SEMISIMPLE LIE ALGEBRAS AND THE CHEVALLEY GROUP CONSTRUCTION

ALEX ZORN

Abstract. This paper illustrates the construction of a Chevalley group for a finite dimensional semisimple Lie algebra over an algebraically complete field.

Contents

Part 1. Introduction

In this paper, the theory of Lie algebras is introduced, with a special focus on the analysis of finite dimensional semisimple Lie algebras. The theory is developed essentially from the ground up, the only prerequisites are a basic knowledge of linear algebra and ring theory. We do not explore the connection between Lie algebras and Lie groups, in particular the definition of Lie algebra does not arise from studying differentiable manifolds. Instead, we present the Lie algebra axioms,

Date: August 22, 2010.

motivated by the algebraic structure of End(V), the ring of endomorphisms of a vector space V.

Part two deals with an introduction to the theory of Lie algebras, which are vector spaces equipped with a bilinear product called the **bracket product** and denoted $[xy]$ or $[x, y]$, satisfying certain axioms. In the case of End(V), $[xy] = xy - yx$ and when we view this space as a Lie algebra we denote it by $\mathfrak{gl}(V)$ and call it the **general linear algebra**. In section two we have the basic algebraic definitions (eg homomorphism, subalgebra) familiar to anyone who has studied algebra. In section three we discuss representations, maps from a Lie algebra L to $\mathfrak{gl}(V)$ for some V. Representations, especially the adjoint representation, will be a central tool for studying Lie algebras, since they allow many theorems about $\mathfrak{gl}(V)$ to be applied to L. Section four deals with some more advanced definitions, including the notion of a semisimple Lie algebra, which will encompass most of our discussion in the rest of the paper. We also prove Engel's theorem, which gives an important criterion for the nilpotency of a Lie algebra. Section 5 discusses Jordan decomposition, which states that a matrix can be written as the sum of a semisimple and nilpotent matrix. Section 6 introduces the **Killing form**, a symmetric bilinear associative form on an arbitrary Lie algebra, which is central to the analysis that follows. Weyl's Theorem, in section seven, is a cruicial theorem about representations of semisimple Lie algebras. Finally, section 8 introduces an analogue of Jordan decomposition for an arbitrary semisimple Lie algebra.

In part 3 we discuss the root space decomposition, root systems, and the construction of a Chevalley group. Section 9 discusses representations of a very important Lie algebra, sections 10 and 11 introduce the idea of a root space decomposition for a semisimple Lie algebra and prove some properties. Section 12 discusses root systems, which arise naturally from considering root space decompositions, but are an interesting geometric concept in their own right (and the analysis requires no Lie algebra concepts, only linear algebra ones). Finally section 13 gives us a much-needed theorem about the existence of isomorphisms between semisimple Lie algebras (and in particular, automorphisms of a particular Lie algebra), and section 14 explains the construction of a Chevalley group (of adjoint type), which is the goal of the paper.

1. Linear Algebra Review

Here we review the interaction of a vector space V (over F) and its endomorphism ring End(V) with the ring of polynomials $F[T]$ over F.

If F is a field and V is a vector space over F, there is a natural homomorphism $F \to \text{End}(V)$ sending each element to scalar multiplication. This allows us to take a polynomial with coefficients in F and evaluate it at some $T \in \text{End}(V)$. To emphasize the fact that the indeterminant will be a linear transformation, we write $F[T]$ for the polynomial ring over F and $p(T)$ for some polynomial in $F[T]$.

Now assume V is finite dimensional. Since $F[T]$ is a principal ideal domain, for any $x \in \text{End}(V)$ we can find a minimal polynomial that has x has a root, this is called the minimal polynomial of x and is denoted $m_x(T)$. We also have the characteristic polynomial $c_x(\lambda) = \det(\lambda I - x)$. In this case we have $\lambda \in F$ and the expression on the right is just the evaluation of the determinant of a linear transformation, which turns out to be a polynomial in λ. We will usually write this as $c_x(T)$, which agrees with our usual notation, but remember $c_x(T) \neq \det(T - x)$. It is a consequence of the Cayley-Hamilton theorem that $m_x(T)$ divides $c_x(T)$.

The Chinese Remainder theorem says given polynomials $p_i(T)$ and $q_i(T)$, we can satisfy the relations $p(T) \equiv p_i(T) \pmod{q_i(T)}$ as long as the $q_i(T)$ are relatively prime.

