Irreducible Representations of Complex Semisimple Lie Algebras

Rush Brown

September 8, 2016

Abstract

In this paper we give the background and proof of a useful theorem classifying irreducible representations of semisimple algebras by their highest weight, as well as restricting what that highest weight can be.

Contents

1 Introduction

In this paper I build up to a useful theorem characterizing the irreducible representations of semisimple Lie algebras, namely that finite-dimensional irreducible representations are defined up to isomorphism by their highest weight ω and that $\omega(H_\alpha)$ is an integer for any root α of R. This is the main theorem Fulton and Harris use for showing that the Weyl construction for $\mathfrak{sl}_n$ gives all (finite-dimensional) irreducible representations.

In the first half of the paper we'll see some of the general theory of semisimple Lie algebras—building up to the existence of Cartan subalgebras—for which we will use a mix of Fulton and Harris and Serre ([1] and [2]), with minor changes where I thought the proofs needed less or more clarification (especially in the proof of the existence of Cartan subalgebras). Most of them are, however, copied nearly verbatim from the source. In the second half of the paper we will describe the roots of a semisimple Lie algebra with respect to some Cartan subalgebra, the weights of irreducible representations, and finally prove the promised result.

Although I've tried to be fairly thorough, a couple of the theorems used are too long or too tangential to include full proofs. For these I've included citations. This paper also leaves out some important parts of the theory that aren't strictly necessary for proving the final theorem (such as the Weyl groups of root systems and their relationship to automorphisms of semisimple Lie algebras fixing a Cartan subalgebra), so this summary should not be considered comprehensive.

In this paper $k = \mathbb{C}$, although any algebraically closed field of characteristic 0 will work for most of the general theory.

2 Lie Algebras

Definition 2.1. *A Lie algebra is a vector space $\mathfrak{g}$ (in this paper it will always be finite-dimensional) together with a binary operation $[\cdot, \cdot]$ called a Lie bracket which is bilinear, skew-symmetric, and satisfies the Jacobi identity: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ for $X, Y, Z \in \mathfrak{g}$.*

Definition 2.2. *An ideal $\mathfrak{g}'$ of a Lie algebra $\mathfrak{g}$ is a sub-Lie algebra such that $[X, Y] \in \mathfrak{g}'$ for $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}'$.*

Although the bracket $[\cdot, \cdot]$ may not be associative, the first isomorphism theorem still holds for Lie algebras.

A Lie algebra is called abelian if $[\cdot, \cdot] \equiv 0$.

For any finite-dimensional vector space W we can form the Lie algebra $\mathfrak{gl}(W)$ (also denoted $\mathfrak{gl}_n$, where $n = \dim W$), which has underlying vector space $M(W) := \mathrm{End}(W)$ and Lie bracket defined $[X, Y] = XY - YX$ for $X, Y \in M(W)$. The traceless endomorphisms of W, denoted $\mathfrak{sl}(W)$ or $\mathfrak{sl}_n$, form a Lie subalgebra of $\mathfrak{gl}(W)$.

Example 2.3. The Lie algebra $\mathfrak{sl}_2$ of traceless 2×2-matrices is spanned by the elements

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \text{ and } Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

which have the relations $[X, Y] = H$, $[H, X] = 2X$, and $[H, Y] = -2Y$.

Definition 2.4. *A representation of a Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism $\rho : \mathfrak{g} \to \mathfrak{gl}(W)$ for some finite-dimensional W.*

We will sometimes drop the ρ and write $X \cdot w$ or $X(w)$ instead of $\rho(X)w$ for $X \in \mathfrak{g}$ and $w \in W$. For any $\mathfrak{g}$ we can construct a representation, called the adjoint representation, $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ by taking $X \mapsto [X, \cdot] \in \mathrm{End}(\mathfrak{g})$.

Definition 2.5. *For any Lie algebra $\mathfrak{g}$, the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ of $\mathfrak{g}$ is defined to be $T\mathfrak{g}/\langle v \otimes w - w \otimes v - [v, w] : v, w \in \mathfrak{g}\rangle$, where $T\mathfrak{g}$ is the tensor algebra of $\mathfrak{g}$.*

A representation of $\mathfrak{g}$ is the same thing as a $\mathcal{U}(\mathfrak{g})$-module. Note that a representation of $\mathfrak{g}$ is not necessarily the same thing as a representation of the image of the adjoint representation, as the latter might not be faithful. The next theorem provides a handy basis for $\mathcal{U}(\mathfrak{g})$:

Theorem 2.6 (Poincaré-Birkhoff-Witt)**.** *Let $\{X_1, \dots X_k\}$ be a basis for $\mathfrak{g}$ and let $\iota : \mathfrak{g} \hookrightarrow \mathcal{U}(\mathfrak{g})$ be the linear inclusion of $\mathfrak{g}$ into $\mathcal{U}(\mathfrak{g})$. If $M = \{\iota(X_1)^{\otimes m_1} \otimes \cdots \otimes \iota(X_k)^{\otimes m_k} : m_i \in \mathbb{N}\}$, then $M \cup \{1\}$ forms a $\mathbb{C}$-basis for $\mathcal{U}(\mathfrak{g})$.*

Proof. See theorem 4.3 of [3]. □

3 Solvable and Semisimple Lie Algebras

Definitions 3.1. *The lower central series $\{\mathscr{D}_n\mathfrak{g}\}$ of a Lie algebra $\mathfrak{g}$ is defined inductively $\mathscr{D}_1\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, $\mathscr{D}_{n+1}\mathfrak{g} = [\mathfrak{g}, \mathscr{D}_n\mathfrak{g}]$. A Lie algebra is called nilpotent if its lower central series stabilizes at 0.*

Similarly,

Definitions 3.2. *The derived series $\{\mathscr{D}^n\mathfrak{g}\}$ is defined $\mathscr{D}^1\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$, $\mathscr{D}^{n+1}\mathfrak{g} = [\mathscr{D}^n\mathfrak{g}, \mathscr{D}^n\mathfrak{g}]$. A Lie algebra is called solvable if its derived series stabilizes at 0; a Lie algebra is called semisimple if it contains no nonzero solvable ideals.*

Note that containing no nonzero solvable ideals is equivalent to containing no nonzero abelian ideals: on the one hand an abelian ideal is solvable, and on the other hand if $\mathfrak{g}' \subset \mathfrak{g}$ is a nonzero solvable ideal, then the last nonzero term of the derived series of $\mathfrak{g}'$ is a nonzero abelian ideal of $\mathfrak{g}$ (the fact that this is an ideal of $\mathfrak{g}$ follows by the Jacobi identity and induction on the length of the derived series).

Semisimple Lie algebras are of interest for two reasons: firstly, every Lie algebra $\mathfrak{g}$ can be put in a short exact sequence

$$0 \to \mathrm{Rad}(\mathfrak{g}) \to \mathfrak{g} \to \mathfrak{g}/\mathrm{Rad}(\mathfrak{g}) \to 0,$$

where $\mathrm{Rad}(\mathfrak{g})$ is the sum of all solvable ideals in $\mathfrak{g}$ (and thus itself a solvable ideal) and $\mathfrak{g}/\mathrm{Rad}(\mathfrak{g})$ is semisimple. Secondly, semisimple Lie algebras contain special subalgebras which make their representations much easier to understand:

Theorem 3.3. *Every semisimple Lie algebra $\mathfrak{g}$ contains a subalgebra $\mathfrak{h}$, called a Cartan subalgebra, which is abelian, acts diagonally on $\mathfrak{g}$, and is maximal among subalgebras with this property.*

Cartan subalgebras are useful for studying representations of semisimple Lie algebras because they are well-behaved, but still large enough in $\mathfrak{g}$ that $\rho|_{\mathfrak{h}}$ determines $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$. However, to prove theorem 3.3 we need a couple of fundamental facts.

The following two theorems will be useful mainly because they apply to the image of the adjoint representation:

Theorem 3.4 (Engel)**.** *If $\mathfrak{g} \subset \mathfrak{gl}(V)$ is a Lie subalgebra such that every $X \in \mathfrak{g}$ is a nilpotent endomorphism of V, then there is a nonzero $v \in V$ such that $X(v) = 0$ for all $X \in \mathfrak{g}$.*

Proof. We will prove this by induction on the dimension of $\mathfrak{g}$, showing that $\mathfrak{g}$ must contain an ideal of codimension one—which by the induction hypothesis annihilates some nonzero subspace W of V—and then showing that an element spanning the rest of $\mathfrak{g}$ must also kill some element of W.

