Classification of Semisimple Lie Algebras

John Austin Charters

Advisor: Dennis Snow
Co-advisor: Antonio Delgado

Contents

1 Introduction

My thesis describes the theory of Lie groups and Lie algebras, named after the Norwegian mathematician Sophus Lie. Lie was interested in groups whose elements depend smoothly on continuous parameters. His work principally focused on transformation groups, differential equations, and differential geometry. I will focus instead on the algebraic theory.

The approach to learning more about Lie groups is to study the linearization of the group at the identity. Such a linearization is called the Lie algebra associated to the Lie group. It is far easier to analyze the algebra, as it takes the structure of a vector space. I will then explain what it means for a Lie algebra to contain a semisimple Lie algebra. The semisimple components are described using geometric structures called root systems, whose classification was completed by the French mathematician Élie Cartan. I will introduce root systems and describe the details of the classification. The principal result of my thesis is the list of diagrams on page 31.

The theory of continuous groups has many applications to physics and other areas of mathematics. I conclude with an introduction to the Lorentz group and the Lorentz algebra, which arise in physics.

2 Lie groups

Definition 1. *A **Lie group** is a smooth manifold G endowed with a group structure, such that the group operation and the inverse map are smooth.*

In general, Lie groups can be defined abstractly. For simplicity, we will only consider **matrix Lie groups**, which are Lie groups that can be expressed as matrices. Denote the set of $n \times n$ matrices with complex entries by $M(n;\mathbb{C})$. The *general linear group* $GL(n;\mathbb{C})$ is the subset of $M(n;\mathbb{C})$ consisting of $n \times n$ invertible matrices. That is, $GL(n;\mathbb{C}) = \{X \in M(n;\mathbb{C}) : \det X \neq 0\}$. To verify that $GL(n;\mathbb{C})$ is a group, it follows from the identity $\det XY = \det X \det Y$ that $GL(n;\mathbb{C})$ is closed under multiplication and inversion. Certainly $GL(n;\mathbb{C})$ contains the identity, and matrix multiplication is always associative. Furthermore, the operations of multiplication and inversion are given by linear functions, so they are smooth. To see why $GL(n;\mathbb{C})$ is also a manifold, first observe that the subset of matrices in $M(n,\mathbb{C})$ with vanishing determinant is closed in $M(n;\mathbb{C}) \simeq \mathbb{C}^{n^2}$. This implies that $GL(n;\mathbb{C})$ is open in $M(n;\mathbb{C})$ and thus it inherits a manifold structure from $\mathbb{C}^{n^2}$. Therefore, $GL(n;\mathbb{C})$ is a Lie group.

There is a powerful result that any closed subgroup of $GL(n;\mathbb{C})$ is a manifold, hence a Lie group. For a proof, see [3]. So all of the matrix Lie groups we will consider are to be thought of as closed subgroups of $GL(n;\mathbb{C})$. For example, consider the *special linear group* $SL(n;\mathbb{C})$, defined as the subset $\{X \in GL(n;\mathbb{C}) : \det X = 1\}$. It is easy to verify that $SL(n;\mathbb{C})$ is a group. Also, $SL(n;\mathbb{C})$ is closed because the determinant is a continuous function, so any convergent sequence in $SL(n;\mathbb{C})$ must converge inside $SL(n;\mathbb{C})$. Therefore, $SL(n;\mathbb{C})$ is a matrix Lie group.

Another example of a matrix Lie group is the *orthogonal group* $O(n)$, which is the set of real matrices that preserve the standard inner product on $\mathbb{R}^n$, hence preserving lengths and relative angles of vectors. This yields the condition $O(n) = \{X \in GL(n;\mathbb{R}) : X^T = X^{-1}\}$. Furthermore, the *special orthogonal group* $SO(n)$ is the subset of $O(n)$ with determinant unity, i.e. $SO(n) = \{X \in O(n) : \det X = 1\}$. It is a matrix Lie group. The group $SO(3)$ is generated by the following 3×3 matrices that rotate $\mathbb{R}^3$ about the standard coordinate axes:

$$R_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},$$

$$R_2(\phi) = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix},$$

$$R_3(\psi) = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We would like to examine what $SO(3)$ looks like near the identity. By expanding the rotation matrices up to first order only, say $R_j(\theta) = I + i\theta\tau^j$, we easily discover the following matrices

$$\tau^1 = i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \tau^2 = i\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \tau^3 = i\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let $[X, Y]$ denote the **commutator** $XY - YX$. Then the τ^j satisfy the commutation relation $[\tau^i, \tau^j] = i\epsilon_{ijk}\tau^k$. These matrices are called the *infinitesimal generators* of $SO(3)$, as they are said to generate the group as follows. For any $n \times n$ matrix X, we define the **matrix exponential** of X as the power series

$$e^X \equiv \sum_{k=0}^{\infty} \frac{X^k}{k!}, \tag{1}$$

which always converges entrywise. It is clear from the definition that the matrix exponential has the property that

$$\frac{d}{dt}e^{tX} = Xe^{tX} = e^{tX}X,$$

and in particular,

$$\left.\frac{d}{dt}e^{tX}\right|_{t=0} = X. \tag{2}$$

Remarkably, the matrix exponential of the infinitesimal generators of $SO(3)$ recovers the rotation matrices:

$$e^{i\theta\tau^j} = R_j(\theta). \tag{3}$$

Apparently, linear deviations from the identity were enough to retain all the information in the generators of $SO(3)$.

Our next examples of matrix Lie groups are the *unitary group* $U(n)$ and the *special unitary group* $SU(n)$, which are analagous to $O(n)$ and $SO(n)$. They are defined as complex matrices that preserve the standard Hermitian inner product on $\mathbb{C}^n$. This gives the condition that $U(n) = \{X \in GL(n;\mathbb{C}) : X^\dagger = X^{-1}\}$, where $X^\dagger$ represents the adjoint—or conjugate transpose—of a matrix X. Likewise, $SU(n) = \{X \in U(n) : \det X = 1\}$. It turns out that the *Pauli matrices*,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which satisfy the commutation relation $[\sigma_i/2, \sigma_j/2] = i\epsilon^{ijk}\sigma_k/2$, are the infinitesimal generators of $SU(2)$. To see this, when we exponentiate the Pauli matrices, we get

$$U_1(\theta) = e^{i\theta\sigma_1/2} = \begin{pmatrix} \cos\theta/2 & i\sin\theta/2 \\ i\sin\theta/2 & \cos\theta/2 \end{pmatrix},$$
$$U_2(\phi) = e^{i\phi\sigma_2/2} = \begin{pmatrix} \cos\phi/2 & \sin\phi/2 \\ -\sin\phi/2 & \cos\phi/2 \end{pmatrix},$$
$$U_3(\psi) = e^{i\psi\sigma_3/2} = \begin{pmatrix} e^{i\psi/2} & 0 \\ 0 & e^{-i\psi/2} \end{pmatrix}.$$

Note that these matrices have determinant 1 and satisfy $U^\dagger = U^{-1}$, which means they are indeed elements of $SU(2)$. Furthermore, $\{U_1, U_2, U_3\}$ are independent, so they generate $SU(2)$. Working backwards, it is easy to check that the Pauli matrices are the first order corrections of the U_j, namely $U_j(\theta) = I + i(\theta/2)\sigma_j$. So once again, linearizing the generators has retained all the original information.

The commutator relations for the infinitesimal generators of $SO(3)$ and $SU(2)$ are identical. This suggests a relationship between $SO(3)$ and $SU(2)$, which we will now introduce. Consider the matrix

$$M = \sum_{i=1}^{3} x_i\sigma_i = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}.$$

We see that M has trace zero and is self-adjoint, i.e. M is equal to its adjoint $M^\dagger$. Let V denote the space consisting of matrices of this form. We can identify V with $\mathbb{R}^3$ using the coordinates (x_1, x_2, x_3) and the inner product $\langle X, Y\rangle = \frac{1}{2}\operatorname{trace} XY$. For each $U \in SU(2)$, define an operator $\Phi_U : V \to V$ by $\Phi_U(X) = UXU^{-1}$. Direct computation shows that each Φ_U preserves the inner product. Indeed,

$$\langle \Phi_U(X), \Phi_U(Y)\rangle = \frac{1}{2}\operatorname{trace} UXU^{-1}UYU^{-1} = \frac{1}{2}\operatorname{trace} UXYU^{-1} = \frac{1}{2}\operatorname{trace} XY = \langle X, Y\rangle,$$

where the second to last equality follows from the cyclic property of the trace. So Φ_U is an orthogonal operator by definition. Moreover, $\Phi_{U_1}\Phi_{U_2} = \Phi_{U_1U_2}$, which implies the map Φ that sends U to Φ_U is a homomorphism. Since an arbitrary matrix in $SU(2)$ is of the form

$$\begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}, \quad |\alpha|^2 + |\beta|^2 = 1, \tag{4}$$

we have that $SU(2)$ is homeomorphic to S^3. It follows that $SU(2)$ is simply connected. And since Φ is continuous, we conclude that the image of Φ lies in the *identity connected component* of $O(3)$, which is $SO(3)$. To see that $\Phi : SU(2) \to SO(3)$ is surjective, take any rotation of $\mathbb{R}^3$ and express the axis of rotation P as

$$P = U_0 \begin{pmatrix} x_3 & 0 \\ 0 & -x_3 \end{pmatrix} U_0^{-1}$$

for some $U_0 \in U(2)$. Then the plane orthogonal to P consists of the matrices

$$W = U_0 \begin{pmatrix} 0 & x_1 - ix_2 \\ x_1 + ix_2 & 0 \end{pmatrix} U_0^{-1}.$$

Let

$$U = U_0 \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix} U_0^{-1}.$$

Then we find that $UPU^{-1} = X$ and UWU^{-1} rotates the x_1 and x_2 components of W by angle θ. Therefore Φ_U agrees with our rotation.

The kernel of Φ is $\{I, -I\}$, and so Φ sends U and $-U$ to the same rotation, for any $U \in SU(2)$. We say that $SU(2)$ is a double cover of $SO(3)$. This result demonstrates that $SO(3)$ is homeomorphic to $\mathbb{R}P^3$, and so it is not simply connected. We conclude that $SU(2)$ is the *universal covering group* of $SO(3)$. In other words, $SU(2)$ is a simply connected topological group which admits a continuous homomorphism onto $SO(3)$ that is locally one-to-one. In fact, there is a more general statement that a universal covering group $\hat{G}$ exists for every connected topological group G, and it is unique up to isomorphism. Moreover, supposing ρ is such a continuous homomorphism from $\hat{G}$ onto G that is locally one-to-one, we have $\hat{G}/\ker\rho \simeq G$.

Here's another example of a matrix Lie group. Consider the skew-symmetric bilinear form ω on $\mathbb{C}^{2n}$ defined by

$$\omega(x, y) = \sum_{j=1}^{n} x_j y_{n+j} - x_{n+j} y_j. \tag{5}$$

This is equivalent to $\omega(x, y) = x^T \Omega y$, where

$$\Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

Clearly the matrices A that leave ω invariant satisfy

$$A^T \Omega A = \Omega. \tag{6}$$

The set of all such A is the **non-compact symplectic group** $Sp(2n;\mathbb{C})$.

3 Lie algebras

We would like to further explore infinitesimal generators of matrix Lie groups. Define a *one-parameter subgroup* as a continuous homomorphism from the additive group $\mathbb{R}^+$ to $GL(n;\mathbb{C})$. It turns out that any one-parameter subgroup can be expressed as an exponential. I will prove the case where the curve is differentiable.

Proposition 1. *Let $\varphi(t)$ be a differentiable one-parameter subgroup, and suppose $A = \varphi'(0)$. Then $\varphi(t) = e^{tA}$ for all t.*

Proof. Since φ is a homomorphism, we have $\varphi(t + \Delta t) = \varphi(t)\varphi(\Delta t)$, and in particular $\varphi(t) = \varphi(t)\varphi(0)$. Then

$$\frac{\varphi(t + \Delta t) - \varphi(t)}{\Delta t} = \varphi(t)\frac{\varphi(\Delta t) - \varphi(0)}{\Delta t}.$$

In the limit as $\Delta t \to 0$, the left hand side becomes $\varphi'(t)$ and the right hand side becomes $\varphi(t)\varphi'(0)$, and so

$$\varphi'(t) = A\varphi(t).$$

Certainly e^{tA} satisfies this differential equation, and since e^{tA} and $\varphi(t)$ pass through the identity at $t = 0$, we conclude (by uniqueness of solutions to systems of ODE's) that $\varphi(t) = e^{tA}$ for all t. To verify that e^{tA} is a one-parameter subgroup, note that for real numbers r and s, the matrices rA and sA commute, and so by a property of the matrix exponential, $e^{(r+s)A} = e^{rA}e^{sA}$. □

I will now provide an abstract definition of a *Lie algebra*, followed by an explanation of how they relate to one-parameter subgroups.

Definition 2. *A* ***Lie algebra*** $\mathfrak{g}$ *is a vector space V over $\mathbb{R}$ or $\mathbb{C}$, equipped with an operation $[\ ,\ ] : V \times V \to V$ called the Lie bracket, which satisfies the following:*
(a) $[\ ,\ ]$ is bilinear.
(b) (skew-symmetry) $[X, Y] = -[Y, X]$ for $X, Y \in V$.
(c) (Jacobi identity) $[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]$ for $X, Y, Z \in V$.

Under this definition, we may associate a Lie algebra to each Lie group. Consider the set of tangent vectors to all smooth curves in G through the identity. Let $\mathfrak{g}$ denote this set of vectors together with the commutator as a Lie bracket. We will now prove that $\mathfrak{g}$ is a Lie algebra in accordance with Definition 2. Suppose X and Y are any two elements in $\mathfrak{g}$, so that there exist smooth curves $U(\theta)$ and $V(\theta)$ in G passing through the identity which satisfy $U'(0) = X$ and $V'(0) = Y$. Consider the curve $U(\theta)V(\theta)$ in G. Its derivative at the identity is

$$\frac{d}{d\theta}U(\theta)V(\theta)\bigg|_{\theta=0} = U'(0)V(0) + U(0)V'(0) = X + Y,$$

which shows that $\mathfrak{g}$ is closed under addition. Also, we can multiply our parameter θ by any scalar λ, so that

$$\frac{d}{d\theta}U(\lambda\theta)\bigg|_{\theta=0} = \lambda U'(0) = \lambda X,$$

and so $\mathfrak{g}$ is closed under multiplication. Thus $\mathfrak{g}$ is a vector space over $\mathbb{R}$ or $\mathbb{C}$, so it makes sense to speak of $\mathfrak{g}$ as a tangent space.

