

# The Representation Theory of Lie Algebras

Jacob Denson

March 29, 2017

# Table Of Contents

<b>1</b>	<b>Basic Definitions</b>	<b>2</b>
1.1	Factorization . . . . .	11
1.2	Solvable Lie Algebras . . . . .	14
1.3	Low Dimensional Classifications . . . . .	16
<b>2</b>	<b>Matrix Lie Algebras</b>	<b>24</b>
2.1	Engel and Lie's Theorem . . . . .	28
2.2	Basic Representation Theory . . . . .	33
2.3	Representations of $\mathfrak{sl}_2(K)$ . . . . .	38
<b>3</b>	<b>Killing Forms</b>	<b>43</b>
3.1	Cartan's Criterion . . . . .	43
3.2	The Casimir Element and Weyl's Theorem . . . . .	51
3.3	Cartan Subalgebras . . . . .	55
3.4	Subalgebras isomorphic to $\mathfrak{sl}_2(K)$ . . . . .	62
3.5	Cartan Subalgebras as Inner Product Spaces . . . . .	68
<b>4</b>	<b>Root systems</b>	<b>70</b>
4.1	The Weyl Group . . . . .	77
4.2	Classification of Root Systems . . . . .	80
<b>5</b>	<b>The Classification of Semisimple Lie Algebras</b>	<b>82</b>

# 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 in a more simple form by considering the linearized approximation of the operation on infinitesimals, and this is how Lie algebras enter the picture.

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. Given a certain Lie group  $G$ , almost every structure related to the group operations on  $G$  can be uniquely characterized on a neighbourhood of the identity. Assuming  $G$  is connected, we find that

- Every neighbourhood of the identity generates the group.
- Every differentiable homomorphism  $\phi : G \rightarrow H$  is uniquely specified on a neighbourhood of the identity. In fact, a homomorphism  $\phi : G \rightarrow H$  is uniquely specified by its differential  $\phi_* : G_e \rightarrow H_e$ .

This tells us that we can likely maintain most of the global information about the Lie group structure by approximating the global information as data on the tangent space  $\mathfrak{g} = G_e$  at the origin. The most important data is the linearization of the group action on the space, and from this operation we obtain the Lie bracket. On the class of *real* Lie groups, we obtain all possible Lie algebras over the real numbers, and over the class of *complex* analytic Lie groups, we obtain all possible Lie algebras over the complex numbers. The theory of analytic groups generalizes the theory of Lie groups to arbitrary complete fields of characteristic zero, and from here we can obtain all Lie algebras over abstract fields of characteristic zero. We shall find that we will soon focus on Lie algebras over the complex numbers. We shall find that since the complex numbers are algebraically closed, complex finite dimensional Lie algebras are particularly simple, and this is the class we shall focus on. The classical explorations of complex Lie algebras are easily generalizable to all Lie algebras over algebraically closed fields of characteristic zero, and this is the scenario we will focus in on.

The first order operations of multiplication and inversion seem to lead to no interesting structure on the tangent space. If  $G$  is a Lie group, the multiplication map  $m : G \times G \rightarrow G$  can be differentiated to obtain an operation  $m_* : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ . Similarly, the inversion map  $i : G \rightarrow G$  can be differentiated to obtain an operation  $i_* : \mathfrak{g} \rightarrow \mathfrak{g}$ . Unfortunately these operations give little information about the Lie group to the first order. Infinitesimal multiplication turns in addition, and infinitesimal inversion turns into negation,

$$m_*(X_e, Y_e) = X_e + Y_e \quad i_*(X_e) = -X_e$$

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. Consider the commutator map  $c(g, h) = ghg^{-1}h^{-1}$ . The first order approximation of this map is trivial around the origin, which is essentially the reason why linearizing Lie groups is difficult, because all Lie groups are commutative up to first order. However, because  $c_*(X_e, Y_e) = 0$  for infinitesimals  $X_e, Y_e$  around the origin, we can consider the second derivative  $c_{**} : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  where if  $\lambda, \gamma \in G_e$  are two curves through the origin, then

$$c_{**}(\lambda, \gamma)(f) = [f \circ c \circ (\lambda, \gamma)]''(0)$$

That  $c_{**}(X, Y)$  is a derivation at the identity, which follows because the first derivative is trivial, so  $(f \circ c \circ (\lambda, \gamma))'(0) = 0$  for all  $X, Y$ , hence by the Leibnitz rule,

$$\begin{aligned} c_{**}(\lambda, \gamma)(fg) &= ((fg) \circ c \circ (\lambda, \gamma))''(0) \\ &= c_{**}(\lambda, \gamma)(f)g(e) + f(e)c_{**}(\lambda, \gamma)(g) \end{aligned}$$

Unlike  $c_*$ ,  $c_{**}$  is non-trivial, and expresses the operation of the commutator up to second order terms. This operation is the Lie bracket operation on the tangent space at the origin, giving us the second order operation of the commutator. If we denote  $c_{**}(\lambda, \gamma)$  by  $[\lambda, \gamma]$ , then we find

$$f(\lambda(t)\gamma(t)\lambda^{-1}(t)\gamma^{-1}(t)) = f(e) + (t^2/2)[\lambda, \gamma](f) + o(t^3)$$

Note that if we instead define the Lie algebra of a group  $G$  to be the space of left-invariant vector fields on the group, then this definition is exactly the Lie bracket operation on vector fields. In Lie algebra theory, we find that the second order terms of the group operation suffice to obtain much of the information about the original Lie group, hence the Lie bracket is the key to connecting the abstract study of Lie algebras and the analytical study of Lie groups.

There is another way to construct the canonical Lie bracket on the tangent space of any Lie group, which is more natural from the study of representation theory. Given any Lie group  $G$ , there is a natural representation  $\text{Adj} : G \rightarrow GL(\mathfrak{g})$  of  $G$  on the vector space formed from tangent vectors at the identity on  $G$ , known as the **adjoint representation** of  $G$ , defined by letting  $\text{Adj}(g)(X) = (c_g)_*(X)$ , where  $c_g : G \rightarrow G$  is the operation  $c_g(h) = ghg^{-1}$ . Since  $\text{Adj}$  is continuous, it is also smooth, hence we may differentiate  $\text{Adj}$  to obtain a map  $\text{adj} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$ , and by calculation we can calculate that  $\text{adj}(X)(Y) = [X, Y]$ . The calculations of the second order terms of the commutator are hidden by taking the tangent bundle of matrices at the tangent bundle of  $G$  at the identity. This makes it more clear why studying the representation theory of Lie algebras naturally gives relations between the representations of Lie groups. Indeed, if  $\rho : G \rightarrow GL(V)$  is any representation, then  $\rho_* \circ \text{Adj}(g) = \text{Adj}(\rho(g)) \circ \rho_*$ , and by differentiating this relation, we find that for any  $X \in \mathfrak{g}$ ,  $\text{adj}(\rho_* X) \circ \rho_* = \rho_* \circ \text{adj}(X)$ , hence  $\rho_* : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$  is a linear map preserving the Lie bracket on  $\mathfrak{g}$ , because the equation says that for any  $Y \in \mathfrak{g}$ ,  $[\rho_* X, \rho_* Y] = \rho_*[X, Y]$ .

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. In many cases, we can recover a global concept about the Lie group from a property of its infinitesimal Lie algebra (in some cases we have to assume the Lie group is simply connected, because the fundamental group cannot be encoded locally on the tangent space. This just means that we have to replace an arbitrary Lie group  $G$  with its covering space  $\tilde{G}$ , which has a canonical Lie group structure). For instance,

- If  $G$  and  $H$  are Lie groups, then every Lie group homomorphism  $f : G \rightarrow H$  gives rise to a Lie algebra homomorphism  $f : \mathfrak{g} \rightarrow \mathfrak{h}$ . Conversely, if  $G$  is a simply connected Lie group, then every Lie algebra homomorphism  $f : \mathfrak{g} \rightarrow \mathfrak{h}$  corresponds to a unique homomorphism  $\phi : G \rightarrow H$  with  $\phi_*(X_e) = f(X_e)$ . In particular, the differentiable representations  $\phi : G \rightarrow GL(V)$  are in one to one correspondence with the linear maps  $f : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$  preserving the Lie bracket, which we call representations of the Lie algebra.
- Every Lie subgroup corresponds to a unique Lie subalgebra, and conversely, in a connected Lie group, a Lie subalgebra gives rise to a unique (immersed) Lie subgroup, so the two classes of objects are in one to one correspondence.

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. Properties of  $\mathfrak{g}$  thereby rise up to properties of  $G$ .

**Example.** The Lie group  $GL_n(K)$  of invertible matrices with coefficients in  $K$  is the most natural non commutative Lie group. 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 Neumann expansion

$$(I + M)^{-1} = \sum_{k=0}^{\infty} (-1)^k M^k$$

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

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

hence the Lie bracket takes the form  $[M, N] = MN - NM$ . We call this Lie algebra the general linear Lie algebra, and denote it by  $\mathfrak{gl}_n(K)$ . In general, the set of invertible endomorphisms  $GL(V)$  on a finite dimensional vector space  $V$  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 a canonical differentiable manifold structure since it is a finite dimensional 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  $[X, Y] = XY - YX$  (the calculation is the same as for the general linear group). We will therefore denote the Lie algebra structure on  $A$  by  $\mathfrak{u}(A)$ .

**Example.** If a Lie group  $G$  is abelian, then  $c(g, h) = e$  for all  $g, h$ , so the corresponding Lie bracket is trivial:  $c_{**}(X_e, Y_e) = 0$ . Given any Lie algebra  $\mathfrak{g}$ , we call the algebra **abelian** if  $[X, Y] = 0$  for all  $X, Y \in \mathfrak{g}$ . If a connected Lie group  $G$  induces an abelian Lie algebra  $\mathfrak{g}$ , then this implies  $G$  is also abelian.

**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  $\partial_X$ . 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  $\partial_X^m(X^m) = m! = 0$  for  $m \geq p$ . But in this scenario we actually find that  $\partial_X^p = 0$ , 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 algebras are very useful in 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 finds 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

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

Thus the set of all 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  $\text{adj}_X$  is actually a derivation on  $\mathfrak{g}$ . The map  $X \mapsto \text{adj}_X$  is actually a homomorphism from  $\mathfrak{g}$  to  $\text{Der}(\mathfrak{g})$  (a homomorphism of Lie algebras is a linear map preserving the Lie bracket), which we have seen before as the differential of the adjoint representation of groups, and

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

The kernel of the adjoint representation is the center of  $\mathfrak{g}$ ,

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

(which, if  $G$  is a connected Lie group, is an ideal of  $\mathfrak{g}$  corresponds to the normal subgroup  $Z(G)$  of  $G$ ). Elements of  $\text{Der}(\mathfrak{g})$  of the form  $\text{adj}_X$  are called **inner derivations**, and other elements are called **outer derivations**. 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  $C^\infty$  real-valued functions on the Lie group. Indeed, if we view an element of the tangent space at the origin as a left-invariant vector field, then these vector field operate as derivations on the space of  $C^\infty$  functions, and we could have defined the Lie bracket operation as  $[X, Y](f) = X(Yf) - Y(Xf)$ .

**Example.** The formula

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

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  $\text{tr}(XY) = \text{tr}(YX)$ , we find

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

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  $\text{char } K \neq 2$ , 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  $\omega$  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  $SP_n(K)$  consist of the 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  $M^t J M = J$ . If we consider a first order approximation at the origin, we find

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

hence the tangent vectors of  $SP_n(K)$  at the origin consist of matrices  $M$  such that  $M^t J + J M = 0$ . Together, these matrices form 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

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

The diagonal coefficients vanish, since they are constant over the entire group, and the other coefficients are allowed to vary arbitrarily, because they were allowed to vary arbitrarily in the original group.

## 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 main strategy in this characterization, like in most algebraic categories, is to consider subalgebras and quotients, which break apart the group into more understandable chunks. A homomorphism 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 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 Lie algebras which are orthogonal to each other with respect to the Lie bracket.

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 by considering the equation  $[g, h] \in H$  infinitesimally, we conclude 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  $G$ . The standard way to form a short exact sequence for a Lie algebra  $\mathfrak{g}$  is to find an ideal  $\mathfrak{a}$ . 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  $\text{tr} : \mathfrak{gl}_n(K) \rightarrow K$  is actually a Lie algebra homomorphism, because  $\text{tr}[x, y] = \text{tr}(xy - yx) = 0 = [\text{tr}(x), \text{tr}(y)]$ . It is the differential of the determinant map  $\det : GL_n(K) \rightarrow K$ , as we saw when determining the elements of the special linear Lie algebra. 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.

**Example.** The adjoint map  $\text{adj} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$  is a homomorphism with kernel  $Z(\mathfrak{g})$ . The image of  $\mathfrak{g}$  is the set of **inner derivations** on  $\mathfrak{g}$ , denoted  $\text{adj } \mathfrak{g}$ . Thus we have an exact sequence

$$0 \rightarrow Z(\mathfrak{g}) \rightarrow \mathfrak{g} \rightarrow \text{adj } \mathfrak{g}$$

And so  $\mathfrak{g}$  is obtained from a twisted product of  $\text{adj } \mathfrak{g}$  and  $Z(\mathfrak{g})$ . For Lie algebras with trivial center, we find that  $\mathfrak{g}$  is isomorphic to  $\text{adj } \mathfrak{g}$ . This is a weak form of

*Ado's theorem, a deep result which shows every finite dimensional Lie algebra over a field of characteristic zero is isomorphic to a subalgebra of  $\mathfrak{gl}(V)$  for some finite dimensional vector space  $V$ .*

**Example.** If  $d : \mathfrak{g} \rightarrow \mathfrak{g}$  is a derivation, and  $X \in \mathfrak{g}$ , then

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

Hence  $[d, \text{adj}_X] = \text{adj}_{dX}$ , and so the subalgebra of inner derivations in  $\text{Der } \mathfrak{g}$  forms an ideal. We can thereby understand the structure of the outer derivations modulo the inner derivations, which can help us understand their structure if we forget about inner derivations.

As in most of algebra, the process of constructing ideals and constructing homomorphisms are dual to one another, 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

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 is 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  $\mathfrak{g}$

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

The subalgebra of  $\mathfrak{g}$  corresponding to the commutator subgroup 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  $\mathfrak{g}_{\text{ab}}$ . 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  $\mathfrak{g}_{\text{ab}}$ , existing somewhat ‘beneath the surface’ of the calculations. If we consider  $D^2\mathfrak{g} = \mathfrak{g}''$ , 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,

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

where  $D^n\mathfrak{g}$  is the  $n$ 'th element of the **derived series**. For this to work, we require that  $D^{n+1}\mathfrak{g} = 0$  for some  $n$ , in which case we say  $\mathfrak{g}$  is a **solvable** Lie algebra. A connected Lie group is solvable if and only if its corresponding Lie algebra is solvable.

