise 21.1.3. Show that for $L = \mathfrak{sl}(3, \mathbb{C})$, the fundamental weights are

$$\lambda_1 = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2, \qquad \lambda_2 = \frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2$$

It is clear that all dominant weights are given by addition of the fundamental dominant weights:

$$\lambda = \sum n_i \lambda_i$$

where n_i are nonnegative integers.

The set of dominant weights is denoted Λ^+. A weight $\lambda = \sum n_i \lambda_i$ is called *strongly dominant* if $n_i > 0$ for all i. One important example is the minimal strongly dominant weight given by

$$\delta = \sum \lambda_i$$

This is characterized in several ways:

(1) $\delta(h_i) = 1$ for all i.
(2)
$$\delta = \frac{1}{2}\sum_{\alpha\in\Phi_+}\alpha$$

To prove the last equation we use the action of the Weyl group W. Let $\mu = \frac{1}{2}\sum\alpha$. Apply the simple reflection s_i given by

$$s_i(x) = x - \langle x, \alpha_i\rangle\,\alpha_i$$

We know that s_i sends α_i to $-\alpha_i$ and permutes the other positive roots. So:

$$s_i(\mu) = \mu - \alpha_i = \mu - \langle \mu, \alpha_i\rangle\,\alpha_i$$

Therefore, $\langle \mu, \alpha_i\rangle = \mu(h_i) = 1$ for all i. So, $\mu = \delta$.

21.2. **finite irreducible modules.**

Theorem 21.2.1. *If $V(\lambda)$ is finite dimensional then*

(1) *λ is a dominant weight.*
(2) *All weights μ of $V(\lambda)$ are integral weights and therefore given as integer linear combinations of the fundamental weights:*
$$\mu = \sum n_i\lambda_i$$
(3) *The set Π of weights μ which occur in $V(\lambda)$ is saturated (defined below).*

A set Π of integral weight $\mu = \sum n_i\lambda_i$ is *saturated* if, for all $\beta\in\Phi$ and all integers $0\le m\le\langle\mu,\beta\rangle$, $\mu - m\beta\in\Pi$. Since every root is a sum of simple roots, it is enough to have this for $\beta = \alpha_i$ in which case $0\le m\le n_i = \langle\mu,\alpha_i\rangle$.

Proof. We view $V(\lambda)$ as a representation of S_i and quote results from Section 7. Each weight of $V(\lambda)_\mu$ becomes $\mu(h_i) = n_i$. Therefore, n_i must be an integer. Thus all weights μ are integral. For the highest weight λ, $\lambda(h_i)$ must be a nonnegative integer. So, λ is dominant.

The action of y_i on $V(\lambda)$ sends $V(\lambda)_\mu$ to $V(\lambda)_{\mu-\alpha_i}$ and, by symmetry of the weights of representations of S_i around 0,

$$y_i^{n_i}(w) \neq 0 \in V(\lambda)_{\mu - n_i\alpha_i}$$

for all $w\neq 0\in V(\lambda)_\mu$. So, Π is saturated. □

Theorem 21.2.2. *If λ is any dominant weight then $V(\lambda)$ is finite dimensional. Furthermore, the set Π of weights μ is invariant under the action of the Weyl group and is minimal W-invariant saturated set of integral weights which contains λ.*

Proof. You can read the proof in the book. Here is an outline.

(1) For each i the sequence of elements
$$v^+, y_iv^+, y_i^2v^+, \cdots, y_i^kv^+$$
for $k = \lambda(h_i)$ forms a finite dimensional S_i submodule of $V(\lambda)$.

(2) Let V' be the sum of all finite dimensional S_i submodules of $V(\lambda)$. Then V' is a a nonzero L-submodule and therefore $V' = V$.
(3) Recall that $\tau_i = \exp(x_i)\exp(-y_i)\exp(x_i)$ is an automorphism of $V(\lambda)$ which lifts the action of the simple reflection σ_i. Thus $\tau_i V(\lambda)_\mu = V(\lambda)_{\sigma_i \mu}$. (Since $V = V'$, $\tau_i V = \tau_i V' = V' = V$.)
(4) All weights are integral by the Key Lemma 20.2.2 we proved last time. Also $V(\lambda)_\mu$ is finite dimensional for all μ.
(5) The symmetry of Π under the Weyl group forces it to be finite.
(6) If Π' is the minimal W-invariant saturated subset of Π then $\bigoplus_{\mu\in\Pi'} V(\lambda)_\mu$ is a submodule of $V(\lambda)$ and therefore the whole thing.

□

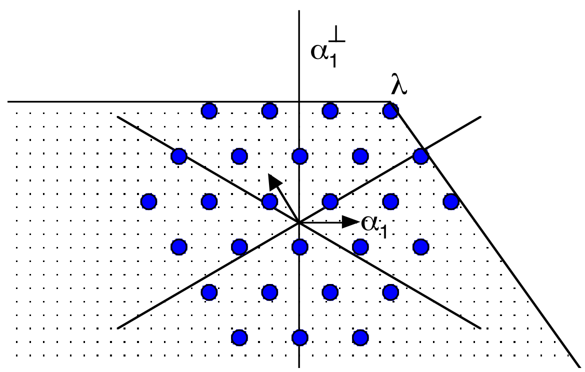


FIGURE 21.2.1. Since λ is the highest weight, Π is confined to the shaded region. Since Π is W invariant it is confined to the blue dots.

I pointed out at the end of the class that, in order to prove that $V(\lambda)$ is finite dimensional it suffices to construct a finite dimensional cyclic module of highest weight λ for any dominant weight λ since $V(\lambda)$ is uniquely determined by λ.

Lemma 21.2.3. *$V(\lambda) \otimes V(\mu)$ contains a maximal vector of highest weight $\lambda + \mu$.*

Proof. Let $v = v_1^+ \otimes v_2^+ \in V(\lambda) \otimes V(\mu)$. Then, for any $h \in H$ we have

$$h(v) = h(v_1^+) \otimes v_2^+ + v_1^+ \otimes h(v_2^+) = \lambda(h)v_1^+ \otimes v_2^+ + v_1^+ \otimes \mu(h)v_2^+ = (\lambda+\mu)(h)v_1^+ \otimes v_2^+$$

Therefore $v = v_1^+ \otimes v_2^+$ has weight $\lambda + \mu$. Also, for any x_α for positive root α we have

$$x_\alpha(v_1^+ \otimes v_2^+) = x_\alpha(v_1^+) \otimes v_2^+ + v_1^+ \otimes x_\alpha(v_2^+) = 0 + 0 = 0$$

So, $v = v_1^+ \otimes v_2^+$ is a maximal vector. □

This elementary lemma implies that, to prove Theorem 21.2.2 it suffices to construct a finite dimensional module containing $V(\lambda_i)$ for the fundamental dominant weights λ_i. We will do this later at least in some case.

22. Formal characters

The main purpose of this section is to set up the notation in order to state the Weyl character formula. The details of the formula, examples and proofs will be given in the remaining lectures of this course.

22.1. **Definition.** (Section 22.5 in book)

Let $\Lambda \subseteq H^*$ be the set of integral weights. This is a free abelian group generated by the fundamental dominant weights λ_i.

$$\Lambda \cong \mathbb{Z}^n$$

If λ is a dominant weight then we recall that $V(\lambda)$ is finite dimensional and the set Π of all weights μ of $V(\lambda)$ is a subset of Λ. We also recall one of the key steps in the proof of these facts:

Proposition 22.1.1. *The action of Weyl group W on H^* leaves Λ invariant and fixes the representation $V(\lambda)$ ($\sigma V(\lambda) \cong V(\lambda)$) and therefore leaves Π invariant.*

We will give several formulas for the dimension of $V(\lambda)_\mu$ which we denote

$$m_\lambda(\mu) := \dim V(\lambda)_\mu$$

Let $\mathbb{Z}[\Lambda]$ be the integer group ring of the group Λ. Additively, this is the free abelian group generated by the elements of Λ which we now need to write multiplicatively. Thus $e(\mu)$ denotes the element of $\mathbb{Z}[\Lambda]$ corresponding to $\mu \in \Lambda$ and

$$e(\lambda + \mu) = e(\lambda)e(\mu)$$

Elements of $\mathbb{Z}[\Lambda]$ are finite formal linear combinations

$$\sum n_i e(\mu_i)$$

The book also writes this as

$$\sum f(\mu)e(\mu)$$

where $f : \Lambda \to \mathbb{Z}$ is a set mapping with finite support (the *support* of a function is the subset of the domain on which it is nonzero). Thus $f(\mu) = 0$ for all but a finite number of μ.

Definition 22.1.2. For any finite dimensional representation V of a semisimple Lie algebra L we define the *(formal) character* ch_V of V to be the element of $\mathbb{Z}[\Lambda]$ given by

$$\mathrm{ch}_V = \sum_\mu \dim V_\mu e(\mu)$$

When $V = V(\lambda)$ we use the notation $\mathrm{ch}_\lambda = \mathrm{ch}_{V(\lambda)}$. Thus

$$\mathrm{ch}_\lambda = \sum_\mu m_\lambda(\mu)e(\mu)$$

22.2. **Basic properties.** The basic properties of the formal character are easy to prove.