It is straightforward to verify that the commutator satisfies the properties required of a Lie bracket. It remains to show that $\mathfrak{g}$ contains the commutator. Since U and V are smooth, they are infinitely differentiable, but all we need to keep are the second order terms. By Taylor's theorem, we are allowed to write

$$U(\theta) = I + \theta X + \frac{\theta^2}{2}U''(0) + \mathcal{O}(\theta^3),$$

for an expansion about the identity. Its inverse is given by

$$U^{-1}(\theta) = I - \theta X - \frac{\theta^2}{2}(U''(0) - 2X^2) + \mathcal{O}(\theta^3),$$

and likewise for V and V^{-1}. You can convince yourself that the inverse formula is correct by multiplying $U(\theta)U^{-1}(\theta)$ and disregarding terms higher than order 2. Define a curve $\gamma(\theta)$ in G by

$$\gamma(\theta) = U(\theta)V(\theta)U^{-1}(\theta)V^{-1}(\theta).$$

In order to expand γ, we first write out $U(\theta)V(\theta)$, which is

$$\begin{aligned}
&(I + \theta X + \frac{\theta^2}{2}U''(0) + \mathcal{O}(\theta^3))(I + \theta Y + \frac{\theta^2}{2}V''(0) + \mathcal{O}(\theta^3)) \\
&= I + \theta(X + Y) + \theta^2(XY + \frac{1}{2}U''(0) + \frac{1}{2}V''(0)) + \mathcal{O}(\theta^3).
\end{aligned}$$

Next we write out $U^{-1}(\theta)V^{-1}(\theta)$, which is

$$\begin{aligned}
&(I - \theta X - \frac{\theta^2}{2}(U''(0) - 2X^2) + \mathcal{O}(\theta^3))(I - \theta Y - \frac{\theta^2}{2}(V''(0) - 2Y^2) + \mathcal{O}(\theta^3)) \\
&= I + \theta(-X - Y) + \theta^2(XY - \frac{1}{2}U''(0) - \frac{1}{2}V''(0) + X^2 + Y^2) + \mathcal{O}(\theta^3).
\end{aligned}$$

Then combining these expressions together, we have

$$\begin{aligned}
\gamma(\theta) &= I + \theta^2(2XY + X^2 + Y^2 - (X + Y)^2) + \mathcal{O}(\theta^3) \\
&= I + \theta^2(XY - YX) + \mathcal{O}(\theta^3) \\
&= I + \theta^2[X, Y] + \mathcal{O}(\theta^3).
\end{aligned}$$

Reparameterize γ with $\tau = \theta^2$, so that $\gamma(\tau) = I + \tau[X, Y] + \mathcal{O}(\tau^{3/2})$. Then we have

$$\frac{d}{d\tau}\gamma(\tau)\bigg|_{\tau=0} = [X, Y],$$

so by definition $[X, Y]$ belongs to $\mathfrak{g}$. Therefore, $\mathfrak{g}$ is a Lie algebra corresponding to G. Suppose I have a basis $\{X_a\}$ of $\mathfrak{g}$ indexed by a. Now that we know $\mathfrak{g}$ is an algebra under commutation, it makes sense to write

$$[X_a, X_b] = if_{abc}X_c, \tag{7}$$

where the summation over c is understood. The f_{abc} are called the **structure constants** of $\mathfrak{g}$ with respect to the chosen basis. The factor of i out front is conventional in physics because it ensures that the structure constants are real in a unitary representation of the algebra. For more information, see [2].

It can be shown that the Lie algebra $\mathfrak{g}$ of a matrix Lie group G—consisting of the tangent vectors to G at the identity—is equivalent to the set $\{X \in M_n(\mathbb{C}) : e^{tX} \in G \ \forall t \in \mathbb{R}\}$. In other words, $\mathfrak{g}$ precisely contains the matrices whose corresponding one-parameter subgroups map to G. Then we may think of the exponential as the canonical map

$$\exp : \mathfrak{g} \to G \tag{8}$$

whose image lies in the identity connected component G_0 of G. According to this alternative formulation, $\mathfrak{g}$ is a real Lie algebra, although it is often the case that $\mathfrak{g}$ ends up being complex as well. It can be shown, by representing group elements around some neighborhood of the identity as exponentials, that in order to preserve the group operation we must ensure the commutator belongs to $\mathfrak{g}$. The calculation is similar to what I performed above, and it involves expanding the logarithm up through second order terms. This further motivates our choice of the commutator as the Lie bracket. The question remains whether expanding to higher order will reveal additional conditions we need to impose on $\mathfrak{g}$ in order to maintain the group operation. Fortunately, that is not the case; knowing the structure constants is sufficient to recover the group operation. This is seen explicitly in the *Baker-Campbell-Hausdorff (BCH) formula*, which I will omit. For reference, see [3].

Lemma 2. *For any $A \in M(n; \mathbb{C})$, we have $e^{\operatorname{trace} A} = \det e^A$.*

Proof. By a standard fact from linear algebra (see [1]), there exists a sequence of matrices $(A_k)_{k\in\mathbb{N}}$ in $M(n; \mathbb{C})$ converging to A such that for each k, A_k has distinct eigenvalues $\lambda_{k,1}, \ldots, \lambda_{k,n}$. Then every A_k is similar to a diagonal matrix D_k with diagonal entries $\lambda_{k,1}, \ldots, \lambda_{k,n}$, meaning $TA_kT^{-1} = D_k$ for some invertible matrix T. Observe that for any $N \in \mathbb{N}$, $(TA_kT^{-1})^N = T(A_k)^NT^{-1}$, and consequently $e^{D_k} = Te^{A_k}T^{-1}$ because their power series agree term-by-term. The eigenvalues of e^{D_k} are $e^{\lambda_{k,1}}, \ldots, e^{\lambda_{k,n}}$ for all k. Thus $e^{\operatorname{trace} D_k} = e^{\lambda_{k,1}+\cdots+\lambda_{k,n}} = e^{\lambda_{k,1}} \cdots e^{\lambda_{k,n}} = \det e^{D_k}$. Note that the determinant and trace of D_k are the same as for A_k, so the result holds for each entry in the sequence. The formula holds in the limit. □

Denote the Lie algebra of a matrix Lie group by corresponding lowercase letters. The Lie algebra $sl(n; \mathbb{C})$ consists of trace 0 matrices, as the above lemma illustrates. It has dimension $n^2 - 1$ and is denoted by A_{n-1} in *Cartan's* *classification*. For completeness, I will list the remaining classical algebras. The Lie algebra $so(n)$ consists of skew-symmetric matrices. Cartan's classification of $so(2n+1)$ is B_n, which has dimension $4n^2 + 2n$. Meanwhile, D_n denotes $so(2n)$, which has dimension $2n^2 - n$. Finally, C_n classifies $sp(2n;\mathbb{C})$, which is the set of matrices

$$\begin{pmatrix} A & B \\ C & -A^T \end{pmatrix}$$

where B and C are symmetric. Its dimension is $2n^2 + n$. See Figure 1 on page 31 for a geometric description of the labels A_{n-1}, B_n, C_n, and D_n.

As a final remark, Lie's third theorem states that any finite-dimensional, real Lie algebra is the Lie algebra of some matrix Lie group.

3.1 Solvable and nilpotent Lie algebras

Definition 3. *Let $\mathfrak{g}$ be a Lie algebra. The* ***adjoint representation*** *of $\mathfrak{g}$ is the map* ad *that sends each $X \in \mathfrak{g}$ to $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$, where $\mathrm{ad}_X(Y) = [X, Y]$.*

The adjoint representation is always a matrix representation. As a consequence of the Jacobi identity, it is easy to see that ad is a *Lie algebra homomorphism* and the image of ad is a *derivation*. This means that ad is a linear map which is compatible with the Lie bracket, i.e. $\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y]$. Furthermore, ad_X is linear and satisfies $\mathrm{ad}_X([Y, Z]) = [\mathrm{ad}_X(Y), Z] + [Y, \mathrm{ad}_X(Z)]$.

We will now begin to study the structure of Lie algebras, mainly following along with [6]. Let's introduce a few concepts first. Suppose $\mathfrak{g}$ is a Lie algebra. Say that $\mathfrak{h}$ is a *subalgebra* of $\mathfrak{g}$ if $\mathfrak{h}$ is a subspace of $\mathfrak{g}$ and $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$, that is, $[H_1, H_2] \in \mathfrak{h}$ for any $H_1, H_2 \in \mathfrak{h}$. A subalgebra $\mathfrak{h}$ is an **ideal** of $\mathfrak{g}$ if it satisfies a stronger condition that $[\mathfrak{g}, \mathfrak{h}] \subset \mathfrak{h}$, that is, $[X, H] \in \mathfrak{h}$ for any $X \in \mathfrak{g}, H \in \mathfrak{h}$. The *center* of $\mathfrak{g}$, denoted $Z(\mathfrak{g})$, is the set of all $Y \in \mathfrak{g}$ such that $[X, Y] = 0$ for all $X \in \mathfrak{g}$. The center is an ideal. Given an ideal $\mathfrak{h}$, we can create a *quotient algebra* as the set of equivalence classes under the equivalence relation $X \sim Y$ whenever $X - Y \in \mathfrak{h}$. The Lie bracket of the quotient algebra is given by $[X + \mathfrak{h}, Y + \mathfrak{h}] = [X, Y] + \mathfrak{h}$, which is well-defined because $\mathfrak{h}$ is an ideal.

The set of commutators $[\mathfrak{g}, \mathfrak{g}]$ is an ideal of $\mathfrak{g}$. Construct a sequence, called the *derived series*, beginning with the space of linear combinations of commutators $\mathfrak{g}^{(1)} = [\mathfrak{g}, \mathfrak{g}]$, and for $n \geq 1$, define $\mathfrak{g}^{(n+1)} = [\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}]$. If the derived series terminates to zero, we say that $\mathfrak{g}$ is **solvable**. Construct another sequence, called the *lower central series*, beginning with the space of linear combinations of commutators $\mathfrak{g}_{(1)} = [\mathfrak{g}, \mathfrak{g}]$, and for $n \geq 1$, define $\mathfrak{g}_{(n+1)} = [\mathfrak{g}, \mathfrak{g}_{(n)}]$. If the lower central series terminates to zero, we say that $\mathfrak{g}$ is **nilpotent**. Both the derived series and lower central series are sequences of ideals. It is clear that nilpotency implies solvability. Finally, say that $\mathfrak{g}$ is *irreducible* if the only ideals of $\mathfrak{g}$ are $\{0\}$ and $\mathfrak{g}$ itself; moreover, $\mathfrak{g}$ is **simple** if it is irreducible and $\dim \mathfrak{g} \geq 2$. Equivalently, $\mathfrak{g}$ is simple if it is irreducible and noncommutative.

Lemma 3. *Suppose $\mathfrak{a}$ is a solvable ideal of $\mathfrak{g}$. If $\mathfrak{g}/\mathfrak{a}$ is solvable, then $\mathfrak{g}$ is solvable. Likewise, suppose $\mathfrak{a}$ is a nilpotent ideal contained in the center of $\mathfrak{g}$. If $\mathfrak{g}/\mathfrak{a}$ is nilpotent, then $\mathfrak{g}$ is nilpotent.*

Proof. It is clear that $(\mathfrak{g}/\mathfrak{a})^{(j)} = \mathfrak{g}^{(j)}/\mathfrak{a}$. Then if $\mathfrak{g}/\mathfrak{a}$ is solvable, there exists an N such that $\mathfrak{g}^{(N)} \subset \mathfrak{a}$. Since $\mathfrak{a}$ is solvable, it follows that $\mathfrak{g}^{(N)}$ is solvable, hence $\mathfrak{g}$ is as well. Similarly, we have that $(\mathfrak{g}/\mathfrak{a})_{(j)} = \mathfrak{g}_{(j)}/\mathfrak{a}$. Then if $\mathfrak{g}/\mathfrak{a}$ is nilpotent, there exists an M such that $\mathfrak{g}_{(M)} \subset \mathfrak{a}$. This implies that $\mathfrak{g}_{M+1} = [\mathfrak{g}, \mathfrak{g}_M] \subset [\mathfrak{g}, \mathfrak{a}] = 0$, since $\mathfrak{a}$ is in the center. Therefore, $\mathfrak{g}$ is nilpotent. □

The notation $gl(\mathfrak{g})$ refers to the Lie algebra of $GL(\mathfrak{g})$—the space of all invertible linear operators on $\mathfrak{g}$—which is simply the space of all linear operators on $\mathfrak{g}$ (not necessarily invertible).

Lemma 4. *Let ρ be a homomorphism from $\mathfrak{g}$ to $gl(\mathfrak{g})$. If $\operatorname{Im}\rho$ and $\ker\rho$ are solvable, then $\mathfrak{g}$ is solvable. In particular, $\operatorname{ad}\mathfrak{g}$ is solvable if and only if $\mathfrak{g}$ is solvable, and $\operatorname{ad}\mathfrak{g}$ is nilpotent if and only if $\mathfrak{g}$ is nilpotent.*

Proof. We know $\mathfrak{g}/\ker\rho \simeq \operatorname{Im}\mathfrak{g}$, and since $\ker\rho$ is an ideal, I can use Lemma 3 and conclude that $\mathfrak{g}$ is solvable. The adjoint representation is one such homomorphism, and $\ker\operatorname{ad}$ is precisely the center of $\mathfrak{g}$. As the center is abelian, $\ker\operatorname{ad}$ is both solvable and nilpotent. Therefore, using the statement we just proved, $\operatorname{ad}\mathfrak{g}$ solvable implies $\mathfrak{g}$ is solvable, where here $\operatorname{ad}\mathfrak{g}$ denotes the image of ad. Another application of Lemma 3 shows that $\operatorname{ad}\mathfrak{g}$ nilpotent implies $\mathfrak{g}$ is nilpotent. The reverse directions of both statements follow from the fact that homomorphic images of solvable algebras are solvable, and likewise for nilpotent algebras. □

Lemma 5. *The sum of two solvable ideals is a solvable ideal.*

Proof. Suppose $\mathfrak{a}$ and $\mathfrak{b}$ are solvable ideals. Then $[\mathfrak{g}, \mathfrak{a}+\mathfrak{b}] = [\mathfrak{g}, \mathfrak{a}]+[\mathfrak{g}, \mathfrak{b}] \subset \mathfrak{a}+\mathfrak{b}$, and thus $\mathfrak{a}+\mathfrak{b}$ is an ideal. Also, $\mathfrak{a}\cap\mathfrak{b}$ is an ideal in $\mathfrak{a}$. Using the relation $(\mathfrak{a}+\mathfrak{b})/\mathfrak{b} \simeq \mathfrak{a}/(\mathfrak{a}\cap\mathfrak{b})$ and knowing that $\mathfrak{a}/(\mathfrak{a}\cap\mathfrak{b})$ is solvable, as it is a homomorphic image of $\mathfrak{a}$, Lemma 3 tells us that $\mathfrak{a}+\mathfrak{b}$ is solvable. □

The above lemma confirms the existence of a unique largest solvable ideal, which we call the *radical* $\mathfrak{R}$. We say that a Lie algebra is **semisimple** if its radical is $\{0\}$. Note that given any Lie algebra $\mathfrak{g}$ and corresponding radical $\mathfrak{R}$, the quotient algebra $\mathfrak{g}/\mathfrak{R}$ is semisimple. The following lemma demonstrates an equivalent notion of semisimplicity.

Lemma 6. *A Lie algebra $\mathfrak{g}$ is semisimple if and only if it contains no nonzero abelian ideals.*

Proof. Suppose $\mathfrak{g}$ is semisimple. Then $\mathfrak{g}$ contains no nonzero solvable ideals. Seeing as abelian ideals are solvable, $\mathfrak{g}$ cannot have any nonzero abelian ideals. Conversely, if $\mathfrak{g}$ were not semisimple, it would have a nonzero solvable ideal, say $\mathfrak{a}$. If k is the smallest natural number for which $\mathfrak{a}^{(k)} = 0$, then $\mathfrak{a}^{(k-1)}$ is a nonzero abelian ideal. □

I will now begin to set up Lie's theorem and Engel's theorem. In reference to an operator (or endomorphism) T, the terms *nilpotent* and *semisimple* should not be confused with their definitions for Lie algebras. By *nilpotent* we mean that $T^N = 0$ for some N. Moreover, say that an endomorphism X is *ad-nilpotent* if ad_X is nilpotent. By *semisimple* we mean that T is diagonalizable over an algebraically closed field. Moving forward, assume that all underlying vector spaces V are nonzero, complex, and finite-dimensional.

