

The Representation Theory of Lie Algebras

Jacob Denson

March 8, 2017

Table Of Contents

1	Basic Definitions	2
1.1	Factorization	9
1.2	Solvable Lie Algebras	11
1.3	Low Dimensional Classifications	13
2	Matrix Lie Algebras	21
2.1	Engel's Theorem	24
2.2	Representation Theory	28
2.3	Constructing Representations	31
2.4	Tensors and Universal Enveloping Algebras	32
2.5	Representations of $\mathfrak{sl}_2(K)$	35
2.6	Weyl's Theorem	39
3	Killing Forms and Root Systems	41
3.1	Cartan's Criterion	41
3.2	Root Systems	48
3.3	Subalgebras isomorphic to $\mathfrak{sl}_2(K)$	53
3.4	Cartan Subalgebras as Inner Product Spaces	57
3.5	Root systems	59
3.6	Root Spaces for the Classical Lie Algebras	63

Chapter 1

Basic Definitions

If K is a field, then a Lie algebra over K is a K vector-space \mathfrak{g} equipped with an alternating, bilinear form $[\cdot, \cdot]$, known as the **Lie bracket**, satisfying the Jacobi identity

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

for any $X, Y, Z \in \mathfrak{g}$. The majority of finite dimensional Lie algebras emerge in geometry and some parts of number theory (normally where $K = \mathbf{R}$ or $K = \mathbf{C}$). The symmetries in these fields can often be characterized as a Lie group with a complicated, nonlinear multiplication operation. As is ritual in differential geometry, we try and understand this operation by considering the linearized approximation of the operation on infinitesimals.

The entire premise behind the study of finite dimensional Lie algebras is that properties of infinitesimals under the linear approximations of group operations rise up to give us properties of the Lie group. To begin with, we can consider ‘infinitesimal multiplication’. Given the multiplication map $m : G \times G \rightarrow G$, we can differentiate the map to obtain $m_* : TG \times TG \rightarrow TG$, and we can consider this a multiplication operation on the set of infinitesimals. Similarly, we can differentiate the inversion map $i : G \rightarrow G$ to obtain an inversion operation on infinitesimals. Unfortunately, these linearized operations give us little information about the overlying Lie group. Infinitesimal multiplication turns in addition, and infinitesimal inversion turns into negation

$$m_*(X_e, Y_e) = X_e + Y_e$$

This means that the first order approximation of multiplication and inversion can only tell us about the dimension of the Lie group.

To obtain a richer structure, we have to consider a second order approximation. Take the commutator map $c(g, h) = ghg^{-1}h^{-1}$. The map is trivial around the origin, because $c(g, e) = c(e, h) = e$. This is essentially the reason why linearizing Lie groups is difficult, because all Lie groups are commutative up to first order. However, because $c_* = 0$ for infinitesimals around the origin, we can consider the second derivative $c_{**} : G_e \times G_e \rightarrow G_e$ with

$$c_{**}(X_e, Y_e)(f) = (f \circ c \circ (x, y))''(0)$$

where $x'(0) = X_e$ and $y'(0) = Y_e$. The fact that $c_{**}(X_e, Y_e)$ lies in G_e depends implicitly on the fact that $c_* = 0$. However, unlike c_* , c_{**} is non-trivial, and expresses the non-commutativity of the group G up to second order terms. It is also a Lie bracket on the tangent space at the origin, and is the key to connecting the abstract study of Lie algebras and the analytical study of Lie groups.

Most algebraic concepts about Lie groups have natural ‘infinitesimal formulations’ on Lie algebras, which are easier to understand because vector spaces tend to have a simpler structure, and in many cases we can recover the global concept from the infinitesimal formulation. Because of the deep entanglement between a Lie group and its Lie algebra (often physicists do not even distinguish the two structures), we normally denote a Lie algebra corresponding to a Lie group by the ‘frakturized’ name of the group. As a first instance of this correspondence, we find that we can identify the Lie algebra corresponding to any Lie subgroup H of a Lie group G with a Lie subalgebra \mathfrak{h} of \mathfrak{g} . Conversely, any Lie subalgebra \mathfrak{h} of \mathfrak{g} corresponds to a unique connected Lie subgroup H of G which has \mathfrak{h} as its Lie subalgebra. Properties of \mathfrak{g} thereby rise up to properties of G .

Example. The Lie group $GL_n(K)$ of invertible matrices with coefficients in K gives our first example of a Lie algebra. The tangent space at the origin can be identified with the space of all matrices, where $M \in M_n(K)$ corresponds to the curve $I + tM$. Now using the expansion $(I + M)^{-1} = \sum_{k=0}^{\infty} (-1)^k M^k$, which

holds for $\|M\| \leq 1$, we find

$$\begin{aligned}
& (I + tM)(I + tN)(I + tM)^{-1}(I + tN)^{-1} \\
&= (I + t(M + N) + t^2MN)(I + t(M + N) + t^2NM)^{-1} \\
&= (I + t(M + N) + t^2MN)(I - t(M + N) + t^2[(M + N)^2 - NM]) + O(t^3) \\
&= I + t^2(MN - NM) + O(t^3)
\end{aligned}$$

hence the Lie bracket takes the form $[X, Y] = XY - YX$. In general, the set of invertible endomorphisms $GL(V)$ on a finite dimensional vector space forms a Lie group, and the Lie algebra can be identified with all endomorphisms $T : V \rightarrow V$, and then the Lie bracket takes the form $[T, S] = T \circ S - S \circ T$. A subalgebra of $\mathfrak{gl}(V)$ will be called a **linear Lie algebra**.

Example. In general, if A is any associative algebra over K , then the commutator $[X, Y] = XY - YX$ gives a Lie algebra structure on A , denoted \mathfrak{a} . From the perspective of differential geometry, if A is an arbitrary finite-dimensional algebra, then A has the structure of a differential manifold since it is a vector space, and $U(A)$ is an open neighbourhood of the origin, hence a Lie group. The induced Lie algebra structure on $U(A)$ can then be identified with the Lie bracket on A defined by $[X, Y] = XY - YX$. We will therefore denote the Lie algebra structure on A by $\mathfrak{u}(A)$.

Example. If \mathfrak{g} is any vector space, then the trivial bracket $[X, Y] = 0$ gives \mathfrak{g} a Lie algebra structure. \mathfrak{g} is known as a commutative Lie algebra, because we then find $[X, Y] = [Y, X]$ for all elements X, Y . For Lie algebras over fields of characteristic $\neq 2$, this is the only circumstance in which this can occur. If a Lie group G is abelian, then \mathfrak{g} is abelian, and the converse holds provided G is connected.

Example. On any field K , the space of linear endomorphisms on $K[X]$ forms an associative algebra over K under composition. Of particular interest are the operators

$$f(X) \mapsto g(X)f(X)$$

for $g \in K[X]$, which we shall identify with the polynomial g , and the differentiation operators

$$f(X) \mapsto f'(X)$$

which we denote ∂_X . If we consider the space of ‘differential operators with polynomial coefficients’, operators which can be written in the form

$$\sum_{k=0}^N g_k(X) \partial_X^k$$

with $g_k(X) \in K[X]$, for some N , then we find these form a subalgebra of the space of endomorphisms, since we have the identity

$$(\partial_X X)(f) = X \partial_X f + f \partial_X X = X \partial_X f + f$$

so that $\partial_X X = X \partial_X + 1$. We may use this identity to rearrange the product of any two differential operators to coincide with an operator of the form above. As an algebra, the space of differential operators has essentially no more relations. If

$$\sum_{k=0}^N g_k(X) \partial_X^k = 0$$

Then successively applying the operator to the monomials X^m gives

$$g_0(X) = g_1(X) = \cdots = g_N(X) = 0$$

assuming that we are working over a field of characteristic 0, in which case we have an isomorphism of the ring of operators with $K\langle X, Y \rangle / (YX - XY - 1)$. In this form, the algebra is useful in quantum mechanics, where the operators represent certain non-commutative physical measurements. If a field has finite characteristic p , this method only allows us to determine that

$$g_0(X) = g_1(X) = \cdots = g_{p-1}(X) = 0$$

because $\partial_X^m(X^m) = m! = 0$. But in this scenario we actually find $\partial_X^p = 0$, so if we add this additional relation $Y^p = 0$, we obtain another isomorphism with a quotient of $K\langle X, Y \rangle$. In general, the ring of differential operators in n variables is called the n 'th Weyl algebra. It is obtained from $K\langle X_1, \dots, X_n, Y_1, \dots, Y_n \rangle$ modulo the relations $[Y_i, X_j] = \delta_{ij}$ and $Y_i^p = 0$, which makes sense if we view Y_i as the operator which is partial differentiation in the i 'th variable. The commutator on these algebras gives a particularly interesting Lie algebra structure. We shall find that the Weyl algebra on a single variable is very useful for characterizing the representations of the Lie algebra $\mathfrak{sl}_2(K)$.

Example. If A is an associative algebra over K , then $M_n(A)$ is an algebra over K , and therefore a Lie algebra, denoted $\mathfrak{gl}_n(A)$. A particularly interesting example occurs if $A = \mathbb{C}[X, X^{-1}]$, in which case we call $\mathfrak{gl}_n(A)$ a Loop Lie algebra. This infinite dimensional algebra have applications in various fields of theoretical physics, including String theory.

Example. Given a (possibly non associative) algebra A , a derivation on A is a linear map $d : A \rightarrow A$ satisfying $d(xy) = xd(y) + d(x)y$. Given two derivations d and d' , $d \circ d'$ may not be a derivation, but the commutator

$$[d_1, d_2] = d_1 \circ d_2 - d_2 \circ d_1$$

is always a derivation, because

$$\begin{aligned} (d_1 \circ d_2 - d_2 \circ d_1)(fg) &= d_1(fd_2(g) + d_2(f)g) - d_2(d_1(f)g + fd_1(g)) \\ &= [d_1(f)d_2(g) + f(d_1 \circ d_2)(g) + d_2(f)d_1(g) + (d_1 \circ d_2)(f)g] \\ &\quad - [(d_2 \circ d_1)(f)g + d_1(f)d_2(g) + d_2(f)d_1(g) + f(d_2 \circ d_1)(g)] \\ &= f(d_1 \circ d_2 - d_2 \circ d_1)(g) - (d_1 \circ d_2 - d_2 \circ d_1)(f)g \end{aligned}$$

Thus the set of derivations on A , denoted $\text{der}(A)$, forms a Lie algebra. We should expect the space of derivations to play a fundamental role in the study of Lie algebras, because if X, Y, Z are elements of any Lie algebra, then the Jacobi identity tells us that

$$[X, [Y, Z]] = -[Y, [Z, X]] - [Z, [X, Y]] = [Y, [X, Z]] + [[X, Y], Z]$$

Introducing the **adjoint map** $\text{adj}_X : \mathfrak{g} \rightarrow \mathfrak{g}$ defined by $\text{adj}_X(Y) = [X, Y]$, we can restate the equation above as

$$\text{adj}_X[Y, Z] = [Y, \text{adj}_X Z] + [\text{adj}_X Y, Z]$$

so that adj_X is actually a derivation on \mathfrak{g} . The map $X \mapsto \text{adj}_X$ is actually a representation of \mathfrak{g} on $\text{der}(\mathfrak{g})$, known as the **adjoint representation** because

$$\begin{aligned} [\text{adj}_X, \text{adj}_Y](Z) &= (\text{adj}_X \text{adj}_Y - \text{adj}_Y \text{adj}_X)(Z) \\ &= [X, [Y, Z]] - [Y, [X, Z]] \\ &= [X, [Y, Z]] + [[X, Z], Y] \\ &= [[X, Y], Z] \\ &= \text{adj}_{[X, Y]}(Z) \end{aligned}$$

The kernel of the representation being the centre of \mathfrak{g} ,

$$Z(\mathfrak{g}) = \{X \in \mathfrak{g} : (\forall Y \in \mathfrak{g} : [X, Y] = 0)\}$$

In general, a bilinear skew-symmetric form is a Lie bracket if and only if the adjoint of every element with respect to this bilinear form is a derivation. Thus the Jacobi identity is exactly the derivation equation in disguise, harkening back to the definition of tangent vectors as derivations on the space of C^∞ real-valued functions on the Lie group.

Example. The formula

$$\det(I + tM) = 1 + t \operatorname{tr}(M) + o(t^2)$$

which can be proved by analyzing the Leibnitz formula and ignoring second order terms, tells us that the trace of a matrix determines the rate of area expansion of a shape under a small perturbation by M . The tangent space of $SL_n(K)$ can be identified as a subspace of the tangent bundle on $GL_n(K)$, and since elements of $SL_n(K)$ are defined by the equation $\det(X) = 1$, the set of tangent vectors at the origin are exactly those which annihilate the determinant – i.e., those matrices M such that

$$\left. \frac{d \det(I + tM)}{dt} \right|_{t=0} = 0$$

This implies that the Lie algebra $\mathfrak{sl}_n(K)$ consists of the matrices $M \in \mathfrak{gl}_n(K)$ with trace zero. Algebraically, we can verify that $\mathfrak{sl}_n(K)$ is a Lie subalgebra of $\mathfrak{gl}_n(K)$, because it is closed under the Lie bracket. Noticing the trace identity $\operatorname{tr}(XY) = \operatorname{tr}(YX)$, we find

$$\operatorname{tr}(XY - YX) = \operatorname{tr}(XY) - \operatorname{tr}(YX) = 0$$

Thus $\mathfrak{sl}_n(K)$ is formally a Lie subalgebra of $\mathfrak{gl}_n(K)$.

Example. The orthogonal group $O_n(K)$ consists of matrices in $GL_n(K)$ satisfying $X^t X = I$. Since

$$(I + tX)^t(I + tX) = I + t(X^t + X) + t^2 X^t X$$

X is a tangent vector at the origin for $O_n(K)$ if and only if $X^t = -X$, which causes the first order term of the expansion of $(I + hX)$ to be zero. If $\operatorname{char} K \neq 2$,

then $X^t = -X$ implies that X vanishes along the diagonal, hence $\mathfrak{o}_n(K) \subset \mathfrak{sl}_n(K)$. This essentially follows because the special orthogonal Lie group $SO_n(K)$ of orthogonal matrices of determinant one forms a connected component of $O_n(K)$, hence $\mathfrak{so}_n(K) = \mathfrak{o}_n(K)$. Algebraically, $\mathfrak{o}_n(K)$ is a subalgebra of $\mathfrak{gl}_n(K)$ because if $X, Y \in \mathfrak{o}_n(K)$, then

$$(XY - YX)^t = Y^t X^t - X^t Y^t = YX - XY = -(XY - YX)$$

so the Lie bracket preserves orthogonality.

Example. If n is even, $n = 2m$, then matrices $M \in M_n(K)$ can represent operators from $K^n \times K^n$ to $K^n \times K^n$. Consider the canonical symplectic form ω on $K^n \times K^n$ (a non-degenerate, antisymmetric bilinear form), defined by

$$\omega(x_0 + y_0, x_1 + y_1) = \omega(x_0, y_1) + \omega(y_0, x_1) = \sum (x_0^i y_1^i - y_0^i x_1^i)$$

We let the symplectic group $SP_n(K)$ be the set of linear operators preserving this symplectic form. In matrix form, if we define the matrix $J \in M_n(K)$ by

$$J = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}$$

then $SP_n(K)$ consists exactly of those matrices M such that $M^t J M = J$, or $M^t J M J^t = I$. If we consider a first order approximation at the origin

$$(I + tM)^t J (I + tM) J^t = I + t(M^t + J M J^t) + t^2 M^t J M J^t$$

we see the tangent vectors consist of matrices M such that $M^t = -J M J^t$, or $M^t J + J M = 0$. Together, these matrices the symplectic Lie algebra $\mathfrak{sp}_n(K)$. If we write

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

Then

$$X^t J = \begin{pmatrix} -C^t & A^t \\ -D^t & B^t \end{pmatrix} \quad J X = \begin{pmatrix} C & D \\ -A & -B \end{pmatrix}$$

Implying that M is in $\mathfrak{sp}_n(K)$ if and only if $C^t = C$, $A^t = -D$, and $B^t = B$.

Example. The Heisenberg group $H_n(K)$ is the Lie group of matrices of the form

$$\begin{pmatrix} 1 & a & c \\ 0 & I_n & b \\ 0 & 0 & 1 \end{pmatrix}$$

where $a \in K^n$ is a row vector, $c \in K$, and $b \in K^n$ is a column vector. It's corresponding Lie algebra $\mathfrak{h}_n(K)$ consists of matrices of the form

$$\begin{pmatrix} 0 & a & c \\ 0 & 0_n & b \\ 0 & 0 & 0 \end{pmatrix}$$

which is flat, since the Heisenberg group is essentially flat.

1.1 Factorization

Provided we can characterize the structure of all Lie algebras, we are well on our way to understanding the structure of Lie groups. The standard way of breaking apart some algebraic structure is, like in most algebraic categories, by considering subalgebras and quotients. A homomorphism f of Lie algebras is, of course, a linear map preserving the Lie bracket operation. As with many other algebraic objects, a basic way to understand a Lie algebra is to factor it into two simpler algebras \mathfrak{h} and \mathfrak{k} by considering a short exact sequence

$$0 \rightarrow \mathfrak{k} \rightarrow \mathfrak{g} \rightarrow \mathfrak{h} \rightarrow 0$$

Then we can identify \mathfrak{g} with the vector space $\mathfrak{h} + \mathfrak{k}$ equipped with a Lie bracket formed from a 'twisted product'

$$[(H_0, K_0), (H_1, K_1)] = ([H_0, H_1], [K_0, K_1] + A(H_0, K_1) - A(H_1, K_0) + B(H_0, H_1))$$

for some bilinear $A : \mathfrak{h} \times \mathfrak{k} \rightarrow \mathfrak{k}$ and $B : \mathfrak{h}^2 \rightarrow \mathfrak{k}$. If we are lucky, then we will find $A = B = 0$, in which case we call \mathfrak{g} the direct sum of X and Y , and denote it by $\mathfrak{h} \oplus \mathfrak{k}$. We have then have literally decomposed \mathfrak{g} into simpler Lie algebras.

The standard way to form a short exact sequence for a Lie algebra \mathfrak{g} is to find a (two-sided) ideal \mathfrak{a} , which is a subspace of \mathfrak{g} such that $[X, Y]$ and $[Y, X]$ are both in \mathfrak{a} for any $X \in \mathfrak{a}$, $Y \in \mathfrak{g}$. We can then consider the quotient Lie algebra $\mathfrak{g}/\mathfrak{a}$, and we have an exact sequence

$$0 \rightarrow \mathfrak{a} \rightarrow \mathfrak{g} \rightarrow \mathfrak{g}/\mathfrak{a} \rightarrow 0$$

The other way to form an exact sequence is to consider a surjective homomorphism $f : \mathfrak{g} \rightarrow \mathfrak{h}$, in which case if we let $\mathfrak{k} = \ker f$, then

$$0 \rightarrow \mathfrak{k} \rightarrow \mathfrak{g} \rightarrow \mathfrak{h} \rightarrow 0$$

is exact.

Example. The trace map $tr : \mathfrak{gl}_n(K) \rightarrow K$ is actually a Lie algebra homomorphism, because

$$tr [x, y] = tr(xy - yx) = 0 = [tr(x), tr(y)]$$

The kernel is the special linear group, which is therefore an ideal of the general linear group. Thus we have an exact sequence

$$0 \rightarrow \mathfrak{sl}_n(K) \rightarrow \mathfrak{gl}_n(K) \rightarrow K \rightarrow 0$$

We conclude that $\mathfrak{gl}_n(K)$ is obtained from a twisted product of $\mathfrak{sl}_n(K)$ and K . But since K is a subset of $Z(\mathfrak{gl}_n(K))$, we find that $\mathfrak{gl}_n(K)$ is actually the direct sum of $\mathfrak{sl}_n(K)$ and K , so understanding the structure of the general linear Lie algebra reduces to studying the Lie algebra structure of $\mathfrak{sl}_n(K)$ and K , which are simpler to understand.

As in most algebraic categories, these two processes are essentially the same by the first isomorphism theorem. In fact, we have a version of all standard isomorphism theorems in the category of Lie algebras.