First we note that if $X \in \mathfrak{gl}(V)$ is any nilpotent endomorphism of V, then $\mathrm{ad}(X)$ is a nilpotent endomorphism (of $\mathfrak{gl}(V)$): if X is nilpotent, then $X^K = 0$ for some K; for any other $Y \in \mathfrak{gl}(V)$, $\mathrm{ad}(X)^{2K}(Y) \in \mathfrak{gl}(V)$ is a homogeneous polynomial in X and Y of degree $2K+1$ in which every term has at most one power of Y, so every term must contain a factor of $X^K = 0$, i.e. $0 = \mathrm{ad}(X)^{2K} \in \mathrm{End}(\mathfrak{gl}(V))$.

Next we produce the promised ideal. Let $\mathfrak{h} \subset \mathfrak{g}$ be any maximal proper subalgebra of $\mathfrak{g}$. Since $\mathfrak{h}$ is a subalgebra, the adjoint action of $\mathfrak{h}$ on $\mathfrak{g}$ preserves $\mathfrak{h}$ and so acts on $\mathfrak{g}/\mathfrak{h}$. On the other hand, by the last paragraph, the adjoint action of each $X \in \mathfrak{h}$ on $\mathfrak{g}$ is nilpotent, so its action on $\mathfrak{g}/\mathfrak{h}$ is nilpotent. By induction, this means that some nonzero $\overline{Y} \in \mathfrak{g}/\mathfrak{h}$ is killed by the adjoint action of $\mathfrak{h}$, so there is some nonzero representative $Y \in \mathfrak{g} \setminus \mathfrak{h}$ such that $[\mathfrak{h}, Y] \subset \mathfrak{h}$. Thus $\mathfrak{h}$ is an ideal of the subalgebra $\mathfrak{g}'$ of $\mathfrak{g}$ generated by Y and $\mathfrak{h}$ (which by the above is spanned by Y and $\mathfrak{h}$). By maximality of $\mathfrak{h}$, $\mathfrak{g}'$ must be all of $\mathfrak{g}$, so $\mathfrak{h}$ is an ideal of $\mathfrak{g}$ of codimension one.

By induction on the dimension of $\mathfrak{g}$, there is a nonzero $v \in V$ such that $X(v) = 0$ for every $X \in \mathfrak{h}$, so the subspace $W := \{w \in V : X(w) = 0 \text{ for every X } \in \mathfrak{h}\}$ is nonzero. To prove the theorem it remains to show that $Y(u) = 0$ for some nonzero $u \in W$. This is where we use that $\mathfrak{h}$ is an ideal: for any $w \in W$ and $X \in \mathfrak{h}$, $X(Y(w)) = Y(X(w)) + [X, Y](w)$. X and $[X, Y]$ are in $\mathfrak{h}$, so both terms on the right-hand side are zero. Thus $X(Y(w)) = 0$ for every $X \in \mathfrak{h}$ and every $w \in W$, so Y carries W into itself. Since Y acts nilpotently on all of V, this means there must be some nonzero $u \in W$ such that $Y(u) = 0$. □

Corollary 3.5. *If $\mathfrak{g}$ is a Lie algebra such that for every $X \in \mathfrak{g}$, ad(X) is a nilpotent endomorphism of $\mathfrak{g}$, then $\mathfrak{g}$ is nilpotent.*

Theorem 3.6 (Lie)**.** *If $\mathfrak{g} \subset \mathfrak{gl}(V)$ is a solvable Lie algebra, then there is a nonzero $v \in V$ such that v is an eigenvector of $\mathfrak{g}$.*

For this theorem we first prove the following lemma:

Lemma 3.7. *Let $\mathfrak{h}$ be an ideal of a Lie algebra $\mathfrak{g}$, let V be a representation of $\mathfrak{g}$ and suppose $\lambda \in \mathfrak{h}^*$. If $W = \{v \in V : X(v) = \lambda(X)v$ for every $X \in \mathfrak{h}\}$, then for every $Y \in \mathfrak{g}$, $Y(W) \subset W$.*

Proof of lemma. Suppose $w \in W$, $w \neq 0$ and $X \in \mathfrak{h}$. Then $X(Y(w)) = Y(X(w)) + [X, Y](w) = \lambda(X)Y(w) + \lambda([X, Y])w$, since $[X, Y] \in \mathfrak{h}$. Thus to show $Y(w) \in W$, it remains to show that $\lambda([X, Y]) = 0$ for all $X \in \mathfrak{h}$:

Let $U \subset V$ be the span of $w, Y(w), Y^2(w), \ldots$. U is certainly preserved by Y, but for any $X \in \mathfrak{h}$, $X(Y^k(w)) = Y(X(Y^{k-1}(w))) + [X, Y](Y^{k-1}(w))$, so by induction on k we see that U is also preserved by X. In fact this shows that the action of $X \in \mathfrak{h}$ on U is upper-triangular with respect to the basis $w, Y(w), Y^2(w), \ldots$, with diagonal entries $\lambda(X)$, so $\mathrm{Tr}(X) = \lambda(X) \cdot \dim U$. On the other hand, $\mathrm{Tr}([X, Y]) = 0$ for any endomorphisms X, Y of V. Thus $\lambda([X, Y]) = 0$ for any $X \in \mathfrak{h}$. □

Proof of theorem 3.6. As in the proof of Engel's theorem we want an ideal $\mathfrak{h} \subset \mathfrak{g}$ of codimension one. This time it's a bit easier: since $\mathfrak{g}$ is solvable, $\mathscr{D}\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] \neq \mathfrak{g}$. The commutator $\mathscr{D}\mathfrak{g}$ is an ideal, so this means the Lie algebra $\mathfrak{a} = \mathfrak{g}/\mathscr{D}\mathfrak{g}$ is nonzero. Any subspace of $\mathfrak{a}$ is an ideal, so by the first isomorphism theorem the preimage of any codimension one subspace of $\mathfrak{a}$ will be a codimension one ideal of $\mathfrak{g}$.

By induction on the dimension of $\mathfrak{g}$, we may assume there is some nonzero $v_0 \in V$ that is an eigenvector for every element of $\mathfrak{h}$. For $X \in \mathfrak{h}$ let $\lambda(X)$ be the eigenvalue of X corresponding to v_0 and let $W = \{v \in V : X(v) = \lambda(X)v \text{ for every } X \in \mathfrak{h}\}$. Again as in the proof of Engel, let $Y \in \mathfrak{g} \setminus \mathfrak{h}$. By the lemma, $Y(W) \subset W$, so Y must have some nonzero eigenvector in W, completing the proof. □

Corollary 3.8. *If $\mathfrak{g} \subset \mathfrak{gl}(V)$ is a solvable Lie algebra, then the elements of $\mathfrak{g}$ can simultaneously be put in upper triangular form.*

Proof. Induct on the dimension of V. □

The next theorem, despite its obvious importance (for example, for motivating the study of irreducible representations of semisimple Lie algebras in the first place), will only be used once or twice in the rest of the paper. Because of this and the length of its proof, this is one of the theorems we will palm off.

Theorem 3.9 (Weyl)**.** *For any semisimple Lie algebra $\mathfrak{g}$ and any finite-dimensional representation V, V is completely reducible (i.e. can be written as a direct sum of irreducible subrepresentations).*

For something like Weyl's original proof, using integration to prove the Peter-Weyl theorem for compact Lie groups and thence the Weyl theorem for semisimple Lie algebras, see [4]. For a more modern algebraic proof, see section 7.8 of [5].

4 Preservation of the Jordan Form

In this section we show what is probably the most useful fact about linear representations of semisimple Lie algebras: the various definitions of the "nilpotent" and "semisimple" parts of an element are equivalent, or rather consistent across all representations

Theorem 4.1. *Suppose $\mathfrak{g}$ is a semisimple Lie subalgebra of $\mathfrak{gl}(V)$. If $X \in \mathfrak{g}$, then $X_s, X_n \in \mathfrak{g}$, where X_s is the semisimple part of X and X_n is the nilpotent part of X as an endomorphism of V.*

Proof. $[X, \mathfrak{g}] \subset \mathfrak{g}$, and likewise $[p(X), \mathfrak{g}] \subset \mathfrak{g}$ for any $p \in \mathbb{C}[T]$. Both X_s and X_n can be expressed as polynomials in X, so $[X_s, \mathfrak{g}] \subset \mathfrak{g}$ and $[X_n, \mathfrak{g}] \subset \mathfrak{g}$. Letting $\mathfrak{n}$ be the subalgebra $\{A \in \mathfrak{gl}(V) : [A, \mathfrak{g}] \subset \mathfrak{g}\}$, this means $X_s, X_n \in \mathfrak{n}$. For each $\mathfrak{g}$-subrepresentation W of V, we define another subalgebra $\mathfrak{s}_W = \{Y \in \mathfrak{gl}(V) : Y(W) \subset W \text{ and } \mathrm{Tr}(Y|_W) = 0\}$.

