

The Representation Theory of Lie Algebras

Jacob Denson

April 25, 2017

Table Of Contents

1	Why Study Lie Algebras?	2
2	Basic Definitions	8
2.1	Simply Connected Lie Groups	17
2.2	Factorization	21
2.3	Solvable Lie Algebras	23
2.4	Low Dimensional Classifications	26
3	Matrix Lie Algebras	39
3.1	Basic Representation Theory	39
3.2	Engel and Lie's Theorem	48
3.3	Representations of $\mathfrak{sl}_2(K)$	56
3.4	Projective Geometry and $\mathfrak{sl}_2(\mathbb{C})$	61
4	Killing Forms	64
4.1	Cartan's Criterion	64
4.2	The Casimir Element and Weyl's Theorem	72
4.3	Cartan Subalgebras	76
4.4	Subalgebras isomorphic to $\mathfrak{sl}_2(K)$	84
4.5	Cartan Subalgebras as Inner Product Spaces	89
5	Root systems	92
5.1	The Weyl Group	101
5.2	Classification of Root Systems	107
6	The Classification of Semisimple Lie Algebras	111

7	Advanced Representation Theory	114
7.1	Universal Enveloping Algebra	116
7.2	Verma Modules	118
7.3	Representations of $\mathfrak{sl}_3(K)$	118
8	Verma Modules	121

Chapter 1

Why Study Lie Algebras?

Many geometric symmetries occur as a group with complicated, nonlinear algebraic operations. However, these symmetries often have a topological structure which makes the operations differentiable. We can therefore linearize the algebraic operations, obtaining simpler maps on the space of infinitesimals. The algebraic study of these linearized operations is known as the theory of Lie algebras. The entire premise behind this study is that algebraic information about the Lie group can be fully realized as a property of the Lie algebra of infinitesimals at the identity, normally in terms of linear operations which are much easier to study than the original operations on the Lie group.

If G is a connected topological group, most of the algebraic data can be sufficiently described locally around the origin of the group. Indeed,

- Every neighbourhood of the identity generates the group.
- Every continuous homomorphism is uniquely determined on a neighbourhood of the identity.

The first point implies, with some intricacy, that we can obtain global information about a group from properties of the group that occur in a neighbourhood of the identity. The second implies that homomorphisms of topological groups essentially only result from local data on the group. Now if G is given the stronger condition of being a connected *Lie group*, then the smooth homomorphisms $\phi : G \rightarrow H$ can be differentiated to obtain linear maps $\phi_* : \mathfrak{g} \rightarrow \mathfrak{h}$ between the space \mathfrak{g} of tangent vectors to G at the identity in G , and the space \mathfrak{h} of tangent vectors to H at the identity.

We would hope that, since the homomorphism is uniquely defined on each neighbourhood of the identity, we can pass one step further, and conclude that the homomorphism is uniquely specified on the infinitesimals at the origin, and indeed, this turns out to be true.

Fact 1.1. *If ϕ and ψ are two homomorphisms with $\phi_* = \psi_*$, then $\phi = \psi$.*

Since the family of homomorphisms of a group tells us all the algebraic structure of the group, it is impressive, and very useful, that we can classify the homomorphisms of a Lie group in terms of linear maps on a vector space, which have a very complete theory. However, not all linear maps at the identity correspond to group homomorphisms on the Lie group, and thus it is desirable to find algebraic conditions which specify which linear maps can be lifted to homomorphisms. We should expect these linear maps to satisfy a condition obtained by preserving some linearized form of the canonical algebraic operations on the Lie group, and indeed, we shall find this is the case. This linearized operation is the Lie bracket on the space of infinitesimals. This is almost the complete reason why the theory of Lie algebras is studied today.

One might expect that since the operations of multiplication and inversion are so fundamental to the definition of the group, their first order approximations alone suffice to describe the infinitesimal homomorphisms. However, if $m : G \times G \rightarrow G$ is a multiplication map, and $i : G \rightarrow G$ an inversion map, then we actually calculate that

$$m_*(X, Y) = X + Y \quad i_*(X) = -X$$

Essentially, this means that all Lie groups are commutative up to first order. This is bad news, because every linear map preserves addition and negation, hence infinitesimally preserving multiplication and inversion up to first order. These operation only classify the dimension of the Lie group, and not any of its interesting algebraic structure.

The correct operation we linearize to obtain a condition describing infinitesimal homomorphisms is the *second order* approximation of conjugation, and this is the Lie bracket. Viewing elements of \mathfrak{g} both as derivations and as curves, the Lie bracket takes curves $\lambda, \gamma \in \mathfrak{g}$, and returns another tangent vector $[\lambda, \gamma] \in \mathfrak{g}$, which acts as a derivation by the equation

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

Surprisingly, the second order terms of conjugation are sufficient to characterize the Lie group operation. Because of this, the Lie bracket operation is essential. Together with this operation, \mathfrak{g} is known as the **Lie algebra** associated with the Lie group G .

If $\phi : G \rightarrow H$ is a group homomorphism, then in two different ways, we calculate that

$$\begin{aligned} f(\phi(\lambda_t \gamma_t \lambda_t^{-1} \gamma_t^{-1})) &= f(e) + t^2[\lambda, \gamma](f \circ \phi) + o(t^3) \\ &= f(e) + t^2\phi_*[\lambda, \gamma](f) + o(t^3) \\ f(\phi(\lambda_t \gamma_t \lambda_t^{-1} \gamma_t^{-1})) &= f((\phi \circ \lambda)_t(\phi \circ \gamma)_t(\phi \circ \lambda)_{-t}(\phi \circ \gamma)_{-t}) \\ &= f(e) + t^2[\phi_*(\lambda), \phi_*(\gamma)](f) + o(t^3) \end{aligned}$$

It follows that $\phi_*[\lambda, \gamma] = [\phi_*(\lambda), \phi_*(\gamma)]$, so all linear maps on infinitesimals induced by homomorphisms of groups preserve the Lie bracket operation. The converse requires some technical machinery, of which we provide a brief overview. Given any Lie group G , there is a unique map $\exp : \mathfrak{g} \rightarrow G$, also denoted $X \mapsto e^X$ and known as the **exponential** of the Lie group, such that

- For any fixed X , $t \mapsto \exp(tX)$ is a homomorphism from \mathbf{R} to G .
- For a fixed X , the exponential map satisfies

$$\frac{de^{tX}}{dt} = X$$

and as we vary t , the map is the unique integral curve through the origin for the left-invariant vector field generated by X .

- If $\phi : G \rightarrow H$ is a homomorphism, then $e^{\phi_*(X)} = \phi(e^X)$, so the exponential is a ‘natural operation’ in the category of Lie groups.
- \exp need not be injective nor surjective, but there is a neighbourhood of the origin in \mathfrak{g} upon which \exp is a diffeomorphism with a neighbourhood of the identity in G . In the case of $GL_n(\mathbf{C})$, the map is surjective, as with any connected, compact Lie group.

The equation $e^{\phi_*(X)} = \phi(e^X)$ gives much stronger conditions on the linear map ϕ_* to be the differential of a homomorphism. Since the image of

the exponential map always contains a neighbourhood of the identity, we can try *defining* a map ϕ on this neighbourhood by the equation $\phi(e^X) = e^{\phi_*(X)}$. Given sufficient conditions on ϕ_* , we can conclude that ϕ is not only well defined, but also *locally a homomorphism* around the identity, in the sense that $\phi(gh) = \phi(g)\phi(h)$ where the equation is defined, and given that the group G is simply connected, we can actually extend ϕ to a homomorphism on the entire group. The sufficient conditions on ϕ_* is that it preserves the Lie bracket on \mathfrak{g} , which is where the study of Lie algebras enters the picture.

That preserving the Lie bracket is sufficient to obtain a local homomorphism follows from a complex analytical expression for the product of two exponentials, known as the Baker-Hausdorff formula. First, note that since \exp is locally a diffeomorphism around the origin, we can locally find a smooth inverse, which we denote by \log , the logarithm of the Lie algebra.

Fact 1.2. *There are universal constants a_α , with α ranging over all multindices such that for any group G , there is a neighbourhood of the origin in G such that on this neighbourhood,*

$$\log(e^X e^Y) = X + Y + (1/2)[X, Y] + \sum_{|\alpha| \geq 2} a_\alpha (X, Y)^\alpha$$

where if $\alpha = (\alpha_1, \dots, \alpha_n)$, then $(X, Y)^\alpha$ is the nasty composition of Lie brackets

$$(X, Y)^\alpha = \underbrace{[X, [X, [\dots, [X, [Y, [\dots, [Y, \dots]]]]]]}_{\alpha_1 \text{ times}} \underbrace{[Y, [\dots, [Y, \dots]]]}_{\alpha_2 \text{ times}}$$

So that the Lie bracket is essentially the second order approximation of the multiplication operation.

First, note that if ϕ_* is a linear map preserving the Lie bracket operation, with $\phi_*[X, Y] = [\phi_*X, \phi_*Y]$, then one can prove by induction that $\phi_*(X, Y)^\alpha = (\phi_*X, \phi_*Y)^\alpha$, and together with the Baker-Hausdorff formula, this implies that

$$\log(e^{\phi_*X} e^{\phi_*Y}) = \phi_* \log(e^X e^Y)$$

This implies that if ϕ_* preserves the Lie bracket, then the map ϕ is well defined on a neighbourhood of the identity by the equation $e^X \mapsto e^{\phi_*X}$, because $e^X = e^Y$ holds if and only if $\log e^X e^{-Y} = 0$, and this means that

$\log e^{\phi_* X} e^{-\phi_* Y} = \phi_*(\log(e^X e^{-Y})) = 0$, so $e^{\phi_* X} = e^{\phi_* Y}$. What's more, ϕ is actually a local homomorphism, because the condition that $\phi(e^X)\phi(e^Y) = \phi(e^X e^Y)$ is the condition that $e^{\phi_* X} e^{\phi_* Y} = \exp(\phi_* \log(e^X e^Y))$, and by taking logarithms, this is exactly the condition implied by the preservation of the Lie bracket. We conclude that *all linear maps preserving the Lie bracket induce local homomorphisms on the Lie group*.

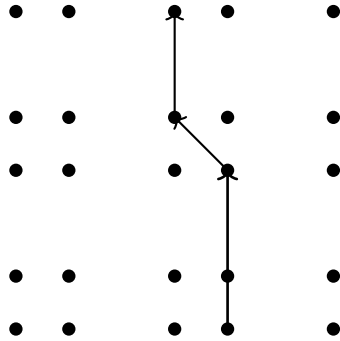
We know that any neighbourhood of the origin generates all elements of the group, so any $g \in G$ can be written as $h_1 h_2 \dots h_n$ for some h_i in the neighbourhood. If ϕ is a local homomorphism, then we can try extending it a global homomorphism by *defining* $\phi(g) = \phi(h_1)\phi(h_2)\dots\phi(h_n)$. Provided this is well defined, this really does extend ϕ to a global homomorphism, and the sufficient condition for this to be well defined is that our group is simply connected.

Theorem 1.1. *If G is a simply connected Lie group, U is a neighbourhood of the origin, and $\phi : U \rightarrow H$ is a local homomorphism, then ϕ can be extended to a global homomorphism.*

Proof. Since G is connected, given any $g \in G$, there is a curve $\lambda : [0, 1] \rightarrow G$ with $\lambda_0 = e$, and $\lambda_1 = g$. Since $[0, 1]$ is compact, there is N such that if $|t - u| < 2/N$, then $\lambda_t \lambda_u^{-1} \in U$, and we define

$$\phi(g) = \prod \phi(\lambda_{k/N} \lambda_{(k-1)/N}^{-1})$$

Provided that ϕ is well defined, ϕ is a homomorphism on all of G . To verify the map is well defined, let λ and γ be two curves from e to g . Since G is simply connected, there is a homotopy $H : [0, 1]^2 \rightarrow G$ with $H_0 = \lambda$ and $H_1 = \gamma$, and since $[0, 1]^2$ is compact, there is N such that if $|t_1 - u_1|, |t_2 - u_2| < 2/N$, then $H_{t_1 u_1} H_{t_2 u_2}^{-1} \in U$. Define a series of paths B^{kl} obtained from H by the curve in $[0, 1]^2$ specified by the diagram



such that it takes $1/N$ time to get between each vertex of the diagram. Then $B_{00} = H_0$, and $B_{N0} = H_1$. Let the definition of $\phi(g)$ relative to B_{kl} be denoted h_{kl} . We shall prove $h_{00} = h_{01}$, then $h_{01} = h_{02}$, and so on until we show $h_{0(N-1)} = h_{0N}$. Then we show $h_{0N} = h_{10}$, and progress from there. Thus in general, it suffices to show that $h_{kl} = h_{k(l+1)}$, and $h_{kN} = h_{(k+1)0}$, and all these essentially follow because the only place in the definition where the two differ is by swapping some $\phi(k)$ in the product formula with $\phi(k_0)\phi(k_1)$, and the equality of these terms follows because $k = k_0k_1$, and ϕ is a homomorphism in the neighbourhood U , and $k, k_0, k_1 \in U$ by the construction of the paths B_{kl} . Thus $\phi(g)$ is well defined. \square

Even though most Lie groups we encounter in practice are rarely simply connected, this theorem still finds significant use throughout the theory, because given any Lie group G , the simply connected covering space \tilde{G} has a canonical Lie group structure, and this simply connected group has the same Lie algebra as G . Provided we can determine the homomorphisms on \tilde{G} , and if $G = \tilde{G}/H$, then we can determine the homomorphisms on G as those homomorphisms on \tilde{G} whose kernel contains H . Thus we now have an effective technique to classify the Lie group homomorphisms, and in the rest of these notes, we work on classifying the different Lie algebra structures which can occur from a Lie group.

Chapter 2

Basic Definitions

Given our discussion in the previous chapter, the goal in this chapter is to isolate the algebraic features of the tangent space \mathfrak{g} which make it the Lie bracket of some Lie group. We can then focus completely on the algebraic problems associated with the definition, isolating the algebraic issues of the subject from the analytic issues. The abstract definition we use works over any field, which may seem frivolous, but it turns out that the functor from Lie groups to Lie algebras over the real numbers generalizes to many different categories of groups.

- Over the category of complex analytic groups, we can associate with each analytic group a Lie algebra over the complex numbers at the tangent space at the origin.
- In the theory of analytic groups over any complete field, we can associate a Lie algebra over that field.
- An algebraic group can be associated with a Lie algebra over the field in which the algebraic group is specified.

The algebraic structure \mathfrak{g} obtained from a Lie group G satisfies three important properties.

- $[\cdot, \cdot]$ is a bilinear operation on \mathfrak{g} .
- If λ is any curve, then $[\lambda_t, \lambda_t] = e$, so $f([\lambda_t, \lambda_t]) = f(e)$, and therefore for any $X \in \mathfrak{g}$, $[X, X] = 0$.

- Less obviously, if X, Y, Z are elements of the Lie algebra, then the Jacobi identity

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

which essentially holds because if g, h, k are elements of G , then

$$[[g, h^{-1}], z]^y [[h, k^{-1}], g]^k [[k, g^{-1}], h]^g = e$$

Though it seems abstract, we will soon see the identity plays a key role in the study of Lie algebras.

It turns out that this is sufficient to characterize the class of Lie algebras. 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 $X, Y, Z \in \mathfrak{g}$. There is a deep theorem of Ado showing that this is the ‘correct’ definition to make.

Fact 2.1. *Every Lie algebra over the real numbers is the Lie algebra of some Lie group. Every Lie algebra over the complex numbers is the Lie algebra of some complex analytic group.*

We shall find that the class of complex Lie groups has a rich structure theory, and this is the class we will begin to focus on as the theory develops. The main reason why the complex numbers is a nice field to work on in Lie algebra theory is that the field has characteristic zero, and is algebraically closed. Thus, though we use the complex numbers to obtain particular examples of Lie algebras, the classification theorem of complex Lie algebras generalizes quite easily to classify Lie algebras over an arbitrary algebraically closed field of characteristic zero. This doesn’t eliminate the study of Lie algebras over the real numbers, since we can always complexify these algebras to obtain a Lie algebra over the complex numbers.

Because of the deep entanglement between a Lie group and its Lie algebra (often physicists do not even distinguish the two structures, which makes interaction between mathematicians *slightly* distracting), we normally denote a Lie algebra corresponding to a Lie group by the ‘frakturized’ name of the group. Thus the Lie algebra of the general linear group $GL_n(K)$ is the general linear Lie algebra $\mathfrak{gl}_n(K)$, and the Lie algebra of an abstract Lie group G under study is denoted \mathfrak{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

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

This means that for any C^∞ function f defined on $GL_n(K)$,

$$f([I + tM, I + tN]) = f(I) + t^2 \nabla_{MN - NM}(f) + o(t^3)$$

hence the Lie bracket on $\mathfrak{gl}_n(K)$ takes the form $[M, N] = MN - NM$. In general, the set of invertible endomorphisms $GL(V)$ on a finite dimensional vector space V forms a Lie group, 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 an open submanifold of Euclidean space. 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 $ghg^{-1}h^{-1} = e$ for all $g, h \in G$, so the 2nd order approximation of this action is trivial, and this tells us that $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$. An arbitrary Lie algebra \mathfrak{g} is called **abelian** if the

equation $[X, Y] = 0$ holds for all $X, Y \in \mathfrak{g}$. If a connected Lie group G has an abelian Lie algebra, then G itself is an abelian group, because the Hausdorff-Campbell theorem then says that for any $X, Y \in \mathfrak{g}$,

$$e^Y e^X = e^{Y+X} = e^{X+Y} = e^X e^Y$$

Since the exponential map is a local diffeomorphism onto a neighbourhood of the origin, we find that G is locally abelian, hence globally abelian since these elements generate the group.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

because $\partial_X^m(X^m) = m! = 0$ 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 a finite dimensional associative algebra over K , then the space $M_n(A)$ of matrices with coefficients in A is an algebra over K . Thus the space $GL_n(A)$ of invertible matrices with coefficients in A forms a Lie group, and the induced Lie algebra $\mathfrak{gl}_n(A)$ is just the space of all matrices with coefficients in A , with the commutator $[X, Y] = XY - YX$. A particularly interesting example occurs if $A = K[X, X^{-1}]$, 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. More concretely, we can consider the Lie group $GL_n(\mathbf{H})$ of invertible quaternion valued matrices, which can be viewed as a subgroup of $GL_{4n}(\mathbf{R})$.