Theorem 1.1 (The First Isomorphism Theorem). *The kernel of a Lie algebra homomorphism $f : \mathfrak{g} \rightarrow \mathfrak{h}$ is an ideal, and the kernel \mathfrak{a} forms an ideal of \mathfrak{g} inducing an injective map $\tilde{f} : \mathfrak{g}/\mathfrak{a} \rightarrow \mathfrak{h}$, such that*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{f} & \mathfrak{h} \\ \downarrow & \nearrow \tilde{f} & \\ \mathfrak{g}/\mathfrak{a} & & \end{array}$$

commutes.

Theorem 1.2 (The Second Isomorphism Theorem). *The set of subalgebras of \mathfrak{h} of $\mathfrak{g}/\mathfrak{a}$ is one to one with the class of subalgebras of \mathfrak{g} containing \mathfrak{a} , and the correspondence maps ideals to ideals, where the subalgebra corresponding to $\mathfrak{a} \subset \mathfrak{h}$ is denoted $\mathfrak{h}/\mathfrak{a}$. If \mathfrak{h} is an ideal, then $(\mathfrak{g}/\mathfrak{a})/(\mathfrak{h}/\mathfrak{a})$ is isomorphic to $\mathfrak{g}/\mathfrak{h}$.*

Theorem 1.3 (The Third Isomorphism Theorem). *If $\mathfrak{a} \subset \mathfrak{b}$ are ideals of \mathfrak{g} , then $\mathfrak{a} + \mathfrak{b}$ and $\mathfrak{a} \cap \mathfrak{b}$ are ideals of \mathfrak{g} , and $(\mathfrak{a} + \mathfrak{b})/\mathfrak{a}$ is isomorphic to $\mathfrak{b}/(\mathfrak{a} \cap \mathfrak{b})$.*

1.2 Solvable Lie Algebras

Recall that a normal subgroup H of a group G is one such that if $g \in G$, $h \in H$, then $ghg^{-1} \in H$. This is equivalent to the fact that $[g, h] \in H$. If H is a Lie subgroup of G , then because $[g, h]$ is the equation which we quadratically approximate to obtain the Lie bracket, we find that if H is a normal Lie subgroup of G , then \mathfrak{h} is an **ideal** of \mathfrak{g} , in the sense that if $X \in \mathfrak{h}$, and $Y \in \mathfrak{g}$, then $[X, Y] \in \mathfrak{h}$. Conversely, if G and H are connected, then the fact that \mathfrak{h} is an ideal implies that H is a normal subgroup of G .

On Lie groups G , we can consider brackets of normal subgroups H ,

$$[H_0, H_1] = \langle h_0 h_1 h_0^{-1} h_1^{-1} : h_0 \in H_0, h_1 \in H_1 \rangle$$

and this subgroup of G will also be normal. Of particular interest in the derived subgroup $[G, G]$, and we can obtain an abelian group $G_{\text{ab}} = G/[G, G]$ by taking the quotient. The corresponding operator for the brackets of normal subgroups are the brackets of ideals in \mathfrak{g}

$$[\mathfrak{a}, \mathfrak{b}] = \text{span}\{[X, Y] : X \in \mathfrak{a}, Y \in \mathfrak{b}\}$$

Analogous to the commutator subgroups of a group is the derived subalgebra $\mathfrak{g}' = D\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, which is the smallest ideal such that $\mathfrak{g}/\mathfrak{g}'$ is abelian, and we call this the **abelianization** of \mathfrak{g} , denoted \mathfrak{g}_{ab} . Since we have the exact diagram

$$0 \rightarrow \mathfrak{g}' \rightarrow \mathfrak{g} \rightarrow \mathfrak{g}_{\text{ab}} \rightarrow 0$$

we can write $\mathfrak{g} = \mathfrak{g}_{\text{ab}} + \mathfrak{g}'$, where

$$[(X_0, Y_0), (X_1, Y_1)] = (0, A(X_0, Y_1) - A(X_1, Y_0) + B(X_0, X_1))$$

for some bilinear A and B . Thus we can think of the elements of \mathfrak{g}' as infinitesimals, since they have no impact on the Lie bracket structure of \mathfrak{g}_{ab} , existing somewhat ‘beneath the surface’ of the calculations. If we consider $\mathfrak{g}'' = (\mathfrak{g}')'$, then we can compute the abelian approximation ‘to a second order’, writing $\mathfrak{g} = \mathfrak{g}_{\text{ab}} + \mathfrak{g}'_{\text{ab}} + \mathfrak{g}''$. Continuing this process, we hope to write \mathfrak{g} as the successive product of abelian infinitesimals,

$$\mathfrak{g} = \mathfrak{g}_{\text{ab}}^{(1)} + \mathfrak{g}_{\text{ab}}^{(2)} + \cdots + \mathfrak{g}_{\text{ab}}^{(n)}$$

where $\mathfrak{g}^{(n)}$ is the n 'th element of the **derived series**. For this to work, we require that $\mathfrak{g}^{(n+1)} = 0$ for some n , in which case we say \mathfrak{g} is a **solvable** Lie

algebra. This is equivalent to the existence of an abelian tower, that is, a decreasing chain $\mathfrak{g}_1 \supset \mathfrak{g}_2 \supset \cdots \supset \mathfrak{g}_n$, with $\mathfrak{g}_1 = \mathfrak{g}$, $\mathfrak{g}_n = 0$, each \mathfrak{g}_{n+1} an ideal in \mathfrak{g}_n , and $\mathfrak{g}_n/\mathfrak{g}_{n+1}$ an abelian Lie algebra.

It is also natural to consider the **lower central series**

$$\mathfrak{g}_1 = \mathfrak{g} \quad \mathfrak{g}_2 = [\mathfrak{g}, \mathfrak{g}] \quad \mathfrak{g}_3 = [\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]] \quad \dots$$

where the derived series is slightly finer, $\mathfrak{g}^{(n)} \subset \mathfrak{g}_n$, so that the series doesn't penetrate deep into the group structure, yet the 'new infinitisimals' have the additional property that if $X \in \mathfrak{g}_n$ and $Y \in \mathfrak{g}_m$, then $[X, Y] \in \mathfrak{g}_{n+m}$ (for the derived series, we can only guarantee that $[X, Y] \in \mathfrak{g}^{(m+1)}$ if $X \in \mathfrak{g}^{(m+1)}$ and $Y \in \mathfrak{g}^{(m)}$). If the lower central series eventually terminates, we call the algebra **nilpotent**. We then have a decomposition

$$\mathfrak{g} = (\mathfrak{g}_1/\mathfrak{g}_2) \times (\mathfrak{g}_2/\mathfrak{g}_3) \times \cdots \times (\mathfrak{g}_n/\mathfrak{g}_{n+1})$$

Of course, this is a very strong condition for a Lie algebra to have.

Example. The Lie group $UT_n(K)$ of unitriangular matrices (upper triangular matrices with ones on the diagonal) is a Lie group, with corresponding Lie algebra $\mathfrak{ut}_n(K)$ consisting of strictly upper triangular matrices. $\mathfrak{ut}_n(K)$ is nilpotent, because any elements X in $[\mathfrak{ut}_n(K)]_m$ has $X_{ij} = 0$ for $i < j + m$, so $[\mathfrak{ut}_n(K)]_n = 0$.

If \mathfrak{a} and \mathfrak{b} are two solvable ideals in a Lie algebra \mathfrak{g} , then $\mathfrak{a} + \mathfrak{b}$ is a solvable ideal, since $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b}$ is isomorphic to $\mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b})$, which is solvable as a quotient of a solvable ideal, and \mathfrak{b} is solvable. As a corollary to this, we see that any Lie algebra \mathfrak{g} has a maximum solvable ideal (add up all of the solvable ideals). We shall denote the maximum solvable ideal of a Lie algebra by $\text{rad}(\mathfrak{g})$, and call it the **radical** of the algebra. The radical essentially separates the approximately commutative section of the algebra from the non-commutative section. The most non-commutative Lie algebras are the **semi-simple** ones, such that $\text{rad}(\mathfrak{g}) = (0)$. For any algebra \mathfrak{g} , $\mathfrak{g}/\text{rad}(\mathfrak{g})$ is semisimple, so the radical efficiently extracts the commutative section of the algebra. It is an essential tool in the description of the finite dimensional Lie algebras, because it means we can write any Lie algebra as the product of a semisimple algebra and a solvable one. The **Levi decomposition** says that this product is actually a direct sum, so that any Lie algebra is the direct sum of a solvable algebra and a semisimple

algebra. Nilpotency does not have as strong a decomposition property, so that it is less useful in the classification problem. If \mathfrak{a} and \mathfrak{b} are nilpotent ideals, then $\mathfrak{a} + \mathfrak{b}$ is nilpotent, so we can consider the maximal nilpotent ideal, called the nilradical and denoted $\text{nil}(\mathfrak{g})$, and we can consider the short exact sequence to the quotient

$$0 \rightarrow \text{nil}(\mathfrak{g}) \rightarrow \mathfrak{g} \rightarrow \mathfrak{g}_{\text{red}}$$

where $\mathfrak{g}_{\text{red}}$ is known as the **reduced algebra** of \mathfrak{g} . Unfortunately, we do not have a direct sum decomposition, though the sequence is useful at times.

A **simple** Lie algebra is a non-abelian Lie algebra \mathfrak{g} having no ideals other than the trivial ideals (0) and \mathfrak{g} . We shall soon prove that semi-simple Lie algebras break down into the direct sum of simple Lie algebras, so that we need only study the simple Lie algebras to understand the semisimple ones. The classification theorem for simple Lie algebras over the complex numbers breaks down the families of simple Lie algebras into the classical matrix algebras \mathfrak{sl}_n , \mathfrak{so}_n , \mathfrak{sp}_n , and some ‘eccentric’ algebras \mathfrak{e}_6 , \mathfrak{e}_7 , \mathfrak{e}_8 , \mathfrak{f}_4 , and \mathfrak{g}_2 . The classification is very strong, because the matrix algebras are very concrete, and we can understand them relatively easily. In our journey into finding this classification, we will require the sophisticated theorems of representation theory.

1.3 Low Dimensional Classifications

There is a useful criterion for verifying some abstractly defined bracket is actually a Lie bracket, which will enable us to classify the low dimensional Lie algebras. Given a vector space \mathfrak{g} , after fixing a basis E_1, \dots, E_n , we find an arbitrary bracket (not necessarily a Lie bracket) $[\cdot, \cdot] : \mathfrak{g} \rightarrow \mathfrak{g}$ can be written as

$$[X, Y] = \sum f_k(X, Y) E_k$$

where each f_i is a bilinear form, and so we may find constants $a_{ij}^k \in K$ such that

$$f_k = \sum a_{ij}^k (E_i^* \otimes E_j^*)$$

which is the same as saying $[E_i, E_j] = \sum a_{ij}^k E_k$. If $[\cdot, \cdot]$ was actually a Lie bracket, then the alternating property of the Lie algebras implies that each

f_k is a skew-symmetric Bilinear form, hence $a_{ij}^k = -a_{ji}^k$. The Jacobi identity takes the form

$$\sum_l a_{jk}^l a_{il}^m + a_{ki}^l a_{jl}^m + a_{ij}^l a_{kl}^m = 0$$

for every fixed i, j, k, m . Conversely, given a set of constants a_{ij}^k over some vector space satisfying these equations, the corresponding bilinear map is a Lie bracket, because if $X = \sum a^i e_i$, $Y = \sum b^i e_i$, and $Z = \sum c^i e_i$, then

$$\begin{aligned} & [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] \\ &= \sum_{i,j,l} a^i b^j c^l ([e_i, [e_j, e_l]] + [e_j, [e_l, e_i]] + [e_l, [e_i, e_j]]) = 0 \end{aligned}$$

Thus Lie algebras and structural constants over a particular set are in one-to-one correspondence.

Example. *It will be useful to know the structural constants of $\mathfrak{sl}_2(K)$ with respect to the basis*

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

To calculate the constants, we just calculate the Lie bracket between distinct elements of the basis

$$[e, f] = h \quad [e, h] = -2e \quad [f, h] = 2f$$

All other structure constants not calculate are zero.

We rarely define Lie algebras by specifying their structural constants, but the coordinatization process is particular useful to understand low dimensional Lie algebras, where the restriction on coefficients enable us to relatively easily classify the algebras up to isomorphism. Note that two algebras \mathfrak{g} and \mathfrak{h} are isomorphic if and only if we can specify a pair of bases in such a way that the structure constants obtained are the same. In the one dimensional case, the only skew symmetric matrix is trivial, so that every one dimensional Lie algebra is abelian. The calculation of the two and three dimensional cases gets progressively more and more involved.

Before we do this though, we should mention a little trick to help verify the Jacobi identity for certain abstract Lie brackets. Given a bilinear map $\omega : \mathfrak{g}^2 \rightarrow \mathfrak{g}$, we consider the trilinear map $T\omega : \mathfrak{g}^3 \rightarrow \mathfrak{g}$ defined by

$$(T\omega)(X, Y, Z) = \omega(X, \omega(Y, Z)) + \omega(Y, \omega(Z, X)) + \omega(Z, \omega(X, Y))$$

Then $T\omega = 0$ if and only if ω satisfies the Jacobi identity. By linearity, each coefficient of $\omega(X, \omega(Y, Z))$ can be written as a polynomial in the variables X^i , Y^j , and Z^k ,

$$\omega(X, \omega(Y, Z)) = \sum b_{ijk}^l X^i Y^j Z^k E_l$$

and so

$$T\omega = \sum (b_{ijk}^l + b_{jki}^l + b_{kij}^l) X^i Y^j Z^k E_l$$

Applying the theory of multivariate polynomials over a field, we see that to verify the Jacobi identity, it suffices to show that the coefficient corresponding to each coefficient vanishes. This is a necessary condition for the Jacobi identity to hold over a field of characteristic zero. In most cases, we will find that only two of the three coefficients will be non-zero, and they will be the negation of one another. Thus we can verify the Jacobi identity by ‘pairing cycles’ in the coefficients of $\omega(X, \omega(Y, Z))$.

Theorem 1.4. *There is a single nonabelian two dimensional Lie algebra.*

Proof. If \mathfrak{g} is a two dimensional Lie algebra, if we fix a basis $\{X, Y\}$, then $[X, Y]$ spans \mathfrak{g}' , hence if \mathfrak{g} is nonabelian, \mathfrak{g}' is one dimensional. Fix a non-zero $X \in \mathfrak{g}'$, and extend X to a basis $\{X, Y\}$. Then $[X, Y] \neq 0$, for otherwise the bracket is abelian. Write $[X, Y] = \lambda X$, for some $\lambda \neq 0$. By scaling Y , we may actually assume $[X, Y] = X$. This shows that there can be at most one non-abelian Lie algebra, because we have found a basis with a particular set of structural constants, and provided these constants specify a Lie algebra, we have determined that there is a single two dimensional Lie algebra. The bracket $[X, Y] = X$ does actually give a Lie algebra structure, since

$$[a_1 X + a_2 Y, [b_1 X + b_2 Y, c_1 X + c_2 Y]] = (a_2 b_2 c_1 - a_2 b_1 c_2) X$$

And $a_2 b_2 c_1$ is obtained from $a_2 b_1 c_2$ by a cycle permutation in the variables a, b , and c , hence the Jacobi holds. \square

As we increase the dimension of our Lie algebras, the classification becomes more and more involved. We shall be able to use structural constants to classify the three dimensional Lie algebras, but the technique becomes infeasible for dimensions much higher than this. We split our discussion of three dimensional Lie algebras into three cases, where the derived subalgebra has dimension 0, 1, 2, and 3. The zero dimensional case is obviously the abelian Lie algebra, and needs no further discussion.

Theorem 1.5. *The Heisenberg Lie algebra is the unique three dimensional Lie algebra whose derived subalgebra is one dimensional, and contained within the center of the algebra.*

Proof. Given such a Lie algebra \mathfrak{g} , take X and Y such that $[X, Y] = Z \neq 0$. Then X , Y , and Z are linearly independent, for if $aX + bY + cZ = 0$, then

$$[aX + bY + cZ, X] = b[Y, X] = -bZ = 0 \quad [aX + bY + cZ, Y] = aZ = 0$$

implying $a = b = 0$, hence $c = 0$. Thus X, Y , and Z are a basis of the Lie algebra, and we have specified the structural constants exactly. These constants are identical to the structural constants of the Heisenberg Lie algebra, where $X = E_{12}$, $Y = E_{23}$, and $Z = E_{33}$. \square

There is one more three dimensional Lie algebra whose derived subalgebra is one dimensional. If we let \mathfrak{g} denote the nonabelian two-dimensional Lie algebra, then $\mathfrak{g} \oplus K$ is a Lie algebra with $(\mathfrak{g} \oplus K)' = \mathfrak{g}' \oplus 0$. Thus the derived subalgebra is one dimensional. However, the derived subalgebra is not contained within the centre of the algebra, because $Z(\mathfrak{g} \oplus K) = Z(\mathfrak{g}) \oplus Z(K) = 0 \oplus K$. We shall now prove that this is the defining three dimensional Lie algebra whose derived subalgebra is one-dimensional, and is not contained within the centre.

Theorem 1.6. *There is a unique three dimensional Lie algebra with one dimensional derived subalgebra not contained within the center of the algebra.*

Proof. Let \mathfrak{h} be such an algebra. Pick $X \neq 0$ spanning the derived \mathfrak{h}' , then pick Y with $[X, Y] = X$. The subalgebra of \mathfrak{h} spanned by X and Y is isomorphic to \mathfrak{g} . To obtain a full isomorphism with $\mathfrak{g} \oplus K$, it suffices to find $Z \in Z(\mathfrak{g})$ which is linearly independent from X and Y . Pick some Z to form a basis $\{X, Y, Z\}$. There must be $a, b \in K$ such that $[X, Z] = aX$, $[Y, Z] = bX$. If we consider the equations

$$\begin{aligned} [\lambda X + \gamma Y + \eta Z, X] &= -(\gamma + \eta a)X \\ [\lambda X + \gamma Y + \eta Z, Y] &= (\lambda - \eta b)X \\ [\lambda X + \gamma Y + \eta Z, Z] &= (a\lambda + b\gamma)X \end{aligned}$$

And since the matrix

$$\begin{pmatrix} 0 & -1 & -a \\ 1 & 0 & -b \\ a & b & 0 \end{pmatrix}$$

has determinant zero, the matrix is non-invertible, and the nullspace therefore contains some non-zero (λ, γ, η) , i.e. $Z(\mathfrak{g})$ is non trivial, and contains some non-zero $Z' = \lambda X + \gamma Y + \eta Z$. If $\eta = 0$, then the equations above would imply $\gamma = \lambda = 0$, which is impossible, hence $\eta \neq 0$, and we have found a nonzero $Z' \in Z(\mathfrak{h})$ not contained in the span of X or Y , hence we have the decomposition

$$\mathfrak{h} = \langle X, Y \rangle \oplus KZ' \cong \mathfrak{g} \oplus K$$

Hence $\mathfrak{g} \oplus K$ is the unique three dimensional Lie algebra whose derived subalgebra is one dimensional, and isn't contained in the centre of the algebra. \square

There is actually an infinite family of non-isomorphic Lie algebras whose derived algebra is two-dimensional in the case where K is a algebraically complete field, but we have a nice classification of this family. We take a basis $\{Y, Z\}$ of the derived subalgebra, and complete it to a basis $\{X, Y, Z\}$ of the entire algebra. We claim that the derived subalgebra is always abelian. To prove this, it suffices to show $[Y, Z] = 0$. Consider the adjoint map adj_Y , which can be written in matrix form with respect to the basis given as

$$\begin{pmatrix} 0 & 0 & 0 \\ * & 0 & \alpha \\ * & 0 & \beta \end{pmatrix}$$

for some constants α and β . We claim that the trace of any adjoint operator from a derived element is zero, so that $\beta = 0$. By consider the adjoint map adj_Z , we conclude that $\alpha = 0$. Thus the derived subalgebra is abelian.

Lemma 1.7. *For any Lie algebra \mathfrak{g} , and $X \in \mathfrak{g}'$, $\text{tr}(\text{adj}_X) = 0$.*

Proof. If $X = [Y, Z]$, then

$$\text{adj}_{[Y, Z]} = [\text{adj}_Y, \text{adj}_Z] \in \mathfrak{gl}(\mathfrak{g})' = \mathfrak{sl}(\mathfrak{g})$$

Hence the trace the adjoint is zero. \square

In fact, adj_X is an isomorphism of \mathfrak{g}' , because $[X, Y]$ and $[X, Z]$ span the derived subalgebra, hence the map is surjective, and therefore an isomorphism. Over K , we may apply the theory of the Jordan normal form, and break our derivation down into two cases.