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,  $D^n \mathfrak{g} \subset \mathfrak{g}_n$ , so that the series doesn't penetrate as 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

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

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 Lie algebras with the most non-commutative substructure are the **semi-simple** ones, non-zero algebras with  $\text{rad}(\mathfrak{g}) = (0)$ . 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. Levi discovered 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 ‘coordinatization process’ 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  $f_i : \mathfrak{g}^2 \rightarrow K$  is a bilinear form, and so we may find constants  $a_{ij}^k \in K$  such that

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

which is the same as saying  $[E_i, E_j] = \sum a_k^{ij} E_k$ . 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  $a_k^{ij} = -a_k^{ji}$ . The Jacobi identity takes the form

$$\sum_l a_l^{jk} a_m^{il} + a_l^{ki} a_m^{jl} + a_l^{ij} a_m^{kl} = 0$$

for every fixed  $i, j, k, m$ . Conversely, given a set of constants  $a_k^{ij}$  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 very useful to memorize 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 calculated 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.

**Example.** *Because every one-dimensional Lie algebra is abelian, all one dimensional real and complex Lie groups are abelian. Over the real numbers, the only connected one dimensional Lie groups are the real numbers  $\mathbf{R}$  under addition, or the torus  $\mathbf{T}$ . Over the complex numbers, the only connected one dimensional Lie groups are the complex numbers  $\mathbf{C}$  under addition, or a quotient of  $\mathbf{C}$  by a discrete subgroup. If the discrete subgroup is one-dimensional, then we obtain a Lie group isomorphic to the group  $\mathbf{C}^\times$  of complex numbers under multiplication. If the discrete subgroup is two-dimensional, we obtain a Lie group isomorphic to the complex torus  $\mathbf{C}/\mathbf{Z}^2 \cong \mathbf{T}^2$ .*

Before we carry out the calculations of Lie algebras of dimension two and three, 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  $\omega$  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 identity 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 center 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 center.

**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  $(\lambda, \gamma, \eta)$ , 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 center of the algebra.  $\square$

To obtain the rest of the three dimensional classification, we need to work over an algebraically complete field. We shall find this is integral to most of the classification theory of Lie algebras. Note that if  $K$  is any field, then it has some algebraic closure  $E$ , and if  $\mathfrak{g}$  is a Lie algebra over  $K$ , we can consider a corresponding algebra  $E \otimes_K \mathfrak{g}$  as a field over  $E$ , and since the bracket is a bilinear map on  $\mathfrak{g}$

There is a large family of non-isomorphic Lie algebras whose derived algebra is two-dimensional. We will focus on the case where  $K$  is an algebraically closed field

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  $\text{adj}_Y$ , 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  $\alpha$  and  $\beta$ . We claim that the trace of any adjoint operator from a derived element is zero, so that  $\beta = 0$ . By consider the adjoint map  $\text{adj}_Z$ , 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,  $\text{adj}_X$  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.

- 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  $\gamma$ , this does indeed define a Lie algebra structure on  $K^3$ , denoted  $\mathfrak{g}_\gamma$ , 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  $\mathfrak{g}_\lambda$  is isomorphic to  $\mathfrak{g}_\gamma$  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  $\mathfrak{g}_\lambda$  and  $\mathfrak{g}_\gamma$  is  $\lambda = 1/\gamma$ . 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  $\mathfrak{g}_\gamma$ ,

$$a_{22}Y + a_{32}Z = [a_{11}X + a_{12}Y + a_{13}Z, a_{22}Y + a_{32}Z] = a_{11}a_{22}Y + a_{11}a_{32}\gamma Z$$

$$\lambda(a_{23}Y + a_{33}Z) = [a_{11}X + a_{12}Y + a_{31}Z, a_{23}Y + a_{33}Z] = a_{11}a_{23}Y + a_{11}a_{33}\gamma Z$$

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  $\lambda = \gamma$ .

- 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  $\lambda$ , 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  $X/\lambda$ , and  $Z$  with  $Z/\lambda$ , 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  $\text{adj}_X$  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  $\text{adj}_X$ . This also implies that  $[X, M] = 0$ , then  $M$  is in the span of  $X$ . We claim that we can choose  $X$  such that  $\text{adj}_X$  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  $\text{adj}_{Y_0}$  has a non-zero eigenvalue.

If  $\text{adj}_X$  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  $\text{adj}_X$  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  $\lambda$  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 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. Given an arbitrary Lie algebra, the adjoint representation embeds  $\mathfrak{g}$  in  $\text{Der}(\mathfrak{g})$ , which is a subalgebra of  $\mathfrak{gl}(V)$ . We may then understand the adjoint action on element of  $\mathfrak{g}$  via the study of matrix Lie algebras. Indeed, the adjoint of a Lie algebra element often gives lots of information about the original element. In this chapter, we will assume all vector spaces are finite dimensional, though we will find that certain infinite dimensional vector spaces will prove important in the forthcoming theory of representations.

**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  $\text{diag}(\lambda_1, \dots, \lambda_n)$ , and if  $E_{ij} : V \rightarrow V$  is the linear operator which maps all  $v_i$  to zero, except mapping  $v_j$  to  $v_i$ , then  $T = \sum \lambda_i E_{ii}$ , and so

$$[\text{diag}(\lambda_1, \dots, \lambda_n), E_{ij}] = \sum_{k=1}^n \lambda_k [E_{kk}, E_{ij}] = (\lambda_i - \lambda_j) E_{ij}$$

And so the  $E_{ij}$  form a basis for  $\mathfrak{gl}(V)$  diagonalizing  $\text{adj}_T$ . □

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

*Proof.* We have an expansion

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

If  $T^m = 0$ , then  $\text{adj}_T^{2m}(X) = 0$  for all  $X$ , because either  $T^k = 0$ , which is implied if  $k \geq m$ , or  $T^{2m-k} = 0$ , because if  $k < m$  then  $2m - k > m$ , so each term  $T^k X T^{2m-k}$  vanishes, and this means that  $\text{adj}_T^{2m}(X)$  vanishes as well.  $\square$

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. We say an element  $X$  of a Lie algebra is *adj-nilpotent* if  $\text{adj}_X$  is nilpotent.

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 linear Lie algebra  $\mathfrak{g}$  acting on a vector space  $V$ , we define an eigenvector of  $\mathfrak{g}$  to be a vector  $x \in V$  such that  $Xx$  is a scalar multiple of  $x$  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  $x$  is an eigenvector for  $\mathfrak{g}$ , then we may define a function  $\lambda : \mathfrak{g} \rightarrow K$  so that the equation  $Xx = \lambda(X)x$  holds for all  $X$ .  $\lambda$  is a linear map, because

$$\lambda(X + Y)x = (X + Y)x = Xx + Yx = [\lambda(X) + \lambda(Y)]x$$

We call such a map a **weight** for  $\mathfrak{g}$ . In general, for a linear functional  $\lambda \in \mathfrak{g}^*$ , we define  $V_\lambda$  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  $V_\lambda$  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 satisfy powerful invariance properties which, like for the theory of the Jordan normal form in linear algebra, 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  $x \in V$  such that  $Xx = 0$  for all  $X \in \mathfrak{a}$  is  $\mathfrak{g}$ -invariant.

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

$$XYx = YXx = Y0 = 0$$

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.

**Lemma 2.4.** If  $\mathfrak{a}$  is an ideal in a linear Lie algebra  $\mathfrak{g}$  over some vector space  $V$ . If we fix  $X \in \mathfrak{g}$  and  $Y \in \mathfrak{a}$ , then for any weight  $\lambda : \mathfrak{a} \rightarrow K$ , and  $x \in V_\lambda$ ,

$$XY^n x = \lambda(X)(Y^n x) + \text{span}(x, Yx, \dots, Y^{n-1}x)$$

*Proof.* For  $n = 0$ , the theorem is just by assumption, since  $x$  is in the weight space of  $\lambda$ . Then, by induction, since  $[X, Y] \in \mathfrak{a}$ , we find

$$\begin{aligned} XY^n x &= YXY^{n-1}x + [X, Y]Y^{n-1}x \\ &= \lambda(X)Y^n x + (\lambda[X, Y])Y^{n-1}x + \text{span}(x, Yx, \dots, Y^{n-1}x) \\ &= \lambda(X)Y^n x + \text{span}(x, Yx, \dots, Y^{n-1}x) \end{aligned}$$

and this verifies the theorem for all  $n$ .  $\square$

**Theorem 2.5.** 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  $x \in V$  satisfies  $Xx = \lambda(X)x$  for all  $X \in \mathfrak{a}$ , then we must show  $XYx = \lambda(X)Yx$  for any  $Y \in \mathfrak{g}$ . Note that

$$XYx = YXx + [X, Y]x = \lambda(X)Yx + \lambda[X, Y]x$$

Thus it suffices to verify that  $\lambda[X, Y] = 0$  for any  $X \in \mathfrak{a}$ , and  $Y \in \mathfrak{g}$ . Now let  $W$  be the  $Y$ -invariant subspace of  $V$  generated by  $x$ . But  $X$  and  $Y$  both restrict to endomorphisms of  $W$ , hence  $[X, Y] = XY - YX \in \mathfrak{sl}(W)$  has trace zero, and we conclude that  $m\lambda[X, Y] = 0$ , hence  $\lambda[X, Y] = 0$ .  $\square$

This is the first theorem where we have to restrict the field upon which Lie algebras were defined. Classically, Lie algebras were only considered over the real and complex fields, and this theorem holds over both fields. But eventually we have to use the fact that the complex numbers are algebraically closed, mainly to ensure that eigenvectors exist for any linear endomorphism, so the classification was over the complex field. We will find that the techniques in this classification extend to any algebraically closed field of characteristic zero. For now, we don't assume any properties of the underlying base field of the Lie algebra, but we will soon have to restrict ourselves to algebraically closed field of characteristic zero. We'll call such fields **Cartan fields** (though this name is completely nonstandard, is it a useful name to shorten the statement of theorems).

The invariance theorem finds uses outside of Lie algebra theory, because it applies to arbitrary operators over finite dimensional vector spaces. For instance, here's a simple consequence about matrix algebras over an algebraically closed field of characteristic zero.

**Theorem 2.6.** *If  $V$  is a vector space over a Cartan field, 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  $\lambda$  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  $\mathfrak{g}$ , and  $\lambda$  acts as a weight on this ideal, hence the space  $V_\lambda$  of vectors  $v$  satisfying  $[X, Y]v = \lambda v$  is  $\mathfrak{g}$  invariant. This implies that  $X$  and  $Y$  restrict to operators on  $V_\lambda$ , and the restriction of  $[X, Y]$  to  $V_\lambda$  is the same as  $XY - YX$ , hence  $[X, Y]$  has trace zero. But  $[X, Y]$  can be put into normal form on  $V_\lambda$ , and we find it has trace  $n\lambda$  if  $V_\lambda$  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 (which acts as the infinitesimals of the normalizer of the subgroup corresponding to  $\mathfrak{h}$ ), 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  $[X, \mathfrak{h}] \subset \mathfrak{h}$  and  $[Y, \mathfrak{h}] \subset \mathfrak{h}$ , then surely  $[aX + bY, \mathfrak{h}] \subset \mathfrak{h} + \mathfrak{h} = \mathfrak{h}$ , and for any  $Z \in \mathfrak{h}$ ,

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

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

**Theorem 2.7.** *If  $\mathfrak{d}$  is the subalgebra of  $\mathfrak{gl}_n(K)$  consisting of diagonal matrices over a field of 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 spanned by some  $E_{ii}$  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 and Lie'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 2.8.** *If  $\mathfrak{g}$  is a linear Lie algebra over a nonzero vector space  $V$  consisting only of nilpotent operators, then there is a nonzero  $x$  for which  $Xx = 0$  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  $X^n x \neq 0$ , then  $X(X^n x) = X^{n+1} x = 0$ , so  $X^n x$  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  $[X, Y] - [X, Z] = [X, Y - Z] \in \mathfrak{a}$ . Because  $X$  is nilpotent,  $\text{adj}_X : \mathfrak{g} \rightarrow \mathfrak{g}$  is also nilpotent, hence so is  $\text{adj}_X : \mathfrak{g}/\mathfrak{a} \rightarrow \mathfrak{g}/\mathfrak{a}$ , and we may apply induction to the set of all  $\text{adj}_X$  conclude that there is  $Y \notin \mathfrak{a}$  such that  $[X, Y] \in \mathfrak{a}$  for all  $X \in \mathfrak{g}$ . This implies that  $\mathfrak{a} + KY$  is a subalgebra of  $\mathfrak{g}$ , and by maximality we see  $\mathfrak{a} + KY = \mathfrak{g}$ , so  $\mathfrak{a}$  has codimension one, and is actually an ideal of  $\mathfrak{g}$ . We can therefore apply induction to conclude that the space  $W$  of  $x \in V$  with  $Xx = 0$  for all  $X \in \mathfrak{a}$  is nontrivial. Since  $\mathfrak{a}$  is an ideal,  $W$  is invariant under multiplication by  $\mathfrak{g}$ , and in particular  $YW \subset W$ . Since  $Y$  is nilpotent, it therefore follows that we can find a nonzero  $x \in W$  with  $Yx = 0$ , and then  $(\mathfrak{a} + KY)x = 0$ .  $\square$

**Theorem 2.9** (Engel). *For any linear 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  $x$  such that  $Xx = 0$  for all  $X \in \mathfrak{g}$ . If  $W = Kx$ , 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  $x_1 + W, \dots, x_m + W$  of  $V/W$  in which all  $X$  are upper triangular. Then  $x, x_1, \dots, x_m$  is a basis for  $V$ , and

$$Xx_i \in \text{span}(x_1, \dots, x_{i-1}) + \text{span}(x)$$

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

There is an analogous formalism in which we can state Engel's theorem. A **flag** on an  $n$  dimensional vector space  $V$  is a strictly increasing chain  $V_0 \subset V_1 \subset \dots \subset V_n$ , with  $V_k$  a  $k$  dimensional subspace of  $V$ . As an example, note that a Lie algebra  $\mathfrak{g}$  is solvable if and only if there is a chain  $\mathfrak{g}_k$  with  $\mathfrak{g}_k$  an ideal in  $\mathfrak{g}_{k+1}$ . Engel's theorem says that for any linear Lie

algebra  $\mathfrak{g}$  of nilpotent linear transformations on a vector space  $V$ , there is a flag  $V_k$  with  $XV_k \subset V_{k-1}$  for each  $V_k$ . If we take some nonzero  $x_1 \in V_1$ , and subsequently extend  $x_1$  to a sequence  $\{x_1, \dots, x_n\}$  with  $\{x_1, \dots, x_k\}$  a basis for  $V_k$ , then each element of  $\mathfrak{g}$  is upper triangular with respect to this basis.