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 adj_X is actually a derivation on \mathfrak{g} . The map $X \mapsto \text{adj}_X$ is actually a homomorphism from \mathfrak{g} to $\text{Der}(\mathfrak{g})$, which we have seen before as the differential of the adjoint representation of groups, and algebraically, we see that

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

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} corresponding to the normal subgroup $Z(G)$ of G). Elements of $\text{Der}(\mathfrak{g})$ of the form 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^∞ 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^∞ functions, and we could have defined the Lie bracket operation as $[X, Y](f) = X(Yf) - Y(Xf)$, a ‘2nd derivative operation’.

Example. The estimate $\det(I + tM) = 1 + t \text{tr}(M) + o(t^2)$, which can be proved by analyzing the Leibnitz formula for the determinant 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 up to first order – i.e., those matrices M such that $\det(I + tM) = 1 + o(t^2)$. 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 that $\text{tr}(XY - YX) = \text{tr}(XY) - \text{tr}(YX) = 0$. Thus $\mathfrak{sl}_n(K)$ is formally a Lie subalgebra of $\mathfrak{gl}_n(K)$.

Example. Let Q be a bilinear form on K^n . Then the group G consisting of linear isomorphisms $M \in GL_n(K)$ preserving the bilinear form Q , in the sense that $Q(Mx, My) = Q(x, y)$ for all $x, y \in K^n$, forms a Lie group. If we find a matrix J such that $Q(x, y) = x^t J y$, then a matrix M is in the group G if $M^t J M = J$. To find the Lie algebra of G , we compute

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

so the Lie algebra \mathfrak{g} consists of matrices satisfying $X^t J + JX = 0$. The various classes of different bilinear forms give rise to numerous classes of different Lie groups.

- If Q is a symmetric, positive-definite bilinear form on \mathbf{R}^n , then the group G preserving Q is called the **orthogonal group** of Q . We can use the absolute value decomposition of positive definite matrices to write $J = N^t N$, and we find that the matrices M in the orthogonal group corresponding to Q satisfy $(NM)^t (NM) = N^t N$, which we can also rewrite as $(NMN^{-1})(NMN^{-1}) = I$, thus the map $M \mapsto NMN^{-1}$ is an isomorphism from the orthogonal group of Q to the standard orthogonal group associated with the Euclidean inner product $\langle x, y \rangle = \sum x_i y_i$, which we denote $O_n(\mathbf{R})$, whose Lie algebra elements X satisfy $X + X^t = 0$, the skew symmetric matrices. In general, every vector space K^n has a bilinear form $\sum x_i y_i$, and we define $O_n(K)$ to be the class of matrices preserving this form, whose Lie algebra is the class of skew symmetric $n \times n$ matrices.
- On \mathbf{C}^n , the matrices preserving a hermitian form Q are called the unitary group of Q . As with the orthogonal groups, any unitary group is conjugate isomorphic to the standard unitary group U_n of matrices preserving the canonical hermitian form $\sum \bar{x}_i y_i = x^* y$ on \mathbf{C}^n . These are the matrices

M with $M^*M = I$, and the Lie algebra \mathfrak{u}_n consists of matrices X with $X^* + X = 0$. The elements of the special unitary group SU_n consist of matrices with $X^*X = I$, and with determinant one, and the Lie algebra \mathfrak{su}_n consist of matrices with $M + M^* = 0$, and with trace zero.

- If n is even, $n = 2m$, then matrices $M \in M_n(K)$ represent operators from $K^m \times K^m$ to $K^m \times K^m$. A symplectic form on such a space is a skew-symmetric, non-degenerate bilinear form, and the symplectic group on the space is the set of matrices preserving this form. The canonical symplectic form is the for

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

and all the symplectic groups are conjugate the symplectic group $SP_n(K)$ over this form. To see this, given any non-degenerate bilinear form Q , we find a basis $\{y_1, \dots, y_m\}$ for K^m , then find a basis $\{x_1, \dots, x_m\}$ for K^m (using the non-degeneracy of Q) such that $Q(x_i, y_j) = \delta_{ij}$. Then $\{x_1, \dots, x_m\} \cup \{y_1, \dots, y_m\}$ forms a basis for $K^m \times K^m$, and in this basis Q takes the form of the canonical 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 $M^t J M = J$, and the symplectic Lie algebra $\mathfrak{sp}_n(K)$ consists of matrices with $M^t J + J M = 0$. If we write

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

Then

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

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

Together with the special linear group, the orthogonal and symplectic groups over the complex numbers will essentially classify all the important Lie algebras over the complex numbers.

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.

Example. The Euclidean group E_n is the space of isometries from \mathbf{R}^n to itself. Every isometry can be described as a composition of a translation and an orthogonal transformation. Thus E_n is isomorphic to the semidirect product $\mathbf{R}^n \ltimes O_n(\mathbf{R})$, where $O_n(\mathbf{R})$ acts as an automorphism of \mathbf{R}^n as it does canonically as an application of the operator. E_n is isomorphic to group of matrices

$$\left\{ \begin{pmatrix} 1 & 0 \\ x & M \end{pmatrix} : x \in \mathbf{R}^n, M \in O_n(\mathbf{R}) \right\}$$

where the matrix acts on \mathbf{R}^n via the embedding in \mathbf{RP}^n . The special Euclidean group SE_n is the space of orientation preserving isometries from \mathbf{R}^n to itself, and is isomorphic to

$$\left\{ \begin{pmatrix} 1 & 0 \\ x & M \end{pmatrix} : x \in \mathbf{R}^n, M \in SO_n(\mathbf{R}) \right\}$$

so these groups are essentially a semidirect product of the additive structure on \mathbf{R}^n and $O_n(\mathbf{R})$, or $SO_n(\mathbf{R})$. The matrix representation lets us see the Lie bracket structure easily. The Lie algebra structure consists of matrices of the form

$$\mathfrak{e}_n = \mathfrak{se}_n = \left\{ \begin{pmatrix} 0 & 0 \\ x & M \end{pmatrix} : M + M^t = 0 \right\}$$

and the Lie bracket is

$$\left[\begin{pmatrix} 0 & 0 \\ x & M \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ y & N \end{pmatrix} \right] = \begin{pmatrix} 0 & 0 \\ My - Nx & MN - NM \end{pmatrix}$$

In general, G and H are Lie groups, and the semidirect product $G \ltimes H$ is formed by some differentiable automorphism representation $\rho : H \rightarrow \text{Aut}(G)$, then we obtain a Lie bracket representation ρ_* mapping elements of \mathfrak{h} to Lie bracket isomorphisms from \mathfrak{g} to itself, and the Lie algebra $\mathfrak{g} \ltimes \mathfrak{h}$ is then formed over the vector space $\mathfrak{g} \oplus \mathfrak{h}$, with the Lie bracket operation

$$[X_0 + Y_0, X_1 + Y_1] = [\rho_*(Y_0)(X_1) - \rho_*(Y_1)(X_0)] + [Y_0, Y_1]$$

which is the ‘semidirect product’ of Lie groups.

2.1 Simply Connected Lie Groups

Because of the connection between representations of Lie groups and Lie algebras, it is useful to know the simply connected cover of any Lie group. The topological structure of the low dimensional classical groups is easy to see in low dimensions, but deriving the fundamental group of more advanced Lie groups requires some advanced theorems of algebraic topology.

Example. $O(2)$ is a one dimensional Lie group, isomorphic to the semidirect product $\mathbf{T} \ltimes \{-1, +1\}$, induced by the embedding $\{-1, +1\} \rightarrow \text{Aut}(\mathbf{T})$ which maps $+1$ to the identity automorphism, and maps -1 to the conjugation automorphism $z \mapsto \bar{z}$. The way that $\mathbf{T} \ltimes \{-1, +1\}$ acts orthogonally on \mathbf{R}^2 , which we think of as \mathbf{C} , is given by complex multiplication. We have $(z, +1)w = zw$, and $(z, -1)w = z\bar{w}$. $O(2)$ is not even connected, but $SO(2)$ is, which is isomorphic to \mathbf{T} when the action is given by complex multiplication, and we see that $\pi_1(SO(2)) = \mathbf{Z}$. Because $SO(2)$ is isomorphic to \mathbf{T} , the simply connected group covering $SO(2)$ is just \mathbf{R} , where $t \mapsto e^{it}$.

Example. Any orthogonal transformation T on $O_3(\mathbf{R})$ fixes some line through the origin, because the orthogonal transformation restricts to a continuous map on S^2 , and any continuous map on S^2 has a fixed point. If $Tx = x$, extend x to an orthogonal basis (x, y, z) . In this basis, T must have a matrix representation

$$\begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix}$$

for some $M \in O_2(\mathbf{R})$. If T is an oriented orthogonal transformation in $SO_3(\mathbf{R})$, then we find that $M \in SO_2(\mathbf{R}) \cong \mathbf{T}$. Thus to obtain any transformation in $SO_3(\mathbf{R})$, we may first 'fix an axis of rotation' by picking $x \in S^2$, and then apply the rotation of M to x^\perp (which acts in a way which is independent of any orthogonal, oriented basis extending x , because the special orthogonal group in two dimensions is commutative), and we denote this transformation $T_{x,M}$. Let B be the ball of radius π , and consider the projection $\pi : B \rightarrow SO_3(\mathbf{R})$, defined by $\pi(x) = T_{\hat{x}, e^{i\|x\|}}$ then the map is injective, except that on the boundary $\pi(x) = \pi(-x)$. Since the map is continuous and open, we find that $SO_3(\mathbf{R})$ has a topology obtained by identifying points on opposite sides of B , which is just the three dimensional projection plane \mathbf{RP}^3 , which therefore has fundamental group \mathbf{Z}_2 .

Example. On the other hand, the unitary group U_2 has fundamental group \mathbf{Z} . First, note that the space has the same topology as the subspace of $\mathbf{C}^2 \times \mathbf{C}^2$ consisting of (x, y) , where x and y are orthonormal. There is a surjective map $f : S_{\mathbf{C}^2} \times \mathbf{T} \rightarrow U_2$ with $f(x, z) = (x, zx^\perp)$, where if $x = (x_1, x_2)$, $x^\perp = (\bar{x}_2, -\bar{x}_1)$. The map $x \mapsto x^\perp$ is a continuous map on \mathbf{C}^2 , as is the map which takes orthonormal (x, y) and returns z such that $y = zx^\perp$, because on the open set of \mathbf{C}^2 where $x_2 \neq 0$, $z = y_1 \bar{x}_2^{-1}$, and on the open set of \mathbf{C}^2 where $x_1 \neq 0$, $z = y_2 \bar{x}_1^{-1}$. From these relations, we also see that the map $f(x, z) = (x, zx^\perp)$ is differentiable, and since SU_2 and $S_{\mathbf{C}^2} \times \mathbf{T}$ are both 4 dimensional real manifolds, this map must also be a diffeomorphism. But $S_{\mathbf{C}^2}$ is diffeomorphic to S^3 , so SU_2 is homeomorphic to $S^3 \times \mathbf{T}$. When we take the special unitary group SU_2 , the torus disappears, we find SU_2 is homeomorphic to S^3 , and therefore the group is simply connected.

Example. We calculate that SU_2 is the simply connected covering group of $SO_3(\mathbf{R})$. Let V denote the space of 2×2 complex self-adjoint matrices. Then V is three dimensional over the real numbers, and has a real-valued inner product

$$\begin{aligned} \langle M, N \rangle &= (1/2) \text{tr}(MN) = (1/2) \sum M_{ij} N_{ji} \\ &= (1/2) \sum M_{ii} N_{ii} + \Re \left(\sum_{i < j} M_{ij} \bar{N}_{ij} \right) \end{aligned}$$

Then V has an orthonormal basis

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & +i \\ -i & 0 \end{pmatrix} \quad C = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}$$

If M is self adjoint, and $N \in SU_2$, then $(NMN^{-1})^* = NMN^{-1}$, so N acts on V as a linear operator, and for any self-adjoint M_0, M_1 ,

$$\langle NM_0N^{-1}, NM_1N^{-1} \rangle = \frac{1}{2} \text{tr}(NM_0M_1N^{-1}) = \frac{1}{2} \text{tr}(M_0M_1) = \langle M_0, M_1 \rangle$$

So the action of SU_2 on V is a representation of SU_2 on $O_3(\mathbf{R})$. Since SU_2 is connected, it is actually a homomorphism into $SO_3(\mathbf{R})$. We find that the induced Lie algebra homomorphism from \mathfrak{su}_2 to $\mathfrak{o}_3(\mathbf{R})$ is just the adjoint representation of \mathfrak{su}_2 on V , and since \mathfrak{su}_2 is spanned by $X = i(E_{11} - E_{22})$, $Y = E_{12} - E_{21}$, and $Z = iE_{12} + iE_{21}$, then we can calculate that

$$\begin{array}{lll} [X, A] = 2B & [X, B] = -2A & [X, C] = 0 \\ [Y, A] = 2C & [Y, B] = 0 & [Y, C] = -2A \\ [Z, A] = 0 & [Z, B] = 2C & [Z, C] = -2B \end{array}$$

Thus the adjoint representation of \mathfrak{su}_2 on $\mathfrak{o}_3(\mathbf{R})$ is injective, hence surjective, and therefore the adjoint representation of SU_2 on V is surjective onto $SO_3(\mathbf{R})$. The map is two to one, because if $NM = MN$ for all $M \in V$, then in particular we see that if $N = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$\begin{pmatrix} b & a \\ d & c \end{pmatrix} = NA = AN = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$$

$$\begin{pmatrix} -ib & +ia \\ -id & +ic \end{pmatrix} = NB = BN = \begin{pmatrix} ic & id \\ -ia & -ib \end{pmatrix}$$

so $b = c$, $a = d$, and $c = -b$, from which we conclude $b = c = 0$. But then $N = z$ for some $z \in \mathbf{C}$, and since the determinant of N is one, we conclude that $z = \pm 1$, so the kernel of the homomorphism from SU_2 to $SO_3(\mathbf{R})$ is $\{\pm 1\}$, and SU_2 is a two to one cover of $SO_3(\mathbf{R})$, which explains why the fundamental group of $SO_3(\mathbf{R})$ is \mathbf{Z}_2 .

Given an arbitrary Lie group G , we can form families of nonisomorphic Lie groups with the same Lie algebra by taking quotients by discrete, normal subgroups H of G . The number of different Lie groups we can form from this process can actually be characterized by the set of all discrete subgroups of $Z(G)$.

Theorem 2.1. *The only normal, discrete subgroups of a connected Lie group are subgroups of the center of the group.*

Proof. Let G be a connected Lie group, and let H be a discrete, normal subgroup. For a fixed $h \in H$, the map $f : G \rightarrow H$ given by $f(g) = ghg^{-1}$ is differentiable, and $f(e) = h$, so by continuity, and the fact that H is discrete, we conclude that $f(g) = h$ for all $g \in G$, so $ghg^{-1} = h$ for all $g \in G$, or $gh = hg$, so $h \in Z(G)$. \square

If $Z(G)$ is a discrete subgroup of G , then all possible quotients G/H is formed by taking any particular subgroup of $Z(G)$. What's more, $G/Z(G)$ is essentially the 'maximal quotient' we can form from the group G .

Theorem 2.2. *If $Z(G)$ is discrete, then $G/Z(G)$ has trivial center.*

Proof. If $gZ(G)$ is a coset in the center of $G/Z(G)$, this means exactly that for any $h \in G$, $ghg^{-1}h^{-1} \in Z(G)$. The map $f(h) = ghg^{-1}h^{-1}$ is therefore a continuous map of G into $Z(G)$, and $f(e) = e$, hence since $Z(G)$ is discrete we conclude that $ghg^{-1}h^{-1} = e$ for all $h \in G$, from which we conclude that $g \in Z(G)$, so $gZ(G)$ is trivial. \square

We say a Lie group homomorphism $f : G \rightarrow H$ is an **isogeny** if it lifts to a homomorphism of the universal covering groups of G and H . The category of isogenies of Lie groups breaks down into certain chains of groups, where the initial object of a chain is the simply connected covering group of any object in the chain, and provided that $Z(G)$ is a discrete subset of G , the final objects of the chain are just $G/Z(G)$. We call the simply connected group corresponding to the group G the **simply connected form** of the group, and $G/Z(G)$ the **adjoint form** of the group.

Example. *The center of $SL_n(K)$ consists of the identity multiplied by any element $x \in K$ with $x^n = 1$. Thus for odd n , $SL_n(\mathbf{R})$ has trivial center, and for even n , $SL_n(\mathbf{R})$ has center $\{\pm 1\}$, hence we may take the quotient to obtain the adjoint form $PSL_n(\mathbf{R})$ of projective matrices. This group may also be described as $GL_n(\mathbf{R})$ modulo the set of real-valued multiples of the identity matrix, and we find $SL_n(\mathbf{R})$ is equal to $GL_n(\mathbf{R})$ modulo the group of positive real-valued multiples of the identity. Similarly, the center of $SL_n(\mathbf{C})$ consists of the n th roots of unity, and we can form the projective matrix group $PSL_n(\mathbf{C})$ by taking the quotient, which is isomorphic to $GL_n(\mathbf{C})$ modulo the group of complex multiples of the identity. We can also consider the projective orthogonal matrices $PSO_n(K)$ where n is odd, and the projective symplectic matrices $PSP_n(K)$.*

Working the opposite way, it is a little bit trickier to construct the simply connected groups of most Lie groups. The simply connected cover of $O_n(K)$ is called the n th spin group, and is denoted $\text{Spin}_n(K)$. These groups arise in certain aspects of quantum mechanics and quantum field theory, where their representations give rise to the proof of existence of the fundamental particles of physics, the bosons and fermions.

2.2 Factorization

We have now motivated our interest in characterizing the structure of all Lie algebras, because many questions about the structure of Lie groups can be reduced to questions about Lie algebras. The main strategy we will find in classifying the Lie algebras up to isomorphism, like in most algebraic categories, is to consider subalgebras and quotients, which break apart the algebra into more understandable chunks. 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 of homomorphisms of the form

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

which are infinitesimal formulations, of course, of an exact sequence in the category of Lie groups

$$0 \rightarrow K \rightarrow G \rightarrow H \rightarrow 0$$

(indeed, though we are unable to prove it here, there is a one to one correspondence of Lie subalgebras of a Lie algebra and connected Lie subgroups of its Lie group, where the correspondence is obtained by differentiating the imbedding map of the group). In the Lie algebra case, 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.