1. If we can choose Y and Z such that $[X, Y] = \lambda Y$, $[X, Z] = \gamma Z$, then by scaling, we may assume $\lambda = 1$. For a fixed γ , this does indeed define a Lie algebra structure on K^3 , denoted \mathfrak{g}_γ , because

$$\begin{aligned} & [a_1X + a_2Y + a_3Z, [b_1X + b_2Y + b_3Z, c_1X + c_2Y + c_3Z]] \\ &= [a_1X + a_2Y + a_3Z, (b_1c_2 - b_2c_1)Y + \gamma(b_1c_3 - b_3c_1)Z] \\ &= (a_1b_1c_2 - a_1b_2c_1)Y + \gamma^2(a_1b_1c_3 - a_1b_3c_1)Z \end{aligned}$$

and $a_1b_1c_2$ is obtained from $a_1b_2c_1$, and $a_1b_1c_3$ is obtained from $a_1b_3c_1$ by a cycle permutation, hence the Jacobi identity holds. These form a large class of distinct Lie algebras. We shall find that \mathfrak{g}_λ is isomorphic to \mathfrak{g}_γ if and only if $\lambda = \gamma$ or $\lambda = 1/\gamma$.

First note that the map $X \mapsto \lambda X$, $Y \mapsto Z$, $Z \mapsto Y$ is an explicit isomorphism between \mathfrak{g}_λ and \mathfrak{g}_γ if $\lambda = 1/\gamma$. Furthermore, if $f : \mathfrak{g}_\lambda \rightarrow \mathfrak{g}_\gamma$ is an isomorphism, it must map $Z(\mathfrak{g}_\lambda)$ to $Z(\mathfrak{g}_\gamma)$, and therefore in the basis X, Y, Z f takes the form

$$\begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

with $a_{11} \neq 0$, $a_{22}a_{33} - a_{32}a_{23} \neq 0$. Yet in \mathfrak{g}_γ ,

$$\begin{aligned} a_{22}Y + a_{32}Z &= [a_{11}X + a_{12}Y + a_{13}Z, a_{22}Y + a_{32}Z] = a_{11}a_{22}Y + a_{11}a_{32}\gamma Z \\ \lambda(a_{23}Y + a_{33}Z) &= [a_{11}X + a_{12}Y + a_{31}Z, a_{23}Y + a_{33}Z] = a_{11}a_{23}Y + a_{11}a_{33}\gamma Z \end{aligned}$$

and since f is a Lie algebra homomorphism, this calculation gives

$$(a_{11} - 1)a_{22} = (\gamma a_{11} - 1)a_{32} = a_{23}(\lambda - a_{11}) = a_{33}(\lambda - \gamma a_{11}) = 0$$

If $a_{22} = 0$ or $a_{33} = 0$, a_{23} and a_{32} are both non-zero, hence $a_{11} = 1/\gamma = \lambda$, hence $\lambda = 1/\gamma$. If a_{22} and a_{33} are both non-zero, then $a_{11} = 1 = \lambda/\gamma$, hence $\lambda = \gamma$.

2. Suppose our operator is not diagonalizable. Since we are working over a complete field, we have a Jordan normal form. That is, there exists an eigenvalue λ , and we may extend X to a basis $\{X, Y, Z\}$ such that $[X, Y] = \lambda Y + Z$, $[X, Z] = \lambda Z$. If we replace X with X/λ , and Z with Z/λ , then $[X, Y] = Y + Z$, and $[X, Z] = Z$, and these define a

specific set of structural constants, which actually give a Lie algebra structure. Indeed, then

$$\begin{aligned}
& [a_1X + a_2Y + a_3Z, [b_1X + b_2Y + b_3Z, c_1X + c_2Y + c_3Z]] \\
&= [a_1X + a_2Y + a_3Z, (b_1c_2 - b_2c_1)Y + (b_1c_3 - b_3c_1 + b_1c_2 - b_2c_1)Z] \\
&= (a_1b_1c_2 - a_1b_2c_1)(Y + Z) + (a_1b_1c_3 - a_1b_3c_1 + a_1b_1c_2 - a_1b_2c_1)Z \\
&= (a_1b_1c_2 - a_1b_2c_1)Y + (2a_1b_1c_2 - 2a_1b_2c_1 + a_1b_1c_3 - a_1b_3c_1)Z
\end{aligned}$$

and we have cycle pairs, hence the Jacobi identity holds.

Finally, suppose that \mathfrak{g} is a Lie algebra such that $\mathfrak{g}' = \mathfrak{g}$. We already know such an algebra exists over fields of characteristic $\neq 2$, $\mathfrak{sl}_2(K)$, because if we consider the basis of the algebra $X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, and $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, then

$$[X, Y] = Z \quad [X, Z] = -2X \quad [Y, Z] = 2Y$$

which span $\mathfrak{sl}_2(K)$. We shall find this is the only such algebra up to isomorphism.

Theorem 1.8. *If \mathfrak{g} is a three dimensional Lie algebra over an algebraically complete field with $\mathfrak{g}' = \mathfrak{g}$, then \mathfrak{g} is isomorphic to $\mathfrak{sl}_2(K)$.*

Proof. Given such an algebra, if $X \in \mathfrak{g}$ is non-zero, then adj_X has rank 2, because \mathfrak{g}' is spanned by $[X, Y]$, $[X, Z]$ and $[Y, Z]$, which must be linearly independent, and $[X, Y]$ and $[X, Z]$ span the range of adj_X . This also implies that $[X, M] = 0$, then M is in the span of X . We claim that we can choose X such that adj_X has a non-zero eigenvalue. If all the eigenvalues of X are non-zero, applying the Jordan canonical form theorem we conclude that there is a basis X_0, Y_0 , and Z_0 such that $[X, X_0] = 0$, $[X, Y_0] = X_0$, and $[X, Z_0] = Y_0$. By rescaling X , we may actually assume $X = X_0$, in which case $[Y_0, X_0] = -[X_0, Y_0] = -X_0$, so that adj_{Y_0} has a non-zero eigenvalue.

If adj_X has one non-zero eigenvalue, it must actually have two non-zero eigenvalues which are the negations of one another (because the trace of the adjoint is zero), and therefore adj_X is diagonalizable. Thus we extend X to a basis X, Y, Z with $[X, Y] = \lambda Y$, $[X, Z] = -\lambda Z$, and X, Y , and Z are a basis of the space. To fully describe the structure of the algebra, we need to determine $[Y, Z]$. Note that

$$[X, [Y, Z]] = [[Z, X], Y] + [[X, Y], Z] = \lambda[Z, Y] + \lambda[Y, Z] = 0$$

Which implies $[Y, Z]$ is a non-zero scalar multiple of X . We may assume that $X = [Y, Z]$ by rescaling Y . Finally, by rescaling X , we can change the value of λ to an arbitrary value, and this shows that the Lie algebra structure is unique, because we have specified all structural constants exactly. For $\lambda = 2$, we obtain the special linear group constants we found above. \square

Chapter 2

Matrix Lie Algebras

An important class of Lie algebras are those which occur as subalgebras of $\mathfrak{gl}(V)$, for some finite dimensional vector space V . This is not only an important class of examples of Lie algebras, but provides a suitable training ground for the understanding of the representation theory of Lie algebras – the idea being that we can completely characterize a Lie algebra by its actions on vector spaces. The most important application of this chapter is by considering the adjoint representation of a Lie algebra \mathfrak{g} on $\mathfrak{gl}(\mathfrak{g})$, where X corresponds to adj_X . This representation even has important applications in the study of concrete matrix Lie algebras.

Theorem 2.1. *If $T : V \rightarrow V$ is diagonalizable, then $\text{adj}_T : \mathfrak{gl}(V) \rightarrow \mathfrak{gl}(V)$ is diagonalizable.*

Proof. Let v_1, \dots, v_n be a basis of eigenvectors of V , with $Tv_i = \lambda_i v_i$. The matrix representation of T with respect to this basis is precisely the diagonal matrix $\text{diag}(v_1, \dots, v_n)$, and

$$[\text{diag}(v_1, \dots, v_n), E_{jk}] = \sum \lambda_i [E_{ii}, E_{jk}] = (\lambda_j - \lambda_k) E_{jk}$$

so the operators $S_{jk}(v_i) = \delta_i^j v_k$ diagonalize the adjoint. \square

Theorem 2.2. *If $T : V \rightarrow V$ is nilpotent, then $\text{adj}_T : \mathfrak{gl}(V) \rightarrow \mathfrak{gl}(V)$ is nilpotent.*

Proof. We have an expansion

$$\text{adj}_T^n(X) = \sum (-1)^k \binom{n}{k} T^k X T^{n-k}$$

If $T^n = 0$, then $\text{adj}_T^{2n}(X) = 0$ for all X , because if $k < n$, then $2n - k > n$, so that each term $T^k X T^{2n-k}$ vanishes. \square

It is an easy mistake to assume the converse of this theorem, which is not necessarily true. For instance, the identity matrix in $\mathfrak{gl}(V)$ is not nilpotent, yet $\text{adj}_I = 0$ is certainly nilpotent. Thus we must distinguish between whether a linear operator is nilpotent, or whether its adjoint is nilpotent on the linear Lie algebra it is contained in.

Eigenvectors and eigenvalues are incredibly important to the classification of linear operators over a finite dimensional vector space. They also play an important part in understanding the action of a Lie algebra on a vector space. Given a Lie algebra \mathfrak{g} acting on a vector space V , we define an eigenvector of \mathfrak{g} to be a vector $v \in V$ such that Xv is a scalar multiple of v for each $X \in \mathfrak{g}$. The scalar multiple may differ depending on the X we choose. For instance, the eigenvectors for the subalgebra of diagonal matrices in $\mathfrak{gl}_n(K)$ consist of exactly the basis vectors e_i . We think of the eigenvectors as a tool to simultaneously diagonalize all the operators in a subalgebra at once. Dual to the concept of generalized eigenvectors are a sort of ‘generalized eigenvalue’. If v is an eigenvector for \mathfrak{g} , then we may define a function $\lambda : \mathfrak{g} \rightarrow K$ so that the equation $Xv = \lambda(X)v$ holds for all X . λ is a linear map, because

$$\lambda(X + Y)v = (X + Y)v = Xv + Yv = [\lambda(X) + \lambda(Y)]v$$

We call such a map a **weight** for \mathfrak{g} . In general, for a linear functional $\lambda \in \mathfrak{g}^*$, we define V_λ to be the set of vectors v such that $Xv = \lambda(X)v$ holds for all $X \in \mathfrak{g}$. A weight is precisely a linear functional for which V_λ is non-trivial.

Example. As we have shown, if \mathfrak{g} is the subalgebra of diagonal matrices in $\mathfrak{gl}_n(K)$, then the eigenvectors consist of the basis vectors e_i , and the corresponding weights are precisely the weights $\varepsilon_i(X) = X_{ii}$, because $Xe_i = X_{ii}e_i$. If the eigenvectors of a matrix subalgebra \mathfrak{g} over V form a basis of V , then \mathfrak{g} is isomorphic to some family of diagonal matrices over V .

The weight spaces have nice invariance properties which, like for the Jordan normal form, allow us to decompose matrix Lie algebras into eigenspaces.

Lemma 2.3. If \mathfrak{a} is an ideal of a subalgebra \mathfrak{g} on $\mathfrak{gl}(V)$, the space of vectors $v \in V$ such that $Xv = 0$ for all $X \in \mathfrak{a}$ is \mathfrak{g} -invariant.

Proof. If $Xv = 0$ for all $X \in \mathfrak{a}$, then for any $Y \in \mathfrak{g}$, $[X, Y] \in \mathfrak{a}$, so $[X, Y]v = 0$, hence

$$XYv = YXv = Y0 = 0$$

and this exactly shows \mathfrak{g} -invariance. \square

We have shown that the set of eigenvectors for \mathfrak{a} with weight 0 form a \mathfrak{g} invariant subspace. We can generalize this result to the case where we have a family of eigenvectors of the same weight, where the weight might not necessarily be equal to zero.

Theorem 2.4. *Over a field of characteristic zero, if \mathfrak{g} is a subalgebra of $\mathfrak{gl}(V)$ with ideal \mathfrak{a} , then the weight space of any $\lambda : \mathfrak{a} \rightarrow K$ is \mathfrak{g} invariant.*

Proof. If $v \in V$ satisfies $Xv = \lambda(X)v$ for all $X \in \mathfrak{a}$, then we must show $XYv = \lambda(X)Yv$ for any $Y \in \mathfrak{g}$. Note that

$$XYv = YXv + [X, Y]v = \lambda(X)Yv + \lambda[X, Y]v$$

Thus it suffices to verify that $\lambda[X, Y] = 0$ for any $X \in \mathfrak{a}$, and $Y \in \mathfrak{g}$. Consider the span of $\{v, Yv, \dots, Y^n v\}$. We claim that for any $Z \in \mathfrak{a}$,

$$ZY^n v = \lambda(Z)(Y^n v) + \text{span}(v, Yv, \dots, Y^{n-1} v)$$

For $n = 0$, the claim is assumed. In general, since $[Z, Y] \in \mathfrak{a}$, we have

$$\begin{aligned} ZY^n v &= YZY^{n-1} v + [Z, Y]Y^{n-1} v \\ &= \lambda(X)Y^n v + [Z, Y]Y^{n-1} v + \text{span}(v, Yv, \dots, Y^{n-2} v) \\ &= \lambda(X)Y^n v + \text{span}(v, Yv, \dots, Y^{n-1} v) \end{aligned}$$

The matrix representation of adj_Z with respect to this basis has trace $n\lambda(Z)$. If we let $Z = [X, Y]$, then we find $n\lambda[X, Y] = 0$, hence $\lambda[X, Y] = 0$. \square

Here's a simple consequence about matrix algebras over an algebraically closed field of characteristic zero.

Theorem 2.5. *If V is a vector space over an algebraically closed field of characteristic zero, and $X, Y \in \mathfrak{gl}(V)$ commute with $[X, Y]$, then $[X, Y]$ is nilpotent.*

Proof. It suffices to show that the only eigenvalue of $[X, Y]$ is 0. Let λ be an eigenvalue of $[X, Y]$, and let \mathfrak{g} be the subalgebra of $\mathfrak{gl}(V)$ generated by X , Y , and $[X, Y]$. Then the span of $[X, Y]$ is a Lie algebra ideal of \mathfrak{g} , hence the space V_λ of vectors v satisfying $[X, Y]v = \lambda v$ is \mathfrak{g} invariant. This implies that X and Y restrict to operators on V_λ , and the restriction of $[X, Y]$ to V_λ is the same as $XY - YX$, hence $[X, Y]$ has trace zero. But $[X, Y]$ is diagonalizable on V_λ , with trace $n\lambda$ if V_λ is n -dimensional, hence $\lambda = 0$. \square

Another consequence tells us about the normalizer of some subalgebra \mathfrak{h} of \mathfrak{g} , which is the largest subalgebra of \mathfrak{g} in which \mathfrak{h} is an ideal, and can be defined as

$$N_{\mathfrak{g}}(\mathfrak{h}) = \{X \in \mathfrak{g} : [X, \mathfrak{h}] \subset \mathfrak{h}\}$$

or just $N(\mathfrak{h})$ for short. This is easily verified to be a subalgebra, because if $X, Y \in N(\mathfrak{h})$, then $X + Y \in N(\mathfrak{h})$, and for any $Z \in \mathfrak{h}$,

$$[[X, Y], Z] = [[X, Z], Y] + [X, [Y, Z]] \in \mathfrak{h}$$

We can measure how much a subalgebra fails to be an ideal by considering how large $N(\mathfrak{h})$ is.

Theorem 2.6. *If \mathfrak{d} is the subalgebra of $\mathfrak{gl}_n(K)$ consisting of diagonal matrices, and if K has characteristic zero, then $N(\mathfrak{d}) = \mathfrak{d}$.*

Proof. Since \mathfrak{d} is an ideal of $N(\mathfrak{d})$, the invariance lemma tells us that the vector spaces KE_{ii} is invariant under $N(\mathfrak{d})$. If we fix some $X \in N(\mathfrak{d})$, then we find that $XE_{ii} = \sum X_{ji}E_{ji}$ is a multiple of E_{ii} , hence $X_{ji} = 0$ for $j \neq i$, and performing this process over all i , we conclude X is diagonal. \square

2.1 Engel's Theorem

In linear algebra, we form classification theorems for linear operators. The theorems show the existence of certain basis elements on the vector space, such that the corresponding matrix representation of the operators have nice structure. But given a family of linear operators, it is a much more difficult problem to find a basis such that *all* the matrix representations of the family of linear operators have nice structure. Any nilpotent operator

on a vector space has a basis with respect to which the matrix representation is upper triangular. We would like to know which families of nilpotent linear operators can be simultaneously ‘strictly upper triangularized’. Also natural is to know which families of linear operators can be simultaneously ‘upper triangularized’. In the case that these families form a Lie subalgebra of operators, then these results essentially constitute Engel’s theorem and Lie’s theorem.

Lemma 2.7. *If \mathfrak{g} is a Lie subalgebra of nilpotent transformations on a finite dimensional vector space V , and $V \neq 0$, then there is $v \neq 0$ for which $Xv = 0$ for all $X \in \mathfrak{g}$.*

Proof. We proceed by induction on the dimension of \mathfrak{g} . If \mathfrak{g} is one-dimensional, spanned by some $X \neq 0$, then there is n such that $X^n \neq 0$, $X^{n+1} = 0$, and if $X^n v \neq 0$, then $X(X^n v) = 0$, so v satisfies the theorem.

In general, let \mathfrak{a} be a maximal proper Lie subalgebra of \mathfrak{g} . For any $X \in \mathfrak{a}$, the adjoint map $\text{adj}_X : \mathfrak{g} \rightarrow \mathfrak{g}$ descends to a map $\text{adj}_X : \mathfrak{g}/\mathfrak{a} \rightarrow \mathfrak{g}/\mathfrak{a}$, because if $Y - Z \in \mathfrak{a}$, then $[X, Y] - [X, Z] = [X, Y - Z] \in \mathfrak{a}$. Note that if X is nilpotent, then $\text{adj}_X : \mathfrak{g} \rightarrow \mathfrak{g}$ is nilpotent, since

$$\text{adj}_X^m(Y) = \sum_{k=0}^m (-1)^k \binom{m}{k} (X^k Y X^{m-k})$$

If $X^n = 0$, then $\text{adj}_X^m(Y) = 0$ when $m \geq 2n$. Since the family adj_X , viewed as endomorphisms of $\mathfrak{g}/\mathfrak{a}$ form a subalgebra of nilpotent transformations on $\mathfrak{gl}(\mathfrak{g}/\mathfrak{a})$ with dimension less than \mathfrak{g} , we conclude that there is $Y + \mathfrak{a} \in \mathfrak{g}/\mathfrak{a}$ with $Y \notin \mathfrak{a}$, such that $[X, Y] \in \mathfrak{a}$ for all $X \in \mathfrak{a}$. This implies that $\mathfrak{a} + KY$ is a subalgebra of \mathfrak{g} , and since \mathfrak{a} is a maximal subalgebra, $\mathfrak{a} + KY = \mathfrak{g}$, so \mathfrak{a} has codimension 1 in \mathfrak{g} , and is actually an ideal of \mathfrak{g} . Now we apply induction again to conclude that the space W of $v \in V$ with $Xv = 0$ for all $X \in \mathfrak{a}$ is non-trivial. We have proved that W is invariant under multiplication by \mathfrak{g} , so in particular $YW \subset W$. Since Y is nilpotent, we can find a non-zero $w \in W$ with $Yw = 0$, and then $(\mathfrak{a} + KY)w = 0$. \square

Theorem 2.8 (Engel). *For any Lie algebra \mathfrak{g} of nilpotent linear transformations on a vector space V , there is a basis which simultaneously upper triangularizes all elements of the algebra.*

Proof. We adapt the proof strategy in the case of a single transformation. First, we find a vector v such that $Xv = 0$ for all $X \in \mathfrak{g}$. If $W = Kv$, then

each X descends to a linear endomorphism on V/W . The image of \mathfrak{g} under this descending process satisfies the hypothesis of Engel's theorem, and therefore by induction there is a basis $v_1 + W, \dots, v_m + W$ of V/W in which all X are upper triangular. Then v, v_1, \dots, v_m is a basis for V , and

$$Xv_i \in \text{span}(v_1, \dots, v_{i-1}) + \text{span}(v)$$

and so the X are upper triangular in this new space. \square

Corollary 2.9. \mathfrak{g} is nilpotent if and only if adj_X is nilpotent for all $X \in \mathfrak{g}$.