Lemma 7. *Every nilpotent element of $gl(V)$ is ad-nilpotent.*

Proof. Let $X \in gl(V)$ be a nilpotent endomorphism. Associate to X the nilpotent endomorphisms $L_X(Y) = XY$, $R_X(Y) = YX$ of left and right translations, respectively. It is clear that L_X and R_X commute. I claim that their difference, namely $L_X - R_X = \mathrm{ad}_X$, is nilpotent. To see this, choose k large enough such that both $(L_X)^k = 0$ and $(R_X)^k = 0$. Consider $(L_X - R_X)^{2k}$. Then because L_X and R_X commute, we can collect terms to obtain

$$(L_X - R_X)^{2k} = \sum_{i=0}^{2k} (-1)^i \binom{2k}{i} (L_X)^{2k-i} (R_X)^i.$$

When $i \leq k$, we have $(L_X)^{2k-i} = 0$, and when $i > k$, we have $(R_X)^i = 0$. Thus every term in the sum vanishes, as desired. $\square$

Proposition 8. *Suppose $\mathfrak{g}$ is a subalgebra of $gl(V)$ consisting of nilpotent endomorphisms. Then there exists a nonzero $v \in V$ such that $Xv = 0$ for all $X \in \mathfrak{g}$.*

Proof. We will proceed by induction on $\dim \mathfrak{g}$. The result is immediate when $\mathfrak{g}$ is one-dimensional. Let $n = \dim \mathfrak{g}$ and suppose the result holds for any proper subalgebra. Let $\mathfrak{a}$ be a maximal proper subalgebra. Consider the action of $\mathfrak{a}$ on the space $\mathfrak{g}/\mathfrak{a}$ given by $X(Y + \mathfrak{a}) = \mathrm{ad}_X(Y) + \mathfrak{a}$ for any $X \in \mathfrak{a}, Y \in \mathfrak{g}$. The set of operators $\{\mathrm{ad}_X : X \in \mathfrak{a}\}$ forms a Lie algebra whose dimension certainly does not exceed that of $\mathfrak{a}$. By Lemma 7, each operator is nilpotent, and so by the induction hypothesis, there exists an element $S + \mathfrak{a} \neq \mathfrak{a}$ in $\mathfrak{g}/\mathfrak{a}$ such that $X(S + \mathfrak{a}) = \mathfrak{a}$ for all $X \in \mathfrak{a}$. This implies that $S \notin \mathfrak{a}$ and $[X, S] \in \mathfrak{a}$ for all $X \in \mathfrak{a}$. Thus $[S] + \mathfrak{a}$ is a subalgebra that properly contains $\mathfrak{a}$, hence $[S] + \mathfrak{a} = \mathfrak{g}$ and $\mathfrak{a}$ is an ideal. Set $W = \{v \in V : Xv = 0 \; \forall X \in \mathfrak{a}\}$, which is nonempty. We see that W is invariant under $\mathfrak{g}$ because given $X \in \mathfrak{a}$, $Y \in \mathfrak{g}$, and $v \in V$, we have $X(Yw) = [X, Y]w + (YX)w = Y(Xw) = 0$. The last equality follows because the action on W is linear. Since S is nilpotent, it has an eigenvalue of 0, and since S stabilizes W, there exists a nonzero $v \in W$ so that $Sv = 0$. But then $Yv = 0$ for any $Y \in \mathfrak{g}$, as we desired to show. $\square$

We are now in a position to prove Engel's theorem, which establishes a connection between the nilpotency of a Lie algebra and nilpotency of its elements. First, here is a closely related statement that applies when $\mathfrak{g} \subset gl(V)$.

Proposition 9. *Let $\mathfrak{g} \subset gl(V)$ be a Lie algebra consisting of nilpotent endomorphisms. Then $\mathfrak{g}$ is isomorphic to a subalgebra of strictly upper triangular complex matrices.*

Proof. Let $n = \dim V$. By Proposition 8, there is a v_1 so that $Xv_1 = 0$ for all $X \in \mathfrak{g}$. Consider the action of $\mathfrak{g}$ on $V/[v_1]$ given by $X(v + [v_1]) = Xv + [v_1]$. By Proposition 8 again, there is a v_2 so that $X(v_2 + [v_1]) = [v_1]$. Repeat until we arrive at a flag of n subspaces $V_i = \text{span}\{v_1, \ldots, v_i\}$. Clearly $\mathfrak{g}V_{i+1} \subset V_i \subset V_{i+1}$, so we say that $\mathfrak{g}$ stabilizes the flag. Hence the matrices of the transformations in $\mathfrak{g}$ are strictly upper triangular with respect to the basis $\{v_1, \ldots, v_n\}$. □

Thus a Lie algebra of nilpotent endomorphisms is itself a nilpotent algebra. Engel's theorem is a stronger statement, as it pertains to any Lie algebra. The difference is that we require its elements to be ad-nilpotent, recalling that the adjoint representation is a matrix representation.

Theorem 10 (Engel)**.** *A Lie algebra $\mathfrak{g}$ is nilpotent if and only if every element of $\mathfrak{g}$ is ad-nilpotent.*

Proof. Suppose $\mathfrak{g}$ is nilpotent. Then $\mathfrak{g}_{(n)} = 0$ for some n, which means that $\text{ad}_{X_1} \cdots \text{ad}_{X_n} = 0$ for any n elements $X_1, \ldots, X_n \in \mathfrak{g}$. In particular, given any $X \in \mathfrak{g}$, we have $(\text{ad}_X)^n = 0$, and so X is ad-nilpotent. To prove the converse, first note that $\text{ad}\,\mathfrak{g}$ is a subalgebra of $gl(\mathfrak{g})$, so by Proposition 8 there is a nonzero $Y \in \mathfrak{g}$ such that $\text{ad}_X(Y) = 0$ for all $X \in \mathfrak{g}$. This means that Y is in the center of $\mathfrak{g}$. The quotient algebra $\mathfrak{g}/Z(\mathfrak{g})$ is therefore strictly smaller than $\mathfrak{g}$, and certainly consists of ad-nilpotent elements. By induction on $\dim \mathfrak{g}$, we have $\mathfrak{g}/Z(\mathfrak{g})$ is nilpotent. But then $\mathfrak{g}_{(n)} \subset Z(\mathfrak{g})$ for some n, and

$$\mathfrak{g}_{(n+1)} = [\mathfrak{g}, \mathfrak{g}_{(n)}] \subset [\mathfrak{g}, Z(\mathfrak{g})] = 0.$$

Therefore, $\mathfrak{g}$ is nilpotent. □

In a similar fashion, we now analyze solvable Lie algebras, but this time we only consider those which are subalgebras of $gl(V)$.

Proposition 11. *Let $\mathfrak{g}$ be a solvable subalgebra of $gl(V)$. There exists a vector $v \in V$ which is a simultaneous eigenvector for every $X \in \mathfrak{g}$.*

Proof. Proceed by induction on $\dim \mathfrak{g}$, with the one-dimensional case being trivial. Since $\mathfrak{g}$ is solvable, it must properly contain $\mathfrak{g}^{(1)}$. Let $\mathfrak{a}$ be a subspace of $\mathfrak{g}$ of codimension 1 that contains $\mathfrak{g}^{(1)}$. Then $[\mathfrak{a}, \mathfrak{g}] \subset [\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}^{(1)} \subset \mathfrak{a}$, and so $\mathfrak{a}$ is an ideal. And certainly $\mathfrak{a}$ is solvable. By our induction hypothesis, there is an element $v_0 \in V$ that is a simultaneous eigenvector for all $X \in \mathfrak{a}$. That is, there is a linear functional λ on $\mathfrak{a}$ where $Xv_0 = \lambda(X)v_0$. Choose any $Y \in \mathfrak{g} \setminus \mathfrak{a}$ and set $v_{j+1} = Yv_j$ for $j \in \mathbb{N}$. Let W denote the subspace spanned by the v_j. Then since W is finite-dimensional, $W = \text{span}\{v_1, \ldots, v_p\}$ for some p. Given $X \in \mathfrak{a}$, we have

$$Xv_1 = XYv_0 = YXv_0 + [X, Y]v_0 = \lambda(X)v_1 + \lambda([X, Y])v_0,$$

and so by induction we easily find that $Xv_j \equiv \lambda(X)v_j \mod v_0, \ldots, v_{j-1}$. Thus, W is invariant under $\mathfrak{a}$, and the matrices corresponding to each X are triangular with respect to the basis $\{v_0, \ldots, v_p\}$. After restricting the trace to W, it is clear that $\operatorname{trace} X = \lambda(X) \dim W$. Now the cyclic property of the trace forces $\operatorname{trace}([X,Y]) = 0$, and consequently $\lambda([X,Y]) = 0$. Then the equation above reduces to $Xv_1 = \lambda(X)v_1$. Now suppose $Xv_j = \lambda(X)v_j$ for some j and for all $X \in \mathfrak{a}$. We have

$$Xv_{j+1} = XYv_j = YXv_j + [X,Y]v_j = \lambda(X)v_{j+1} + \lambda([X,Y])v_j = \lambda(X)v_{j+1},$$

hence we conclude $Xv_j = \lambda(X)v_j$ for all j and for all $X \in \mathfrak{a}$. Thus the elements in $\mathfrak{a}$ are multiples of the identity over the basis $\{v_0, \ldots, v_p\}$. Restricting Y to an endomorphism of W over $\mathbb{C}$, we realize that Y has an eigenvector $w \in W$. This w is the desired eigenvector for every element in $\mathfrak{g}$. □

Proposition 11 tells us it makes sense to say there exists a linear functional λ on $\mathfrak{g}$ with $Xv = \lambda(X)v$ for some v.

Theorem 12 (Lie)**.** *Let $\mathfrak{g}$ be a subalgebra of $gl(V)$. Then the elements of $\mathfrak{g}$ are simultaneously upper triangularizable if and only if $\mathfrak{g}$ is solvable.*

Proof. Let $\mathfrak{N}_k$ denote the set of complex upper triangular square matrices such that $a_{ij} = 0$ for $j < k + i$. The reader may verify the commutation relations $[\mathfrak{N}_0, \mathfrak{N}_0] \subset \mathfrak{N}_1$, while $[\mathfrak{N}_k, \mathfrak{N}_l] \subset \mathfrak{N}_{k+l}$ for $k \geq 0, l \geq 1$. This demonstrates that the set of upper triangular matrices $\mathfrak{N}_0$ is solvable, and that the set of strictly upper triangular matrices $\mathfrak{N}_1$ is nilpotent. Now suppose $\mathfrak{g}$ is solvable, and consider the action of $\mathfrak{g}$ on V. Let v_1 be a simultaneous eigenvector for every $X \in \mathfrak{g}$, which is guaranteed to exist by Proposition 11. Let $V_1 = V/[v_1]$, and note that all transformations on V can be carried over to V_1 becuase they leave v_1 invariant. Proposition 11 again implies there is a vector v_2 so that $v_2 + [v_1]$ is a simultaneous eigenvector for all $X \in \mathfrak{g}$ acting on V_1. This means that $X(v_2 + [v_1]) = Xv_2 + [v_1] = \lambda_2(X)v_2 + [v_1]$. Let $V_2 = V_1/[v_2]$, and repeat this procedure. We arrive at a set of vectors $v_1, \ldots, v_n$ where $Xv_i \equiv \lambda_i(X)v_i \mod v_1, \ldots, v_{i-1}$. Thus every X is represented by an upper triangular matrix with respect to the basis $\{v_1, \ldots, v_n\}$. □

3.2 Killing form

We define the **Killing form** K as the symmetric bilinear form

$$K(X,Y) \equiv \operatorname{trace} \operatorname{ad}_X \operatorname{ad}_Y . \tag{9}$$

The Killing form has the nice associative property that

$$K(X,[Y,Z]) = K([X,Y],Z). \tag{10}$$

Using structure constants, it can be shown that the Killing form of an ideal $\mathfrak{a} \subset \mathfrak{g}$, considered as its own Lie algebra, is just the Killing form of $\mathfrak{g}$ restricted to $\mathfrak{a}$. In general, a form $\langle \cdot, \cdot \rangle$ on a vector space V is called *nondegenerate* if its nullspace is $\{0\}$, i.e. given any nonzero vector $v \in V$, there exists a $v' \in V$ such that $\langle v, v' \rangle \neq 0$. The *radical* of a form B is the set $\operatorname{rad} B = \{u \in V : B(u, v) = 0 \ \forall v \in V\}$.

I will now build more criteria for illustrating solvability and semisimplicity. We first need the following theorem, whose proof can be found in [4].

Theorem 13 (Jordan-Chevalley)**.** *Let $X \in gl(V)$ be an endomorphism of a complex vector space V. Then X decomposes uniquely as $X = S + N$, where $S, N \in gl(V)$ are commuting polynomials in X without constant terms, S is semisimple, and N is nilpotent.*

The Jordan-Chevalley decomposition of elements in a Lie algebra is critical in some of the proofs that follow. In particular, it is needed for Cartan's criterion, which I will turn to next. Interestingly, there is even an *abstract Jordan decomposition* which holds for elements in an arbitrary semisimple Lie algebra $\mathfrak{g}$; it agrees with the Jordan-Chevalley decomposition presented above whenever $\mathfrak{g} \subset gl(V)$. I will use abstract Jordan decomposition when setting up Cartan subalgebras.

Theorem 14 (Cartan's criterion)**.** *Let $\mathfrak{g}$ be a subalgebra of $gl(V)$. Then $\mathfrak{g}$ is solvable if and only if* $\operatorname{trace} XY = 0$ *for all $X \in \mathfrak{g}$, $Y \in \mathfrak{g}^{(1)}$.*

Proof. Suppose $\mathfrak{g}$ is solvable. By Lie's theorem, $\mathfrak{g}$ is isomorphic to a subalgebra of upper triangular matrices. Then $\mathfrak{g}^{(1)} = [\mathfrak{g}, \mathfrak{g}]$ consists of strictly upper triangular matrices. Furthermore, given any $Y \in \mathfrak{g}^{(1)}$, XY is strictly upper triangular for all $X \in \mathfrak{g}$, and so $\operatorname{trace} XY = 0$. To prove the converse, it suffices to show that $\mathfrak{g}^{(1)}$ is nilpotent, since $\mathfrak{g}^{(1)}$ nilpotent implies that $\mathfrak{g}^{(1)}$ is solvable, hence so is $\mathfrak{g}$. By Proposition 9, we are done if we show that each $X \in \mathfrak{g}^{(1)}$ is nilpotent. Choose any such element X. Let $X = S + N$ be its Jordan-Chevalley decomposition, with $S = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. The reader may verify that $\operatorname{ad}_S(E_{ij}) = (\lambda_i - \lambda_j)E_{ij}$, which demonstrates ad_S is diagonal with respect to the basis $\{E_{ij} : 1 \leq i, j \leq n\}$. By Lemma 7, ad_N is nilpotent. Thus, $\operatorname{ad}_S + \operatorname{ad}_N$ is the unique Jordan-Chevalley decomposition for ad_X, and consequently ad_S is a polynomial in ad_X without a constant term. Taking the complex conjugate of S, that is, $\overline{S} = \operatorname{diag}(\overline{\lambda_1}, \ldots, \overline{\lambda_n})$, it is clear that $\operatorname{ad}_{\overline{S}}$ is also a polynomial in ad_X without a constant term, say $\operatorname{ad}_{\overline{S}} = p(\operatorname{ad}_X)$. Then for any $Y \in \mathfrak{g}$, $\operatorname{ad}_{\overline{S}}(Y) = [\overline{S}, Y] = p(\operatorname{ad}_X)(Y) \in \mathfrak{g}^{(1)}$. Furthermore, N nilpotent implies N is strictly upper triangular (apply Proposition 9 to the space spanned by N). Then $\overline{S}N$ is strictly upper triangular as well. This implies $\operatorname{trace} \overline{S}X = \operatorname{trace} \overline{S}(S + N) = \operatorname{trace} S\overline{S} = \sum_i |\lambda_i|^2$. But we also know that X can be expressed as $X = \sum_j [A_j, B_j]$ since $X \in \mathfrak{g}^{(1)}$. Thus

$$\begin{aligned}\operatorname{trace} \overline{S}X &= \sum_j \operatorname{trace} \overline{S}[A_j, B_j] = \sum_j \operatorname{trace} \overline{S}A_jB_j - \sum_j \operatorname{trace} \overline{S}B_jA_j \\ &= \sum_j \operatorname{trace} \overline{S}A_jB_j - \sum_j \operatorname{trace} A_j\overline{S}B_j = \sum_j \operatorname{trace}[\overline{S}, A_j]B_j.\end{aligned}$$

But since each $[\overline{S}, A_j]$ is in $\mathfrak{g}^{(1)}$, by our hypothesis, the last expression above vanishes. It follows that $\sum_i |\lambda_i|^2 = 0$, and thus each $\lambda_i = 0$. This implies that $S = 0$, and we conclude that $X = N$, meaning X is nilpotent, as desired. □

Note that the trace calculations above verify associativity of the Killing form.