Now we claim that

$$\mathfrak{g}' := \mathfrak{n} \cap \left(\bigcap_{W \subset V \text{ a subrep}} \mathfrak{s}_W \right) = \mathfrak{g}.$$

Since $\mathfrak{g}$ is semisimple, $\mathscr{D}\mathfrak{g} = \mathfrak{g}$. Thus every $X \in \mathfrak{g}$ is traceless, so $\mathfrak{g} \subset \mathfrak{g}'$. Since $\mathfrak{g} \subset \mathfrak{g}' \subset \mathfrak{n}$, $\mathfrak{g}$ is an ideal of $\mathfrak{g}'$. By theorem 3.9, $\mathfrak{g}$ has a complementary $\mathfrak{g}$-representation U in $\mathfrak{g}'$, i.e. $\mathfrak{g}' = \mathfrak{g} \oplus U$, so it remains to show that $U = 0$; for this it suffices to show that for any $Y \in U$, $Y|_W = 0$ for any irreducible $\mathfrak{g}$-subrepresentation $W \subset V$, since any $Y \in \mathfrak{g}'$ must preserve each such W and by theorem 3.9 we know that V is a sum of irreducible representations. Now, $[\mathfrak{g}, \mathfrak{g}'] \subset \mathfrak{g}$, so $[\mathfrak{g}, U] = 0$. Thus Y commutes with $\mathfrak{g}$, so $Y|_W$ is a $\mathfrak{g}$-intertwining endomorphism of W. By Schur's lemma for Lie algebras (which is just the same as Schur's lemma for finite groups), this means $Y|_W = \lambda \cdot I$ for some $\lambda \in \mathbb{C}$. But $Y|_W$ is traceless, so $Y|_W = 0$. □

For any semisimple $\mathfrak{g}$ we have the adjoint representation $\mathrm{ad} : \mathfrak{g} \hookrightarrow \mathfrak{gl}(\mathfrak{g})$, so by the above lemma for any $X \in \mathfrak{g}$ we have $X_s, X_n \in \mathfrak{g}$, whose adjoint actions on $\mathfrak{g}$ are the semisimple and nilpotent parts of $\mathrm{ad}(X)$ respectively.

Observation 4.2. *For $\mathfrak{g}, \mathfrak{g}'$ semisimple and $\rho : \mathfrak{g} \to \mathfrak{g}'$ a Lie algebra homomorphism, $\rho(X_s) = \rho(X)_s$ and $\rho(X_n) = \rho(X)_n$.*

Proof. By the lemma, $\rho(X)_s \in \mathrm{im}(\rho)$, so it suffices to show that the restriction of $\mathrm{ad}(\rho(X_s))$ to $\mathrm{im}(\rho)$ is diagonal. Thus it suffices to consider the case $\mathrm{im}(\rho) = \mathfrak{g}'$, i.e. $\mathfrak{g}' = \mathfrak{g}/\ker(\rho)$; $\ker(\rho)$ is an ideal of $\mathfrak{g}$, so $\mathrm{ad}(X_s)$ must act (diagonally) on $\ker(\rho)$ and thus (diagonally) on $\mathfrak{g}/\ker(\rho)$. An identical argument works for X_n. □

Corollary 4.3. *For any representation $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ of a semisimple Lie algebra $\mathfrak{g}$ and any $X \in \mathfrak{g}$, $\rho(X_s)$ is the semisimple part of $\rho(X)$ and $\rho(X_n)$ is the nilpotent part of $\rho(X)$.*

Proof. By the above, $\rho(X_s)$ is the semisimple part of $\rho(X)$ regarded as an element of the semisimple Lie algebra $\mathfrak{g}' := \rho(\mathfrak{g}) \subset \mathfrak{gl}(V)$, i.e. $\rho(X_s)|_{\mathfrak{g}'}$ is semisimple. On the other hand, by theorem 4.1, the semisimple part $\rho(X)_s$ of $\rho(X)$, as an endomorphism of V, is in $\mathfrak{g}'$, so to show that the semisimple part $\rho(X)_s$ of $\rho(X)$ is $\rho(X_s)$ it suffices to show that the adjoint action of the semisimple part of $\rho(X)$ on $\mathfrak{g}'$ is diagonal.

If any $Y \in \mathfrak{gl}(V)$ is diagonal with respect to some basis B of V then $\mathrm{ad}(Y)$ is diagonal with respect to the standard basis[1] of $\mathfrak{gl}(V)$ induced by B, so $\mathrm{ad}(Y_s) = \mathrm{ad}(Y)_s$. Thus the adjoint action of the semisimple part of $\rho(X)$ on $\mathfrak{g}$ is diagonal. However, by the above, the semisimple part of $\rho(X)$ acts on $\mathfrak{g}'$, so it must act diagonally on $\mathfrak{g}'$. An identical argument works for $\rho(X_n)$. □

Thus for $\mathfrak{g}$ semisimple, representations respect the Jordan decomposition of elements of $\mathfrak{g}$.

5 The Killing Form

Now we introduce a bilinear form B on Lie algebras, called the Killing form:

Definition 5.1. *$B(X, Y) := Tr(ad(X) \circ ad(Y))$ for $X, Y \in \mathfrak{g}$ any Lie algebra.*

Proposition 5.2. *For $X, Y, Z \in \mathfrak{g}$ a Lie algebra and B defined as above, $B([X, Y], Z) = B(X, [Y, Z])$*

Proof. For any endomorphisms of a vector space A, B, C,

$$\mathrm{Tr}((AB - BA)C) = \mathrm{Tr}((AB - BA)C) + \mathrm{Tr}(B(AC) - (AC)B) = \mathrm{Tr}(A(BC - CB)).$$

□

[1] The matrices with one entry of 1 and the rest 0, usually denoted $\{E_{ij}\}$

The Killing form is handy for detecting semisimplicity and for separating the root spaces of semisimple Lie algebras, which we will encounter later in this section. To show these properties we first define B_V on $\mathfrak{gl}(V)$ by $B_V(X, Y) = \mathrm{Tr}(X \circ Y)$; as usual we do this to prove something about the adjoint action of a general Lie algebra.

Theorem 5.3 (Cartan's criterion)**.** *If $\mathfrak{g}$ is a subalgebra of $\mathfrak{gl}(V)$ and $B_V(X, Y) = 0$ for every $X, Y \in \mathfrak{g}$, then $\mathfrak{g}$ is solvable.*

Proof. By Engel's theorem it suffices to show that every element of $\mathscr{D}\mathfrak{g}$ is nilpotent. Let $X \in \mathscr{D}\mathfrak{g}$, take a basis B for V that puts X in Jordan form, and let $\lambda_1, \ldots, \lambda_r$ be the diagonal entries of X. We need to show that the λ_i are all zero, for which it suffices to show that $\overline{\lambda}_1\lambda_1 + \cdots + \overline{\lambda}_r\lambda_r = 0$.

Let $D = X_s$ be the semisimple part of X, which by the above is the diagonal transformation with entries $\lambda_1, \ldots, \lambda_r$ and an element of $\mathfrak{g}$ by theorem 4.1. Let $\overline{D}$ be the diagonal transformation with entries $\overline{\lambda}_1, \ldots, \overline{\lambda}_r$. Since $B_V(\overline{D}, X) = \overline{\lambda}_1\lambda_1 + \cdots + \overline{\lambda}_r\lambda_r$, we want to show that $B_V(\overline{D}, X) = 0$. Since X is a sum of elements of the form $[Y, Z]$, $Y, Z \in \mathfrak{g}$, $B_V(\overline{D}, X)$ is a sum of terms of the form $B_V(\overline{D}, [Y, Z]) = B_V([\overline{D}, Y], Z)$. Our hypothesis is that $B_V(Y, Z) = 0$ for $Y, Z \in \mathfrak{g}$, so it suffices to show $[\overline{D}, Y] \in \mathfrak{g}$.