Proof. A Lie algebra \mathfrak{g} is nilpotent if and only if there is a value n such that

$$\text{adj}_{X_1} \circ \text{adj}_{X_2} \cdots \circ \text{adj}_{X_n} = 0$$

for any $X_i \in \mathfrak{g}$. In particular, $\text{adj}_X^n = 0$ for all X . Conversely, if the adj_X are nilpotent, then Engel's theorem implies the existence of a basis of X_i such that all adj are strictly upper triangular. But if the Lie algebra is m dimensional, this implies that the nilpotency condition above holds for $n = m$. \square

Corollary 2.10. If \mathfrak{g} is a Lie algebra over an algebraically closed field, then \mathfrak{g} is nilpotent if and only if every 2 dimensional subalgebra of \mathfrak{g} is abelian.

Proof. If \mathfrak{g} is nilpotent, let $\mathfrak{h} = KX + KY$ be a 2-dimension subalgebra. Then either \mathfrak{h} is abelian, or we can choose X and Y such that $[X, Y] = X$. But then $\text{adj}_Y^n(X) = X$, so adj_Y is not nilpotent, hence \mathfrak{g} cannot be nilpotent. Thus every 2 dimensional subalgebra of \mathfrak{g} is abelian. Conversely, let \mathfrak{g} be an algebra such that every two dimensional algebra is abelian. Fix $X \in \mathfrak{g}$. Because we are working over \mathbb{C} , the adjoint adj_X has some eigenvector Y , and it suffices to show that the eigenvalue corresponding to that eigenvector is zero. But X and Y span a subalgebra of \mathfrak{g} of dimension two, and therefore $[X, Y] = 0$. \square

Engel's theorem holds for Lie algebras over arbitrary fields. The analogous theorem for solvable Lie algebras is not so easy to generalize.

Lemma 2.11. If $\mathfrak{g} \subset \mathfrak{gl}(V)$ is a solvable Lie algebra over an algebraically closed field K of characteristic zero, and $V \neq 0$, there there is $v \neq 0$ which is an eigenvector for every element of \mathfrak{g} .

Proof. We prove this theorem by induction on the dimension of \mathfrak{g} . If \mathfrak{g} has dimension 1, it suffices to show that every matrix has an eigenvector, and this follows because K is an algebraically complete field. For an arbitrary \mathfrak{g} , we note that \mathfrak{g}' is a proper subalgebra of \mathfrak{g} , hence we may enlarge it to a subspace \mathfrak{a} of \mathfrak{g} of codimension 1, choosing X such that $\mathfrak{g} = \mathfrak{a} + KX$. Since $\mathfrak{g}/\mathfrak{g}'$ is abelian, every subspace is an ideal, and hence every subspace of \mathfrak{g} containing \mathfrak{g}' is an ideal. Since every ideal of a solvable Lie algebra is solvable in of itself, we can apply induction and conclude \mathfrak{a} has an eigenvector v with corresponding weight $\lambda : \mathfrak{a} \rightarrow \mathbb{C}$. We know that V_λ is \mathfrak{g} invariant, so that there is a vector $w \in V_\lambda$ which is an eigenvector of X with eigenvalue γ . We claim that this implies that w is an eigenvector for all elements of \mathfrak{g} , because if we write $Z = \alpha X + \beta Y$, where $Y \in \mathfrak{a}$, then $Zw = (\alpha\gamma + \beta\lambda(Y))w$. \square

Lie's theorem is proved using essentially the same techniques as Engel's theorem.

Theorem 2.12. *If $\mathfrak{g} \subset \mathfrak{gl}(V)$ is a solvable finite dimensional Lie algebra over a field of characteristic zero, then there is a basis for V such that the matrix representations of \mathfrak{g} are upper triangular.*

Proof. We choose an eigenvector v for \mathfrak{g} , and then consider V/W for $W = Kv$. Every element of \mathfrak{g} descends to an operator on V/W , and by induction we can write each element as an upper triangular matrix with a certain basis $v_1 + W, \dots, v_n + W$. Passing back up, we find that each element of \mathfrak{g} is upper triangular with respect to the basis v, v_1, \dots, v_n . \square

If \mathfrak{g} is a solvable Lie algebra over an algebraically closed field of characteristic zero, which is a subset of some $\mathfrak{gl}(V)$, then Lie's theorem tells us there is a basis of V in which every element of \mathfrak{g} is upper triangular. If A and B are upper triangular matrices, then

$$(AB - BA)_{ii} = \sum_{j=1}^n A_{ij}B_{ji} - B_{ij}A_{ji} = A_{ii}B_{ii} - B_{ii}A_{ii} = 0$$

Hence the matrix representations of \mathfrak{g}' are strictly upper triangular, and therefore nilpotent! Conversely, if \mathfrak{g}' is nilpotent, then \mathfrak{g}' is solvable, and $\mathfrak{g}/\mathfrak{g}'$ is abelian, hence \mathfrak{g} is solvable. Thus for Lie algebras which are concretely represented as subalgebras of $\mathfrak{gl}(V)$, \mathfrak{g} is solvable if and only if \mathfrak{g}' is nilpotent.

Lie's theorem does not hold over all fields, unlike Engel's theorem. Over a field K of characteristic n , the matrices $X = E_{n1} + \sum_{k=1}^{n-1} E_{k(k+1)}$ and $Y = \sum_{k=1}^n (k-1)E_{kk}$ satisfy

$$\begin{aligned} [X, Y] &= \sum_{k=1}^n (k-1)[E_{n1}, E_{kk}] + \sum_{i=1}^n \sum_{j=1}^n (j-1)[E_{i(i+1)}, E_{jj}] \\ &= -(n-1)E_{n1} + \sum_{i=1}^n (i - (i-1))E_{i(i+1)} = X \end{aligned}$$

Thus X and Y span a solvable subalgebra of \mathfrak{g} , yet X and Y have no common eigenvector, because any such v with eigenvalues λ and γ must satisfy $v_1 = \lambda v_n$, and $v_{k+1} = \lambda v_k$ for $k < n$. This shows that $\lambda^n = \lambda = 1$. Looking at the definition of Y , we conclude that v is a scalar multiple of e_2 , yet $Xe_2 = e_1$. Thus the key lemma behind Lie's theorem fails here, and we find Lie's theorem does not hold either.

Because Lie's theorem is so crucial to the classical classification of Lie algebras, our techniques from now on will almost always apply only to Lie algebras over an algebraically closed field of characteristic zero.

2.2 Representation Theory

Using representation theory, we can view abstract Lie algebras as subalgebras of endomorphisms over a finite dimensional vector space. Just as groups can be represented as automorphisms over some category, reflecting their construction as the mathematical representation of symmetries, Lie algebras act on vector spaces, reflecting their construction as left-invariant vector fields on a Lie group. We define a **representation** of a Lie algebra \mathfrak{g} to be a homomorphism $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$. The homomorphism induces an **action** of \mathfrak{g} on V , defined by $Xv = \rho(X)(v)$. Conversely, if we have an action $\mathfrak{g} \times V \rightarrow V$ of \mathfrak{g} on V , satisfying

$$(\lambda X + \gamma Y)v = \lambda Xv + \gamma Yv$$

$$X(\mu v + \nu w) = \mu Xv + \nu Xw$$

$$[X, Y]v = X(Yv) - Y(Xv)$$

Then this induces a representation of \mathfrak{g} on $\mathfrak{gl}(V)$. Thus representation theory is a particular language with which to discuss module theory over

a Lie algebra. A representation is **faithful** if it is injective, and **transitive** if we can map any $v \in V$ to any $w \in V$ by some $X \in \mathfrak{g}$.

Example. We have already seen that the adjoint map $\text{adj} : \mathfrak{g} \rightarrow \text{Der}(\mathfrak{g})$ is a representation of \mathfrak{g} . It occurs very often in the theory of Lie algebras. The kernel of the homomorphism is the center $Z(\mathfrak{g})$. Since $Z(\mathfrak{g}) \subset \text{rad}(\mathfrak{g})$, this representation is faithful for the semisimple Lie algebras.

Example. If \mathfrak{g} is a subalgebra of $\mathfrak{gl}(V)$, then the inclusion of \mathfrak{g} in $\mathfrak{gl}(V)$ is a faithful representation, known as the natural representation of the subalgebra.

Example. For any Lie algebra \mathfrak{g} over a field K , we have a trivial representation of \mathfrak{g} on $\mathfrak{gl}(V)$ for any K vector space V , where we let $Xv = 0$ for all $v \in V$.

Example. Given any vector space V , the \mathbb{C} representations on V are in one-to-one correspondence with the endomorphisms $T : V \rightarrow V$. Fixing a particular T , we let \mathbb{C} act on V by defining $zv = zT(v)$.

A **submodule** of a \mathfrak{g} module V is a subspace which is closed under multiplication by elements of \mathfrak{g} . As examples, the submodules of \mathfrak{g} with the adjoint representation are exactly the ideals. Lie's theorem says that every module over a complex solvable Lie algebra has a one dimensional submodule. Given a submodule W , we can form the factor module V/W , and consider the exact sequence

$$0 \rightarrow W \rightarrow V \rightarrow V/W \rightarrow 0$$

which allows us to write V as a twisted product of W and V/W . The other isomorphism theorems hold here as well.

We wish to break modules down into their simpler representations, so that the action of \mathfrak{g} on the module becomes clear. Given two modules V and W over \mathfrak{g} , we can make the direct sum $V \oplus W$ into a module, by letting $X(v + w) = Xv + Xw$. A module V is called **irreducible**, or **simple**, if it has no submodules other than (0) and V itself. If V is not irreducible, we can find a submodule W of minimal dimension. Then V/W will have an irreducible submodule, and we may proceed inductively to write any module as the twisted product of irreducible modules. The irreducible submodules essentially form building blocks for all modules. Another reason why irreducible modules are useful is that they are precisely the modules for which the action of the Lie algebra is transitive.

Example. The only Lie algebras which are irreducible with respect to their adjoint representation are the simple Lie algebras, since the submodules of a Lie algebra correspond exactly to the ideals.

Example. The only irreducible modules over a field are one dimensional, because the field is a solvable Lie algebra. More generally, the only irreducible modules over an algebraically complete field are one-dimensional.

If we can write a module as the direct sum of other submodules, then we have said to decompose the module. A module is **indecomposable** if it cannot be broken down into a direct sum. A module is **completely reducible** if it can be written as the direct sum of irreducible submodules. This is the ideal situation to be in for understanding the structure of the action. Being completely reducible is a very strong condition. Nonetheless, so is the condition of being a semisimple Lie algebra, and we shall soon find that all modules over a complex semisimple Lie algebra are completely irreducible. For other fields, the situation is not nearly so complete.

As can be expected, a module homomorphism between two \mathfrak{g} modules is exactly one preserving the action of the algebra. The three isomorphism theorems continue to hold here, as they do for g -actions in group theory.

Example. Let V and W be complex vector spaces, and suppose we have a representation of an algebraically closed field K on V and W , corresponding to linear maps $T : V \rightarrow V$, and $S : W \rightarrow W$. Then the two modules are isomorphic if and only if there is a linear isomorphism $U : V \rightarrow W$ such that $U \circ T = S \circ U$, or $U \circ T \circ U^{-1} = S$. Thus there are bases under which S and T correspond to the same matrix. Thus the isomorphism classes of representations of K are represented by the family of Jordan normal forms.

Homomorphisms provide the best way at determining the structure of modules, and it is natural to begin looking at homomorphisms of irreducible modules. Since the image of a module homomorphism is always a submodule of that module, these homomorphisms must either be trivial, or isomorphisms.

Lemma 2.13 (Schur). If \mathfrak{g} is a Lie algebra over an algebraically closed field, and V is a finite dimensional irreducible \mathfrak{g} module, then the only module endomorphisms on V are scalar multiples of the identity.

Proof. Let $T : V \rightarrow V$ be a Lie algebra homomorphism. If $T \neq 0$, then T has an eigenvector v , satisfying $Tv = \lambda v$. Then $T - \lambda$ is a module homomorphism mapping v to zero, hence $T - \lambda = 0$, so $T = \lambda$. \square

It follows that if $X \in Z(\mathfrak{g})$, and V is an irreducible \mathfrak{g} algebra, then there exists λ such that $Xv = \lambda v$ for all $v \in V$, because X acts as a module endomorphism on V with respect to the action of \mathfrak{g} . Thus the irreducible modules for an abelian Lie algebra are all one dimensional.

2.3 Constructing Representations

The direct sum is one of a family of techniques for constructing modules out of other modules. Given a \mathfrak{g} module V , we can take the dual, and V^* has a natural \mathfrak{g} module structure, if we let $\langle Xf, v \rangle = -\langle f, Xv \rangle$, because

$$\begin{aligned}\langle [X, Y]f, v \rangle &= -\langle f, [X, Y]v \rangle = -\langle f, X(Yv) - Y(Xv) \rangle \\ &= \langle Xf, Yv \rangle - \langle Yf, Xv \rangle = \langle X(Yf) - Y(Xf), v \rangle\end{aligned}$$

This is why the negation is necessary. A module is self-dual if V is isomorphic to V^* .

Example. The adjoint representation of $\mathfrak{so}_3(\mathbf{R})$ is self dual. If we let $\rho : \mathfrak{so}_3(\mathbf{R}) \rightarrow \mathfrak{gl}_3(\mathbf{R})$ denote this representation, and if we consider a basis of $\mathfrak{so}_3(\mathbf{R})$, then

$$a = E_{12} - E_{21} \quad b = E_{13} - E_{31} \quad c = E_{23} - E_{32}$$

then $[a, b] = -c$, $[a, c] = b$, and $[b, c] = -a$. With respect to this basis, the adjoint operation has

$$\rho(a) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \quad \rho(b) = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \rho(c) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

If we consider the dual basis a^* , b^* , and c^* of $\mathfrak{so}_3(\mathbf{R})^*$, then the dual representation $\nu : \mathfrak{so}_3(\mathbf{R}) \rightarrow \mathfrak{so}_3(\mathbf{R})^*$ satisfies $\nu(X) = -\rho(X)^t$, if we identify $\mathfrak{so}_3(\mathbf{R}^*)$ with $\mathfrak{so}_3(\mathbf{R})$ by mapping a basis to its dual basis. Since all the matrix representations of a , b , and c are skew symmetric, $\nu(X) = \rho(X)$. More generally, there is a basis for a module V and a basis for \mathfrak{g} such that every element of the Lie algebra basis is skew symmetric if and only if the dual representation on V^* is isomorphic to V .

If V and W are \mathfrak{g} modules, then the space $\text{Hom}(V, W)$ of all linear maps between V and W can be given a \mathfrak{g} module structure by defining $(Xf)(v) = X(fv) - f(Xv)$, because then

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

If $Xf = 0$ for all $X \in \mathfrak{g}$, then f is a module homomorphism.

2.4 Tensors and Universal Enveloping Algebras

Given vector spaces V and W , we define the tensor product $V \otimes W$ to be equal to the free vector space on the set $V \times W$, modulo the relations

$$(\lambda v + \gamma w) \otimes (\mu u) = \lambda \mu (v \otimes u) + \gamma \mu (w \otimes u)$$

$$(\lambda v) \otimes (\gamma w + \mu u) = \lambda \gamma (v \otimes w) + \lambda \mu (v \otimes u)$$

We can iteratively form $V \otimes (W \otimes U)$ and $(V \otimes W) \otimes U$, and we find they are isomorphic as vector spaces. A nice fact about the tensor product is it effectively ‘curries’ linear maps on vector spaces. Given three vector spaces V, W, U , there is a canonical homomorphism from $\text{Hom}(V \otimes W, U)$ to $\text{Hom}(V, \text{Hom}(W, U))$. Given $f : V \otimes W \rightarrow U$, and given $v \in V$, then we obtain a homomorphism from W to U by $w \mapsto f(v, w)$. Conversely, given any homomorphism $T : V \rightarrow \text{Hom}(W, U)$, the map $f(v_i \otimes w_j) = T(v_i)(w_j)$ is a linear homomorphism from $V \otimes W$ to U . For a fixed vector space V , we let $V^{\otimes n}$ denote the n fold tensor product of V with itself. We then can define the tensor algebra over V to be $TV = \bigoplus_{k=0}^{\infty} V^{\otimes k}$. Then TV is not only a vector space, but also an algebra, if we define the tensor product $v \otimes w \in V^{\otimes n+m}$ for every $v \in V^{\otimes n}$ and $w \in V^{\otimes m}$.

The space $\text{Sym}(\mathfrak{g})$ of **symmetric tensors**, is formed from $T\mathfrak{g}$ by considering the additional relation $X \otimes Y = Y \otimes X$. It can be described as the direct sum of $\bigoplus_{k=0}^{\infty} \text{Sym}(\mathfrak{g})^k$, where each $\text{Sym}(\mathfrak{g})^k$ is formed from $\mathfrak{g}^{\otimes k}$ modulo the relation generated by $X \otimes Y = Y \otimes X$. Under this trajectory, we also introduce the set of **alternating tensors** $\Lambda(\mathfrak{g})$, which is formed from $T\mathfrak{g}$ by letting $X \otimes Y = -Y \otimes X$, and similar to the space of symmetric tensors, we consider the space of alternating n tensors $\Lambda^n(\mathfrak{g})$. The tensor product in

$\Lambda(V)$ is often changed to the alternating product $X \wedge Y$ to make a distinction between the two spaces. We have an embedding of $\text{Sym}^n(V)$ in $V^{\otimes n}$ given by

$$X_1 \dots X_n = \frac{1}{n!} \sum_{\sigma \in S_n} X_{\sigma(1)} \otimes \dots \otimes X_{\sigma(n)}$$

and an embedding of $\Lambda^n(V)$ in $V^{\otimes n}$ with

$$X_1 \wedge \dots \wedge X_n = \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma) X_{\sigma(1)} \otimes \dots \otimes X_{\sigma(n)}$$

If V is m dimensional, then $\Lambda^n(V)$ is $\binom{m}{n}$ dimensional, and $\text{Sym}^n(V)$ is $\binom{n+m-1}{m}$ dimensional. Thus we find that $\Lambda(V)$ is actually a finite dimensional vector space if V is finite dimensional, whereas TV and $\text{Sym}(V)$ are infinite dimensional. The embeddings of the symmetric and alternating tensors combine to give an isomorphism.

Theorem 2.14. *For any vector space V , $V \otimes V \cong \text{Sym}^2(V) \oplus \Lambda^2(V)$.*

Proof. The embeddings in the two dimensional case are given by

$$XY \mapsto (1/2)(X \otimes Y + Y \otimes X) \quad X \wedge Y = (1/2)(X \otimes Y - Y \otimes X)$$

and so $XY + X \wedge Y = X \otimes Y$. □

Tensors appear in almost every facet of linear algebra. Here we will apply them to Lie algebras. Given a Lie algebra \mathfrak{g} , we can form the tensor algebra $T\mathfrak{g}$. As it stands, $T\mathfrak{g}$ has no relation to the Lie bracket structure on \mathfrak{g} . But we can add this relation in by forming the **universal enveloping algebra** $\mathfrak{U}\mathfrak{g}$, which is obtained by quotienting $T\mathfrak{g}$ by the two sided ideal generated by $X \otimes Y - Y \otimes X - [X, Y]$. The embedding of \mathfrak{g} in $T\mathfrak{g}$ induces an embedding of \mathfrak{g} in $\mathfrak{U}\mathfrak{g}$, and the map is injective since the ideal generated by $X \otimes Y - Y \otimes X - [X, Y]$ contains no elements of the image of \mathfrak{g} . There is a nice property which describes the bilinear extension of \mathfrak{g} .