Finally, if we fix $x \in \text{End}(V)$ we can regard V as an $F[T]-$module, that is for a polynomial $p(T)$, $v \in V$, the action of $p(T)$ on v is $p(x)(v)$. Then it is a theorem of module theory that if $p(T) = p_1(T) \cdots p_n(T)$ is a decomposition into pairwise relatively prime polynomials, the kernel (In V) of the action by $p(T)$ is the direct sum of the kernels of the actions of the $p_i(T)$. In particular, if $p(T) = c_x(T)$, and F is algebraically closed, we can write $c_x(T) = \prod_i (T - a_i)^{m_i}$. Thus ker $c_x(x) = \oplus$ ker $(x - a_i)^{m_i}$ where the polynomials on the right are the factors of $c_x(T)$ to their linear powers. Finally, by the Cayley-Hamilton theorem, $c_x(x) = 0$ so we get a decomposition of V into subspaces which are invariant under x.

Lastly, recall that the **trace** of an endomorphism T of a finite dimensional vector space V (written $Tr(T)$) is the sum of the diagonal entries of the matrix representation

Part 2. Lie Algebras

2. Introduction to Lie Algebras

We begin our discussion with a definition:

Definition 2.1. A **Lie Algebra** is a vector space L endowed with an operation $[\cdot\cdot] : L \times L \to L$, called a **bracket product**. The bracket product satisfies the following three axioms:

(L1) The bracket product is bilinear.
(L2) $[xx] = 0$ for all $x \in L$
(L3) $[x[yz]] + [y[zx]] + [z[xy]] = 0$ for all $x, y, z \in L$

Note that the bracket product is not necessarily associative, instead we have axiom (L3) which is called the **Jacobi identity**. It is also worth remarking at this point that we will occasionally write $[x, y]$ instead of $[xy]$ for clarity.

The motivation for the above axioms comes from the following idea: Consider, for a vector space V, the vector space $\text{End}(V)$ consisting of endomorphisms of V (that is, linear operators $x : V \to V$). For $x, y \in \text{End}(V)$, define $[xy] = xy - yx$. The reader should check that $\text{End}(V)$ endowed with this bracket product is a Lie algebra, which we denote by $\mathfrak{gl}(V)$.We call it the **general linear algebra**. (Note that $\text{End}(V)$ and $\mathfrak{gl}(V)$ are the same objects. We use the latter notation when we want to emphasize the Lie algebra structure).

We now introduce some terminology, which is standard for any kind of algebraic analysis and should feel familiar:

Definition 2.2. Let L be a Lie algebra over a field F.

- A (vector) subspace K of L is a **subalgebra** if it is closed under the bracket product: $x \in K$ and $y \in K$ imply $[xy] \in K$.

- A subspace I of L is an **ideal** if for $x \in L$, $y \in I$, $[xy] \in I$. (obviously ideals are subalgebras)

- If A and B are subspaces of L, $A + B = \{x + y | x \in A, y \in B\}$ and $[AB] = \{\sum_i [x_i y_i] | x_i \in A, y_i \in B\}$. These are also subspaces of L. If I and J are ideals of L, then $I + J$ and $[IJ]$ are also ideals (proof left to reader).

- $[LL]$ is called the **derived algebra** of L, and L is **abelian** if $[LL] = 0$.

- L is **simple** if it is nonabelian and has exactly two ideals: itself and 0.

- If K is a subspace of L, $N_L(K) = \{x \in L | [xk] \in K \text{ for all } k \in K\}$ is called the **normalizer** of K in L. $N_L(K)$ is a subalgebra of L, and If K is a subalgebra of L, $N_L(K)$ is the largest subalgebra of L having K as an ideal (proof left to reader).

- If K is a subalgebra of L, K is **self-normalizing** if $K = N_L(K)$

- If X is a subset of L, $C_L(X) = \{y \in L | [xy] = 0 \text{ for all } x \in X\}$ is called the **centralizer** of X in L. $C_L(X)$ is a subalgebra of L, and is an ideal if X is an ideal (proof left to reader). If K is a subalgebra of $C_L(X)$, we say K **centralizes** X.

- $C_L(L)$ is written $Z(L)$ and is called the **center** of L.