Proposition 22.2.1. *The formal character is additive:*

$$\mathrm{ch}_{V\oplus W} = \mathrm{ch}_V + \mathrm{ch}_W$$

Proof. $(V \oplus W)_\mu = V_\mu \oplus W_\mu$. So, $\dim(V \oplus W)_\mu = \dim V_\mu + \dim W_\mu$. □

Proposition 22.2.2. *The formal character is multiplicative:*

$$\mathrm{ch}_{V\otimes W} = \mathrm{ch}_V \mathrm{ch}_W$$

Proof. As I pointed out last time:

$$V_\lambda \otimes W_\mu \subseteq (V \otimes W)_{\lambda+\mu}$$

Therefore,

$$(V \otimes W)_\lambda = \bigoplus_{\mu+\nu=\lambda} V_\mu \otimes W_\nu$$

Taking dimensions, we get:

$$m_{V\otimes W}(\lambda) = \sum_{\mu+\nu=\lambda} m_V(\mu) m_W(\nu)$$

which implies $\mathrm{ch}_{V\otimes W} = \mathrm{ch}_V \mathrm{ch}_W$ by definition of multiplication of elements of $\mathbb{Z}[\Lambda]$. □

Proposition 22.2.3. *The character of any finite dimensional representation V is fixed under the action of the Weyl group:*

$$\sigma \mathrm{ch}_V = \mathrm{ch}_{\sigma V} = \mathrm{ch}_V$$

for any $\sigma \in W$.

Proof. Since V is finite dimensional, it is a direct sum of $V(\lambda)$ where λ is dominant. But then $\sigma V = \bigoplus \sigma V(\lambda) \cong \bigoplus V(\lambda) = V$ for all $\sigma \in W$ since $\sigma V(\lambda) \cong V(\lambda)$. □

Lemma 22.2.4. *$\lambda \in \Lambda$ is dominant iff it lies in the fundamental Weyl chamber. In particular, every integral weight μ is in the W-orbit of a dominant weight.*

Proof. This follows from the definition since $\lambda(h_i) = \langle \lambda, \alpha_i \rangle$ is ≥ 0 for all i iff λ lies on the nonnegative side of the hyperplane perpendicular to each simple root α_i. □

Proposition 22.2.5. *Any element of $\mathbb{Z}[\Lambda]$ which is fixed by the action of W can be expressed uniquely as an integer linear combination of ch_λ where $\lambda \in \Lambda^+$.*

Proof. Take any element $f \in \mathbb{Z}[\Lambda]$ which is invariant under the action of W. View f as a function $f : \Lambda \to \mathbb{Z}$ with finite support. Let M_f be the set of all dominant weights μ so that $f(\lambda) \neq 0$ for a dominant weight $\lambda \geq \mu$. We will show by induction on the size of M_f that f is an integer linear combination of the ch_λ. If M_f is empty then $f = 0$ by the lemma. So, suppose that M_f is nonempty and we know the existence statement for all smaller M_f. Let λ be a maximal element of M_f. Then $f(\lambda) = a \in \mathbb{Z}, a \neq 0$. Let $g = f - a\mathrm{ch}_\lambda$ then $M_g \subset M_f$ but $\lambda \notin M_g$. Therefore, g is an integer linear combination of ch_μ's. So, $f = g + \mathrm{ch}_\lambda$.

Uniqueness is easy. Suppose that we have two expressions for f as an integer linear combination of ch_λ's. Then the difference is a linear combination which gives 0. Putting all positive terms on one side of the equation and negative terms on the other, this can be written as a equality $\mathrm{ch}_V = \mathrm{ch}_W$ between two representations V, W. Take $\lambda \in \Lambda^+$ maximal so that λ occurs as a weight of V and W. Then $V(\lambda)$ must be a direct summand of both V and W and by induction on their size we get $V \cong W$. So, the expression is unique. □

22.3. **Representation ring.** The properties of the formal character ch_V can be summarized by saying that it gives an isomorphism

$$\mathrm{ch} : Rep(L) \cong \mathbb{Z}[\Lambda]^W$$

between the representation ring of L and the ring of W-invariant elements of the integer group ring of Λ.

Definition 22.3.1. The *representation ring* $Rep(L)$ of L is defined as follows. As an additive group, $Rep(\Lambda)$ is the free additive group generated by isomorphism classes $[V]$ of finite dimensional representations V of L modulo the relation:

$$[V] + [W] = [V \oplus W]$$

The multiplication on $Rep(L)$ is given by:

$$[V] \cdot [W] = [V \otimes W]$$

The additive and multiplicative properties of ch implty that $\mathrm{ch} : Rep(L) \to \mathbb{Z}[\Lambda]$ is a ring homomorphism. (And ch sends the unit $\mathbb{C}$ of $Rep(L)$ to the unit $1 = e(0)$ of $\mathbb{Z}[\Lambda]$ implying that ch is unital.) The W-invariance property of ch_V imply that the image of ch is contained in $\mathbb{Z}[\Lambda]^W$ and the last proposition implies that ch is a monomorphism with image $\mathbb{Z}[\Lambda]^W$.

Example 22.3.2. In the case $L = \mathfrak{sl}(2, F)$, the irreducible representations are $V(m), m \geq 0$ with highest weight m. So, the group ring $\mathbb{Z}[\Lambda]$ is isomorphic to $\mathbb{Z}[t, t^{-1}]$ with $t^m = e(m)$. The nontrivial element of the Weyl group $\mathbb{Z}/2$ acts by $t \leftrightarrow t^{-1}$. So, the W-invariant subring consists of all linear combinations of the form:

$$a_0 + a_1(t + t^{-1}) + a_2(t^2 + t^{-2}) + \cdots + a_n(t^n + t^{-n})$$

with multiplication of the generators $s_n = t^n + t^{-n}$ given by

$$s_n s_m = (t^n + t^{-n})(t^m + t^{-m}) = t^{n+m} + t^{n-m} + t^{m-n} + t^{-n-m} = s_{n+m} + s_{|n-m|}$$

The representation ring of $\mathfrak{sl}(2, F)$ is the additive group of all linear combinations

$$b_0[V(0)] + b_1[V(1)] + b_2[V(2)] + \cdots + b_n[V(n)]$$

with integer coefficients b_i with multiplicative structure to be explained. The weight space decomposition of $V(m)$ gives

$$\mathrm{ch}_{V(m)} = t^m + t^{m-2} + t^{m-4} + \cdots + t^{-m} = \frac{t^{m+1} - t^{-m-1}}{t - t^{-1}}$$

Using the fact that ch is a ring isomorphism we get:

$$\begin{aligned}\mathrm{ch}_{V(n)\otimes V(m)} = \mathrm{ch}_{V(n)}\mathrm{ch}_{V(m)} &= \frac{(t^{n+1}-t^{-n-1})(t^{m+1}-t^{-m-1})}{(t-t^{-1})^2}\\ &= \frac{t^{m+n+2}-t^{m-n}+t^{-m-n-2}-t^{n-m}}{(t-t^{-1})^2}\end{aligned}$$

This implies the following.

Theorem 22.3.3 (Clebsch-Gordon)**.** *If* $m \geq n$ *then*

$$V(n)\otimes V(m) \cong V(n+m)+V(n+m-2)+V(n+m-4)+\cdots+V(m-n)$$

Proof. The formal character of the right hand side is:

$$\frac{t^{m+n+1}-t^{-m-n-1}}{t-t^{-1}}+\frac{t^{m+n-1}-t^{-m-n+1}}{t-t^{-1}}+\cdots+\frac{t^{m-n+1}-t^{-m+n-1}}{t-t^{-1}}$$

which is equal to the formal character of $V(n)\otimes V(m)$ given above. □

22.4. **Weyl character formula.** Recall that $\delta \in \Lambda$ is the integral weight given by

$$\delta = \sum \lambda_i = \frac{1}{2}\sum_{\alpha\in\Phi_+}\alpha$$

Definition 22.4.1.

$$q = \sum_{\sigma\in W}\mathrm{sgn}(\sigma)e(\sigma\delta) \in \mathbb{Z}[\Lambda]$$

where $\mathrm{sgn}(\sigma) = (-1)^\ell$ where $\ell = \ell(\sigma)$ is the length of σ (the minimum number of reflections that gives σ. (This is also $\mathrm{sgn}(\sigma) = \det(\sigma)$ since each reflection has det $= -1$.)

Example 22.4.2. For $L = \mathfrak{sl}(2,F)$, $\delta = \lambda_1 = \frac{1}{2}\alpha_1$ is equal to $t = t^1$ in the notation above. The nontrivial element of W is a reflection with sign -1. So, the element q is equal to

$$q = t - t^{-1}$$

Definition 22.4.3. For any integral weight μ let

$$\omega(\mu) = \sum_{\sigma\in W}\mathrm{sgn}(\sigma)e(\sigma\mu) \in \mathbb{Z}[\Lambda]$$

Example 22.4.4. For $L = \mathfrak{sl}(2,F)$, $e(m) = t^m$ and $\omega(m) = t^m - t^{-m}$. In particular, $q = \omega(\delta) = \omega(1) = t - t^{-1}$.