**Corollary 2.10.**  *$\mathfrak{g}$  is nilpotent if and only if  $\text{adj}_X$  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  $\text{adj}_X$  are nilpotent, then Engel's theorem implies the existence of a basis of  $X_i$  such that all  $\text{adj}$  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 2.11.** *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 two dimensional subalgebra. Then either  $\mathfrak{h}$  is abelian, or we can choose  $X$  and  $Y$  such that  $[X, Y] = X$ . In the second case,  $\text{adj}_Y^n(X) = X$ , so  $\text{adj}_Y$  is not nilpotent, hence  $\mathfrak{g}$  cannot be nilpotent. Thus every two 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  $K$ , the adjoint  $\text{adj}_X$  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. However, the analogous theorem for solvable Lie algebras is not so easy to generalize, and is key to the classification of Lie algebras.

**Lemma 2.12.** *If  $\mathfrak{g}$  is a solvable linear Lie algebra over a nonzero vector space  $V$ , whose base field is Cartan, then there is a nonzero  $x$  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 one, it suffices to show that every matrix has an eigenvector, and this follows because  $K$  is algebraically complete. 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 one. Since  $\mathfrak{g}/\mathfrak{g}'$  is abelian, every subspace is an ideal, hence every subspace of  $\mathfrak{g}$  containing  $\mathfrak{g}'$  is an ideal of  $\mathfrak{g}$ . We can therefore apply induction to find an eigenvector  $x$  of  $\mathfrak{a}$  with corresponding weight  $\lambda : \mathfrak{a} \rightarrow K$ . We know that  $V_\lambda$  is  $\mathfrak{g}$  invariant, so that if we let  $\mathfrak{g} = \mathfrak{a} + KX$ , then  $Xx \in V_\lambda$  for each  $x \in V_\lambda$ , and we may apply algebraic completeness again to determine that there is  $x \in V_\lambda$  with  $Xx = \gamma x$  for some  $\gamma \in K$ . It then follows that  $x$  is an eigenvector for all elements of  $\mathfrak{g}$ , because we can write any element of  $\mathfrak{g}$  as  $\alpha X + \beta Y$ , with  $Y \in \mathfrak{a}$ , and then  $(\alpha X + \beta Y)x = [\alpha\gamma + \beta\lambda(Y)]x$ .  $\square$

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

**Theorem 2.13.** *If  $\mathfrak{g}$  is a solvable linear Lie algebra on a vector space  $V$  over a Cartan field, then there is a basis for  $V$  such that the matrix representations of  $\mathfrak{g}$  are upper triangular.*

*Proof.* We choose an eigenvector  $x$  for  $\mathfrak{g}$ , and then consider  $V/W$  for  $W = Kx$ . 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  $x_1 + W, \dots, x_n + W$ . Passing back up, we find that each element of  $\mathfrak{g}$  is upper triangular with respect to the basis  $x, x_1, \dots, x_n$ .  $\square$

Equivalently, in the language of chains,  $V$  has a flag  $\{V_k\}$  such that each element of  $\mathfrak{g}$  **stabilizes** the flag, in the sense that if  $X \in \mathfrak{g}$ , then  $XV_k \subset V_k$  for each  $k$ .

**Corollary 2.14.** *If  $\mathfrak{g}$  is a solvable Lie algebra over a Cartan field, then there is a flag  $\{\mathfrak{g}_k\}$  where each  $\mathfrak{g}_k$  is an ideal in  $\mathfrak{g}$ .*

*Proof.* Apply Lie's theorem to the adjoint representation  $\text{Der } \mathfrak{g} \subset \mathfrak{gl}(\mathfrak{g})$ , and note that an ideal of  $\mathfrak{g}$  is exactly one stabilized by  $\text{Der } \mathfrak{g}$ .  $\square$

If  $\mathfrak{g}$  is a solvable linear Lie algebra over an algebraically closed field of characteristic zero, then Lie's theorem tells us there is a basis of the



underlying vector space 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! This strategy can be applied to obtain a more general result.

**Theorem 2.15.** *Over a Cartan field,  $\mathfrak{g}$  is solvable if and only if  $\mathfrak{g}'$  is nilpotent.*

*Proof.* There is a basis of  $\mathfrak{g}$  such that each element of  $\text{adj } \mathfrak{g}$  is upper triangularized. It therefore follows that  $\text{adj}_{[X,Y]}$  is strictly upper triangular, hence nilpotent, and therefore by Engel's theorem, we find that  $\mathfrak{g}'$  is nilpotent. Conversely, if  $\mathfrak{g}'$  is nilpotent, then  $\mathfrak{g}'$  is solvable, and  $\mathfrak{g}/\mathfrak{g}'$  is abelian, hence  $\mathfrak{g}$  is solvable.  $\square$

Lie's theorem does not hold over all fields, unlike Engel's theorem.

**Example.** *Over a field  $K$  of characteristic  $p$ , the  $p \times p$  matrices*

$$X = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p-1 \end{pmatrix}$$

which can be written  $X = E_{p1} + \sum_{k=1}^{p-1} E_{p(p+1)}$  and  $Y = \sum_{k=1}^p (k-1)E_{kk}$ , satisfy

$$\begin{aligned} [X, Y] &= \sum_{k=1}^p (k-1)[E_{p1}, E_{kk}] + \sum_{i=1}^p \sum_{j=1}^p (j-1)[E_{i(i+1)}, E_{jj}] \\ &= -(p-1)E_{p1} + \sum_{i=1}^p (i - (i-1))E_{i(i+1)} = X \end{aligned}$$

Thus  $X$  and  $Y$  span a solvable subalgebra of  $\mathfrak{g}$ , yet  $X$  and  $Y$  have no common eigenvector, because any such  $x$  with eigenvalues  $\lambda$  and  $\gamma$  over  $X$  and  $Y$  must

satisfy  $x_1 = \lambda x_n$ , and  $x_{k+1} = \lambda x_k$  for  $k < p$ . This shows that  $\lambda^p = \lambda = 1$ . Looking at the definition of  $Y$ , we conclude that  $x$  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 Basic 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  $Xx = \rho(X)(x)$ . Conversely, if we have an action  $\mathfrak{g} \times V \rightarrow V$  of  $\mathfrak{g}$  on  $V$ , satisfying

$$\begin{aligned}(\lambda X + \gamma Y)v &= \lambda Xv + \gamma Yv \\ X(\mu v + \nu w) &= \mu Xv + \nu Xw \\ [X, Y]v &= X(Yv) - Y(Xv)\end{aligned}$$

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  $x \in V$  to any  $y \in V$  by some  $X \in \mathfrak{g}$ . Lie algebras were invented to study Lie groups, and in the representation theory of Lie algebras the correspondence remains stronger than ever. Any Lie group representation  $\rho : G \rightarrow GL(V)$  gives rise to a Lie algebra representation  $\rho_* : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ , and if  $G$  is simply connected, every Lie algebra representation gives rise to a unique Lie group representation.

**Example.** For any Lie group  $G$ , the adjoint representation  $Adj : G \rightarrow GL(\mathfrak{g})$  is canonical, and gives rise to a representation  $adj : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$ , which is just the adjoint representation of the Lie algebra given by  $adj(X)(Y) = [X, Y]$ . The

adjoint representation of any Lie algebra is well defined by the same formula on any Lie algebra. We have already seen that the adjoint map is really a map into the Lie algebra  $\text{Der}(\mathfrak{g})$  of derivations on  $\mathfrak{g}$ . 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 class of semisimple Lie algebras.

**Example.** If  $\mathfrak{g}$  is a subalgebra of  $\mathfrak{gl}(V)$ , then the inclusion of  $\mathfrak{g}$  in  $\mathfrak{gl}(V)$  gives a faithful Lie algebra representation. In terms of Lie groups,  $\mathfrak{g}$  corresponds to a Lie subgroup  $G$  of  $GL(V)$ , and the inclusion of  $G$  in  $GL(V)$  can be differentiated to yield the inclusion of  $\mathfrak{g}$  in  $\mathfrak{gl}(V)$ .

**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  $v \in V$ . This corresponds to the trivial Lie group representation with  $gv = v$  for all  $g \in G$  and  $v \in V$ .

**Example.** On any field  $K$ , the Lie algebra representations of  $K$  over a vector space  $V$  are classified by the endomorphisms  $T : V \rightarrow V$ , by defining  $\rho(x) = xT$ . If  $K = \mathbf{C}$ , then the representations of the additive Lie group  $\mathbf{C}$  over  $V$  are given by  $\rho(z) = e^{zT}$ . It follows that  $\rho(tz) = e^{tzT} = 1 + tzT + o(t^2)$ , and this gives us all representations of the Lie algebra  $\mathbf{C}$  on  $V$ .

**Example.** The one dimensional representations of a Lie group  $G$  on  $\mathbf{C}$  are classified by the differentiable characters  $\chi : G \rightarrow \mathbf{C}$  of the group. The one dimensional representations of a Lie algebra  $\mathfrak{g}$  on  $\mathbf{C}$  are given by the linear functionals  $\lambda : \mathfrak{g} \rightarrow \mathbf{C}$ , which vanish on the derived subalgebra  $\mathfrak{g}'$ , defined by letting  $Xz = \lambda(X)z$ . Thus if  $G$  is simply connected, for every linear functional  $\lambda$  which vanishes on  $\mathfrak{g}'$ , there is a character  $\chi$  on  $G$  whose differential at  $e$  is  $\lambda$ .

**Example.** Recall that if  $\rho : G \rightarrow GL(V)$  and  $\eta : G \rightarrow GL(W)$  are two representations of a Lie group  $G$ , we can define a representation  $\rho \otimes \eta : G \rightarrow GL(V \otimes W)$  by the formula

$$(\rho \otimes \eta)(g)(v \otimes w) = \rho(g)(v) \otimes \eta(g)(w)$$

Since the tensor product is a bilinear operation, it follows that

$$(\rho \otimes \eta)_*(X)(v \otimes w) = \rho_*(X)(v) \otimes w + v \otimes \eta_*(X)(w)$$

Given an arbitrary action of a Lie algebra  $\mathfrak{g}$  on two spaces  $V$  and  $W$ , we can define the tensor product of the actions to be the action of  $\mathfrak{g}$  on  $V \otimes W$  given

by  $g(v \otimes w) = (gv) \otimes w + v \otimes (gw)$ . Similar processes give us representations of other tensors. Indeed, we can consider successive tensor products  $V_1 \otimes V_2 \otimes \cdots \otimes V_n$ , and we find that since the Lie group representation on the product is invariant of which order the tensor products are taken, the corresponding Lie algebra representation is also invariant of the order, and is just

$$g(v_1 \otimes \cdots \otimes v_n) = (gv_1) \otimes \cdots \otimes v_n + v_1 \otimes (gv_2) \otimes \cdots \otimes v_n + \cdots + v_1 \otimes \cdots \otimes (gv_n)$$

and this generalizes to giving a representation of  $\mathfrak{g}$  on the algebra  $\bigotimes V$  of tensors over  $V$ , though this representation is not finite dimensional.

**Example.** If  $\rho : G \rightarrow GL(V)$  is a representation of a Lie group, and  $W$  is a subspace of  $V$  invariant under the action of  $G$ , then we have a natural quotient representation  $\rho' : G \rightarrow GL(V/W)$ , and by differentiating this relation, we find that if  $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$  is a representation that fixes  $W$ , then  $\rho' : \mathfrak{g} \rightarrow \mathfrak{gl}(V/W)$  is a representation of the Lie algebra. Noting that for any representation of  $\mathfrak{g}$  on  $V$ ,  $\mathfrak{g}$  fixes the span of all  $v \otimes w - w \otimes v$  in  $V \otimes V$ , as well as the span of all  $v \otimes w + w \otimes v$ , and so we obtain canonical representations of  $\mathfrak{g}$  over  $\text{Sym}(V)$  and  $\Lambda V$ . In  $\text{Sym}(V)$ , we find that  $X(v^2) = (Xv)v + v(Xv) = 2v(Xv)$ , which corresponds quite elegantly to the corresponding Lie group relation  $g(v^2) = (gv)^2$ .

**Example.** If  $\rho : G \rightarrow GL(V)$  is a representation, then we define the dual representation as  $\rho^* : G \rightarrow GL(V^*)$  by  $\rho^*(g)(f)(v) = f(\rho(g^{-1})(v))$ . It therefore follows that  $\rho^*(X)(f)(v) = -f(\rho_*(X))$ , and so given an arbitrary action of a Lie algebra  $\mathfrak{g}$  on  $V$ , we define the dual action of  $\mathfrak{g}$  on  $V^*$  to be  $(gf)(v) = -f(gv)$ .

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. 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. If  $G$  is a connected Lie group, then a representation of the Lie group is irreducible if and only if the corresponding Lie algebra representation is irreducible.

Lie's theorem says that every module over a complex solvable Lie algebra has a one dimensional submodule, so Lie's theorem essentially implies that the only irreducible representations of a solvable Lie algebra are one-dimensional. We can generalize this to arbitrary complex Lie algebras, we find that the only irreducible representations of a solvable Lie algebra are one-dimensional. We can use the decomposition of any Lie algebra by its radical.