- If L and L' are Lie algebras over a common field F, a linear map $\psi : L \to L'$ is a **homomorphism** if $\psi[xy] = [\psi(x)\psi(y)]$ for all $x, y \in L$. If ψ is injective it is called a **monomorphism** and if it is surjective it is called a **epimorphism**. If it is both mono- and epi- it is an **isomorphism**, and L and L' are said to be **isomorphic** (written $L \cong L'$)

The following elementary proposition is should also look familiar:

Proposition 2.3. *Suppose L and L' are Lie algebras over F, and $\psi : L \to L'$ is a homomorphism. Then Ker $\psi = \{x \in L | \psi(x) = 0\}$ is an ideal of L, and $\psi(L)$, the image of L, is a subalgebra of L'.*

Proof. Left to reader. □

If I is an ideal of a Lie algebra L, then we can construct the quotient algebra L/I, and the familiar homomorphism theorems all hold. This is the content of the following proposition:

Proposition 2.4. *Let L be a Lie algebra, and $I \subset L$ an ideal. Then we can define a bracket product on the quotient space L/I by $[a+I, b+I] = [ab] + I$. With this bracket product, L/I is a Lie algebra, and the canonical map $\psi : L \to L/I$ is a homomorphism. The following results also hold:*

a) If L and L' are Lie algebras over a field F and $\psi : L \to L'$ is a homomorphism, then Ker ψ, the kernel of ψ, is an ideal of L. We have $L/$ Ker $\psi \cong L'$, in particular there is a unique isomorphism $\phi : L/$ Ker $\psi \to L'$ such that $\psi = \phi \circ \pi$.

b) If I and J are ideals of L with $I \subset J$, then J/I is an ideal of L/I and $(L/I)/(J/I) \cong L/J$.

c) If I and J are ideals of L, $(I+J)/J \cong I/(I \cap J)$.

Proof. Straightforward. This imitates the proof for other algebraic structures (eg. rings) nearly exactly. □

We end this section with a final elementary proposition (again, the reader should supply the proof if he or she so desires):

Proposition 2.5. *If L and L' are Lie algebras over F, $\psi : L \to L'$ a homomorphism, and I, J ideals of L, then $\psi(I+J) = \psi(I) + \psi(J)$ and $\psi([IJ]) = [\psi(I)\psi(J)]$. Also, these are all ideals in $\psi(L)$.*

Proof. Easy □

3. Representations

As we shall see, an incredibly useful tool for studying a lie algebra is the notion of a representation:

Definition 3.1. If L is a Lie algebra over a field F, a **representation** of L is a homomorphism $\phi : L \to \mathfrak{gl}(V)$, where V is a vector space over F. ϕ is said to be **finite dimensional** if V is.

If $\phi : L \to \mathfrak{gl}(V)$ is a representation, we will sometimes say L *acts on* V via ϕ. What this means is that given an element $x \in L$ we can interpret x as taking elements of V and moves them (this is the action) to other elements of V, explicitly x takes $v \in V$ to $\phi(x)(v)$ (as $\phi(x)$ is a function from V to V). For clarity of notation, if T is an element of $\mathfrak{gl}(V)$ we will occasionally write $T.v$ for $T(v)$, so when $\phi : L \to \mathfrak{gl}(V)$ is a representation we can write $\phi(x).v$ for $\phi(x)(v)$.

Now if L is any Lie algebra, we have a representation called "ad" as follows: for $x \in L$, $\operatorname{ad} x : L \to L$ is defined by $\operatorname{ad} x(y) = [xy]$. Then one can easily verify that $\operatorname{ad} : L \to \mathfrak{gl}(L)$ is a representation, it is called the **adjoint representation** of L. Every Lie algebra has an adjoint representation, and we can add in a subscript, as in ad_L, to distinguish the adjoint representation of L from, say, the adjoint representation ad_K of a subalgebra K of L.

Using ad, we can transfer notions from vector spaces of the form $\operatorname{End}(V)$ to arbitrary Lie algebras. For example, a linear operator $x \in \operatorname{End}(V)$ is said to be **nilpotent** if $x^n = 0$ for some $n \in \mathbb{N}$. Then if L is any Lie algebra, $x \in L$ is **ad-nilpotent** if $\operatorname{ad} x \in \operatorname{End}(L)$ is nilpotent. Similarly, if V is finite dimensional over F, $x \in \operatorname{End}(V)$ is said to be **semisimple** if the roots in F of $m_x(T)$, the minimal polynomial of x, are all distinct. Then if L is any finite dimensional Lie algebra, $x \in L$ is **ad-semisimple** if $\operatorname{ad} x \in \operatorname{End}(L)$ is semisimple. Note that, if F is an algebraically closed field, $x \in \operatorname{End}(V)$ semisimple if and only if it is diagonalizable.

Proposition 3.2. *Let V be a vector space over a field F, and $x \in \mathfrak{gl}(V)$.*

(a) If x is nilpotent, then x is ad-nilpotent.
(b) If V is finite dimensional, F algebraically closed, x semisimple, then x is ad-semisimple.