Theorem 22.4.5 (Weyl Character Formula)**.** *For any dominant weight* $\lambda \in \Lambda^+$*, the formal character of* $V(\lambda)$ *is given by*

$$\boxed{\mathrm{ch}_\lambda = \frac{\omega(\lambda+\delta)}{\omega(\delta)}}$$

Example 22.4.6. For $\mathfrak{sl}(2,F)$ this gives

$$\mathrm{ch}_{V(m)} = \frac{\omega(m+1)}{\omega(1)} = \frac{t^{m+1}-t^{-m-1}}{t-t^{-1}}$$

22.5. **Example in $\mathfrak{sl}(3,F)$.** We will examine what the Weyl character formula says in the case of $\mathfrak{sl}(n+1,F)$ (with rank = dim $H = n$). First, we continue the "superman" example from before for $\mathfrak{sl}(3,F)$. In this case $W = S_3$ has six elements, three have positive sign and three have negative sign. This means that

$$\omega(\lambda+\delta) = \sum_{\sigma\in W=S_3} \mathrm{sgn}(\sigma)e(\sigma(\lambda+\delta))$$

has six terms, three positive and three negative as shown in red in Figure 22.5.1. The support of ch_λ for $\lambda = 3\lambda_1 + 2\lambda_2$ is shown in blue.

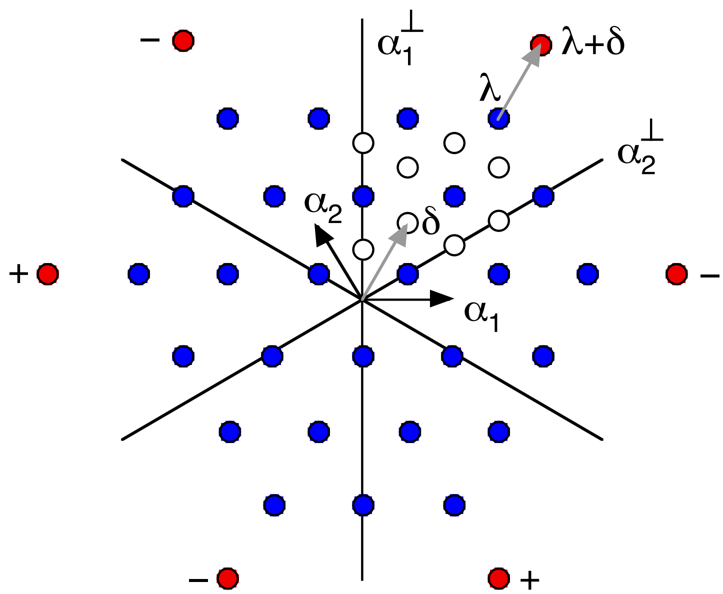


FIGURE 22.5.1. Support of ch_λ for $\lambda = 3\lambda_1 + 2\lambda_2$ in blue. The white dots are the other dominant weights including $\delta = \lambda_1 + \lambda_2$. They lie in the fundamental chamber between $\alpha_1^\perp$ and $\alpha_2^\perp$. Support of $\omega(\lambda+\delta)$ is in red.

Here we make the important observation that $\lambda+\delta$ is always strongly dominant. (So, the red spots are always distinct and do not cancel.)

The Weyl character formula says $\mathrm{ch}_\lambda\,\omega(\delta) = \omega(\lambda+\delta)$ or

$$\mathrm{ch}_\lambda \sum_{\sigma\in W} \mathrm{sgn}(\sigma)e(\sigma\delta) = \sum_{\sigma\in W} \mathrm{sgn}(\sigma)e(\sigma(\lambda+\delta))$$

Comparing the coefficient of $e(\mu)$ on both sides, we get:

$$\sum_{\sigma\in W} \mathrm{sgn}(\sigma)m_\lambda(\mu-\sigma\delta) = \begin{cases} \mathrm{sgn}(\sigma) & \text{if } \mu = \sigma(\lambda+\delta) \\ 0 & \text{otherwise} \end{cases}$$

For example, around the point marked μ ($= \lambda - \delta$) in Figure 22.5.2, the values of $\mathrm{sgn}(\sigma) m_\lambda(\mu - \sigma\delta)$ are

$$+3 - 2 + 1 - 1 + 1 - 2 = 0$$

Using the fact that the support of ch_λ is in the convex hull of the points $\sigma\lambda, \sigma \in W$ (the corners of the blue spot region) we can determine the value of $m_\lambda(\mu)$ at all points: starting with the value of $m_\lambda(\mu) = 1$ at the six corners (including $\mu = \lambda$) and working inward, we see that the value of $m_\lambda(\mu)$ is as given in Figure 22.5.2.

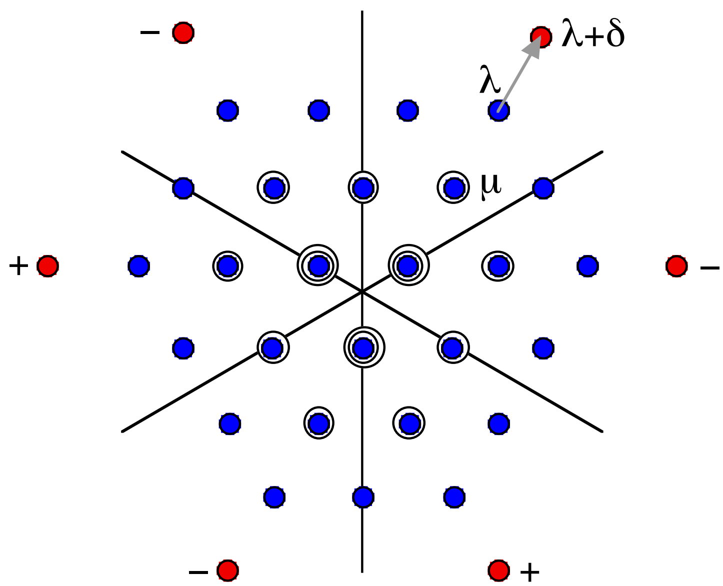


FIGURE 22.5.2. $m_\lambda(\mu)$ as given by the Weyl character formula.

22.6. **Character formula for $\mathfrak{sl}(n+1, F)$.** The simple roots of $L = \mathfrak{sl}(n+1, F)$ are $\alpha_i = \epsilon_i - \epsilon_{i+1}$, $i = 1, \cdots, n$, where $\epsilon_i : H \to F$ is projection to the ith entry. The corresponding basis elements in H are $h_i = e_i - e_{i+1}$ where e_i is the diagonal matrix with 1 in the ith position and 0 elsewhere. Thus $\lambda = \sum_{i=1}^n a_i \epsilon_i$ is dominant iff $a_1 \geq a_2 \geq \cdots a_n \geq 0$. The fundamental dominant weights are:

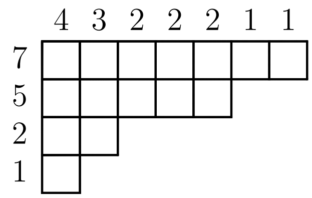
$\lambda_1 = \epsilon_1$

$\lambda_2 = \epsilon_1 + \epsilon_2$

$\lambda_i = \epsilon_1 + \cdots + \epsilon_i$

For example, $\lambda = 2\lambda_1 + 3\lambda_2 + \lambda_3 + \lambda_4$ is equal to $7\epsilon_1 + 5\epsilon_2 + 2\epsilon_3 + \epsilon_4$. This is illustrated by the Young diagram below. Also, $\delta = \sum \lambda_i = \sum (n + 1 - i)\epsilon_i$. Thus, if $n = 5$, then

$$\lambda + \delta = (7, 5, 2, 1, 0) + (5, 4, 3, 2, 1) = 12\epsilon_1 + 9\epsilon_2 + 5\epsilon_3 + 3\epsilon_4 + \epsilon_5$$



The Weyl group $W = S_{n+1}$ permutes the elements ϵ_i, $i = 1, \cdots, n+1$. The weights μ which occur in ch_λ for $\lambda = \sum a_i \epsilon_i$ correspond to points $(b_1, \cdots, b_{n+1})$ in the affine n plane given by $\sum b_i = \sum a_i$ ($= 15$ in this case). The correspondence is given by

$$\mu(b) = \sum_{i=1}^{n} (b_i - b_{i+1}) \lambda_i$$

This correspondence is additive in the sense that $\mu(b + b') = \mu(b) + \mu(b')$.

Theorem 22.6.1. *The coefficient $m_\lambda(\mu)$ of $e(\mu)$ in ch_λ for $\mu = \mu(b)$ as above is equal to the coefficient of $\prod x_i^{b_i}$ in the Schur polynomial $s_\lambda(x_1, \cdots, x_{n+1})$.*

The proof of this is simply that the definition of the Schur polynomial gives the Weyl character formula under the multiplicative correspondence

$$e(\mu(b)) \leftrightarrow \prod x_i^{b_i}$$

22.7. **Schur polynomials.**

Definition 22.7.1. If $\lambda = (a_1, \cdots, a_n)$ where $a_1 \geq a_2 \geq \cdots \geq a_n \geq 0$ then

$$s_\lambda(x_1, \cdots, x_{n+1}) := \frac{W(x_1^{a_1+n} x_2^{a_2+n-1} \cdots x_n^{a_n+1})}{W(x_1^n x_2^{n-1} \cdots x_n)}$$

where

$$W(x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}) = \sum_{\sigma \in S_{n+1}} \mathrm{sgn}(\sigma) x_{\sigma(1)}^{a_1} \cdots x_{\sigma(n)}^{a_n}$$

Remark 22.7.2. The components of λ are usually called $\lambda_1 \geq \lambda_2 \geq \cdots$. However, we have a clash of notation since λ_i is the ith fundamental dominant weight. So we use a_i.