Now 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 $g \in G, h \in H$. 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. We also obtain from this fact that $GL_n(K)$ is isomorphic to the direct product $K \times SL_n(K)$ by the map $(x, M) = xM$, which can be verified by a direct, algebraic calculation.

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 $\text{adj } \mathfrak{g}$.

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 by consider their algebraic relations modulo the inner derivations.

As in most of algebra, the process of constructing ideals and constructing homomorphisms are dual to one another, a fact implied by the first isomorphism theorem. In fact, we have a version of all standard isomorphism theorems in the category of Lie algebras.

Theorem 2.3 (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 the diagram below commutes.*

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

Theorem 2.4 (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 2.5 (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 $\mathfrak{b}/(\mathfrak{a} \cap \mathfrak{b})$.*

2.3 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 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}_{ab} . Since we have the exact diagram

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

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

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

for some bilinear A and B . Thus we can think of the elements of \mathfrak{g}' as infinitesimals, since they have no impact on the Lie bracket structure of \mathfrak{g}_{ab} , existing somewhat ‘beneath the surface’ of the calculations. If we consider $D^2\mathfrak{g} = \mathfrak{g}''$, 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. However, the ‘new infinitesimals’ 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{o}_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.

2.4 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 $f^k : \mathfrak{g}^2 \rightarrow K$ is a bilinear form, and so we may find constants $a_{ij}^k \in K$ such that

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

which is the same as saying $[E_i, E_j] = \sum a_{ij}^k E_k$. If $[\cdot, \cdot]$ was actually a Lie bracket, then the alternating property of the Lie algebras implies that each f_k is a skew-symmetric Bilinear form, hence $a_{ij}^k = -a_{ji}^k$. The Jacobi identity takes the form

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

for every fixed i, j, k, m . Conversely, given a set of constants a_{ij}^k over some vector space satisfying these equations, the corresponding bilinear map is a Lie bracket, because if $X = \sum a^i E_i$, $Y = \sum b^j E_j$, and $Z = \sum c^l E_l$, 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 algebra structures over a vector space and structural constants over a fixed basis 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. What's more, as we turn up the dimension, we shall begin to see certain techniques which we will advance to apply in the more general context. For fun, we shall also use our classification to classify the low dimensional Lie groups, and the low dimensional analytic groups over the complex numbers.

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

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

and so

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

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

The only zero dimensional Lie algebra is the trivial Lie algebra. There is only a single connected Lie/analytic group which has this Lie algebra

as its tangent space – the trivial group. Every one dimensional Lie algebra is abelian, because if a Lie algebra is spanned by a vector X , then $[aX, bX] = ab[X, X] = 0$. Thus there is a unique one dimensional Lie algebra up to isomorphism. The simply connected Lie group corresponding to the one dimensional Lie algebra is the real number line \mathbf{R} under addition, and all other groups with this Lie algebra are obtained from discrete quotients of \mathbf{R} . Any such quotient is isomorphic to the toral group \mathbf{T} of complex numbers of norm one under multiplication. The analytic groups look slightly different. The simply connected group is the complex plane \mathbf{C} under addition, and the discrete quotients are formed by a discrete subgroup Λ . If $\Lambda = w\mathbf{Z}$ has rank one, then we have an isomorphism between \mathbf{C}/Λ and the group \mathbf{C}^* of non-zero complex numbers under multiplication, given by $z + w\mathbf{Z} \mapsto e^{2\pi iz/w}$. If $\Lambda = z\mathbf{Z} \oplus w\mathbf{Z}$ has rank two, then \mathbf{C}/Λ is isomorphic to $\mathbf{C}/(\mathbf{Z} \oplus (w/z)\mathbf{Z})$, but as w/z varies over the upper half complex plane, and at the infinity point, we obtain a continuous family of analytic groups which are not all isomorphic to each other, but instead have isomorphism classes characterized by the j invariant

$$j(q) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + \dots$$

where $q = e^{2\pi iz}$, defined for z in the upper half plane, and for $z = \infty$.

In two dimensions, we will show there are exactly two isomorphism classes of Lie algebras – one abelian Lie algebra, and one nonabelian Lie algebra. The simply connected group corresponding to the abelian Lie algebra in one dimension is \mathbf{R}^2 , whose quotient groups are either isomorphic to $\mathbf{T} \times \mathbf{R}^2$, or to \mathbf{T}^2 . The two dimensional analytic groups with abelian Lie algebra are very difficult to classify. The rank one quotients are isomorphic to $\mathbf{C} \times \mathbf{C}^*$, whereas the rank two quotients are either isomorphic to $\mathbf{C}^* \times \mathbf{C}^*$ or one of the family of analytic groups $\mathbf{C} \times \mathbf{C}/(\mathbf{Z} \oplus w\mathbf{Z})$, depending on whether the generators for the rank two subgroup lie in a complex line or not. The rank three and four cases are very mysterious, and there is no satisfactory classification theory for these analytic groups.

There is one other two dimensional Lie algebra, which is unique, and classifies the remaining Lie and analytic groups.

Theorem 2.6. *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

$$[a_1X + a_2Y, [b_1X + b_2Y, c_1X + c_2Y]] = (a_2b_2c_1 - a_2b_1c_2)X$$

And $a_2b_2c_1$ is obtained from $a_2b_1c_2$ by a cycle permutation in the variables a, b , and c , hence the Jacobi identity holds. \square

One way to obtain a Lie group corresponding to the unique nonabelian two dimension real Lie algebra is obtained from exponentiating the adjoint matrix representations of the algebra. Since the basis elements have the matrix representation

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

we can exponentiate the set of matrices of the form $\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$ and we find this is the Lie group

$$SE_1 = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a > 0 \right\}$$

which can be seen as the class of orientation preserving Euclidean transformations on the real line, where the matrix with coefficients a and b corresponds to the transformation $x \mapsto ax + b$. The group is simply connected, and the center is trivial, so SE_1 is the only real two dimensional Lie group with a nonabelian two dimensional Lie algebra. In the category of analytic manifolds, the group of matrices we obtain can be viewed as the group of all transformations and dilations of \mathbf{C} of the form $x \mapsto ax + b$, where $a \neq 0$. This can be seen as the group of complex matrices of the form

$$\left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a \neq 0 \right\}$$

This group is isomorphic to $\mathbf{C}^* \ltimes \mathbf{C}$ where $(a, b)(c, d) = (ac, ad + b)$. It is not simply connected, having fundamental group \mathbf{Z} , but the simply connected

cover is $\mathbf{C} \ltimes \mathbf{C}$, where $(a, b)(c, d) = (a + c, e^{ia}d + b)$. The center of this group is $2\pi\mathbf{Z} \times \{0\}$, and the integer subgroups correspond to a family of nonisomorphic quotient groups, whose adjoint form is the group of translations and dilations of the plane.

Theorem 2.7. *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 independant, 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 $Z = E_{13}$. \square

The Heisenberg Lie algebra obviously has a group representation as the Heisenberg group

$$H = \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

This group is simply connected, and its center consists of the matrices of the form

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

which is isomorphic to \mathbf{R} or \mathbf{C} . Over the real Lie groups, all discrete quotients of this algebra are isomorphic to a version of the Heisenberg group where we let b range over \mathbf{T} rather than \mathbf{R} , and where the corresponding operation is

$$\begin{pmatrix} 1 & a_0 & b_0 \\ 0 & 1 & c_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & a_1 & b_1 \\ 0 & 1 & c_1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a_0 + a_1 & b_0 b_1 e^{ia_0 c_1} \\ 0 & 1 & c_0 + c_1 \\ 0 & 0 & 1 \end{pmatrix}$$

If we view the Heisenberg group as acting on $\mathbf{R}^2 \subset \mathbf{RP}^2$, then we obtain the set of transformations of the form

$$T : (x, y) \mapsto \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b \\ c \end{pmatrix}$$

which is just a skew transform in the x axis combined with a translation. The maps preserve the distance in the y axis between two points, and thus the Heisenberg group acts on the quotient space $\mathbf{R}^2/[\{0\} \times \mathbf{Z}]$. The elements of the Heisenberg group that act trivially on this quotient space are exactly the matrices

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

where b is an integer. But this implies that the quotient of the Heisenberg group we obtain can be viewed as the set of skews on $\mathbf{R}^2/[\{0\} \times \mathbf{Z}]$. Similarly, over the complex numbers, we can take either rank one or rank two abelian subgroups of matrices whose coefficients all vanish except for b , and we find that the rank one quotient correspond to letting b be an element of \mathbf{C}^* , or we can let b range over some quotient of a lattice over the complex plane.

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

Proof. Let \mathfrak{h} be such an algebra. Pick $X \neq 0$ spanning the derived \mathfrak{h}' , then pick Y with $[X, Y] = X$. The subalgebra of \mathfrak{h} spanned by X and Y is isomorphic to \mathfrak{g} . To obtain a full isomorphism with $\mathfrak{g} \oplus K$, it suffices to find

$Z \in Z(\mathfrak{g})$ which is linearly independent from X and Y . Pick some Z to form a basis $\{X, Y, Z\}$. There must be $a, b \in K$ such that $[X, Z] = aX$, $[Y, Z] = bX$. If we consider the equations

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

And since the matrix

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

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

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

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

The Lie algebra $\mathfrak{g} \oplus \mathbf{R}$ has a simply connected Lie group $SE_1 \times \mathbf{R}$, and subgroups correspond to the Lie group $SE_1 \times \mathbf{T}$. The Lie algebra $\mathfrak{g} \oplus \mathbf{C}$ has a simply connected Lie group which is the direct product of the complex numbers with the group of translations and dilations of the complex plane, and we obtain all families of these Lie groups by taking discrete quotients in each set.

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, because it implies that we can always find enough eigenvectors to restrict the structure of certain actions of Lie algebras on spaces. 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 Lie algebra over E by extending the Lie algebra operation to be bilinear in elements of E . The most important

application of this tool is to take a Lie algebra \mathfrak{g} over the real numbers, and to then consider its ‘complexification’ $\mathbf{C} \otimes_{\mathbf{R}} \mathfrak{g}$, often denoted $\mathfrak{g}_{\mathbf{C}}$, which will be a Lie algebra whose elements consist of $X + iY$, where $X, Y \in \mathfrak{g}$. If \mathfrak{h} is a complex Lie algebra isomorphic to $\mathbf{C} \otimes_{\mathbf{R}} \mathfrak{g}$ for some real Lie algebra \mathfrak{g} , then we say \mathfrak{g} is a **real form** of \mathfrak{h} .

Example. The complexification of the real Lie algebras $\mathfrak{gl}_n(\mathbf{R})$ and \mathfrak{u}_n are isomorphic to $\mathfrak{gl}_n(\mathbf{C})$. The complexification of \mathfrak{su}_n and $\mathfrak{sl}_n(\mathbf{R})$ are isomorphic to $\mathfrak{sl}_n(\mathbf{C})$. The complexification of $\mathfrak{so}_n(\mathbf{R})$ is isomorphic to $\mathfrak{so}_n(\mathbf{C})$. The complexification of $\mathfrak{sp}_n(\mathbf{R})$ is isomorphic to $\mathfrak{sp}_n(\mathbf{C})$.

In order to understand the representations of a Lie group G , it is often easier to switch to studying representations of the Lie algebra $\mathfrak{g}_{\mathbf{C}}$, and then to attempt to recover the representations of G from these representations. Indeed, if $\rho : G \rightarrow GL(V)$ is a representation, then $\rho_* : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ is a representation of Lie algebras, and we can extend this map to a complexification map $\rho_* : \mathfrak{g}_{\mathbf{C}} \rightarrow \mathfrak{gl}(V_{\mathbf{C}})$. At least with respect to the complex representations of G , this process can fully recover all representations, in which case V is already a complex vector space, and $V_{\mathbf{C}}$ is just V .

Returning to our classification, there is a large family of non-isomorphic Lie algebras whose derived algebra is two-dimensional. We take a basis $\{Y, Z\}$ of the derived subalgebra, and complete it to a basis $\{X, Y, Z\}$ of the entire algebra. We claim that the derived subalgebra is always abelian. To prove this, it suffices to show $[Y, Z] = 0$. Consider the adjoint map adj_Y , which can be written in matrix form with respect to the basis given as

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

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

Lemma 2.9. 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. □

In fact, adj_X is an isomorphism of \mathfrak{g}' , because $[X, Y]$ and $[X, Z]$ span the derived subalgebra, hence the map is surjective, and therefore an isomorphism. Over an algebraically complete field, we may apply the theory of the Jordan normal form, and we 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 γ , this does indeed define a Lie algebra structure on K^3 , denoted \mathfrak{g}_γ , because

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

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

First note that the map $X \mapsto \lambda X$, $Y \mapsto Z$, $Z \mapsto Y$ is an explicit isomorphism between \mathfrak{g}_λ and \mathfrak{g}_γ 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 , the map f takes the form

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

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

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

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

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 λ , and we may extend X to a basis $\{X, Y, Z\}$ such that $[X, Y] = \lambda Y + Z$, $[X, Z] = \lambda Z$. If we replace X with X/λ , and Z with Z/λ , then $[X, Y] = Y + Z$, and $[X, Z] = Z$, and these define a specific set of structural constants, which actually give a Lie algebra structure. Indeed, then

$$\begin{aligned}
& [a_1 X + a_2 Y + a_3 Z, [b_1 X + b_2 Y + b_3 Z, c_1 X + c_2 Y + c_3 Z]] \\
&= [a_1 X + a_2 Y + a_3 Z, (b_1 c_2 - b_2 c_1)Y + (b_1 c_3 - b_3 c_1 + b_1 c_2 - b_2 c_1)Z] \\
&= (a_1 b_1 c_2 - a_1 b_2 c_1)(Y + Z) + (a_1 b_1 c_3 - a_1 b_3 c_1 + a_1 b_1 c_2 - a_1 b_2 c_1)Z \\
&= (a_1 b_1 c_2 - a_1 b_2 c_1)Y + (2a_1 b_1 c_2 - 2a_1 b_2 c_1 + a_1 b_1 c_3 - a_1 b_3 c_1)Z
\end{aligned}$$

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

To find the analytic groups G_γ with Lie algebra \mathfrak{g}_γ , we exponentiate the adjoint representations. We have

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

Hence a general element of the Lie algebra will be of the form

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

and exponentiating this elements, we obtain an open subset of matrices around the origin of a Lie subgroup of $GL_3(\mathbb{C})$ whose Lie algebra is \mathfrak{g}_γ . This is exactly the set of matrices

$$G_\gamma = \begin{pmatrix} 1 & 0 & 0 \\ -b[(e^a - 1)/a] & e^a & 0 \\ -c[(e^{\gamma a} - 1)/a] & 0 & e^{\gamma a} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ y & e^x & 0 \\ z & 0 & e^{\gamma x} \end{pmatrix}$$

for any $x, y, z \in \mathbb{C}$, and these matrices do form a Lie subgroup of $GL_3(\mathbb{C})$. If γ is not a rational number, then we find that the exponential map is a one-to-one map from \mathfrak{g}_γ to this set of matrices, and the map is actually a

holomorphic embedding of \mathfrak{g}_γ into this embedding, so the induced analytic group G_γ is simply connected. If γ is not rational, then this analytic group has trivial center, and so the analytic group G_γ is the unique analytic group with this Lie algebra. On the other hand, if γ is a rational number, then G_γ is not simply connected, and we have an exact sequence

$$0 \rightarrow \left\{ \begin{pmatrix} 1 & 0 & 0 \\ y & 1 & 0 \\ z & 0 & 1 \end{pmatrix} \right\} \rightarrow G_\gamma \rightarrow \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^x & 0 \\ 0 & 0 & e^{\gamma x} \end{pmatrix} \right\} \rightarrow 0$$

Hence by the first isomorphism theorem G_γ is isomorphic to the direct product of the left hand side with the quotient group of the right hand. The left hand side is always simply connected, whereas when γ is a rational number the group on the right hand side is isomorphic to \mathbf{C}^* , and has fundamental group \mathbf{Z} . This implies that G_γ is isomorphic to the semidirect product of the two groups, and therefore also has fundamental group \mathbf{Z} . The connected cover is the group

$$\begin{pmatrix} 1 & 0 & 0 \\ y & e^x & 0 \\ z & 0 & e^{\gamma x} \end{pmatrix}$$

where we now view two matrices distinct if they are defined by different x . That is, the group is really \mathbf{C}^3 , where the group operation is

$$(x_0, y_0, z_0)(x_1, y_1, z_1) = (x_0 + x_1, y_0 + e^{2\pi i x_0} y_1, z_0 + e^{2\pi i \gamma x_0} z_1)$$

If $\gamma = m/n$, where n does not divide m , then the center of this group is $(x, 0, 0)$, where $x \in n\mathbf{Z}$. The quotients of the subgroups of the center are all nonisomorphic, and are all the groups with \mathfrak{g}_γ as a Lie algebra.

For real three dimensional Lie algebras, the proof above has a third possibility. Suppose that the eigenvalues of adj_X consist of zero, and two complex eigenvalues z and \bar{z} .

Though we can't necessarily classify all real three dimensional Lie algebras using this method, the Lie algebra constructed in the proofs above can be taken with real scalars, so we can at least classify the Lie groups with this algebra, and at least we can classify the Lie algebras such that adj_X can be put into Jordan normal form like in the proof. The exponential map above tells us that the Lie groups with Lie algebra \mathfrak{g}_λ , for $\lambda \in \mathbf{R}$

consist of matrices of the form

$$\left\{ \begin{pmatrix} 1 & 0 & 0 \\ y & x & 0 \\ z & 0 & x^\gamma \end{pmatrix} : x > 0, y, z \in \mathbf{R} \right\}$$

In this case this group is always simply connected, and has trivial center so it is the unique Lie group with this Lie algebra.