Theorem 2.15. *If A is an associative \mathbb{C} algebra, then any Lie algebra homomorphism $f : \mathfrak{g} \rightarrow A$ extends to a unique algebra homomorphism $f : \mathfrak{U}\mathfrak{g} \rightarrow A$.*

Proof. We can surely extend f to an algebra map from $T\mathfrak{g} \rightarrow A$ by letting $f(X \otimes Y) = f(X)f(Y)$, noting that the map is well defined, for

$$f(\lambda X + \gamma Y)f(Z) = \lambda f(X)f(Z) + \gamma f(Y)f(Z)$$

This follows if f is only a linear transformation. Now we find that

$$f(X \otimes Y - Y \otimes Z - [X, Y]) = f(X)f(Y) - f(Y)f(Z) - f([X, Y]) = 0$$

so f descends to a map on $\mathfrak{U}\mathfrak{g}$. The uniqueness follows because the map on $T\mathfrak{g}$ is uniquely defined. \square

It follows that if $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ is a representation, then ρ extends to an algebra map from $\mathfrak{U}\mathfrak{g}$ to $\mathfrak{gl}(V)$, hence modules on \mathfrak{g} are exactly the same as modules on $\mathfrak{U}\mathfrak{g}$. For convenience, we shall now denote the product operation on $\mathfrak{U}\mathfrak{g}$ as normal multiplication.

Theorem 2.16 (Poincare-Birkhoff-Witt). *Let \mathfrak{g} have an ordered basis consisting of elements $\{X_0, X_1, \dots, X_n\}$. Then $\mathfrak{U}\mathfrak{g}$ has a vector space basis consisting of $X_{i_1}^{j_1} \dots X_{i_n}^{j_n}$, for some positive indices j_i , and monotone $i_1 < \dots < i_n$.*

Another way to describe the Poincare-Birkhoff-Witt theorem is through the language of filtrations. A **filtration** on an algebra A is a monotone increasing family $\{A_i\}$ with $A_0 = (0)$, $A = \lim A_i$, and $A_i A_j \subset A_{i+j}$. Associated with A is the graded algebra $\text{gr}(A) = \bigoplus A_{i+1}/A_i$ with multiplication operation defined by $(X + A_i)(Y + A_j) = (XY + A_{i+j})$. The graded algebra is often easier to study than the original filtered algebra. Now given $\mathfrak{U}\mathfrak{g}$, we can form the filtration $\mathfrak{U}\mathfrak{g}_n = \text{span}(X_1 \dots X_m : X_i \in \mathfrak{g}, m \leq n)$. The Poincare-Birkhoff theorem can be restated as saying that the grade algebra with respect to this filtration is just the set of symmetric tensors.

Like the direct sum and dual space construction, tensor products are very useful because they enable us to construct representations from other representations. Given two \mathfrak{g} modules V and W , the tensor product $V \otimes W$ is a representation if we let $X(v \otimes w) = X(v) \otimes w + v \otimes X(w)$. The space of all tensors TV becomes an infinite dimensional representation, with each $V^{\otimes k}$ a subrepresentation. The ideal from which we form $\text{Sym}(V)$ forms a submodule of TV , and therefore $\text{Sym}(V)$ is also a representation of V . By a similar process, we find ΛV is also a representation of V .

Example. $\mathfrak{sl}_n(\mathbb{C})$ has a natural representation on $\text{Sym}^m(\mathbb{C}^n)$, which we can view as the space of homogenous polynomials of degree m in the variables e_1, \dots, e_n . The action is defined by letting

$$X(P(e_1, \dots, e_n)) = \sum X(e_i) \frac{\partial P}{\partial e_i}(e_1, \dots, e_n)$$

This is a representation, since

$$\begin{aligned}
& X(Y(P)) - Y(X(P)) \\
&= \sum X(e_j) \frac{\partial}{\partial e_j} \left(Y(e_i) \frac{\partial P}{\partial e_i} \right) - Y(e_j) \frac{\partial}{\partial e_j} \left(X(e_i) \frac{\partial P}{\partial e_i} \right) \\
&= \sum [X(e_j)Y(e_i) - Y(e_j)X(e_i)] \frac{\partial^2 P}{\partial e_j \partial e_i} + \left[X(e_j) \frac{\partial Y(e_i)}{\partial e_j} - Y(e_j) \frac{\partial X(e_i)}{\partial e_j} \right] \frac{\partial P}{\partial e_i} \\
&= \sum [X, Y](e_i) \frac{\partial P}{\partial e_i} = [X, Y](P)
\end{aligned}$$

The next section shows that these representations essentially characterize all the representations of $\mathfrak{sl}_2(\mathbb{C})$.

2.5 Representations of $\mathfrak{sl}_2(K)$

Many of the ideas which occur in the general representation theory of Lie algebras occur in the theory of representations of $\mathfrak{sl}_2(K)$. What's more, we will find that the representations of this Lie algebra control a large part of the representation theory. Recall the basis $e = E_{12}$, $f = E_{21}$, and $h = E_{11} - E_{22}$ spans the algebra. We begin by constructing a family of irreducible representations.

For each n , define V_n to be the subspace of $K[X, Y]$ generated by the homogenous polynomials with degree n

$$X^n, X^{n-1}Y, \dots, XY^{n-1}, Y^n$$

V_n is therefore a space of dimension $n + 1$. We represent $\mathfrak{sl}_2(K)$ on $\mathfrak{gl}(V_n)$ with a homomorphism $\rho : \mathfrak{sl}_2(K) \rightarrow \mathfrak{gl}(V_n)$ by defining

$$\rho(e) = X \frac{\partial}{\partial Y} \quad \rho(f) = Y \frac{\partial}{\partial X} \quad \rho(h) = X \frac{\partial}{\partial X} - Y \frac{\partial}{\partial Y}$$

These are well defined operators since they carry degree n monomials into degree n monomials. Note that now $\rho(h)(X^a Y^b) = (a - b)X^a Y^b$, so that each monomial is an eigenvector of $\rho(h)$. By construction, ρ is linear. To verify that brackets are preserved, we need only verify this for the basis.

We calculate $[e, f] = h$, and using the relations we calculated for the Weyl algebra, we find

$$\begin{aligned} [\rho(e), \rho(f)] &= \left[X \frac{\partial}{\partial Y}, Y \frac{\partial}{\partial X} \right] = X \frac{\partial}{\partial Y} Y \frac{\partial}{\partial X} - Y \frac{\partial}{\partial X} X \frac{\partial}{\partial Y} \\ &= \left(1 + Y \frac{\partial}{\partial Y} \right) \left(X \frac{\partial}{\partial X} \right) - \left(1 + X \frac{\partial}{\partial X} \right) \left(Y \frac{\partial}{\partial Y} \right) \\ &= X \frac{\partial}{\partial X} - Y \frac{\partial}{\partial Y} = \rho(h) \end{aligned}$$

Next, we calculate $[e, h] = -2e$, and find

$$\begin{aligned} [\rho(e), \rho(h)] &= \left[X \frac{\partial}{\partial Y}, X \frac{\partial}{\partial X} - Y \frac{\partial}{\partial Y} \right] \\ &= X^2 \frac{\partial}{\partial X} \frac{\partial}{\partial Y} - X \left(1 + X \frac{\partial}{\partial X} \right) \frac{\partial}{\partial Y} \\ &\quad - X \left(1 + Y \frac{\partial}{\partial Y} \right) \frac{\partial}{\partial Y} + XY \frac{\partial^2}{\partial Y^2} \\ &= -2X \frac{\partial}{\partial Y} = -2\rho(e) \end{aligned}$$

Finally, $[f, h] = 2f$, and we can reduce the calculation of the product to the calculation above by exchanging variables, so

$$\begin{aligned} [\rho(f), \rho(h)] &= \left[Y \frac{\partial}{\partial X}, X \frac{\partial}{\partial X} - Y \frac{\partial}{\partial Y} \right] = - \left[Y \frac{\partial}{\partial X}, Y \frac{\partial}{\partial Y} - X \frac{\partial}{\partial X} \right] \\ &= 2Y \frac{\partial}{\partial X} = 2\rho(f) \end{aligned}$$

In the basis $X^n, X^{n-1}Y, \dots, Y^n$, we have matrix representations

$$\rho(e) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & n \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \quad \rho(f) = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ n & 0 & \dots & 0 & 0 \\ 0 & n-1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

$$\rho(h) = \begin{pmatrix} n & 0 & \dots & 0 & 0 \\ 0 & n-2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 2-n & 0 \\ 0 & 0 & \dots & 0 & -n \end{pmatrix}$$

It is easy to see that any submodule containing a monomial contains the entire space, since we can use the representations of the basis elements to permute the monomials around, so the submodule contains all monomials, and hence the entire space. In fact, each V_n is irreducible.

Theorem 2.17. *V_n is an irreducible $\mathfrak{sl}_n(K)$ module if K is algebraically closed.*

Proof. Let W be a non-zero submodule of V_n . An eigenvector of $\rho(h)$ lies in W , because $\rho(h)$ restricts to an operation on this space. But this implies that W contains a monomial, since the eigenvalues of $\rho(h)$ are all distinct, and since $\rho(e)$ and $\rho(f)$ permute the monomials to the left and the right (disregarding scalars), we find $W = V_n$. \square

It turns out that the class of V_n represent all irreducible modules of $\mathfrak{sl}_2(K)$. The trick, given a general module V , is to look at the eigenvectors of h .

Lemma 2.18. *If $v \in V$ is an eigenvector of h with eigenvalue λ , then*

- (i) *$ev = 0$, or ev is an eigenvector of h with eigenvalue $\lambda + 2$.*
- (ii) *$fv = 0$, or fv is an eigenvector of h with eigenvalue $\lambda - 2$.*

Proof. (i) is proved by calculation,

$$h(ev) = e(\lambda v) + [h, e]v = \lambda ev + 2ev = (\lambda + 2)ev$$

and essentially the same calculation shows (ii) holds as well. \square

Lemma 2.19. *If V is a finite dimensional $\mathfrak{sl}_2(K)$ module, where K is algebraically complete, then V contains an eigenvector v for h such that $ev = 0$.*

Proof. Using the last lemma, we find that if v is an eigenvector with eigenvalue λ , then either $ev = 0$, or ev is an eigenvector with eigenvalue $\lambda + 2$, and since the eigenvalue is different, ev is independent of v . Continuing this process, we either find that $e^m v$ is an eigenvector for v with eigenvalue

$\lambda + 2m$ for all m , or there is an eigenvector w satisfying $ew = 0$. But if each $e^m v$ is an eigenvector, then they form an infinite family of independent vectors, contradicting finite dimensionality. \square

Theorem 2.20. *If V is a finite dimensional irreducible $\mathfrak{sl}_2(K)$ module over an algebraically closed field of characteristic zero, then V is isomorphic to some V_n .*

Proof. Find an eigenvector v for h such that $ev = 0$, with eigenvalue λ . Consider the sequence v, fv, f^2v, \dots . The last lemma essentially implies that $f^{m+1}v = 0$, $f^m v \neq 0$. We claim that $v, fv, \dots, f^m v$ form a basis for a submodule of V . They are certainly linearly independent, since they are eigenvectors of h with different eigenvalues. It remains to show that $e(f^k v) \in \text{span}(f^l v)$, for $l \leq k$. For $k = 0$, we know $ev = 0$. For the induction, we find

$$\begin{aligned} e(f^k v) &= (fe + h)(f^{k-1}v) \\ &= f(e(f^{k-1}v)) + (\lambda + 2(k-1))f^{k-1}v \end{aligned}$$

and by induction, $e(f^{k-1}v) \in \text{span}(f^l v)$, for $l < k$, and therefore $e(f^{k-1}v) \in \text{span}(f^l v)$, for $l \leq k$. By irreducibility, V is the span of the $f^k v$. The matrix of h with respect to the basis $f^k v$ is diagonal, with trace

$$\lambda + (\lambda - 2) + \dots + (\lambda - 2m) = (m+1)\lambda - m(m+1)$$

hence $\lambda = m$, since the image of f is in the derived subgroup, and therefore has trace zero.

We now have enough information to provide an explicit homomorphism with V_m . Note that V_m is spanned by $X^m, fX^m, \dots, f^m X^m$. If we set $\psi(f^k v) = f^k X^m$ this defines a vector space isomorphism which commutes with the action of h and f . It remains to show that ψ commutes with e , we use induction. For $k = 0$,

$$eX^m = 0 = \psi(ev)$$

Now by induction,

$$\begin{aligned}
ef^k X^m &= fef^{k-1} X^m + hf^{k-1} X^m \\
&= \psi(fef^{k-1} v) + (\lambda - 2(k-1))f^{k-1} X^m \\
&= \psi(ef^k v) - \psi(hf^{k-1} v) + (\lambda - 2(k-1))f^{k-1} X^m \\
&= \psi(ef^k v) + (\lambda - 2(k-1))(f^{k-1} X^m - \psi(f^{k-1} v)) \\
&= \psi(ef^k v)
\end{aligned}$$

and this completes the correspondence. \square

Corollary 2.21. *If V is a finite dimensional $\mathfrak{sl}_2(K)$ where K is algebraically closed, and $v \in V$ is an eigenvector of h such that $ev = 0$, then $fv = mv$ for some integer m , and the submodule of V generated by v is isomorphic to V_m .*

Proof. We have argued that $v, fv, \dots, f^m v$ span the submodule generated by v . Now we just apply irreducibility to conclude that the module generated is isomorphic to V_m . \square

Given a finite dimensional $\mathfrak{sl}_2(K)$ module V , we can find the eigenvector v for f with the highest eigenvalue, subject to the constraint that $ev = 0$. The associated eigenvalue m is known as the **highest weight**, and uniquely characterizes the module V up to isomorphism if V is irreducible.

2.6 Weyl's Theorem

Weyl discovered that the representations of complex semisimple Lie algebras take a very nice form. This will be key to the classification of semisimple Lie algebras.

Theorem 2.22 (Weyl). *If \mathfrak{g} is a semisimple Lie algebra over a field of characteristic zero, then every finite dimensional module over \mathfrak{g} is completely reducible.*

We have classified the irreducible modules of $\mathfrak{sl}_2(K)$, and since $\mathfrak{sl}_2(K)$ is semisimple, Weyl's theorem tells us that we have classified all irreducible modules, and therefore all modules over $\mathfrak{sl}_2(K)$. This makes it very easy to understand the module structure of all finite dimensional representations of $\mathfrak{sl}_2(K)$.

Example. The trivial representation of $\mathfrak{sl}_2(K)$ on K is isomorphic to V_0 .

Example. To understand the embedding of $\mathfrak{sl}_2(K)$ in $\mathfrak{gl}_2(K)$ as a representation of $\mathfrak{sl}_2(K)$, and the eigenvalues of h with respect to this representation are calculated by taking the characteristic polynomial, $(X - 1)(X + 1)$. Thus the highest weight vector of this representation has eigenvalue 1, so the natural representation of $\mathfrak{sl}_2(K)$ is isomorphic to V_1 , because it is two dimensional.

Example. With respect to the adjoint representation of $\mathfrak{sl}_2(K)$, we find that

$$[h, e] = 2e \quad [h, f] = -2f \quad [h, h] = 0$$

hence e is a highest weight vector with eigenvalue 2, and we find the adjoint representation of $\mathfrak{sl}_2(K)$ is isomorphic to V_2 , since $\mathfrak{sl}_2(K)$ has dimension 3.

Example. Consider the embedding $\rho : \mathfrak{sl}_2(K) \rightarrow \mathfrak{sl}_3(K)$, with

$$\rho \begin{pmatrix} a & b \\ c & -a \end{pmatrix} = \begin{pmatrix} a & b & 0 \\ c & -a & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Then $\mathfrak{sl}_3(K)$ can be considered a $\mathfrak{sl}_2(K)$ module, where $XY = [\rho(X), Y]$. As a matrix, $\rho(h)$ has characteristic polynomial $(X - 1)(X + 1)X$, hence as an element of the adjoint representation on $\mathfrak{sl}_3(K)$, $\rho(h)$ has an eigenvalue 0 of weight 3, eigenvalues ± 1 of weight 2, and an eigenvalues ± 2 of weight 1, so that as a representation, the module is isomorphic to $V_2 \oplus V_1 \oplus V_1 \oplus V_0$.

Chapter 3

Killing Forms and Root Systems

3.1 Cartan's Criterion

From the definition, it is very difficult to verify that an algebra is semisimple. In this chapter, we develop simple methods to decide if an algebra is semisimple, or, on the other extreme, solvable. Recall that every linear operator T on a finite dimensional algebraically closed vector space can be uniquely written as $T = T_s + T_n$, where T_s is a diagonalizable operator, T_n is nilpotent, and T_s and T_n commute. Then $T_n = 0$ precisely when T is diagonalizable, in which case we say T is **semisimple**, and $T_s = 0$ precisely when T is nilpotent. The only semisimple nilpotent operator is the zero operator. Because of our extensive use of this decomposition, and the other results we have proved, from now on we will implicitly assume that all fields in question are algebraically closed, with characteristic zero.

If \mathfrak{g} is a solvable Lie algebra, which is a subalgebra of $\mathfrak{gl}(V)$, then we know there is a basis in which all elements of \mathfrak{g} have upper triangular matrix representations, and the representations of \mathfrak{g}' are strictly upper triangular. It follows that $\text{tr}(XY) = 0$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}'$, because

$$\sum_{i,j} X_{ij} Y_{ji} = \sum_{\substack{i \leq j \\ j < i}} X_{ij} Y_{ji} = 0$$

Thus we have a necessary condition for a concrete subalgebra to be solvable. Cartan found that this is essentially a sufficient condition.

Theorem 3.1. *If \mathfrak{g} is a Lie subalgebra of $\mathfrak{gl}(V)$ over an algebraically closed field such that $\text{tr}(XY) = 0$ for $X, Y \in \mathfrak{g}$ then \mathfrak{g} is solvable.*

Proof. We prove that every linear operator in \mathfrak{g}' is nilpotent. Engel's theorem then says that \mathfrak{g}' is a nilpotent Lie algebra, and therefore \mathfrak{g} is solvable. Using the Jordan decomposition, write any $X \in \mathfrak{g}'$ as a sum $X_s + X_n$. Fix a basis where X_s is diagonal, and X_n is strictly upper triangular. Let X_s have entries $\lambda_1, \dots, \lambda_n$. It will suffice to show that $\sum |\lambda_i|^2 = 0$, so that $D = 0$, and so X is nilpotent. If we consider the diagonal matrix $\overline{X_s}$ obtained by taking the complex conjugate on the diagonal, then $\sum |\lambda_i|^2$ is just the trace of $\overline{X_s}X$. Write

$$X = \gamma_1[Y_1, Z_1] + \dots + \gamma_m[Y_m, Z_m]$$

It suffices to show that the trace of $\overline{X_s}[Y_i, Z_i]$ is zero. But since

$$\text{tr}(\overline{X_s}(Y_i Z_i - Z_i Y_i)) = \text{tr}([\overline{X_s}, Y_i]Z_i)$$

provided we can show that $[\overline{X_s}, Y_i] \in \mathfrak{g}$, we can apply our hypothesis to conclude the trace is zero. Note that the Jordan decomposition of adj_X is $\text{adj}_{X_s} + \text{adj}_{X_n}$, and therefore there is a polynomial $f \in K[X]$ with

$$f(\text{adj}_X) = \overline{\text{adj}_{X_s}} = \text{adj}_{\overline{X_s}}$$

Since adj_X maps \mathfrak{g} into itself, so too does $\text{adj}_{\overline{X_s}}$. Thus the proof is completed. \square

Since \mathfrak{g} is solvable if and only if its image in the adjoint representation is solvable, we conclude that \mathfrak{g} is solvable if and only if $\text{tr}(\text{adj}_X \circ \text{adj}_Y) = 0$ for all $X \in \mathfrak{g}$, $Y \in \mathfrak{g}'$, for this guarantees that \mathfrak{g}' is solvable, and therefore \mathfrak{g} is solvable as well. Because of this discussion, it appears that the symmetric, bilinear form on \mathfrak{g} defined by

$$\kappa(X, Y) = \text{tr}(\text{adj}_X \circ \text{adj}_Y)$$

is of interest. It is known as the **Killing form**. Another very nice property of the Killing form is that it is in fact associative,

$$\kappa([X, Y], Z) = \kappa(X, [Y, Z])$$

which follows from the trace identity. The theorem above can be stated that \mathfrak{g} is solvable if and only if $\kappa(X, Y) = 0$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}'$. This is known as Cartan's first criterion.

Example. Consider the two dimensional non-abelian Lie algebra with basis X, Y such that $[X, Y] = X$. Then the adjoints have matrix representations

$$\text{adj}_X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{adj}_Y = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

and so we see $\kappa = Y^* \otimes Y^*$. Since $\kappa(X, X) = \kappa(X, Y) = 0$, our algebra is solvable. In this case, it is easy to check that the derived subalgebra is abelian.