For this it suffices to show that $\mathrm{ad}(\overline{D})$ is a polynomial in $\mathrm{ad}(D)$, since $\mathrm{ad}(D)^k Y \in \mathfrak{g}$. But with respect to the standard basis of $\mathfrak{gl}(V)$ induced by B, $\mathrm{ad}(D)$ and $\mathrm{ad}(\overline{D})$ are diagonal complex conjugate matrices, and thus polynomials in each other. □

Corollary 5.4. *For any Lie algebra $\mathfrak{g}$ and B defined on $\mathfrak{g}$ as above, B is nondegenerate if and only if $\mathfrak{g}$ is semisimple.*

Proof. Let $\mathfrak{s} = \{X \in \mathfrak{g} : B(X, Y) = 0 \text{ for all } Y \in \mathfrak{g}\}$. If $X \in \mathfrak{s}$ and $Y \in \mathfrak{g}$ then for any $Z \in \mathfrak{g}$, $B([X, Y], Z) = B(X, [Y, Z]) = 0$, so $\mathfrak{s}$ is an ideal. Further, by Cartan's criterion, the image of the adjoint action of $\mathfrak{s}$, $\mathrm{ad}(\mathfrak{s}) \subset \mathfrak{gl}(\mathfrak{g})$ is solvable, so $\mathfrak{s}$ is solvable. Now suppose $\mathfrak{g}$ is semisimple. By the above, $\mathfrak{s} \subset \mathfrak{g}$ is a solvable ideal, so $\mathfrak{s} = 0$.

Going the other way, suppose $\mathfrak{s} = 0$. Any Lie algebra containing a nonzero solvable ideal must contain a nonzero abelian ideal, namely the last nonzero element of the solvable ideal's derived series, so it suffices to show that any abelian ideal $\mathfrak{a} \subset \mathfrak{g}$ must be zero. If $X \in \mathfrak{a}$ and $Y \in \mathfrak{g}$, then $\mathrm{ad}(X) \circ \mathrm{ad}(Y)$ maps $\mathfrak{g}$ into $\mathfrak{a}$ (since the latter is an ideal) and $\mathfrak{a}$ to zero, so $B(X, Y) = \mathrm{Tr}(\mathrm{ad}(X) \circ \mathrm{ad}(Y)) = 0$, so $\mathfrak{a} \subset \mathfrak{s} = 0$. □

6 Cartan Subalgebras

Proof of Theorem 3.3. For $X \in \mathfrak{g}$, let $P_X(T) = \det(T - \mathrm{ad}(X))$, i.e. the characteristic polynomial of $\mathrm{ad}(X)$. If $n = \dim \mathfrak{g}$, we can write the coefficients of $P_X(T)$ as functions of X: $P_X(T) = \sum_{i=0}^{i=n} a_i(X)T^i$. If $X = X_1e_1 + \cdots + X_ne_n$ for some basis $\{e_i\}$ of $\mathfrak{g}$, $a_i(X)$ is a homogeneous polynomial of degree $n - i$ in $X_1, \ldots, X_n$. We call the least index l such that a_l is not identically zero on $\mathfrak{g}$ the rank of $\mathfrak{g}$.

Definition 6.1. *We call an element $H \in \mathfrak{g}$ regular if $a_l(H) \neq 0$; equivalently, an element $H \in \mathfrak{g}$ is regular if the dimension of the nilspace of ad(H) (i.e. the subspace of $\mathfrak{g}$ killed by some power of ad(H)) is minimal among the elements of $\mathfrak{g}$.*

$\mathfrak{g}$ must contain regular elements,[2] and our eventual goal will be to show that for $H \in \mathfrak{g}$ regular, $\mathfrak{g}_0(H) \subset \mathfrak{g}$ is a Cartan subalgebra. First, however, some observations about how the Lie bracket and the Killing form behave on the Jordan blocks of an element $H \in \mathfrak{g}$:

For any $H \in \mathfrak{g}$ we can decompose $\mathfrak{g}$ according to the Jordan blocks of $\mathrm{ad}(H)$:

$$\mathfrak{g} = \bigoplus_{\lambda} \mathfrak{g}_\lambda(H), \tag{1}$$

where $\mathfrak{g}_\lambda(H) = \{X \in \mathfrak{g} : \exists k \text{ st for } k' \geq k, (\mathrm{ad}(H) - \lambda I)^{k'}(X) = 0\}$.

Observation 6.2. *For any $H \in \mathfrak{g}$, $[\mathfrak{g}_\lambda(H), \mathfrak{g}_\mu(H)] \subset \mathfrak{g}_{\lambda+\mu}(H)$*

Proof. For $X \in \mathfrak{g}_\lambda(H)$ and $Y \in \mathfrak{g}_\mu(H)$, we claim that

$$(\mathrm{ad}(H) - (\lambda+\mu)I)^k([X,Y]) = \sum_{j=0}^{k} \binom{k}{j} [(\mathrm{ad}(H) - \lambda I)^j(X), (\mathrm{ad}(H) - \mu I)^{k-j}(Y)]. \tag{2}$$

All the terms on the right vanish for k sufficiently large, so this implies $[X,Y] \in \mathfrak{g}_{\lambda+\mu}(H)$. We prove the equation by induction on k. By the induction hypothesis,

$$\begin{aligned}
&(\mathrm{ad}(H) - (\lambda+\mu)I)^k([X,Y]) \\
&\qquad = (\mathrm{ad}(H) - (\lambda+\mu)I) \sum_{j=0}^{k-1} \binom{k-1}{j} [(\mathrm{ad}(H) - \lambda I)^j(X), (\mathrm{ad}(H) - \mu I)^{k-1-j}(Y)],
\end{aligned}$$

but by the Jacobi identity,

$$\begin{aligned}
&(\mathrm{ad}(H) - (\lambda+\mu)I)[(\mathrm{ad}(H) - \lambda I)^j(X), (\mathrm{ad}(H) - \mu I)^{k-1-j}(Y)] \\
&\qquad = -(\lambda+\mu)[(\mathrm{ad}(H) - \lambda I)^j(X), (\mathrm{ad}(H) - \mu I)^{k-1-j}(Y)] + \\
&\qquad\quad [\mathrm{ad}(H)(\mathrm{ad}(H) - \lambda I)^j(X), (\mathrm{ad}(H) - \mu I)^{k-1-j}(Y)] + \\
&\qquad\quad [(\mathrm{ad}(H) - \lambda I)^j(X), \mathrm{ad}(H)(\mathrm{ad}(H) - \mu I)^{k-1-j}(Y)] \\
&\qquad = [(\mathrm{ad}(H) - \lambda I)^{j+1}(X), (\mathrm{ad}(H) - \mu I)^{k-1-j}(Y)] + \\
&\qquad\qquad [(\mathrm{ad}(H) - \lambda I)^j(X), (\mathrm{ad}(H) - \mu I)^{k-j}(Y)],
\end{aligned}$$

[2] In fact, the regular elements of $\mathfrak{g}$ form a Zariski open subset.

so the above becomes

$$\sum_{j=0}^{k-1}\binom{k-1}{j}[(\mathrm{ad}(H)-\lambda I)^{j+1}(X),(\mathrm{ad}(H)-\mu I)^{k-1-j}(Y)]+$$
$$\sum_{j=0}^{k-1}\binom{k-1}{j}[(\mathrm{ad}(H)-\lambda I)^{j}(X),(\mathrm{ad}(H)-\mu I)^{k-j}(Y)]$$
$$=\sum_{j=0}^{k}\binom{k}{j}[(\mathrm{ad}(H)-\lambda I)^{j}(X),(\mathrm{ad}(H)-\mu I)^{k-j}(Y)].$$

□

Thus if $Y \in \mathfrak{g}_\lambda(H)$ for $\lambda \neq 0$ then by the above observation, $\mathrm{ad}(Y)$ takes each $\mathfrak{g}_\mu(H)$ (μ possibly 0) to a different $\mathfrak{g}_{\mu'}(H)$; $X \in \mathfrak{g}_0(H)$ preserves the $\mathfrak{g}_\mu(H)$ (note that this means $\mathfrak{g}_0(H)$ is a subalgebra of $\mathfrak{g}$), so $\mathrm{ad}(X)\circ\mathrm{ad}(Y)$ also takes each $\mathfrak{g}_\mu(H)$ to a different $\mathfrak{g}_{\mu'}(H)$, so $B(X,Y) = \mathrm{Tr}(\mathrm{ad}(X)\circ\mathrm{ad}(Y)) = 0$, i.e. $\mathfrak{g}_0(H) \perp \bigoplus_{\lambda\neq 0} \mathfrak{g}_\lambda(H)$. By corollary 5.4, however, B is nondegenerate, so the restriction of B to $\mathfrak{g}_0(H)$ must be nondegenerate.