Proof. (a): Fix x, $(x^n = 0)$ and consider two endomorphisms of $\operatorname{End}(V)$: λ_x and ρ_x where for $y \in \operatorname{End}(V)$, $\lambda_x(y) = xy$ and $\rho_x(y) = yx$. Then $\lambda_x^n(y) = x^n y = 0$, and $\rho_x^n(y) = yx^n = 0$, so they are both nilpotent endomorphisms of $\operatorname{End}(V)$. They also commute: $\lambda_x \rho_x(y) = \rho_x \lambda_x(y) = xyx$. Then $\operatorname{ad} x = \lambda_x - \rho_x$, and since the sum or difference of nilpotent endomorphisms is nilpotent, $\operatorname{ad} x$ is nilpotent.

(b): Let $v_1, v_2, \ldots, v_n$ be a basis for V that diagonalizes x, and say $a_1, \ldots, a_n$ are the eigenvalues. Now consider the basis e_{ij} for $\mathfrak{gl}(V)$, where $e_{ij}(v_k) = \delta_{ik} v_j$, δ the Kronecker delta. We will see that the basis e_{ij} diagonalizes $\operatorname{ad} x$. We have $\operatorname{ad} x(e_{ij})(v_k) = (xe_{ij} - e_{ij}x)v_k = x\delta_{ik}v_j - e_{ij}a_k v_k = a_j \delta_{ik} v_j - a_k \delta_{ik} v_j$. But $a_k \delta_{ik} = a_i \delta_{ik}$ (since if $i \neq k$ both sides are zero), so $\operatorname{ad} x(e_{ij})(v_j) = (a_j - a_i)\delta_{ik} v_j = (a_j - a_i) e_{ij}(v_k)$. We get that e_{ij} is an eigenvector of $\operatorname{ad} x$, with eigenvalue $a_j - a_i$. □

Now if L acts on V via some representation ϕ, and W is a subspace of V stabilized by L ($\phi(x).w \in W$ for all $x \in L$, $w \in W$), then by restricting the action of L to W we get L acting on W. Similarly, if L stabilizes $U \subset W$, with U and W subspaces of V, then the action of L on W/U via ϕ is well-defined (as the reader should check). The most general case of this phenomenon is explained in the proposition below:

Proposition 3.3. *L acts on V via ϕ. K is a subalgebra of L and I is an ideal of K. $U \subset W$ are subspaces of V. K stabilizes U and W, and the action of I maps U into 0. Then K/I acts on W/U via ϕ.*

Remark 3.4. Since restricted actions satisfy the same equations as the original actions, (formally, restricting the range to a subspace or quotient subspace is a ring homomorphism) this restriction to a subspace/quotient preserves nilpotency (characterized by the equation $x^n = 0$ for some n) and semisimplicity (characterized by x being the root of a polynomial with no duplicate roots in F). Examples:

- If K is a subalgebra of L, and $\operatorname{ad}_L x$ is nilpotent for some $x \in K$, so is $\operatorname{ad}_K x$
- x as above. K stabilizes L and K so (even when K is not an ideal) K acts on (the vector space) L/K via ad, and the action of x is nilpotent.
- I is an ideal of L, and $x \in L$ with $\operatorname{ad}_L x$ nilpotent, then $\operatorname{ad}_{L/I} \overline{x}$ is nilpotent.
- Respectively, semisimple in each of the above cases (L finite dimensional)

We end this section with a final definition which will be useful later:

Definition 3.5. If $\phi : L \to \mathfrak{gl}(V)$ is a representation, we say ϕ is **irreducible** if there are no proper subspaces of V stabilized by L. ϕ is **completely reducible** if V is the direct sum of subspaces V_i stabilized by L such that the restricted action of L on V_i is irreducible.

4. Solvable, Semisimple, and Nilpotent

We now introduce three more concepts which are slightly more intricate than those discussed in Definition 2.3. The first is the notion of solvability of a Lie algebra, which mimics solvability in group theory. Given a Lie algebra L, we define a sequence of ideals of L as follows: $L^{(0)} = L$, $L^{(1)} = [LL]$, $L^{(2)} = [L^{(1)}L^{(1)}]$, and in general $L^{(i)} = [L^{(i-1)}L^{(i-1)}]$.

Definition 4.1. A Lie algebra L is **solvable** if $L^{(n)} = 0$ for some n.

We have the following proposition:

Proposition 4.2. *Let L be a Lie algebra over a field F.*

(a) If L is solvable and K is a subalgebra of L, K is solvable.
(b) Homomorphic images of L are solvable.
(c) If I is an ideal of L such that I and L/I are solvable, then L is solvable.
(d) If I, J are solvable ideals of L, then so is $I + J$.