Corollary 15. *A Lie algebra $\mathfrak{g}$ is solvable if and only if $K(X,Y) = 0$ for all $X \in \mathfrak{g}, Y \in \mathfrak{g}^{(1)}$.*

Proof. Apply Cartan's criterion to $\operatorname{ad} \mathfrak{g} \subset gl(V)$ to conclude that $\operatorname{ad} \mathfrak{g}$ is solvable if and only if $K(X,Y) = 0$ for all $X \in \mathfrak{g}$, $Y \in \mathfrak{g}^{(1)}$. Recall that $\operatorname{ad} \mathfrak{g}$ solvable is equivalent to $\mathfrak{g}$ solvable by Lemma 4. □

Given a Lie algebra $\mathfrak{g}$, Cartan's criterion states that the ideal $\mathfrak{g}^{(1)}$ is contained in $\operatorname{rad} K$. We conclude this section with further applications of the Killing form. I will explain a similar criterion for semisimplicity and show that semisimple Lie algebras are composed of simple subalgebras. Then we will move on to studying semisimple Lie algebras.

Theorem 16. *A Lie algebra $\mathfrak{g}$ is semisimple if and only if the Killing form is nondegenerate.*

Proof. Suppose that $\mathfrak{g}$ is semisimple, i.e. its radical $\mathfrak{R}$ is zero. Let $\mathfrak{S} = \operatorname{rad} K$. By definition, given any $X \in \mathfrak{S}$, we have $K(X,Y) = 0$ for all $Y \in \mathfrak{g}$, in particular for $Y \in \mathfrak{S}^{(1)}$. Using the corollary to Cartan's criterion, we have that $\mathfrak{S}$ is solvable. It is not hard to show that $\mathfrak{S}$ is an ideal of $\mathfrak{g}$. Consequently, $\mathfrak{S} \subset \mathfrak{R} = 0$, and thus $\mathfrak{S} = 0$. By definition, K is nondegenerate. Conversely, suppose that $\mathfrak{S} = 0$. Let $\mathfrak{a}$ be an abelian ideal of $\mathfrak{g}$. Given $X \in \mathfrak{a}$ and $Y \in \mathfrak{g}$, the map $(\operatorname{ad}_X \operatorname{ad}_Y)^2$ sends arbitrary elements in $\mathfrak{g}$ to elements in $[\mathfrak{a}, \mathfrak{a}] = 0$. So the transformation $\operatorname{ad}_X \operatorname{ad}_Y$ is nilpotent, which implies that $K(X,Y) = 0$. In turn, this implies that $\mathfrak{a} \subset \mathfrak{S} = 0$, and so $\mathfrak{g}$ contains no nonzero abelian ideals. By Lemma 6, $\mathfrak{g}$ is semisimple. □

Say that a Lie algebra $\mathfrak{g}$ is the *Lie algebra direct sum* of subalgebras $\mathfrak{a}$ and $\mathfrak{b}$, denoted $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$, if $\mathfrak{g}$ is the vector space direct sum of $\mathfrak{a}$ and $\mathfrak{b}$ (also denoted $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$) and if $[A, B] = 0$ for any $A \in \mathfrak{a}$ and $B \in \mathfrak{b}$.

Theorem 17. *Let $\mathfrak{g}$ be a semisimple Lie algebra. Then $\mathfrak{g}$ decomposes as a Lie algebra direct sum $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$, where each $\mathfrak{g}_i \subset \mathfrak{g}$ is a simple subalgebra. Moreover, the decomposition is unique up to order.*

Proof. If $\mathfrak{g}$ is simple, we are done. Otherwise, there exists a proper ideal $\mathfrak{a}$ in $\mathfrak{g}$. Let $\mathfrak{a}^\perp$ denote its orthogonal complement relative to the Killing form. To see that $\mathfrak{a}^\perp$ is an ideal, simply use associativity of K:

$$K(\mathfrak{a}, [\mathfrak{g}, \mathfrak{a}^\perp]) = K([\mathfrak{a}, \mathfrak{g}], \mathfrak{a}^\perp) = K(\mathfrak{a}, \mathfrak{a}^\perp) = 0,$$

which implies $[\mathfrak{g}, \mathfrak{a}^\perp] \subset \mathfrak{a}^\perp$, as desired. We have that $\mathfrak{g}$ decomposes as a vector space direct sum $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$. Since $\mathfrak{a}$ and $\mathfrak{a}^\perp$ are ideals, it follows that $[\mathfrak{a}, \mathfrak{a}^\perp] \subset$

$\mathfrak{a} \cap \mathfrak{a}^\perp = \{0\}$, which means that $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$ is a Lie algebra direct sum. Repeat this argument on $\mathfrak{a}$ and $\mathfrak{a}^\perp$, if necessary, and continue the process until every ideal in the direct sum is irreducible. If $\dim \mathfrak{g}_j$ were commutative for some j, then $\mathfrak{g}_j \subset Z(\mathfrak{g})$. But this cannot happen, since the center of any semisimple Lie algebra is trivial. Thus each $\mathfrak{g}_j$ is simple.

Now if $\mathfrak{h}$ is any simple ideal of $\mathfrak{g}$, then we can bracket $\mathfrak{h}$ with $\mathfrak{g}$ to obtain

$$[\mathfrak{h}, \mathfrak{g}] = [\mathfrak{h}, \mathfrak{g}_1] \oplus \cdots \oplus [\mathfrak{h}, \mathfrak{g}_n].$$

The left hand side is $[\mathfrak{h}, \mathfrak{g}] = \mathfrak{h}$ since $\mathfrak{h}$ is simple and $\mathfrak{g}$ is centerless. This forces the right hand side to contain one term, which means $\mathfrak{h} = [\mathfrak{h}, \mathfrak{g}_k]$ for some k. But then $\mathfrak{h}$ must be contained in $\mathfrak{g}_k$. Since $\mathfrak{g}_k$ is simple, we have $\mathfrak{h} = \mathfrak{g}_k$. Therefore every simple subalgebra coincides with one of the $\mathfrak{g}_i$, and so the decomposition is unique. $\square$

4 Root spaces

4.1 Irreducible representations of $sl(2;\mathbb{C})$

An example that will help motivate our discussion of semisimple Lie algebras is that of $sl(2;\mathbb{C})$. The Lie algebra $sl(2;\mathbb{C})$ is important because it is the complexification of $su(2) \simeq so(3)$, which are of physical significance. By complexification, I mean $sl(2;\mathbb{C})$ is the space of formal linear combinations $v_1 + iv_2$ with $v_1, v_2 \in su(2) \simeq so(3)$. The calculations we will perform parallel the raising and lowering operators in the quantum-mechanical treatment of angular momentum.

Recall that matrices in $sl(2;\mathbb{C})$ have trace zero. Consider the following basis for $sl(2;\mathbb{C})$:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

They satisfy the commutation relations

$$[H, X] = 2X, \quad [H, Y] = -2Y, \quad [X, Y] = H. \tag{11}$$

Let π be any representation of $sl(2;\mathbb{C})$ acting on a finite-dimensional complex vector space V. Then π will map the basis elements to operators which satisfy the same commutation relations. Suppose u is an eigenvector of $\pi(H)$ with eigenvalue α, which we know exists because we are working over $\mathbb{C}$. Then since $[\pi(H), \pi(X)] = 2\pi(X)$, we have $\pi(H)\pi(X) = \pi(X)\pi(H) + 2\pi(X)$. Acting on u gives

$$\pi(H)\pi(X)u = \pi(X)\alpha u + 2\pi(X)u = (\alpha + 2)\pi(X)u.$$

This implies that either $\pi(X)u = 0$ or $\pi(X)u$ is an eigenvector for $\pi(H)$ with eigenvalue $\alpha + 2$. Similarly,

$$\pi(H)\pi(Y)u = (\alpha - 2)\pi(Y)u,$$

so that either $\pi(Y)u = 0$ or $\pi(Y)u$ is an eigenvector for $\pi(H)$ with eigenvalue $\alpha-2$. We see that $\pi(X)$ and $\pi(Y)$ shift the eigenvalues of $\pi(H)$ up and down by 2, or possibly annihilate the eigenvector. This observation is crucial to determining the irreducible representations of $sl(2;\mathbb{C})$.

Assume that π is irreducible. As before, let u be an eigenvector with eigenvalue α. Repeatedly apply $\pi(X)$, so that

$$\pi(H)\pi(X)^k u = (\alpha + 2k)\pi(X)^k u.$$

Since V is finite-dimensional, $\pi(H)$ can only have finitely many eigenvalues. So it must be the case that $\pi(X)^N u \neq 0$ but $\pi(X)^{N+1}u = 0$ for some N. Define $u_0 = \pi(X)^N u$ and $\lambda = \alpha + 2N$. Then $\pi(H)u_0 = \lambda u_0$ and $\pi(X)u_0 = 0$. Next, define $u_k = \pi(Y)^k u_0$. It is clear that

$$\pi(H)u_k = (\lambda - 2k)u_k.$$

I will now use induction to show that

$$\pi(X)u_k = k[\lambda - (k-1)]u_{k-1}, \quad k \geq 1.$$

When $k = 1$, we have $\pi(X)u_1 = \pi(X)\pi(Y)u_0$. Using the relation $[\pi(X), \pi(Y)] = \pi(H)$, we have

$$\pi(X)\pi(Y)u_0 = \pi(Y)\pi(X)u_0 + \pi(H)u_0 = \pi(H)u_0 = \lambda u_0,$$

as desired. Now assume $\pi(X)u_k = k[\lambda - (k-1)]u_{k-1}$ for some $k \geq 1$. We want to show that $\pi(X)u_{k+1} = (k+1)(\lambda - k)u_k$. We have

$$\begin{aligned}
\pi(X)u_{k+1} &= \pi(X)\pi(Y)u_k = \pi(Y)\pi(X)u_k + \pi(H)u_k \\
&= \pi(Y)k[\lambda - (k-1)]u_{k-1} + (\lambda - 2k)u_k \\
&= k[\lambda - (k-1)]\pi(Y)u_{k-1} + (\lambda - 2k)u_k \\
&= \{k[\lambda - (k-1)] + (\lambda - 2k)\}u_k \\
&= (k+1)(\lambda - k)u_k.
\end{aligned}$$

Once again, since $\pi(H)$ has finitely many eigenvalues, there exists an m such that $u_m \neq 0$ but $u_{m+1} = 0$. For $u_{m+1} = 0$, we have $\pi(X)u_{m+1} = 0$. By our inductive formula, this implies that

$$(m+1)(\lambda - m)u_m = 0.$$

Since m is nonnegative and u_m is nonzero, we see that $\lambda = m$. We conclude that for every irreducible representation π of $sl(2;\mathbb{C})$, there exists a nonnegative integer m and nonzero vectors $u_0, \ldots, u_m$ such that

$$\pi(H)u_k = (m - 2k)u_k \tag{12}$$

$$\pi(Y)u_k = \begin{cases} u_{k+1} & k < m \\ 0 & k = m \end{cases} \tag{13}$$

$$\pi(X)u_k = \begin{cases} k[m - (k-1)]u_{k-1} & k > 0 \\ 0 & k = 0 \end{cases}. \tag{14}$$

Notice that $u_0, \dots, u_m$ are eigenvectors of $\pi(H)$ with distinct eigenvalues, which means they are linearly independent. Then $u_0, \dots, u_m$ span an $(m+1)$-dimensional subspace which is clearly invariant under $\pi(H)$, $\pi(X)$, and $\pi(Y)$. Hence the subspace is invariant under $\pi(Z)$ for all $Z \in sl(2;\mathbb{C})$. Since π is irreducible, $\text{span}\{u_0, \dots, u_m\}$ must be all of V. Therefore, every irreducible representation of dimension $m+1$ is governed by Equations 12, 13, and 14. Moreover, any two irreducible representations of the same dimension are isomorphic.

4.2 Semisimple Lie algebras

We will now proceed to study semisimple Lie algebras in more detail, first referencing [4]. Let $\mathfrak{g}$ denote a semisimple Lie algebra. If $\mathfrak{g}$ consisted of only nilpotent elements, then they would be ad-nilpotent by Lemma 7, and so Engel's theorem tells us that $\mathfrak{g}$ would be nilpotent. Consequently, $\mathfrak{g}$ would be solvable and its radical $\mathfrak{R}$ would be all of $\mathfrak{g}$. This cannot be the case if $\mathfrak{g}$ is semisimple. Hence we can find an element $X \in \mathfrak{g}$ with a nonzero semisimple component $S \in \mathfrak{g}$ in its abstract Jordan decomposition. The span of S provides a straightforward subalgebra of semisimple (i.e. diagonalizable) elements. Therefore, $\mathfrak{g}$ contains **toral** subalgebras, which are algebras consisting of semisimple elements. We will need the following lemma.

Lemma 18. *Any toral subalgebra of* $\mathfrak{g}$ *is abelian.*

Proof. Let $\mathfrak{t}$ be toral, restrict the adjoint representation to $\mathfrak{t}$, and take $X \in \mathfrak{t}$ to be nonzero. Seeing as ad_X is semisimple, we need to show that ad_X has no nonzero eigenvalues. Suppose there is a nonzero $Y \in \mathfrak{t}$ such that $[X, Y] = \lambda Y$ for $\lambda \neq 0$. Now, as ad_Y is also semisimple, there is a basis of $\mathfrak{g}$ consisting of eigenvectors for ad_Y, say $\{v_1, \dots, v_n\}$. Write X as a linear combination $X = c_1v_1 + \dots + c_nv_n$. Consider $\text{ad}_Y(X) = -\lambda Y$, which on the one hand is an eigenvector of ad_Y with eigenvalue 0. On the other hand, $\text{ad}_Y(X) = [Y, c_1v_1 + \dots + c_nv_n] = c_1[Y, v_1] + \dots + c_n[Y, v_n]$ is a linear combination of basis vectors that have nonzero eigenvalues. So applying ad_Y to $c_1[Y, v_1] + \dots + c_n[Y, v_n]$ produces yet another linear combination of basis vectors with nonzero eigenvalues, which must be equal to 0. Since X is arbitrary, this yields a contradiction. □

Fix a maximal toral subalgebra $\mathfrak{h}$, i.e. a toral subalgebra not properly contained in another. Since every $H \in \mathfrak{h}$ is semisimple, we have that every ad_H is semisimple as well. We have thus shown the existence of *Cartan subalgebras*, defined as follows.

Definition 4. *A* ***Cartan subalgebra*** *of a semisimple Lie algebra* $\mathfrak{g}$ *is a maximal abelian subalgebra* $\mathfrak{h}$ *such that* ad_H *is semisimple for all* $H \in \mathfrak{h}$.

In general, $\mathfrak{h}$ need not be unique. But if $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are Cartan subalgebras of $\mathfrak{g}$, then they are conjugates, so there exists an automorphism $\varphi : \mathfrak{g} \to \mathfrak{g}$ such that $\varphi(\mathfrak{h}_1) = \mathfrak{h}_2$. Without loss of generality, we may speak of *the* Cartan subalgebra of $\mathfrak{g}$ instead.

It is easy to check that since $\mathfrak{h}$ is abelian, $\operatorname{ad}\mathfrak{h}$ is a family of commuting operators. Seeing as each of these operators is diagonalizable, a standard result in linear algebra allows us to conclude that over a finite-dimensional vector space, every ad_H is simultaneously diagonalizable. So there exists a basis of $\mathfrak{g}$ such that each basis vector is a simultaneous eigenvector for every ad_H. If $X \in \mathfrak{g}$ is one such eigenvector, then the eigenvalues for each ad_H form a linear functional on $\mathfrak{h}$. If the functional is nonzero, we call it a *root*, as the following definition declares.