On the other hand, if our operators are not diagonalizable, then the Lie algebra has an adjoint form

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

So the adjoint form of a general matrix takes the form

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

and by exponentiating, we find the analytic group consisting of the matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ -b(e^a - 1)a & e^a & 0 \\ -(1/a)(be^a + c(e^a - 1)) & ae^a & e^a \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ y & x & 0 \\ z & x \log x & x \end{pmatrix}$$

The exponential map is injective, and so the group is simply connected. The group has trivial center, so it is the unique analytic group with the properties required. The same is true of the real Lie group corresponding to this Lie algebra.

For the final part of our classification, 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 $\neq 2$, because $\mathfrak{sl}_2(K)$ satisfies this exactly. We shall find this is the only such algebra up to isomorphism.

Theorem 2.10. *If \mathfrak{g} is a three dimensional Lie algebra over an algebraically closed field of characteristic zero with $\mathfrak{g}' = \mathfrak{g}$, then \mathfrak{g} is isomorphic to $\mathfrak{sl}_2(K)$.*

Proof. Given such an algebra, if $X \in \mathfrak{g}$ is non-zero, then adj_X has rank 2, because \mathfrak{g}' is spanned by $[X, Y]$, $[X, Z]$ and $[Y, Z]$, which must be linearly

independent, and $[X, Y]$ and $[X, Z]$ span the range of adj_X . This also implies that $[X, M] = 0$, then M is in the span of X . We claim that we can choose X such that adj_X has a non-zero eigenvalue. If all the eigenvalues of X are non-zero, applying the Jordan canonical form theorem we conclude that there is a basis X_0, Y_0 , and Z_0 such that $[X, X_0] = 0$, $[X, Y_0] = X_0$, and $[X, Z_0] = Y_0$. By rescaling X , we may actually assume $X = X_0$, in which case $[Y_0, X_0] = -[X_0, Y_0] = -X_0$, so that adj_{Y_0} has a non-zero eigenvalue.

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

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

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

The group $SL_2(\mathbb{C})$ is simply connected (this requires a little bit of advanced algebraic topology, so we take it as a fact), and has kernel $\{\pm I\}$, so $PSL_2(\mathbb{C})$ is the only other analytic group with this fundamental group. This is a group which is very fundamental to the study of modular forms, since $PSL_2(\mathbb{C})$ acts as the set of biholomorphisms of the upper half complex plane.

Chapter 3

Matrix Lie Algebras

3.1 Basic Representation Theory

Most algebraic concepts about Lie groups have natural ‘infinitesimal formulations’ on Lie algebras, which are easier to understand because bilinear operations have a nicer structure theory than diffeomorphisms do. For instance,

- Every Lie subgroup of a Lie group corresponds to a unique Lie subalgebra of the Lie group, and conversely, in a connected Lie group, a Lie subalgebra corresponds to a unique (immersed) Lie subgroup, so the two classes of objects are in one to one correspondence.

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

The study of representation theory is fundamental to understanding Lie algebras. Indeed, we find that we can define the Lie bracket of a Lie group by passing through certain representations of the Lie group. Indeed, this foreshadows the fact that we will classify all of the complex semisimple Lie algebras as special cases of their 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 Adj is continuous, it is also smooth, hence we may differentiate Adj to obtain a map $\text{adj} : \mathfrak{g} \rightarrow \mathfrak{gl}(\mathfrak{g})$, where

$$\text{adj}(X)(Y) = \frac{d\text{Adj}(e^{tX})(Y)}{dt} = \frac{d(c_{e^{tX}})_*(Y)}{dt} = \frac{d^2 e^{tX} e^{tY} e^{-tX}}{dt^2}$$

and using the Baker-Hausdorff formula, we find that for small enough t ,

$$\log(e^{tX} e^{tY}) = tX + tY + (t^2/2)[X, Y] + \sum t^{|\alpha|} a_\alpha(X, Y)^\alpha$$

and hence

$$\begin{aligned} \log(e^{tX} e^{tY} e^{-tX}) &= tX + tY + (t^2/2)[X, Y] - tX + (t^2/2)[X, Y] + o(t^3) \\ &= tY + t^2[X, Y] + o(t^3) \end{aligned}$$

Thus, differentiating this equation twice, we find that

$$\frac{d^2 e^{tX} e^{tY} e^{-tX}}{dt^2} = 2[X, Y]$$

Thus the adjoint map is really the Lie bracket in disguise, in the sense that $\text{adj}(X)(Y) = [X, Y]$. We know that if $\rho : G \rightarrow GL(V)$ is any group representation, then $\rho_* : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ will be a Lie algebra representation, as a general case of the theorem for group homomorphisms. However, we can prove this special case with ease without any advanced Lie theory. We find that for any $g \in G$,

$$(\rho \circ c_g)(h) = \rho(ghg^{-1}) = \rho(g)\rho(h)\rho(g^{-1}) = (c_{\rho(g)} \circ \rho)(h)$$

If we first differentiate this equation over the variable h , we find that $\rho_* \circ \text{Adj}(g) = \text{Adj}(\rho(g)) \circ \rho_*$, and then by differentiating this equation over the variable g , we find that $\rho_* \circ \text{adj}(X) = \text{adj}(\rho_*(X)) \circ \rho_*$ which we can restate in terms of the Lie bracket as $\rho_*[X, Y] = [\rho_*(X), \rho_*(Y)]$ for any $X, Y \in \mathfrak{g}$.

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 $K = K$ or $K = \mathbf{R}$, then the representations of the additive Lie group K 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 K on V .

Example. The one dimensional representations of a Lie group G on K are classified by the differentiable characters $\chi : G \rightarrow K$ of the group. The one dimensional representations of a Lie algebra \mathfrak{g} on K are given by the linear functionals $\lambda : \mathfrak{g} \rightarrow K$, which vanish on the derived subalgebra \mathfrak{g}' , defined by letting $Xz = \lambda(X)z$. This makes sense, because χ vanishes on the commutator

subgroup $[G, G]$. Thus if G is simply connected, for every linear functional λ which vanishes on \mathfrak{g}' , there is a character χ on G whose differential at e is λ , and this correspondence is one-to-one.

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 $\bigwedge V$. 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 $(gf)(v) = -f(gv)$.

Example. The adjoint representation of $\mathfrak{o}_3(\mathbf{R})$ is self dual. If we denote the adjoint representation by $\rho : \mathfrak{o}_3(\mathbf{R}) \rightarrow \mathfrak{gl}_3(\mathbf{R})$, and if we consider a basis of $\mathfrak{o}_3(\mathbf{R})$ of the form $a = E_{12} - E_{21}$, $b = E_{13} - E_{31}$, and $c = E_{23} - E_{32}$, then $[a, b] = -c$, $[a, c] = b$, and $[b, c] = -a$. With respect to this basis, the adjoint operation has

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

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

Before we get to some serious representation theory, we need some elementary results about classes of operators in $\mathfrak{gl}(V)$ which form Lie algebras, which will enable us to restrict the structure of any representation.

Theorem 3.1. If $T : V \rightarrow V$ is diagonalizable, then $\text{adj}_T : \mathfrak{gl}(V) \rightarrow \mathfrak{gl}(V)$ is diagonalizable, viewed as a linear operator over $\mathfrak{gl}(V)$.

Proof. Let v_1, \dots, v_n be a basis of eigenvectors of V , with $Tv_i = \lambda_i v_i$. If we define the canonical basis $E_{ij} : V \rightarrow V$ of linear operators on V with respect to the basis v_i , for which $E_{ij}(v_k) = \delta_{ik} v_j$, then $[E_{ij}, E_{kl}] = \delta_{jk} E_{il} - \delta_{il} E_{kj}$, and $T = \sum \lambda_i E_{ii}$, so

$$[T, E_{ij}] = \sum_{k=1}^n \lambda_k [E_{kk}, E_{ij}] = (\lambda_i - \lambda_j) E_{ij}$$

Therefore the E_{ij} form a basis for $\mathfrak{gl}(V)$ diagonalizing adj_T . \square

The construction in this theorem tells us something stronger. If we can count the multiplicities of the eigenvalues of T , then we can also count the multiplicities of the eigenvalues of adj_T , by subtracting the values of the eigenvalues.

Theorem 3.2. *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 1 in $\mathfrak{gl}(V)$ is not nilpotent, yet $1 \in Z(\mathfrak{gl}(V))$ and so $\text{adj}_1 = 0$ 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 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 representations of a Lie algebra. 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 . λ 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_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 $Xe_i = X_{ii}e_i$. Conversely, 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 obtain a powerful decomposition theory for Lie algebra representations. A **submodule** or **subrepresentation** of a \mathfrak{g} representation V is a subspace which is closed under multiplication by elements of \mathfrak{g} . As examples, the submodules of \mathfrak{g} with respect to the adjoint representation are exactly the ideals of \mathfrak{g} . Given a subrepresentation W , we can form the factor representation 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 representations 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 representation V is called **irreducible**, or **simple**, if it has no subrepresentations other than (0) and V itself. 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. If a finite dimensional representation V is not irreducible, we can find a subrepresentation W of minimal dimension. Then V/W will have an irreducible submodule by induction, and we may proceed 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. These properties transfer back and forth from the Lie algebra to the Lie group.

Lemma 3.3. *If G is connected, then a representation of the Lie group is irreducible if and only if the corresponding Lie algebra representation is irreducible.*

Proof. Let $\rho : G \rightarrow GL(V)$ is a representation of a Lie group, then for a fixed nonzero $v \in V$, the map $f : g \mapsto \rho(g)(v)$ is differentiable, and the map certainly has constant rank because

$$f(g) = f(h^{-1}hg) = \rho(h^{-1})f(hg)$$

so $f_* = \rho(h^{-1})_* \circ f_* \circ (m_h)_*$, and $\rho(h^{-1})$ and m_h have full rank, and $(m_h)_*$ maps the tangent space at the identity in G to the tangent space at h . Since all connected Lie groups are second countable, we conclude that f is a submersion. This implies that $\rho_* : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ is actually surjective, hence ρ_* is an irreducible representation. Conversely, if ρ_* is surjective, then it implies that ρ is actually an open map in a range of the identity so in particular we can map any vector v to a sufficiently small scalar multiple of any other vector w , implying that ρ is also irreducible. \square

Lemma 3.4. *If \mathfrak{g} acts on V , and \mathfrak{a} is an ideal of \mathfrak{g} , then the space of vectors $x \in V$ such that $Xx = 0$ for all $X \in \mathfrak{a}$ is a subrepresentation of V .*

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

Corollary 3.5. *If V is an irreducible representation of \mathfrak{g} , and \mathfrak{a} is a proper, nontrivial ideal of \mathfrak{g} , then for every $x \in V$ there is $X \in \mathfrak{a}$ with $Xx \neq 0$.*

In the language of weights, 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 3.6. *If \mathfrak{a} is an ideal in a Lie algebra \mathfrak{g} with a representation 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 λ . 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 3.7. *Over a field of characteristic zero, if V is a representation of a Lie algebra \mathfrak{g} with ideal \mathfrak{a} , then the weight space of any $\lambda : \mathfrak{a} \rightarrow K$ is a subrepresentation of V .*

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

Corollary 3.8. *For any irreducible representation V of \mathfrak{g} over a field of characteristic zero, and for any $\lambda \in \mathfrak{g}^*$, either $V_\lambda = 0$ or $V_\lambda = V$.*

This is one of the first theorems of the theory of Lie algebras 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** (this name is completely nonstandard, but I can't find a short name for such a field).

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 3.9. *If \mathfrak{g} is a linear Lie algebra over a Cartan field, and $X, Y \in \mathfrak{g}$ commute with $[X, Y]$ in the representation, then $[X, Y]$ is nilpotent.*

Proof. It suffices to show that the only eigenvalue of $[X, Y]$ is 0. Let λ be an eigenvalue of $[X, Y]$, and let \mathfrak{h} 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{h} , and

λ acts as a weight on this ideal, hence the space V_λ of vectors v satisfying $[X, Y]v = \lambda v$ is \mathfrak{g} invariant. This implies that X and Y restrict to operators on V_λ , and the restriction of $[X, Y]$ to V_λ is the same as $XY - YX$, hence $[X, Y]$ has trace zero. But $[X, Y]$ can be put into normal form on V_λ , and we find it has trace $n\lambda$ if V_λ is n -dimensional, hence $\lambda = 0$. \square

Another consequence tells us about the normalizer of some subalgebra \mathfrak{h} of \mathfrak{g} , which is the largest subalgebra of \mathfrak{g} in which \mathfrak{h} is an ideal (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 3.10. *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

3.2 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 3.11. *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 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 3.12 (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

Corollary 3.13. *The only irreducible representations of a nilpotent Lie algebra \mathfrak{g} are one dimensional.*

Proof. If $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ is an irreducible representation, and \mathfrak{g} is nilpotent, then $\rho(\mathfrak{g})$ is nilpotent, because it is isomorphic to a quotient of \mathfrak{g} , and every quotient of a nilpotent Lie algebra is nilpotent. Engel's theorem tells us that there is a basis v_1, \dots, v_n for V such that for any $X \in \mathfrak{g}$, $Xv_i \in \text{span}(v_{i+1}, \dots, v_n)$. But this implies that $n = 1$, for otherwise v_2, \dots, v_n span a subrepresentation of V . \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 3.14. *\mathfrak{g} is nilpotent if and only if adj_X is nilpotent for all $X \in \mathfrak{g}$.*

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

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

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

Corollary 3.15. *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 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 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 3.16. *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_λ 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, where we apply the one dimensional case, and then apply a particular attempt to extend the technique by induction.

Theorem 3.17. *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 3.18. *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 3.19. *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 λ and γ 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. This is what makes the study of Lie algebras over fields of positive characteristic so difficult.

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.

Corollary 3.20. *Over a Cartan field, all irreducible representations of a solvable Lie algebra are one-dimensional.*

This can be generalized to a result about the irreducible representations of all Lie algebras, by noting that $\mathfrak{g}/\text{rad}(\mathfrak{g})$ is always solvable. First, we consider a trick to modify the definition of a representation of a linear functional. Given a representation $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$, and a one-dimensional representation specified by some linear functional $\lambda : V \rightarrow K$, then $V \otimes K$ can be identified with the representation $\rho \otimes \lambda : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ where $(\rho \otimes \lambda)(X) = \rho(X) + \lambda(X)$. In particular, since K^* corresponds to the representation over K corresponding to the linear functional $-\lambda$, $K \otimes K^*$ is always the trivial representation, and we can use tensoring by K to remove linear relations in the group among eigenvectors.

Theorem 3.21. *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 $\text{tr}(X) = \dim(V)\lambda(X)$, λ vanishes on $\text{rad}(\mathfrak{g}) \cap \mathfrak{g}'$. If λ is extended to all of V , and vanishes on \mathfrak{g}' , then we take the one dimensional representation K of \mathfrak{g} , and then write $V = (V \otimes K^*) \otimes K$. Since the representation on K^* satisfies $Xf = -\lambda(X)f$, so if $X \in \mathfrak{g}'$, then $\lambda(X) = 0$, so $Xf = 0$ for $f \in K^*$, and $Xv = 0$ for $v \in V$. It follows that $V \otimes K^*$ 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 $V \otimes K^*$ factors into a representation of $\mathfrak{g}/\text{rad}(\mathfrak{g})$, and K is a one-dimensional representation of \mathfrak{g} . \square

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 representations over a semisimple Lie algebra over a Cartan field 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, and 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 3.22 (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

We can use Schur's lemma to prove that, unlike the class of Lie algebras, not all Lie groups can be embedded into a matrix group.

Example. *The group H_0 obtained from the Heisenberg group in two dimensions by letting b range over \mathbf{T} is interesting because it is a Lie group with no faithful representation (refer to the notes on the classification of low dimensional Lie algebras for this construction). We see the group as $\mathbf{R}^2 \times \mathbf{T}$ with the group multiplication operation*

$$(a_0, b_0, c_0)(a_1, b_1, c_1) = (a_0 + a_1, e^{ia_0 c_1} b_0 b_1, c_0 + c_1)$$

Since the group has no faithful representation, it cannot be isomorphic to a Lie subgroup of matrices over the complex numbers, and therefore Ado's theorem fails for Lie groups. If ρ is any irreducible, finite dimensional representation of the group on a vector space V , then since $Z(H_0) = \{(0, z, 0) : z \in \mathbf{T}\}$ form a compact, abelian subgroup, they must be simultaneously diagonalizable (it is a general fact from harmonic analysis that every compact abelian group has one dimensional irreducible representations). If we pass to the underlying Lie algebra, we see that this implies that $Z = E_{13}$ acts diagonalizably on V . But since $Z \in Z(\mathfrak{h})$, this means that X is actually a morphism of the representation, and by Schur's lemma, this implies that Z acts as scalar multiplication. That is, there is λ such that $Zv = \lambda v$ for all vectors $v \in V$. But since Z is in the derived algebra of $Z(\mathfrak{h})$, the trace of Z is zero, so that $Zv = 0$ for all vectors v , and this implies that elements of the form $(0, b, 0)$ act trivially on V as well, so the representation cannot be faithful. Nonetheless, this group does have faithful infinite dimensional representations, though this is beyond the scope of these notes.

3.3 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 $\mathfrak{sl}_2(K)$, where K is algebraically closed and has characteristic zero. What's more, we will find that the representations of this Lie algebra control a large part of the representation theory of the semisimple Lie algebras. It is easy to conjure up a representation of $\mathfrak{sl}_2(K)$, because $SL_2(K)$ acts on K^2 via its standard representation, and differentiating this action gives us an action of $\mathfrak{sl}_2(K)$ on K^2 by standard matrix multiplication. Recalling the basis $e = E_{12}$, $f = E_{21}$, and $h = E_{11} - E_{22}$, we find that if $X = (1, 0)$, and $Y = (0, 1)$, then

$$eX = 0 \quad eY = X \quad fX = Y \quad fY = 0 \quad hX = X \quad hY = -Y$$

To formally verify this is a representation, we need to verify the bracket property for $[e, f] = h$, $[e, h] = -2e$, and $[f, h] = 2h$, and this is done by a trivial calculation for X and Y separately, so V is really a representation of Lie algebras. We can now use this representation to generate a family of interesting representations by taking finite subrepresentations of the space $\text{Sym}(V)$, which can be viewed as the space of all polynomials in X and Y . We find

$$e(X^n Y^m) = mX^{n+1} Y^{m-1} \quad f(X^n Y^m) = nX^{n-1} Y^{m+1}$$

$$h(X^n Y^m) = (n - m)X^n Y^m$$

Thus, viewing $\text{Sym}(V)$ as the space of polynomials in X and Y , we see that $\mathfrak{sl}_2(K)$ acts as the differential operators

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

We let $V_n = \text{Sym}^n(V)$ be the space of homogenous polynomials of degree n , which is preserved under the action of $\mathfrak{sl}_2(K)$. It is spanned by

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

and is therefore a space of dimension $n + 1$. In the basis $X^n, X^{n-1}Y, \dots, Y^n$, we have matrix representations

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

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

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

Theorem 3.23. *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 modules V_n classify all irreducible modules of $\mathfrak{sl}_2(K)$. The trick, given a general module V , is to look at how the eigenvectors of h are modified by the action of $\mathfrak{sl}_2(K)$.