Example. If \mathfrak{g} is a nilpotent Lie algebra, then adj_X is nilpotent for any $X \in \mathfrak{g}$, and there is a basis for \mathfrak{g} upon which the adj_X are strictly upper triangularized, and therefore $\kappa(X, Y) = \text{tr}(\text{adj}_X \circ \text{adj}_Y) = 0$ for any $X, Y \in \mathfrak{g}$. The converse is not true, however; there are non-nilpotent Lie algebras with trivial Killing form.

Lemma 3.2. If \mathfrak{a} is an ideal of \mathfrak{g} , then the Killing form on \mathfrak{a} is the restriction of the Killing form on \mathfrak{g} .

Proof. Let κ denote the Killing form on \mathfrak{g} , and $\kappa_{\mathfrak{a}}$ the Killing form on \mathfrak{a} . If $X \in \mathfrak{a}$, then $\text{adj}_X : \mathfrak{g} \rightarrow \mathfrak{g}$ maps \mathfrak{g} into \mathfrak{a} , and therefore if we take a basis v_1, \dots, v_n for \mathfrak{a} and extend it to a basis on \mathfrak{g} , the matrix representation of adj_X will be

$$\begin{pmatrix} A_X & B_X \\ 0 & 0 \end{pmatrix}$$

If we then consider $\text{adj}_Y : \mathfrak{g} \rightarrow \mathfrak{g}$, for $Y \in \mathfrak{a}$, then the matrix representation will be

$$\begin{pmatrix} A_Y & B_Y \\ 0 & 0 \end{pmatrix}$$

and so

$$\text{adj}_X \circ \text{adj}_Y = \begin{pmatrix} A_X A_Y & A_X B_Y \\ 0 & 0 \end{pmatrix}$$

and so the trace of the restriction will agree with the trace of the adjoint on all of \mathfrak{g} . \square

The Killing form not only gives us a criterion for solvability, but also a criterion for semisimplicity. As should be expected, it is essentially the opposite of saying the Killing form vanishes. We say κ is **non-degenerate** if there is no $X \in \mathfrak{g}$ such that $\kappa(X, Y) = 0$ for all Y .

Lemma 3.3. *A Lie algebra is semisimple if and only if it has no non-zero abelian ideals.*

Proof. If a Lie algebra is semisimple, it surely contains no non-zero abelian ideals, for abelian Lie algebras are solvable. Conversely, suppose \mathfrak{g} has no non-zero abelian ideals. If \mathfrak{a} is an ideal of \mathfrak{g} , then $Z(\mathfrak{a})$ is an ideal of \mathfrak{g} , because if $[X, Z] = 0$ for all $Z \in \mathfrak{a}$, then for any $Y \in \mathfrak{g}$, $[Y, Z] \in \mathfrak{a}$, and hence

$$[[X, Y], Z] = [X, [Y, Z]] + [Y, [Z, X]] = 0 + [Y, 0] = 0$$

Thus if \mathfrak{a} is an ideal of a semisimple Lie algebra, then we conclude $Z(\mathfrak{a}) = 0$. Similarly, if \mathfrak{a} is an ideal, then \mathfrak{a}' is an ideal, because given $X \in \mathfrak{g}$, $Y, Z \in \mathfrak{a}$,

$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]] \in \mathfrak{a}' + \mathfrak{a}'$$

Now if \mathfrak{g} contains a solvable ideal \mathfrak{a} , Lie's theorem shows \mathfrak{a}' is nilpotent, and we know it is also an ideal of \mathfrak{g} . If $\mathfrak{a}'_{n+1} = [\mathfrak{a}', \mathfrak{a}_n] = 0$, and $\mathfrak{a} \neq 0$, we find $Z(\mathfrak{a}') \neq 0$, and so $Z(\mathfrak{a}')$ is an abelian ideal of \mathfrak{g} , contradicting the fact that \mathfrak{g} has no non-zero abelian ideals. \square

Theorem 3.4 (Cartan's Second Criterion). *\mathfrak{g} is semisimple if and only if the Killing form is nondegenerate.*

Proof. The set of $X \in \mathfrak{g}$ such that $\kappa(X, Y) = 0$ for all Y forms an ideal of \mathfrak{g} because the Killing form is associative. Since $\kappa_\alpha = 0$, we conclude that this ideal is solvable. Thus if \mathfrak{g} is semisimple, $\alpha = 0$, and so κ is nondegenerate. Conversely, if \mathfrak{g} has a non-trivial abelian ideal \mathfrak{a} , and if $Y \in \mathfrak{a}$ is non-zero, then $\text{adj}_Y \circ \text{adj}_X \circ \text{adj}_Y = 0$ for any $X \in \mathfrak{g}$, hence $(\text{adj}_Y \circ \text{adj}_X)^2 = 0$. Nilpotent maps have trace zero, so $\kappa(Y, X) = 0$ for all X , and so κ is degenerate. \square

Example. *Let us compute the Killing form on $\mathfrak{sl}_2(K)$. If*

$$X = \begin{pmatrix} a_0 & b_0 \\ c_0 & -a_0 \end{pmatrix} \quad Y = \begin{pmatrix} a_1 & b_1 \\ c_1 & -a_1 \end{pmatrix}$$

Then

$$[X, e] = 2a_0e - c_0h \quad [X, f] = b_0h - 2a_0f \quad [X, h] = 2c_0f - 2b_0e$$

Hence with respect to the basis $\{e, f, g\}$, we have matrix representations

$$\text{adj}_X = \begin{pmatrix} +2a_0 & 0 & -2b_0 \\ 0 & -2a_0 & +2c_0 \\ -c_0 & +b_0 & 0 \end{pmatrix} \quad \text{adj}_Y = \begin{pmatrix} +2a_1 & 0 & -2b_1 \\ 0 & -2a_1 & +2c_1 \\ -c_1 & +b_1 & 0 \end{pmatrix}$$

Hence

$$\kappa(X, Y) = \text{trace}(\text{adj}_X \circ \text{adj}_Y) = 8a_0a_1 + 4b_0c_1 + 4c_0b_1$$

Since $\mathfrak{sl}_2(K)$ is simple, we see that κ is non-degenerate.

These criterions are incredibly powerful to deriving structural results for Lie algebras. We will start by showing that every semisimple Lie algebra is the direct sum of simple Lie algebras (so the algebras really determine the name semisimple). Define the perpendicular \mathfrak{h}^\perp of a subalgebra \mathfrak{h} of a \mathfrak{g} to be the perpendicular with respect to the Killing form

$$\mathfrak{h}^\perp = \{x \in \mathfrak{g} : (\forall y \in \mathfrak{h} : \kappa(y, x) = 0)\}$$

The killing form is non-degenerate on \mathfrak{g} if $\mathfrak{g}^\perp = (0)$.

Lemma 3.5. *If \mathfrak{g} is a semisimple Lie algebra, and \mathfrak{a} is an ideal, then $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$. \mathfrak{a} is also a semisimple Lie algebra.*

Proof. Because $\kappa : \mathfrak{a} \cap \mathfrak{a}^\perp \rightarrow \mathfrak{a} \cap \mathfrak{a}^\perp$ is equal to zero, $\mathfrak{a} \cap \mathfrak{a}^\perp$ is a solvable subalgebra of a semisimple algebra, hence $\mathfrak{a} \cap \mathfrak{a}^\perp = (0)$, and by dimension counting we certainly have the vector space structure of \mathfrak{g} as a direct sum of \mathfrak{a} and \mathfrak{a}^\perp . If $X \in \mathfrak{a}$, and $Y \in \mathfrak{a}^\perp$, then $[X, Y] \in \mathfrak{a} \cap \mathfrak{a}^\perp$, because \mathfrak{a} and \mathfrak{a}^\perp are both ideals, hence $[X, Y] = 0$, and so \mathfrak{g} is the direct sum of \mathfrak{a} and \mathfrak{a}^\perp . Finally, if \mathfrak{a} was not semisimple, then $\kappa_{\mathfrak{a}}$ would be degenerate, but this would imply the degeneracy of κ . \square

We verify that \mathfrak{a}^\perp is also an ideal of \mathfrak{g} , so by recursively performing this calculation, we can break any semisimple Lie algebra down into the direct sum of simple Lie algebras. Conversely, if we consider a Lie algebra $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_m$, where each \mathfrak{g}_m is simple, and $\mathfrak{a} \subset \mathfrak{g}$ is a solvable ideal of \mathfrak{g} , then $[\mathfrak{a}, \mathfrak{g}_i] \subset \mathfrak{a} \cap \mathfrak{g}_i$ is a solvable ideal of \mathfrak{g}_i . But this implies $[\mathfrak{a}, \mathfrak{g}_i] = 0$, hence

$$[\mathfrak{a}, \mathfrak{g}] = [\mathfrak{a}, \mathfrak{g}_1 + \mathfrak{g}_2 + \cdots + \mathfrak{g}_m] = [\mathfrak{a}, \mathfrak{g}_1] + \cdots + [\mathfrak{a}, \mathfrak{g}_m] = 0$$

so $\mathfrak{a} \subset Z(\mathfrak{g})$. But $Z(\mathfrak{g}) = \bigoplus Z(\mathfrak{g}_i) = \bigoplus 0 = 0$, hence $\mathfrak{a} = 0$. Using very similar ideals, we can prove that the quotient of any semisimple Lie algebra is semisimple, because if \mathfrak{a} is any ideal of \mathfrak{g} , then $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$, and $\mathfrak{g}/\mathfrak{a}$ is isomorphic to \mathfrak{a}^\perp , which is semisimple.

Theorem 3.6. *If \mathfrak{g} is a semisimple Lie algebra, then the adjoint representation of \mathfrak{g} is an isomorphism with the space of all derivations on \mathfrak{g} .*

Proof. Denote the image of \mathfrak{g} by $\text{adj } \mathfrak{g}$. If d is any derivation, then

$$[\text{adj}_X, d](Y) = [X, dY] - d[X, Y] = -[dX, Y] = \text{adj}_{-dX}(Y)$$

Thus $\text{adj } \mathfrak{g}$ is an ideal of $\text{Der } \mathfrak{g}$. Note that this fact is true for any Lie algebra. For a semisimple Lie algebra, the adjoint representation is injective, since the center of such an algebra is trivial. The Killing form on $\text{adj } \mathfrak{g}$ is the restriction of the killing form on $\text{Der } \mathfrak{g}$. Since $\text{adj } \mathfrak{g}$ is semisimple, $(\text{adj } \mathfrak{g}) \cap (\text{adj } \mathfrak{g})^\perp = 0$. Thus $\text{Der } \mathfrak{g} = (\text{adj } \mathfrak{g}) \oplus (\text{adj } \mathfrak{g})^\perp$. This implies that if $d \in (\text{adj } \mathfrak{g})^\perp$, then $[d, \text{adj}_X] = \text{adj}_{dX} = 0$ for all X , hence $dX = 0$ for all X , because the center of \mathfrak{g} is trivial, so $d = 0$. \square

The Killing form has a certain uniqueness property which is very useful. Let $f : \mathfrak{g} \times \mathfrak{g} \rightarrow K$ be a symmetric, associative, bilinear form on \mathfrak{g} . Then f induces a map $X \mapsto X^*$, with $X \in \mathfrak{g}$ and $X^* \in \mathfrak{g}^*$, such that $X^*(Y) = f(X, Y)$. This is a vector space isomorphism, and also a \mathfrak{g} module homomorphism over the adjoint action (and the induced action over the dual space), because

$$[X, Y]^*(Z) = -f([Y, X], Z) = -f(Y, [X, Z]) = -Y^*[X, Z]$$

Thus \mathfrak{g} and \mathfrak{g}^* are isomorphic as \mathfrak{g} modules if f is non-degenerate. Now if \mathfrak{g} is simple, we can consider the Killing form κ , which is non-degenerate, and induces another dual map, which we will distinguish from the one induced by f by denoting the former by X^{*f} , and the latter by $X^{*\kappa}$. Now if $F : X \mapsto X^{*f}$ is the dual map, and $G : X \mapsto X^{*\kappa}$ is the other dual map, then $G \circ F^{-1}$ is a \mathfrak{g} module isomorphism, and \mathfrak{g} is an irreducible module since it is simple, so we conclude that there is λ such that $(G \circ F^{-1})(X) = \lambda X$ for all X , or that $G(X) = \lambda F(X)$ for all X . We conclude that $\kappa(X, Y) = \lambda f(X, Y)$ for some $\lambda \neq 0$. Thus the Killing form is essentially the only associative, symmetric, non-degenerate bilinear form on a simple Lie algebra.

We can also use the Killing form and its structure of semisimple Lie algebras to classify the representations of the algebra as operators. An important property of the semisimple Lie algebras is that the ‘Jordan decomposition’ is invariant of the representation. Note that in general, the Jordan decomposition of a representation can be fairly arbitrary. For instance, the representations of K over some vector space V are obtained by mapping 1 to an arbitrary element of V , so that the Jordan decomposition for elements of K take the form of arbitrary matrices. However, for semisimple Lie algebras, we have a nice relation.

Lemma 3.7. *If \mathfrak{g} is a complex Lie algebra, and $D \in \text{Der}(\mathfrak{g})$ has Jordan decomposition $D = D_s + D_n$, then D_s and D_n are both derivations on \mathfrak{g} .*

Proof. For each $\lambda \in \mathbb{C}$, let \mathfrak{g}_λ be the set of X such that $(D - \lambda)^m X = 0$ for some m . Then the vector space structure of \mathfrak{g} decomposes into the sum of the subspaces \mathfrak{g}_λ . We have $[\mathfrak{g}_\lambda, \mathfrak{g}_\gamma] \subset \mathfrak{g}_{\lambda+\gamma}$, because

$$(D - (\lambda + \gamma))^m (XY - YX) = \sum \binom{m}{k} [(D - \lambda)^k X, (D - \gamma)^{m-k} Y]$$

Since D_s is diagonalizable, the λ eigenspace for D_s is \mathfrak{g}_λ . If $X \in \mathfrak{g}_\lambda$ and $Y \in \mathfrak{g}_\gamma$, then $[X, Y] \in \mathfrak{g}_{\lambda+\gamma}$, hence

$$X_s[X, Y] = (\lambda + \gamma)[X, Y] = [X_s X, Y] + [X, X_s Y]$$

hence X_s is a derivation. X_n is then a derivation, because it is the difference of two derivations. \square

Theorem 3.8. *If \mathfrak{g} is semisimple, then any $X \in \mathfrak{g}$ can be written uniquely as $D + N$, where adj_D is diagonalizable, adj_N is nilpotent, and $[D, N] = 0$. If $[X, Y] = 0$, then $[D, Y] = [N, Y] = 0$.*

Proof. We apply the last theorem to \mathfrak{g} , viewed as $\text{adj } \mathfrak{g}$ by an isomorphism. Given X , we have the Jordan decomposition $\text{adj}_X = \text{adj}_D + \text{adj}_N$ for some $D, N \in \mathfrak{g}$, and hence $X = D + N$. Since adj_D and adj_N commute, $[D, N] = 0$. If Y commutes with X , then $\text{adj}_X(Y) = 0$, then because adj_D and adj_N are polynomials in adj_X , we see

$$\text{adj}_N = \sum c_k \text{adj}_X^k$$

hence $\text{adj}_N(Y) = c_0 Y$. But since adj_N is nilpotent, $c_0 = 0$. \square

The unique decomposition $X = D + N$ into semisimple and nilpotent elements is known as the **abstract Jordan decomposition** of X . We say X is semisimple if $X = D$, and X is nilpotent if $X = N$ (this new characterization of nilpotency begins with the last one, and X is semisimple if and only if adj_X is diagonalizable). In the case where our semisimple Lie algebra is a matrix Lie algebra, the abstract decomposition agrees with the concrete decomposition.

Theorem 3.9. *If $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ is a representation of a semisimple Lie algebra, and $X \in \mathfrak{g}$ has an abstract Jordan decomposition $D + N$, then $\rho(X)_s = D$ and $\rho(X)_n = N$.*

Proof. The image of any representation of a complex semisimple Lie algebra is semisimple, since if \mathfrak{h} is the kernel of ρ , then $\mathfrak{g}/\mathfrak{h}$ is isomorphic to $\rho(\mathfrak{g})$. We can therefore talk about the abstract Jordan decomposition of elements of $\rho(\mathfrak{g})$. The concrete decomposition of $\rho(X)$ agrees with the abstract decomposition, because if Y is nilpotent/diagonalizable it implies adj_Y is as well. If adj_D is diagonalizable, then it is diagonalizable on \mathfrak{h} by some basis X_1, \dots, X_n , and if we extend this basis of eigenvectors to the whole space, $X_1, \dots, X_n, Y_1, \dots, Y_m$, then, viewed as a map on the quotient $\mathfrak{g}/\mathfrak{h}$, we have $[D, Y_i + \mathfrak{h}] = \lambda_i Y_i + \mathfrak{h}$. If adj_N is nilpotent, then adj_N is nilpotent on the quotient as well. This verifies that $\rho(X)$ has abstract Jordan decomposition $\rho(D) + \rho(N)$, because the decomposition is unique. \square

Because of this theorem, it is fair to denote both the abstract and concrete Jordan decomposition of an element X_s and X_n of a Lie algebra by the same notation. They agree where they are both defined.

3.2 Root Systems

Recall the proof that $\mathfrak{sl}_2(K)$ is the only simple 3 dimensional Lie algebra. Given a particular simple Lie algebra \mathfrak{g} :

1. We found $X \in \mathfrak{g}$ such that adj_X is diagonalizable.
2. We took a basis of \mathfrak{g} consisting of eigenvectors for adj_X , and showed that the structural constants can be chosen to coincide with $\mathfrak{sl}_2(K)$.

To classify the semisimple Lie algebras, we will utilize a very important generalization of this technique. In order to apply this technique to higher dimensional Lie algebras, we will need to generalize the first step to finding a sufficiently large simultaneously diagonalizable subalgebra of the Lie algebra. This subalgebra is known as the Cartan subalgebra.

Let us try and generalize the technique to $\mathfrak{sl}_n(K)$. Consider the subalgebra \mathfrak{h} of $\mathfrak{sl}_n(K)$ consisting of the diagonal matrices. We find that for any diagonal matrix $X = (\sum a_i E_{ii})$, if $m \neq n$, then

$$[X, E_{mn}] = \sum a_i [E_{ii}, E_{mn}] = (a_m - a_n) E_{mn}$$

Since the diagonal matrices form an abelian subalgebra of $\mathfrak{sl}_n(K)$, we can extend the E_{ij} to a full basis diagonalizing all diagonal matrices. In particular, we see the weights for this action are precisely the maps $\varepsilon_{ij}(M) = M_{ii} - M_{jj}$. The weight space for ε_{ij} for $i \neq j$ corresponds precisely to the span of e_{ij} , so the space is 1 dimensional, and for $i = j$, the weight space is precisely the space of diagonal matrices, which is $n - 1$ dimensional. Since the weight spaces form \mathfrak{h} -invariant subspaces, we have a module decomposition

$$\mathfrak{sl}_n(K) = \mathfrak{h} \oplus \bigoplus_{i \neq j} K e_{ij}$$

which effectively characterizes the action of the adjoint maps on \mathfrak{h} .

It seems that the adjoint representation abelian subalgebras are very useful for decomposing Lie algebras. Indeed, let \mathfrak{h} be an arbitrary abelian subalgebra of \mathfrak{g} consisting of semisimple elements. If $X, Y \in \mathfrak{h}$, then adj_X and adj_Y are commutative, because

$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]] = [Y, [X, Z]]$$

hence $\text{adj}_X \circ \text{adj}_Y = \text{adj}_Y \circ \text{adj}_X$. Since each adj_X is diagonalizable, the elements of \mathfrak{h} are *simultaneously diagonalizable* – there is a basis of \mathfrak{g} consisting of eigenvectors. We thus have a decomposition of \mathfrak{g} as

$$\mathfrak{g} = \bigoplus_{\lambda} \mathfrak{g}_{\lambda}$$

where λ ranges over the set of weights for \mathfrak{h} , and the \mathfrak{g}_{λ} are \mathfrak{h} invariant under the Lie bracket. The elements of \mathfrak{g}_0 are exactly the elements of the centralizer $C(\mathfrak{h})$.