**Theorem 2.16.** *If  $\mathfrak{g}$  is a Lie algebra over a Cartan field, then every irreducible representation of  $\mathfrak{g}$  is isomorphic to a representation of the form  $V = V_0 \otimes W$ , where  $V_0$  is an irreducible representation of  $\mathfrak{g}/\text{rad}(\mathfrak{g})$ , and  $W$  is a one-dimensional representation.*

*Proof.* By Lie's theorem, there is  $\lambda \in \text{rad}(\mathfrak{g})$  such that the space  $W$  of all  $v$  such that  $Xv = \lambda(X)v$  for all  $X \in \text{rad}(\mathfrak{g})$  is nonzero. Since  $\text{rad}(\mathfrak{g})$  is an ideal of  $\mathfrak{g}$ ,  $W$  is a submodule of  $V$ , and by irreducibility, we conclude that  $W = V$ . Since  $\text{tr}(X) = \dim(V)\lambda(X)$ ,  $\lambda$  vanishes on  $\text{rad}(\mathfrak{g}) \cap \mathfrak{g}'$ . If  $\lambda$  is extended to all of  $V$ , and vanishes on  $\mathfrak{g}'$ , then we may define a one-dimensional irreducible representation  $U$  on  $\mathfrak{g}$  by letting  $Xv = \lambda(X)v$ . Then the representation on  $U^*$  satisfies  $Xf = -\lambda(X)f$ , so if  $X \in \mathfrak{g}'$ , then  $\lambda(X) = 0$ , so  $Xf = 0$  for  $f \in U^*$ , and  $Xv = 0$  for  $v \in V$ . It follows that  $V \otimes U^*$  is a representation upon which  $\text{rad}(\mathfrak{g})$  acts trivially, because if  $X \in \text{rad}(\mathfrak{g})$ ,

$$X(v \otimes f) = (Xv) \otimes f + v \otimes (Xf) = \lambda(X)(v \otimes f) - \lambda(X)(v \otimes f) = 0$$

Thus  $V \otimes U^*$  factors into a representation of  $\mathfrak{g}/\text{rad}(\mathfrak{g})$ , and  $(V \otimes U^*) \otimes U = V \otimes (U^* \otimes U)$  is really just a new representation on the vector space  $V$ , where if  $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$  is the original representation, the new representation is  $\rho' : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ , where  $\rho'(X) = \rho(X) + \lambda(X)$ . In particular,  $U^* \otimes U$  is the trivial representation, so  $(V \otimes U^*) \otimes U$  is  $\mathfrak{g}$  isomorphic to  $V$ .  $\square$

**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 2.17 (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  $\lambda$  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 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,  $\rho$  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

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

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 2.18.**  $V_n$  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 2.19.** *If  $v \in V$  is an eigenvector of  $h$  with eigenvalue  $\lambda$ , 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 2.20.** *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  $\lambda$ , 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 2.21.** *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  $\lambda$ . 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

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

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  $\psi$  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 2.22.** *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  $ev = 0$ , and  $hv = mv$  for some integer  $m$ , then 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.

# Chapter 3

## Killing Forms

### 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. This map is not very easy to work with. We don't necessarily have  $(T + S)_s = T_s + S_s$  or  $(T + S)_n = T_n + S_n$ , unless  $T$  and  $S$  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 Cartan*.

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.

**Lemma 3.1.** *If  $V$  and  $W$  are two subspaces of endomorphisms in  $\mathfrak{gl}(U)$ , where  $U$  is finite dimensional, and any  $X \in \mathfrak{gl}(V)$  with  $[X, V] \subset W$  satisfies  $\text{tr}(XY) = 0$  for all  $Y$  with  $[Y, V] \subset W$ , then  $X$  is nilpotent.*

*Proof.* Using the Jordan decomposition, write any  $X = X_s + X_n$ , and fix a basis such that  $X_s$  is diagonalized, and  $X_n$  is upper triangular. Since we can write  $X_s$  as a polynomial in  $X$  with no constant coefficient, then  $\text{adj}_{X_s}$  maps  $V$  into  $W$ . It will suffice to show that  $X_s = 0$ . Let  $\lambda_1, \dots, \lambda_n$  be the eigenvalues of  $X_s$  in  $K$ , and consider the rational vector subspace  $E$  of  $K$  generated by the  $\lambda_i$ . We shall show  $E = 0$ , so that each  $\lambda_i = 0$ . To prove this, it suffices to show  $E^* = 0$ . Given a particular linear  $f : E \rightarrow \mathbf{Q}$ , consider the diagonal matrix  $Y = \text{diag}(f(\lambda_1), \dots, f(\lambda_n))$  in  $\mathfrak{gl}(V)$ . Then

$$[Y, E_{ij}] = (f(\lambda_i) - f(\lambda_j))E_{ij}$$

Let  $g(X) \in K[X]$  be a polynomial such  $g(\lambda_i - \lambda_j) = f(\lambda_i) - f(\lambda_j)$  (constructed by Lagrange interpolation). Then  $g(0) = 0$  has constant coefficients, so that if  $\text{adj}_{X_s} : V \rightarrow W$ . We find  $\text{adj}_Y = g(\text{adj}_{X_s})$ , and since  $\text{adj}_{X_s} : V \rightarrow W$ , we find that  $\text{adj}_Y : V \rightarrow W$  as well, hence

$$\text{tr}(XY) = \text{tr}(X_s Y) = \sum \lambda_i f(\lambda_i) = 0$$

which implies that  $f(\text{tr}(XY)) = \sum f(\lambda_i)^2 = 0$ , so  $f(\lambda_i) = 0$  for all  $i$ , and hence  $f = 0$  since the  $\lambda_i$  form a basis for  $E$ . This shows that  $E^* = 0$ , and therefore that  $E = 0$ .  $\square$

**Theorem 3.2.** *If  $\mathfrak{g}$  is a Lie subalgebra of  $\mathfrak{gl}(V)$  such that  $\text{tr}(XY) = 0$  for  $X \in \mathfrak{g}'$ ,  $Y \in \mathfrak{g}$ , 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. Let  $V = \mathfrak{g}$ , and  $W = \mathfrak{g}'$ . Our hypothesis says that  $\text{tr}(XY) = 0$  for  $X \in \mathfrak{g}$ ,  $Y \in \mathfrak{g}'$ , and to apply our last lemma we need to prove that  $\text{tr}(XM) = 0$  for all  $X \in \mathfrak{g}'$  with  $[M, \mathfrak{g}] \subset \mathfrak{g}'$ . But for any  $X, Y \in \mathfrak{g}$ ,

$$\text{tr}([X, Y]M) = \text{tr}(X[Y, M]) = 0$$

which is zero by hypothesis, since  $[Y, M] \in \mathfrak{g}'$ . Thus any  $[X, Y]$  is verified to be nilpotent by the last lemma. But since any element of  $\mathfrak{g}'$  is a linear span of elements of this form, we include that every element of  $\mathfrak{g}'$  is nilpotent.  $\square$

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  $\text{adj}_X$  is nilpotent for any  $X \in \mathfrak{g}$ , and there is a basis for  $\mathfrak{g}$  upon which the  $\text{adj}_X$  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 3.3.** 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  $\kappa$  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  $\text{adj}_X$  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  $\kappa$  is **non-degenerate** if there is no  $X \in \mathfrak{g}$  such that  $\kappa(X, Y) = 0$  for all  $Y$ .

**Lemma 3.4.** *A Lie algebra is semisimple iff 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 3.5** (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  $\kappa$  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  $\kappa$  is degenerate.  $\square$

**Example.** The Killing form on  $\mathfrak{gl}_n(K)$  is found by noting that

$$\text{adj}_X(E_{ij}) = [X, E_{ij}] = \sum_k X_{ki} E_{kj} - X_{jk} E_{ik}$$

so

$$\begin{aligned} (\text{adj}_Y \circ \text{adj}_X)(E_{ij}) &= \sum_k X_{ki} [Y, E_{kj}] - X_{jk} [Y, E_{ik}] \\ &= \sum_{k,l} X_{ki} (Y_{lk} E_{lj} - Y_{jl} E_{kl}) - X_{jk} (Y_{li} E_{lk} - Y_{kl} E_{il}) \end{aligned}$$

Hence

$$\begin{aligned} \kappa(X, Y) &= \text{tr}(\text{adj}_X \circ \text{adj}_Y) \\ &= \sum_{i,j} \left( \sum_k X_{ki} Y_{ik} \right) - X_{ii} Y_{jj} - X_{jj} Y_{ii} + \left( \sum_k X_{jk} Y_{kj} \right) \\ &= \sum_{i,j} (YX)_{ii} + (XY)_{jj} - X_{ii} Y_{jj} - Y_{ii} X_{jj} \\ &= 2[\text{tr}(XY) - \text{tr}(X)\text{tr}(Y)] \end{aligned}$$

The Killing form on  $\mathfrak{sl}_n(K)$  is the same, since  $\mathfrak{sl}_n(K)$  is an ideal of  $\mathfrak{gl}_n(K)$ . The form is non-degenerate, because if there was  $X$  such that  $\text{tr}(XY) = \text{tr}(X)\text{tr}(Y)$  for all  $Y$ , then if we let  $Y = E_{ij}$  we find  $X_{ji} = 0$  for  $j \neq i$ , from which we conclude  $X$  is a diagonal matrix, and then  $X_{ii} = nX_{ii}$  for any  $i$ , hence  $X_{ii} = 0$ . Thus  $\mathfrak{gl}_n(K)$  and  $\mathfrak{sl}_n(K)$  are both semisimple.

These criteria 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 and only if  $\mathfrak{g}^\perp = (0)$ . Note that if  $\mathfrak{a}$  is an ideal, then  $\mathfrak{a}^\perp$  is also an ideal, because the Killing form is associative.



**Lemma 3.6.** *If  $\mathfrak{a}$  is an ideal of a semisimple Lie algebra  $\mathfrak{g}$ , then  $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$ .*

*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  $\mathfrak{a}^\perp$ .  $\square$

**Corollary 3.7.** *Every nonzero ideal of a semisimple Lie algebra is semisimple.*

*Proof.* If  $\mathfrak{g}$  is semisimple, with an ideal  $\mathfrak{a}$ , then  $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$ , and so every ideal of  $\mathfrak{a}$  is also an ideal of  $\mathfrak{g}$ , hence  $\mathfrak{a}$  can have no nonzero abelian ideals.  $\square$

If  $\mathfrak{g}$  is a semisimple Lie algebra which is not simple, we may find a nonzero ideal  $\mathfrak{a}$  not equal to  $\mathfrak{g}$ . We then find  $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$ , where  $\mathfrak{a}$  and  $\mathfrak{a}^\perp$  are both semisimple Lie algebras of dimension less than  $\mathfrak{g}$ . Recursively then, we find that every semisimple Lie algebra is the direct sum of simple Lie algebras. Conversely, the direct sum of two semisimple Lie algebras  $\mathfrak{g}$  and  $\mathfrak{h}$  is semisimple, because if  $\mathfrak{a}$  is a solvable ideal, then  $\mathfrak{a} \cap \mathfrak{g}$  and  $\mathfrak{a} \cap \mathfrak{h}$  are both solvable, so  $\mathfrak{a} \cap \mathfrak{g} = \mathfrak{a} \cap \mathfrak{h} = 0$ , and so we find that for any  $X \in \mathfrak{g}$ ,  $Y \in \mathfrak{h}$ ,  $Z \in \mathfrak{a}$ ,  $[Z, X + Y] = [X, Z] + [Y, Z] = 0$ , so that  $\mathfrak{a} \subset Z(\mathfrak{g} \oplus \mathfrak{h}) = Z(\mathfrak{g}) \oplus Z(\mathfrak{h}) = 0$ . Thus a semisimple Lie algebra is exactly a Lie algebra which can be written as the direct sum of simple Lie algebras.

**Theorem 3.8.** *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  $\text{Der } \mathfrak{g}$ . 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 on a semisimple Lie algebra. It is essentially the unique bilinear form of its kind on simple ideals. 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 by associativity

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

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  $\kappa$ , 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 may apply Schur's lemma to conclude that there is  $\lambda$  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 3.9.** *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  $\mathfrak{g}_\lambda$  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  $\mathfrak{g}_\lambda$ . 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  $\lambda$  eigenspace for  $D_s$  is  $\mathfrak{g}_\lambda$ . 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 3.10.** *If  $\mathfrak{g}$  is semisimple, then any  $X \in \mathfrak{g}$  can be written uniquely as  $D + N$ , where  $\text{adj}_D$  is diagonalizable,  $\text{adj}_N$  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  $\text{adj}_D$  and  $\text{adj}_N$  commute,  $[D, N] = 0$ . If  $Y$  commutes with  $X$ , then  $\text{adj}_X(Y) = 0$ , then because  $\text{adj}_D$  and  $\text{adj}_N$  are polynomials in  $\text{adj}_X$ , we see

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

hence  $\text{adj}_N(Y) = c_0$ . But since  $\text{adj}_N$  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  $\text{adj}_X$  is diagonalizable). In the case where our semisimple Lie algebra is a matrix Lie algebra, the abstract decomposition agrees with the concrete decomposition.

**Theorem 3.11.** *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  $\rho(X)_s = \rho(D)$  and  $\rho(X)_n = \rho(N)$ .*

*Proof.* The image of any representation of a complex semisimple Lie algebra is semisimple, since if  $\mathfrak{h}$  is the kernel of  $\rho$ , 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  $\text{adj}_Y$  is as well. If  $\text{adj}_D$  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  $\text{adj}_N$  is nilpotent, then  $\text{adj}_N$  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 The Casimir Element and Weyl's Theorem

The Killing form  $\kappa : \mathfrak{g} \times \mathfrak{g} \rightarrow K$  is only one of a family of bilinear forms which can be generated on a Lie algebra. We note that it can be easily generalized based on its dependence on the adjoint representation. Given an arbitrary representation  $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ , we define the **trace form** of the representation to be the symmetric bilinear map  $\beta(X, Y) = \text{tr}(\rho(X) \circ \rho(Y))$ . The form is associative, for the exact same reason that the Killing form is associative.

**Lemma 3.12.** *If  $\rho$  is a faithful representation of a semisimple Lie algebra, then  $\beta$  is nondegenerate.*

*Proof.* Define the radical  $\text{rad}(\beta)$  of the bilinear map  $\beta$  to be the set of all  $X$  such that  $\beta(X, Y) = 0$  for all  $Y$ . Then  $\text{rad}(\beta)$  is an ideal of  $\mathfrak{g}$ , because of the associativity of  $\beta$ . But this implies that  $\rho(\text{rad}(\beta))$  is a solvable algebra, since the trace vanishes on products in this algebra, and since  $\rho$  is faithful,  $\text{rad}(\beta)$  is solvable. But this means that  $\text{rad}(\beta) = 0$ , since the Lie algebra is semisimple.  $\square$

We can therefore use  $\beta$  to identify  $\mathfrak{g}$  and  $\mathfrak{g}^*$  canonically. Given  $X \in \mathfrak{g}$ , we let  $X^* \in \mathfrak{g}^*$  be the map such that  $X^*(Y) = \beta(X, Y)$ . Given a basis  $X_1, \dots, X_n$  of  $\mathfrak{g}$ , find  $Y_1, \dots, Y_n$  such that  $\beta(X_i, Y_j) = \delta_{ij}$ .