Lemma 3.24. *If $v \in V$ is an eigenvector of h with eigenvalue λ , then*

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

Proof. (i) is proved by calculation,

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

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

Now if V is an irreducible $\mathfrak{sl}_2(K)$ module, this lemma implies that all eigenvalues of $\mathfrak{sl}_2(K)$ are congruent modulo 2, and the eigenvalues must appear as a string of numbers of the form $x, x+2, \dots, x+2k$.

Lemma 3.25. *If V is a finite dimensional $\mathfrak{sl}_2(K)$ module, then V contains an eigenvector v for h such that $ev = 0$.*

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

Theorem 3.26. *If V is a finite dimensional irreducible $\mathfrak{sl}_2(K)$ module over an algebraically closed field of characteristic zero, then V is isomorphic to some V_n .*

Proof. Find an eigenvector v for h such that $ev = 0$, with eigenvalue λ for h . Consider the sequence $v, f v, f^2 v, \dots$. The last lemma essentially implies that $f^{m+1} v = 0$, $f^m v \neq 0$. We claim that $v, f v, \dots, f^m v$ 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 $e(f^k v) \in \text{span}(f^l v : k \leq l)$. For $k = 0$, we know $ev = 0$. For the induction, we find

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

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

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

hence $\lambda = m$, since the image of f is in the derived subgroup, and therefore has trace zero. We now have enough information to provide an explicit homomorphism with V_m . Note that V_m is spanned by $X^m, fX^m, \dots, f^m X^m$. If we set $\psi(f^k v) = f^k X^m$ this defines a vector space isomorphism which

commutes with the action of h and f . It remains to show that ψ commutes with e , we use induction. For $k = 0$, $eX^m = 0 = \psi(ev)$. Now by induction,

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

and this completes the correspondence. \square

Corollary 3.27. *If V is a finite dimensional $\mathfrak{sl}_2(K)$, 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 $\mathfrak{sl}_2(K)$ representation 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** V of the representation, and uniquely characterizes the module V up to isomorphism if V is irreducible. We can use this fact to decompose the various representations obtained by the standard constructions of representations from the other representations, because we will soon find that the class of $\mathfrak{sl}_2(K)$ modules is semisimple – every finite dimensional module can be decomposed into the direct sum of irreducible modules. This constitutes Weyl's theorem.

Example. *Consider the representation $V_2 \otimes V_2$, which has dimension 9. If $x, y \in V_2$ satisfy $hx = \lambda x$ and $hy = \gamma y$, then $h(x \otimes y) = (\lambda + \gamma)(x \otimes y)$. Since V_2 has a basis x_1, x_2, x_3 with eigenvalues $-2, 0, 2$, so $V_2 \otimes V_2$ has eigenvalues $\pm 4, \pm 2$, and 0 , where ± 4 has multiplicity 1, ± 2 have multiplicity 2, and 0 has multiplicity 3. Thus $V_2 \otimes V_2$ is isomorphic as an $\mathfrak{sl}_2(K)$ module to $V_4 \oplus V_2 \oplus V_2$. In general, the eigenvalues of $V \otimes W$ are obtained by convolving the eigenvalues of V and W together by a summation. For instance, we find that $V_2 \otimes V_3$ has eigenvalues $+2, 0, -2$ and $+3, +1, -1, -3$, so that $V_2 \otimes V_3$ has eigenvalues ± 5*

has multiplicity 1, ± 3 has multiplicity 2, ± 1 has multiplicity 3. In general, one can argue by the same strategy that for $a \geq b$, $V_a \otimes V_b$ can be decomposed as $V_{a+b} \oplus V_{a+b-2} \oplus \cdots \oplus V_{a-b}$.

Example. Consider $\text{Sym}^2(V_2)$. We know that $V_2 \otimes V_2 = V_4 \oplus V_2 \oplus V_0$. Since the vector space spanned by $x \otimes y - y \otimes x$, for $x, y \in V_2$ is a subrepresentation of $V_2 \otimes V_2$, it is equal to the direct sum over some subset of the terms in the direct sum. If $x \neq y$, where $hx = \lambda x$ and $hy = \gamma y$, then $x \otimes y - y \otimes x$ is a non-zero vector, which is an eigenvector of eigenvalue $\lambda + \gamma$. The space spanned by all symmetric differences has dimension 3, and this therefore implies that the space is just V_2 , and therefore $\text{Sym}^2(V_2)$ is isomorphic to $V_4 \oplus V_0$. A more direct way to see this is to define a $\mathfrak{sl}_2(K)$ homomorphism from $\text{Sym}^2(V_2)$ to $\text{Sym}^4(V_2)$ by evaluating the elements of V_2 out. Thus $(XY)(Y^2)$ evaluates to XY^3 , and $(Y^2)(X^2)$ evaluates to X^2Y^2 . That this is a $\mathfrak{sl}_2(K)$ homomorphism follows from the product rule for differentiation of polynomials. The map is surjective, so the kernel is one dimensional, generated by $(X^2)(Y^2) - (XY)(XY)$. We essentially remove an eigenvector for each fourth integer, so that in general

$$\text{Sym}^n(V_2) = \bigoplus_{k=0}^{\lfloor n/2 \rfloor} V_{2n-4k}$$

We find that the same decomposition holds for $\text{Sym}^2(V_n)$.

Example. Switching to looking at the wedge products of V_2 , we note that $\bigwedge^2 V_2$ is three dimensional, so is either 3 copies of V_0 , a copy of V_0 and V_1 , or a copy of V_2 . We can obtain the wedge product as a quotient under the span of $v \otimes w + w \otimes v$, and we find that since V_2 has eigenvalues $\{-2, 0, 2\}$, we take out one copy of ± 4 , one copy of ± 2 , and two copies of zero, so that since $V_2 \otimes V_2 = V_4 \oplus V_2 \oplus V_0$, $\bigwedge^2 V_2$ is obtained by removing the copy of V_4 and the copy of V_0 , and hence is isomorphic to V_2 . This implies that $\bigwedge^3 V_2$ is also isomorphic to V_2 .

Now we know that since the complexification of \mathfrak{su}_2 and $\mathfrak{sl}_2(\mathbf{R})$ is isomorphic to $\mathfrak{sl}_2(\mathbf{C})$, we know that we can obtain all irreducible complex representations of these Lie algebras by restricting the irreducible representations on $\mathfrak{sl}_2(\mathbf{C})$ to these groups. We obtain the complexification by embedding $\mathfrak{sl}_2(\mathbf{R})$ and \mathfrak{su}_2 in $\mathfrak{sl}_2(\mathbf{C})$ by inclusion, so that the actions on V_1 are given by the standard actions of $\mathfrak{sl}_2(\mathbf{R})$ and \mathfrak{su}_2 , and the remaining actions are obtained by taking symmetric products. This implies that all

the complex irreducible representations of SU_2 are given by the symmetric products of the standard representation on \mathbf{C}^2 . We know there is a surjective homomorphism $f : SU_2 \rightarrow SO_3(\mathbf{R})$ which is two to one, whose kernel is ± 1 . This implies that the irreducible complex representations of $SO_3(\mathbf{R})$ are in correspondence with those irreducible representations of SU_2 which act trivially on ± 1 . Now

$$(-1)(X^n Y^m) = (-X)^n (-Y)^m = (-1)^{n+m} X^n Y^m$$

Which implies that the irreducible representations $V_n = \text{Sym}^n(V)$ descend to $SO_3(\mathbf{R})$ only if n is an even number. Indeed, since $\mathfrak{so}_3(\mathbf{R}) \cong \mathfrak{su}_2$, we can take certainly define all irreducible representations on the Lie algebra, but these only lift when n is an even number. In fact, if n is an odd number, then the standard technique to lift Lie algebra homomorphisms give us a two-valued homomorphism $f : SO_3(\mathbf{R}) \rightarrow GL(V)$. If one ‘applies f continuously’ by fixing a vector $v \in V$ and applying $f(X(t))(v)$ as we continuously vary $X(t)$, then one finds that the path of rotations induced by a 360 degree rotation in some plane actually give us $f(X(1))v = -f(X(0))v$, which reflects the fact that the fundamental group of $SO_3(\mathbf{R})$ is \mathbf{Z}_2 . In the theory of quantum field theory, one encounters both ‘integer spin particles’, whose behaviour corresponds to an even-degree representation of V , and a ‘half spin particles’, such as the electron (a spin 1/2 particle). If we perform a 360 degree rotation of the wave function of the electron, we find that we obtain the negation of the wave function.

3.4 Projective Geometry and $\mathfrak{sl}_2(\mathbf{C})$

There is an interesting interpretation the decomposition of the tensor products of irreducible representations of $\mathfrak{sl}_2(\mathbf{C})$ which occurs in projective geometry. If $V = \{X_1, \dots, X_m\}$, then the space $\text{Sym}^n(V)$ can be viewed as the space of homogenous polynomials of degree n in the variables X_1, \dots, X_m . The space $\mathbf{P}(\text{Sym}^n V)$ is then the set of hypersurfaces of degree n in $\mathbf{P}V$. Given a representation of a group G on a vector space V , it is natural to switch to studying the induced group action of G on $\mathbf{P}V$, known as a **projective representation**. Here we will look at the geometric interpretations of the decomposition properties of projective representations of $SL_2(\mathbf{C})$.

Since $SL_2(\mathbf{C})$ is a simply connected group, the representations V_n constitute all irreducible representations of $SL_2(\mathbf{C})$. They are constructed

from the standard representation of $SL_2(\mathbb{C})$ on the two dimensional vector space V by taking symmetric powers. The representation on homogenous polynomials of degree n descends to $PSL_2(\mathbb{C})$ only when n is even. Regardless, any projective representation of $SL_2(\mathbb{C})$ on $\mathbf{P}V$ descends to $PSL_2(\mathbb{C})$, because multiplying by a non-zero scalar is idempotent. In the case of the one dimensional projective space $\mathbf{P}V$, the action is transitive, because

$$\begin{pmatrix} a & 0 \\ b & a^{-1} \end{pmatrix} X = aX + bY \quad \begin{pmatrix} a & -b^{-1} \\ b & 0 \end{pmatrix} X = aX + bY$$

allows us to map X to any nonzero vector in V . However, the action on $\mathbf{P}V_n$ will not always be transitive, and the orbits of the action give many interesting shapes from the perspective of algebraic geometry.

Example. For any n , the veronese embedding embeds $\mathbf{P}V$ into $\mathbf{P}Sym^n(V)$ by mapping the line through $v \in V$ to the line through v^n . The image of this map is known as the **rational normal curve** of degree n . $\mathbf{P}V$ is naturally isomorphic to \mathbf{CP}^1 under the map $[a : b] \mapsto [aX + bY]$, so the rational normal curves are parameterized by points on the Riemann sphere. Similarly, \mathbf{CP}^n is isomorphic to $\mathbf{P}Sym^n(V)$ where

$$[a_1 : \cdots : a_{n+1}] \mapsto [a_1 X^n + a_2/n X^{n-1} Y + \cdots + a_{n+1} Y^n] = \left[\sum \binom{n}{k} a_k X^{n-k} Y^k \right]$$

The rational normal curves therefore are parameterized by

$$[a : b] \mapsto [(aX + bY)^n] = \sum \binom{n}{k} a^k b^{n-k} X^k Y^{n-k} = [a^n : a^{n-1} b : \cdots : ab^{n-1} : b^n]$$

For $n = 1$, the curve is all of \mathbf{PC} . For $n = 2$, the curve is obtained from the equation $X = Y^2$ by adding the point $(1 : 0 : 0)$ at infinity. For $n = 3$, the curve is obtained from the algebraic equations $X = Z^3, Y = Z^2$ by adding the point $(1 : 0 : 0 : 0)$ at infinity, and so on and so forth. Returning to our discussion of the action of $PSL_n(\mathbb{C})$ on $Sym^n(V)$, we find that the rational normal curve is preserved under the action of $PSL_2(\mathbb{C})$, because

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} [1 : 0 : \cdots : 0] = [c^n : ac^{n-1} : \cdots : a^n]$$

$$\begin{aligned}
\begin{pmatrix} a & b \\ c & d \end{pmatrix} [x^n : x^{n-1} : \cdots : 1] &= \sum_{k=0}^n \binom{n}{k} [(axX + cxY)^k (bX + dY)^{n-k}] \\
&= [((ax + b)X + (cx + d)Y)^n] \\
&= [(ax + b)^n : (ax + b)^{n-1}(cx + d) : \cdots : (cx + d)^n]
\end{aligned}$$

Conversely, any isomorphism of \mathbf{P}^n which fixes every point on the rational normal curve is trivial. To prove this, we rely on the fact that these isomorphisms are described exactly by the class PGL_{n+1} (the isomorphisms are either the biholomorphisms of \mathbf{P}^n as a complex manifold, or the isomorphisms of an algebraic variety). This follows because any biholomorphism fixing infinitely many points in \mathbf{C}^n is the identity. This implies that the action of PSL_2 on \mathbf{P}^n describes all isomorphisms preserving the rational normal curve, since the action of $\mathrm{PSL}_2(\mathbf{C})$ is transitive on the rational normal curve *FIX LATER*.

Now let us interpret the decomposition properties of symmetric powers of representations of $\mathrm{SL}_2(\mathbf{C})$. To begin with, we have the decomposition

$$\mathrm{Sym}^2(\mathrm{Sym}^2(V)) \cong \mathrm{Sym}^4(V) \oplus V_0$$

We can obtain this decomposition from the surjective map mapping elements of $\mathrm{Sym}^2(\mathrm{Sym}^2(V))$ to $\mathrm{Sym}^4(V)$ obtained from evaluating polynomials, whose kernel is $(X^2)(Y^2) - (XY)^2$.

Chapter 4

Killing Forms

4.1 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 4.1. *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 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 λ_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 λ_i form a basis for E . This shows that $E^* = 0$, and therefore that $E = 0$. \square

Theorem 4.2. *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 adj_X is nilpotent for any $X \in \mathfrak{g}$, and there is a basis for \mathfrak{g} upon which the adj_X are strictly upper triangularized, and therefore $\kappa(X, Y) = \text{tr}(\text{adj}_X \circ \text{adj}_Y) = 0$ for any $X, Y \in \mathfrak{g}$. The converse is not true, however; there are non-nilpotent Lie algebras with trivial Killing form.

Lemma 4.3. 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 κ denote the Killing form on \mathfrak{g} , and $\kappa_{\mathfrak{a}}$ the Killing form on \mathfrak{a} . If $X \in \mathfrak{a}$, then $\text{adj}_X : \mathfrak{g} \rightarrow \mathfrak{g}$ maps \mathfrak{g} into \mathfrak{a} , and therefore if we take a basis v_1, \dots, v_n for \mathfrak{a} and extend it to a basis on \mathfrak{g} , the matrix representation of adj_X will be

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

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

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

and so

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

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

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

Lemma 4.4. *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 4.5 (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 κ is nondegenerate.

Conversely, if \mathfrak{g} has a non-trivial abelian ideal \mathfrak{a} , and if $Y \in \mathfrak{a}$ is non-zero, then $\text{adj}_Y \circ \text{adj}_X \circ \text{adj}_Y = 0$ for any $X \in \mathfrak{g}$, hence $(\text{adj}_Y \circ \text{adj}_X)^2 = 0$. Nilpotent maps have trace zero, so $\kappa(Y, X) = 0$ for all X , and so κ is degenerate. \square

Example. 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 4.6. *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 4.7. *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 4.8. *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 κ , which is non-degenerate, and induces another dual map, which we will distinguish from the one induced by f by denoting the former by X^{*f} , and the latter by $X^{*\kappa}$. Now if $F : X \mapsto X^{*f}$ is the dual map, and $G : X \mapsto X^{*\kappa}$ is the other dual map, then $G \circ F^{-1}$ is a \mathfrak{g} module isomorphism, and \mathfrak{g} is an irreducible module since it is simple, so we may apply Schur's lemma to conclude that there is λ such that $(G \circ F^{-1})(X) = \lambda X$ for all X , or that $G(X) = \lambda F(X)$ for all X . We conclude that $\kappa(X, Y) = \lambda f(X, Y)$ for some $\lambda \neq 0$. Thus the Killing form is essentially the only associative, symmetric, non-degenerate bilinear form on a simple Lie algebra.

