The Representation Theory of Lie Algebras

Jacob Denson

February 12, 2017

Table Of Contents

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 Lie algebras emerge in geometry and some parts of number theory (normally where $K = \mathbf{R}$ or $K = \mathbf{C}$), where a Lie group G gives rise to a natural Lie algebra structure on the space $\mathfrak{g}$ of tangent vectors at the origin (There is a natural theory of 'Lie groups' for fields other than the real and complex numbers, though we won't discuss them here). We obtain the Lie algebra by differentiating the operation on the group, so we can think of a Lie algebra as an infinitisimal linear approximation to the action of the group. Indeed, most algebraic concepts about Lie groups have 'infinitisimal formulations' on Lie algebras. Because of the correspondence between G and $\mathfrak{g}$, we normally denote a Lie algebra corresponding to a Lie group by the 'frakturized' name of the group.

Example. *If we differentiate the Lie group $GL_n(K)$ of invertible matrices with coefficients in K, we obtain our first example of a Lie algebra. The tangent space at the origin can be identified with the space of all matrices, since $GL_n(K)$ is an open submanifold of $M_n(K)$, and with some work we show that the Lie bracket takes the form $[X, Y] = XY - YX$. In general, the set of invertible endomorphisms $GL(V)$ on a finite dimensional vector space forms a Lie group, the Lie algebra can be identified with the set of all endomorphisms $T : V \to V$, and the Lie bracket takes the form $[T, S] = T \circ S - S \circ T$.*

Example. *In general, if A is any associative algebra over K, then the commutator $[X, Y] = XY - YX$ gives a Lie algebra structure on A, denoted $\mathfrak{a}$. From the perspective of differential geometry, if A is an arbitrary finite-dimensional algebra, then A has the structure of a differential manifold since it is a vector space, and $U(A)$ is an open neighbourhood of the origin, hence a Lie group. The induced Lie algebra structure on $U(A)$ can then be identified with the Lie bracket on A defined by $[X, Y] = XY - YX$. Note, however, that this is no more general than our discussion of the general linear group, because any finite dimensional algebra A can be faithfully represented as endomorphisms $End(A)$ over the vector space structure of A by the map $\rho : A \to End(A)$, $\rho(X)(Y) = XY$, and this essentially identifies $U(A)$ as a subalgebra of $\mathfrak{gl}(V)$.*

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

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

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

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

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

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

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

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

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

so that $\partial_X X = X\partial_X + 1$, and we may use this identity to rearrange the product of any two differential operators to coincide with an operator of the form above. As

3

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

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

provided 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 physical measurements, which are non-commutative. If a field has finite characteristic p, this method only allows us to determine that

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

but $\partial_X^p = 0$ in this case, 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 $K\langle X_1,\ldots,X_n, Y_1,\ldots,Y_n\rangle$ modulo the relations $Y_iX_j - X_jY_i = \delta_{ij}$ 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.

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 $A = \mathbf{C}[X, X^{-1}]$, in which case we call $\mathfrak{gl}_n(A)$ a Loop Lie algebra. These algebras have applications in various fields of theoretical physics, including String theory.*

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

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

is always a derivation, because

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

4

Thus the set of derivations on A, denoted 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** $adj_X : \mathfrak{g} \to \mathfrak{g}$ defined by $adj_X(Y) = [X, Y]$, we can restate the equation above as*

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

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

$$
\begin{aligned}
[adj_X, adj_Y](Z) &= [adj_X adj_Y - adj_Y adj_X](Z) \\
&= [X, [Y, Z]] - [Y, [X, Z]] \\
&= [X, [Y, Z]] + [Y, [Z, X]] \\
&= [[X, Y], Z] \\
&= adj_{[X,Y]}(Z)
\end{aligned}
$$

The kernel of the representation being the centre of $\mathfrak{g}$, $Z(\mathfrak{g})$. In general, a bilinear skew-symmetric map is a Lie bracket if and only if the corresponding adjoint maps are all derivations; thus the Jacobi identity is exactly the derivation equation in disguise, harkening back to the definition of vector fields as derivations on the space of C^∞ functions on the Lie group.

Example. *The formula*

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

which can be proved by using the Leibnitz formula to expand the determinant out algebraically, and then only taking into the account the first-order terms. 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

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

5

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. Because $tr(XY) = tr(Y,X)$,

$$tr(XY - YX) = tr(XY) - tr(YX) = 0$$

Though this follows automatically because $SL_n(K)$ is a subgroup of $GL_n(K)$.

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

$$(I + hX)^t(I + hX) = I + h(X^t + X) + h^2 X^t X$$

X is a tangent vector at the origin for $O_n(K)$ if and only if $X^t = -X$. If char $K \neq 2$, then $X^t = -X$ implies that X vanishes along the diagonal, hence $\mathfrak{o}_n(K) \subset \mathfrak{sl}_n(K)$. This essentially follows because the special orthogonal Lie group $SO_n(K)$ of orthogonal matrices of determinant one forms a connected component of $O_n(K)$, hence $\mathfrak{so}_n(K) = \mathfrak{o}_n(K)$. Algebraically, $\mathfrak{o}_n(K)$ is a subalgebra of $\mathfrak{gl}_n(K)$ because if $X, Y \in \mathfrak{o}_n(K)$, then

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

so the Lie bracket preserves orthogonality.