**Lemma 3.13.** *If  $X \in \mathfrak{g}$ , and  $[X_i, X] = \sum a_{ij} X_j$ , then  $[X, Y_j] = \sum a_{ji} Y_i$ .*

*Proof.* We find

$$\beta([X_i, X], Y_j) = \sum a_{ik} \beta(X_k, Y_j) = a_{ij}$$

If  $[X, Y_j] = \sum b_{ji} Y_i$ , then

$$a_{ij} = \beta([X_i, X], Y_j) = \beta(X_i, [X, Y_j]) = \sum b_{jk} \beta(X_i, Y_k) = b_{ji}$$

$\square$

Now given the trace form  $\beta$ , we find two basis  $X_1, \dots, X_n$  and  $Y_1, \dots, Y_n$  as described above, and construct the **Casimir operator** of the representation, which is the linear map

$$c_\rho(x) = \sum X_i(Y_i x)$$

The map is a  $\mathfrak{g}$  module homomorphism, for if  $[X_i, X] = \sum a_{ij} X_j$ ,

$$\begin{aligned} c_\rho(Xx) &= \sum_{i=1}^n X_i(Y_i(Xx)) = \sum_{i=1}^n X_i(X(Y_i x)) - X_i([Y_i, X]x) \\ &= \sum_{i=1}^n X(X_i(Y_i x)) - [X_i, X](Y_i x) - X_i([Y_i, X]x) \\ &= \sum_{i=1}^n X(X_i(Y_i x)) - \sum_{j=1}^n a_{ij} X_j(Y_i x) + a_{ji} X_i(Y_j x) = X(c_\rho(x)) \end{aligned}$$

and

$$\text{tr}(c_\rho) = \sum \text{tr}(\rho(X_i) \circ \rho(Y_i)) = \sum \beta(X_i, Y_i) = \dim(\mathfrak{g})$$

It is satisfying to hear that the Casimir operator is basis independent.

**Theorem 3.14.** *If  $X_1, \dots, X_n$  and  $X'_1, \dots, X'_n$  are two bases of  $\mathfrak{g}$ , then the induced Casimir operators  $c_\rho^X$  and  $c_\rho^{X'}$  are the same.*

*Proof.* If  $Y_1, \dots, Y_n$  and  $Y'_1, \dots, Y'_n$  are the induced dual basis, then  $X_i = \sum \beta(X_i, Y'_j) X'_j$ , and  $Y_i = \sum \beta(Y_i, X'_j) Y'_j$ , hence

$$\begin{aligned} \beta(Y_j, X_l) &= \sum_{il} \beta(X'_l, Y_j) \beta(Y'_i, X_k) \beta(X'_i, Y'_l) \\ &= \sum \beta(X'_i, Y_j) \beta(Y'_i, X_k) = \delta_{jl} \end{aligned}$$

Then we just carry out the calculation

$$\begin{aligned} c_\rho^{X'} &= \sum_i \rho(X'_i) \circ \rho(Y'_i) = \sum_{ijk} \beta(X'_i, Y_j) \beta(Y'_i, X_k) [\rho(X_j) \circ \rho(Y_k)] \\ &= \sum \rho(X_i) \circ \rho(Y_i) = c_\rho^X \end{aligned}$$

hence the Casimir operator is basis independent. □

The Casimir element is key to proving Weyl's theorem about the complete reducibility of the representations of semisimple Lie algebras.

**Theorem 3.15** (Weyl). *If  $\mathfrak{g}$  is semisimple, every finite dimensional  $\mathfrak{g}$  module is completely reducible.*

*Proof.* Let  $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$  be the representation. If  $W$  is a submodule of  $V$ , it suffices to show there is a submodule  $U$  of  $V$  such that  $V = W \oplus U$ , because by induction on the dimension of  $V$  we may break down the module into irreducible components. We may assume that the representation is faithful (otherwise consider representations over  $\mathfrak{g}/\ker(\rho)$ ).

A simple first case we will analyze is where  $W$  has codimension one in  $V$ . Then  $V/W$  is the trivial one dimensional  $\mathfrak{g}$  module, because  $\mathfrak{g}'$  acts trivially on one dimensional submodules of a module, and if  $\mathfrak{g}$  is semisimple then  $\mathfrak{g}' = \mathfrak{g}$ . Thus if  $X \in \mathfrak{g}$ , and  $x \in V$ , then  $Xx \in W$ .

- If  $W$  is simple, let  $U$  be the kernel of the Casimir operator  $c_\rho$ . If  $X \in \mathfrak{g}$  and  $x \in V$ , then  $c(x) \in W$  for all  $x$ , and this shows  $c$  has nontrivial kernel. This also shows  $c$  restricts to an endomorphism of  $W$ , and therefore by Schur's lemma there is  $\lambda \in K$  such that  $c(x) = \lambda x$  for all  $x \in W$ . Then  $\lambda$  is nonzero, for the trace of  $c$  over the whole space (since  $c(x) \in W$  for all  $x$ ) is  $m\lambda$ , where  $m$  is the dimension of  $W$ , and also  $\dim V$ , which is nonzero. Thus  $U \cap W = 0$ , and by dimension counting we find that  $U \oplus W = V$ , because  $W$  is  $m$  dimensional, and the kernel of  $c$  is  $n - m$  ( $c$  has rank  $m$ , because the image is  $W$ , which is  $m$  dimensional).
- Now suppose  $W$  has a proper submodule  $W_1$ . Then  $W/W_1$  is a submodule of  $V/W_1$  of codimension one, since  $(V/W_1)/(W/W_1) \cong V/W$ . By induction, we find that  $V/W_1 = W/W_1 \oplus L$ , for some one dimensional submodule  $L$  of  $V/W_1$ . Let  $L_1$  be the submodule of  $V$  containing  $W_1$  such that  $L_1/W_1 = L$ . Then  $\dim L_1 = 1 + \dim W_1 < \dim W$ , so we find  $L_1 = W_1 \oplus U$  for some submodule  $U$  of  $L_1$  by induction. We claim that  $V = W \oplus U$ . The direct sum decomposition of  $V/W_1$  implies that  $U \cap W \subset W_1$ , but the direct sum decomposition of  $L_1$  implies that  $U \cap W_1 = 0$ , hence  $U \cap W = 0$ . By dimension counting, we find  $V = W \oplus U$ , because  $U \neq 0$ .

Finally, we consider the general case. Let  $W$  be a  $\mathfrak{g}$  submodule of  $V$ . Then  $\text{Hom}(V, W)$  becomes a  $\mathfrak{g}$  module under the action

$$(Xf)(x) = X(f(x)) - f(Xx)$$

Let  $A$  be the set of homomorphisms which restrict to scalar multiples on  $W$ , and  $B$  the set of homomorphisms which vanish on  $W$ . Then  $A$  and  $B$  are both submodules of  $\text{Hom}_{\mathfrak{g}}(V, W)$ . What's more,  $A/B$  is one dimensional, so we may now apply Weyl's theorem in this special case to find  $C$  such that  $A = B \oplus C$ . Then  $C$  is one dimensional, and trivial as a  $\mathfrak{g}$  module, so we may find  $f \in C$  such that  $f(x) = 1$  for all  $x \in W$ , and  $Xf = 0$ . Since  $f$  is a  $\mathfrak{g}$ -module homomorphism, its kernel  $U$  is a submodule of  $V$ . We claim  $V = W \oplus U$ . Since  $f$  restricts to the identity on  $W$ , no element of  $W$  is in  $U$ . We also find that since  $f(V) = W$ ,  $f$  has rank  $m$  if  $W$  is  $m$  dimensional, and therefore  $U$  has dimension  $n - m$ , if  $n$  is the dimension of  $V$ . But then  $W \oplus U$  is  $n$  dimensional, so  $V = W \oplus U$ .  $\square$

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  $\pm 1$  of weight 2, and an eigenvalues  $\pm 2$  of weight 1, so that as a representation, the module is isomorphic to  $V_2 \oplus V_1 \oplus V_1 \oplus V_0$ .

### 3.3 Cartan Subalgebras

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  $\text{adj}_X$  is diagonalizable.
2. We took a basis of  $\mathfrak{g}$  consisting of eigenvectors for  $\text{adj}_X$ , 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 commutative elements of the Lie algebra. This subalgebra is known as the Cartan subalgebra.

Let us try and generalize the technique. Let  $\mathfrak{h}$  be an arbitrary abelian subalgebra of  $\mathfrak{g}$  consisting of semisimple elements. If  $X, Y \in \mathfrak{h}$ , then  $\text{adj}_X$  and  $\text{adj}_Y$  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  $\text{adj}_X$  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  $\lambda$  ranges over the set of weights for  $\mathfrak{h}$ , and the  $\mathfrak{g}_{\lambda}$  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 3.16.** *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. □



Since there are only finitely many weights, we conclude that if  $X \in \mathfrak{g}_\lambda$ , with  $\lambda \neq 0$ , then  $\text{adj}_X$  is nilpotent. Given any element  $Y \in \mathfrak{g}$ , write  $Y = \sum Y_\lambda$ , where  $Y_\lambda \in \mathfrak{g}_\lambda$ . Then  $[X, Y] = \sum [X, Y_\lambda] \in \sum \mathfrak{g}_{\lambda+\gamma}$ , 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  $\gamma$  with  $\mathfrak{g}_\gamma \neq 0$ , then  $\text{adj}_X^n = 0$ . This implies each  $\mathfrak{g}_\lambda$  is nilpotent, so we cannot expect any of these weight spaces to be ideals on a semisimple Lie algebra.

**Theorem 3.17.** *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  $\kappa$  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  $\kappa$  is non-degenerate on  $\mathfrak{g}_0$ , 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  $X = 0$ . If  $\mathfrak{h}$  was an ideal, and  $\mathfrak{g}$  was semisimple we would conclude by Cartan's criterion that it was semisimple. However, this never occurs, because  $\mathfrak{h}$  is abelian so this would imply  $\mathfrak{g}$  was solvable.

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  $\kappa$  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 one dimensional Lie subalgebra of  $\mathfrak{sl}_n(K)$  spanned by  $E_{11} - E_{22}$ , 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 3.18.** *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 a linear Lie algebra which is a subalgebra of  $\mathfrak{gl}_n(K)$ , the diagonal matrices form a canonical abelian Lie subalgebra of semisimple elements, which we would hope to be a Cartan subalgebra. Of course, we cannot expect this always to be the case, because the subalgebra might not contain any diagonal matrices, but there is a nice condition which will show that we can choose the diagonal matrices to be the Cartan subalgebra of all the linear Lie algebras we will be considering. Given a linear Lie algebra  $\mathfrak{g}$ , let  $\mathfrak{h}$  be the diagonal matrices in  $\mathfrak{g}$ , and  $\mathfrak{k} = \mathfrak{g} \cap \text{span}(e_{ij} : i \neq j)$ . Then  $\mathfrak{k}$  is invariant under  $\mathfrak{h}$ , hence  $\mathfrak{h}$  are simultaneously diagonalizable over  $\mathfrak{k}$ , and we find

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

where  $\mathfrak{k}_\alpha$  is the eigenspace with weight  $\alpha$  corresponding to  $\mathfrak{h}$  on  $\mathfrak{k}$ . We will find that there is an easy condition which shows  $\mathfrak{k}_0 = 0$ , so that  $\mathfrak{h}$  is really a Cartan subalgebra for  $\mathfrak{g}$ .

**Lemma 3.19.** *Suppose that, for all  $\alpha \in \Phi$ , there is  $X \in \mathfrak{h}$  such that  $\alpha(X) \neq 0$ . Then  $\mathfrak{h}$  is a Cartan subalgebra for  $\mathfrak{g}$ .*

*Proof.* We need only show that  $\mathfrak{h}$  is a maximal abelian subalgebra of semisimple elements, since we already know it is abelian and semisimple. Suppose that  $X \in \mathfrak{g}$  and  $[X, Y] = 0$  for all  $Y \in \mathfrak{h}$ . We may take the direct sum decomposition, and write  $X = Y + \sum X_\alpha$  for  $X_\alpha \in \mathfrak{k}_\alpha$ , and  $Y \in \mathfrak{h}$ . Then for any  $Z \in \mathfrak{h}$ ,

$$0 = [Z, X] = \sum \alpha(Z) X_\alpha$$

For each  $\alpha$ , there is  $Z$  with  $\alpha(Z) \neq 0$ , hence  $X_\alpha = 0$ , and so  $X \in \mathfrak{h}$ .  $\square$

Essentially, what we have shown is that if  $\mathfrak{g}$  has an abelian subalgebra  $\mathfrak{h}$  for which  $C(\mathfrak{h}) = \mathfrak{h}$ , then  $\mathfrak{h}$  is a Cartan subalgebra. In this language, the lemma is essentially obvious.

There is an easy way to show the classical Lie algebras are semisimple.

**Theorem 3.20.** *Let  $\mathfrak{g}$  be a Lie algebra with Cartan subalgebra  $\mathfrak{h}$ . Let*

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

*Suppose that*

- *For each nonzero  $X$ , there is a weight  $\alpha$  with  $\alpha(X) \neq 0$ .*
- *For each  $\alpha$ ,  $\mathfrak{g}_\alpha$  is one dimensional, spanned by some  $X_\alpha$ .*
- *If  $\alpha \in \Phi$ , then  $-\alpha \in \Phi$ , and  $[[X_\alpha, X_{-\alpha}], X_\alpha] \neq 0$  (Note that this is equivalent to  $\alpha[X_\alpha, X_{-\alpha}]$  being nonzero).*

*Then  $\mathfrak{g}$  is semisimple.*

*Proof.* It is enough to show that  $\mathfrak{g}$  has no nonzero abelian ideals. Let  $\mathfrak{a}$  be some abelian ideal. Since  $\mathfrak{h}$  diagonalizes  $\mathfrak{g}$ , and  $[\mathfrak{h}, \mathfrak{a}] \subset \mathfrak{a}$ , so  $\mathfrak{h}$  acts diagonally on  $\mathfrak{a}$ , and we can write

$$\mathfrak{a} = (\mathfrak{a} \cap \mathfrak{h}) \oplus \bigoplus_{\alpha \in \Phi_0} \mathfrak{g}_\alpha$$

for some subset  $\Phi_0$  of  $\Phi$ . If  $\alpha \in \Phi_0$ , then  $\mathfrak{a}$  contains an element of the form  $Y = [X_\alpha, X_{-\alpha}] \in \mathfrak{a}$ , and we know  $[[X_\alpha, X_{-\alpha}], X_\alpha] = 0$ , since  $\mathfrak{a}$  is abelian. However, we assumed this was not the case, hence  $\mathfrak{a} \subset \mathfrak{h}$ . If  $\mathfrak{a}$  contains any nonzero element  $Y \in \mathfrak{h}$ , then we know there is  $\alpha \in \Phi$  with  $\alpha(Y) \neq 0$ , and then if  $X \in \mathfrak{g}_\alpha$ , then  $[Y, X] = \alpha(Y)X \in \mathfrak{a} \cap \mathfrak{g}_\alpha$ , which is impossible. We conclude  $\mathfrak{a} = 0$ .  $\square$

**Example.** *The diagonal matrices  $\mathfrak{h}$  in  $\mathfrak{sl}_n(K)$  are spanned by the elements  $H_k = E_{kk} - E_{(k+1)(k+1)}$ , and are diagonalized by the off-diagonal elements, with weights  $\alpha_{ij}(X) = X_{ii} - X_{jj}$ , for  $i \neq j$ . Since  $\alpha_{ij} \neq 0$  for any  $i, j$ , we find that the diagonal matrices form a Cartan subalgebra of  $\mathfrak{sl}_n(K)$ . The root space corresponding to  $\alpha_{ij}$  is spanned by  $E_{ij}$ . We know that  $\mathfrak{sl}_n(K)$  is semisimple already, but we can use the theorem above to conclude this as well because each weight space is one dimensional, and  $\alpha_{ij}[E_{ij}, E_{ji}] = \alpha_{ij}(E_{ii} - E_{jj}) = 2$ .*

**Example.** Recall that  $\mathfrak{sp}_n(K)$ , for  $n = 2m$ , consists of the matrices  $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$  with  $A = -D^t$ ,  $B^t = B$ , and  $C^t = C$ . We will let  $\mathfrak{h}$  consist of the diagonal matrices, which has basis  $H_k = E_{kk} - E_{(k+m)(k+m)}$ . There are three classes of matrices which diagonalize  $\mathfrak{h}$ , which can be partitioned into the elements which vanish outside of  $A$  and  $D$ , or vanish outside of  $B$ , or vanish outside of  $C$ .