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

Lemma 4.9. *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 $\lambda \in K$, let \mathfrak{g}_λ be the set of X such that $(D - \lambda)^m X = 0$ for some m . Then the vector space structure of \mathfrak{g} decomposes into the sum of the subspaces \mathfrak{g}_λ . We have $[\mathfrak{g}_\lambda, \mathfrak{g}_\gamma] \subset \mathfrak{g}_{\lambda+\gamma}$, because

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

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

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

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

Theorem 4.10. *If \mathfrak{g} is semisimple, then any $X \in \mathfrak{g}$ can be written uniquely as $D + N$, where adj_D is diagonalizable, adj_N is nilpotent, and $[D, N] = 0$. If $[X, Y] = 0$, then $[D, Y] = [N, Y] = 0$.*

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

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

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

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

Theorem 4.11. *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 ρ , then $\mathfrak{g}/\mathfrak{h}$ is isomorphic to $\rho(\mathfrak{g})$. We can therefore talk about the abstract Jordan decomposition of elements of $\rho(\mathfrak{g})$. The concrete decomposition of $\rho(X)$ agrees with the abstract decomposition, because if Y is nilpotent/diagonalizable it implies adj_Y is as well. If adj_D is diagonalizable, then it is diagonalizable on \mathfrak{h} by some basis X_1, \dots, X_n , and if we extend this basis of eigenvectors to the whole space, $X_1, \dots, X_n, Y_1, \dots, Y_m$, then, viewed as a map on the quotient $\mathfrak{g}/\mathfrak{h}$, we have $[D, Y_i + \mathfrak{h}] = \lambda_i Y_i + \mathfrak{h}$. If adj_N is nilpotent, then adj_N is nilpotent on the quotient as well. This verifies that $\rho(X)$ has abstract Jordan decomposition $\rho(D) + \rho(N)$, because the decomposition is unique. \square

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

4.2 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 4.12. *If ρ is a faithful representation of a semisimple Lie algebra, then β is nondegenerate.*

Proof. Define the radical $\text{rad}(\beta)$ of the bilinear map β 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 β . But this implies that $\rho(\text{rad}(\beta))$ is a solvable algebra, since the trace vanishes on products in this algebra, and since ρ 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 β 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 4.13. *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 β , 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 4.14. *If X_1, \dots, X_n and X'_1, \dots, X'_n are two bases of \mathfrak{g} , then the induced Casimir operators c_ρ^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 4.15 (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_ρ . 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 λ 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 ± 1 of weight 2, and an eigenvalues ± 2 of weight 1, so that as a representation, the module is isomorphic to $V_2 \oplus V_1 \oplus V_1 \oplus V_0$.

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

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

Let us try and understand why this technique will be so useful. Let \mathfrak{h} be an arbitrary abelian subalgebra of \mathfrak{g} consisting of semisimple elements. If $X, Y \in \mathfrak{h}$, then adj_X and adj_Y are commutative, because

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

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

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

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

Theorem 4.16. *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. □

As a visual example, we need to cheat a little. We will soon prove that $\mathfrak{sl}_3(K)$ has an abelian subalgebra \mathfrak{h} with weights α_{ij} , for $1 \leq i \neq j \leq 3$, where $\alpha_{ij}(X) = X_{ii} - X_{jj}$. If we look at $X \in \mathfrak{sl}_2(K)_{\alpha_{12}}$, and consider adj_X on this space, then the theorem above tells us that we can see the motion of adj_X on the root spaces by the diagram

$$\begin{array}{ccccc}
 & \alpha_{23} & & & \\
 & \cdot & \xrightarrow{\quad} & \cdot & \xrightarrow{\alpha_{13}} & \cdot & 0 \\
 \\
 \alpha_{21} & & & & & & \\
 \cdot & \xrightarrow{\quad} & \cdot & \xrightarrow{\quad} & \cdot & \xrightarrow{\alpha_{12}} & \cdot & 0 \\
 \\
 & & & & & & \\
 & \cdot & \xrightarrow{\quad} & \cdot & \xrightarrow{\quad} & \cdot & 0 \\
 \alpha_{31} & & & & & \alpha_{32} &
 \end{array}$$

Thus if $Y \in \mathfrak{sl}_3(K)_{\alpha_{13}}$, or $Y \in \mathfrak{sl}_3(K)_{\alpha_{12}}$, or $Y \in \mathfrak{sl}_3(K)_{\alpha_{32}}$, then $[X, Y] = 0$. This implies strong orthogonality properties for the Killing form, since the adjoint representation always pushes weight spaces around as the sum of two weights, and provided this sum does not equal zero, there is no way to return to the original weight space without annihilating the vector.

Theorem 4.17. *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 κ that

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

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

κ acts very simply on \mathfrak{h} , because if $X, Y \in \mathfrak{h}$, and $Z \in \mathfrak{g}_\lambda$, then $(\text{adj}_X \circ \text{adj}_Y)(Z) = \lambda(X)\lambda(Y)Z$, and because \mathfrak{h} is abelian, we find

$$\kappa(X, Y) = \sum_{\lambda} \dim(\mathfrak{g}_\lambda) \lambda(X) \lambda(Y)$$

We shall find that if \mathfrak{h} is a ‘maximal’ abelian subalgebra, which we will introduce as a family of Cartan subalgebras, then we find that the dimension

of each weight space is one, and so $\kappa(X, Y) = \sum \lambda(X)\lambda(Y)$. This shows that κ is non-degenerate on \mathfrak{h} , because if there was X such that $\kappa(X, Y) = 0$ for all $Y \in \mathfrak{h}$, then 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, and no semisimple Lie algebra is 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 κ and how the bracket operates on weights not powerful enough to classify the algebra. It turns out that every semisimple Lie algebra has a maximal abelian subalgebra containing semisimple elements, called a **Cartan subalgebra**. This algebra has the property that the Lie algebra elements of weight zero correspond precisely to elements of the Cartan subalgebra, so that the weight decomposition is very tight.

Example. If \mathfrak{h} is the 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 4.18. *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 the \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}_{α} is the eigenspace with weight α 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 4.19. *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 α , there is Z with $\alpha(Z) \neq 0$, hence $X_{\alpha} = 0$, and so $X \in \mathfrak{h}$. □

All of our results about the Killing form has focused on the fact that the Lie algebras we are trying to understand are semisimple. But we have not proved that any of the interesting Lie algebras we want to study *are* semisimple. It turns out that if we can find a decomposition of a Lie algebra into weight spaces, like the result in this chapter show, then the Lie algebra must be semisimple. This also shows that the techniques we discuss likely have no generalization to non semisimple Lie algebras.

Theorem 4.20. *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 α with $\alpha(X) \neq 0$.*

- For each α , \mathfrak{g}_α is one dimensional, spanned by some X_α .
- 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 $[\mathfrak{h}, \mathfrak{a}] \subset \mathfrak{a}$, we find \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 Φ_0 of Φ . 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 α_{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, and here we will have to distinguish the structure of $\mathfrak{o}_n(K)$ depending on whether n is odd or even. For $n = 2m$, define the matrix J by

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

There is a matrix N such that $J = N^t N$, since J is nondegenerate, and then the map $M \mapsto NMN^{-1}$ is an isomorphism between the set of matrices preserving the bilinear form corresponding to J , and the matrices preserving the bilinear form $\sum x_i y_i$, which is really just $O_n(K)$. This implies that the Lie algebras of these two Lie groups are isomorphic, and so we will often find that is more convenient to think of $\mathfrak{o}_n(K)$ as the set of 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 $\mathfrak{o}_n(K)$. 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 $X^t J + JX = 0$ are the matrices $X = \begin{pmatrix} A & B & C \\ D & E & F \\ G & K & L \end{pmatrix}$, with

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

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 α_{ij} , β_{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 4.21. *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 adj_Y , and consider an alternate extension of $\{Z'_1, \dots, Z'_n\}$ which diagonalize adj_X on $C(Y)$. Then the triple of basis are linearly independent, for if

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

Then for some λ_i

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

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

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

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

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

Theorem 4.22. *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 adj_Y is equal to adj_N on $C(X)$, because $S \in \mathfrak{h} = C(X)$. This implies adj_Y is nilpotent, hence Y is nilpotent, and since Y was arbitrary we use Engel's theorem to conclude that $C(\mathfrak{h})$ is nilpotent. Next, we claim that every element of $C(\mathfrak{h})$ is semisimple. If $Y \in C(\mathfrak{h})$, write $Y = S + N$. $C(\mathfrak{h})$ is nilpotent, hence solvable, so by Lie's theorem there is a basis for \mathfrak{g} in which case adj_Y is represented by an upper triangular matrix. Since adj_N is nilpotent, the matrix is strictly upper triangular, hence if $Z \in C(X)$,

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

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

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

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

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

4.4 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 4.23. *If α is a non-zero root, then $-\alpha$ is a non-zero root, and for every non-zero $X \in \mathfrak{g}_{\alpha}$ there is $Y \in \mathfrak{g}_{-\alpha}$ such that $\text{span}(X, Y, [X, Y])$ is a Lie subalgebra of \mathfrak{g} isomorphic to $\mathfrak{sl}_2(K)$.*

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

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

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

We let $\mathfrak{sl}(\alpha)$ denote the subalgebra obtained in the lemma. As we now know, $\mathfrak{sl}(\alpha)$ might not be unique, but we will soon show it is, because each weight space is one dimensional, so the choice of X and Y are effectively uniquely determined. We let $e_\alpha \in \mathfrak{g}_\alpha$, $f_\alpha \in \mathfrak{g}_{-\alpha}$, and $h_\alpha \in \mathfrak{h}$ be elements of $\mathfrak{sl}(\alpha)$ which map to e, f , and h in the isomorphism with $\mathfrak{sl}_2(K)$. We can let $e_\alpha = X$ in the last theorem, and f_α an appropriate scalar multiple of Y 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 4.24. *The choices of e_α , f_α , and h_α are unique up to inversely scaling e_α and f_α as λe_α , $\lambda^{-1} f_\alpha$, h_α for some non-zero $\lambda \in K$. To obtain a basis for $\mathfrak{sl}(-\alpha)$ from $\mathfrak{sl}(\alpha)$, swap e_α with f_α , and replacing h_α with $-h_\alpha$.*

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

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

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

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

Lemma 4.25. *If α is a root, let $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$. Then $[X, Y] = \kappa(X, Y)t_\alpha$. In particular, $h_\alpha = [e_\alpha, f_\alpha] \in \text{span}(t_\alpha)$.*

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

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

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

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

Lemma 4.26. *If V is a $\mathfrak{sl}(\alpha)$ submodule of \mathfrak{g} , then the eigenvalues of h_α acting on V are integers.*

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

Example. 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 $\alpha_{ij}(X) = X_{ii} - X_{jj}$,

$$\begin{aligned} e_{\alpha_{ij}} &= E_{ij} - E_{(j+m)(i+m)} & f_{\alpha_{ij}} &= E_{ji} - E_{(i+m)(j+m)} \\ h_{\alpha_{ij}} &= E_{ii} - E_{jj} + E_{(j+m)(j+m)} - E_{(i+m)(i+m)} \end{aligned}$$

- 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

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

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

Theorem 4.27. The root space of any non-zero α on a semisimple Lie algebra is one-dimensional, and the only multiples of α which are roots are $\pm\alpha$.

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

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

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

Theorem 4.28. 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_α on some element of \mathfrak{g}_β , hence it must be an integer by the classification of $\mathfrak{sl}_2(K)$ representations. If we consider the root string V , which is the direct sum of $\mathfrak{g}_{\beta+k\alpha}$ where k is chosen such that $\beta + k\alpha$ is a root, then we have $(\beta + k\alpha)(h_\alpha) = \beta(h_\alpha) + 2k$, thus the eigenvalues of h_α on the root string are either all even or all odd, and we therefore find that the root string is an irreducible $\mathfrak{sl}(\alpha)$ module, hence V is isomorphic to some V_n . On V_n , h_α acts diagonally with eigenvalues $n, n-2, \dots, -(n-2), -n$, and this must be paired up with $\beta(h_\alpha) + 2k$. Thus we may let $-n = \beta(h_\alpha) - 2a$, $n = \beta(h_\alpha) + 2b$, in which case $a - b = \beta(h_\alpha)$. This proves (ii). (iv) essentially follows from (ii) in the same manner. If $X \in \mathfrak{g}_\beta$, then X belongs to the h_α eigenspace of eigenvalue $\beta(h_\alpha)$, and if $[e_\alpha, e_\beta] = 0$, then e_β is the highest weight vector in the irreducible representation V isomorphic to V_n . If $\alpha + \beta$ is a root, then h_α has eigenvalue $(\alpha + \beta)(h_\alpha) = 2 + \beta(h_\alpha)$, hence e_β is not the highest weight vector, hence $[e_\alpha, e_\beta] \neq 0$, and this proves (iii). \square

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

4.5 Cartan Subalgebras as Inner Product Spaces

Lemma 4.29. *If $X \in \mathfrak{h}$ is non-zero, then there is a root α with $\alpha(X) \neq 0$, and therefore the roots span \mathfrak{h}^* .*

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

Lemma 4.30. 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_α by multiplying by $\kappa(e_\alpha, f_\alpha)$ on both sides of the equation

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

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

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

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

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

□

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

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

and the previous lemmas imply

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

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

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

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

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

Lemma 4.31. *If β is a root, then β is a linear combination of the α_i with rational coefficients.*

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

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

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

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

Theorem 4.32. *(\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 α , hence $t_{\lambda} = 0$, so $\lambda = 0$. Now the theorem follows for all elements of $E_{\mathbf{Q}}$, because if $(x \otimes \lambda, x \otimes \lambda) = 0$, then either $(\lambda, \lambda) = 0$, or $x^2 = 0$, in which case either $x = 0$ or $\lambda = 0$, and hence $x \otimes \lambda = 0$. \square

Chapter 5

Root systems

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

- $0 \notin \Phi$.
- The only scalar multiples of $\alpha \in \Phi$ which are in Φ are $\pm\alpha$.
- If $\alpha, \beta \in \Phi$, if we define $\langle \alpha, \beta \rangle = 2(\alpha, \beta)/(\beta, \beta)$, then

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

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

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

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

Example. In \mathbf{R}^n with the standard inner product, for $n \geq 2$, let R be the set of $e_i - e_j$, for $i \neq j$. Then R is a root system over the span of the $e_i - e_j$, which is

the set of vectors x such that $\sum x_i = 0$. The first and second properties of root systems are trivial, and

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

and fixing $\alpha = e_k - e_l$

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

Then

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

Thus R is a root system.

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

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 α_{ij} , and we find

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

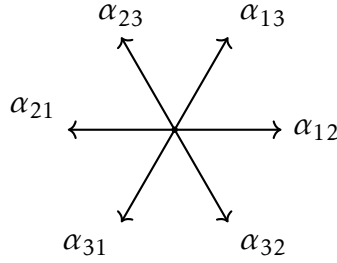
and if (ij) is the permutation of i and j , we find that

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

So, for instance, $s_{\alpha_{ij}}(\alpha_{ij}) = \alpha_{ji}$, and $s_{\alpha_{ij}}(\alpha_{ik}) = \alpha_{jk}$. The root system of $\mathfrak{sl}_2(K)$ can be isometrically embedded in one dimension,

$$\alpha_{21} \longleftrightarrow \alpha_{12}$$

and the root system of $\mathfrak{sl}_3(K)$ can be isometrically embedded in two dimensions.



From this diagram it is easy to see that the root strings of $\mathfrak{sl}_3(K)$ are two or three dimensional.

Example. On $\mathfrak{sp}_n(K)$, as on $\mathfrak{sl}_n(K)$, using the values for h_α , 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

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

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

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

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

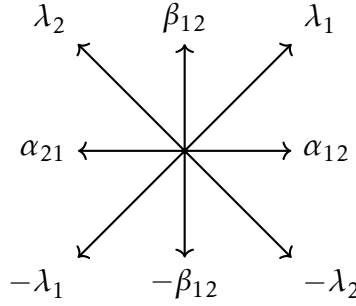
$$s_{\beta_{kl}}(\beta_{ki}) = s_{\beta_{kl}}(\beta_{ik}) = \alpha_{il} \quad s_{\beta_{kl}}(\beta_{kl}) = -\beta_{kl}$$

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

we only list the points which aren't fixed by the maps. The space $\mathfrak{sp}_2(K)$ is one dimensional, and can be isometrically embedded in one dimension.

$$-\lambda_1 \longleftrightarrow \lambda_1$$

This is the same root system as form $\mathfrak{sl}_2(K)$, and we actually find that $\mathfrak{sp}_2(K)$ is isomorphic to $\mathfrak{sl}_2(K)$. The space $\mathfrak{sp}_4(K)$ can be isometrically embedded in two dimensions.



It appears that the λ roots form a distinct class different than the α and β roots, because there is no way to get one set from the other, and indeed, these roots have a different length than the roots α and β .

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 $\mathfrak{o}_{2n+1}(K)$, we add the extra γ roots, and we find

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

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

$$\langle \gamma_i, \gamma_r \rangle = 2\delta_{ir}$$

Hence

$$s_{\gamma_r}(\alpha_{rj}) = s_{\gamma_r}(\alpha_{jr}) = \gamma_i \quad s_{\gamma_r}(\beta_{rj}) = s_{\gamma_r}(\beta_{jr}) = -\gamma_j$$

$$s_{\gamma_r}(\gamma_r) = -\gamma_r$$

$$s_{\alpha_{ij}}(\gamma_i) = \gamma_i - 2\alpha_{ij} = 2\gamma_j - \gamma_i$$

$\mathfrak{o}_2(K)$ has no roots, because it is actually an abelian Lie algebra, corresponding to the fact that $SO_2(\mathbf{C})$ and $SO_2(\mathbf{R})$ are commutative. $\mathfrak{o}_3(K)$ is embeddable in one dimensional space, with root diagram

$$-\gamma_1 \longleftrightarrow \bullet \longrightarrow \gamma_1$$

and we should expect that $\mathfrak{o}_3(K)$ is isomorphic to $\mathfrak{sl}_3(K)$ and $\mathfrak{sp}_2(K)$. $\mathfrak{o}_4(K)$ is embeddable in two dimensional space, with root diagram

$$\begin{array}{c} \alpha_{21} \\ \uparrow \\ \alpha_{21} \leftarrow \bullet \rightarrow \alpha_{12} \\ \downarrow \\ -\beta_{12} \end{array}$$

we see that $\mathfrak{o}_4(K)$ essentially contains two ‘subdiagrams’ $\{\alpha_{12}, \alpha_{21}\}$ and $\{\beta_{12}, -\beta_{12}\}$ which are orthogonal to one another, and this reflects the fact that $\mathfrak{o}_4(K)$ is isomorphic to $\mathfrak{sl}_2(K) \oplus \mathfrak{sl}_2(K)$. Then, $\mathfrak{o}_5(K)$ is embeddable in two dimensional space, with diagram

$$\begin{array}{c} \beta_{12} \\ \uparrow \\ \lambda_2 \swarrow \quad \nearrow \lambda_1 \\ \alpha_{21} \leftarrow \bullet \rightarrow \alpha_{12} \\ \swarrow -\lambda_1 \quad \searrow -\lambda_2 \\ \downarrow -\beta_{12} \end{array}$$

which is essentially the same root diagram of $\mathfrak{sp}_4(K)$ if we rotate the diagram by 90 degrees and scale up, and we can actually find an isomorphism between these Lie algebras.