Now suppose H is a regular element of $\mathfrak{g}$.

Claim 6.3. *$\mathfrak{g}_0(H)$ is nilpotent.*

Proof. By corollary 3.5 to Engel's theorem it suffices to show that for each $Y \in \mathfrak{g}_0(H)$, the restriction of $\mathrm{ad}(Y)$ to $\mathfrak{g}_0(H)$ is a nilpotent endomorphism. Let $\mathrm{ad}'(Y) = \mathrm{ad}(Y)|_{\mathfrak{g}_0(H)}$ and let $\mathrm{ad}''(Y)$ be the induced endomorphism of $\mathrm{ad}(Y)$ on the vector space quotient $\mathfrak{g}/\mathfrak{g}_0(H)$. We define subsets $U = \{Y \in \mathfrak{g}_0(H) : \mathrm{ad}(Y) \text{ is not nilpotent}\}$ and $V = \{Y \in \mathfrak{g}_0(H) : \mathrm{ad}''(Y) \text{ is invertible}\}$.

Suppose that U is nonempty. U is an open subset of $\mathfrak{g}_0(H)$ and V is a nonempty open subset (since $X \in V$). Further, V is the complement of an algebraic variety, so it's dense in $\mathfrak{g}_0$ and so U and V intersect, say at some element Y. The multiplicity of 0 as an eigenvalue of $\mathrm{ad}(Y)$—i.e., the dimension of $\mathfrak{g}_0(Y)$—is the multiplicity a of 0 as an eigenvalue of $\mathrm{ad}'(Y)$ plus the multiplicity b of 0 as an eigenvalue of $\mathrm{ad}''(Y)$. Since $Y \in U$, $a < \dim \mathfrak{g}_0(H)$. On the other hand, since $Y \in V$, $b = 0$. Thus $\dim \mathfrak{g}_0(Y) < \dim \mathfrak{g}_0(H)$, contradicting the regularity of H. Thus $U = \emptyset$. □

The claim shows that $\mathfrak{g}_0(H)$ is solvable, so by Lie's theorem there is a basis for $\mathfrak{g}$ in which $\mathrm{ad}(X)$ is upper-triangular for every $X \in \mathfrak{g}_0(H)$. Thus for $X, Y, Z \in \mathfrak{g}_0(H)$, $\mathrm{ad}([X,Y])$ is strictly upper-triangular and $\mathrm{ad}(Z)$ is upper-triangular with respect to the same basis, so $\mathrm{Tr}(\mathrm{ad}([X,Y])\circ\mathrm{ad}(Z)) = B([X,Y],Z) = 0$. However, B is nondegenerate on $\mathfrak{g}_0(H)$, so it must be that $[X,Y] = 0$ for any $X, Y \in \mathfrak{g}_0(H)$, i.e. $\mathfrak{g}_0(H)$ is abelian.

Next we show that any element of $\mathfrak{g}_0(H)$ is semisimple:

Claim 6.4. *If $\mathfrak{n}(\mathfrak{g}_0(H))$ is the "normalizer" of $\mathfrak{g}_0(H)$, defined as $\{X \in \mathfrak{g} : [X, \mathfrak{g}_0(H)] \subset \mathfrak{g}_0(H)\}$, then $\mathfrak{n}(\mathfrak{g}_0(H)) = \mathfrak{g}_0(H)$.*

Proof. Let $Z \in \mathfrak{n}(\mathfrak{g}_0(H))$, so in particular $[Z, H] \in \mathfrak{g}_0(H)$, meaning that $\mathrm{ad}(H)^k[Z, H] = 0$ for some $k \in \mathbb{N}$. Thus $\mathrm{ad}(H)^{k+1}Z = 0$, so $Z \in \mathfrak{g}_0(H)$. $\square$

If $\mathfrak{c}(\mathfrak{g}_0(H)) := \{X \in \mathfrak{g} : [X, \mathfrak{g}_0(H)] = 0\}$ (called the commutator of $\mathfrak{g}_0(H)$) then, since $\mathfrak{g}_0(H)$ is abelian, $\mathfrak{g}_0(H) \subset \mathfrak{c}(\mathfrak{g}_0(H))$. On the other hand, $\mathfrak{c}(\mathfrak{g}_0(H)) \subset \mathfrak{n}(\mathfrak{g}_0(H))$, which by the above is contained in $\mathfrak{g}_0(H)$, so $\mathfrak{c}(\mathfrak{g}_0(H)) = \mathfrak{g}_0(H)$. Now, if $X, Y \in \mathfrak{g}_0(H)$ and X_n is the nilpotent part of X, then $[X, Y] = 0$ so $[X_n, Y] = 0$. Thus $X_n \in \mathfrak{c}(\mathfrak{g}_0(H)) = \mathfrak{g}_0(H)$. With respect to the same basis of $\mathfrak{g}_0(H)$ as we used to show that $\mathfrak{g}_0(H)$ is abelian, $\mathrm{ad}(X_n)$ is strictly upper-triangular, so $B(X_n, Y) = 0$ for any $Y \in \mathfrak{g}_0(H)$. However, the restriction of B to $\mathfrak{g}_0(H)$ is nondegenerate, so $X_n = 0$, so X is semisimple. Since the elements of $\mathfrak{g}_0(H)$ commute, this shows that the adjoint action of $\mathfrak{g}_0(H)$ on $\mathfrak{g}$ is diagonal.

Since $\mathfrak{c}(\mathfrak{g}_0(H)) = \mathfrak{g}_0(H)$, maximality is immediate. $\square$

7 Representations of $\mathfrak{sl}_2$

In this section we take a break from theory to look at a semisimple Lie algebra, $\mathfrak{sl}_2$, and use a Cartan subalgebra to characterize all of its irreducible representations. This will give a concrete example of how Cartan subalgebras are used to study representations by decomposing them into root spaces, as well as motivating the definition of roots in section 8. Further, the subalgebras $\mathfrak{s}_\alpha \cong \mathfrak{sl}_2$ of a general semisimple Lie algebra $\mathfrak{g}$ (which we will construct in section 8) and particularly their representations will be useful to the study of the roots and irreducible representations of $\mathfrak{g}$.

Recall the basis X, Y, H from example 2.3.

Claim 7.1. *$\mathfrak{sl}_2$ is semisimple.*

Proof. First suppose that $\mathfrak{sl}_2$ has a one-dimensional ideal, spanned by some element $Z = aX + bY + cH$. Then $[Z, X] = -(bH + 2cX)$ is in the span of Z, so $b = 0$ and likewise $[Z, Y] = aH - 2cY$ is in the span of Z, so $a = 0$ (the apparent alternatives $[Z, X] = 0$ and $[Z, Y] = 0$ are not in fact alternatives, as they imply $b = 0$ and $a = 0$ anyway). Thus $Z = cH$, but the span of H is not an ideal.

Thus the only way $\mathfrak{sl}_2$ could fail to be semisimple is if it had a two-dimensional abelian ideal $\mathfrak{a}$, spanned by some nonzero U, V. Since U and V commute, they can simultaneously be put in upper-triangular form with respect to some basis B' of $\mathbb{C}^2$. Thus, defining X', Y', H' according to this new basis, $\mathfrak{a}$ must be the span of X' and H'. However, X' and H' do not commute, so $\mathfrak{a}$ cannot be abelian. $\square$

In this case we can see that H spans a Cartan subalgebra $\mathfrak{h}$, and its adjoint action on $\mathfrak{sl}_2$ has eigenvalues $2, 0, -2$ on the eigenvectors X, H, Y respectively.

Now suppose V is an irreducible $\mathfrak{sl}_2$-representation. By corollary 4.3, H acts semisimply on V, so we can decompose $V = \bigoplus_i V_{\omega_i}$, where each ω_i is an element of $\mathfrak{h}^*$ such that $H(v) = \omega_i(H) \cdot v$ for $v \in V_{\omega_i}$. We call the ω_i the weights of V.

Let $T := \{\omega_i(H)\}$. Since V is finite-dimensional, T must be finite, so we can pick $n = \omega(H)$ to be some element of T with maximal real part and v to be some nonzero vector in V_ω.

Claim 7.2. $\{v, Yv, Y^2v, \dots\}$ *span* V.

Proof. Firstly, we must have $Xv = 0$, since $H(Xv) = [H,X]v + X(Hv) = (2+n)Xv$, so if $Xv \neq 0$ then $2 + n \in T$, contradicting our choice of n.