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

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

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

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

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

$$(I + hM)^t J(I + hM)J^t = I + h(M^t + JMJ^t) + h^2 M^t JMJ^t$$

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

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

Then

$$X^tJ = \begin{pmatrix} -C^t & A^t \\ -D^t & B^t \end{pmatrix} \qquad JX = \begin{pmatrix} C & D \\ -A & -B \end{pmatrix}$$

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

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

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

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

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

Then we can identity $\mathfrak{g}$ 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 $A, B : \mathfrak{g}^2 \to \mathfrak{g}$. In the nicest case, we will have $A = B = 0$, 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 literally decomposed $\mathfrak{g}$ into more simple Lie algebras.

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

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

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

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

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

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

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 $(\mathfrak{g}/\mathfrak{a})/(\mathfrak{b}/\mathfrak{a})$ is isomorphic to $\mathfrak{g}/\mathfrak{b}$.*

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 $\mathfrak{b}/(\mathfrak{a} \cap \mathfrak{b})$.*

Example. *Consider the homomorphism $f : \mathfrak{gl}_n(K) \to K$ obtained by taking the trace of the matrix, which is a homomorphism because*

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

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

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

Hence the vector space structure of $\mathfrak{gl}_n(K)$ is the direct sum of $\mathfrak{sl}_n(K)$ and $K \cdot I$, and since $K \subset Z(K)$, the direct sum agrees with the Lie algebra operations as well.

1.1   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 in the derived subgroup $[G, G]$, and when we take the quotient we have an abelian group $G_{textab} := G/[G,G]$. On Lie algebras, we can take brackets of ideals to form new ideals

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

which is essentially the infinitisimal version of the product of two groups. Analogous to the commutator subgroups of a group is the derived subalgebra $\mathfrak{g}' = D\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, which is the smallest ideal such that $\mathfrak{g}/\mathfrak{g}'$ is trivial (we call this the *abelianization* of $\mathfrak{g}$, denoted $\mathfrak{g}_{ab}$). Since we have the exact diagram

$$0 \to \mathfrak{g}' \to \mathfrak{g} \to \mathfrak{g}_{ab} \to 0$$

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

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

for some Bilinear A and B. Thus we can think of the elements of $\mathfrak{g}'$ as infinitisimals, since they have no impact on the Lie bracket structure of $\mathfrak{g}_{ab}$, existing somewhat 'beneath the surface' of the calculations. If we consider $\mathfrak{g}'' = (\mathfrak{g}')'$, then we can compute the abelian approximation 'to a second order', writing $\mathfrak{g} = \mathfrak{g}_{ab} + \mathfrak{g}'_{ab} + \mathfrak{g}''$. Continuing this process, we hope to write $\mathfrak{g}$ as the successive product of abelian infinitisimals, $\mathfrak{g} = \mathfrak{g}_{ab}^{(1)} + \mathfrak{g}_{ab}^{(2)} + \cdots + \mathfrak{g}_{ab}^{(n)}$,

where $\mathfrak{g}^{(n)}$ is the n'th element of the **derived series**. For this to work, we require that $\mathfrak{g}^{(n+1)} = 0$ for some n, in which case we say $\mathfrak{g}$ is a **solvable** Lie algebra. This is equivalent to the existence of an abelian tower, and by refinement, a cyclic tower as well. Solvability on Lie algebras is a tower property, in the sense that if $\mathfrak{g}/\mathfrak{h}$ and $\mathfrak{h}$ are solvable, then $\mathfrak{g}$ is also solvable. It is natural to consider the **lower central series**

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

where the derived series is slightly finer, $\mathfrak{g}^{(n)} \subset \mathfrak{g}_n$, yet the infinitisimals have the additional property that if $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}_n$, then $[X,Y] \in \mathfrak{g}_{n+1}$ (for the derived series, we can only guarantee that $[X,Y] \in \mathfrak{g}^{(1)}$ if $X \in \mathfrak{g}$). If the lower central series eventually terminates, we call the algebra **nilpotent**. We then have a decomposition

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

Of course, this is a very strong condition for a Lie algebra to have. An important example of a nilpotent Lie algebra is the set of strictly upper triangular matrices.

As a corollary to this, we see that any Lie algebra $\mathfrak{g}$ has a maximum solvable ideal. If $\mathfrak{a}$ is a *maximal* solvable ideal, and $\mathfrak{b}$ is any other solvable ideal, then $\mathfrak{a} + \mathfrak{b}$ is solvable, hence $\mathfrak{a} + \mathfrak{b} = \mathfrak{a}$, and so $\mathfrak{b} \subset \mathfrak{a}$. We shall denote the maximum solvable ideal of a Lie algebra by $\mathrm{rad}(\mathfrak{g})$, and call it the **radical** of the algebra. The radical essentially separates the approximately commutative section of the algebra from the non-commutative section. The most non-commuatative Lie algebras are the **semi-simple** ones, such that $\mathrm{rad}(\mathfrak{g}) = (0)$. For any algebra $\mathfrak{g}$, $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$ is semisimple, so the radical efficiently extracts the commutative section of the algebra. It is an essential tool in the description of the finite dimensional Lie algebras, because it means we can write any Lie algebra as the product of a semisimple algebra and a solvable one.