It is easy to remember these identities, if we remember that the inner product is given by the standard Euclidean inner product induced by the

basis $\{E_{ij}^*\}$. We know that we have some orthonormal basis with respect to the induced inner product the roots, but in the case of the classical Lie algebras this orthonormal basis is just the dual basis to the $\{E_{ij}\}$.

Lemma 5.1. *If R is a root system, for two roots α and β , with $\beta \neq \pm\alpha$,*

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

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

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

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

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

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

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

Theorem 5.2. *If α and β are roots, and the angle between α and β is strictly obtuse, then $\alpha + \beta$ is a root. If the angle is strictly acute, and $(\beta, \beta) \geq (\alpha, \alpha)$, then $\alpha - \beta$ is a root.*

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

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

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

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

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

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

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

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

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

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

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

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

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

and this system is denoted G_2 .

In the case $A_1 \times A_1$, the perpendicular vectors don't act on one another. We say a root system R is reducible if it can be written as the disjoint union of two sets R_0 and R_1 , where vectors of R_1 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 5.3. *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 λ . If $\alpha, \beta \in R_i$, and $\beta \neq \pm\alpha$, then $s_\alpha(\beta)$ is in the plane spanned by α and β , and therefore cannot be perpendicular to both α and β 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 $c_\alpha \in \mathbf{Z}$ and such that all non-zero c_α 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 α and β 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 5.4. *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 α . Let R^+ consist of the roots α 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 β which minimizes the inner product (x, β) . 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 β_0 and β_1 can be written as the sum of positive elements. But this implies that β 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 α and β 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 α or β 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 $(\alpha, \beta) \leq 0$ (α and β 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 α . \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_α 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 α_{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 λ_n , since we may surely describe all α_{ij} in this manner, where α_{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 α_{ij} for $i < j$, and β_{ij} and λ_i .

Example. Over $\mathfrak{o}_{2n}(K)$, 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 α_{ij} for $i < j$, and β_{ij} . Over \mathfrak{o}_{2n+1} , we choose a basis of $\alpha_{12}, \dots, \alpha_{(n-1)n}$ and λ_n , for then we have all λ_i , and therefore also all $\beta_{ij} = \alpha_{ij} + 2\lambda_i$. The positive roots are α_{ij} for $i < j$, and λ_i and β_{ij} . When we add the additional roots in $\mathfrak{o}_{2n+1}(K)$, we replace $\beta_{(n-1)n}$ in the basis γ_n , and we find that the β_{ij} and γ_i are positive, and the α_{ij} for $i < j$.

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 λ 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, and for the root systems of the classical Lie algebras, we have fixed this basis. We call a root **simple** if it is in this basis.

5.1 The Weyl Group

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

Theorem 5.5. *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 (because these roots span the space), hence there can only be finitely many transformations in the Weyl group. \square

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 α is any root, then $\alpha = f(\beta)$ for some simple root β , 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 5.6. *If α is simple, s_α permutes all positive roots but α itself.*

Proof. Suppose β is a positive root other than α . Then $\beta = \sum c_\kappa \kappa$, where κ 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_κ , which is positive, hence $s_\alpha(\beta)$ is positive. \square

Theorem 5.7. *If β is any root, there is a Weyl group permutation $f \in W_0$ and a simple root α 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 β 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 γ , then $(\beta, \beta) = \sum c_\gamma (\beta, \gamma) \leq 0$, so $\beta = 0$, which is impossible. Thus there is λ with $(\beta, \lambda) > 0$, and so $s_\lambda(\beta)$ fixes the coefficients of all c_α with all $\alpha \neq \lambda$, and decreases the coefficient c_λ , hence by induction $s_\lambda(\beta) = f(\alpha)$ for some simple root α , and then $\beta = (s_\lambda \circ f)(\alpha)$. Now suppose that β is a negative root. Then $-\beta$ is positive, and hence $-\beta = f(\alpha)$ for some simple root α , and then $f(-\alpha) = \beta$. The reflection s_α maps α to $-\alpha$, hence $(f \circ s_\alpha)(\alpha) = \beta$. \square

Corollary 5.8. *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 β , then

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

and since

$$\begin{aligned}
\langle x, \alpha - \langle \alpha, \beta \rangle \beta \rangle &= \frac{\kappa(x, \alpha - \langle \alpha, \beta \rangle \beta)}{\kappa(\alpha - \langle \alpha, \beta \rangle \beta, \alpha - \langle \alpha, \beta \rangle \beta)} \\
&= \frac{[\langle x, \alpha \rangle - \langle \beta, \alpha \rangle \langle x, \beta \rangle] \kappa(\alpha, \alpha)}{\kappa(\alpha - \langle \alpha, \beta \rangle \beta, \alpha - \langle \alpha, \beta \rangle \beta)} \\
&= \frac{\langle x, \alpha \rangle - \langle x, \beta \rangle \langle \beta, \alpha \rangle}{1 - \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle}
\end{aligned}$$

$$\begin{aligned}
s_{s_\beta(\alpha)}(x) &= x - \langle x, \alpha - \langle \alpha, \beta \rangle \beta \rangle (\alpha - \langle \alpha, \beta \rangle \beta) \\
&= x - \frac{\langle x, \alpha \rangle - \langle x, \beta \rangle \langle \beta, \alpha \rangle}{1 - \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle} (\alpha - \langle \alpha, \beta \rangle \beta)
\end{aligned}$$

TODO: FIX THIS CALCULATION. The theorem then follows by letting the elements s_α generate W . Since the Weyl group is generated by all reflections, it suffices to show that each s_α is a product of simple reflections, even if α is not simple. But there is a simple root β such that $\alpha = f(\beta)$, and then $s_\alpha = f \circ s_\beta \circ f^{-1}$. \square

Example. Over $\mathfrak{sl}_n(K)$, the Weyl group is equal to the entire symmetric group over the roots, because $s_{\alpha_{ij}}$ acts as the transposition (ij) on the indices of the α_{ij} , hence the Weyl group can be used to permute the indices of the α_{ij} in an arbitrary way, since transpositions generate the symmetric group.

Example. If $n = 2m$, then the action of the $s_{\alpha_{ij}}$ acts as a permutation of the indices of the roots of $\mathfrak{sp}_n(K)$, so the Weyl group contains a subgroup isomorphic to S_m .

A **Weyl chamber** is a connected component in the set of all connected components of V when we remove all points in a hyperplane perpendicular to the roots.

Theorem 5.9. *For any two Weyl chambers, there is an element of the Weyl group mapping one chamber to the other.*

Proof. TODO □

An element $\lambda \in V$ is known as **dominant** if $(\lambda, \alpha) \geq 0$ for all positive roots α . That is, λ lies within an angle of $\pm\pi/2$ from each root α . We shall find that dominant roots are essentially to characterizing the irreducible representations of arbitrary semisimple Lie algebras.

Theorem 5.10. *For any element $\lambda \in V$, there is an element f of the Weyl group such that $f(\lambda)$ is dominant.*

Proof. TODO □

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

Theorem 5.11. *If B and B' are any two bases, then there is a transformation $f \in W$ with $f(B) = B'$.*

Proof. TODO □

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 β , $\langle s_\beta(\alpha_i), s_\beta(\alpha_j) \rangle = \langle \alpha_i, \alpha_j \rangle$, 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 Δ , 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 α_i and α_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**.

Example. The Dynkin diagram for $\mathfrak{sl}_2(K)$, $\mathfrak{o}_3(K)$, and $\mathfrak{sp}_2(K)$ consists of a single node, with no edges. The Dynkin diagram for $\mathfrak{sl}_3(K)$ is



The Dynkin diagram for $\mathfrak{sp}_4(K)$ and $\mathfrak{o}_5(K)$ is



The Dynkin diagram for $\mathfrak{o}_4(K)$ is just the graph on two vertices with no nodes.



which has two components reflecting the fact that the algebra can be decomposed into the direct sum of two simple Lie algebras.

Example. We calculated that on $\mathfrak{sl}_n(K)$,

$$\langle \alpha_{i(i+1)}, \alpha_{j(j+1)} \rangle = \alpha_{i(i+1)}(H_j) = \begin{cases} 2 & i = j \\ -1 & |i - j| = 1 \\ 0 & \text{elsewise} \end{cases}$$

Thus the Dynkin diagram of $\mathfrak{sl}_n(K)$ is

Example. On $\mathfrak{sp}_n(K)$, $\langle \alpha_{i(i+1)}, \alpha_{j(j+1)} \rangle$ behaves as in $\mathfrak{sl}_m(K)$, and

$$\langle \alpha_{i(i+1)}, \lambda_m \rangle = \alpha_{i(i+1)}(E_{mm} - E_{(2m)(2m)}) = -\delta_{im}$$

$$\langle \lambda_m, \alpha_{i(i+1)} \rangle = \lambda_m(E_{ii} - E_{(i+1)(i+1)} + E_{(i+1+m)(i+1+m)} - E_{(i+m)(i+m)}) = -2\delta_{m(i+1)}$$

Thus the Dynkin diagram of $\mathfrak{sp}_n(K)$ is

Example. On \mathfrak{o}_{2m} , $\langle \alpha_{i(i+1)}, \alpha_{j(j+1)} \rangle$ remains the same, and we find

$$\langle \alpha_{i(i+1)}, \beta_{(m-1)m} \rangle = \alpha_{i(i+1)}(E_{(m-1)(m-1)} + E_{mm}) = -\delta_{i(m-2)}$$

$$\langle \beta_{(m-1)m}, \alpha_{i(i+1)} \rangle = \beta_{(m-1)m}(E_{ii} - E_{(i+1)(i+1)}) = -\delta_{i(m-2)}$$

so the Dynkin diagram of $\mathfrak{so}_{2m}(K)$ is except for the edge case $\mathfrak{so}_4(K)$, in which the Dynkin diagram is just two disconnected nodes. On $\mathfrak{o}_{2m+1}(K)$, we add the additional roots γ_i , and swap out the $\beta_{(m-1)m}$ in the basis for γ_m , in which case we find

$$\langle \alpha_{i(i+1)}, \gamma_m \rangle = 2\alpha_{i(i+1)}(E_{mm}) = -2\delta_{i(m-1)}$$

$$\langle \gamma_m, \alpha_{i(i+1)} \rangle = \gamma_m(E_{ii} - E_{(i+1)(i+1)}) = -\delta_{i(m-1)}$$

and therefore the Dynkin diagram looks like

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 5.12. 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 α_i is larger than α_j , then β_i is larger than β_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_{α_i} generate the Weyl group of R , for any v , $f(W_R v) \subset W_T f(v)$, hence in particular if α_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

Thus we have reduced the classification of root systems to the classifications of the possible Dynkin diagrams we can obtain from root systems. Indeed, it is much easier to read off certain information from the diagram than from the inner product specification of the system. For instance, the graph is connected if and only if the root system is irreducible.

5.2 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 classify these ‘simpler’ root systems, which is exactly the family of root systems whose Dynkin diagrams are connected. We will find that classifying the Dynkin diagrams are much simpler than classifying the root systems from the ground up.

First, let us try and classify the graphs obtained by ‘forgetting the arrows’ in Dynkin diagrams. These correspond to a slightly more general family of vector sets than root systems. Define a subset X of linearly independent vectors in an inner product space to be **admissible** if $\|x\| = 1$ for all $x \in X$, $(x, y) \leq 0$ if $x \neq y$, with $4(x, y)^2 \in \{0, 1, 2, 3\}$ for $x \neq y$. Given an admissible set, we associate the graph Γ whose vertices are the points in X , and we draw $D_{xy} = 4(x, y)^2$ edges between x and y for $x \neq y$. Given a root system, we obtain an admissible set by normalizing all the vectors in some basis, and the difference between the Dynkin diagram and the graph from the admissible set is just the removal of the arrows.

Lemma 5.13. *The number of pairs of vertices joined in Γ is at most $|X| - 1$.*

Proof. If $X = \{x_1, \dots, x_n\}$, let $v = \sum x_i$. Then $v \neq 0$, because the x_i are linearly independent, and so

$$(v, v) = n + 2 \sum_{i < j} (x_i, x_j) > 0$$

and so if there are N pair with $(x_i, x_j) \neq 0$, then

$$n > -2 \sum_{i < j} (x_i, x_j) > \sum_{i < j} \sqrt{D_{x_i x_j}} \geq N$$

and this completes the proof. \square

Corollary 5.14. *No graph Γ obtained from an admissible set has a cycle.*

Proof. Since every subset of an admissible set is an admissible set, a cycle of n vectors must form an admissible set, and therefore there must be at most $n - 1$ edges between the vectors in this set, hence such a cycle cannot exist. \square

Thus every graph obtained from an admissible set must have a tree structure. But we can weaken this condition more.

Lemma 5.15. *No vertex of Γ is incident to four or more edges.*

Proof. Fix some vertex $y \in \Gamma$, and suppose y is attached to x_1, \dots, x_n . The vectors x_1, \dots, x_n must be orthogonal, because an admissible set cannot have a cycle. Consider a vector x_0 such that $\{x_0, \dots, x_n\}$ is the span of the vectors x_i and y . If $y = \sum (y, x_i) x_i$, then $\sum (y, x_i)^2 = 1$, since $\|y\| = 1$. Since $(y, x_0)^2 > 0$, this means that

$$\sum_{i=1}^n (y, x_i)^2 < 1$$

But $(y, x_i)^2 \geq 1/4$ for all i , so $n \leq 3$. \square

We conclude that if Γ is a connected graph with a vertex of degree three, then Γ is isomorphic to the two vertex graph. These small graphs are particularly easy to understand, and we have a lemma which allows us to focus our analysis on ‘irreducible’ graphs.

Lemma 5.16. *Suppose Γ has a subgraph x_1, \dots, x_n which forms a line where there are no multiple edges. If we remove x_1, \dots, x_n from the admissible set, and add the vector $\sum x_i$, then the new set is admissible, and the graph Γ' is obtained from the original graph by contracting the line in the graph.*

Proof. By assumption, $2(x_i, x_{i+1}) = -1$. and $(x_i, x_j) = 0$ otherwise, so

$$\left(\sum x_i, \sum x_i \right) = \sum (x_i, x_j) = \sum (x_i, x_i) + 2 \sum (x_i, x_{i+1}) = n - (n - 1) = 1$$

For any over vertex v in the graph, the vertex can only be connected to at most one of the vertices in the x_i , and so $(v, \sum x_i) = (v, x_i)$, so we conclude that the integrality properties still hold. \square

A **branch vertex** is a vertex incident to three vertices. We know that no vertex can be incident to more than three vertices.

Lemma 5.17. *Γ has at most one branch vertex, no more than one double edge, and cannot have both a branch vertex and a double edge.*

Proof. Suppose a graph has two double edges x and y . Then there is a line going from x to y . If this forms a line of singly connected vertices, and we can contract this line to obtain an admissible graph with a vertex of degree four, which is impossible. If there is a branch vertex on this line, then we can contract to find a vertex with two singly connected incident vertices, and one doubly connected vertex, which is impossible. Similarly, we can contract the line between two branch vertices to obtain a contradiction. \square

Lemma 5.18. *If Γ has a double edge, it is either of the form or*

Proof. We require a small calculation. If we have a line of vertices x_1, \dots, x_n , and we let $y = \sum kx_k$, then $(y, y) = n(n+1)/2$, because

$$(y, y) = \sum_{i,j=1}^n ij(x_i, x_j) = \sum_{k=1}^n k^2 - \sum_{k=1}^{n-1} k(k+1) = n^2 - \sum_{k=1}^{n-1} k = \frac{n(n+1)}{2}$$

Now to do the actual work of the proof. We know that since we cannot have a triple branch or another double edge, that any graph Γ has the form where the vertices are labelled $x_1, \dots, x_n, y_m, \dots, y_1$. Let $v = \sum kx_k$, and $w = \sum ky_k$. Then $(v, v) = n(n+1)/2$, and $(w, w) = m(m+1)/2$. Now $(x_n, y_m) = -1/\sqrt{2}$, and $(x_i, y_j) = 0$ for all i, j . Thus

$$(v, w) = \sum ij(v_i, w_j) = -nm/\sqrt{2}$$

Since v and w are linearly independent, Cauchy Schwarz implies that $(v, w) < (v, v)(w, w)$, and this gives the inequality $2nm < (n+1)(m+1)$, and therefore

$$(n-1)(m-1) = nm - nm + 1 < 2$$

Thus either $n = 1$ or $n = m = 2$. \square

Lemma 5.19. *If Γ has a triple branch, then Γ is either or the graph is one of E_6 , E_7 , or E_8*

Proof. We know that the graph breaks into for some vertices x_1, \dots, x_n , y_m, \dots, y_1 , and z_k, \dots, z_1 connected at a central vertex a , and we assume $n \geq m \geq k$. It suffices to show that either $m = k = 1$, or $m = 2, k = 1$, and $n \leq 4$. As in the last lemma, let $v = \sum kx_k$, $w = \sum ky_k$, and $u = \sum kz_k$. Then v , w , and u are orthogonal to one another, and can be normalized to \hat{v} , \hat{w} , \hat{u} , and we can consider another orthogonal vector a_0 such that these vectors span the span of v, w, u , and a . If we write $a = (a, \hat{v})\hat{v} + (a, \hat{w})\hat{w} + (a, \hat{u})\hat{u} + (a, a_0)a_0$. Then since $(a, a_0) \neq 0$ we conclude that

$$(a, \hat{v})^2 + (a, \hat{w})^2 + (a, \hat{u})^2 < 1$$

We know the lengths of v, w , and u , and we calculate that $(a, v)^2 = n^2/4$, $(a, w)^2 = m^2/4$, and $(a, u)^2 = k^2/4$, so we find that

$$\frac{2n^2}{4n(n+1)} + \frac{2m^2}{4m(m+1)} + \frac{2k^2}{4k(k+1)} < 1$$

and this reduces to

$$\frac{1}{n+1} + \frac{1}{m+1} + \frac{1}{k+1} > 1$$

This implies that $3/(k+1) > 1$, hence $k < 2$, so we actually must have $k = 1$. Then this argument repeats to show $m < 3$, so $m = 1$ or $m = 2$. If $m = 2$, then this gives $n < 5$, giving us the excentric diagrams, or if $m = 1$, we have no restriction on n , giving us the other graphs. \square

This is a sufficient classification of the diagrams for admissible sets. All that remains is to determine ‘which way the arrows go’, to obtain all possible Dynkin diagrams. Without loss of generality, the two vertices connected by a triple vertex can be oriented one way or another (one must vertex must be longer than the other by our discussion of root systems). The order of the arrow given to the one double edge in gives two different root systems, and the choice of orientation in is arbitrary.