- $E_{ij} - E_{(j+m)(i+m)}$  for  $i \neq j$  has weight  $\alpha_{ij}(X) = X_{ii} - X_{jj}$ .
- $E_{i(i+m)}$  has weight  $\lambda_i(X) = 2X_{ii}$ .
- $E_{i(j+m)} + E_{j(i+m)}$  has weight  $\beta_{ij}(X) = X_{ii} + X_{jj}$ .
- $E_{(i+m)i}$  has weight  $-\lambda_i$ .
- $E_{(i+m)j} + E_{(j+m)i}$  has weight  $-\beta_{ij}$ .

Thus  $\mathfrak{h}$  is a Cartan subalgebra for  $\mathfrak{sp}_n(K)$ . We see each weight space is one dimensional, and

- $\alpha_{ij}[E_{ij} - E_{(j+m)(i+m)}, E_{ji} - E_{(i+m)(j+m)}] = 2$ .
- $\lambda_i[E_{i(i+m)}, E_{(i+m)i}] = E_{ii} - E_{(i+m)(i+m)} = 2$ .
- $\beta_{ij}[E_{i(j+m)} + E_{j(i+m)}, E_{(i+m)j} + E_{(j+m)i}] = 2$ .

We conclude that  $\mathfrak{sp}_n(K)$  is semisimple.

**Example.** It is harder to find a Cartan subalgebra for  $\mathfrak{o}_n(K)$ , because elements of  $\mathfrak{o}_n(K)$  vanish on the diagonal. We will find a linear Lie algebra isomorphic to  $\mathfrak{o}_n(K)$  which does not vanish on the diagonal, so we can consider a Cartan subalgebra consisting of diagonal elements. We shall find that the structure of  $\mathfrak{o}_n(K)$  is very different depending on whether  $n$  is odd or even.

For even  $n = 2m$ , define the matrix  $J$  by

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

and let  $\mathfrak{g}$  consist of the matrices  $X$  in  $\mathfrak{gl}_n(K)$  with  $X^t J + JX = 0$ . If we write  $X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ , then

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

so that  $X \in \mathfrak{g}$  if and only if  $C^t = -C$ ,  $B^t = -B$ , and  $A^t = -D$ . Then  $\mathfrak{g}$  is isomorphic to  $\mathfrak{o}_n(K)$  (TODO: Prove this). We can then consider a Cartan subalgebra  $\mathfrak{h}$  of  $\mathfrak{g}$  consisting of diagonal matrices, and if  $X$  is diagonal, then we can calculate weights.

- $E_{ij} - E_{(j+m)(i+m)}$  has weight  $\alpha_{ij}(X) = X_{ii} - X_{jj}$ , like for  $\mathfrak{sp}_n(K)$ .
- $E_{i(j+m)} - E_{j(i+m)}$  has weight  $\beta_{ij}(X) = X_{ii} + X_{jj}$ .
- $E_{(i+m)j} - E_{(j+m)i}$  has weight  $-\beta_{ij}$ .

Thus  $\mathfrak{h}$  really is a Cartan subalgebra. We find

- $\alpha_{ij}[E_{ij} - E_{(j+m)(i+m)}, E_{ji} - E_{(i+m)(j+m)}] = 2$ .
- $\beta_{ij}[E_{i(j+m)} - E_{j(i+m)}, E_{(i+m)j} - E_{(j+m)i}] = -2$ .

Thus we conclude  $\mathfrak{g}$  is semisimple, and by isomorphism, that  $\mathfrak{o}_n(K)$  is also semisimple. For odd  $n$ , we have to use a slightly different matrix  $J$  to define an algebra with nonvanishing diagonals. If  $n = 2m + 1$ , we let

$$J = \begin{pmatrix} 0_m & I_m & 0 \\ I_m & 0_m & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Then the matrices  $X$  satisfying  $X^t J + J X = 0$  are the matrices  $X = \begin{pmatrix} A & B & C \\ D & E & F \\ G & K & L \end{pmatrix}$ , with

$$X^t J = \begin{pmatrix} D^t & A^t & G^t \\ E^t & B^t & K^t \\ F^t & C^t & L^t \end{pmatrix} \quad J X = \begin{pmatrix} D & E & F \\ A & B & C \\ G & K & L \end{pmatrix}$$

Hence  $X \in \mathfrak{g}$  if and only if  $D^t = -D$ ,  $B^t = -B$ ,  $L = 0$ ,  $A^t = -E$ ,  $G^t = -F$ , and  $K^t = -C$ . We let  $\mathfrak{h}$  consist of the diagonal elements of this algebra. Using the same elements as for  $\mathfrak{o}_n(K)$  for  $n$  even, we find that we have roots  $\alpha_{ij}$ ,  $\beta_{ij}$ , and  $-\beta_{ij}$ . It only remains to find eigenvectors on  $C, K, G$ , and  $F$ . We find

- $E_{in} - E_{n(i+m)}$  is an eigenvector with weight  $\gamma_i(X) = X_{ii}$ .
- $E_{(i+m)n} - E_{ni}$  is an eigenvector with weight  $-\gamma_i$ .

And  $\gamma_i[E_{in} - E_{n(i+m)}, E_{(i+m)n} - E_{ni}] = -1$ . Thus  $\mathfrak{o}_n(K)$  is semisimple in any dimension.

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 3.21.** *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  $\text{adj}_Y$ , and consider an alternate extension of  $\{Z'_1, \dots, Z'_n\}$  which diagonalize  $\text{adj}_X$  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  $\lambda_i$

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

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  $\lambda_i$ ,  $\gamma_j$ ,  $\sigma_k$  and  $\nu_k$  are non-zero, and if we choose a non-zero  $\mu$  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 3.22.** *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  $\text{adj}_Y$  is equal to  $\text{adj}_N$  on  $C(X)$ , because  $S \in \mathfrak{h} = C(X)$ . This implies  $\text{adj}_Y$  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  $\text{adj}_Y$  is represented by an upper triangular matrix. Since  $\text{adj}_N$  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  $\kappa$  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  $\Phi$  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.

### 3.4 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 3.23.** *If  $\alpha$  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  $\kappa$  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  $f_\alpha$  an appropriate scalar multiple of  $Y$  such that  $\alpha[e_\alpha, f_\alpha] = 2$ . We then 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 3.24.** *The choices of  $e_\alpha$ ,  $f_\alpha$ , and  $h_\alpha$  are unique up inversely scaling  $e_\alpha$  and  $f_\alpha$  as  $\lambda e_\alpha$ ,  $\lambda^{-1} f_\alpha$ ,  $h_\alpha$  for some non-zero  $\lambda \in K$ . To obtain a basis for  $\mathfrak{sl}(-\alpha)$  from  $\mathfrak{sl}(\alpha)$ , swap  $e_\alpha$  with  $f_\alpha$ , and replacing  $h_\alpha$  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  $\mathfrak{g}_\alpha$  is one dimensional), that  $e'_\alpha = e_\alpha$ . Let  $f'_\alpha = \lambda f_\alpha$ . It then follows that  $h'_\alpha$  is a scalar multiple of  $h_\alpha$ ,  $h'_\alpha = \gamma h_\alpha$ . 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  $\alpha$  there is  $t_\alpha \in \mathfrak{h}$  with  $\kappa(t_\alpha, X) = \alpha(X)$  for all  $X \in \mathfrak{h}$ .

**Lemma 3.25.** *If  $\alpha$  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  $\kappa$  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 3.26.** *If  $V$  is a  $\mathfrak{sl}(\alpha)$  submodule of  $\mathfrak{g}$ , then the eigenvalues of  $h_\alpha$  acting on  $V$  are integers.*

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

**Example.** On  $\mathfrak{sl}_n(K)$ , the weights take the form  $\alpha_{ij}(X) = X_{ii} - X_{jj}$ , and we may take

$$\begin{aligned} e_{\alpha_{ij}} &= E_{ij} & f_{\alpha_{ij}} &= E_{ji} \\ h_{\alpha_{ij}} &= E_{ii} - E_{jj} \end{aligned}$$

Considering  $\mathfrak{sl}_3(K)$  as a  $\mathfrak{sl}_2(\alpha)$  module, where  $\alpha(X) = X_{11} - X_{22}$ , we find this is exactly the  $\mathfrak{sl}_2(K)$  module 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  $\mathfrak{sl}_2(\alpha)$  module.

**Example.** On  $\mathfrak{sp}_{2m}(K)$ , the  $\mathfrak{sl}_2(\alpha)$  are defined by

- For the weights  $\alpha_{ij}(X) = X_{ii} - X_{jj}$ ,

$$e_{\alpha_{ij}} = E_{ij} - E_{(j+m)(i+m)} \quad f_{\alpha_{ij}} = E_{ji} - E_{(i+m)(j+m)}$$

$$h_{\alpha_{ij}} = E_{ii} - E_{jj} + E_{(j+m)(j+m)} - E_{(i+m)(i+m)}$$

- For the weights  $\lambda_i(X) = 2X_{ii}$ ,

$$e_{\lambda_i} = E_{i(i+m)} \quad f_{\lambda_i} = E_{(i+m)i}$$

$$h_{\lambda_i} = E_{ii} - E_{(i+m)(i+m)}$$

- For the weights  $\beta_{ij}(X) = X_{ii} + X_{jj}$ ,

$$e_{\beta_{ij}} = E_{i(j+m)} + E_{j(i+m)} \quad f_{\beta_{ij}} = E_{(i+m)j} + E_{(j+m)i}$$

$$h_{\beta_{ij}} = E_{ii} + E_{jj} - E_{(j+m)(j+m)} - E_{(i+m)(i+m)}$$

**Example.** On the algebra corresponding to  $\mathfrak{o}_{2m}(K)$ , we find

- For the weights  $\alpha_{ij}(X) = X_{ii} - X_{jj}$ ,

$$e_{\alpha_{ij}} = E_{ij} - E_{(j+m)(i+m)} \quad f_{\alpha_{ij}} = E_{ji} - E_{(i+m)(j+m)}$$

$$h_{\alpha_{ij}} = E_{ii} - E_{jj} + E_{(j+m)(j+m)} - E_{(i+m)(i+m)}$$

- For the weights  $\beta_{ij}(X) = X_{ii} + X_{jj}$ ,

$$e_{\beta_{ij}} = E_{i(j+m)} - E_{j(i+m)} \quad f_{\beta_{ij}} = E_{(j+m)i} - E_{(i+m)j}$$

$$h_{\beta_{ij}} = E_{ii} + E_{jj} - E_{(j+m)(j+m)} - E_{(i+m)(i+m)}$$

When we move up to  $\mathfrak{o}_{2m+1}(K)$ , in addition to the weights for  $\mathfrak{o}_{2m}(K)$ , we also have additional weights.

- For the weights  $\gamma_i(X) = X_{ii}$ ,

$$e_{\gamma_i} = E_{in} - E_{n(i+m)} \quad f_{\gamma_i} = 2(E_{ni} - E_{(i+m)n})$$

$$h_{\gamma_i} = 2(E_{ii} - E_{(i+m)(i+m)})$$

**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  $\alpha$ -root string through  $\beta$ , 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 3.27.** The root space of any non-zero  $\alpha$  on a semisimple Lie algebra is one-dimensional, and the only multiples of  $\alpha$  which are roots are  $\pm\alpha$ .

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

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

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  $h_\alpha$  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  $h_\alpha$  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  $\pm 2$  with respect to  $h_\alpha$ , which would imply  $W$  contains a submodule isomorphic to some  $V_n$  with  $n$  even. Thus if  $\alpha$  is a root,  $2\alpha$  is never a root. Finally, assume  $W$  has a submodule

isomorphic to  $V_n$  with  $n$  odd. Then  $V$  contains a  $h_\alpha$  eigenvector  $X$  with eigenvalue 1, which must be contained in  $W$ . Thus there is a root  $\beta$  with  $\beta(h_\alpha) = 1$ , but then  $2\beta = \alpha$ , since  $2\beta$  and  $\alpha$  agree at  $h_\alpha$ , 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 3.28.** *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  $h_\alpha$  on some element of  $\mathfrak{g}_\beta$ , 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  $h_\alpha$  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  $V_n$ ,  $h_\alpha$  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  $h_\alpha$  eigenspace of eigenvalue  $\beta(h_\alpha)$ , and if  $[e_\alpha, e_\beta] = 0$ , then  $e_\beta$  is the highest weight vector in the irreducible representation  $V$  isomorphic to  $V_n$ . If  $\alpha + \beta$  is a root, then  $h_\alpha$  has eigenvalue  $(\alpha + \beta)(h_\alpha) = 2 + \beta(h_\alpha)$ , hence  $e_\beta$  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  $h_\alpha$ . We are well on our way to classifying the semisimple Lie algebras!