The problem with using nilpotency to verify properties of a group is that it is not as amenable to algebraic manipulations than solvability. Just because $\mathfrak{a}$ and $\mathfrak{g}/\mathfrak{a}$ is nilpotent, this does not imply that $\mathfrak{g}$ is nilpotent. Thus we can't prove properties as easily by induction on the size of towers. Thus solvability is more natural. In conclusion, we have broken down study of an arbitrary Lie algebra $\mathfrak{g}$ into the study of a solvable Lie algebra $\mathrm{rad}(\mathfrak{g})$,

and a semisimple algebra $\mathfrak{g}/\mathrm{rad}(\mathfrak{g})$. Most of the study of the classification of Lie algebras concerns itself with understanding the solvable and semisimple Lie algebras.

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 semisimple 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. As we've stated before, the classification theorem for simple Lie algebras over the complex numbers gives a very simple set of families to understand, $\mathfrak{sl}_n$, $\mathfrak{so}_n$, $\mathfrak{sp}_n$, and some 'eccentric' algebras $\mathfrak{e}_6$, $\mathfrak{e}_7$, $\mathfrak{e}_8$, $\mathfrak{f}_4$, and $\mathfrak{g}_2$. This will require us to know some very sophisticated theorems of representation theory.

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 it's actions on vector spaces. This theory will become a powerful tool to classify all the Lie algebras. For now, we assume our Lie algebras occur as matrix subalgebras of $\mathfrak{gl}(V)$, but we note that our theory will soon be used to talk about general actions of Lie algebras on vector spaces.

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

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

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

$X \in \mathfrak{g}$. A weight is precisely a linear functional for which V_λ is non-trivial.

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

Example. *The vector e_1 is an eigenvector for the subalgebra of upper triangular matrices, and the corresponding weight is $\varepsilon(X) = X_{11}$. The corresponding weight space V_ε is just the span of e_1, because if v is any vector such that $Xv = X_{11}v$, then $v_1 \neq 0$, for otherwise Xv could not depend on X_{11}. If $v_i \neq 0$ for $i = 1$, we could change the i'th column of X, which would not adjust X_{11}, but it would change Xv.*

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

Lemma 2.1. *If $\mathfrak{a}$ is an ideal of a subalgebra $\mathfrak{g}$ on $\mathfrak{gl}(V)$, the space of vectors $v \in V$ such that $Xv = 0$ for all $X \in \mathfrak{a}$ is $\mathfrak{g}$ invariant.*

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

$$XYv = YXv = Y0 = 0$$

hence we have verified $\mathfrak{g}$-invariance. □

We can generalize this result to the case where v is an eigenvector, and the eigenvalue is non-zero. This is very important to the decomposition results of Lie algebras we will prove in the future.

Theorem 2.2. *Over a field K of characteristic zero, let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$ with ideal $\mathfrak{a}$. Then the weight space of any $\lambda : \mathfrak{a} \to K$ is $\mathfrak{g}$ invariant.*

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

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

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

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

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

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

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

Here's a simple consequence.

Theorem 2.3. *If $X, Y \in \mathfrak{gl}(V)$ commute with $[X,Y]$, then $[X,Y]$ is nilpotent.*

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

2.1   Engel's Theorem

In linear algebra, we form classification theorems for linear operators. The theorems show the existence of certain basis elements on the vector space, such that the corresponding matrix representation of the operators have nice structure. But given a family of linear operators, it is a much more difficult problem to find a basis such that *all* the matrix representations of the family of linear operators all 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 2.4. *If $\mathfrak{g}$ is a Lie subalgebra of nilpotent transformations on V, then there is $v \neq 0$ for which $Xv = 0$ for all $X \in \mathfrak{g}$.*

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

In general, let $\mathfrak{a}$ be a maximal proper Lie subalgebra of $\mathfrak{g}$. For any $X \in \mathfrak{a}$, the adjoint map $\mathrm{adj}_X : \mathfrak{g} \to \mathfrak{g}$ descends to a map from $\varphi_X : \mathfrak{g}/\mathfrak{a} \to \mathfrak{g}/\mathfrak{a}$, because if $Y - Z \in \mathfrak{a}$, then $[X,Y] - [X,Z] = [X, Y - Z] \in \mathfrak{a}$. Thus $\varphi(\mathfrak{a})$ is a subalgebra of $\mathfrak{gl}(V)$, and $\dim \varphi(\mathfrak{a}) < \dim \mathfrak{g}$. To apply the induction hypothesis, we require that $\varphi(X)$ is nilpotent when X is nilpotent, but this follows because adj_X is nilpotent when X is nilpotent, because we have

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

and if $X^n = 0$, then $\mathrm{adj}_X^m(Y) = 0$ when $m \geqslant 2n$. Thus, by induction, there is $Y + \mathfrak{a} \in \mathfrak{g}/\mathfrak{a}$ with $Y \notin \mathfrak{a}$ such that $[X,Y] \in \mathfrak{a}$ for all $X \in \mathfrak{a}$. This implies that $\mathfrak{a} + KY$ is a subalgebra of $\mathfrak{g}$, and by maximality, $\mathfrak{a} + KY = \mathfrak{g}$. But $\mathfrak{a}$ is an ideal of $\mathfrak{a} + KY$, hence $\mathfrak{a}$ is an ideal of $\mathfrak{g}$ with codimension 1.