Definition 5. *A nonzero element $\alpha \in \mathfrak{h}^*$ is a* ***root*** *of $\mathfrak{g}$ if there exists a nonzero $X \in \mathfrak{g}$ such that $[H, X] = \alpha(H)X$ for all $H \in \mathfrak{h}$. Given a root α, the* ***root space*** *$\mathfrak{g}_\alpha$ is the set of all $X \in \mathfrak{g}$ such that $[H, X] = \alpha(H)X$ for all $H \in \mathfrak{h}$.*

A nonzero element in a root space is referred to as a *root vector*. Even if α is not a root, the subspace $\mathfrak{g}_\alpha$ can still be defined accordingly. Note that $\mathfrak{g}$ decomposes as a vector space direct sum of the $\mathfrak{g}_\alpha$ precisely because the image of $\operatorname{ad}\mathfrak{h}$ is simultaneously diagonalizable. We can say even more, however. Observe that $\mathfrak{g}_0$ is the set of elements that commute with $\mathfrak{h}$, meaning $\mathfrak{g}_0$ is the *centralizer* of $\mathfrak{h}$. It turns out that $\mathfrak{g}_0$ and $\mathfrak{h}$ are equivalent, as I will now prove. First, we need to show nondegeneracy of the Killing form on $\mathfrak{g}_0$.

Lemma 19. *The restriction of K to $\mathfrak{g}_0$ is nondegenerate.*

Proof. Let $H \in \mathfrak{g}_0$ be given. By Lemma 21, we have $K(H, X_\alpha) = 0$ for any root vector $X_\alpha \in \mathfrak{g}_\alpha$. Suppose that $K(H, H') = 0$ for all $H' \in \mathfrak{g}_0$. Then it follows that $K(H, X) = 0$ for all $X \in \mathfrak{g}$ by the decomposition of $\mathfrak{g}$ into $\mathfrak{g}_0$ and its root spaces $\mathfrak{g}_\alpha$. Since K is nondegenerate on $\mathfrak{g}$, this implies that $H = 0$. Thus K restricted to $\mathfrak{g}_0$ is nondegenerate as well. □

Proposition 20. *Suppose $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$, and let $\mathfrak{g}_0$ denote its centralizer. Then $\mathfrak{g}_0 = \mathfrak{h}$.*

Proof. I will first show that $\mathfrak{g}_0$ is nilpotent. Restrict the adjoint representation to $\mathfrak{g}_0$. If $X \in \mathfrak{g}_0$ is nilpotent, then ad_X is nilpotent by Lemma 4. If $X \in \mathfrak{g}_0$ is semisimple, then the subalgebra $\mathfrak{h} + \mathbb{C}X$ of $\mathfrak{g}$ is toral. By maximality of $\mathfrak{h}$, we have $\mathfrak{h} + \mathbb{C}X = \mathfrak{h}$, which implies $X \in \mathfrak{h}$. But then $\operatorname{ad}_X = 0$ is nilpotent. Now choose an arbitrary $X \in \mathfrak{g}_0$ and consider its Jordan-Chevlley form $X = S + N$. As we've seen before, $\operatorname{ad}_X = \operatorname{ad}_S + \operatorname{ad}_N$ is the Jordan-Chevalley form of ad_X. By definition, $X \in \mathfrak{g}_0$ means that ad_X maps $\mathfrak{h}$ to the subspace $\{0\}$. By the Jordan-Chevalley theorem, ad_S and ad_N are commuting polynomials in ad_X without constant term, so they also map $\mathfrak{h}$ to $\{0\}$. Then S and N both belong to $\mathfrak{g}_0$, and so by the above, ad_S and ad_N are nilpotent. Thus ad_X is the sum of commuting nilpotent endomorphisms, and so ad_X is itself nilpotent by the argument we used in Lemma 7. According to Engel's theorem, $\mathfrak{g}_0$ is nilpotent.

Next, we will show that K is nondegenerate on $\mathfrak{h}$, which we do not know a priori. Suppose $K(H, \mathfrak{h}) = 0$ for some $H \in \mathfrak{h}$. Recall from the proof of Cartan's criterion that if A and B are commuting endomorphisms on a finite-dimensional vector space, with B nilpotent, then AB is nilpotent, so that $\operatorname{trace} AB = 0$. For any $N \in \mathfrak{g}_0$ nilpotent, since $[N, \mathfrak{h}] = 0$, we have that ad_N commutes with any element of $\mathrm{ad}\,\mathfrak{h}$. This implies that $K(N, \mathfrak{h}) = 0$. Then for any $X \in \mathfrak{g}_0$, $K(H, \mathfrak{g}_0)$ can be broken up into $K(H, S) + K(H, N)$; since S also lives in $\mathfrak{h}$, the left term is zero by hypothesis, and we just showed why the right term is zero. Thus $K(H, \mathfrak{g}_0) = 0$, which forces $H = 0$ by nondegeneracy on $\mathfrak{g}_0$.

It is clear that $[\mathfrak{g}_0, \mathfrak{h}] = 0$, and by associativity of the Killing form, we have $K([\mathfrak{g}_0, \mathfrak{h}], \mathfrak{g}_0) = K(\mathfrak{h}, [\mathfrak{g}_0, \mathfrak{g}_0])$. This implies that $\mathfrak{h} \cap \mathfrak{g}_0^{(1)} = 0$. We will use this to show that $\mathfrak{g}_0$ is abelian. Suppose not, so that $\mathfrak{g}_0^{(1)} \neq 0$. Since $\mathfrak{g}_0$ acts on the ideal $\mathfrak{g}_0^{(1)}$ via the adjoint representation, we can think of $\mathfrak{g}_0$ as a subalgebra of $gl(\mathfrak{g}_0^{(1)})$. By Proposition 8, there exists a nonzero $Y \in \mathfrak{g}_0^{(1)}$ such that $[\mathfrak{g}_0, Y] = 0$. Therefore, $Z(\mathfrak{g}_0) \cap \mathfrak{g}_0^{(1)} \neq 0$. Choose any nonzero element in this intersection, say Y again. Then Y cannot be semisimple, else $Y \in \mathfrak{g}_0$ would be in $\mathfrak{h}$ by the above. Then the nilpotent part of Y, say N', must be nonzero, and since N' is a polynomial in Y without constant term, it also belongs to $Z(\mathfrak{g}_0)$. Like before, we then see that $\mathrm{ad}_{N'}$ commutes with ad_X for all $X \in \mathfrak{g}_0$. And $\mathrm{ad}_{N'}$ nilpotent implies that $K(N', \mathfrak{g}_0) = 0$. This contradicts the nondegeneracy of K on $\mathfrak{g}_0$, meaning $\mathfrak{g}_0$ must actually be abelian. Now if $\mathfrak{g}_0 \neq \mathfrak{h}$, then $\mathfrak{g}_0$ would contain a nonzero nilpotent element X. But since $\mathfrak{g}_0$ is abelian, ad_X nilpotent commutes with any element in $\mathrm{ad}\,\mathfrak{g}_0$, which again implies $K(X, \mathfrak{g}_0) = 0$. Therefore $\mathfrak{g}_0 = \mathfrak{h}$. □

Let R denote the set of roots. We are now able to assert that $\mathfrak{g}$ decomposes as a vector space direct sum of the Cartan subalgebra and its corresponding root spaces:

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha. \tag{15}$$

We will proceed to analyze root spaces, following along with [6].

Lemma 21. *For any α and β in $\mathfrak{h}^*$, we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$. If $\alpha + \beta \neq 0$, then $\mathfrak{g}_\alpha$ is orthogonal to $\mathfrak{g}_\beta$.*

Proof. Suppose that X is in $\mathfrak{g}_\alpha$ and Y is in $\mathfrak{g}_\beta$. Since ad_H is a derivation, we have $\mathrm{ad}_H[X, Y] = [\mathrm{ad}_H X, Y] + [X, \mathrm{ad}_H Y] = \alpha(H)[X, Y] + \beta(H)[X, Y] = (\alpha + \beta)(H)[X, Y]$. Thus, $[X, Y]$ is contained in $\mathfrak{g}_{\alpha+\beta}$. For the second assertion, choose some $H \in \mathfrak{h}$ with $(\alpha+\beta)(H) \neq 0$, and let X and Y be as above. By associativity of the Killing form, we have $\alpha(H)K(X, Y) = K([H, X], Y) = -K([X, H], Y) = -K(X, [H, Y]) = -\beta(H)K(X, Y)$, which implies that $(\alpha + \beta)(H)K(X, Y) = 0$. Since the functional $\alpha + \beta$ is nonzero, $K(X, Y) = 0$, and thus the root spaces are orthogonal. □

Note that if X is in $\mathfrak{g}_\alpha$ and Y is in $\mathfrak{g}_{-\alpha}$, then $[X, Y]$ is in $\mathfrak{h}$. Also note that if $\alpha + \beta \neq 0$ is not a root, then we must have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = 0$. Moreover, given any root α and $X_\alpha \in \mathfrak{g}_\alpha$ a corresponding root vector, the map ad_{X_α} sends $\mathfrak{g}_\beta$ to $\mathfrak{g}_{\beta+\alpha}$, $\mathfrak{g}_{\beta+\alpha}$ to $\mathfrak{g}_{\beta+2\alpha}$, and so on. Eventually the functional $\beta + n\alpha$ will not be a root, which shows that ad_X is nilpotent. Lemma 21 also implies that 2α

cannot be a root, else $\mathfrak{g}$ would be orthogonal to itself, which is absurd. We will explore multiples of roots in greater detail below.

Lemma 22. *For each root α, there exists a unique $T_\alpha \in \mathfrak{h}$ such that $\alpha(H) = K(H, T_\alpha)$.*

Proof. The restriction of K to $\mathfrak{h}$ is nondegenerate. This implies that the Killing form induces an injective linear map $T : \mathfrak{h} \to \mathfrak{h}^*$ defined by $T(H) = K(H, \cdot)$. Since any finite-dimensional vector space has the same dimension as its dual space, we see that T is an isomorphism. We can therefore identify $\mathfrak{h}^*$ with $\mathfrak{h}$ in the desired way. □

A nice application of this result is recorded in the following lemma.

Lemma 23. *The set of roots R spans $\mathfrak{h}^*$. Equivalently, the set $\{T_\alpha : \alpha \in R\}$ spans $\mathfrak{h}$.*

Proof. Suppose the set of T_α failed to span $\mathfrak{h}$. Then we can find a nonzero $T \in \mathfrak{h}$ in the orthogonal complement of the subspace spanned by the T_α. Then $K(T, T_\alpha) = 0$ for all T_α. By our identification above, this means that $\alpha(T) = 0$ for all $\alpha \in R$. Thus, for each root α, we have $[T, X_\alpha] = 0$ for any corresponding root vector X_α. But since $[T, H] = 0$ for all $H \in \mathfrak{h}$, we have that T commutes with all of $\mathfrak{g}$, and so T lies in the center $Z(\mathfrak{g})$. But $Z(\mathfrak{g})$ is trivial in semisimple Lie algebras, so we get a contradiction. □

Lemma 24. *If α is a root, then so is $-\alpha$. For any $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$, we have $[X, Y] = K(X, Y)T_\alpha$.*

Proof. Choose a root vector $X_\alpha \in \mathfrak{g}_\alpha$ and suppose $-\alpha$ is not a root. Then the functional $\alpha + \beta$ is nonzero for any root β. By Lemma 21, $K(X_\alpha, X_\beta) = 0$ for all $X_\beta \in \mathfrak{g}_\beta$, hence $K(X_\alpha, X) = 0$ for all $X \in \mathfrak{g}$. This forces X_α to be 0, a contradiction. To prove the second assertion, recall that $[X, Y] \in \mathfrak{h}$, and calculate $K(H, [X, Y]) = K([H, X], Y) = \alpha(H)K(X, Y) = K(H, T_\alpha)K(X, Y) = K(H, K(X, Y)T_\alpha)$. Thus

$$K(H, [X, Y] - K(X, Y)T_\alpha) = 0$$

for all $H \in \mathfrak{h}$. Since K is nondegenerate on $\mathfrak{h}$, it follows that $[X, Y] - K(X, Y)T_\alpha = 0$, as desired. □

Lemma 25. *The number $\alpha(T_\alpha)$ is nonzero for any root α.*

Proof. Choose a root vector $X_\alpha \in \mathfrak{g}_\alpha$. We know $K(X_\alpha, Y) = 0$ for all $Y \notin \mathfrak{g}_\alpha$, and if the Killing form vanishes for every element in $\mathfrak{g}$, nondegeneracy would force X_α to be 0, a contradiction. Thus there is some element $Y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $K(X_\alpha, Y_\alpha)$ is nonzero, and we may as well normalize it. By Lemma 24, we conclude that $[X_\alpha, Y_\alpha] = T_\alpha$. If $\alpha(T_\alpha) = 0$, then $[T_\alpha, X_\alpha] = 0$, and by the Jacobi identity, $[T_\alpha, Y_\alpha] = 0$ as well. So the subspace $\mathfrak{s}$ spanned by X_α, Y_α, and T_α is a three-dimensional solvable subalgebra. By the proof of Lie's theorem,

$T_\alpha \in \mathfrak{s}^{(1)}$ must be nilpotent, which implies ad_{T_α} is nilpotent by Lemma 7. But ad_{T_α} also semisimple forces ad_{T_α} to be 0. This means that T_α lies in the center of $\mathfrak{g}$, which is trivial. Thus $\alpha(T_\alpha)$ is actually nonzero. □

The above lemma allows us to define a vector

$$H_\alpha \equiv \frac{2}{K(T_\alpha, T_\alpha)} T_\alpha \tag{16}$$

for any root α. Furthermore, given a root vector $X_\alpha \in \mathfrak{g}_\alpha$, I know I can find a Y_α such that $K(X_\alpha, Y_\alpha)$ is nonzero. Rescale either of the vectors so that

$$K(X_\alpha, Y_\alpha) = \frac{2}{K(T_\alpha, T_\alpha)}.$$

It is not hard to check that H_α, X_α, and Y_α satisfy the commutation relations

$$[X_\alpha, Y_\alpha] = H_\alpha, \quad [H_\alpha, X_\alpha] = 2X_\alpha, \quad [H_\alpha, Y_\alpha] = -2Y_\alpha. \tag{17}$$

We see that $\mathfrak{s}_\alpha = \mathrm{span}\{H_\alpha, X_\alpha, Y_\alpha\}$ is actually a subalgebra isomorphic to $sl(2;\mathbb{C})$. This is an incredible result. We discovered that every semisimple Lie algebra $\mathfrak{g}$ contains a copy of $sl(2;\mathbb{C})$ for each $\alpha \in R$. In light of Lemma 21, $\mathfrak{s}_\alpha$ acts on the string

$$\mathfrak{g}_{\beta\alpha} \equiv \bigoplus_{n\in\mathbb{Z}} \mathfrak{g}_{\beta+n\alpha}$$

via the adjoint representation. We do not yet know that the action is irreducible, but we can still use our analysis of the irreducible representations of $sl(2;\mathbb{C})$; we just cannot conclude the vectors $u_0, \ldots, u_m$ span V.

Proposition 26. *Suppose $\alpha, \beta \in R$. Let p and q denote the largest integers for which $\beta - p\alpha$ and $\beta + q\alpha$ are roots, respectively. Then $\beta + n\alpha$ is a root for every $-p \leq n \leq q$ and is not a root otherwise.*

Proof. The string $\mathfrak{g}_{\beta\alpha}$ consists of the vectors X satisfying $\mathrm{ad}_H X = (\beta + n\alpha)(H)X$. Thus the eigenvalues of ad_{H_α} are $\beta(H_\alpha) + n\alpha(H_\alpha) = \beta(H_\alpha) + 2n$ for finitely many values of n. By our analysis of the representations of $sl(2;\mathbb{C})$, the largest and smallest eigenvalues of H_α are $\pm m$ for some m. Thus, p and q must satisfy $\beta(H_\alpha) - 2p = -m$ and $\beta(H_\alpha) + 2q = m$. Therefore $\beta(H_\alpha) = p - q$, which is an integer. The numbers $\beta(H_\alpha)$ are called the **Cartan integers**. It follows that every integer n between p and q corresponds to another eigenvalue, and so the ***α*-string through *β*** $\{\beta + n\alpha : -p \leq n \leq q\}$ is an uninterrupted sequence of roots. There are no other roots $\beta + n\alpha$ for n outside the range $-p \leq n \leq q$, else we could find a second interval bounded by some other p' and q', and in that representation $\beta(H_\alpha)$ would equal $p' - q'$, which the reader can easily verify is a contradiction. □

Note that in particular, $\beta - \beta(H_\alpha)\alpha = \beta - (p - q)\alpha$ is an element of the α-string through β, meaning $\beta - \beta(H_\alpha)\alpha$ is a root.