The numerator and denominator of s_λ are homogeneous *alternating polynomials* in the sense that

$$W(x_{\sigma(1)}^{a_1} x_{\sigma(2)}^{a_2} \cdots x_{\sigma(n)}^{a_n}) = \mathrm{sgn}(\sigma) W(x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n})$$

Furthermore, the denominator is:

$$\Delta := W(x_1^n x_2^{n-1} \cdots x_n) = \prod_{1 \leq i < j \leq n+1} (x_i - x_j)$$

which has the property that it divides every alternating polynomial since $x_i - x_j$ clearly divides every alternating polynomial. Therefore, the ratio is a homogeneous *symmetric polynomial* in the sense that

$$s_\lambda(x_{\sigma(1)}, \cdots, x_{\sigma(n+1)}) = s_\lambda(x_1, \cdots, x_{n+1})$$

for all $\sigma \in S_{n+1}$.

Example 22.7.3. In the first example we have $n = 2, \lambda = 3\lambda_1 + 2\lambda_2 = 5\epsilon_1 + 2\epsilon_2$ (and $\lambda + \delta = 7\epsilon_1 + 3\epsilon_2$). So,

$$s_{(5,2)}(x_1, x_2, x_3) = \frac{W(x_1^7 x_2^3)}{W(x_1^2 x_2)} = \frac{x_1^7x_2^3 - x_1^3x_2^7 + x_2^7x_3^3 - x_2^3x_3^7 + x_1^3x_3^7 - x_1^7x_3^3}{x_1^2x_2 - x_1x_2^2 + x_2^2x_3 - x_2x_3^2 + x_1x_3^2 - x_1^2x_3}$$

This is a homogeneous polynomial of degree 7 in which every monomial corresponds to a weight $\mu(b) = \sum(b_i - b_{i+1})\lambda_i$ of $V(\lambda)$ with $b_1 + b_2 + b_3 = 7$ as illustrated in Figure 22.7 below.

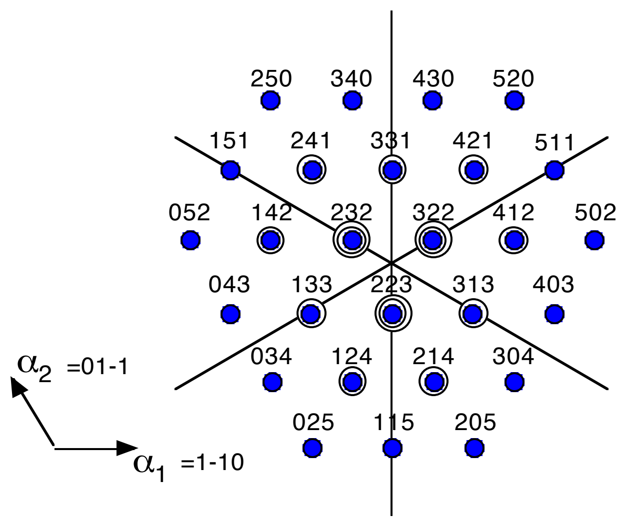


FIGURE 22.7.1. The points (b_1, b_2, b_3) with nonnegative integer b_i are in the plane given by $\sum b_i = 7$. The Weyl group $W = S_3$ permutes the b_i.

Figure (22.7) illustrates the equation: $s_\lambda(x_1, x_2, x_3) =$

$$x_1^2x_2^5 \quad + x_1^3x_2^4 \quad + x_1^4x_2^3 \quad + x_1^5x_2^2$$

$$+x_1x_2^5x_3 + 2x_1^2x_2^4x_3 + 2x_1^3x_2^3x_3 + 2x_1^4x_2^2x_3 + x_1^5x_2x_3$$

$$+x_2^5x_3^2 + 2x_1x_2^4x_3^2 + 3x_1^2x_2^3x_3^2 + 3x_1^3x_2^2x_3^2 + 2x_1^4x_2x_3^2 + x_1^5x_3^2$$

$$+x_2^4x_3^3 + 2x_1x_2^3x_3^3 + 3x_1^2x_2^2x_3^3 + 2x_1^3x_2x_3^3 + x_1^4x_3^3$$

$$+x_2^3x_3^4 + 2x_1x_2^2x_3^4 + 2x_1^2x_2x_3^4 + x_1^3x_3^4$$

$$+x_2^2x_3^5 \;\; + x_1x_2x_3^5 \;\; + x_1^2x_3^5$$

22.8. Dimension formula.

Theorem 22.8.1. *For any dominant weight λ, the dimension of the irreducible module $V(\lambda)$ is given by*

$$\dim V(\lambda) = \prod_{\sigma\in\Phi_+} \frac{\langle \lambda+\delta,\alpha\rangle}{\langle \delta,\alpha\rangle} = \prod_{\sigma\in\Phi_+} \frac{(\lambda+\delta,\alpha)}{(\delta,\alpha)}$$

In the case $L = \mathfrak{sl}(n+1,F)$ we get:

Corollary 22.8.2. *If $\lambda = (a_1,\cdots,a_{n+1})$ where $a_1 \geq \cdots \geq a_n \geq a_{n+1} = 0$ then*

$$s_\lambda(1,1,\cdots,1) = \prod_{1\leq i<j\leq n+1} \frac{a_i - a_j + j - i}{j-i}$$

Proof. This follows from the fact that the positive roots of type A_n are $\alpha_{ij} = \epsilon_i - \epsilon_j$ and $\langle \lambda, \alpha_{ij}\rangle = a_i - a_j$. ☐

Example 22.8.3. For $n = 2$ and $\lambda = (5,2,0)$ we have:

$$s_{(5,2)}(1,1,1) = \frac{1}{2}(a_1 - a_3 + 2)(a_1 - a_2 + 1)(a_2 - a_3 + 1)$$

$$\frac{1}{2}(5-0+2)(5-2+1)(2-0+1) = \frac{7\cdot 4\cdot 3}{2} = 42$$

which is equal to the number of blue spots with multiplicity calculated above.

Proof of Theorem 22.8.1. Let $v : \mathbb{Z}[\Lambda] \to \mathbb{Z}$ be the *augmentation map* given by $v(e(\mu)) = 1$ for all $\mu \in \Lambda$. Then

$$\dim V(\lambda) = v(\mathrm{ch}_\lambda)$$

However, the augmentation of

$$\omega(\delta) = \sum_{\sigma\in W} \mathrm{sgn}(\sigma) e(\sigma\delta) = \prod_{\alpha\in\Phi_+} \left(e\left(\frac{\alpha}{2}\right) - e\left(-\frac{\alpha}{2}\right)\right)$$

is zero. So, we get $\dim V(\lambda) = \frac{0}{0}$.

The equation $e(\lambda+\mu) = e(\lambda)\,e(\mu)$ implies that, for all $\alpha \in \Phi_+$,

$$\partial_\alpha e(\lambda) := \langle \lambda, \alpha\rangle\, e(\lambda)$$

is a derivation. Take the equation

$$\mathrm{ch}_\lambda\, \omega(\delta) = \omega(\lambda+\delta)$$

and apply the product $\partial = \prod_{\alpha\in\Phi_+} \partial_\alpha$ of all of these derivations. Then take the augmentation v.

Claim: For any nonempty subset $S \subset \Phi_+$, $v \prod_{\alpha\in S} \partial_a \omega(\delta) = 0$.

This follows from the fact that $\omega(\delta)$ is a product of $|\Phi_+|$ factors each of which evaluates to 0. So, we must differentiate each factor at least once. This gives:

$$\partial\omega(\delta) = \sum_{\sigma\in W} \mathrm{sgn}(\sigma) \prod_{\alpha\in\Phi_+} \langle \sigma\delta, \alpha\rangle\, e(\sigma\delta)$$

$$= \sum_{\sigma \in W} \operatorname{sgn}(\sigma) \prod_{\alpha \in \Phi_+} \left\langle \delta, \sigma^{-1}\alpha \right\rangle e(\sigma\delta)$$

However, the collection of roots $\{\sigma^{-1}\alpha\}$ is, up to sign, the set of positive roots with exactly $\ell(\alpha)$ negative roots. Since $\operatorname{sgn}(\sigma) = (-1)^{\ell(\sigma)}$, for each $\sigma \in W$ we have:

$$\operatorname{sgn}(\sigma) \prod_{\alpha \in \Phi_+} \left\langle \delta, \sigma^{-1}\alpha \right\rangle = \prod_{\alpha \in \Phi_+} \langle \delta, \alpha \rangle$$

Therefore,

$$v(\omega(\delta)) = |\Phi_+| \prod_{\alpha \in \Phi_+} \langle \delta, \alpha \rangle$$

Similarly,

$$v(\omega(\lambda + \delta) = |\Phi_+| \prod_{\alpha \in \Phi_+} \langle \lambda + \delta, \alpha \rangle$$

This gives

$$\dim V(\lambda) |\Phi_+| \prod_{\alpha \in \Phi_+} \langle \delta, \alpha \rangle = |\Phi_+| \prod_{\alpha \in \Phi_+} \langle \lambda + \delta, \alpha \rangle$$

and the theorem follows. □

23. Schur functors

In this lecture we will examine the Schur polynomials and Young tableaux to obtain irreducible representations of $\mathfrak{sl}(n+1,\mathbb{C})$. The construction is based on Fulton's book [3] which is highly recommended for beginning students because it is a beautifully written book with minimal prerequisites. This will be only in the case of $L = \mathfrak{sl}(n+1, F)$. Other cases and the proof of the Weyl character formula will be next.