Proof. (a): Suppose $L^{(n)} = 0$. If K is a subalgebra of L, then clearly $[KK]$ is a subalgebra of $[LL]$. It follows by induction that $K^{(i)} \subset L^{(i)}$ for all i, hence $K^{(n)} \subset L^{(n)} = 0$, so K is solvable.

(b): Suppose $L^{(n)} = 0$, L' is a Lie algebra over F, and $\psi : L \to L'$ is a homomorphism. Write $M = \psi(L)$, a subalgebra of L'. By proposition 2.6, it is clear that $\psi(L^{(i)}) = M^{(i)}$ for all i, hence $M^{(n)} = \psi(L^{(n)}) = 0$, so M is solvable.

(c): Suppose $(L/I)^{(n)} = 0$ and $I^{(m)} = 0$. Let $\pi : L \to L/I$ be the canonical homomorphism. Then by part (b), $\pi(L^{(n)}) = \pi(L)^{(n)} = (L/I)^{(n)} = 0$, so $L^{(n)} \subset I$. By part (a), $(L^{(n)})^{(m)} \subset I^{(m)} = 0$. It is clear that $(L^{(n)})^{(m)} = L^{(n+m)}$, hence L is solvable.

(d): By part (c) of proposition 2.5, $(I + J)/J \cong I/(I \cap J)$, the latter of which is the homomorphic image of the solvable Lie algebra I. Part (a) above then implies $(I + J)/J$ is solvable. On the other hand, J is also solvable. So part(c) above implies $I + J$ is solvable. □

The concept of solvability also inspires the following definition:

Definition 4.3. A Lie algebra L is **semisimple** if is finite dimensional and its only solvable ideal is the 0 ideal.

We now move on to the notion of nilpotency, which at first glance seems quite similar to solvability. We define a sequence of ideals as follows: $L^0 = L$, $L^1 = [LL]$, $L^2 = [LL^1]$, and in general $L^i = [LL^{i-1}]$.

Definition 4.4. A Lie algebra L is **nilpotent** if $L^n = 0$ for some n.

It is easy to see that $L^{(i)} \subset L^i$, hence nilpotent Lie algebras are solvable. We have the following elementary proposition for nilpotent Lie algebras, analogous to proposition 4.2 above:

Proposition 4.5. *Let L be a Lie algebra over a field F.*
(a) If L is nilpotent and K is a subalgebra of L, K is nilpotent.
(b) Homomorphic images of L are solvable.
(c) If $L/Z(L)$ is nilpotent, then so is L.
(d) If L is nilpotent and nonzero, then $Z(L) \neq 0$.

Proof. (a): Suppose $L^n = 0$. We can show by induction that $K^i \subset L^i$. By assumption, $K^0 \subset L^0$. Then by induction $K^{i+1} = [KK^i] \subset [LK^i] \subset [LL^i] = L^{i+1}$. Then $K^n \subset L^n = 0$, so K is nilpotent.

(b): Suppose $L^n = 0$, L' is a Lie algebra over F, and $\psi : L \to L'$ is a homomorphism. Write $M = \psi(L)$, a subalgebra of L'. By proposition 2.6, it is clear that $\psi(L^i) = M^i$ for all i, hence $M^n = \psi(L^n) = 0$, so M is solvable.

(c): Suppose $(L/Z(L))^n = 0$, and let $\pi : L \to L/Z(L)$ be the canonical homomorphism. By (b), $\pi(L^n) = (L/Z(L))^n = 0$, hence $L^n \subset Z(L)$, so $L^{n+1} \subset [LZ(L)] = 0$.

(d): Pick m such that L^m is nonzero, but $L^{m+1} = 0$. Then for $x \neq 0 \in L^m$, $y \in L$ we have $[xy] \in L^{m+1} = 0$, so $x \in Z(L)$. □

The reader will surely notice an overlap in terminology: We have used the words "semisimple" and "nilpotent" to describe both Lie algebras and elements of $\mathfrak{gl}(V)$. One may wonder why this is so. In fact, there is a beautiful relationship between nilpotent Lie algebras and nilpotent endomorphisms: If L is a finite-dimensional Lie algebra, then L is nilpotent iff all its elements are ad-nilpotent. This is Engel's theorem, which we will prove shortly. First we introduce a theorem which is a valuable result in and of itself:

Theorem 4.6. *If V is any nonzero vector space and L is a finite dimensional subalgebra of $\mathfrak{gl}(V)$ consisting of nilpotent elements, then there exists nonzero $v \in V$ such that $x.v = 0$ for all $x \in L$.*

Proof. Induct on the dimension of L. If $L = 0$ there is nothing to prove. If dim $L = 1$, pick $x \neq 0 \in L$. By hypothesis x is nilpotent so find $n \in \mathbb{N}$ such that $x^n \neq 0$ but $x^{n+1} = 0$. Then there exists $w \in V$ such that $x^n.w \neq 0$. Letting $v = x^n.w.$, we have $x.v = 0$.