### 3.5 Cartan Subalgebras as Inner Product Spaces

**Lemma 3.29.** *If  $X \in \mathfrak{h}$  is non-zero, then there is a root  $\alpha$  with  $\alpha(X) \neq 0$ , and therefore the roots span  $\mathfrak{h}^*$ .*

*Proof.* If  $\alpha(X) = 0$  for all roots  $\alpha$ , 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 3.30.** *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  $t_\alpha$  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

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

$\square$

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  $\alpha, \beta$ , 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  $\alpha, \beta$  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 3.31.** *If  $\beta$  is a root, then  $\beta$  is a linear combination of the  $\alpha_i$  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  $\lambda_i$ , and since each  $(\alpha_i, \alpha_j)$  is rational, and  $(\beta, \alpha_j)$  is rational, we conclude the  $\lambda_i$  are rational.  $\square$

Thus the *rational* subspace generated by the  $\alpha_i$  contains all the roots of  $\Phi$ , and doesn't depend on the particular choice of basis roots  $\alpha_i$ . 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 3.32.**  *$(\cdot, \cdot)$  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  $\alpha$ , 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  $\lambda$ , and hence  $x \otimes \lambda = 0$ .  $\square$

# Chapter 4

## Root systems

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

- $0 \notin \Phi$ .
- The only scalar multiples of  $\alpha \in \Phi$  which are in  $\Phi$  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  $\theta$  is the angle between  $\alpha$  and  $\beta$ , then  $2\|\alpha\|\cos(\theta)$  (twice the length of the projection of  $\alpha$  on  $\beta$ ) is an integer multiple of  $\|\beta\|$ .

- If  $s_\alpha$  is the reflection about the perpendicular to  $\alpha$ , which can be defined as  $s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha$  (subtracting twice the projection of  $\beta$  onto  $\alpha$ ), then  $s_\alpha$  permutes  $\Phi$ .

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  $x \in \mathbf{R}$ .

On Lie algebras, the normalized inner product  $\langle \cdot, \cdot \rangle$  is calculated to be  $\langle \alpha, \beta \rangle = \alpha(h_\beta)$ . This is how we'll calculate the inner product on the classical Lie algebras.

**Example.** On  $\mathfrak{sl}_n(K)$ , we have roots  $\alpha_{ij}$ , and we find

$$\langle \alpha_{ij}, \alpha_{kl} \rangle = \alpha_{ij}(h_{\alpha_{kl}}) = \alpha_{ij}(E_{kk} - E_{ll}) = \delta_{ik} + \delta_{jl} - \delta_{jk} - \delta_{il}$$

and if  $(ij)$  is the permutation of  $i$  and  $j$ , we find that

$$s_{\alpha_{ij}}(\alpha_{kl}) = \alpha_{(ij)k (ij)l}$$

So, for instance,  $s_{\alpha_{ij}}(\alpha_{ij}) = \alpha_{ji}$ , and  $s_{\alpha_{ij}}(\alpha_{ik}) = \alpha_{jk}$ .



**Example.** On  $\mathfrak{sp}_n(K)$ , as on  $\mathfrak{sl}_n(K)$ , using the values for  $h_\alpha$ , we find

$$\langle \alpha_{ij}, \alpha_{kl} \rangle = \delta_{ik} - \delta_{il} - \delta_{jk} + \delta_{jl} \quad \langle \lambda_i, \lambda_j \rangle = 2\delta_{ij} \quad \langle \beta_{ij}, \beta_{kl} \rangle = \delta_{ik} + \delta_{jk} + \delta_{il} + \delta_{jl}$$

$$\langle \alpha_{ij}, \beta_{kl} \rangle = \delta_{ik} + \delta_{il} - \delta_{jk} - \delta_{jl} \quad \langle \beta_{kl}, \alpha_{ij} \rangle = \delta_{ik} + \delta_{il} - \delta_{jk} - \delta_{jl}$$

$$\langle \alpha_{ij}, \lambda_r \rangle = \delta_{ir} - \delta_{jr} \quad \langle \lambda_r, \alpha_{ij} \rangle = 2(\delta_{ir} - \delta_{jr})$$

$$\langle \beta_{kl}, \lambda_r \rangle = \delta_{kr} + \delta_{lr} \quad \langle \lambda_r, \beta_{kl} \rangle = 2(\delta_{kr} + \delta_{lr})$$

and so

$$s_{\alpha_{ij}}(\alpha_{kl}) = \alpha_{(ij)k} (ij)l \quad s_{\alpha_{ij}}(\lambda_r) = \lambda_{(ij)r} \quad s_{\alpha_{ij}}(\beta_{kl}) = \beta_{(ij)k} (ij)l$$

$$s_{\lambda_r}(\alpha_{ir}) = \beta_{ir} \quad s_{\lambda_r}(\alpha_{rj}) = -\beta_{rj}$$

$$s_{\lambda_r}(\beta_{rl}) = s_{\lambda_r}(\beta_{lr}) = \alpha_{lr} \quad s_{\lambda_r}(\lambda_r) = -\lambda_r$$

$$s_{\beta_{kl}}(\alpha_{ik}) = s_{\beta_{lk}}(\alpha_{ik}) = \beta_{il} \quad s_{\beta_{kl}}(\alpha_{kj}) = s_{\beta_{lk}}(\alpha_{kj}) = -\beta_{jl}$$

$$s_{\beta_{kl}}(\beta_{ki}) = s_{\beta_{lk}}(\beta_{ik}) = \alpha_{il} \quad s_{\beta_{kl}}(\beta_{kl}) = -\beta_{kl}$$

$$s_{\beta_{kl}}(\lambda_k) = s_{\beta_{lk}}(\lambda_k) = -\lambda_k$$

we only list the points which aren't fixed by the maps. It appears that the  $\lambda$  roots form a distinct class different than the  $\alpha$  and  $\beta$  roots, because there is no way to get one set from the other.

**Example.** On  $\mathfrak{o}_{2n}(K)$ , we have

$$\langle \alpha_{ij}, \alpha_{kl} \rangle = \delta_{ik} + \delta_{il} - \delta_{il} - \delta_{jk} \quad \langle \beta_{ij}, \beta_{kl} \rangle = \delta_{ik} + \delta_{il} + \delta_{jk} + \delta_{jl}$$

$$\langle \alpha_{ij}, \beta_{kl} \rangle = \delta_{ik} + \delta_{il} - \delta_{jk} - \delta_{jl} \quad \langle \beta_{kl}, \alpha_{ij} \rangle = \delta_{ik} + \delta_{il} - \delta_{jk} - \delta_{jl}$$

Thus

$$s_{\alpha_{ij}}(\alpha_{kl}) = \alpha_{(ij)k} (ij)l$$

$$s_{\alpha_{ij}}(\beta_{il}) = s_{\alpha_{ij}}(\beta_{li}) = -\beta_{jl} \quad s_{\alpha_{ij}}(\beta_{jl}) = s_{\alpha_{ij}}(\beta_{lj}) = \beta_{il}$$

$$s_{\beta_{kl}}(\alpha_{kj}) = s_{\beta_{lk}}(\alpha_{kj}) = -\beta_{jl} \quad s_{\beta_{kl}}(\alpha_{jk}) = s_{\beta_{lk}}(\alpha_{jk}) = \beta_{jl}$$

$$s_{\beta_{kl}}(\beta_{kj}) = s_{\beta_{lk}}(\beta_{kj}) = \alpha_{lj} \quad s_{\beta_{kl}}(\beta_{lj}) = s_{\beta_{lk}}(\beta_{lj}) = \alpha_{kj}$$

On  $\mathfrak{o}_{2n+1}$ , we add the extra  $\gamma$  roots, and we find

$$\langle \alpha_{ij}, \gamma_r \rangle = 2(\delta_{ir} - \delta_{jr}) \quad \langle \gamma_r, \alpha_{ij} \rangle = \delta_{ir} - \delta_{jr}$$

$$\begin{aligned}\langle \beta_{kl}, \gamma_r \rangle &= 2(\delta_{kr} + \delta_{lr}) & \langle \gamma_r, \beta_{kl} \rangle &= \delta_{kr} + \delta_{lr} \\ \langle \gamma_i, \gamma_r \rangle &= 2\delta_{ir}\end{aligned}$$

Hence

$$\begin{aligned}s_{\gamma_r}(\alpha_{rj}) &= s_{\gamma_r}(\alpha_{jr}) = \gamma_i & s_{\gamma_r}(\beta_{rj}) &= s_{\gamma_r}(\beta_{jr}) = -\gamma_j \\ s_{\gamma_r}(\gamma_r) &= -\gamma_r \\ s_{\alpha_{ij}}(\gamma_i) &= \gamma_i - 2\alpha_{ij} = 2\gamma_j - \gamma_i\end{aligned}$$

**Lemma 4.1.** *If  $R$  is a root system, for two roots  $\alpha$  and  $\beta$ , 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  $\alpha$  and  $\beta$  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.

$\langle \alpha, \beta \rangle$	$\langle \beta, \alpha \rangle$	$\theta$	$\frac{(\beta, \beta)}{(\alpha, \alpha)}$
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 4.2.** *If  $\alpha$  and  $\beta$  are roots, and the angle between  $\alpha$  and  $\beta$  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  $\langle \alpha, \beta \rangle = 1$ .  $\square$

The angle between  $\alpha$  and  $\beta$  is strictly obtuse if and only if  $\cos(\theta) < 0$ , hence if  $(\alpha, \beta) < 0$ . Then  $\alpha + \beta$  is a root, and provided that  $(\alpha + \beta, \beta) = (\alpha, \beta) + (\beta, \beta) < 0$ , then  $\alpha + 2\beta$  is a root, and so on and so forth.

**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  $\alpha$  have minimum length. Consider some other root  $\beta$  linearly independent from  $\alpha$  ( $\beta$  must exist, since the root system spans  $\mathbf{R}^2$ ). We may assume that  $\beta$  makes an obtuse angle with  $\alpha$  (otherwise, consider  $-\beta$ ), and that this angle is as large as possible.

- If  $\beta$  lies at an angle  $\pi/2$  for  $\alpha$ , 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  $\alpha$ , in which case  $x = \pm\beta$ . This is the root system  $A_1 \times A_1$ .

- If  $\beta$  lies at an angle  $2\pi/3$  from  $\alpha$ , 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  $\beta$ , hence the roots are exactly these vectors. We call this system  $A_2$ .

- If  $\beta$  lies at an angle  $3\pi/4$  from  $\alpha$ , then  $\alpha + \beta$  and  $-(\alpha + \beta)$  are roots.  $2\alpha + \beta$  is then obtained from reflecting  $\beta$  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  $\beta$  is the largest angle. We can't add any other roots, and this classifies the root system, which we denote by  $B_2$ .

- If  $\beta$  lies at an angle  $5\pi/6$  from  $\alpha$ , 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  $R_1$  are perpendicular to vectors of  $R_0$ . 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 4.3.** *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 complementary subspaces  $E_i$  which sum up to  $E$ , because each element of the span of  $R_j$  is perpendicular to every element in the span of some  $R_i$ , for  $j \neq i$ , and hence cannot be an element of this space. Clearly each  $R_i$  contains the negation of each root, since  $(\lambda, -\lambda) = -(\lambda, \lambda) \neq 0$  for all  $\lambda$ . If  $\alpha, \beta \in R_i$ , and  $\beta \neq \pm\alpha$ , then  $s_\alpha(\beta)$  is in the plane spanned by  $\alpha$  and  $\beta$ , and therefore cannot be perpendicular to both  $\alpha$  and  $\beta$  at the same time, hence  $s_\alpha(\beta) \in R_i$ .  $\square$

We may surely consider linearly independent elements of a root system, but we also have something stronger, which guarantees that these elements can be added together in a useful way. Define a **base** for the root system  $R$  to be  $B \subset R$  which forms a vector space basis for the extension field, such that for every  $\lambda \in R$ , we can write  $\lambda = \sum_{\alpha \in B} c_\alpha \alpha$  with integral coefficients  $c_\alpha \in \mathbb{Z}$  and such that all non-zero  $c_\alpha$  are either all negative or all positive. We say a root is **positive** with respect to a base  $B$  if the root can be written as the sum of elements of the basis with positive coefficients, and **negative** if the root is the sum of elements with negative coefficients.

It follows that the angle between any two elements of a basis is obtuse, because if the angle between two roots  $\alpha$  and  $\beta$  is less than  $\pi/2$ , then either

$\alpha - \beta$  or  $\beta - \alpha$  is a root, contradicting the fact that roots are the sum of all positive coefficients or all negative coefficients.

**Example.** The root system  $e_i - e_j$  on  $\mathbf{R}^n$  has a basis consisting of the elements  $e_{i+1} - e_i$ . A root  $e_i - e_j$  is positive if and only if  $i > j$ , and negative if  $i < j$ .

There is an easy way to construct a base for a root system, which is obtained by fixing some hyperplane through the origin upon which none of the roots lie, and then defining one side of the hyperplane to consist of positive elements.

**Theorem 4.4.** Every root system has a basis.

*Proof.* Let  $R$  be a root system. We may assume  $R$  lies in  $n$  dimensional space, for  $n > 1$ , since the 1 dimensional case is trivial. Then we may pick  $x \in E$  such that  $x \notin \alpha^\perp$  for any root  $\alpha$ . Let  $R^+$  consist of the roots  $\alpha$  with  $(x, \alpha) > 0$ , and  $R^-$  the roots with  $(x, \alpha) < 0$ . Then  $R^+$  is mapped onto  $R^-$  by negation, and we will find a basis in  $R^+$  in which  $R^+$  is exactly the positive elements. Let  $B$  be the set of roots  $\alpha \in R^+$  such that  $\alpha \neq \beta + \gamma$  for any  $\beta, \gamma \in R^+$ . We claim  $B$  is a basis for  $R$ . It suffices to show that any  $\alpha \in R^+$  can be written as the positive sum of elements of  $B$ . If there is a root which cannot be written in this way, consider such a root  $\beta$  which minimizes the inner product  $(x, \beta)$ . As  $\beta \notin B$ , there are  $\beta_0, \beta_1 \in R^+$  with  $\beta = \beta_0 + \beta_1$ . Then  $(x, \beta) = (x, \beta_0) + (x, \beta_1)$ , and we conclude that by minimality, both  $\beta_0$  and  $\beta_1$  can be written as the sum of positive elements. But this implies that  $\beta$  can be written in this way.

All that remains is to show that  $B$  is linearly independent. First, note that if  $\alpha, \beta \in B$ , then the angle between  $\alpha$  and  $\beta$  is greater than or equal to  $\pi/2$ , for otherwise if  $(\beta, \beta) \geq (\alpha, \alpha)$ , then  $\alpha - \beta$  is a root, and either  $\alpha - \beta$  or  $\beta - \alpha$  is in  $R^+$ , in which case either  $\alpha$  or  $\beta$  can be written as the sum of two elements of  $R^+$ . If  $\sum_{\alpha \in B} c_\alpha \alpha = 0$ , where  $c_\alpha \in \mathbf{R}$ , then we find  $\sum_{c_\alpha > 0} c_\alpha \alpha = \sum_{c_\alpha < 0} (-c_\alpha) \alpha$ . Denote this element by  $y$ . Then, since  $(\alpha, \beta) \leq 0$  ( $\alpha$  and  $\beta$  lie at an obtuse angle),

$$(y, y) = \sum_{\substack{c_\alpha > 0 \\ c_\beta < 0}} c_\alpha (-c_\beta) (\alpha, \beta) \leq 0$$

Thus  $y = 0$ . Therefore  $0 = (x, y) = \sum_{c_\alpha > 0} c_\alpha (x, \alpha)$ , and since each  $(x, \alpha) > 0$ , we conclude  $c_\alpha = 0$  for all  $\alpha$ .  $\square$

Given a base  $B$ , we let  $R^+$  denote the set of all positive roots, and  $R^-$  the set of all negative roots. The elements of  $B$  are known as the **simple roots**. The reflections  $s_\alpha$  for  $\alpha \in B$  are known as the **simple reflections**.