Secondly, both Y and H fix the subspace $W \subset V$ spanned by $\{v, Yv, Y^2v, \dots\}$, so (since V is irreducible) it remains to show that $X(W) \subset W$ as well. We do this we will show a slightly stronger fact:

$$X(Y^m v) = m(n-m+1) \cdot Y^{m-1}(v). \tag{3}$$

We prove equation 3 by induction on m. The relation holds for $m = 1$, since $X(Yv) = [X,Y]v + Y(Xv) = Hv = n \cdot v$. On the other hand,

$$X(Y^m v) = [X,Y](Y^{m-1}v) + Y(X(Y^{m-1}v)),$$

which by the induction hypothesis is

$$\begin{aligned} &H(Y^{m-1}v) + (m-1)(n-m+2) \cdot Y^{m-1}v \\ &\quad = (n-2(m-1)) \cdot Y^{m-1}v + (m-1)(n-m+2) \cdot Y^{m-1}v = m(n-m+1) \cdot Y^{m-1}(v). \end{aligned}$$

□

$v, Yv, Y^2v, \dots$ have H-eigenvalues $n, n-2, n-4, \dots$ respectively, so the above claim shows that T must be of the form $n - K$, where K is an initial segment of $2\mathbb{N}$.

The next claim will even further restrict the possible weights of V. We will use it in section 9 to restrict the possible weights of irreducible representations of semisimple Lie algebras everywhere.

Claim 7.3. T *contains only integers and is symmetric about* 0.

Proof. Since V is finite-dimensional, some power of Y must annihilate v. Let m be the smallest such power, so by equation 3, $0 = X(0) = m(n-m+1) \cdot Y^{m-1}(v)$. Since $Y^{m-1}v \neq 0$, we must have $n - m + 1 = 0$. Thus n is an integer, so $T \subset \mathbb{Z}$, and the smallest element of T is $n - 2m = -n$. □

Thus V is totally determined by the largest element of T.

8 Roots and Weights

For any semisimple $\mathfrak{g}$ we have some Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$, and thus a linear decomposition $\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus \mathfrak{g}_\alpha$, where α ranges over some subset $R \subset \mathfrak{h}^* \setminus \{0\}$ and for $X \in \mathfrak{g}_\alpha, H \in \mathfrak{h}$, $\mathrm{ad}(H)(X) = \alpha(H)X$. For the sake of brevity we will assume that the $\mathfrak{h}$ we are using is of the form $\mathfrak{g}_0(H)$ for some regular $H \in \mathfrak{g}$—if you like, take this as a new, possibly stronger definition of Cartan subalgebra—so that we can use the properties of $\mathfrak{g}_0(H)$ shown in the proof of theorem 3.3.[3]

Proposition 8.1. $\mathfrak{h} = \mathfrak{g}_0$.

Proof. Since the adjoint action of $\mathfrak{h}$ on $\mathfrak{g}$ is diagonal, $\mathfrak{g}_0 = \mathfrak{c}(\mathfrak{h})$. However, we showed in the proof of theorem 3.3 that $\mathfrak{h} = \mathfrak{c}(\mathfrak{h})$. □

Definitions 8.2. *The elements of R are called roots and the $\mathfrak{g}_\alpha$ are called root spaces (of $\mathfrak{g}$ with respect to $\mathfrak{h}$).*

The roots form a "reduced root system" in $\mathfrak{h}^*$ in the following sense:

Proposition 8.3. *For $R, \mathfrak{h}$ as above,*

1. *R is finite, does not contain 0 and spans $\mathfrak{h}^*$,*
2. *For each $\alpha \in R$ there is a symmetry s_α with vector α which leaves R invariant,*
3. *If $\alpha, \beta \in R$, then $s_\alpha(\beta) - \beta$ is an integer multiple of α, and*
4. *For each $\alpha \in R$, the only elements of $R \cap \mathbb{C} \cdot \alpha$ are $\pm\alpha$.*

Where "a symmetry s_α with vector α" means an automorphism $s_\alpha = 1 - \alpha^* \otimes \alpha$ of $\mathfrak{h}^*$, for some $\alpha^* \in \mathfrak{h}^{**}$ such that $\alpha * (\alpha) = 2$

We will prove these facts, along with those of theorem 8.4 below. Most of the power of this proposition will go to waste in this paper—outside of the proof of the next theorem, we will use it directly only to define positive and negative roots.

These roots in turn give us some information about the structure of $\mathfrak{g}$:

Theorem 8.4. *For $R, \mathfrak{g}, \mathfrak{h}$ as above and $\alpha \in R$,*

1. *$\mathfrak{g}_\alpha$ is one-dimensional, as is the subspace $\mathfrak{h}_\alpha := [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$. There is a unique $H_\alpha \in \mathfrak{h}_\alpha$ such that $\alpha(H_\alpha) = 2$.*
2. *For each nonzero $X_\alpha \in \mathfrak{g}_\alpha$ there is a unique $Y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $[X_\alpha, Y_\alpha] = H_\alpha$. Further, $[H_\alpha, X_\alpha] = 2X_\alpha$ and $[H_\alpha, Y_\alpha] = -2Y_\alpha$, so the subalgebra $\mathfrak{s}_\alpha := \mathfrak{h}_\alpha \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$ is isomorphic to $\mathfrak{sl}_2$.*

To prove this we first need some facts about the $\mathfrak{g}_\alpha$ and their relationship to the Killing form, several of which we have seen versions of in previous proofs.

Proposition 8.5.
1. *If $\alpha \neq -\beta$ then $\mathfrak{g}_\alpha$ and $\mathfrak{g}_\beta$ are orthogonal (with respect to the Killing form). Further, $B(\mathfrak{g}_{-\alpha}, \cdot) = \mathfrak{g}_\alpha^*$, i.e. $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$ are dual. The restriction of $B(\cdot, \cdot)$ to $\mathfrak{h}$ is nondegenerate.*
2. *If $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$ and $H \in \mathfrak{h}$, then $B(H, [X, Y]) = \alpha(H) \cdot B(X, Y)$.*

[3]There are two slightly longer ways around this. One would be to take Serre's equivalent definition of Cartan subalgebra (although more useful, I find this definition a bit more opaque), and the other would be to show that every Cartan subalgebra is of the form $\mathfrak{g}_0(H)$ for some regular element H.

3. *If $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$ and $\alpha \in R$ and h_α is the element of $\mathfrak{h}$ such that $\alpha = B(h_\alpha, \cdot)$, whose existence is guaranteed by the second statement of 1 and the fact that $\mathfrak{g}_0 = \mathfrak{h}$, then $[X,Y] = B(X,Y) \cdot h_\alpha$.*

Proof. Suppose $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$, and $H \in \mathfrak{h}$. Then $0 = B([H,X],Y) + B(X,[H,Y]) = (\alpha(H) + \beta(H)) \cdot B(X,Y)$; if $\alpha + \beta \neq 0$, then we can pick H so the first factor on the right is nonzero (for all $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$), proving the first statement of 1. Thus $\mathfrak{g} = \mathfrak{h} \oplus \bigoplus(\mathfrak{g}_\alpha + \mathfrak{g}_{-\alpha})$ is an orthogonal decomposition of $\mathfrak{g}$, and since $B(\cdot,\cdot)$ is nondegenerate on $\mathfrak{g}$ it must be nondegenerate on each component of the decomposition, proving the second and third statements of 1.

Suppose $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_{-\alpha}$, and $H \in \mathfrak{h}$. Then $B(H,[X,Y]) = B([H,X],Y) = \alpha(H)B(X,Y)$, proving 2.

Finally suppose $H \in \mathfrak{h}$. Then $B(H,h_\alpha) = \alpha(H)$, so by 2 we have $B(H,[X,Y]) = B(H, B(X,Y) \cdot h_\alpha)$. This holds for any $H \in \mathfrak{h}$ and $[X,Y], B(X,Y) \cdot h_\alpha \in \mathfrak{h}$, so by the second statement of 1 we conclude that $[X,Y] = B(X,Y) \cdot h_\alpha$. □

Proof of 8.4 and 8.3. First note that by the Jacobi identity, $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$, so in particular $\mathfrak{h}_\alpha \subset \mathfrak{h}$.

First we prove 8.3.1: R is clearly finite and by definition doesn't contain 0. Now suppose R doesn't span $\mathfrak{h}^*$, so there is some nonzero $H \in \mathfrak{h}$ such that $R(H) = \{0\}$. Then $\mathrm{ad}(H) = 0$, so $\mathfrak{c}(\mathfrak{g}) := \{X \in \mathfrak{g} : [X, \mathfrak{g}] = 0\} \neq 0$. But $\mathfrak{c}(\mathfrak{g}) \subset \mathfrak{g}$ is an abelian ideal, contradicting the semisimplicity of $\mathfrak{g}$.