Recall that α in R implies $-\alpha$ is also in R. The following lemma declares that no other multiple of α is a root. This allows us to conclude that $\mathfrak{s}_\alpha$ acts on $\mathfrak{g}_{\beta\alpha}$ irreducibly.

Proposition 27. *Every root space is one-dimensional. For any $\alpha \in R$, the only multiples of α in R are $\pm\alpha$.*

Proof. Consider the subspace $\mathfrak{s}$ spanned by H_α, Y_α, and $\mathfrak{g}_{n\alpha}$ for $n \geq 1$. Then $\mathfrak{s}$ is invariant under $\mathrm{ad}_{X_\alpha}, \mathrm{ad}_{Y_\alpha}$, and ad_{H_α}. Since ad_X is nilpotent for any root vector X, we have that $\mathrm{trace}\,\mathrm{ad}_{H_\alpha} = \mathrm{trace}\,\mathrm{ad}_{[X_\alpha, Y_\alpha]} = \mathrm{trace}[\mathrm{ad}_{X_\alpha}, \mathrm{ad}_{Y_\alpha}] = 0$, where the trace is computed relative to $\mathfrak{s}$. On the other hand, we have that $\mathrm{trace}\,\mathrm{ad}_{H_\alpha} = -\alpha(H_\alpha) + \sum_n n\alpha(H_\alpha) \dim \mathfrak{g}_{n\alpha}$. Thus we need to have

$$\alpha(H_\alpha) = \sum_n n\alpha(H_\alpha) \dim \mathfrak{g}_{n\alpha},$$

hence $\dim \mathfrak{g}_\alpha = 1$ and $\dim \mathfrak{g}_{n\alpha} = 0$ for $n \geq 2$. For the second statement, suppose that $\beta = c\alpha$ is a root for some complex number c. We know that $\beta(H_\alpha) = c\alpha(H_\alpha) = 2c$ is an integer k. It suffices to assume c is positive, else we could use the root $-\alpha$ instead. Since we ruled out the case where c is a positive integer greater than 1, we take c to be a half-integer $k/2$ for k odd. Once again considering the action of $\mathfrak{s}_\alpha$ on $\mathfrak{g}_{\beta\alpha}$, we have that the α-string through $k\alpha/2$ is an uninterrupted sequence of roots. While we cannot deduce the value of q, we know p is at least k, since $k\alpha/2 - k\alpha = -k\alpha/2$ is a root by Lemma 24. Thus, we know that $-k\alpha/2$, $-k\alpha/2 + \alpha$, $\ldots$, $k\alpha/2 - \alpha$, and $k\alpha/2$ are all roots. In particular, $-k\alpha/2 + (k+1)/2\alpha = \alpha/2$ is a root. But this implies $2(\alpha/2) = \alpha$ is not a root, which yields a contradiction. Therefore, c can only be ± 1. □

We can now refine Lemma 21 into the following statement:

Proposition 28. *Given $\alpha, \beta \in \mathfrak{h}^*$ such that $\alpha + \beta \neq 0$, we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$.*

Proof. If $\alpha + \beta \in R$, the result immediately follows from the fact that each root space is one-dimensional, provided ad_{X_α} does not annihilate X_β. This would only occur if β is at the top of the α-string through β, which holds if $q = 0$. But in this case, $\alpha + \beta \notin R$, and so both sides are equal to 0. □

5 Root systems

Let $\mathfrak{h}$ and R be defined as before. There is a natural way to extend the inner product on $\mathfrak{h}$ to its dual space $\mathfrak{h}^*$, namely $(\alpha, \beta) \equiv K(T_\alpha, T_\beta)$. Since the roots span $\mathfrak{h}^*$, it suffices to define the form on the roots only. It is left to the reader to check that (α, β) is positive-definite. Observe that

$$\beta(H_\alpha) = K(H_\alpha, T_\alpha) = 2\frac{K(T_\alpha, T_\beta)}{K(T_\alpha, T_\alpha)},$$

so that the Cartan integers now emerge as

$$a_{\beta\alpha} \equiv \beta(H_\alpha) = 2\frac{(\beta, \alpha)}{(\alpha, \alpha)}. \tag{18}$$

Choose a basis of $\mathfrak{h}^*$ consisting of roots, say $\{\alpha_1, \ldots, \alpha_\ell\}$. Then for any $\beta \in R$, we have $\beta = \sum_{i=1}^{\ell} c_i \alpha_i$ for $c_i \in \mathbb{C}$. It turns out that every c_i is actually in $\mathbb{Q}$. Indeed, for each $j = 1, \ldots, \ell$, we have

$$(\beta, \alpha_j) = \sum_{i=1}^{\ell} c_i (\alpha_i, \alpha_j).$$

Now multiply both sides by $2/(\alpha_j, \alpha_j)$ to obtain

$$2\frac{(\beta, \alpha_j)}{(\alpha_j, \alpha_j)} = \sum_{i=1}^{\ell} 2\frac{(\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)} c_i.$$

Interpret this as a system of ℓ equations in ℓ unknowns c_i with Cartan integers as coefficients. The form being nondegenerate implies the *Cartan matrix*, whose ij^{th} entry is the Cartan integer $a_{\alpha_i \alpha_j}$, for this system of equations is invertible. Thus, a unique solution for the c_i exists over $\mathbb{Q}$. Let $E_{\mathbb{Q}}$ denote the $\mathbb{Q}$-subspace of $\mathfrak{h}^*$ spanned by R.

To generalize our results, we are now concerned with a fixed Euclidean space E, i.e. a finite-dimensional vector space over $\mathbb{R}$ with a positive-definite symmetric bilinear form (α, β). Let E be the real subalgebra obtained by canonically extending the base field of $E_{\mathbb{Q}}$ to $\mathbb{R}$, that is, $E = \mathbb{R} \otimes_{\mathbb{Q}} E_{\mathbb{Q}}$. All of the results we have obtained for roots still hold in the real extension. For $\alpha \in R$, we may define a geometric reflection of $\gamma \in E$ by

$$\sigma_\alpha \cdot \gamma = \gamma - 2\frac{(\beta, \alpha)}{(\alpha, \alpha)}\alpha. \tag{19}$$

The reflection σ_α fixes a hyperplane, i.e. a subspace of codimension one, given by $\{\beta \in E : (\beta, \alpha) = 0\}$. It is interesting that the formula for σ_α contains the Cartan integer $a_{\beta\alpha}$, for if γ happens to be a root, we immediately know that $\sigma_\alpha \cdot \gamma$ is also a root. As a special case, if $\gamma = \alpha$, the reflection sends α to $-\alpha$. The group of isometries generated by each σ_α is called the *Weyl group* W. See [3] and [4] for a treatment of the Weyl group. We now define the abstract notion of a *root system*, of which the set of roots we've considered so far is an example.

Definition 6. *Let E be a Euclidean space. A subset $R \subset E$ of nonzero vectors is called a* ***root system*** *in E provided that R satisfies the following axioms:*
(R1) R is finite and spans E.
(R2) For any $\alpha \in R$, the only multiples of α in R are $\pm\alpha$.
(R3) For any $\alpha \in R$, the reflection σ_α leaves R invariant.
(R4) If $\alpha, \beta \in R$, then $2\frac{(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$.

The elements of R are called *roots*. We now state some geometric properties of roots, as outlined in [3].

Proposition 29. *Suppose α and β are roots, α is not a multiple of β, and $(\alpha, \alpha) \geq (\beta, \beta)$. Let θ denote the angle between α and β. Then one of the*

following holds:
1. $(\alpha, \beta) = 0$.
2. $(\alpha, \alpha) = (\beta, \beta)$ *and* θ *is* $\pi/3$ *or* $2\pi/3$.
3. $(\alpha, \alpha) = 2(\beta, \beta)$ *and* θ *is* $\pi/4$ *or* $3\pi/4$.
4. $(\alpha, \alpha) = 3(\beta, \beta)$ *and* θ *is* $\pi/6$ *or* $5\pi/6$.

Proof. All we need is to observe that

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

and

$$\frac{a_{\alpha\beta}}{a_{\beta\alpha}} = \frac{(\alpha, \alpha)}{(\beta, \beta)} \geq 1.$$

Thus $0 \leq a_{\alpha\beta}a_{\beta\alpha} < 4$, since $\alpha \neq \pm\beta$. If $a_{\alpha\beta}a_{\beta\alpha} = 0$, then α and β are orthogonal. If $a_{\alpha\beta}a_{\beta\alpha} = 1$, then $\theta = \pi/3$ or $2\pi/3$, and α and β have the same length. If $a_{\alpha\beta}a_{\beta\alpha} = 2$, then $\theta = \pi/4$ or $3\pi/4$, and α is longer than β by a factor of $\sqrt{2}$. Finally, if $a_{\alpha\beta}a_{\beta\alpha} = 3$, then $\theta = \pi/6$ or $5\pi/6$, and α is longer than β by a factor of $\sqrt{3}$. □

Furthermore, by considering the sign of the Cartan integers in relation to the reflections they produce, we can convince ourselves that $(\alpha, \beta) > 0$ when θ is acute, $(\alpha, \beta) = 0$ when $\theta = \pi/2$, and $(\alpha, \beta) < 0$ when θ is obtuse.

Lemma 30. *Suppose* α *and* β *are roots, and let* θ *be the angle between the two. If* θ *is acute, then* $\alpha - \beta$ *and* $\beta - \alpha$ *are roots. If* θ *is obtuse, then* $\alpha + \beta$ *is a root.*

Proof. As before, take $(\alpha, \alpha) \geq (\beta, \beta)$. Whenever θ is acute, analyzing each case reveals that the projection of β onto α is $\alpha/2$, hence $\sigma_\alpha \cdot \beta = \beta - \alpha$ is a root. Thus, $-(\beta - \alpha) = \alpha - \beta$ is also a root. If θ is obtuse, then the projection of β onto α is $-\alpha/2$, hence $\sigma_\alpha \cdot \beta = \beta + \alpha$ is a root. □

Given roots α and β, we define the **α-string through β** just like before. Since R is finite, there exists integers p and q that denote the largest integers for which $\beta - p\alpha$ and $\beta + q\alpha$ are roots, respectively. Suppose the α-string through β is broken, i.e. there is an integer i, where $-p \leq i \leq q$, such that $\beta + i\alpha$ is not a root. Then there exist integers $r < s$, where $-p \leq r, s \leq q$ such that $\beta + r\alpha$ and $\beta + s\alpha$ are in R, but $\beta + (r+1)\alpha$ and $\beta + (s-1)\alpha$ are not in R. According to Lemma 30, this implies that $(\alpha, \beta + r\alpha) \geq 0$ and $(\alpha, \beta + s\alpha) \leq 0$. Combining these expressions yields

$$(s - p)(\alpha, \alpha) \leq 0.$$

Since the form on E is positive-definite, we get a contradiction. Therefore, the α-string through β $\{\beta + n\alpha : -p \leq n \leq q\}$ is unbroken. The reflection σ_α adds or subtracts a multiple of α to any root it acts on, and since σ_α sends roots to other roots, we conclude the α-string through β is invariant under σ_α. In fact, σ_α reverses the string. In particular,

$$\sigma_\alpha \cdot (\beta + q\alpha) = \beta - p\alpha.$$

The left side evaluates to $\beta - (a_{\beta\alpha} + q)\alpha$, which implies

$$a_{\beta\alpha} = p - q. \tag{20}$$

We conclude that root strings are of length at most four.

Definition 7. *Let R be a root system. A subset $\Delta \subset R$ is called a* ***base*** *if Δ is a basis for E and if each $\alpha \in R$ can be expressed as an integer combination of elements in Δ such that the coefficients are either all nonnegative or all nonpositive.*

A **positive root** is a root whose integer combination of elements in Δ has all nonnegative coefficients. The set of positive roots is labelled R^+. A positive root is **decomposable** if it is the sum of two other positive roots.

Proposition 31. *If α, β are distinct elements in a base Δ, then $(\alpha, \beta) \leq 0$.*

Proof. If $(\alpha, \beta) > 0$, then the angle between α and β would be acute, in which case $\alpha - \beta \in R$ by Lemma 30. Then the unique integer combination for $\alpha - \beta$ using elements of Δ should have either all nonnegative or all nonpositive coefficients. But $\alpha - \beta$ has one positive coefficient and one negative coefficient, and so $\alpha - \beta$ cannot be a root. □

Consequently, the angle between two distinct roots in a base is either right or obtuse.

Lemma 32. *There exists a hyperplane V through the origin in E that does not contain any roots.*

Proof. For each $\alpha \in R$, define a hyperplane $V_\alpha = \{H \in E : (\alpha, H) = 0\}$. Since the set of V_α is finite, it can be shown that their union is not all of E. So there exists some $H \in E$ not contained in any V_α. This means that H is not orthogonal to any root. Let V be the hyperplane through the origin that is orthogonal to H. Then certainly V cannot contain any roots. □

Theorem 33. *Suppose R is a root system, V a hyperplane through the origin not containing any roots, and R^+ the set of roots lying on a fixed side of V. Then the indecomposable elements in R^+ form a base.*

Proof. Choose a nonzero vector $H \in E$ so that the fixed side of V consists of $\mu \in E$ such that $(\mu, H) > 0$. Let Δ consist of the indecomposable elements of R^+. I will first show that any positive root is a nonnegative integer combination of elements in Δ. If not, then among the roots where this fails, choose the root α so that (α, H) is as small as possible. Now α is decomposable because $\alpha \notin \Delta$, so $\alpha = \beta + \gamma$ for some $\beta, \gamma \in R^+$. Then $(\alpha, H) = (\beta, H) + (\gamma, H)$, where (β, H) and (γ, H) are both greater than 0. But β and γ cannot both be expressed as nonnegative integer combinations of Δ, else α would be as well, contradicting the minimality of α.

Next, we need to show that for distinct elements $\alpha, \beta \in \Delta$, we have $(\alpha, \beta) \leq 0$. Well, if $(\alpha, \beta) > 0$, then $\alpha - \beta$ and $\beta - \alpha$ are roots. One of them must be in

R^+, and so without loss of generality suppose $\alpha - \beta \in R^+$. But then α would be decomposable, since $\alpha = (\alpha - \beta) + \beta$, which gives a contradiction.

Third, we must demonstrate that the elements of Δ are linearly independent. Suppose

$$\sum_{\alpha \in \Delta} c_\alpha \alpha = 0$$

for some constants c_α. Partition Δ depending on the sign of the coefficient, so that

$$\sum c_\alpha \alpha = \sum d_\beta \beta.$$

for nonnegative constants c_α and d_β. Let u denote this vector. Then

$$(u, u) = (\sum c_\alpha \alpha, \sum d_\beta \beta) = \sum \sum c_\alpha d_\beta (\alpha, \beta) \leq 0,$$

which forces u to vanish. Then $(u, H) = \sum c_\alpha (\alpha, H) = 0$, and since (α, H) is always positive, we must have that the c_α are all 0. Likewise, the d_β are all 0.

We are now able to conclude that Δ is a base. Its elements are linearly independent, and every positive root can be expressed as integer combinations of elements in Δ with nonnegative coefficients. The remaining roots are in R^-, and since they are the negatives of the positive roots, they can be expressed as integer combinations of elements in Δ with nonpositive coefficients. And since R spans E, it follows that Δ also spans E, hence Δ is a basis for E. □

This theorem motivates the term **positive simple roots** to refer to the elements of Δ, where here *simple* refers to being indecomposable.

5.1 Dynkin diagrams

Given a base Δ for a root system R, it is helpful to construct the **Dynkin diagram** for R. First, denote every positive simple root α_i by a vertex v_i. Then connect every pair of vertices v_i, v_j with $a_{\alpha_i \alpha_j} a_{\alpha_j \alpha_i}$ edges, which we've shown signifies the square of the relative lengths of the roots. If it happens that one root is longer than another, draw an arrow in the direction of the shorter root.