Now we apply induction again to conclude that the space W of $v \in V$ with $Xv = 0$ for all $X \in \mathfrak{a}$ is non-trivial. Now W is invariant under multiplication by $\mathfrak{g}$, so in particular $YW \subset W$. Since Y is nilpotent, we find a non-zero $w \in W$ with $Yw = 0$, and then $(\mathfrak{a} + KY)w = 0$. $\square$

Theorem 2.5 (Engel). *For any algebra $\mathfrak{g}$ of nilpotent linear transformations on $\mathfrak{gl}(V)$, there is a basis which simultaneously upper triangularizes all elements of the algebra.*

Proof. We adapt the proof strategy in the case of a single transformation. First, we find a vector v such that $Xv = 0$ for all $X \in \mathfrak{g}$. If $W = Kv$, then each X descends to a linear endomorphism on V/W. The image of $\mathfrak{g}$ under this descending process satisfies the hypothesis of Engel's theorem, and therefore by induction there is a basis $v_1 + W, \ldots, v_m + W$ of V/W in which all X are upper triangular. Then $v, v_1, \ldots, v_m$ is a basis for V, But then

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

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

15

Corollary 2.6. *A Lie algebra* $\mathfrak{g}$ *is nilpotent iff* 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

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

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

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

Lemma 2.7. *If* $\mathfrak{g} \subset \mathfrak{gl}(V)$ *is a solvable Lie algebra over the complex numbers, and* $V \neq 0$, *there there is* $v \neq 0$ *which is an eigenvector for every element of* $\mathfrak{g}$.

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

The work of Lie's theorem is essentially the same techniques as Engel's theorem.

Theorem 2.8. *If* $\mathfrak{g} \subset \mathfrak{gl}(V)$ *is a solvable Lie algebra, then there is a basis for* V *such that the matrix representations of* $\mathfrak{g}$ *are upper triangular.*

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

If $\mathfrak{g}$ is a complex solvable Lie algebra, which is a subset of some $\mathfrak{gl}(V)$, then Lie tells us there is a basis of V in which every element of $\mathfrak{g}$ is upper triangular. If A and B are upper triangular matrices, then

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

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

Chapter 3

Representation Theory

In this chapter, we 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 $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$. The homomorphism induces an **action** of $\mathfrak{g}$ on V, defined by $Xv = \rho(X)(v)$. Conversely, if we have an action $\mathfrak{g} \times V \to V$ of $\mathfrak{g}$ on V, satisfying

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

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

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

Then this induces a representation of $\mathfrak{g}$ on $\mathfrak{gl}(V)$. A representation is **faithful** if it is injective, and **transitive** if we can map any $v \in V$ to any $w \in V$ by some $X \in \mathfrak{g}$.

Example. *We have already seen that the adjoint map $ad : \mathfrak{g} \to \mathbf{gl}(\mathfrak{g})$ is a representation of $\mathfrak{g}$. It occurs very often in the theory. The kernel of the homomorphism is the center $Z(\mathfrak{g})$.*

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 $\mathfrak{g}$ are exactly the ideals. Lie's theorem says that every module over a complex solvable Lie algebra has a one dimensional submodule. Given a submodule W, we can form the factor module V/W.

We wish to break modules down into their simpler representations, so that the action of $\mathfrak{g}$ on the modul becomes clear. 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. Thus 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. The Lie algebras which are irreducible modules over their adjoint representation are exactly the simple ones. The only irreducible modules over a complex solvable algebra 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.

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. 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 3.1 (Schur)**.** *If $\mathfrak{g}$ is a complex Lie algebra, 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 $T : V \to V$ be a Lie algebra homomorphism. Then 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 $T = \lambda$. □

It follows that if $X \in Z(\mathfrak{g})$, and V is an irreducible $\mathfrak{g}$ algebra, then there exists λ such that $Xv = \lambda v$ for all $v \in V$, because X acts as a module endomorphism. We can also conclude that the only irreducible **C** modules are one-dimensional.

3.1 Representations of $\mathfrak{sl}_2(\mathbf{C})$

Many of the ideas which occur in the general representation theory of complex Lie algebras occurs in the theory of representations of $\mathfrak{sl}_2(\mathbf{C})$. 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_{12}, $E_{11} - E_{22}$, and E_{21} spans the algebra. We begin by constructing a family of irreducible representations.

For each n, define V_n to be the subspace of $\mathbf{C}[X, Y]$ generated by the homogenous polynomials with degree n

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

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

$$\rho \begin{pmatrix} a & b \\ c & -a \end{pmatrix} = a \left(X \frac{\partial}{\partial X} - Y \frac{\partial}{\partial Y} \right) + bX \frac{\partial}{\partial Y} + cY \frac{\partial}{\partial X}$$