Chapter 6

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 6.1. *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 Φ_0 is some subset of Φ , 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 Φ_0 and Φ_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 Φ into two perpendicular root systems,

so either $\Phi_0 = \emptyset$ or $\Phi_1 = \emptyset$. If Φ_1 is empty, then all the root spaces \mathfrak{g}_α are contained in \mathfrak{a} , but since \mathfrak{g} is generated by its root spaces we conclude $\mathfrak{a} = \mathfrak{g}$. If $\Phi_0 = \emptyset$, then $\mathfrak{a} \subset \mathfrak{h}$, but then \mathfrak{a} is a solvable ideal in a semisimple Lie algebra, so $\mathfrak{a} = 0$. \square

Corollary 6.2. *All the classical Lie algebras are simple, except for the following ‘degenerate’ cases: $\mathfrak{sl}_1(K)$, $\mathfrak{o}_1(K)$, and $\mathfrak{o}_2(K)$, which are all commutative, and $\mathfrak{o}_4(K)$, which is isomorphic to $\mathfrak{o}_2(K) \oplus \mathfrak{o}_2(K)$.*

Our first result we can obtain from the simplicity of the classical Lie algebras is a way to determine their Killing form. First note that the symmetric bilinear form $\beta(X, Y) = \text{tr}(XY)$ is nondegenerate on any simple linear Lie algebra, because, since the trace is associative, the set of X such that $\beta(X, Y) = 0$ for all Y in the Lie algebra forms an ideal, and since $\beta \neq 0$, the ideal must be trivial. Now we have shown that there is a unique non-degenerate symmetric, associative bilinear form on a simple Lie algebra, so it follows that β is some multiple of the Killing form on each of the corresponding Lie algebras. This is easy to prove on a case by case basis, by plugging in some easy non-orthogonal matrices into the Killing form.

Example. On $\mathfrak{sl}_n(K)$, given diagonal matrices X, Y , we find that

$$\kappa(X, Y) = \text{tr}(\text{adj}_X \text{adj}_Y) = \sum [X, [Y, E_{ij}]]_{ij} = \sum (X_{ii} - X_{jj})(Y_{ii} - Y_{jj})$$

If we set $X_{11} = 1$, $X_{22} = -1$, and $X_{ii} = 0$ for all $i > 2$, then we find $\kappa(X, X) = 2[4 + 2(n-2)] = 4(n+1)$, and $\text{tr}(X^2) = 2$, we find that the trace on $\mathfrak{sl}_n(K)$ can be expressed as

$$\kappa(X, Y) = 2(n+1)\text{tr}(XY) = \sum X_{ii}Y_{ii}$$

Similarly, we find that on \mathfrak{o}_{2n} , $\kappa(X, Y) = 2(n-1)\text{tr}(XY)$, on \mathfrak{o}_{2n+1} , $\kappa(X, Y) = (2n-1)\text{tr}(XY)$, and on \mathfrak{sp}_{2n} , $\kappa(X, Y) = 2(n+1)\text{tr}(XY)$.

Next, we justify that all we have done so far is universal. If we choose a different base in the root system, we obtain the same Dynkin diagram, so this specification is fine. However, we haven’t argued that choosing a different Cartan subalgebra will change the theory.

Theorem 6.3. *If \mathfrak{g} is a semisimple Lie algebra with two Cartan subalgebras \mathfrak{h}_0 and \mathfrak{h}_1 , with associated root systems Φ_0 and Φ_1 , then Φ_0 is isomorphic to Φ_1 .*

Proof. TODO. □

Thus two Lie algebras can only be isomorphic if their associated root systems are isomorphic. This shows that the classical Lie algebras are all non isomorphic, except for the exceptional isomorphisms

- $\mathfrak{so}_3(K) \cong \mathfrak{sp}_2(K) \cong \mathfrak{sl}_2(K)$
- $\mathfrak{so}_4(K) \cong \mathfrak{sl}_2(K) \oplus \mathfrak{sl}_2(K)$
- $\mathfrak{so}_5(K) \cong \mathfrak{sp}_4(K)$
- $\mathfrak{so}_6(K) \cong \mathfrak{sl}_4(K)$

We have not explicitly proved these correspondences. The first two could be done as exercise, but it is a theorem of Serre that if two semisimple Lie algebras have isomorphic root systems (which we know for these Lie algebras), then they must necessarily be isomorphic.

Chapter 7

Advanced Representation Theory

The root space decomposition of a Lie algebra \mathfrak{g} into its root spaces with respect to its Cartan subalgebra \mathfrak{h} is a powerful way to classify the adjoint representation of \mathfrak{g} on itself. With a little generalization, we can use the structural results about the root space decomposition to obtain results about arbitrary representations of semisimple Lie algebras.

If V is any representation of a semisimple Lie algebra \mathfrak{g} , it can also be considered a representation of \mathfrak{h} . Since \mathfrak{h} is abelian, the elements of the Cartan subalgebra act commutatively on V . Since \mathfrak{g} is semisimple, each element of \mathfrak{h} acts diagonalizably on V , and they are *simultaneously diagonalizable*. Thus V decomposes into weight spaces for the representation of V with respect to \mathfrak{h} . Using the representation theory of $\mathfrak{sl}_2(K)$, for any root α , we can also consider V as a representation of $\mathfrak{sl}_2(K)$, because the subalgebra of \mathfrak{g} spanned by e_α , f_α , and h_α is isomorphic. In particular, we find that $\lambda(h_\alpha) \in \mathbf{Z}$. If $x \in V_\alpha$, then as should be expected, $h_\alpha x = \lambda(h_\alpha)x$, and also $e_\alpha x \in V_{\lambda+\alpha}$ and $f_\alpha x \in V_{\lambda-\alpha}$, because if $X \in \mathfrak{h}$, then

$$\begin{aligned} X(e_\alpha x) &= \lambda(X)(e_\alpha x) + [X, e_\alpha]x = \lambda(X)(e_\alpha x) + \alpha(X)(e_\alpha x) \\ X(f_\alpha x) &= \lambda(X)(f_\alpha x) + [X, f_\alpha]x = \lambda(X)(f_\alpha x) - \alpha(X)(f_\alpha x) \end{aligned}$$

Since V is finite dimensional, the set of weights is finite, and since the set of roots is finite, there must exist a weight λ such that for any *positive* root α , $\alpha + \lambda$ is not a weight of V . We call λ a **highest weight**, and an eigenvector for λ a **highest weight vector**. This generalizes the definition of highest weights we gave for representations of $\mathfrak{sl}_2(K)$, where roots and weights are just eigenvalues for the one dimensional Cartan subalgebra.

bra generated by h , and the only positive root acts on the eigenspaces by shifting the eigenvalue up by a value of 2.

Lemma 7.1. *If V is a irreducible representation of a Lie algebra \mathfrak{g} , then there is a unique highest weight λ , and V_λ is one dimensional. All other weights of V are equal to $\lambda - \sum_{\alpha \in \Phi^+} n_\alpha \alpha$ for some non-negative integers n_α .*

Proof. We know that some highest weight λ exists. Let $x \in V_\lambda$ be a nonzero vector, and let W denote the subspace of V spanned by all vectors of the form $f_{\alpha_1} \dots f_{\alpha_n} x$, where $\alpha_1, \dots, \alpha_n$ are positive roots of \mathfrak{g} . It suffices to verify this when e_α, f_α , or h_α are applied to these vectors, for some vector α , and the result then lies in W . This is clear for f_α and h_α . We show that $e_\alpha f_{\alpha_1} \dots f_{\alpha_n} x \in W$ by induction on n . For $n = 0$, we find $e_\alpha x = 0$, since $\lambda + \alpha$ is not a weight of V . In general, we find

$$e_\alpha f_{\alpha_1} \dots f_{\alpha_n} x = f_{\alpha_1} (e_\alpha f_{\alpha_2} \dots f_{\alpha_n} x) + [e_\alpha, f_{\alpha_1}] (f_{\alpha_2} \dots f_{\alpha_n} x)$$

The first element is in W by induction. The second follows because $[e_\alpha, f_{\alpha_1}]$ is a scalar multiple of $e_{\alpha+\alpha_1}$, and $\alpha + \alpha_1$ is positive if α and α_1 are positive, in which we can apply induction again. Thus $V = W$, and this implies that the weights are obtained by subtracting from the highest weight by multiples of positive roots. \square

If λ is a weight of any representation V , and $\lambda(h_\alpha) < 0$, then viewing V as a representation of $\mathfrak{sl}(\alpha)$, we find that if $x \in V_\lambda$, then $e_\alpha x \neq 0$. Thus if λ is a highest weight vector for V , then $\lambda \geq 0$.

Theorem 7.2. *For any $\lambda \in \mathfrak{h}^*$ such that $\lambda(h_\alpha) \in \mathbb{Z}$ for all α , and $\lambda \geq 0$, then there is a unique finite dimensional irreducible representation V where λ is a highest weight, and any irreducible representation can be constructed this way.*

Proof. TODO \square

The elements of \mathfrak{h}^* of this form are called the **integral weights** of the Lie algebra \mathfrak{g} . If we define the **fundamental dominant weights** $\lambda_1, \dots, \lambda_n$ such that $\lambda_i(H_{\alpha_j}) = \delta_{ij}$, then the integral weights are exactly those in the integral cone generated by the fundamental dominant weights. If we write $\lambda_i = \sum c_{ik} \alpha_k$ then $\lambda_i(h_{\alpha_j}) = \sum c_{ik} \langle \alpha_k, \alpha_j \rangle$, so the coefficients c_{ik} are obtained by inverting the Cartan matrix corresponding to the group \mathfrak{g} .

7.1 Universal Enveloping Algebra

Thus all that remains is to be able to provide a constructive procedure to generate all irreducible representations of a given semisimple Lie algebra. This is provided by the **universal enveloping algebra**. Given a Lie algebra \mathfrak{g} , we first form the tensor algebra $\bigotimes \mathfrak{g}$ generated by the vector space structure of \mathfrak{g} . In order to connect the tensor product operation on $\bigotimes \mathfrak{g}$ with the Lie bracket operation, we quotient the tensor algebra by adding the relation $X \otimes Y - Y \otimes X - [X, Y]$. The resulting algebra is called the **universal enveloping algebra** $\mathfrak{U}\mathfrak{g}$, which is obtained by quotienting $T\mathfrak{g}$ by the two sided ideal generated by $X \otimes Y - Y \otimes X - [X, Y]$. The embedding of \mathfrak{g} in $T\mathfrak{g}$ induces an embedding of \mathfrak{g} in $\mathfrak{U}\mathfrak{g}$. There is a nice universal property which tells us why the enveloping algebra is useful.

Theorem 7.3. *If A is an associative K algebra, then any Lie algebra homomorphism $f : \mathfrak{g} \rightarrow A$ extends to a unique algebra homomorphism $f : \mathfrak{U}\mathfrak{g} \rightarrow A$.*

Proof. Using the universal property of the tensor product, we can extend f to an algebra homomorphism from $\bigotimes \mathfrak{g} \rightarrow A$ by letting $f(X \otimes Y) = f(X)f(Y)$. Now we find that

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

so f descends to a map on $\mathfrak{U}\mathfrak{g}$. But since the elements of \mathfrak{g} generate the algebra, this map must be unique. \square

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

Example. *Given the Lie algebra K , every element of the tensor product $\bigotimes K$ can be written uniquely as*

$$a_0 + a_1(1 \otimes 1) + \cdots + a_n(1 \otimes 1 \otimes \cdots \otimes 1)$$

where $a_0, \dots, a_n \in K$, so $\bigotimes K$ is isomorphic to $K[X]$. The ideal used to obtain $\mathfrak{U}K$ is generated by elements of the form

$$(X \otimes Y) - (Y \otimes X) - [X, Y] = XY(1 \otimes 1) - YX(1 \otimes 1) = 0$$

so the ideal is trivial and the universal enveloping algebra of K is K . More generally, the universal enveloping algebra of the abelian Lie algebra structure on K^n is just the polynomial ring $K[X_1, \dots, X_n]$.

Example. Consider the Lie algebra $\mathfrak{sl}_2(K)$ with its basis e, f, h . The relations in the ideal generated the universal enveloping algebra are of the form

$$e \otimes f - f \otimes e = h \quad e \otimes h - h \otimes e = -2e \quad f \otimes h - h \otimes f = 2f$$

By induction on the number of terms in some monomial $X_1 \otimes \dots \otimes X_n$, we can use these relations to show that the set of monomials of the form $e^{\otimes n} f^{\otimes m} h^{\otimes k}$ span the space. However, it is much more difficult to show these monomials are independent.

The Poincare-Birkhoff-Witt theorem generates a standard basis for $\mathfrak{U}\mathfrak{g}$ from a basis of \mathfrak{g} . The theorem is quite difficult, and we won't prove it here.

Theorem 7.4. Let \mathfrak{g} have an ordered basis $\{X_0, X_1, \dots, X_n\}$. Then $\mathfrak{U}\mathfrak{g}$ has a vector space basis consisting of monomials of the form

$$X_{i_1}^{j_1} \dots X_{i_n}^{j_n}$$

for some positive indices j_i , and monotone $i_1 < \dots < i_n$.

A simple corollary of the PBW theorem is that the embedding of \mathfrak{g} in $\mathfrak{U}\mathfrak{g}$ is injective. This isn't even easy to see from the definition, since the two-sided ideal upon which we take the quotient isn't a very easily understood object.

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

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

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

This is a representation, since

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

7.2 Verma Modules

The advantage of the theory of modules over an algebra is that we have an easy way to find a family of modules over a given algebra, because we can take quotients of left ideals in an algebra to obtain left modules. If \mathfrak{g} has a root system Φ , and given some basis $\alpha_1, \dots, \alpha_n$ of the root system, $h_{\alpha_1}, \dots, h_{\alpha_n}$ is a basis of \mathfrak{h} . For each linear function $\lambda \in \mathfrak{h}^*$, let $I(\lambda)$ be the left ideal of $\mathfrak{U}\mathfrak{g}$ generated by the elements e_α for $\alpha \in \Phi$, and $h\alpha_i - \lambda(h_{\alpha_i})$. We define the Verma module associated with the weight λ to be the module $M(\lambda) = \mathfrak{U}\mathfrak{g}/I(\lambda)$.

Theorem 7.5.

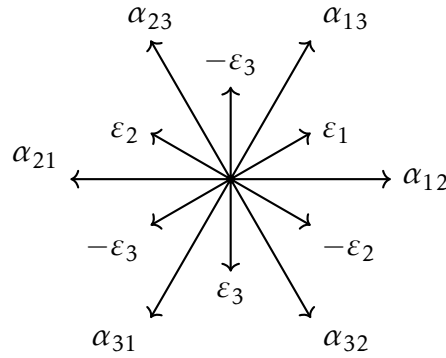
7.3 Representations of $\mathfrak{sl}_3(K)$

The representation theory of the Lie algebra $\mathfrak{sl}_3(K)$ is useful to understand, because it is still feasible to determine all possible representations of the algebra, though $\mathfrak{sl}_3(K)$ has the complexities present in the analysis of the representation theory of the more general class of semisimple Lie algebras. In this section, we will denote α_{12} by α , and α_{23} by β . At

times, it will be useful to use the fact that \mathfrak{h}^* is spanned by $\varepsilon_1(X) = X_{11}$, $\varepsilon_2(X) = X_{22}$, and $\varepsilon_3(X) = X_{33}$, though this is not an independent set. We will also use the basis $\{\alpha, \beta\}$ which generates the positive weights. To find ε_i with respect to the basis of α_{ij} , we note that $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0$, so

$$\begin{aligned}\varepsilon_1 &= (1/3)(2\varepsilon_1 - \varepsilon_2 - \varepsilon_3) = (1/3)(2\alpha + \beta) \\ \varepsilon_2 &= (1/3)(2\varepsilon_2 - \varepsilon_1 - \varepsilon_3) = (1/3)(\beta - \alpha) \\ \varepsilon_3 &= (1/3)(2\varepsilon_3 - \varepsilon_1 - \varepsilon_2) = (1/3)(-\alpha - 2\beta)\end{aligned}$$

In terms of the root diagram, we can see that the ε_i fit in geometrically as



We find that the fundamental dominant weights for $\mathfrak{sl}_3(K)$ is $\{\varepsilon_1, -\varepsilon_3\}$, and the cone generated by these weights gives all irreducible representations of $\mathfrak{sl}_3(K)$.

Example. The standard representation of $\mathfrak{sl}_3(K)$ on K^3 induces weights on \mathfrak{h} , and we find that if X is a diagonal matrix, then

$$Xe_1 = X_{11}e_1 \quad Xe_2 = X_{22}e_2 \quad Xe_3 = X_{33}e_3$$

Hence the weights are $\varepsilon_1, \varepsilon_2$, and ε_3 . A highest weight vector for this representation is e_1 , and we find the standard representation is irreducible.

On the dual representation $(K^3)^*$, we find

$$Xe_1^* = -X_{11}e_1^* \quad Xe_2^* = -X_{22}e_2^* \quad Xe_3^* = -X_{33}e_3^*$$

hence the weights of $(K^3)^*$ are $-\varepsilon_1, -\varepsilon_2, -\varepsilon_3$. This shows that K^3 cannot be isomorphic to $(K^3)^*$ as representations, for an isomorphism of representations preserves the weight vectors on the algebra.

Example. *The adjoint representation of $\mathfrak{sl}_3(K)$ on itself gives the decomposition into weight spaces. Each root has multiplicity one, and zero has multiplicity two.*

A **highest weight vector** for an $\mathfrak{sl}_3(K)$ representation V is a weight vector v such that $E_{ij}(v) = 0$ for $i < j$. The weight of v is then called a highest weight vector. A weight μ is **dominant** if $(\mu, \alpha_{12}), (\mu, \alpha_{23}) \geq 0$. This means exactly that the angle between μ lies less than a right angle between α_{12} and α_{23} .

Lemma 7.6. *If V is finite dimensional, a highest weight μ is dominant and integral.*

Proof. k

□

Chapter 8

Verma Modules

Bibliography

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