First, we construct the irreducible representations $V(\lambda_i)$ corresponding to the fundamental dominant weights $\lambda_i = \epsilon_1 + \cdots + \epsilon_i$. Then we use Young tableaux to find the modules $V(\lambda)$ inside tensor products of these fundamental representations.

23.1. **Fundamental representations.**

Theorem 23.1.1. *The fundamental representation $V(\lambda_j)$ is equal to the ith exterior power of $V = F^{n+1}$:*

$$V(\lambda_j) = \wedge^j V$$

$L = \mathfrak{sl}(n+1, F)$ acts by the isomorphism $\mathfrak{sl}(n+1, F) \cong \mathfrak{sl}(V)$.

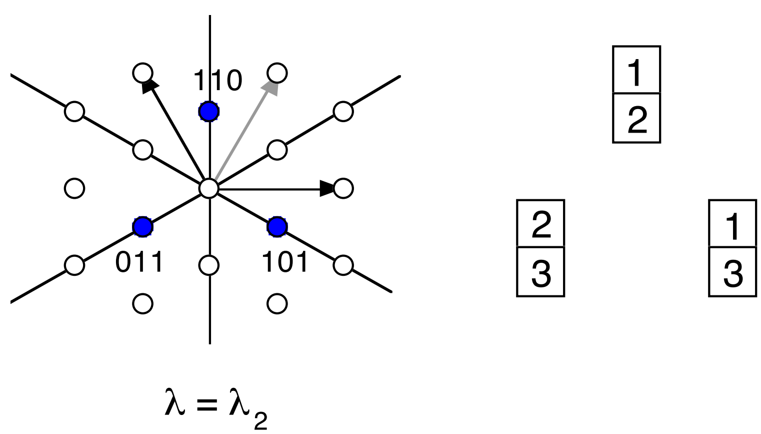
Definition 23.1.2. Recall that the *jth exterior power* of a vector space V is the quotient of the jth tensor power by all elements of the form $w_1 \otimes \cdots \otimes w_j$ where the w_i are not distinct. The image of $w_1 \otimes \cdots \otimes w_j$ in $\wedge^j V$ is denoted $w_1 \wedge \cdots \wedge w_j$. When the characteristic of F is not equal to 2, this is equivalent to saying that the wedge is skew symmetric:

$$w_{\sigma(1)} \wedge \cdots \wedge w_{\sigma(j)} = \operatorname{sgn}(\sigma) w_1 \wedge \cdots \wedge w_j$$

If $v_1, \cdots, v_{n+1}$ is a basis for V then

$$v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_j}$$

for $1 \le i_1 < i_2 < \cdots < i_j \le n+1$ is a basis for $\wedge^j V$.



Example 23.1.3. Take $n = 2$ and $\lambda = \lambda_2 = \epsilon_1 + \epsilon_2$. The Schur polynomial is

$$s_{11} = x_1x_2 + x_2x_3 + x_1x_3$$

The Young diagram is filled in with numbers $1, \cdots, n+1$ so that the numbers are increasing as we go down and nondecreasing as we go across. Let $V = F^{n+1} = F^3$ in this case with basis v_1, v_2, v_3. Then the corresponding representation is $\wedge^2 V$ with basis

$$v_1 \wedge v_2, \quad v_2 \wedge v_3, \quad v_1 \wedge v_3$$

The action of the Cartan subalgebra H is given on basis elements by

$$h(v_i \wedge v_j) = h(v_i) \wedge v_j + v_i \wedge h(v_j) = \epsilon_i(h)v_i \wedge v_j + v_i \wedge \epsilon_j(h)v_j$$

Therefore, $v_i \wedge v_j$ lies in the $\epsilon_i + \epsilon_j$ weight space. Since $\epsilon_1 + \epsilon_2 + \epsilon_3 = 0$ we get:

vector	$v_1 \wedge v_2$	$v_2 \wedge v_3$	$v_1 \wedge v_3$
weight	$\epsilon_1 + \epsilon_2$	$\epsilon_2 + \epsilon_3$	$\epsilon_1 + \epsilon_3$
	λ_2	$-\lambda_1$	$\lambda_1 - \lambda_2$
	maximal		

Proof of Theorem. In the general case, the maximal vector in $\wedge^j V$ is $v_1 \wedge v_2 \wedge \cdots \wedge v_j$ with highest weight λ_j. □

One of the key points is that the wedge $V \mapsto \wedge^j V$ is a *functor* (i.e., natural). In particular, any endomorphism $g : V \to V$ induces an endomorphism $\wedge^j g : \wedge^j V \to \wedge^j V$ making $\wedge^j V$ into a module over $\mathfrak{sl}(V) \cong \mathfrak{sl}(n+1, F)$.

23.2. **Symmetric powers.** As I pointed out before, the irreducible module $V(\lambda)$ for $\lambda = \sum c_i\lambda_i$ is a direct summand of a tensor product of $V(\lambda_i)$, taking c_i copies of $V(\lambda_i)$. For example, take $n = 2$ and $\lambda = \lambda_1 + \lambda_2 = 2\epsilon_1 + \epsilon_2$. But, the tensor product of $V(\lambda_2) = \wedge^2 V$ and $V(\lambda_1) = \wedge^1 V = V$ is $3 \times 3 = 9$ dimensional, whereas, $V(\lambda_1 + \lambda_2)$ is 8 dimensional (see worksheet). The Schur-Weyl construction tells us which 1-dimensional subspace to mod out.

Example 23.2.1. Take the example $\lambda = 2\lambda_1$. This Young diagram

and Schur polynomial

$$s_{20} = x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_2x_3 + x_1x_3$$

This is the sum of all monomials of degree 2 in the variables x_1, x_2, x_3. Therefore, it corresponds to the second symmetric power of V:

$$V(2\lambda_1) = \mathcal{S}^2(V)$$

A basis for $\mathcal{S}^2(V)$ is given by $v_i \widetilde{\otimes} v_j$ where $1 \le i \le j \le n+1$. These basis elements are represented by filling in the boxes:

i	j

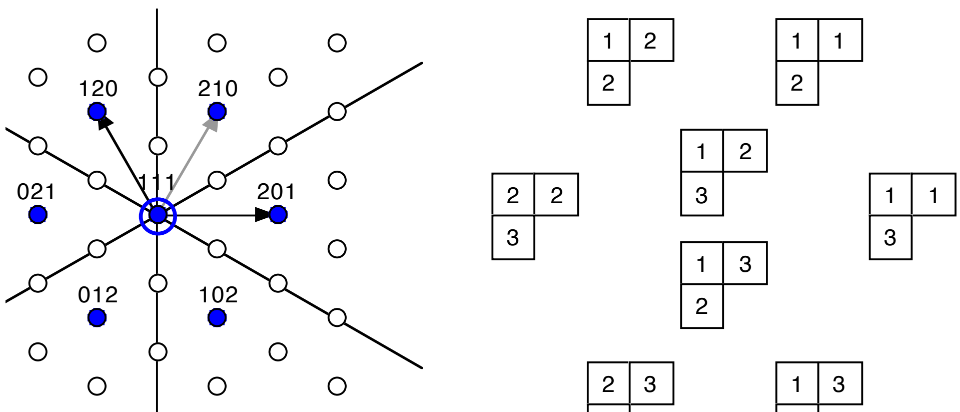
More generally, for any n and k, we have

$$V(k\lambda_1) = \mathcal{S}^k(V)$$

23.3. **Schur functors.** Let $\lambda = \sum \lambda_{b_j}$, where $b_1 \geq b_2 \geq \cdots$, be given by a Young diagram D (whose jth column has b_j boxes). Then the Schur functor $\mathcal{S}_\lambda(V)$ is given as a quotient of

$$\wedge^{b_1} V \otimes \wedge^{b_2} V \otimes \wedge^{b_3} V \otimes \cdots$$

by *exchange relation* which I will explain in class using the two examples on the worksheet.



$\lambda = \lambda_1 + \lambda_2$

Example 23.3.1. Take the example $\lambda = \lambda_1 + \lambda_2 = 2\epsilon_1 + \epsilon_2$. The irreducible module $V(\lambda)$, which is 8 dimensional by the calculation in the diagrams above, is a direct summand of $V(\lambda_2) \otimes V(\lambda_1) = \wedge^2 V \otimes \wedge^1 V = \wedge^2 V \otimes V$ which is $3 \times 3 = 9$ dimensional. What is the missing basis vector?

The missing vector is $(v_2 \wedge v_3) \otimes v_1$ which corresponds to the Young tableau

$$\begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 & \\ \cline{1-1} \end{array}$$

This is not admissible since $2 > 1$ but it is a linear combination of admissible Young tableaux by the exchange relation which says that the contents of any box can be exchanged with the contents of all boxed in any other column to the left of the first box:

$$\begin{array}{|c|c|}\hline A & D \\ \hline B & \\ \cline{1-1} C & \\ \cline{1-1} \end{array} = \begin{array}{|c|c|}\hline D & A \\ \hline B & \\ \cline{1-1} C & \\ \cline{1-1} \end{array} + \begin{array}{|c|c|}\hline A & B \\ \hline D & \\ \cline{1-1} C & \\ \cline{1-1} \end{array} + \begin{array}{|c|c|}\hline A & C \\ \hline B & \\ \cline{1-1} D & \\ \cline{1-1} \end{array}$$

In this case we have:

$$\begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 & \\ \cline{1-1} \end{array} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \\ \cline{1-1} \end{array} + \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & \\ \cline{1-1} \end{array}$$

The second term is equal to

$$\begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & \\ \cline{1-1} \end{array} = - \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \\ \cline{1-1} \end{array}$$

since the first column represents $v_2 \wedge v_1 = -v_1 \wedge v_2$. Thus:

$$(v_2 \wedge v_3)\widetilde{\otimes}v_1 = (v_1 \wedge v_3)\widetilde{\otimes}v_2 - (v_1 \wedge v_2)\widetilde{\otimes}v_3$$

where $\widetilde{\otimes}$ indicates tensor product symmetrised by the exchange relations.