Now let dim $L > 1$ and let K be a maximal proper subalgebra of L. Then (Remark 3.4) K acts on L/K via ad, and the action of each element is nilpotent.

Let $M \subset \mathfrak{gl}(L/K)$ be the image of K under this representation, then dim M ¡ dim L so by induction there exists $z + K \in L/K$ sent to zero by M. So for $x \in K$, $0 + K = \operatorname{ad} x(z + K) = [xz] + K$, or $[xz] \in K$.

So consider $K + Fz$. (F the underlying field). This a subalgebra of L properly containing K, hence it equals L (by maximality). Let W be the subspace of V such that $x.w = 0$ for all $x \in K$, $w \in W$, by induction W is nonzero. If $w \in W$ and $x \in K$, $x(z.w) = [xz].w + z(x.w) = 0$, since $[xz] \in K$. This implies $z.w \in W$, so Fz can be regarded as a subspace of $\mathfrak{gl}(W)$. Since Fz has dimension 1 and z is nilpotent, we know there exists nonzero $w \in W$ such that $z.w = 0$. Then for all $y = x + \alpha z \in K + Fz$, $y.w = x.w + \alpha z.w = 0$, but $K + Fz = L$ so we are done. □

Corollary 4.7. *If L is a finite-dimensional Lie algebra and $\phi : L \to \mathfrak{gl}(V)$ is a representation such that $\phi(x)$ is nilpotent for all $x \in L$, there exists nonzero $v \in V$ such that $\phi(x).v = 0$ for all $x \in L$.*

Proof. Immediately clear from Theorem 4.6. □

Theorem 4.8. Engel's Theorem *If L is a finite dimensional Lie algebra, L is nilpotent if and only if all elements of L are ad-nilpotent.*

Proof. First assume L is nilpotent. if $x \in L$, $y \in L^i$, $\operatorname{ad} x(y) \in L^{i+1}$. In other words, $\operatorname{ad} x(L^i) \subset L^{i+1}$, so $\operatorname{ad} x^n(L) \subset L^n = 0$. Therefore, $\operatorname{ad} x$ is nilpotent.

On the other hand, assume L consists of ad-nilpotent elements. We proceed by induction. If dim L equals 0 or 1, L is nilpotent. If dim $L > 1$, by Corollary 3.8 there exists nonzero $x \in L$ such that $\operatorname{ad} y(x) = 0$ for all $y \in L$. This means $x \in Z(L)$, so $Z(L) \neq 0$. Then $L/Z(L)$ consists of ad-nilpotent endomorphisms (Remark 3.4), hence by induction $L/Z(L)$ is nilpotent, and by proposition 4.5 we are done. □

We end this section with another application of Theorem 4.6 (via Corollary 4.7):

Lemma 4.9. *Let L is a nilpotent finite dimensional Lie algebra, K a nonzero ideal of L. Then $K \cap Z(L) \neq 0$.*

Proof. L acts on K via ad, so Corollary 4.7 gives nonzero $z \in K$ such that $\operatorname{ad} x(z) = 0$ for all $x \in L$. This implies $z \in Z(L)$. □

5. Jordan Decomposition

Recall from linear algebra that if V is a finite dimensional vector space over an algebraically closed field we can write any linear transformation T as an upper-triangular matrix with respect to some basis. Then we can write $T = S + N$, where S is a diagonal matrix and N is strictly upper-triangular (zeroes on the diagonal). Note also that S and N commute. In fact, we have the following general proposition:

Proposition 5.1. *Let V be a finite dimensional vector space over F, F an algebraically closed field, $x \in$ End(V)*

(a) There exist unique $x_s, x_n \in$ End(V) satisfying $x = x_s + x_n$, x_s semisimple, x_n nilpotent, and x_s and x_n commute.
(b) x_s and x_n commute with any endomorphism commuting with x.
(c) If A and B are subspaces of V such that x maps A into B, then x_s and x_n both map A into B.

Proof. (If necessary, the reader should consult the introduction for a brief review of End(V) as an $F[x]-$module).