**Example.** We can give a basis of  $\mathfrak{sl}_n(K)$  of the form  $\alpha_{12}, \alpha_{23}, \dots, \alpha_{(n-1)n}$ , since if  $i < j$ ,

$$\alpha_{ij} = \sum_{k=i}^j \alpha_{k(k+1)}$$

and if  $i > j$ , then

$$\alpha_{ij} = \sum_{k=j}^i (-\alpha_{k(k+1)})$$

The positive roots are therefore the  $\alpha_{ij}$  for  $i < j$ .

**Example.** Over  $\mathfrak{sp}_{2n}(K)$ , the canonical choice of a basis is  $\alpha_{12}, \dots, \alpha_{(n-1)n}$ , and  $\lambda_n$ , since we may surely describe all  $\alpha_{ij}$  in this manner, where  $\alpha_{ij}$  has positive coefficients when  $i < j$ . For  $i < n$ ,  $\beta_{in} = \alpha_{in} + \lambda_n$  is positive, and then  $\beta_{ij} = \alpha_{in} + \beta_{jn}$ , and  $\lambda_i = 2\alpha_{in} + 2\lambda_n$ . The positive roots are then  $\alpha_{ij}$  for  $i < j$ , and  $\beta_{ij}$  and  $\lambda_i$ .

**Example.** Over  $\mathfrak{so}_{2n}(K)$ , the canonical choice of a basis is  $\alpha_{12}, \dots, \alpha_{(n-1)n}$ , and  $\beta_{(n-1)n}$ , for then  $\beta_{in} = \alpha_{i(n-1)} + \beta_{(n-1)n}$ , and so  $\beta_{ij} = \alpha_{in} + \beta_{jn}$ . The positive roots are  $\alpha_{ij}$  for  $i < j$ , and  $\beta_{ij}$ . Over  $\mathfrak{so}_{2n+1}$ , we choose a basis of  $\alpha_{12}, \dots, \alpha_{(n-1)n}$  and  $\lambda_n$ , for then we have all  $\lambda_i$ , and therefore also all  $\beta_{ij} = \alpha_{ij} + 2\lambda_i$ . The positive roots are  $\alpha_{ij}$  for  $i < j$ , and  $\lambda_i$  and  $\beta_{ij}$ .

We note that the choice of base  $B$  is not canonical, and the roots depend on the choice of basis. Indeed, if  $B$  is any basis, and  $\lambda$  is a root, then  $s_\lambda(B)$  is also a basis. Regardless, from now on we will assume a base is fixed over any particular root system.

## 4.1 The Weyl Group

The **Weyl Group** of a root system is the set of all transformations of  $E$  obtained by the reflections  $s_\alpha$ , for any root  $\alpha$ , denoted  $W$ , or  $W(R)$  if the root system in question is undetermined.

**Theorem 4.5.** *The Weyl group of any root system is finite.*

*Proof.* Each reflection permutes the roots of the system. Since there are only finitely many roots, there are only finitely many permutations of the roots. The action of the Weyl group on  $E$  is determined by the action of the Weyl group on the roots, hence there can only be finitely many transformations in the Weyl group.  $\square$

Thus we can faithfully represent the Weyl group over the permutations of the roots.

**Example.** Over  $\mathfrak{sl}_n(K)$ , the Weyl group contains the entire symmetric group. For each  $\alpha_{ij}$ ,

It turns out that a root system is uniquely determined by any of its bases. To prove this, we will use the Weyl group. In particular, we will show that if  $\alpha$  is any root, then  $\alpha = f(\beta)$  for some simple root  $\beta$ , and some  $f$  in the Weyl group. What's more, we will find that the Weyl group is generated by the reflections on the simple roots, so that the simple roots contain the information about the entire root system. For now, we denote the subgroup of the Weyl group generated by the reflections on the simple roots by  $W_0$ .

**Lemma 4.6.** *If  $\alpha$  is simple,  $s_\alpha$  permutes all positive roots but  $\alpha$  itself.*

*Proof.* Suppose  $\beta$  is a positive root other than  $\alpha$ . Then  $\beta = \sum c_\kappa \kappa$ , where  $\kappa$  are simple roots and  $c_\kappa \geq 0$ . We know  $s_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha$  is a root, and provided if  $\beta \neq \alpha$ , there is some  $\kappa \neq 0$  with  $c_\kappa > 0$ , in which case the coefficient corresponding to  $s_\alpha(\beta)$  is  $c_\kappa$ , which is positive, hence  $s_\alpha(\beta)$  is positive.  $\square$

**Theorem 4.7.** *If  $\beta$  is any root, there is a Weyl group permutation  $f \in W_0$  and a simple root  $\alpha$  such that  $\beta = f(\alpha)$ .*

*Proof.* We will prove this first for any positive root. Given some positive root  $\beta = \sum c_\alpha \alpha$ , we will proceed by induction on  $\sum c_\alpha$ . If  $\sum c_\alpha = 1$ , then  $\beta$  is a simple root, and the theorem is trivial. For the induction, first note that if  $(\beta, \gamma) \leq 0$  holds for all simple roots  $\gamma$ , then  $(\beta, \beta) = \sum c_\gamma (\beta, \gamma) \leq 0$ , so  $\beta = 0$ , which is impossible. Thus there is  $\lambda$  with  $(\beta, \lambda) > 0$ , and so  $s_\lambda(\beta)$  fixes the coefficients of all  $c_\alpha$  with all  $\alpha \neq \lambda$ , and decreases the coefficient  $c_\lambda$ , hence by induction  $s_\lambda(\beta) = f(\alpha)$  for some simple root  $\alpha$ , and then  $\beta = (s_\lambda \circ f)(\alpha)$ .

Now suppose that  $\beta$  is a negative root. Then  $-\beta$  is positive, and hence  $-\beta = f(\alpha)$  for some simple root  $\alpha$ , and then  $f(-\alpha) = \beta$ . The reflection  $s_\alpha$  maps  $\alpha$  to  $-\alpha$ , hence  $(f \circ s_\alpha)(\alpha) = \beta$ .  $\square$

**Corollary 4.8.** *The Weyl group generated by simple reflections is equal to the entire Weyl group.*

*Proof.* First, we show that for any  $g \in W$ ,  $g \circ s_\alpha \circ g^{-1} = s_{g(\alpha)}$ . If  $g = s_\beta$  for some  $\beta$ , then

$$\begin{aligned} (s_\beta \circ s_\alpha \circ s_\beta)(x) &= (s_\beta \circ s_\alpha)(x - \langle x, \beta \rangle \beta) \\ &= s_\beta(x - \langle x, \beta \rangle \beta - \langle x, \alpha \rangle \alpha + \langle x, \beta \rangle \langle \beta, \alpha \rangle \alpha) \\ &= x + \langle x, \beta \rangle \beta - \langle x, \alpha \rangle \alpha + \langle x, \beta \rangle \langle \beta, \alpha \rangle \alpha - \langle x, \beta \rangle \beta \\ &\quad + \langle x, \alpha \rangle \langle \alpha, \beta \rangle \beta - \langle x, \beta \rangle \langle \beta, \alpha \rangle \langle \alpha, \beta \rangle \beta \end{aligned}$$

TODO: FINISH THIS PROOF. Since the Weyl group is generated by all reflections, it suffices to show that each  $s_\alpha$  is a product of simple reflections, even if  $\alpha$  is not simple. But there is a simple root  $\beta$  such that  $\alpha = f(\beta)$ , and then  $s_\alpha = f \circ s_\beta \circ f^{-1}$ .  $\square$

With respect to the Weyl group, all bases of a root system are the same.

**Theorem 4.9.** *If  $B$  and  $B'$  are any two bases, then there is a transformation  $f \in W$  with  $f(B) = B'$ .*

Now given a base  $B$ , fix an ordering  $B = \{\alpha_1, \dots, \alpha_n\}$ . The **Cartan matrix** with respect to  $B$  is the  $n \times n$  matrix  $C$  with  $C_{ij} = \langle \alpha_i, \alpha_j \rangle$ . For any root  $\beta$ ,  $\langle s_\beta(\alpha_i), s_\beta(\alpha_j) \rangle = \langle \alpha_i, \alpha_j \rangle$ , so the Cartan matrix does not depend on the base chosen, except for the ordering we pick for the base.

**Example.** Consider the root system  $\mathbf{R}^n$  with basis  $e_{i+1} - e_i$ . Since

$$\langle e_{i+1} - e_i, e_{j+1} - e_j \rangle = (e_{i+1} - e_i, e_{j+1} - e_j) = 2\delta_i^j - \delta_i^{j+1} - \delta_{i+1}^j$$

Thus the Cartan matrix is

$$\begin{pmatrix} 2 & -1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & 0 & \dots & -1 & 2 \end{pmatrix}$$



Since each  $\langle \alpha_i, \alpha_j \rangle$  is an integer between -3 and 3, there is a discrete way to represent the Cartan matrix graphically. Given a root system, we define a graph  $\Delta$ , whose vertices are simple roots, and such that we draw  $\langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle$  edges between each pair of simple roots  $\alpha_i$  and  $\alpha_j$ . If one of the roots is longer than the other roots, we draw an arrow from the longer root to the shorter root. The resulting diagram is known as the **Dynkin diagram**, and up to graph isomorphism the diagram is unique. The graph without the arrows is known as the **Coexeter graph**.

An **isomorphism** between two root systems is a vector space isomorphism which preserves the normalized inner product  $\langle \cdot, \cdot \rangle$  on the roots. Thus the isomorphism preserves angles, but not distances between vectors. It is clear that isomorphic root systems have the same Dynkin diagram. It turns out that the converse is also true.

**Theorem 4.10.** *A graph isomorphism between two Dynkin diagrams induces an isomorphism between two root systems.*

*Proof.* Let  $R$  and  $T$  be two basis with isomorphic Dynkin diagrams. We may choose basis  $\{\alpha_1, \dots, \alpha_n\}$  for  $R$  and a basis  $\{\beta_1, \dots, \beta_n\}$  for  $T$  with

$$\langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle = \langle \beta_i, \beta_j \rangle \langle \beta_j, \beta_i \rangle$$

And if  $\alpha_i$  is larger than  $\alpha_j$ , then  $\beta_i$  is larger than  $\beta_j$ , and vice versa. Consider the map  $f : \alpha_i \mapsto \beta_i$ . It is linear, and preserves  $\langle \cdot, \cdot \rangle$  on the simple roots. We have

$$f(s_{\alpha_i}(\lambda)) = f(\lambda - \langle \lambda, \alpha_i \rangle \alpha_i) = f(\lambda) - \langle f(\lambda), \beta_i \rangle \beta_i = s_{\beta_i}(f(\lambda))$$

It then follows that  $f \circ s_{\alpha_i} = s_{\beta_i} \circ f$ . Since the  $s_{\alpha_i}$  generate the Weyl group of  $R$ , for any  $v$ ,  $f(W_R v) \subset W_T f(v)$ , hence in particular if  $\alpha_j$  is simple, that

$$f(R) = \bigcup_{j=1}^n f(W_R \alpha_j) \subset \bigcup_{j=1}^n W_T \beta_j \subset T$$

The same argument shows  $f^{-1}(T) = R$ , hence  $f$  is an isomorphism.  $\square$

## 4.2 Classification of Root Systems

Just like with Lie algebras, it is of interest to classify root systems up to isomorphism. We shall find this nicely characterizes the set of all simple

Lie algebras. We can always break down root systems into irreducible root systems, so we might as well focus on this family of root systems, which is exactly the family of root systems whose Dynkin diagrams are connected.

## Chapter 5

# The Classification of Semisimple Lie Algebras

We can now connect our newfound understanding of root systems with our deep theory of Lie algebras to obtain a very satisfying classification result for the semisimple Lie algebras.

**Theorem 5.1.** *If  $\mathfrak{g}$  is a semisimple Lie algebra whose root system is irreducible, then  $\mathfrak{g}$  is simple.*

*Proof.* Using the root system decomposition, write

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

If  $\mathfrak{a}$  is an ideal of  $\mathfrak{g}$ , then we may write

$$\mathfrak{a} = \mathfrak{h}_0 \oplus \bigoplus_{\alpha \in \Phi_0} \mathfrak{g}_{\alpha}$$

where  $\Phi_0$  is some subset of  $\Phi$ , and  $\mathfrak{h}_0$  is a subalgebra of  $\mathfrak{h}$ . By the decomposition results about the Killing form of semisimple Lie algebras, we may write

$$\mathfrak{a}^{\perp} = \mathfrak{h}_1 \oplus \bigoplus_{\alpha \in \Phi_1} \mathfrak{g}_{\alpha}$$

where  $\Phi_1 = \Phi_0^c$ , and  $\mathfrak{h} = \mathfrak{h}_0 \oplus \mathfrak{h}_1$ . Then  $\Phi_0$  and  $\Phi_1$  are both root systems, and if  $\alpha \in \Phi_0$ , and  $\beta \in \Phi_1$ , then  $t_{\alpha} \in \mathfrak{h}_0 \subset \mathfrak{a}$ ,  $t_{\beta} \in \mathfrak{h}_1 \subset \mathfrak{a}^{\perp}$ , and so  $(\alpha, \beta) = \kappa(t_{\alpha}, t_{\beta}) = 0$ . Thus we have broken  $\Phi$  into two perpendicular root systems,

so either  $\Phi_0 = \emptyset$  or  $\Phi_1 = \emptyset$ . If  $\Phi_1$  is empty, then all the root spaces  $\mathfrak{g}_\alpha$  are contained in  $\mathfrak{a}$ , but since  $\mathfrak{g}$  is generated by its root spaces we conclude  $\mathfrak{a} = \mathfrak{g}$ . If  $\Phi_0$  is emptyset, then  $\mathfrak{a} \subset \mathfrak{h}$ , but then  $\mathfrak{a}$  is a solvable ideal in a semisimple Lie algebra, so  $\mathfrak{a} = 0$ .  $\square$

**Example.**  $s$

# Bibliography

- [1] William Fulton, Joe Harris *Representation Theory: A First Course*