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 the $X^a Y^b$ are eigenvectors of $\rho(h)$. By construction, ρ is linear. To verify that brackets are preserved, we need only verify this for the basis. We calculate $[E_{12}, E_{21}] = E_{11} - E_{22}$, and using the relations we calculated for the Weyl algebra, we find

$$[\rho(E_{12}), \rho(E_{21})] = \left[X \frac{\partial}{\partial Y}, Y \frac{\partial}{\partial X} \right] = X \frac{\partial}{\partial Y} Y \frac{\partial}{\partial X} - Y \frac{\partial}{\partial X} X \frac{\partial}{\partial Y}$$

$$= \left(1 + Y \frac{\partial}{\partial Y} \right) \left(X \frac{\partial}{\partial X} \right) - \left(1 + X \frac{\partial}{\partial X} \right) \left(Y \frac{\partial}{\partial Y} \right)$$

$$= X \frac{\partial}{\partial X} - Y \frac{\partial}{\partial Y} = \rho(E_{11} - E_{22})$$

Next, we calculate $[E_{12}, E_{11} - E_{22}] = -2E_{12}$, and find

$$[\rho(E_{12}), \rho(E_{11} - E_{22})] = \left[X\frac{\partial}{\partial Y}, X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y} \right]$$

$$= X^2 \frac{\partial}{\partial X}\frac{\partial}{\partial Y} - X\left(1 + X\frac{\partial}{\partial X}\right)\frac{\partial}{\partial Y}$$

$$- X\left(1 + Y\frac{\partial}{\partial Y}\right)\frac{\partial}{\partial Y} + XY\frac{\partial^2}{\partial Y^2}$$

$$= -2X\frac{\partial}{\partial Y} = -2\rho(E_{12})$$

Finally, $[E_{21}, E_{11} - E_{22}] = 2E_{21}$, and we can reduce the calculation of the product to the calculation above by exchanging variables, so

$$[\rho(E_{21}), \rho(E_{11} - E_{22})] = \left[Y\frac{\partial}{\partial X}, X\frac{\partial}{\partial X} - Y\frac{\partial}{\partial Y} \right] = -\left[Y\frac{\partial}{\partial X}, Y\frac{\partial}{\partial Y} - X\frac{\partial}{\partial X} \right]$$

$$= 2Y\frac{\partial}{\partial X} = 2\rho(E_{21})$$

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

$$\rho(E_{12}) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & n \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix} \quad \rho(E_{21}) = \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \\ n & 0 & \ldots & 0 & 0 \\ 0 & n-1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}$$

$$\rho(E_{11} - E_{22}) = \begin{pmatrix} n & 0 & \ldots & 0 & 0 \\ 0 & n-2 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 2-n & 0 \\ 0 & 0 & \ldots & 0 & -n \end{pmatrix}$$

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

Theorem 3.2. V_n is an irreducible $\mathfrak{sl}_n(\mathbf{C})$ module.

Proof. Let W be a non-zero submodule of V_n. An eigenvector of $\rho(E_{11} - E_{22})$ lies in W, because $\rho(E_{11} - E_{22})$ restricts to an operation on this space. But this implies that W contains a monomial, since the eigenspaces of $\rho(E_{11} - E_{22})$ are all one dimensional. $\qquad\square$

It turns out that the class of V_n represent all irreducible modules of $\mathfrak{sl}_2(\mathbf{C})$. The trick, given a general module V, is to look at the eigenvectors of $E_{11} - E_{22}$.

Lemma 3.3. *If $v \in V$ is an eigenvector of $E_{11} - E_{22}$ with eigenvalue λ, then*

  (i) $E_{12}v = 0$, or $E_{12}v$ is an eigenvector of $E_{11} - E_{22}$ with eigenvalue $\lambda + 2$.

  (ii) $E_{21}v = 0$, or $E_{21}v$ is an eigenvector of $E_{11} - E_{22}$ with eigenvalue $\lambda - 2$.

Proof. (i) is proved by calculation,

$$(E_{11} - E_{22})(E_{12}v) = E_{12}(\lambda v) + [E_{11} - E_{22}, E_{12}]v$$
$$= \lambda E_{12}v + 2E_{12}v = (\lambda + 2)E_{12}v$$

and essentially the same calculation shows (ii). $\qquad\square$

Lemma 3.4. *If V is a finite dimensional $\mathfrak{sl}_2(\mathbf{C})$ module, then V contains an eigenvector v for $E_{11} - E_{22}$ such that $E_{12}v = 0$.*

Proof. Using the last lemma, we find that if v is an eigenvector with eigenvalue λ, then either $E_{12}v = 0$, or $E_{12}v$ is an eigenvector with eigenvalue $\lambda + 2$, and since the eigenvalue is different, $E_{12}v$ is independent of v. Continuing this process, we either find that $E_{12}^m v$ is an eigenvector for v with eigenvalue $\lambda + 2m$ for all m, or there is an eigenvector w satisfying $E_{12}w = 0$. But if each $E_{12}^m v$ is an eigenvector, then they form an infinite family of independent vectors, contradicting finite dimensionality. $\qquad\square$

Theorem 3.5. *If V is a finite dimensional irreducible $\mathfrak{sl}_2(\mathbf{C})$ module, then V is isomorphic to some V_d.*

Proof. Find an eigenvector v for $E_{11} - E_{22}$ such that $E_{12}v = 0$, with eigenvalue λ. Consider the sequence $v, E_{21}v, E_{21}^2 v, \ldots$. The last lemma essentially implies that $E_{21}^{m+1} v = 0$, $E_{21}^m v \neq 0$. We claim that $v, E_{21}v, \ldots, E_{21}^m v$ form a basis for a submodule of V. They are certainly linearly independant, since they are eigenvectors of $E_{11} - E_{22}$ with different eigenvalues. It