Theorem 3.10. *If $X \in \mathfrak{g}_{\lambda}$ and $Y \in \mathfrak{g}_{\gamma}$, then $[X, Y] \in \mathfrak{g}_{\lambda+\gamma}$.*

Proof. Using the Jacobi identity, we find that for any $Z \in \mathfrak{h}$,

$$\begin{aligned} [Z, [X, Y]] &= [[Z, X], Y] + [X, [Z, Y]] \\ &= \lambda(Z)[X, Y] - \gamma(Z)[Y, X] \\ &= (\lambda(Z) + \gamma(Z))[X, Y] \end{aligned}$$

In particular, if $\lambda + \gamma$ is not a root for \mathfrak{h} , then X and Y commute. \square

Since there are only finitely many weights, we conclude that if $X \in \mathfrak{g}_\lambda$, with $\lambda \neq 0$, then adj_X is nilpotent. Given any element $Y \in \mathfrak{g}$, write $Y = \sum Y_\gamma$, where $Y_\gamma \in \mathfrak{g}_\gamma$. Then $[X, Y] = \sum [X, Y_\gamma] \in \sum \mathfrak{g}_{\lambda+\gamma}$, and more generally, $\text{adj}_X^n(Y) \in \sum \mathfrak{g}_{\gamma+n\lambda}$. If we choose n such that $\mathfrak{g}_{\gamma+n\lambda} = 0$ for all γ with $\mathfrak{g}_\gamma \neq 0$, then adj_X is nilpotent.

Theorem 3.11. *If $\lambda + \gamma \neq 0$, and $X \in \mathfrak{g}_\lambda$, $Y \in \mathfrak{g}_\gamma$, then $\kappa(X, Y) = 0$.*

Proof. If $Z \in \mathfrak{h}$, then using the Jacobi identity, we find using linearity and commutativity of κ that

$$\lambda(Z)\kappa(X, Y) = \kappa([Z, X], Y) = -\kappa(X, [Z, Y]) = -\gamma(Z)\kappa(X, Y)$$

subtracting one side of this equation from the other, we find that $(\lambda + \gamma)(Z)\kappa(X, Y) = 0$. If $(\lambda + \gamma)(Z) \neq 0$, then $\kappa(X, Y) = 0$, and we can always choose Z such that $(\lambda + \gamma)(Z) \neq 0$ if $\lambda + \gamma \neq 0$. \square

This shows that κ is non-degenerate on \mathfrak{h} , because if there was $X \in \mathfrak{g}_0$ such that $\kappa(X, Y) = 0$ for all $Y \in \mathfrak{g}_0$, then for any $Z \in \mathfrak{g}$, we may write $Z = \sum Z_\lambda$, and then $\kappa(X, Z) = \kappa(X, Z_0) = 0$, so that X annihilates all elements of \mathfrak{g} , and thus $X = 0$. It follows from Cartan's criterion that \mathfrak{h} is a semisimple Lie algebra.

In general, our aim will be to identify an abelian subalgebra of semisimple elements of every semisimple Lie algebra, in which case we can find a structural decomposition of the algebra. If the abelian subalgebra is too small, the weight decomposition is likely to be too coarse, and the results about orthogonality on κ and how the bracket operates on weights not powerful enough to classify the algebra. It turns out that every semisimple Lie algebra has a maximal abelian subalgebra containing semisimple elements, called a **Cartan subalgebra**. This algebra has the property that the Lie algebra elements of weight zero correspond precisely to elements of the Cartan subalgebra, so that the weight decomposition is very tight.

Example. *If \mathfrak{h} is the Lie subalgebra of $\mathfrak{sl}_n(K)$ spanned by h , then \mathfrak{g}_0 consists precisely of the matrices X with $X_{1n} = X_{n1} = 0$ for $n \neq 1$, and $X_{2n} = X_{n2} = 0$ for $n \neq 2$, so $\mathfrak{g}_0 = \mathfrak{h}$ precisely when $n = 2$. The \mathfrak{g}_0 has dimension $n^2 - 1 - 2(n - 1) - 2(n - 2) = n^2 - 4n + 5$, so the decomposition is very coarse for $n > 2$.*

Theorem 3.12. *Cartan subalgebras exist on semisimple Lie algebras.*

Proof. First, we note that semisimple elements must exist on a semisimple Lie algebra \mathfrak{g} . We can write any $X \in \mathfrak{g}$ uniquely as the sum of a semisimple element and a nilpotent element, so if a semisimple Lie algebra contains no semisimple elements, every element of the algebra is nilpotent, and therefore \mathfrak{g} is nilpotent. It follows that \mathfrak{g} is solvable, yet no semisimple Lie algebras are solvable. If X is a semisimple element, then the span of X is abelian and semisimple, and by finite dimensionality we may enlarge this algebra to be maximum abelian and semisimple, in which case the algebra obtained is a Cartan subalgebra. \square

In the case of the classical semisimple Lie algebras, there are canonical choices of a Cartan subalgebra, which will correspond to the diagonal matrices of the algebra.

Example. The canonical Cartan subalgebra of $\mathfrak{sl}_n(\mathbb{C})$ consists of the diagonal matrices, and is spanned by the matrices $H_k = E_{kk} - E_{(k+1)(k+1)}$, for $k < n$. We then find that for $i \neq j$,

$$[H_k, E_{ij}] = (\delta_k^i - \delta_{k+1}^i + \delta_{k+1}^j - \delta_k^j)E_{ij}$$

Hence the weights on $\mathfrak{sl}_n(\mathbb{C})$ are $\alpha_{ij}(H_k) = \delta_k^i - \delta_{k+1}^i + \delta_{k+1}^j - \delta_k^j = (\varepsilon_i - \varepsilon_j)(H_k)$.

Example. For $\mathfrak{sp}_n(\mathbb{C})$, with $n = 2m$, the Cartan subalgebra is the space of diagonal matrices, which has basis $H_k = E_{kk} - E_{(k+m)(k+m)}$. The corresponding eigenvectors are $E_{ij} - E_{(j+m)(i+m)}$, with eigenvalues s . TODO: BLAH BLAH BLAH FINISH LATER.

If $S \subset \mathfrak{g}$, then we define the **centralizer** of S with respect to \mathfrak{g} , denoted $C_{\mathfrak{g}}(S)$ or just $C(S)$, to be the set of all $X \in \mathfrak{g}$ such that $[X, Y] = 0$ for all $Y \in S$. The centralizer of Cartan subalgebra combined with the maximality of the Cartan algebra will turn out to be very useful. The centralizer is always a subalgebra of \mathfrak{g} , because if $X, Y \in C(S)$, and $Z \in S$, then $[[X, Y], Z] = [X, [Y, Z]] + [[X, Z], Y] = 0 + 0 = 0$. Using our previous notation, we have $Z(\mathfrak{g}) = C(\mathfrak{g})$, and for any subalgebra of \mathfrak{h} of \mathfrak{g} consisting of semisimple elements, $C(\mathfrak{h}) = \mathfrak{g}_0$.

Lemma 3.13. If \mathfrak{h} is a Cartan subalgebra of \mathfrak{g} . If $X \in \mathfrak{h}$ is chosen such that the dimension of $C(X)$ is minimized, then $\mathfrak{h} \subset Z(C(X))$, so $C(X) = C(\mathfrak{h})$.

Proof. We claim that if $Y \in C(X) \cap \mathfrak{h}$ is not in $Z(C(X))$, then there is some element of \mathfrak{h} of the form $aX + bY$ whose centralizer has smaller dimension than X . First, consider some basis $\{Z_1, \dots, Z_k\}$ on $C(X) \cap C(Y)$. Extend this basis with vectors $\{Z'_1, \dots, Z'_n\}$ on $C(X)$ to diagonalize adj_Y , and consider an alternate extension of $\{Z''_1, \dots, Z''_m\}$ which diagonalize adj_X on $C(Y)$. Then the triple of basis are linearly independent, for if

$$\sum a_i Z_i + \sum b_i Z'_i + \sum c_i Z''_i = 0$$

Then for some λ_i

$$\left[X, \sum a_i Z_i + \sum b_i Z'_i + \sum c_i Z''_i \right] = \sum c_i \lambda_i Z''_i$$

Hence $c_i \lambda_i = 0$, and by assumption $Z''_i \notin C(X)$, hence $\lambda_i \neq 0$, hence $c_i = 0$ for all i . Similarly, considering the bracket with Y , we find $b_i = 0$, and then $a_i = 0$ as well. Thus the triple is a basis for $C(X) + C(Y)$. Finally, since X and Y commute, we can simultaneously diagonalize X and Y , so we consider a final set of independent elements $\{Z'''_1, \dots, Z'''_l\}$ such that the quadruple forms a basis for \mathfrak{h} , and simultaneously diagonalizes X and Y .

If $[Y, Z'_i] = \lambda_i$, $[X, Z''_j] = \gamma_j$, $[X, Z'''_k] = \sigma_k$, and $[Y, Z'''_k] = \nu_k$, then all λ_i , γ_j , σ_k and ν_k are non-zero, and if we choose a non-zero μ such that $\mu \neq -\nu_i/\sigma_i$ for any i , we find

$$\begin{aligned} & \left[Y + \mu X, \sum a_i Z_i + \sum b_i Z'_i + \sum c_i Z''_i + \sum d_i Z'''_i \right] \\ &= \sum b_i \lambda_i Z'_i + \mu \sum c_i \gamma_i Z'''_i + \sum d_i (\nu_i + \mu \sigma_i) Z'''_i \end{aligned}$$

Hence $C(Y + \mu X) = C(X) \cap C(Y)$, a strictly smaller set than $C(X)$ because $Y \in C(Y)$, $Y \notin C(X)$. \square

Theorem 3.14. *If \mathfrak{h} is a Cartan subalgebra, and $X \in \mathfrak{h}$ satisfies $C(X) = C(\mathfrak{h})$, then $C(X) = \mathfrak{h}$. Thus \mathfrak{h} is self-centralizing.*

Proof. Certainly $\mathfrak{h} \subset C(X)$. If $Y \in C(X)$, consider the Jordan decomposition $Y = S + N$. Since X commutes with Y , S and N both commute with X . Thus it suffices to show that $N = 0$, in which case Y is semisimple, and therefore lies in \mathfrak{h} . We already know $S \in \mathfrak{h}$ by maximality.

We claim the only nilpotent element in $C(\mathfrak{h})$ is zero. First, we claim that $C(X)$ is nilpotent. If we take any $Y \in C(X)$, and write $Y = S + N$, then

adj_Y is equal to adj_N on $C(X)$, because $S \in \mathfrak{h} = C(X)$. This implies adj_Y is nilpotent, hence Y is nilpotent, and since Y was arbitrary we use Engel's theorem to conclude that $C(\mathfrak{h})$ is nilpotent.

Next, we claim that every element of $C(\mathfrak{h})$ is semisimple. If $Y \in C(\mathfrak{h})$, write $Y = S + N$. $C(\mathfrak{h})$ is nilpotent, hence solvable, so by Lie's theorem there is a basis for \mathfrak{g} in which case adj_Y is represented by an upper triangular matrix. Since adj_N is nilpotent, the matrix is strictly upper triangular, hence if $Z \in C(X)$,

$$\kappa(N, Z) = \text{tr}(\text{adj}_N \circ \text{adj}_Z) = 0$$

But we know that κ is non-degenerate on $C(\mathfrak{h})$, because it contains \mathfrak{h} , so $N = 0$. \square

Thus a general semisimple Lie algebra \mathfrak{g} has a Cartan subalgebra \mathfrak{h} , and since the eigenvector of \mathfrak{h} of weight zero are precisely $C_{\mathfrak{g}}(\mathfrak{h})$, and we have shown $\mathfrak{h} = C_{\mathfrak{g}}(\mathfrak{h})$, we may write \mathfrak{g} via the **root space decomposition**

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

where Φ is the set of roots of \mathfrak{h} , the set of non-zero weights. The roots and weights depend on the Cartan subalgebra \mathfrak{h} we choose, but for the classical algebras the Cartan subalgebra is canonical – it is simply the diagonal matrices in the algebra. We shall assume that a particular Cartan subalgebra \mathfrak{h} has been fixed over the space, so we can discuss the roots of \mathfrak{g} without ambiguity.

3.3 Subalgebras isomorphic to $\mathfrak{sl}_2(K)$

It turns out that semisimple Lie algebras contain an abundance of subalgebras isomorphic to $\mathfrak{sl}_2(K)$, which enables us to bring the representation theory we have developed for $\mathfrak{sl}_2(K)$ to bear, thereby obtaining structural results for all semisimple Lie algebras.

Lemma 3.15. *If α is a non-zero root, then $-\alpha$ is a non-zero root, and for every non-zero $X \in \mathfrak{g}_{\alpha}$ there is $Y \in \mathfrak{g}_{-\alpha}$ such that $\text{span}(X, Y, [X, Y])$ is a Lie subalgebra of \mathfrak{g} isomorphic to $\mathfrak{sl}_2(K)$.*

Proof. Fixing some $X \in \mathfrak{g}_\alpha$, the non-degeneracy of κ implies that there is some $Z \in \mathfrak{g}$ with $\kappa(X, Z) \neq 0$. If we write the weight decomposition of Z as $\sum Z_\alpha$, then this implies that $Z_{-\alpha} \neq 0$, hence $-\alpha$ is a root of \mathfrak{h} . Let $Y = Z_{-\alpha}$. Then $[X, Y]$ is a weight zero element, hence $[X, Y] \in \mathfrak{h}$. Since $\alpha \neq 0$, find $Z \in \mathfrak{h}$ with $\alpha(Z) \neq 0$, so that

$$\kappa(Z, [X, Y]) = \kappa([Z, Y], X) = \alpha(Z)\kappa(Y, X) \neq 0$$

hence $[X, Y] \neq 0$. The span of X, Y , and $[X, Y]$ is therefore a 3-dimensional subalgebra \mathfrak{k} of \mathfrak{g} . We claim that $\alpha[X, Y] \neq 0$, which would imply $\mathfrak{k}' = \mathfrak{k}$, and therefore that \mathfrak{k} is isomorphic to $\mathfrak{sl}_2(K)$. If $\alpha[X, Y] = 0$, then $[X, [X, Y]] = [Y, [X, Y]] = 0$. Then X and Y commute with $[X, Y]$, so $[X, Y]$ is nilpotent on \mathfrak{k} . But then $[X, Y]$ is both nilpotent and semisimple, which can only occur if $[X, Y] = 0$, which we know not to be the case. \square

We let $\mathfrak{sl}(\alpha)$ denote the subalgebra obtained in the lemma. As we now know, $\mathfrak{sl}(\alpha)$ might not be unique, but we will soon show it is, because each weight space is one dimensional, so the choice of X and Y are effectively uniquely determined. We let $e_\alpha \in \mathfrak{g}_\alpha$, $f_\alpha \in \mathfrak{g}_{-\alpha}$, and $h_\alpha \in \mathfrak{h}$ be elements of $\mathfrak{sl}(\alpha)$ which map to e, f , and h in the isomorphism with $\mathfrak{sl}_2(K)$. We can let $e_\alpha = X$ in the last theorem, and f_α an appropriate scalar multiple of Y . We know

$$[e_\alpha, f_\alpha] = h_\alpha \quad [e_\alpha, h_\alpha] = -2e_\alpha \quad [f_\alpha, h_\alpha] = 2f_\alpha$$

The choices aren't unique, but we will find they are unique enough, and fixed over each $\mathfrak{sl}(\alpha)$.

Theorem 3.16. *The choices of e_α , f_α , and h_α are unique up inversely scaling e_α and f_α as λe_α , $\lambda^{-1} f_\alpha$, h_α for some non-zero $\lambda \in K$. To obtain a basis for $\mathfrak{sl}(-\alpha)$ from $\mathfrak{sl}(\alpha)$, swap e_α with f_α , and replacing h_α with $-h_\alpha$.*

Proof. Consider some other $e'_\alpha \in \mathfrak{g}_\alpha$, $f'_\alpha \in \mathfrak{g}_{-\alpha}$, and $h'_\alpha \in \mathfrak{h}$. We may assume, by inverse scaling (since \mathfrak{g}_α is one dimensional), that $e'_\alpha = e_\alpha$. Let $f'_\alpha = \lambda f_\alpha$. It then follows that h'_α is a scalar multiple of h_α , $h'_\alpha = \gamma h_\alpha$. Then

$$[e_\alpha, f'_\alpha] = \lambda h_\alpha \quad [e_\alpha, f'_\alpha] = [e'_\alpha, f'_\alpha] = h'_\alpha = \gamma h_\alpha$$

Hence $\lambda = \gamma$. Also $[e_\alpha, h'_\alpha] = -2e_\alpha$, but also $[e_\alpha, h'_\alpha] = -2\gamma e_\alpha$, so $\lambda = \gamma = 1$. \square

The Killing form is non-degenerate, and it therefore gives a canonical isomorphism of \mathfrak{h} with \mathfrak{h}^* , mapping X to $X^* \in \mathfrak{h}^*$ defined by $X^*(Y) = k(X, Y)$. In particular, for each α there is $t_\alpha \in \mathfrak{h}$ with $\kappa(t_\alpha, X) = \alpha(X)$ for all $X \in \mathfrak{h}$.

Lemma 3.17. *If α is a root, let $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$. Then $[X, Y] = \kappa(X, Y)t_\alpha$. In particular, $h_\alpha = [e_\alpha, f_\alpha] \in \text{span}(t_\alpha)$.*

Proof. If $Z \in \mathfrak{h}$, then

$$\kappa(Z, [X, Y]) = \kappa([Z, X], Y) = \alpha(Z)\kappa(X, Y) = \kappa(Z, \kappa(X, Y)t_\alpha)$$

Since κ is non-degenerate on \mathfrak{h} , $[X, Y] = \kappa(X, Y)t_\alpha$. \square

Here is where all our work comes to fruition. Since any semisimple Lie algebra \mathfrak{g} contains some $\mathfrak{sl}(\alpha)$, we can view \mathfrak{g} as a $\mathfrak{sl}_2(K)$ module by first considering the isomorphism of $\mathfrak{sl}_2(K)$ with $\mathfrak{sl}(\alpha)$, and then taking the adjoint representation. The submodules of \mathfrak{g} with respect to this representation are precisely the vector subspaces V with $[X, Y] \in V$ for all $X \in \mathfrak{sl}_2(K)$ and $Y \in V$.

Lemma 3.18. *If V is a $\mathfrak{sl}(\alpha)$ submodule of \mathfrak{g} , then the eigenvalues of h_α acting on V are integers.*

Proof. This follows from the classification of modules over $\mathfrak{sl}_2(K)$. \square

Example. The Cartan subalgebra of $\mathfrak{sl}_3(K)$ consists of all diagonal matrices with trace zero. The weights are $\varepsilon_{ij}(X) = X_{ii} - X_{jj}$, for $i \neq j$, and the elements of the weight space for ε_{ij} are the scalar multiples of E_{ij} . Let $\alpha = \varepsilon_{12}$. Then $\mathfrak{sl}(\alpha)$ is generated by E_{12} , E_{21} , and $E_{11} - E_{22}$. We can let $E_{12} = e_\alpha$, $E_{21} = f_\alpha$, and $E_{11} - E_{22} = h_\alpha$. This is exactly the embedding we considered when we were considering the classification of finite dimensional representations of $\mathfrak{sl}_2(K)$, so that we know $\mathfrak{sl}_3(K)$ is isomorphic to $V_2 \oplus V_1 \oplus V_1 \oplus V_0$ as an $\mathfrak{sl}_2(K)$ module.

Example. Let $\mathfrak{u} = \mathfrak{h} + \mathfrak{sl}(\alpha)$. Let $\mathfrak{k} = \ker \alpha \subset \mathfrak{h}$. Then \mathfrak{k} has codimension 1 in \mathfrak{h} . As \mathfrak{h} is abelian, $[h_\alpha, X] = 0$ for all $X \in \mathfrak{k}$, and

$$[e_\alpha, X] = -[X, e_\alpha] = -\alpha(X)e_\alpha = 0$$

and similarly $[f_\alpha, X] = 0$. Thus every element of $\mathfrak{sl}(\alpha)$ acts trivially on \mathfrak{k} , and we can decompose \mathfrak{u} into the direct sum of \mathfrak{k} and $\mathfrak{sl}(\alpha)$. $\mathfrak{sl}(\alpha)$ is isomorphic as an $\mathfrak{sl}(\alpha)$ module to V_2 , so \mathfrak{u} is isomorphic to the direct sum of $\dim \mathfrak{h} - 1$ copies of the trivial representation V_0 , and one copy of V_2 .