More generally, the exchange relation says that any set of squares in any column can be exchanged with squares in one other column as long as one takes the sum of all ways to do that and, also, the second column is to the left of the first column. For example, if two columns have the same number of squares then they can be switched. Another example is:

$$\begin{array}{|c|c|}\hline A & D \\ \hline B & E \\ \hline C \\ \cline{1-1}\end{array} = \begin{array}{|c|c|}\hline D & A \\ \hline E & B \\ \hline C \\ \cline{1-1}\end{array} + \begin{array}{|c|c|}\hline A & B \\ \hline D & C \\ \hline E \\ \cline{1-1}\end{array} + \begin{array}{|c|c|}\hline D & A \\ \hline B & C \\ \hline E \\ \cline{1-1}\end{array}$$

Theorem 23.3.2. *If λ is the sum of fundamental dominant weights λ_{j_i} then $V(\lambda)$ is the quotient of $\bigotimes V(\lambda_{j_i})$ by the exchange relations.*

This quotient is the *Schur functor* $\mathcal{S}_\lambda(V)$.

Proof. The quotient $\mathcal{S}_\lambda(V)$ is a representation of $L = \mathfrak{sl}(V)$. By a combinatorial argument, we can see that it has a basis given by the admissible Young tableaux. The Schur polynomial s_λ is the sum of the corresponding monomials. So, the character of $\mathcal{S}_\lambda(V)$ is equal to ch_λ as given by the Weyl character formula. But we know that representations are uniquely determined by their formal characters. So, we conclude that $\mathcal{S}_\lambda(V) = V(\lambda)$. $\square$

24. Proof of Weyl character formula

(1) Casimir operator
(2) Review α-root strings.
(3) Formula for $m_V(\mu) = \dim V_\mu$
(4) Freudenthal multiplicity formula
(5) Weyl character formula
(6) Kostant's formula

This proof is from Fulton and Harris [4], Lecture 25, following the notation of Humphreys [5] Section 22.

24.1. **Casimir operator.** (from subsection 6.3). If V is a representation of a semisimple Lie algebra L we have a homomorphism $\varphi : L \to \mathfrak{gl}(V)$. The *Casimir operator* $c \in \mathrm{End}_L(V)$ of φ is given by

$$c_V = \sum \varphi(x_i)\varphi(x_i^*)$$

where $\{x_i\}$ is a basis for L and $\{x_i^*\}$ is the dual basis with respect to the Killing form κ. We proved the following:

(1) c_V is independent of the choice of basis $\{x_i\}$.
(2) $c_V : V \to V$ is a homomorphism of L-modules.
(3) If V is irreducible, then, by Schur's Lemma, c_V is multiplication by a scalar (which we also call c).

Lemma 24.1.1. *If V is irreducible, the trace of $c_V = \sum \varphi(x_i)\varphi(x_i^*)|V_\mu$ is $c\, m_V(\mu) = c \dim V_\mu$.*

24.1.1. *choice of basis.* Take a root space decomposition $L = H \oplus \bigoplus L_\alpha$ and take the basis $h_i, i = 1, \cdots, n$ for H. For every root $\alpha \in \Phi$ choose an element $x_\alpha \in L_\alpha$. This gives a basis for L.

Let $\{h_i^*\}$ be the basis of H dual to $\{h_i\}$.

Lemma 24.1.2.

$$\sum_{i=1}^{n} \mathrm{Tr}(\varphi(h_i)\varphi(h_i^*)|V_\mu) = \sum_{i=1}^{n} \mu(h_i)\mu(h_i^*) m_V(\mu) = (\mu, \mu) m_V(\mu)$$

Proof. Suppose $t_\mu = \sum a_i h_i$. By definition of t_μ we have

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

$$\mu(h_j^*) = \sum_i a_i \kappa(h_i, h_j^*) = a_j$$

So,

$$\sum_j \mu(h_j)\mu(h_j^*) = \sum_{i,j} a_i a_j \kappa(h_i, h_j) = \kappa(t_\mu, t_m u) = (\mu, \mu)$$

□

Lemma 24.1.3. *Let z_α be dual to x_α. (So, $\kappa(x_\alpha, z_\alpha) = 1$.) Then*

$$z_\alpha = \frac{(\alpha,\alpha)}{2} y_\alpha \qquad \text{and} \qquad [x_\alpha, z_\alpha] = t_\alpha$$

Proof. Since $z_\alpha \in L_{-\alpha}$ which is 1-dimensional, and

$$[x_\alpha, y_\alpha] = h_\alpha = \frac{2}{(\alpha,\alpha)} t_\alpha$$

the two equations are equivalent. Associativity of κ forces the second equation to be true:

$$\kappa(h_\alpha, [x_\alpha, z_\alpha]) = \kappa([h_\alpha, x_\alpha], z_\alpha) = \kappa(2x_\alpha, z_\alpha) = 2$$

$$\kappa(h_\alpha, t_\alpha) = \alpha(h_\alpha) = 2.$$

Since $[x_\alpha, z_\alpha]$ is a multiple of t_α, they must be equal. □

Lemma 24.1.4.

$$\mathrm{Tr}(\varphi(x_\alpha)\varphi(z_\alpha)|V_\mu) = \sum_{i=0}^{\infty} (\mu + i\alpha, \alpha) m_V(\mu + i\alpha)$$

Multiplying both sides by $\frac{2}{(\alpha,\alpha)}$ we see that this is equivalent to:

$$\mathrm{Tr}(\varphi(x_\alpha)\varphi(y_\alpha)|V_\mu) = \sum_{i=0}^{\infty} \langle \mu + i\alpha, \alpha \rangle\, m_V(\mu + i\alpha) \tag{24.1}$$

To prove this we need to look at the α root string through μ.

24.2. **α-root strings.** Recall (subsection 8.4.2) that the α string through μ is

$$V_{\mu+q\alpha} \oplus V_{\mu+(q-1)\alpha} \oplus \cdots \oplus V_{\mu+\alpha} \oplus V_\mu \oplus V_{\mu-\alpha} \oplus \cdots \oplus V_{\mu-r\alpha}$$

where $\mu(h_\alpha) = \langle \mu, \alpha \rangle = r - q$ and

$$q + r = m = (\mu + q\alpha)(h_\alpha) = \langle \mu + q\alpha, \alpha \rangle = \langle \mu, \alpha \rangle + 2q$$

Consider the α string as a module over $S_\alpha \cong \mathfrak{sl}(2, F)$. Then $m_V(\mu + q\alpha) = \dim V_{\mu+q\alpha}$ is the number of copies of the indecomposable S_α module $V(m)$ in V where $V(m)$ is the S_α module which starts at $V_{\mu+q\alpha}$ and goes to $V_{\mu-r\alpha}$.

By Lemma 7.0.4, the action of $x_\alpha y_\alpha$ on the weight space $V(m)_\mu$ is

$$(q+1)(m-q) = (q+1)(\langle \mu, \alpha \rangle + q)$$

Since this occurs $m_V(\mu + q\alpha)$ times, the contribution of $V(m)$ to the trace of $\varphi(x_\alpha)\varphi(y_\alpha)$ is

$$m_V(\mu + q\alpha)(q+1)(\langle \mu, \alpha \rangle + q)$$

Consider the components of the α string which go from $V_{\mu+i\alpha}$ to $V_{\mu-j\alpha}$. The number of such components is $m_V(\mu + i\alpha) - m_V(\mu + (i+1)\alpha)$. Here $m = \langle \mu, \alpha \rangle + 2i$. So, the action of $x_\alpha y_\alpha$ on the weight space $V(m)_\mu$ is

$$(i+1)(m-i) = (i+1)(\langle \mu, \alpha \rangle + i)$$

Claim:

$$\mathrm{Tr}(\varphi(x_\alpha)\varphi(y_\alpha)|V_\mu) = \sum_{i=0}^{\infty}(i+1)(\langle \mu, \alpha\rangle + i)[m_V(\mu+i\alpha) - m_V(\mu+(i+1)\alpha)]$$

Furthermore, this expression simplifies to give Equation (24.1) since the coefficient of $m_V(\mu+i\alpha)$ is

$$(i+1)(\langle \mu,\alpha\rangle + i) - i(\langle \mu,\alpha\rangle + i - 1) = (\langle \mu,\alpha\rangle + i) + i = \langle \mu + i\alpha, \alpha\rangle$$

To prove the claim we note two things.

(1) The terms with $i > q$ are zero since $m_V(\mu+i\alpha) = 0$ in those cases.
(2) The terms corresponding to α strings which do not contain μ add up to zero.