22

remains to show that $E_{12}(E_{21}^k v) \in \text{span}(E_{21}^l v)$, for $l \leqslant k$. For $k = 0$, we know $E_{12} v = 0$. For the induction, we find

$$E_{12}(E_{21}^k v) = (E_{21} E_{12} + (E_{11} - E_{22}))(E_{21}^{k-1} v)$$
$$= E_{21}(E_{12}(E_{21}^{k-1} v)) + (\lambda + 2(k-1))E_{21}^{k-1} v$$

and by induction, $E_{12}(E_{21}^{k-1} v) \in \text{span}(E_{21}^l v)$, for $l < k$, and therefore $E_{12}(E_{21}^{k-1} v) \in \text{span}(E_{21}^l v)$, for $l \leqslant k$.

By irreducibility, V is the span of the $E_{21}^k v$. The matrix of $E_{11} - E_{22}$ with respect to the basis $E_{21}^k v$ is diagonal, with trace

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

hence $\lambda = m$, since the image of E_{21} 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, $E_{21} X^m, \ldots, E_{21}^m X^m$. If we set $\psi(E_{21}^k v) = E_{21}^k X^m$ this defines a vector space isomorphism which commutes with the action of $E_{11} - E_{22}$ and E_{21}. It remains to show that ψ commutes with E_{12}, we use induction. For $k = 0$,

$$E_{12} X^m = 0 = \psi(E_{12} v)$$

Now by induction,

$$
\begin{aligned}
E_{12}(E_{21}^k X^m) &= E_{21}(E_{12}(E_{21}^{k-1} X^m)) + (E_{11} - E_{22})(E_{21}^{k-1} X^m) \\
&= \psi(E_{21}(E_{12}(E_{21}^{k-1} v))) + (\lambda - 2(k-1))E_{21}^{k-1} X^m \\
&= \psi(E_{12}(E_{21}^k v)) - \psi((E_{22} - E_{11})E_{21}^{k-1} v) + (\lambda - 2(k-1))E_{21}^{k-1} X^m \\
&= \psi(E_{12}(E_{21}^k v)) + (\lambda - 2(k-1))(E_{21}^{k-1} X^m - \psi(E_{21}^{k-1} v)) \\
&= \psi(E_{12}(E_{21}^k v))
\end{aligned}
$$

and this completes the correspondence. $\qquad\square$

Corollary 3.6. *If V is a finite dimensional $\mathfrak{sl}_2(\mathbf{C})$ and $v \in V$ is an eigenvector of $E_{11} - E_{22}$ such that $E_{12} v = 0$, then $E_{21} v = mv$ for some integer m, and the submodule of V generated by v is isomorphic to V_m.*

23

Proof. We have argued that $v, E_{21}v, \ldots, E_{21}^m v$ span the submodule generated by v. Now we just apply irreducibility to conclude that the module generated is isomorphic to V_m. □

The vector v we constructed with the largest eigenvalue is known as the **heighest weight vector**. The associated eigenvalue m is known as the **highest weight**.

3.2   Weyl's Theorem

Theorem 3.7 (Weyl). *If $\mathfrak{g}$ is a complex, semisimple Lie algebra, then every finite dimensional module over $\mathfrak{g}$ is completely reducible.*

3.3   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 complex can be uniquely written $T = D + N$, where D is a diagonalizable operator, N is a nilpotent operator, and D and N commute. $N = 0$ precisely when T is diagonalizable, in which case we say T is **semisimple**.

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 $\mathrm{tr}(XY) = 0$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}'$, because

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

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

Theorem 3.8. *If $\mathfrak{g}$ is a complex Lie subalgebra of $\mathfrak{gl}(V)$ such that $\mathrm{tr}(XY) = 0$ for $X, Y \in \mathfrak{g}$ then $\mathfrak{g}$ is solvable.*

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

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

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

$$ \mathrm{tr}(\overline{D}(Y_i Z_i - Z_i Y_i)) = \mathrm{tr}([\overline{D}, Y_i] Z_i) $$

provided we can show that $[\overline{D}, Y_i] \in \mathfrak{g}$, we can apply our hypothesis to conclude the trace is zero. Note that the Jordan decomposition of adj_X is $\mathrm{adj}_D + \mathrm{adj}_N$, and therefore there is a polynomial $f \in \mathbf{C}[X]$ with $f(\mathrm{adj}_X) = \overline{adj_D} = \mathrm{adj}_{\overline{D}}$. Since adj_X maps $\mathfrak{g}$ into itself, so too does $\mathrm{adj}_{\overline{D}}$. Thus the proof is completed. $\square$

Since $\mathfrak{g}$ is solvable if and only if it's image in the adjoint representation is solvable, we conclude that $\mathfrak{g}$ is solvable if and only if $\mathrm{tr}(\mathrm{adj}_X \circ \mathrm{adj}_Y) = 0$ 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

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

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 adjoint have matrix representations*

$$ adj_X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad adj_Y = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} $$

and by looking at these representations we see $\kappa = Y^* \otimes Y^*$. *The Killing form vanishes on the span of X, and therefore our algebra is solvable. Indeed, the derived subalgebra is abelian.*