See [3] for an explanation that the Weyl group W acts transitively on the bases of a root system, which implies that the Dynkin diagrams corresponding to different bases for the same root system are isomorphic. Furthermore, if the Dynkin diagrams of two root systems R_1 and R_2 are isomorphic, then R_1 and R_2 are isomorphic. This tells us that a Dynkin diagram uniquely determines its root system.

A root system R is said to be *reducible* if it is a direct sum of two other root systems, and R is *irreducible* otherwise. Also, a Dynkin diagram is *connected* if there is a path of edges between every pair of vertices, and it is *disconnected* otherwise.

Proposition 34. *A root system is irreducible if and only if its Dynkin diagram is connected.*

Proof. Suppose R reduces into root systems R_1 and R_2. Then we quickly see that a base Δ for R decomposes into respective bases for R_1 and R_2 as $\Delta = \Delta_1 \cup \Delta_2$. Since the elements of R_1 are orthogonal to those of R_2, the elements of Δ_1 are likewise orthogonal to those of Δ_2. We conclude that the Dynkin diagram of Δ_1 is disconnected from the Dynkin diagram of Δ_2, hence Δ is disconnected. Conversely, suppose the Dynkin diagram of R is disconnected, so that $\Delta = \Delta_1 \cup \Delta_2$ as before. Then E is the direct sum of span$\{\Delta_1\}$ and span$\{\Delta_2\}$. Set $R_1 = R \cap \text{span}\{\Delta_1\}$ and $R_2 = R \cap \text{span}\{\Delta_2\}$. Then we easily see that R_1 and R_2 are root systems with bases Δ_1 and Δ_2, respectively. It remains to check that each element of R is in either R_1 or R_2. One can show that the Weyl group of any root system is generated by the set of reflections by elements of its base. Furthermore, one can show that every element of R is part of some base. Let W_1 and W_2 denote the Weyl group of R_1 and R_2, respectively. Since W acts transitively on the bases of R, it is clear that

$$W \cdot \Delta = (W_1 \cdot \Delta_1) \cup (W_2 \cdot \Delta_2),$$

and therefore R is the direct sum of R_1 and R_2. □

Proposition 35. *Let $\mathfrak{g}$ be a simple Lie algebra with corresponding root system R. Then R is irreducible.*

Proof. Suppose R reduces into R_1 and R_2. Take $\alpha \in R_1$ and $\beta \in R_2$. Then neither $(\alpha+\beta, \alpha)$ nor $(\alpha+\beta, \beta)$ is zero, which means $\alpha+\beta$ cannot belong to either R_1 or R_2, i.e. $\alpha+\beta$ is not a root. This implies $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = 0$. Then the subalgebra $\mathfrak{s}$ of $\mathfrak{g}$ generated by the root spaces associated to R_1 is a proper subalgebra of $\mathfrak{g}$; it is nonzero because the center of $\mathfrak{g}$ is trivial. Moreover, $\mathfrak{s}$ is *normalized* by all of $\mathfrak{g}$, which means $\mathfrak{s}$ is a proper ideal of $\mathfrak{g}$. This contradicts the simplicity of $\mathfrak{g}$. □

Now suppose $\mathfrak{g}$ is a semisimple Lie algebra with Cartan subalgebra $\mathfrak{h}$. Then $\mathfrak{g}$ can be decomposed into simple subalgebras $\mathfrak{g}_1, \dots, \mathfrak{g}_n$ as in Theorem 17. One can show that $\mathfrak{h} = \mathfrak{h}_1 \oplus \cdots \oplus \mathfrak{h}_n$, where $\mathfrak{h}_i = \mathfrak{g}_i \cap \mathfrak{h}$. Each $\mathfrak{h}_i$ is a maximal toral subalgebra of $\mathfrak{g}_i$. Indeed, any toral subalgebra of $\mathfrak{g}_i$ larger than $\mathfrak{h}_i$ would yield a toral subalgebra larger than $\mathfrak{h}$. Let R_i be the root system of $\mathfrak{g}_i$ relative to $\mathfrak{h}_i$. Then if $\alpha \in R_i$, we can extend α to a root of $\mathfrak{g}$ relative to $\mathfrak{h}$ by declaring $\alpha(\mathfrak{h}_j) = 0$ whenever $j \neq i$. Conversely, if $\alpha \in R$, then we must have $[\mathfrak{h}_\mathfrak{i}, \mathfrak{g}_\alpha] \neq 0$ for some i, otherwise $\mathfrak{h}$ would centralize $\mathfrak{g}_\alpha$, contradicting Proposition 28. But then $\mathfrak{g}_\alpha$ is contained in $\mathfrak{g}_i$, so that α restricted to $\mathfrak{h}_i$ is a root of $\mathfrak{g}_i$ relative to $\mathfrak{h}_i$. Therefore, R decomposes accordingly into $R_1 \cup \cdots \cup R_n$. We arrive at the following corollary:

Corollary 36. *Let $\mathfrak{g}$ be a semisimple Lie algebra with Cartan subalgebra $\mathfrak{h}$ and root system R. If $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$ is the Lie algebra direct sum of simple subalgebras, then each $\mathfrak{h}_i = \mathfrak{h} \cap \mathfrak{g}_i$ is a Cartan subalgebra of $\mathfrak{g}_i$ with corresponding root system R_i such that $R = R_1 \cup \cdots \cup R_n$ is the decomposition of R into irreducible root systems.*

I will now classify all irreducible root systems using their associated Dynkin diagrams: see Figure 1. Note that in the figure, in place of arrows we use filled vertices to represent longer roots. By Proposition 34, we only need to analyze connected Dynkin diagrams. The proof relies on [4].

Theorem 37. *An irreducible root system of rank ℓ is isomorphic to one of A_ℓ ($\ell \geq 1$), B_ℓ ($\ell \geq 2$), C_ℓ ($\ell \geq 3$), D_ℓ ($\ell \geq 4$), G_2, F_4, E_6, E_7, or E_8.*

Proof. Consider a set $\mathfrak{U} = \{\varepsilon_1, \dots, \varepsilon_n\}$ of n linearly independent unit vectors that satsify $(\varepsilon_i, \varepsilon_j) \leq 0$ and $4(\varepsilon_i, \varepsilon_j)^2 = 0$, 1, 2, or 3, for $i \neq j$. Call such a set **admissable**. As we now know, positive simple elements are admissable, once normalized. Create a graph Γ corresponding to $\mathfrak{U}$ by drawing n vertices and connecting pairs of vertices i and j with $4(\varepsilon_i, \varepsilon_j)^2$ edges. My first claim is that the number of pairs of vertices in Γ connected by at least one edge is strictly less than n. To see this, let $\varepsilon = \sum_{i=1}^n \varepsilon_i$, which is nonzero since the ε_i are linearly independent. Then $0 < (\varepsilon, \varepsilon) = n + 2\sum_{i<j}(\varepsilon_i, \varepsilon_j)$. Let i and j be a pair of vertices connected by at least one edge. Then $4(\varepsilon_i, \varepsilon_j)^2 = 1$, 2, or 3, and so $2(\varepsilon_i, \varepsilon_j) \leq -1$. The number of such pairs cannot exceed $n-1$. The claim immediately implies that Γ contains no cycles. Indeed, if $\Gamma' \subset \Gamma$ were a cycle with n' vertices, it would correspond to an admissable subset $\mathfrak{U}' \subset \mathfrak{U}$, and Γ' would contain at least n' pairs of vertices connected by at least one edge.

My second claim is that no more than three edges can originate from a vertex in Γ. Choose some $\varepsilon \in \mathfrak{U}$, and suppose the vectors in $\mathfrak{U}$ connected to ε are $\eta_1, \dots, \eta_k$. Seeing as there are no cycles, we must have that the η_i are orthonormal. Since $\mathfrak{U}$ is a linearly independent set, the $(k+1)$-dimensional span of $\varepsilon, \eta_1, \dots, \eta_k$ contains a unit vector η_0 orthogonal to each η_i. Then we have $\varepsilon = \sum_{i=0}^k (\varepsilon, \eta_i)\eta_i$. Hence $1 = (\varepsilon, \varepsilon) = \sum_{i=0}^k (\varepsilon, \eta_i)^2$, and so $\sum_{i=1}^k (\varepsilon, \eta_i)^2 < 1$. But then $\sum_{i=1}^k 4(\varepsilon, \eta_i)^2 < 4$, where the left hand side is the number of edges connected to ε. An immediate consequence of this result is that G_2 is the only connected graph of an admissable set to contain a triple edge.

Third, suppose a subset $\{\eta_1, \dots, \eta_k\}$ of $\mathfrak{U}$ yields a subgraph of Γ which is a *simple chain.* I claim that the set $\mathfrak{U}' = (\mathfrak{U} \setminus \{\eta_1, \dots, \eta_k\}) \cup \eta$, where $\eta = \sum_{i=1}^k \eta_i$, is admissable. Linear independence of $\mathfrak{U}'$ follows at once from linear independence of $\mathfrak{U}$. Since we are dealing with a simple chain, we have $2(\eta_i, \eta_{i+1}) = -1$, for $1 \leq i \leq k-1$, and so $(\eta, \eta) = k + 2\sum_{i<j}(\eta_i, \eta_j) = k - (k-1) = 1$, which shows that η is a unit vector. For any $\varepsilon \in \mathfrak{U}' \setminus \eta$, we must have that ε is connected to at most one of $\eta_1, \dots, \eta_k$, else we would draw a cycle. So either $(\varepsilon, \eta) = (\varepsilon, \eta_i)$ for some $1 \leq i \leq k$ or $(\varepsilon, \eta) = 0$ altogether. This implies not only that $\mathfrak{U}'$ is admissable, but also that the graph of $\mathfrak{U}'$ is obtained from shrinking the simple chain to a point. Therefore, we cannot have subgraphs of Γ containing a simple chain with two additional edges emerging from both endpoints, else shrinking the chain would yield a point connected to four edges.

Any connected graph Γ of an admissable set is already quite restricted. It is either G_2, a simple chain A_ℓ, a simple chain with a second edge connecting one pair of adjacent vertices, or three simple chains that intersect at one vertex. Indeed, if Γ contained more than one double-edge or branch point, then there

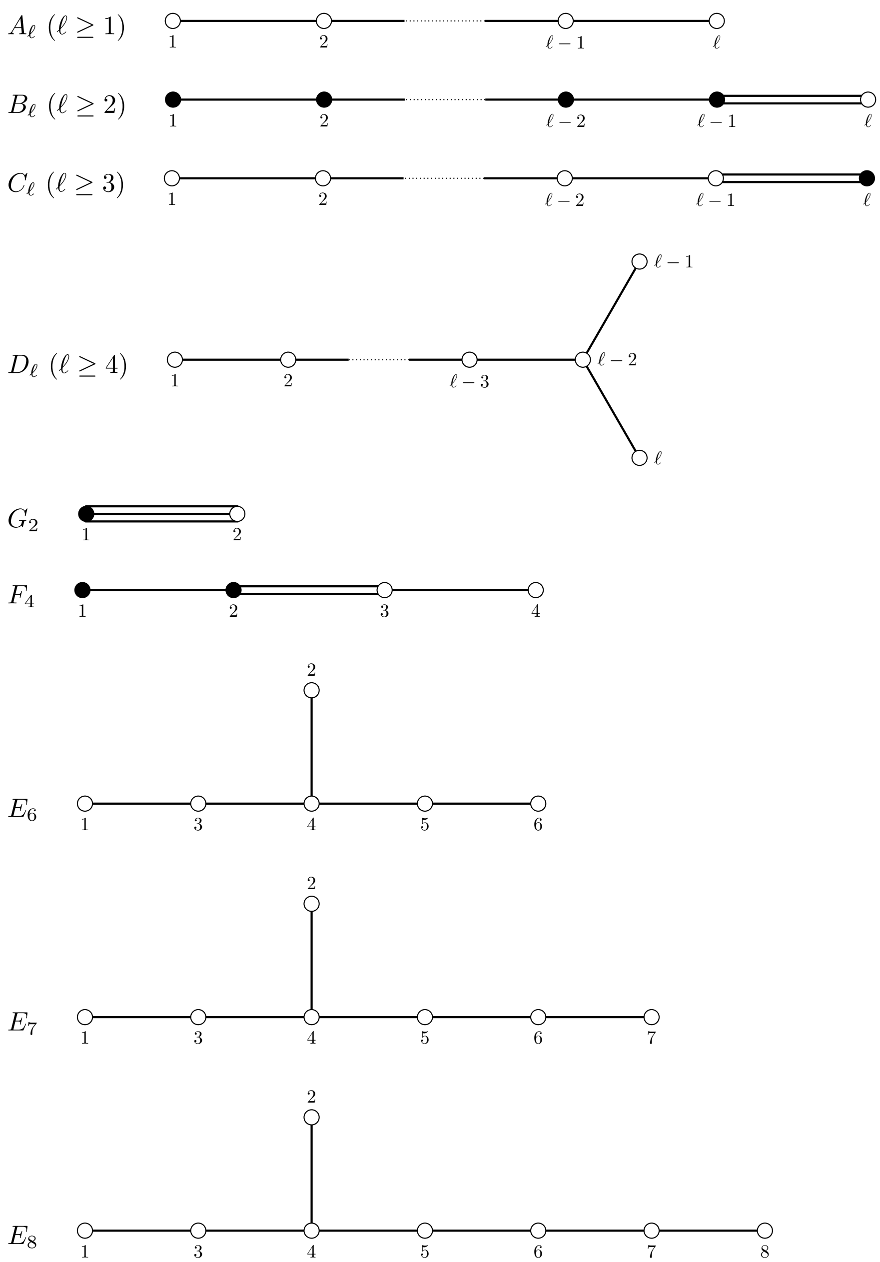

Figure 1: A complete list of the Dynkin diagrams associated to irreducible root systems. All complex semisimple Lie algebras are classified by these diagrams.

would be a subgraph consisting of a simple chain with two additional edges emerging from each endpoint, which is exactly what I just showed cannot occur.

Consider first the case of a simple chain from ε_1 to ε_p, a double-edge between ε_p and η_q, and a simple chain from η_q to η_1. Of course, the vertices still arise from an admissable set. Say that $\varepsilon = \sum_{i=1}^{p} i\varepsilon_i$ and $\eta = \sum_{i=1}^{q} i\eta_i$. Since $2(\varepsilon_i, \varepsilon_{i+1}) = -1$, for $1 \leq i \leq p-1$, and since the other pairs are orthogonal, it is easily seen that $(\varepsilon, \varepsilon) = \sum_{i=1}^{p} i^2 - \sum_{i=1}^{p-1} i(i+1) = p(p+1)/2$. Likewise, $(\eta, \eta) = q(q+1)/2$. Also, since $4(\varepsilon_p, \eta_q)^2 = 2$, we have $(\varepsilon, \eta)^2 = p^2q^2(\varepsilon_p, \eta_q)^2 = p^2q^2/2$. By the Cauchy-Schwarz inequality, $(\varepsilon, \eta)^2 < (\varepsilon, \varepsilon)(\eta, \eta)$, and thus

$$\frac{p^2q^2}{2} < \frac{p(p+1)}{2}\frac{q(q+1)}{2}.$$

Simplifying this expression gives $2pq < (p+1)(q+1)$, from which we can quickly deduce that $(p-1)(q-1) < 2$. One solution is $p = 2$ and $q = 2$, which yields the graph F_4. The remaining solutions are $p = 1$ with q arbitrary, and $q = 1$ with p arbitrary. By the symmetry of the graph, these solutions both correspond to B_ℓ and C_ℓ, where the only difference between B_ℓ and C_ℓ is the relative lengths of the roots.