Since F is algebraically closed we can completely factor the characteristic polynomial of x: $c_x(T) = \prod_i (T - a_i)^{m_i}$ where a_i are the roots and m_i are the multiplicities. Then we can write V as a direct sum of $V_i = \text{Ker } (x - a_i)^{m_i}$, each stable under x. Using the Chinese Remainder Theorem (in $F[T]$), find $p(T)$ such that $p(T) \equiv a_i \pmod{(T - a_i)^{m_i}}$, $p(T) \equiv 0 \pmod{T}$. Note that the last equation is unnecessary if $a_i = 0$ for some i, and if all the a_i are nonzero then T is relatively prime to the other moduli. Let $q(T) = T - p(T)$.

Now set $x_n = q(x)$, $x_s = p(x)$. Since they are each polynomials in x they commute with each other (and every endomorphism commuting with x), and their sum is clearly x. Now if $v \in V_i$, we know we can write $x_s = a_i + r(x)(x - a_i)^{m_i}$ for some $r(x)$, hence $x_s(v) = a_i v + r(x)(x - a_i)^{m_i}(v) = a_i v$, so any basis of V_i will consist of eigenvectors of x_s (with eigenvalue a_i). Hence we can diagonalize x_s, so x_s is semisimple. Lastly, if $v \in V_i$, we have $x_n(v) = (x - x_s)(v) = x(v) - x_s(v) = x(v) - a_i v = (x - a_i)(v)$, hence $x_n^{m_i}(v) = (x - a_i)^{m_i}(v) = 0$. So $x_n^M = 0$, where $M = max(m_i)$, therefore x_n is nilpotent.

(c) is now clear, since $q(T)$ and $p(T)$ are both polynomials with no constant term.

We have shown everything but the "uniqueness" clause in part (a). Suppose $x = s + n$ is a different decomposition of x. Then s and n commute with x, hence with x_s and x_n (part (b)). The sum of commuting nilpotent endomorphisms is nilpotent (clear) and the sum of commuting semisimple endomorphisms is semisimple, since they can be simultaneously diagonalized (a standard linear algebra fact, see Lemma 9.3: Simultaneous Diagonalization for details). So we have $x_s - s = n - x_n$ is a semisimple, nilpotent endomorphism, hence it must be identically zero. □

This decomposition is known as the **Jordan decomposition**, or **Chevalley-Jordan decomposition**. Note that part (a) of the theorem justifies the word "the" (as the decomposition is unique). We also have:

Lemma 5.2. *Let $x \in$ End(V), V a finite dimensional vector space over an algebraically closed field F, $x = x_s + x_n$ its Jordan decomposition. Then $ad\, x = ad\, x_s + ad\, x_n$ is the Jordan decomposition in End(End(V)).*

Proof. By Proposition 3.2, $\text{ad}\, x_s$ is semisimple, $\text{ad}\, x_n$ is nilpotent, and we know $[\text{ad}\, x_s, \text{ad}\, x_n] = \text{ad}\, [x_s, x_n] = 0$. □

6. Killing Form

From here on, we will be primarily working with algebraically closed fields of characteristic zero. We will call such fields ”typical”. In addition, we will use ”typical Lie algebra” to denote a finite-dimensional Lie algebra over a typical field.

Jordan decomposition is one of the central tools we will use in our analysis of Lie algebras. The second central tool is known as the **Killing form**, and is defined as follows:

Definition 6.1. Let L be a finite dimensional Lie algebra. If $x, y \in L$, define $\kappa(x, y) = Tr(\operatorname{ad} x \operatorname{ad} y)$ (The trace of the linear transformation- see the introduction). κ is called the **Killing form**.

Recall that a **form** on a vector space is simply a function from $L \times L$ to the underlying field F (an example is the dot product in $\mathbb{R}^n$). The Killing form enjoys some useful properties:

Proposition 6.2. *The Killing form κ of L is:*
(a) Symmetric: $\kappa(x, y) = \kappa(y, x)$.
(b) Bilinear
(c) Associative: $\kappa([xy], z) = \kappa(x, [yz])$

Proof. All of these proofs will hinge on an elementary fact about trace established in the introduction: that if $A, B \in \operatorname{End}(V)$, $Tr(AB) = Tr(BA)$.

(a): Obvious from the fact stated above.

(b): We know that $\operatorname{ad} : L \to \mathfrak{gl}(L)$ is linear, and that the trace functional is linear.