To prove the second point, we look at the case where the α string goes from $V_{\mu+i\alpha}$ to $V_{\mu+(j+1)\alpha}$ where $j \geq 0$. The number of copies of this α string in V is

$$[m_V(\mu+i\alpha) - m_V(\mu+(i+1)\alpha)]$$
$$= -[m_V(\mu+j\alpha) - m_V(\mu+(j+1)\alpha)]$$

Also, $\langle \mu,\alpha\rangle = -i-j-1$. This implies that the $i=i$ and $i=j$ terms in the summation cancel:

$$(i+1)(-i-j-1+i)[m_V(\mu+i\alpha) - m_V(\mu+(i+1)\alpha)]$$
$$+(j+1)(-i-j-1+j)[m_V(\mu+j\alpha) - m_V(\mu+(j+1)\alpha)] = 0$$

Point (2) also applies to the case where μ is not in the support of V (i.e., when $V_\mu = 0$). In that case, all terms in the sum will cancel. So, the formula in the Lemma applies to all μ.

Lemma 24.2.1. *For all weights μ we have:*

$$\sum_{i=-\infty}^{\infty}(\mu+i\alpha,\alpha)m_V(\mu+i\alpha) = 0$$

Proof. The sum does not change if we replace μ by $\mu - N\alpha$. Making N large enough, the sum will be zero for $i < 0$ and Point (2) applies. □

24.3. **First formula for $m_V(\mu)$.** Putting together the lemmas, we get the following for all weights μ.

$$c\, m_V(\mu) = \mathrm{Tr}(c_V|V_\mu) = (\mu,\mu)m_V(\mu) + \sum_{\alpha\in\Phi}\sum_{i=0}^{\infty}(\mu+i\alpha,\alpha)m_V(\mu+i\alpha)$$

The terms in the sum with $i=0$ are $(\mu,\alpha)m_V(\mu)$ which changes sign when α becomes $-\alpha$. So, these terms cancel and we can start the summation with $i=1$. By Lemma 24.2.1, the $-\alpha$ term in the sum is now

$$\sum_{i=1}^{\infty}(\mu-i\alpha,-\alpha)m_V(\mu-i\alpha) = \sum_{i=0}^{\infty}(\mu+i\alpha,\alpha)m_V(\mu+i\alpha)$$

This gives

$$c\, m_V(\mu) = (\mu,\mu)m_V(\mu) + \sum_{\alpha\in\Phi_+}(\mu,\alpha)m_V(\mu) + 2\sum_{\alpha\in\Phi_+}\sum_{i=1}^{\infty}(\mu+i\alpha,\alpha)m_V(\mu+i\alpha)$$

Using $\delta = \frac{1}{2}\sum_{\alpha\in\Phi_+}\alpha$, and collecting the $m_V(\mu)$ term on the left, we get:

$$(c-(\mu,\mu)-2(\mu,\delta))m_V(\mu) = 2\sum_{\alpha\in\Phi_+}\sum_{i=1}^{\infty}(\mu+i\alpha,\alpha)m_V(\mu+i\alpha)$$

There is one case in which we know all of these numbers. Suppose $V = V(\lambda)$ has highest weight λ and we take $\mu = \lambda$. Then $m_V(\lambda+i\alpha) = m_\lambda(\lambda+i\alpha) = 0$ for all $\alpha\in\Phi_+$ by definition of highest weight. So,

$$c = (\lambda,\lambda)+2(\lambda,\delta) = (\lambda+\delta,\lambda+\delta)-(\delta,\delta) = ||\lambda+\delta||^2 - ||\delta||^2$$

Therefore,

$$\begin{aligned} c-(\mu,\mu)-2(\mu,\delta) &= c-(\mu+\delta,\mu+\delta)+(\delta,\delta) \\ &= (\lambda+\delta,\lambda+\delta)-(\mu+\delta,\mu+\delta) \\ &= ||\lambda+\delta||^2 - ||\mu+\delta||^2 \end{aligned}$$

This gives the following theorem.

Theorem 24.3.1 (Freudenthal multiplicity formula)**.** $m_\lambda(\lambda) = 1$ *and for all* $\mu < \lambda$ *we have*

$$m_\lambda(\mu) = \frac{2\sum_{\alpha\in\Phi_+}\sum_{i=1}^{\infty}(\mu+i\alpha,\alpha)m_\lambda(\mu+i\alpha)}{||\lambda+\delta||^2-||\mu+\delta||^2}$$

This is a recursive formula for $m_\lambda(\mu)$ since the terms on the right involve only $m_\lambda(\nu)$ for $\mu < \nu \le \lambda$.

Example 24.3.2. Let $L = \mathfrak{sl}(3,F)$ and $\lambda = \lambda_1+\lambda_2 = \delta = \alpha_1+\alpha_2$. Then λ is a positive root and all roots in L have length 1. So,

$$||\lambda+\delta||^2 = 4||\lambda||^2 = 4$$

For $\mu = \alpha_2 = \lambda-\alpha_1$, $\mu+\delta = \alpha_1+2\alpha_2$ and $(\mu+\delta,\mu+\delta) = 3$. So, the formula gives

$$m_\lambda(\alpha_2) = 2(\lambda,\alpha_1)m_\lambda(\lambda) = 1$$

The case $\mu = \alpha_1$ also gives $m_\lambda(\alpha_1) = 1$ by symmetry.

For $\mu = 0$ we get

$$m_\lambda(0) = \frac{2((\alpha_1,\alpha_1)m_\lambda(\alpha_1)+(\alpha_2,\alpha_2)m_\lambda(\alpha_2)+(\lambda,\lambda)m_\lambda(\lambda))}{4-1} = 2$$

etc.

24.4. **Outline of the proof.** We want to prove the Weyl character formula

$$\omega(\delta)\mathrm{ch}_\lambda = \omega(\lambda+\delta)$$

In this form, both sides are alternating with highest weight $\lambda+\delta$ and the coefficient of $e(\lambda+\delta)$ is 1 on both sides.

Lemma 24.4.1. *The elements $\omega(\lambda)$, for λ strongly dominant, form a basis for $\mathbb{Z}[\Lambda]^-$, the additive group of alternating elements of $\mathbb{Z}[\Lambda]$.*

Proof. Every nonzero element of $\mathbb{Z}[\Lambda]$ has at least one highest weight λ which is dominant. But, if λ is dominant but not strongly dominant, then the coefficient of $e(\lambda)$ must be zero for any alternating element since $s_i\lambda = \lambda$ for some simple reflection s_i and $e(s_i\lambda) = e(\lambda)$ have coefficients of opposite sign (being alternating).

The lemma is now an easy induction. Choose a highest weight λ and subtract the appropriate multiple of $\omega(\lambda)$ to reduce the size of the support. □

Definition 24.4.2 (Dominance order)**.** If $\lambda,\mu \in \Lambda$ are integer weights, we say $\lambda > \mu$ if $\lambda - \mu$ is dominant.

Exercise 24.4.3. If λ, μ are strongly dominant weights with $\lambda > \mu$ then show that $||\lambda|| > ||\mu||$. (Since $(\mu, \lambda-\mu) \geq 0$.)

Lemma 24.4.4. *Let $\Delta : \mathbb{Z}[\Lambda] \to \mathbb{Z}[\Lambda]$ be the linear function given by*

$$\Delta\left(\sum n_\mu e(\mu)\right) = \sum n_\mu(\mu,\mu)e(\mu)$$

then

$$\Delta\left(\sum a_\lambda \omega(\lambda)\right) = \sum a_\lambda ||\lambda||^2\omega(\lambda)$$

To prove the character formula it is enough to show that

$$\Delta(\omega(\delta)\mathrm{ch}_\lambda) = ||\lambda+\delta||^2\omega(\delta)\mathrm{ch}_\lambda$$

The reason is that, $\omega(\delta)\mathrm{ch}_\lambda$ is alternating and has highest weight $\lambda+\delta$. The other weights that occur in $\omega(\delta)\mathrm{ch}_\lambda$ are

$$\mu + \sigma(\delta) < \lambda + \delta$$

since $\mu \leq \lambda$ and $\sigma(\delta) \leq \delta$.

24.4.1. *Formulas for $\omega(\delta)$.* Recall that $\delta = \sum_{\alpha>0}\frac{1}{2}\alpha$ where $\alpha > 0$ means $\alpha \in \Phi_+$. Thus

$$e(\delta) = \prod_{\alpha>0} e(\alpha/2)$$

We used the following formula in the proof of the dimension formula (Theorem 22.8.1) and we need it again.

Lemma 24.4.5.

$$\omega(\delta) = \prod_{\alpha>0}\left(e\left(\frac{\alpha}{2}\right) - e\left(\frac{-\alpha}{2}\right)\right) = \prod_{\alpha>0} e(\delta)(1-e(-\alpha))$$

Proof. Both sides are alternating with the same highest weight term $e(\delta) = \prod_{\alpha>0} e(\alpha/2)$. But δ is the minimal strongly dominant weight. □

Let $r = |\Phi_+|$ be the number of positive roots. Then

$$(-1)^r \omega(\delta) = \prod_{\alpha<0} \left(e\left(\frac{\alpha}{2}\right) - e\left(\frac{-\alpha}{2}\right)\right)$$

since each of the r terms on the RHS is the negative of the corresponding term in the equation above. Therefore,

$$(-1)^r \omega(\delta)^2 = \prod_{\alpha} \left(e\left(\frac{\alpha}{2}\right) - e\left(\frac{-\alpha}{2}\right)\right)$$

where the product is now over all roots α. Call this expression P. Since $\prod_\alpha e(\alpha/2) = e(\sum \alpha/2) = e(0) = 1$ we have

$$P = \prod_{\alpha} (e(\alpha) - 1)$$

The plan is now the following. To prove the Weyl character formula, we will take the equation that we used to prove the Freudenthal formula, multiply both sides by P, then differentiate using the operator d defined below.