The Killing form descends on ideals consistantly

Lemma 3.9. *If $\mathfrak{a}$ is an ideal of $\mathfrak{g}$, then the Killing form on $\mathfrak{a}$ is the restriction of the Killing form on $\mathfrak{g}$.*

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

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

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

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

and so

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

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

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

Lemma 3.10. *A Lie algebra is semisimple if and only if it has no non-zero abelian ideals.*

Proof. If an algebra $\mathfrak{g}$ has a non-zero abelian ideal, the ideal is solvable, and thus $\mathrm{rad}(\mathfrak{g}) \neq 0$. Note that $Z(\mathfrak{g})$ is always an ideal of $\mathfrak{g}$, because if $[X, Y] = 0$ for all Y, then $[X, Y] = 0 \in Z(\mathfrak{g})$. Thus in a semisimple algebra, $Z(\mathfrak{g}) = 0$. 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, $[Y, Z] \in \mathfrak{a}$, and hence

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

In any solvable Lie algebra $\mathfrak{h}$, $\mathfrak{h}'$ is nilpotent, hence $\mathfrak{h}'_{k+1} = [\mathfrak{h}, \mathfrak{h}_k] = 0$ for some k, implying that $Z(\mathfrak{h}')$ is a non-trivial abelian subalgebra of $\mathfrak{h}$, and we have also show it is in fact an ideal of $\mathfrak{h}$. □

Theorem 3.11 (Cartan's Second Criterion). $\mathfrak{g}$ *is semisimple if and only if the Killing form is nondegenerate.*

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

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

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

Lemma 3.12. *If $\mathfrak{g}$ is a complex semisimple Lie algebra, and $\mathfrak{a}$ is an ideal, then $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$. $\mathfrak{a}$ is also a semisimple Lie algebra.*

Proof. Because $\kappa : \mathfrak{a} \cap \mathfrak{a}^\perp \to \mathfrak{a} \cap \mathfrak{a}^\perp$ is equal to zero, $\mathfrak{a} \cap \mathfrak{a}^\perp$ is solvable, 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 $[X, Y] = 0$, because $\mathfrak{a}$ and $\mathfrak{a}^\perp$ are both ideals so $[X, Y] \in \mathfrak{a} \cap \mathfrak{a}^\perp$. Thus $\mathfrak{g}$ really is the Lie algebra sum of the two ideals. Finally, if $\mathfrak{a}$ was not semisimple, then $\kappa_\mathfrak{a}$ would be degenerate, but this would imply the degeneracy of κ. □

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

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

27

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

Theorem 3.13. *If $\mathfrak{g}$ is a finite dimensional complex 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 adj $\mathfrak{g}$. This subalgebra is actually an ideal, for if d is any derivation, then

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

Thus adj $\mathfrak{g}$ is an ideal on Der $\mathfrak{g}$. This is true for any Lie algebra.

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

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 $\mathbf{C}$ over some vector space V are obtained my mapping 1 to an arbitrary element of V.

Theorem 3.14. *If $\mathfrak{g}$ is a complex Lie algebra, and $D \in Der(\mathfrak{g})$ has Jordan decomposition $D = A + N$, then A and N are both derivations on $\mathfrak{g}$.*

Proof. For each $\lambda \in \mathbf{C}$, 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

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

28

Since A is diagonalizable, the λ eigenspace for A 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

$$A[X,Y] = (\lambda + \gamma)[X,Y] = [AX,Y] + [X,AY]$$

hence A is a derivation. N is then a derivation, because it is the difference of two derivations. $\qquad\square$

Theorem 3.15. *If $\mathfrak{g}$ is semisimple, then any $X \in \mathfrak{g}$ can be written uniquely as $D + N$, where adj_D is diagonalizable, adj_N is nilpotent, and $[D,N] = 0$. If Y commutes with X, then $[D,Y] = [N,Y] = 0$.*

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

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

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

We call the decomposition $X = D + N$ the **abstract Jordan decomposition** of X.

Theorem 3.16. *If $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation of a complex semisimple Lie algebra, and $X \in \mathfrak{g}$ has a Jordan decomposition $D + N$, then the (concrete) Jordan decomposition of $\rho(X)$ is $\rho(D) + \rho(N)$.*

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

Chapter 4

Root Systems

Recall the proof that $\mathfrak{sl}_2(\mathbf{C})$ is the only simple 2 dimensional Lie algebra.

1. We found $X \in \mathfrak{sl}_2(\mathbf{C})$ such that adj_X is diagonalizable.

2. We took a basis of $\mathfrak{sl}_2(\mathbf{C})$ consisting of eigenvectors for adj_X, and showed that the corresponding structural constants

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

Let us try and generalize the technique to $\mathfrak{sl}_n(\mathbf{C})$. Consider the subalgebra $\mathfrak{h}$ of $\mathfrak{sl}_n(\mathbf{C})$ consisting of the diagonal matrices. Since

$$[E_{ij}, E_{kl}] = \delta_j^k E_{il} - \delta_l^i E_{kj}$$

We find that for any diagonal matrix $X = (\sum a_i E_{ii})$, if $m \neq n$, then

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

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

precisely the space of diagonal matrices, which is $n-1$ dimensional. Since the weight spaces form $\mathfrak{h}$-invariant subspaces, we have a module decomposition

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

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

In general, our aim is to find a maximal abelian subalgebra consisting only of elements X with adj_X diagonalizable (we call these **semisimple** elements), and then to decompose $\mathfrak{g}$ into the corresponding weight spaces. We shall call a maximal abelian subalgebra containing only semisimple elements a **Cartan subalgebra**.