Finally, we consider the case of a simple chain from ε_1 to ε_{p-1}, a branching vertex which we will label ψ, followed by two simple chains $\eta_{q-1}, \ldots, \eta_1$ and $\zeta_{r-1}, \ldots, \zeta_1$. Similar to before, set $\varepsilon = \sum_{i=1}^{p} i\varepsilon_i$, $\eta = \sum_{i=1}^{q} i\eta_i$, and $\zeta = \sum_{i=1}^{r} i\zeta_i$. Replacing p with $p-1$ reveals that $(\varepsilon, \varepsilon) = p(p-1)/2$, and likewise for η and ζ. Let θ_1, θ_2, and θ_3 denote the angles between ψ and ε, η, and ζ, respectively. We compute

$$\cos^2\theta_1 = \frac{(\varepsilon, \psi)^2}{(\varepsilon, \varepsilon)(\psi, \psi)} = \frac{(p-1)^2(\varepsilon_{p-1}, \psi)^2}{(\varepsilon, \varepsilon)} = \frac{p-1}{2p} = \frac{1}{2}(1 - \frac{1}{p}),$$

and likewise for η and ζ. Now ε, η, and ζ are mutually orthogonal, while ψ is linearly independent from them, so $\text{span}\{\varepsilon, \eta, \zeta, \psi\}$ is a four-dimensional subspace. Let the orthonormal basis for this space be $\{\varepsilon, \eta, \zeta, \kappa\}$, where κ is a unit vector orthogonal to ε, η, and ζ. Then $\psi = (\psi, \varepsilon)\varepsilon + \cdots + (\psi, \kappa)\kappa$, from which we deduce that $1 = (\psi, \psi) = (\psi, \varepsilon)^2 + \cdots + (\psi, \kappa)^2$. So then

$$1 > (\psi, \varepsilon)^2 + (\psi, \eta)^2 + (\psi, \zeta)^2 = \cos^2\theta_1 + \cos^2\theta_2 + \cos^2\theta_3.$$

We conclude that

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1,$$

where p, q, and r are at least 2, else the graph reduces to A_ℓ. If each variable were at least 3, then the inequality would fail; without loss of generality, say that $r = 2$. Thus $1/p + 1/q > 1/2$, and again without loss of generality, take $p \geq q$. This implies that $q < 4$. If $q = 2$, then p is arbitrary. This solution corresponds to D_ℓ. If $q = 3$, then $p = 3$, 4, or 5, which correspond to E_6, E_7, and E_8, respectively.

I will not delve into the details of constructing root systems associated to the Dynkin diagrams that remain, but it is possible to do so. Therefore, there are no more restrictions on the Dynkin diagrams, and the proof terminates. □

See [4] for further explanation that for every root system R there exists a semisimple Lie algebra that gives rise to a root system isomorphic to R.

To summarize our findings, we know that any semisimple Lie algebra $\mathfrak{g}$, with root system R, decomposes into a Lie algebra direct sum of simple subalgebras. And each simple subalgebra produces a decomposition of R into irreducible root systems. Finally, each of these irreducible root systems corresponds to a connected Dynkin diagram, which must be listed in Figure 1. Therefore, every complex semisimple Lie algebra is classified by a finite union of connected Dynkin diagrams of the type A_ℓ ($\ell \geq 1$), B_ℓ ($\ell \geq 2$), C_ℓ ($\ell \geq 3$), D_ℓ ($\ell \geq 4$), G_2, F_4, E_6, E_7, and E_8.

6 An application to physics

6.1 Lorentz group

An important Lie group that arises in physics is the Lorentz group. Our discussion will rely heavily on [5]. The invariance of the speed of light c enforces a *metric tensor*

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Space-time, or *Minkowski space*, is said to be $\mathbb{R}^4$ with metric tensor $\eta_{\mu\nu}$, which is usually called the *Minkowski metric*. In four-vector notation, we denote points in Minkowski space as x^μ. We say that $x^\mu = (ct, \boldsymbol{x})$ and $x_\mu = (ct, -\boldsymbol{x})$. Then the inner product of x^μ with itself, or rather the norm squared of x^μ, is given by $x_\mu x^\mu = \eta_{\mu\nu} x^\nu x^\mu = (ct)^2 - |\boldsymbol{x}|^2$. If a linear transformation $x^\mu \mapsto x'^\mu = \Lambda^\mu{}_\nu x^\nu$ preserves the norm squared, then

$$\eta_{\mu\nu} x^\mu x^\nu = \eta_{\rho\sigma} x'^\rho x'^\sigma = \eta_{\rho\sigma} \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu x^\mu x^\nu,$$

which implies that

$$\eta_{\mu\nu} = \eta_{\rho\sigma} \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu = \Lambda^\rho{}_\mu \eta_{\rho\sigma} \Lambda^\sigma{}_\nu. \tag{21}$$

Alternatively,

$$\eta = \Lambda^T \eta \Lambda. \tag{22}$$

Such transformations are called *Lorentz transformations*, and they form a real matrix Lie group called the *Lorentz group* $O(3,1)$. From Equation 22, it is clear that $\det \Lambda = \pm 1$ for any Λ. Also, setting $\mu, \nu = 0$ in Equation 21 yields

$$(\Lambda^0{}_0)^2 - \sum_{i=1}^{3} (\Lambda^i{}_0)^2 = 1,$$

which implies that $(\Lambda^0{}_0)^2 \geq 1$. The Lorentz group may then be written as a union of four connected components, depending on the signs of $\det \Lambda$ and $\Lambda^0{}_0$.

The set $SO(3,1)^\uparrow = \{\Lambda \in O(3,1) : \det \Lambda = 1,\ \Lambda^0{}_0 = 1\}$ is called the *proper orthochronous* Lorentz group. It is a normal subgroup, as it is the kernel of the map that takes Λ to the pair $(\det \Lambda, \operatorname{sgn} \Lambda^0{}_0)$. The remaining three components of $O(3,1)$ are simply cosets of $SO(3,1)^\uparrow$. Frequently, when we speak of the Lorentz group we mean just the proper orthochronous Lorentz group.

We will now motivate why the universal covering group of $SO(3,1)^\uparrow$ is $SL(2;\mathbb{C})$. Letting σ_0 be the 2×2 identity matrix, we can represent any point x^μ in Minkowski space by

$$x^\mu \sigma_\mu = X = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix}.$$

The representation is invertible, as we can recover x^μ via

$$x^\mu = \frac{1}{2} \operatorname{trace} X \sigma_\mu.$$

Moreover, the norm squared of x^μ is simply given by $\det X$. For any $M \in SL(2;\mathbb{C})$, the map $X \mapsto X' = MXM^\dagger$ preserves the determinant, and thus it represents a Lorentz transformation. So there exists a $\Lambda_M \in SO(3,1)^\uparrow$ such that

$$X' = \Lambda_M X.$$

We immediately see that M and $-M$ both give rise to the same Lorentz transformation, so under this identification there is a two-to-one homomorphism of $SL(2,\mathbb{C})$ into $SO(3,1)^\uparrow$. It can be shown that such a homomorphism sending $\pm M$ to Λ_M is onto, continuous, and locally one-to-one. Furthermore, by analyzing the polar decomposition of a matrix in $SL(2;\mathbb{C})$, one can show that $SL(2;\mathbb{C})$ is simply connected. Therefore, $SL(2;\mathbb{C})$ is the universal covering group of $SO(3,1)^\uparrow$.

Now consider an infinitesimal Lorentz transformation of the form

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu. \tag{23}$$

Substituting this transformation into Equation 21 gives

$$\eta_{\rho\sigma} = \eta_{\mu\nu}(\delta^\mu{}_\rho + \omega^\mu{}_\rho)(\delta^\nu{}_\sigma + \omega^\nu{}_\sigma) = (\eta_{\rho\nu} + \omega_{\nu\rho})(\delta^\nu{}_\sigma + \omega^\nu{}_\sigma).$$

Neglecting the quadratic term in ω, we arrive at

$$\eta_{\rho\sigma} = \eta_{\rho\sigma} + \omega_{\sigma\rho} + \omega_{\rho\sigma},$$

and consequently ω is antisymmetric. Any 4×4 antisymmetric matrix has six independent entries, which means that Lorentz group has six continuous parameters. Three transformations correspond to the usual rotations of $SO(3)$, which leave t invariant, and they are parameterized by the three rotation angles. The remaining transformations leave $t^2 - j^2$ invariant for $j = x, y, z$, and each is called a *boost* along its respective axis. In natural units, boosts can be written as

$$t \mapsto \gamma(t + vx), \quad x \mapsto \gamma(x + vt), \tag{24}$$

where

$$\gamma \equiv \frac{1}{\sqrt{1-v^2}}, \quad -1 < v < 1.$$

These transformations are parameterized by the components of the velocity $\boldsymbol{v}$. It is often convenient to reparameterize v in terms of *rapidity* η, so that $v = \tanh \eta$ with $-\infty < \eta < \infty$. This illustrates that boosts are hyperbolic transformations given by

$$t \mapsto (\cosh \eta)t + (\sinh \eta)x, \quad x \mapsto (\sinh \eta)t + (\cosh \eta)x. \tag{25}$$

Seeing as the boost velocity $\boldsymbol{v}$ ranges over the non-compact interval $0 \leq |\boldsymbol{v}| < 1$, it follows that the Lorentz group is non-compact. This is troubling, as there is a theorem that states non-compact groups have no nontrivial finite-dimensional unitary representations. Unitary representations are desired in physics because their generators are Hermitian operators, which correspond to observables. In order to identify non-compact groups with observables, we require infinite-dimensional representations. This problem is overcome using the Hilbert space of one-particle states.

6.2 Lorentz algebra

Now we proceed to analyze the Lorentz algebra. We know that the Lorentz group has six continuous parameters, which are entries in the antisymmetric matrix $\omega_{\mu\nu}$. Let's label their corresponding generators as $J^{\mu\nu}$, such that $J^{\mu\nu} = -J^{\nu\mu}$. Then an element Λ of the Lorentz group is expressed as

$$\Lambda = e^{-\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}}, \tag{26}$$

where the factor of $1/2$ arises because each generator is counted twice in our implicit summation. Suppose we have a representation D of the Lorentz group, of dimension n. Then a collection of objects ϕ^i, for $1 \leq i \leq n$, transforms under D whenever

$$\phi^i \mapsto \left[e^{-\frac{i}{2}\omega_{\mu\nu}J_D^{\mu\nu}}\right]^i{}_j \phi^j, \tag{27}$$

where the exponential is an $n \times n$ matrix representation of Λ and $J_D^{\mu\nu}$ are the $n \times n$ Lorentz generators in D. Therefore, the variation of ϕ^i is

$$\delta\phi^i = -\frac{i}{2}\omega_{\mu\nu}[J_D^{\mu\nu}]^i{}_j\phi^j. \tag{28}$$

Say that a *contravariant four-vector* V^μ is an object that satisfies the Lorentz transformation law $V^\mu \mapsto \Lambda^\mu{}_\nu V^\nu$. A *covariant four-vector* V_μ is an object that transforms as $V_\mu \mapsto \Lambda_\mu{}^\nu V_\nu$, where $\Lambda_\mu{}^\nu = \eta_{\mu\rho}\eta^{\nu\sigma}\Lambda^\rho{}_\sigma$. Given a contravariant four-vector V^μ, it can be shown that the four-vector $V_\mu \equiv \eta_{\mu\nu}V^\nu$ is a covariant four-vector. We say that a *scalar* is a quantity that is invariant under the Lorentz transformation. The rest mass of a particle is an example of a scalar. For a scalar ϕ, the index i has only one value, meaning the representation is one-dimensional. But in order for the Lorentz transformation on ϕ to be the identity, we must have $J^{\mu\nu} = 0$. Thus, the representation is trivial.

Now we discuss the four-vector representation (which is four-dimensional). Consider the variation of a contravariant four-vector V^μ under an infinitesimal Lorentz transformation as given by Equation 23. We have

$$V^\mu \mapsto \Lambda^\mu{}_\nu V^\nu = (\delta^\mu{}_\nu + \omega^\mu{}_\nu)V^\nu,$$

so that the variation is

$$\delta V^\mu = \omega^\mu{}_\nu V^\nu.$$

But using Equation 28, we know that

$$\delta V^\mu = -\frac{i}{2}\omega_{\mu\nu}[J^{\mu\nu}]^\rho{}_\sigma V^\sigma,$$

where the matrix indices in the four-vector representation are conventionally replaced by four-vector indices. In order to make both equations compatible, we require the solution

$$[J^{\mu\nu}]^\rho{}_\sigma = i(\eta^{\mu\rho}\delta^\nu{}_\sigma - \eta^{\nu\rho}\delta^\mu{}_\sigma). \tag{29}$$

Note that this matrix is antisymmetric with respect to exchanging $\mu \leftrightarrow \nu$. To verify the solution is correct, substitute it into our expression for the variation to get

$$\begin{aligned}
\delta V^\rho &= \frac{1}{2}\omega_{\mu\nu}(\eta^{\mu\rho}\delta^\nu{}_\sigma - \eta^{\nu\rho}\delta^\mu{}_\sigma)V^\sigma \\
&= \frac{1}{2}\omega_{\mu\nu}\eta^{\mu\rho}\delta^\nu{}_\sigma V^\sigma + \frac{1}{2}\omega_{\nu\mu}\eta^{\nu\rho}\delta^\mu{}_\sigma V^\sigma \\
&= \frac{1}{2}\omega^\rho{}_\nu V^\nu + \frac{1}{2}\omega^\rho{}_\mu V^\mu \\
&= \omega^\rho{}_\nu V^\nu,
\end{aligned}$$

as desired. Note that in the second equality I utilized the antisymmetry of ω.

The four-vector representation is irreducible. Using Equation 29, we find that the commutator is

$$[J^{\mu\nu}, J^{\rho\sigma}] = i(\eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\rho}J^{\nu\sigma} - \eta^{\nu\sigma}J^{\mu\rho} + \eta^{\mu\sigma}J^{\nu\rho}), \tag{30}$$

which completely determines the Lie algebra of the Lorentz group. Define a new set of vectors by breaking up the generators as follows:

$$J^i \equiv \epsilon_{ijk}J^{jk}, \quad K^i \equiv J^{i0}. \tag{31}$$

Their commutation relations are:

$$\begin{aligned}
[J^i, J^j] &= i\epsilon_{ijk}J^k, \\
[J^i, K^j] &= i\epsilon_{ijk}K^k, \\
[K^i, K^j] &= -i\epsilon_{ijk}J^k.
\end{aligned}$$

The first relation implies the J^i form an $sl(2;\mathbb{C})$ algebra, and so we interpret J^i as the angular momentum. Observe that the K^i do not form an algebra. Finally, define

$$\theta^i \equiv \frac{1}{2}\epsilon_{ijk}\omega^{jk}, \quad \eta^i \equiv \epsilon_{ijk}\omega^{i0}. \tag{32}$$

Then since $\omega_{i0} = -\omega^{i0} = -\eta^i$, else $\omega_{jk} = \omega^{jk}$, we can write

$$\begin{aligned}\frac{1}{2}\omega_{\mu\nu}J^{\mu\nu} &= \omega_{12}J^{12} + \omega_{13}J^{13} + \omega_{23}J^{23} + \sum_{i=1}^{3}\omega_{i0}J^{i0} \\ &= \boldsymbol{\theta}\cdot\boldsymbol{J} - \boldsymbol{\eta}\cdot\boldsymbol{K}.\end{aligned}$$

Consequently, an element of the Lorentz group may be written as

$$\Lambda = e^{-i\boldsymbol{\theta}\cdot\boldsymbol{J}+i\boldsymbol{\eta}\cdot\boldsymbol{K}}. \tag{33}$$

We interpret $\boldsymbol{K}$ as a spatial vector and $\boldsymbol{\theta}$ as a rotation angle.

Our treatment of the Lorentz group and algebra will allow one to explore further topics in general relativity and quantum field theory.

References

[1] Michael Artin. *Algebra*. 2nd ed. Pearson Modern Classics. Pearson, 2018.

[2] Howard Georgi. *Lie Algebras in Particle Physics: From Isospin to Unified Theories*. 2nd ed. Frontiers in Physics. Westview Press, 1999.

[3] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. 2nd ed. Graduate Texts in Mathematics. Springer, 2015.

[4] James E. Humphreys. *Introduction to Lie Algebras and Representation Theory*. Graduate Texts in Mathematics. Springer-Verlag, 1972.

[5] Michele Maggiore. *A Modern Introduction to Quantum Field Theory*. Oxford Master Series in Physics. Oxford University Press, 2005.

[6] D.H. Sattinger and O.L. Weaver. *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*. Applied Mathematical Sciences. Springer-Verlag, 1986.