24.5. **The derivation d.** Let d be the linear differential operator on $\mathbb{Z}[\Lambda]$ defined by

$$d \sum a_\mu e(\mu) = \sum a_\mu e(\mu)\, d\mu \in \mathbb{Z}[LL] \otimes dH^*$$

where $d\mu$ is a symbol which is linear in $\mu \in H^*$. More precisely, $d\mu$ is an element of dH^* which is a free H^* module generated by the single element d. In particular, we have

$$d(\mu + \nu) = d\mu + d\nu$$

Lemma 24.5.1. *$d : \mathbb{Z}[LL] \to \mathbb{Z}[LL] \otimes dH^*$ is a derivation.*

Proof. Let $A = \sum a_\mu e(\mu)$ and $B = \sum b_\nu e(\nu)$. Then

$$AB = \sum_{\mu,\nu} a_\mu b_\nu e(\mu + \nu)$$

So,

$$d(AB) = \sum_{\mu,\nu} a_\mu b_\nu e(\mu+\nu) d(\mu+\nu)$$

$$= \sum_{\mu,\nu} a_\mu b_\nu e(\mu)e(\nu) d\mu + \sum_{\mu,\nu} a_\mu b_\nu e(\mu)e(\nu) d\nu$$

$$= BdA + AdB$$

□

Example 24.5.2. Since $P = \prod_\alpha (e(\alpha) - 1)$ and $d(e(\alpha) - 1) = e(\alpha)d\alpha$, we have

$$dP = \sum_\alpha P_\alpha e(\alpha)\, d\alpha$$

where $P_\alpha = \prod_{\beta \neq \alpha}(e(\beta) - 1)$. Since $P = (-1)^r \omega(\delta)^2$ we also have

$$dP = (-1)^r 2\omega(\delta)\, d\omega(\delta)$$

Definition 24.5.3. Let $\Omega = \mathbb{Z}[\Lambda] \otimes dH^*$ considered as a left $\mathbb{Z}[\Lambda]$ module and let $(-,-) : \Omega \otimes \Omega \to \mathbb{C}[\Lambda]$ be given by

$$\left(\sum A_\mu d\mu, \sum B_\nu d\nu\right) = \sum_{\mu,\nu} A_\mu B_\nu (\mu, \nu)$$

Lemma 24.5.4. *For any $A, B \in \mathbb{Z}[\Lambda]$ we have*

$$\Delta(AB) = B\Delta A + A\Delta B + 2(dA, dB)$$

Proof. Since both sides are bilinear in A, B it suffices to consider the case $A = e(\mu), B = e(\nu)$. Then

$$\Delta(e(\mu)e(\nu)) = ||\mu + \nu||^2 e(\mu + \nu)$$

On the other side we have

$$(||\mu||^2 + ||\nu||^2 + 2(\mu, \nu))e(\mu + \nu)$$

which are equal. □

24.6. **Proof of Weyl character formula.** Recall the formula:

$$c\, m_\lambda(\mu) = (\mu, \mu)m_\lambda(\mu) + \sum_{\alpha \in \Phi} \sum_{i=0}^{\infty} (\mu + i\alpha, \alpha) m_\lambda(\mu + i\alpha)$$

where $c = ||\lambda + \delta||^2 - ||\mu + \delta||^2$. Multiply by $e(\mu)$ and sum over all weights μ to get

$$\underbrace{c \sum_\mu m_\lambda(\mu) e(\mu)}_{c\,\mathrm{ch}_\lambda} = \underbrace{\sum_\mu ||\mu||^2 m_\lambda(\mu) e(\mu)}_{\Delta \mathrm{ch}_\lambda} + \sum_\mu \sum_\alpha \sum_{i \geq 0} (\mu + i\alpha, \alpha) m_\lambda(\mu + i\alpha) e(\mu)$$

The last term simplifies if we multiply by $P = P_\alpha(e(\alpha) - 1)$ since

$$(e(\alpha) - 1) \sum_\mu \sum_{i \geq 0} (\mu + i\alpha, \alpha) m_\lambda(\mu + i\alpha) e(\mu) = \sum_\mu (\mu, \alpha) m_\lambda(\mu) e(\mu + \alpha)$$

Therefore, P times the triple summation is

$$\sum_\mu \sum_\alpha (\mu, \alpha) P_a m_\lambda(\mu) e(\mu + \alpha) = (dP, d\mathrm{ch}_\lambda)$$

This gives:

$$cP\mathrm{ch}_\lambda = P\Delta \mathrm{ch}_\lambda + (dP, d\mathrm{ch}_\lambda)$$

Now expand $P = (-1)^r \omega(\delta)^2$.

$$(-1)^r c\omega(\delta)^2 \mathrm{ch}_\lambda = (-1)^r \omega(\delta)^2 \Delta \mathrm{ch}_\lambda + (-1)^r 2\omega(\delta)(d\omega(\delta), d\mathrm{ch}_\lambda)$$

Since $\mathbb{C}[\Lambda] \cong \mathbb{C}[x_1, x_1^{-1}, \cdots, x_n, x_n^{-1}]$ is an integral domain, we can cancel $(-1)^r\omega(\delta)$ from both sides to get

$$\begin{aligned} c\omega(\delta)\mathrm{ch}_\lambda &= \omega(\delta)\Delta\mathrm{ch}_\lambda + 2(d\omega(\delta), d\mathrm{ch}_\lambda) \\ &= \Delta(\omega(\delta)\mathrm{ch}_\lambda) - \mathrm{ch}_\lambda\Delta(\omega(\delta)) \end{aligned}$$

But,

$$c = ||\lambda+\delta||^2 - ||\delta||^2$$
$$\Delta\omega(\delta) = ||\delta||^2$$

So, we get

$$||\lambda+\delta||^2\omega(\delta)\mathrm{ch}_\lambda = \Delta(\omega(\delta)\mathrm{ch}_\lambda)$$

proving that

$$\omega(\delta)\mathrm{ch}_\lambda = \omega(\lambda+\delta)$$

as I explained earlier.

24.7. **Kostant's formula.** For any integer weight ν let $p(\nu)$ be the number of ways that ν can be written as a sum of positive roots $\alpha \in \Phi_+$. This is also equal to the number of ways that $-\nu$ can be written as a sum of negative roots. So,

$$\sum_\nu p(\nu)e(-\nu) = \prod_{\alpha>0} \frac{1}{1-e(-\alpha)}$$

Theorem 24.7.1.

$$m_\lambda(\mu) = \sum_{\sigma\in W} \mathrm{sgn}(\sigma)p(\sigma(\lambda+\delta) - (\mu+\delta))$$

Proof. Since $\omega(\delta) = e(\delta)\prod_{\alpha>0}(1-e(-\alpha))$, we have

$$\omega(\delta)^{-1} = e(-\delta)\prod_{\alpha>0}\frac{1}{1-e(-\alpha)} = \sum_\nu p(\nu)e(-\nu-\delta)$$

Therefore,

$$\mathrm{ch}_\lambda = \omega(\delta)^{-1}\omega(\lambda+\delta) = \sum_\nu p(\nu)e(-\nu-\delta)\sum_{\sigma\in W}\mathrm{sgn}(\sigma)e(\sigma(\lambda+\delta))$$

$$\sum_\mu m_\lambda(\mu)e(\mu) = \sum_{\sigma\in W}\sum_\nu \mathrm{sgn}(\sigma)p(\nu)e(\sigma(\lambda+\delta)-(\nu+\delta))$$

Thus $m_\lambda(\mu)$ is the coefficient of $e(\mu)$ on the right. But

$$\mu = \sigma(\lambda+\delta) - (\nu+\delta)$$

is equivalent to

$$\nu = \sigma(\lambda+\delta) - (\mu+\delta)$$

So, the coefficient of $e(\mu)$ is

$$m_\lambda(\mu) = \sum_{\sigma\in W} \mathrm{sgn}(\sigma)p(\sigma(\lambda+\delta) - (\mu+\delta))$$

□

References

1. R. W. Carter, *Lie algebras of finite and affine type*, Cambridge Studies in Advanced Mathematics, vol. 96, Cambridge University Press, Cambridge, 2005.
2. Karin Erdmann and Mark J. Wildon, *Introduction to Lie algebras*, Springer Undergraduate Mathematics Series, Springer-Verlag London Ltd., London, 2006.
3. William Fulton, *Young tableaux*, London Mathematical Society Student Texts, vol. 35, Cambridge University Press, Cambridge, 1997, With applications to representation theory and geometry.
4. William Fulton and Joe Harris, *Representation theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991, A first course, Readings in Mathematics.
5. James E. Humphreys, *Introduction to Lie algebras and representation theory*, Graduate Texts in Mathematics, vol. 9, Springer-Verlag, New York, 1978, Second printing, revised.
6. Claudio Procesi, *Lie groups*, Universitext, Springer, New York, 2007, An approach through invariants and representations.
7. Patrice Tauvel and Rupert W. T. Yu, *Lie algebras and algebraic groups*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2005.