Next we prove 8.4.1: by 8.5.3, $\mathfrak{h}_\alpha$ is spanned by $\{h_\alpha\}$, so $\mathfrak{h}_\alpha$ is one-dimensional, and so if $H_\alpha \in \mathfrak{h}_\alpha$ exists then it must be unique. Now suppose that no such H_α exists, so $\alpha \equiv 0$ on $\mathfrak{h}_\alpha$. Pick $X_\alpha \in \mathfrak{g}_\alpha$, $Y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $Z := [X,Y] \neq 0$. Since $\alpha(Z) = 0$, $\mathrm{ad}(Z)(\mathfrak{g}_{\pm\alpha}) = 0$, so $[Z,X] = [Z,Y] = 0$, so the subalgebra $\mathfrak{a} \subset \mathfrak{g}$ generated by X, Y, Z is solvable. Thus, by Lie's theorem, there is a basis of $\mathfrak{g}$ such that the adjoint action of $\mathfrak{a}$ on $\mathfrak{g}$ is upper-triangular. Since $Z \in [\mathfrak{a}, \mathfrak{a}]$, $\mathrm{ad}(Z)$ has zeros along the diagonal in this basis, so $Z = Z_n$. On the other hand, $Z \in \mathfrak{h}$, so $Z = Z_s$. Thus $Z = 0$, contradicting our assumption.

To finish 8.4.1 it remains to show that the $\mathfrak{g}_\alpha$ (for $\alpha \in R$) are one-dimensional. To do so, we first show the first statement of 8.4.2 (existence but not uniqueness of a Y_α). Suppose X_α is some nonzero element of $\mathfrak{g}_\alpha$. By 8.5.1, there is some $Y \in \mathfrak{g}_{-\alpha}$ such that $B(X_\alpha, Y) \neq 0$, so $\mathbb{C} \cdot [X_\alpha, Y] = \mathfrak{h}_\alpha$, and scaling Y appropriately we get a Y_α. Now suppose $\dim \mathfrak{g}_\alpha > 1$, so (since $B(\mathfrak{g}_{-\alpha}, \cdot)$ contains all of $\mathfrak{g}_\alpha^*$), there is some nonzero $Y \in \mathfrak{g}_{-\alpha}$ such that $B(Y, X_\alpha) = 0$, and so by 8.5.3, $[X_\alpha, Y] = 0$. On the other hand, since $Y \in \mathfrak{g}_{-\alpha}$, $[H_\alpha, Y] = -\alpha(H_\alpha) \cdot Y = -2Y$, so $-2Y = [[X_\alpha, Y_\alpha], Y] = [Y_\alpha, [Y, X_\alpha]] + [X_\alpha, [Y_\alpha, Y]] = [Y_\alpha, 0] + [X_\alpha, 0] = 0$, contradicting our assumption that $Y \neq 0$.

We've pretty much proved 8.4.2. The uniqueness of Y_α is guaranteed by the fact that $\dim \mathfrak{g}_{-\alpha} = 1$, to it remains to show the two commutator relations. But these follow from the choice of H_α and the fact that $X \in \mathfrak{g}_\alpha, Y \in \mathfrak{g}_{-\alpha}$. Thus by the relations in example 2.3, $(X, Y, H) \mapsto (X_\alpha, Y_\alpha, H_\alpha)$ gives an isomorphism $\phi_\alpha : \mathfrak{sl}_2 \to \mathfrak{s}_\alpha$, and $\mathfrak{s}_\alpha = \mathfrak{h}_\alpha \oplus \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$ by the above. In particular this means that for any $\alpha \in R$, the adjoint action of $\mathfrak{s}_\alpha$ on $\mathfrak{g}$ makes $\mathfrak{g}$ into an $\mathfrak{sl}_2$-module.

Now we claim that for $\alpha, \beta \in R$, $\beta(H_\alpha)$ is an integer and $\beta - \beta(H_\alpha)\alpha \in R$; $\gamma \mapsto \gamma - \gamma(H_\alpha)\alpha$ for $\gamma \in \mathfrak{h}^*$ is a symmetry with vector α, in the sense of proposition 8.3, so this will prove 8.3.2 and 8.3.3. Let Y be a nonzero element of $\mathfrak{g}_\beta$ and let $p = \beta(H_\alpha)$. Since $Y \in \mathfrak{g}_\beta$, $[H_\alpha, Y] = p \cdot Y$. This means that, considered as an $\mathfrak{sl}_2$-module in the way of the last paragraph, Y has weight p. Thus by claim 7.3, p is an integer. Now we set $Z := \mathrm{ad}(Y_\alpha)^p(Y)$ if $p \geq 0$ and $Z := \mathrm{ad}(X_\alpha)^{-p}(Y)$ if $p < 0$. By the second half of claim 7.3, $Z \neq 0$, and by the observation we made at the very beginning of this proof, Z has weight $\beta - p\alpha$. Thus $\beta - p\alpha \in R$.

For 8.3, only 8.3.4 remains. Suppose 8.3.4 does not hold, so there exist $\alpha', \alpha \in R$ such that $\alpha = c\cdot\alpha'$ for some $c \in \mathbb{C}\setminus\{\pm 1\}$. WLOG we may assume that $|c| \leq 1$, as otherwise we can swap the roles of α' and α. By the last paragraph, $\alpha(H_{\alpha'})$ is an integer (and $\alpha(H_{\alpha'}) = c\cdot\alpha'(H_{\alpha'}) = 2c$), so $2c$ must be an integer. Thus $c = \pm\frac{1}{2}$. Since $-\alpha'$ is also a root, this means 2α is a root. By the same argument this means 3α cannot be a root, as then we would have $2\alpha = \frac{2}{3}(3\alpha)$, even though $|\frac{2}{3}| < 1$ and $2\cdot\frac{2}{3}$ is not an integer. Let Y be a nonzero element of $\mathfrak{g}_{2\alpha}$, so $[H_\alpha, Y] = 2\alpha(H_\alpha)Y = 4Y$. On the other hand, since 3α isn't a root, $[X_\alpha, Y] = 0$, so by the Jacobi identity $[H_\alpha, Y] = [[X_\alpha, Y_\alpha], Y] = [X_\alpha, [Y_\alpha, Y]]$. However, $[Y_\alpha, Y] \in \mathfrak{g}_\alpha$, so by the above it's the span of X_α, so $4Y = [X_\alpha, [Y_\alpha, Y]] = 0$, contradicting our assumption. □

Since R is a (complex) root system, there is a subset $S \subset R$, called a base of R, such S is a basis for $\mathfrak{h}^*$ and every element of R can be written in the form $\sum_{\alpha\in S} m_\alpha \alpha$, where the m_α are all integers of the same sign. We will not prove this fact, but the reasoning goes like this: prove the existence of bases geometrically for root systems over $\mathbb{R}$, then show that every complex root system is the complexification of a real root system. For details see sections V.8 and V.17 of [2]. Thus R can be partitioned into R^+ and R^-, the elements with at least one positive m_α or one negative m_α respectively.

We define the Borel subalgebra $\mathfrak{b}$ (of $\mathfrak{g}$) corresponding to $\mathfrak{h}$ and S to be $\mathfrak{h} \oplus \mathfrak{n}$, where $\mathfrak{n} = \mathfrak{n}^+ := \sum_{\alpha\in R^+} \mathfrak{g}_\alpha$ and likewise $\mathfrak{n}^- := \sum_{\alpha\in R^-} \mathfrak{g}_\alpha$. The only fact about Borel subalgebras that we will use is that $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{b}$.

9 Representations of Semisimple Lie Algebras

Suppose $\mathfrak{g}$ is semisimple, $\mathfrak{h} \subset \mathfrak{g}$ is a Cartan subalgebra, V is a $\mathfrak{g}$-representation, and $S = \{\alpha_i\}$ is a fixed base of R.