Example. If $\beta \in \Phi$, or $\beta = 0$, let

$$V = \bigoplus_{z \in \mathbb{C}} \mathfrak{g}_{z\alpha + \beta}$$

Then V is a $\mathfrak{sl}(\alpha)$ submodule of \mathfrak{g} , known as the α -root string through β , because if $X \in \mathfrak{g}_{z\alpha + \beta}$, then $[e_\alpha, X] \in \mathfrak{g}_{(z+1)\alpha + \beta}$, $[f_\alpha, X] \in \mathfrak{g}_{(z-1)\alpha + \beta}$, and $[h_\alpha, X] = (z\alpha + \beta)(h_\alpha)X \in \mathfrak{g}_{z\alpha + \beta}$.

Theorem 3.19. The root space of any non-zero α on a semisimple Lie algebra is one-dimensional, and the only multiples of α which are roots are $\pm\alpha$.

Proof. If $z\alpha$ is a root, then h_α has $z\alpha(h_\alpha) = 2z$ as an eigenvalue. As the eigenvalues of h_α are integral, $2z \in \mathbb{Z}$. Consider the root string module

$$V = \mathfrak{k} \oplus \bigoplus_{n \neq 0} \mathfrak{g}_{(n/2)\alpha}$$

Let $\mathfrak{k} = \ker \alpha \subset \mathfrak{h}$. Then $\mathfrak{k} \oplus \mathfrak{sl}(\alpha)$ is a $\mathfrak{sl}(\alpha)$ submodule of V containing \mathfrak{h} . Since modules over $\mathfrak{sl}(\alpha)$ are completely reducible, we can find a complementary submodule W such that $V = \mathfrak{k} \oplus \mathfrak{sl}(\alpha) \oplus W$. If the conclusion of this theorem was false, then there would be situations where we would find $W \neq 0$, hence it suffices to prove that $W = 0$. Note that the eigenvectors of h_α on V of eigenvalue zero are exactly the elements of \mathfrak{k} , so W cannot contain any eigenvectors of eigenvalue zero, thus W cannot contain any irreducible submodule isomorphic to V_n , where n is even. This already gives us an interesting consequence. If $2\alpha \in \Phi$, then h_α has $(2\alpha)(h_\alpha) = 4$, hence V has an eigenvector of eigenvalue 4. But every element of $\mathfrak{k} \oplus \mathfrak{sl}(\alpha)$ has eigenvalue 0 and ± 2 with respect to h_α , which would imply W contains a submodule isomorphic to some V_n with n even. Thus if α is a root, 2α is never a root. Finally, assume W has a submodule isomorphic to V_n with n odd. Then V contains a h_α eigenvector X with eigenvalue 1, which must be contained in W . Thus there is a root β with $\beta(h_\alpha) = 1$, but then $2\beta = \alpha$, since 2β and α agree at h_α , and we have just verified this we cannot obtain a root by doubling another root. Thus W has no irreducible submodules, and as such $W = 0$. \square

Theorem 3.20. Let $\alpha, \beta \in \Phi$, with $\beta \neq \pm\alpha$.

(i) $\beta(h_\alpha) \in \mathbb{Z}$.

- (ii) There are positive integers $a, b \geq 0$ such that $\beta + k\alpha \in \Phi$ if and only if $k \in \mathbb{Z}$ and $-a \leq k \leq b$, and $a - b = \beta(h_\alpha)$.
- (iii) If $\alpha + \beta \in \Phi$, then $[e_\alpha, e_\beta]$ is a non-zero scalar multiple of $e_{\alpha+\beta}$.
- (iv) $\beta - \beta(h_\alpha)\alpha \in \Phi$.

Proof. Note that $\beta(h_\alpha)$ is the eigenvalue of h_α on some element of \mathfrak{g}_β , hence it must be an integer by the classification of $\mathfrak{sl}_2(K)$ representations. If we consider the root string V , which is the direct sum of $\mathfrak{g}_{\beta+k\alpha}$ where k is chosen such that $\beta + k\alpha$ is a root, then we have $(\beta + k\alpha)(h_\alpha) = \beta(h_\alpha) + 2k$, thus the eigenvalues of h_α on the root string are either all even or all odd, and we therefore find that the root string is an irreducible $\mathfrak{sl}(\alpha)$ module, hence V is isomorphic to some V_n . On V_n , h_α acts diagonally with eigenvalues $n, n-2, \dots, -(n-2), -n$, and this must be paired up with $\beta(h_\alpha) + 2k$. Thus we may let $-n = \beta(h_\alpha) - 2a$, $n = \beta(h_\alpha) + 2b$, in which case $a - b = \beta(h_\alpha)$. This proves (ii). (iv) essentially follows from (ii) in the same manner. If $X \in \mathfrak{g}_\beta$, then X belongs to the h_α eigenspace of eigenvalue $\beta(h_\alpha)$, and if $[e_\alpha, e_\beta] = 0$, then e_β is the highest weight vector in the irreducible representation V isomorphic to V_n . If $\alpha + \beta$ is a root, then h_α has eigenvalue $(\alpha + \beta)(h_\alpha) = 2 + \beta(h_\alpha)$, hence e_β is not the highest weight vector, hence $[e_\alpha, e_\beta] \neq 0$, and this proves (iii). \square

Thus we find that the roots of a Lie algebra essentially give us all the needed structural constants between brackets of the weight space decomposition. It determines the brackets $[e_\alpha, e_\beta]$ for up to a scalar constant, and tells us that $[e_\alpha, e_{-\alpha}]$ is in the span of h_α . We are well on our way to classifying the semisimple Lie algebras!

3.4 Cartan Subalgebras as Inner Product Spaces

Lemma 3.21. *If $X \in \mathfrak{h}$ is non-zero, then there is a root α with $\alpha(X) \neq 0$, and therefore the roots span \mathfrak{h}^* .*

Proof. If $\alpha(X) = 0$ for all roots α , then for any $Y \in \mathfrak{g}$, we can write $Y = Y_0 + \sum_{\alpha \in \Phi} Y_\alpha$, then $[X, Y] = \sum \alpha(X) Y_\alpha = 0$, hence $X \in Z(\mathfrak{g}) = 0$, hence $X = 0$ by semisimplicity. If V is the span of the roots on \mathfrak{h}^* , and if $V \neq \mathfrak{h}^*$, then the annihilator $W^\circ = \{X \in \mathfrak{h} : (\forall \lambda \in V : \lambda(X) = 0)\}$ has non-zero dimension, which we have proved is impossible. \square

Lemma 3.22. For any $\alpha \in \Phi$,

$$t_\alpha = \frac{h_\alpha}{\kappa(e_\alpha, f_\alpha)} \quad h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)} \quad \kappa(t_\alpha, t_\alpha)\kappa(h_\alpha, h_\alpha) = 4$$

Proof. We obtain the formula for t_α by multiplying by $\kappa(e_\alpha, f_\alpha)$ on both sides of the equation

$$h_\alpha = [e_\alpha, f_\alpha] = \kappa(e_\alpha, f_\alpha)t_\alpha$$

which we have already proved. Now $\alpha(h_\alpha) = 2$, hence

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

The second formula then follows by substitution. Using this formula, we find

$$\kappa(h_\alpha, h_\alpha) = \kappa\left(\frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}, \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}\right) = \frac{4}{\kappa(t_\alpha, t_\alpha)}$$

□

As a corollary, we find that $\kappa(h_\alpha, h_\beta) \in \mathbf{Z}$ and $\kappa(t_\alpha, t_\beta) \in \mathbf{Q}$ for all α, β , because by the root space decomposition

$$\kappa(h_\alpha, h_\beta) = \text{tr}(\text{adj}_{h_\alpha} \circ \text{adj}_{h_\beta}) = \sum_{\gamma \in \Phi} \gamma(h_\alpha)\gamma(h_\beta)$$

and the previous lemmas imply

$$\kappa(t_\alpha, t_\beta) = \frac{\kappa(t_\alpha, t_\alpha)\kappa(t_\beta, t_\beta)}{4} \kappa(h_\alpha, h_\beta) \in \mathbf{Q}$$

Hence the Killing form doesn't seem to take many complex values on the roots. This is important, for we shall find that the roots form an interesting real subspace on the roots.

The Killing form on \mathfrak{h} translates to a non-degenerate symmetric bilinear form on \mathfrak{h}^* , denoted (\cdot, \cdot) . We can define the form as

$$(\lambda, \gamma) = \kappa(t_\lambda, t_\gamma)$$

We have verified that $(\alpha, \beta) \in \mathbf{Q}$ if α, β are roots. Since the roots of \mathfrak{h}^* span \mathfrak{h} , \mathfrak{h}^* has a basis of roots $\{\alpha_1, \dots, \alpha_n\}$. Something stronger can be said on this front.

Lemma 3.23. *If β is a root, then β is a linear combination of the α_i with rational coefficients.*

Proof. Certainly we may write $\beta = \sum \lambda_i \alpha_i$, with $\lambda_i \in K$. We have

$$(\beta, \alpha_j) = \sum \lambda_i (\alpha_i, \alpha_j)$$

This is a system of linear equations with rational coefficients in the values λ_i , and since each (α_i, α_j) is rational, and (β, α_j) is rational, we conclude the λ_i are rational. \square

Thus the *rational* subspace generated by the α_i contains all the roots of Φ , and doesn't depend on the particular choice of basis roots α_i . Let $E_{\mathbf{Q}}$ denote the rational subspace generated by the roots, and let $E = \mathbf{R} \otimes_{\mathbf{Q}} E_{\mathbf{Q}}$ denote the vector space obtained by extending the basefield of $E_{\mathbf{Q}}$ to the real numbers. Then (\cdot, \cdot) extends to a bilinear form on E , in the obvious way, by defining $(x \otimes \lambda, y \otimes \gamma) = xy(\lambda, \gamma)$.

Theorem 3.24. *(\cdot, \cdot) is a real inner product on E .*

Proof. Since (\cdot, \cdot) takes the value of rational numbers on the roots, we know that the bilinear map is rational-valued on $E_{\mathbf{Q}}$, and if $\lambda \in E_{\mathbf{Q}}$ is a given element, then for any $X \in \mathfrak{g}_{\alpha}$, $\text{adj}_{t_{\lambda}}(X) = \alpha(t_{\lambda})X$, and if $X \in \mathfrak{h}$, $\text{adj}_{t_{\lambda}}(X) = 0$, hence

$$(\lambda, \lambda) = \kappa(t_{\lambda}, t_{\lambda}) = \text{tr}(\text{adj}_{t_{\lambda}}^2) = \sum_{\alpha} [\alpha(t_{\lambda})]^2$$

hence if $(\lambda, \lambda) = 0$, $\alpha(t_{\lambda}) = 0$ for all α , hence $t_{\lambda} = 0$, so $\lambda = 0$. Now the theorem follows for all elements of $E_{\mathbf{Q}}$, because if $(x \otimes \lambda, x \otimes \lambda) = 0$, then either $(\lambda, \lambda) = 0$, or $x^2 = 0$, in which case either $x = 0$ or λ , and hence $x \otimes \lambda = 0$. \square

3.5 Root systems

Let E be the real vector space spanned by the roots Φ of some semisimple Lie algebra. Let us summarize what we know about the roots Φ in E .

- $0 \notin \Phi$.

- The only scalar multiples of $\alpha \in \Phi$ which are in Φ are $\pm\alpha$.
- If $\alpha, \beta \in \Phi$, if we define $\langle \alpha, \beta \rangle = 2(\alpha, \beta)/(\beta, \beta)$, then

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

So that if θ is the angle between α and β , then $2\|\alpha\|\cos(\theta)\beta$ (twice the projection of α onto β) is an integer multiple of β .

- If s_α is the reflection about the perpendicular to α , which can be defined as $s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha$, then s_α permutes Φ .

It turns out that finite subsets of real inner product spaces with the properties above are very useful throughout mathematics. They are known as **root systems**, and we will attempt to classify them. We will find that root systems give another way to represent semisimple Lie algebras, so that by classifying the root systems, we classify the semisimple Lie algebras.

Example. In \mathbf{R}^n with the standard inner product, for $n \geq 2$, let R be the set of $e_i - e_j$, for $i \neq j$. Then R is a root system over the span of the $e_i - e_j$, which is the set of vectors x such that $\sum x_i = 0$. The first and second properties of root systems are trivial, and

$$\langle e_i - e_j, e_k - e_l \rangle = \frac{2(e_i - e_j, e_k - e_l)}{\|e_k - e_l\|^2} = (e_i - e_j, e_k - e_l) = \delta_i^k + \delta_j^l - \delta_i^l - \delta_j^k$$

and fixing $\alpha = e_k - e_l$

$$s_\alpha(e_i - e_j) = (e_i - e_j) - (\delta_i^k + \delta_j^l - \delta_i^l - \delta_j^k)(e_k - e_l)$$

Then

- If $(\delta_i^k + \delta_j^l - \delta_i^l - \delta_j^k) = 2$, then $i = k, j = l$, and $s_\alpha(e_i - e_j) = e_j - e_i$.
- If $(\delta_i^k + \delta_j^l - \delta_i^l - \delta_j^k) = 1$, then either $i = k$ and $j \neq l$, in which case $s_\alpha(e_i - e_j) = (e_i - e_j) - (e_i - e_l) = e_l - e_j$, or $j = l$ and $i \neq k$, in which case $s_\alpha(e_i - e_j) = (e_i - e_j) - (e_k - e_j) = e_i - e_k$.
- If $(\delta_i^k + \delta_j^l - \delta_i^l - \delta_j^k) = 0$, then $s_\alpha(\beta) = \beta$.

- If $(\delta_i^k + \delta_j^l - \delta_i^l - \delta_j^k) = -1$, then either $i = l$ and $j \neq k$, in which case $s_\alpha(e_i - e_j) = (e_i - e_j) + (e_k - e_i) = e_k - e_j$, or $j = k$ and $i \neq l$, in which case $s_\alpha(e_i - e_j) = (e_i - e_j) + (e_j - e_l) = e_i - e_l$.
- If $(\delta_i^k + \delta_j^l - \delta_i^l - \delta_j^k) = -2$, then $i = l$ and $j = k$, and $s_\alpha(e_i - e_j) = e_j - e_i$.

Thus R is a root system.

Example. The only root systems on \mathbf{R} , with the standard inner product consist of pairs $\{\pm x\}$, for any non-zero $x \in \mathbf{R}$.

Lemma 3.25. If R is a root system, for two roots α and β , with $\beta \neq \pm\alpha$,

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \in \{0, 1, 2, 3\}$$

Proof. A relationship of the angle between two vectors $x, y \in E$ is

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = \frac{4(\alpha, \beta)^2}{(\alpha, \alpha)(\beta, \beta)} = 4\cos^2 \theta \leq 4$$

Thus $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle$ is an integer between 0 and 4, and if $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4$, then $\cos^2 \theta = 1$, hence α and β are linearly dependent, which we know is impossible. \square

There is therefore only particularly many cases for $\langle \alpha, \beta \rangle$, since each of these values must also be integral. They are given in the chart below.

$\langle \alpha, \beta \rangle$	$\langle \beta, \alpha \rangle$	θ	$\frac{(\beta, \beta)}{(\alpha, \alpha)}$
0	0	$\pi/2$	underdetermined
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
1	2	$\pi/4$	2
-1	-2	$3\pi/4$	2
1	3	$\pi/6$	3
-1	-3	$5\pi/6$	3

Note that the angle between two vectors determines the ratio of length between the two vectors.

Theorem 3.26. If α and β are roots, and the angle between α and β is strictly obtuse, then $\alpha + \beta$ is a root. If the angle is strictly acute, and $(\beta, \beta) \geq (\alpha, \alpha)$, then $\alpha - \beta$ is a root.

Proof. In either case, we may assume $(\beta, \beta) \geq (\alpha, \alpha)$. It therefore follows that $s_\beta(\alpha) = \alpha - \langle \alpha, \beta \rangle \beta$ is an element of the root system. If the angle is strictly obtuse, then $\langle \alpha, \beta \rangle = -1$, and if the angle is strictly acute, then $\langle \alpha, \beta \rangle = 1$. \square

Example. We already have enough information to classify all root systems in \mathbf{R}^2 with respect to the standard inner product. Let R be a root system, and let α have minimum length. Consider some other root β linearly independent from α (β must exist, since the root system spans \mathbf{R}^2). We may assume that β makes an obtuse angle with α (otherwise, consider $-\beta$), and that this angle is as large as possible.

- If β lies at an angle $\pi/2$ for α , then $\|\beta\| = \|\alpha\|$, and $R = \{\alpha, \beta, -\alpha, -\beta\}$, because it is a root system, with

$$s_\alpha(\beta) = -\beta \quad s_\beta(\alpha) = -\alpha$$

and given any vector $x \in R$, either $x = \pm\alpha$, or x lies at an obtuse angle from α , in which case $x = \pm\beta$. This is the root system $A_1 \times A_1$.

- If β lies at an angle $2\pi/3$ from α , then $\|\alpha\| = \|\beta\|$, and we find $\alpha + \beta$ is also a root, as is $-(\alpha + \beta)$. What's more, the set of roots

$$\{\alpha, -\alpha, \beta, -\beta, \alpha + \beta, -(\alpha + \beta)\}$$

form a root system. It is impossible to add any other root vectors to the system without introducing a root which lies at a larger obtuse angle than β , hence the roots are exactly these vectors. We call this system A_2 .

- If β lies at an angle $3\pi/4$ from α , then $\alpha + \beta$ and $-(\alpha + \beta)$ are roots. $2\alpha + \beta$ is then obtained from reflecting β across the line $\alpha + \beta$, so the root system contains

$$\{\pm\alpha, \pm\beta, \pm(\alpha + \beta), \pm(2\alpha + \beta), -(2\alpha + \beta)\}$$

and any other root would make an angle smaller than $\pi/4$, contradicting that β is the largest angle. We can't add any other roots, and this classifies the root system, which we denote by B_2 .

- If β lies at an angle $5\pi/6$ from α , then by similar reasoning we find the root system consists of 12 vectors,

$$\{\pm\alpha, \pm\beta, \pm(\alpha + \beta), \pm(2\alpha + \beta), \pm(3\alpha + \beta), \pm(3\alpha + 2\beta)\}$$

and this system is denoted G_2 .

In the case $A_1 \times A_1$, the perpendicular vectors don't act on one another. We say a root system R is reducible if it can be written as the disjoint union of two sets R_0 and R_1 , where vectors of R_1 . If this is impossible, we call the root system irreducible. Note that if this is possible, then both R_0 and R_1 are already root systems over the vector spaces they span.

Lemma 3.27. *Every root system R on a vector space E can be decomposed as the disjoint union of irreducible root systems $R_1 \cup \dots \cup R_n$, where each R_i is an irreducible root system over some subspace E_i of E , and E decomposes as the direct sum of E_i .*

Proof. Consider the transitive, symmetric, reflexive closure of the relation $(\lambda, \gamma) \neq 0$. Let R_i be an equivalence class of this relation. It is clear that the span of the R_i form subspaces E_i which form E . Clearly each R_i contains the negation of each root, since $(\lambda, -\lambda) = -(\lambda, \lambda) \neq 0$ for all λ . If $\alpha, \beta \in R_i$, then $s_\alpha(\beta) \in R_j$ for some j , because

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

□

3.6 Root Spaces for the Classical Lie Algebras

It will be important for us to consider the Cartan subalgebras and root spaces of the classical Lie algebras, that occur as spaces of matrices.

We have already seen the Cartan subalgebra \mathfrak{h} of $\mathfrak{sl}_n(\mathbb{C})$, which consists of the diagonal matrices. Indeed, if $XY = YX$ for all diagonal matrices $Y \in \mathfrak{sl}_n(\mathbb{C})$, then in particular

$$0 = [X, E_{ii} - E_{(i+1)(i+1)}] = \sum_j X_{ji} E_{ji} - X_{ij} E_{ij} - X_{j(i+1)} E_{j(i+1)} + X_{(i+1)j} E_{(i+1)j}$$

which implies $X_{ij} = 0$ for $i \neq j$, hence \mathfrak{h} is maximal. For $i \neq j$,

$$\left[\sum_k a_k E_{kk}, E_{ij} \right] = (a_i - a_j) E_{ij}$$

so the weights consist of the maps $\varepsilon_{ij}(X) = X_{ii} - X_{jj}$, for $i \neq j$.