It is natural to ask why abelian subalgebras consisting of semisimple elements are useful. Given any particular abelian subalgebra $\mathfrak{h}$, the adjoint maps adj_X for $X \in \mathfrak{h}$ commute with one another – if $X, Y \in \mathfrak{h}$, then

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

hence $\mathrm{adj}_X \circ \mathrm{adj}_Y = \mathrm{adj}_Y \circ \mathrm{adj}_X$. Since each adj_X is diagaonalizable, 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

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\lambda \neq 0} \mathfrak{g}_\lambda$$

where λ ranges over the set of weights for $\mathfrak{h}$. Given $\mathfrak{g}_\lambda$ and $\mathfrak{g}_\gamma$, $[\mathfrak{g}_\lambda, \mathfrak{g}_\gamma] \subset \mathfrak{g}_{\lambda+\gamma}$, because if $X \in \mathfrak{g}_\lambda$, $Y \in \mathfrak{g}_\gamma$, then by using the Jacobi identity we find that for any $Z \in \mathfrak{h}$,

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

If $\lambda + \gamma \neq 0$, then $\kappa(\mathfrak{g}_\lambda, \mathfrak{g}_\gamma) = 0$. Indeed, if $X \in \mathfrak{g}_\lambda$ and $Y \in \mathfrak{g}_\gamma$, and $Z \in \mathfrak{h}$

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

If we choose Z such that $\lambda(Z) + \gamma(Z) \neq 0$, we conclude $\kappa(X, Y) = 0$. This shows that κ is non-degenerate on $\mathfrak{g}_0$, because if there is $X \in \mathfrak{g}_0$ such that $\kappa(X, Y) = 0$ for all $Y \in \mathfrak{g}_0$, then any $Z \in \mathfrak{g}$ can be expanded as $\sum Z_\lambda$, with

$Z_\lambda \in \mathfrak{g}_\lambda$, and $\kappa(X, Z) = \kappa(X, Z_0)$, so that X annihilates all elements of $\mathfrak{g}$, and hence $X = 0$. It follows that $\mathfrak{h}$ is semisimple.

If $\mathfrak{h}$ is small, then the weight decomposition of $\mathfrak{g}$ is likely to be rather coarse. If we choose $\mathfrak{h}$ too small, then $\mathfrak{h}$ may be a proper subset of $\mathfrak{g}_0$, hence the decomposition and nice properties of the Killing form may be not so useful.

Example. *If $\mathfrak{h}$ is the Lie subalgebra of $\mathfrak{sl}_n(\mathbf{C})$ 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 $X_{2n} = X_{2n} = 0$ 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.*

Theorem 4.1. *Complex semisimple Lie algebras have a non-zero Cartan subalgebra $\mathfrak{h}$. If $X \in \mathfrak{h}$ minimizes the dimension of $C_\mathfrak{g}(X)$, then every $Y \in \mathfrak{h}$ is central in $C_\mathfrak{g}(X)$, and so $C_\mathfrak{g}(X) \subset C_\mathfrak{g}(Y)$ and therefore $C_\mathfrak{g}(X) = C_\mathfrak{g}(\mathfrak{h})$.*

Proof. Nah. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Theorem 4.2. *If $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$ then $C_\mathfrak{g}(\mathfrak{h}) = \mathfrak{h}$, hence $\mathfrak{h}$ is self centralizing, and if we consider the weight decomposition with respect to $\mathfrak{h}$, then $\mathfrak{g}_0 = \mathfrak{h}$.*

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

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

where Φ is the set of roots of $\mathfrak{g}$, 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.

Lemma 4.3. *If α is a non-zero root, then $-\alpha$ is a root, and for every non-zero $X \in \mathfrak{g}_\alpha$ there is $Y \in \mathfrak{g}_{-\alpha}$ such that $\mathrm{span}(X, Y, [X, Y])$ is a Lie subalgebra of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}_2(\mathbf{C})$.*

4.1 Cartan Subalgebras as Inner Product Spaces

Lemma 4.4. *If $X \in \mathfrak{h}$ is non-zero, then there is a root α with $\alpha(X) \neq 0$, and the roots span $\mathfrak{h}^*$.*

Proof. If $\alpha(X) = 0$ for all roots X, then for any $Y \in \mathfrak{g}$, we can write $Y = Y_0 + \sum_{\alpha \in \Phi} Y_\alpha$, then $[X, Y] = \sum \alpha(X) Y = 0$, hence $X \in Z(\mathfrak{g}) = 0$, hence $X = 0$. That the roots span $\mathfrak{h}^*$ is essentially a restatement of what we have proved. $\qquad\square$

Since the Killing form on $\mathfrak{h}$ is non-degenerate, it induces a canonical isomorphism of $\mathfrak{h}$ with $\mathfrak{h}^*$, and therefore the Killing form can be applied to elements of the dual as well.

4.2   Root Spaces for the Classical Lie Algebras

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

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

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

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

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

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