For $\omega \in \mathfrak{h}^*$, let $V_\omega = \{v \in V : \forall H \in \mathfrak{h}, Hv = \omega(H)v\}$. We call $\dim V_\omega$ the multiplicity of ω in V. The sum $V' := \sum_\omega V_\omega$ is a subrepresentation of V, since for any $X \in \mathfrak{g}_\alpha, v \in V_\omega$,

$$H(Xv) = [H, X]v + X(Hv) = (\omega + \alpha)(H)(Xv),$$

i.e. $\mathfrak{g}_\alpha V_\omega \subset V_{\alpha+\omega}$. This sum is direct, since eigenspaces with distinct eigenvalues are linearly independent. If $V_\omega \neq 0$ then ω is called a weight of V. The image $\rho(\mathfrak{h})$ must be abelian, so by corollary 4.3 $\rho(\mathfrak{h})$ is simultaneously diagonalizable, so $V \neq 0$ implies $V' \neq 0$. Thus if V is irreducible then it must be of the form $\bigoplus_{\omega\in L} V_\omega$ for some set L.

Definition 9.1. *We call $v \in V$ a primitive element of weight ω if $v \neq 0$, $Hv = \omega(H)v$ for all $H \in \mathfrak{h}$, and $X_\alpha v = 0$ for all $\alpha \in R^+$.*

Proposition 9.2. *If $v \in V$ is a primitive element of weight ω and $E = \mathfrak{g}(v)$, i.e. the $\mathfrak{g}$-representation generated by v, then*

1. *If $\beta_1, \dots, \beta_k$ are the different positive roots, E is spanned as a vector space by elements of the form $Y_{\beta_1}^{m_1} \dots Y_{\beta_k}^{m_k} v$, $m_i \in \mathbb{N}$ ($Y_{\beta_i} \in \mathfrak{g}_{-\beta_i}$ as above),*
2. *The weights of E have the form $\omega - \sum_{i=1}^{n} m_i \alpha_i$, $m_i \in \mathbb{N}$,*
3. *ω has multiplicity 1 in E, and*
4. *E is indecomposable as a $\mathfrak{g}$-representation (so by Weyl's theorem, E is irreducible).*

Proof. Let $A = \mathcal{U}(\mathfrak{g})$ be the universal enveloping algebra of $\mathfrak{g}$, let $B = \mathcal{U}(\mathfrak{b})$ and let $C = \mathcal{U}(\mathfrak{n}^-)$. As we noted at the end of the last section, $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{b}$, so $A = C \cdot B$ and thus $E = A \cdot v = C \cdot B \cdot v$. By definition of a primitive element, v is an eigenvector of $\mathfrak{b}$, so $B \cdot v \subset \mathbb{C}v$, and so $E = C \cdot v$. By the Poincaré-Birkhoff-Witt theorem, C is spanned as a vector space by the monomials $Y_{\beta_1}^{m_1} \dots Y_{\beta_k}^{m_k}$, proving 1. As we saw at the beginning of this section, $Y_{\beta_1}^{m_1} \dots Y_{\beta_k}^{m_k} v \in V_{w - \sum m_i \beta_i}$ and this together with 1 gives us 2. Exactly one of the $Y_{\beta_1}^{m_1} \dots Y_{\beta_k}^{m_k} v$ has weight ω, namely the one with $m_1 = \dots = m_k = 0$, giving 3. Finally, suppose we can write $E = F \oplus G$, F, G some $\mathfrak{g}$-representations. Then $E_\omega = F_\omega \oplus G_\omega$, one of which (by 3) must be 0, meaning the other addend must contain v; WLOG say $v \in F$. Then $E \subset \mathfrak{g} \cdot F = F$, so $G = 0$, proving 4. □

This proposition is nearly where we want to be. Any irreducible V must have at least one primitive element (as otherwise it wouldn't be finite-dimensional) and by the above that primitive element generates V. However, we have not yet excluded the possibility of V having primitive elements of two different weights. Thankfully this is impossible:

Theorem 9.3. *Suppose V is an irreducible $\mathfrak{g}$-representation and $v \in V$ is a primitive element of weight ω. Then*

1. *Up to scalar multiplication, v is the only primitive element of V; we call ω the highest weight of V,*
2. *The weights π of V are all of the form $\omega - \sum m_i \alpha_i$, where $m_i \in \mathbb{N}$,*
3. *$\pi(H_\alpha)$ is an integer for any $\alpha \in R$, and*
4. *For two irreducible $\mathfrak{g}$-modules V_1 and V_2, $V_1 \cong V_2$ if and only if their highest weights, ω_1 and ω_2, are equal.*

Proof. By 9.2.4, $\mathfrak{g}(v)$ is an irreducible, nonzero subrepresentation of V, so $\mathfrak{g}(v) = V$. Thus 2 follows from 9.2.2.

For 1, suppose v' is another primitive element of V of weight ω', so by 2 it is of the form $\omega - \sum m_i \alpha_i$ for some $m_i \in \mathbb{N}$. However, swapping the roles of v' and v, we also have $\omega = \omega' - \sum m_i' \alpha_i$ for some $m_i' \in \mathbb{N}$. These two can only hold if $m_i = m_i' = 0$ for all i. Thus ω is unique in its status as a weight with a primitive element in V, and by 9.2.3, $\dim V_\omega = 1$, so v is the only primitive element of V up to scalar multiplication.

To prove 3, consider V as an $\mathfrak{s}_\alpha$-representation. By claim 7.3 and the isomorphism ϕ_α of theorem 8.4, $\pi(H_\alpha)$ is an integer.

Finally, suppose V_1 and V_2 are irreducible representations with primitive elements v_1 and v_2 of highest weights ω_1 and ω_2 respectively. Clearly if V_1 and V_2 are isomorphic then $\omega_1 = \omega_2$, so it remains to show the converse. Suppose $\omega_1 = \omega_2 = \omega$ and consider the $\mathfrak{g}$-representation $V = V_1 \oplus V_2$. V has $v = v_1 + v_2$ as a primitive element of weight ω. Let E be $\mathfrak{g}(v) \subset V$ and let Π_i be the projection from V to V_i; $\Pi_2|_E := f_2$ gives a homomorphism of $\mathfrak{g}$-representations $E \to V_2$. Now, $f_2(v) = v_2$ and $V_2 = \mathfrak{g}(v_2)$, so f_2 is surjective. Further, $\ker f_2 = V_1 \cap E = 0$: On the one hand E does not contain v_1, because v_1 is not in the span of v and by 9.2.3, ω has multiplicity 1 in E. On the other hand $\ker f_2$ is a $\mathfrak{g}$-subrepresentation of V_1, which is irreducible and so would have to contain v_1 if it were to be nonzero. Thus f_2 is an isomorphism, so $E \cong V_2$. By the same argument $E \cong V_1$, so $V_2 \cong V_1$. □

Now, at last, we have the promised result. This theorem greatly simplifies both the problem of finding all irreducible representations of a given semisimple Lie algebra—which it reduces to finding representations generated by a primitive element of each possible highest weight—as well as the study of representations built out of other representations, for example by tensor operations, as weights and primitive elements (like characters of representations of finite groups) are easier to keep track of under such operations than the $\mathfrak{g}$-action itself.

This, as we have said, is one of the things Fulton and Harris use this theorem for: to show that applying operations called Schur functors to the standard representation $\mathbb{C}^n$ of $\mathfrak{sl}_n$ gives all irreducible representations of $\mathfrak{sl}_n$, they first show[4] that the resulting representations are irreducible and then calculate their highest weights combinatorially to show that every possible highest weight allowed by 9.3.3 occurs.

This concludes the paper. If you want to learn more about Lie algebras and their representations, try starting with either [3] or chapters 7-9 of [1], both of which give more information about the structure of general Lie algebras and their connection to Lie groups.

Acknowledgments

Before ending I'd like to thank Drew Moore, who was an awesome mentor, and the faculty members who donated their time to the 2016 REU.

[4]by an application of Schur-Weyl duality; see chapter 5 of [6] for a thorough explanation.

References

[1] William Fulton and Joe Harris. *Representation Theory: A First Course*. Springer Science+Business Media Inc., New York, NY, 2003.

[2] Jean-Pierre Serre. *Complex Semisimple Lie Algebras*. Springer-Verlag Berlin Heidelberg New York, 1987.

[3] Jean-Pierre Serre. *Lie Algebras and Lie Groups: 1964 Lectures given at Harvard University*, 2nd ed. Springer-Verlag Berlin Heidelberg New York, 2006.

[4] Theodor Bröcker and Tammo tom Dieck. *Representations of Compact Lie Groups*. Springer-Verlag Berlin Heidelberg New York, 1985.

[5] Charles A. Weibel. *An Introduction to Homological Algebra*. Cambridge University Press, Cambridge, UK, 1994.

[6] Pavel Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, and Elena Yudovina. *Introduction to Representation Theory*. American Mathematical Society, Providence, RI, 2011.