(c): We see for any $x, y, z \in \mathfrak{gl}(V)$, V finite dimensional, $Tr([xy]z) = Tr(xyz) - Tr(y(xz)) = Tr(xyz) - Tr((xz)y) = Tr(x[yz])$. Hence $\kappa([xy], z) = Tr(\operatorname{ad}[xy] \operatorname{ad} z) = Tr([\operatorname{ad} x \operatorname{ad} y] \operatorname{ad} z) = Tr(\operatorname{ad} x[\operatorname{ad} y \operatorname{ad} z]) = Tr(\operatorname{ad} x \operatorname{ad}[yz]) = \kappa(x, [yz])$. □

As with the adjoint representation, when K is a subalgebra of L we can distinguish between the Killing form of K and the Killing form of L with a subscript: κ_K versus κ_L. We do, however, have the following useful lemma:

Lemma 6.3. *Let L be a finite dimensional Lie algebra, I an ideal of L. Then if $x, y \in I$, $\kappa_I(x, y) = \kappa_L(x, y)$*

Proof. Create a basis for L by first creating a basis for I and extending it to $L : (x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_n)$ (say dim $I = m$). Now if $x \in I$, consider the matrix representation for $\operatorname{ad}_L x$. For $i > 0$, $x_{m+i} \notin I$ but since I is an ideal $\operatorname{ad}_L x(x_{m+i}) \in I$. This implies the $m + i$ diagonal entry in the matrix is zero, so the terms that contribute to the trace correspond to the basis elements of I: $Tr(\operatorname{ad}_L x) = Tr(\operatorname{ad}_I x)$. Similarly, $Tr((\operatorname{ad}_L x)(\operatorname{ad}_L y)) = Tr(((\operatorname{ad}_L x)(\operatorname{ad}_L y)) |_I) = Tr((\operatorname{ad}_I x)(\operatorname{ad}_I y))$. □

We also have the following definition:

Definition 6.4. Let L be a finite dimensional Lie algebra, κ its Killing form. The **radical** S of κ is the set $\{x \in L | \kappa(x, y) = 0 \text{ for all } y \in L\}$. κ is **nondegenerate** if its radical is 0.

Since κ is associative, its radical S is an ideal: If $x \in S$, $y \in L$, we want to show $[xy] \in S$. This amounts to showing $\kappa([xy], z) = 0$ for any $z \in L$, but $\kappa([xy], z) = \kappa(x, [yz]) = 0$. We want to find conditions on L for the Killing form κ to be nondegenerate. It turns out that (provided char $F = 0$) this is exactly the same as L being semisimple. This result is the goal for the rest of this section. We begin with a lemma:

Lemma 6.5. *Let $A \subset B$ be two subspaces of $\mathfrak{gl}(V)$, V a finite dimensional vector space over a typical field F. Let $M = \{x \in \mathfrak{gl}(V) | [x, b] \subset A \text{ for all } b \in B\}$. If $x \in M$ satisfies $Tr(xy) = 0$ for all $y \in M$, x is nilpotent.*

Proof. Let $x = x_s + x_n$ be the Jordan decomposition of x. Then x_s is diagonalizable, so fix a basis $v_1, \dots, v_m$ of V consisting of eigenvectors of x_s and let $a_1, \dots, a_m$ be the corresponding eigenvalues. If we can show each $a_i = 0$, then $x_s = 0$ so $x = x_n$ is nilpotent, as desired. To do this, consider the smallest subfield E of F containing all the a_i. Then E is a vector space over $\mathbb{Q}$, the prime subfield of F, of finite dimension (at most m). We need to show $E = 0$, equivalently the dual space $E^* = 0$. In other words, any linear function $f : E \to \mathbb{Q}$ is zero.

To show this, pick $f : E \to \mathbb{Q}$ and let $y \in \mathfrak{gl}(V)$ be defined by $y(v_i) = f(a_i)v_i$. We can find a polynomial without constant term $r(T) \in F[T]$ satisfying $r(a_i - a_j) = f(a_i - a_j)$ for all pairs i, j (this follows from Lagrange interpolation). Remember from proposition 3.2 that we have a basis e_{ij} of $\mathfrak{gl}(V)$, where $\operatorname{ad} x_s(e_{ij}) = (a_i - a_j)e_{ij}$ and $\operatorname{ad} y(e_{ij}) = (f(a_i) - f(a_j))e_{ij} = f(a_i - a_j)e_{ij} = r(a_i - a_j)e_{ij}$. It follows that $\operatorname{ad} y = r(\operatorname{ad} s)$.

Now since $\operatorname{ad} x_s$ is the semisimple part of $\operatorname{ad} x$ and $\operatorname{ad} x$ maps B into A, $\operatorname{ad} x_s$ maps B into A. Then since $\operatorname{ad} y$ is a polynomial in $\operatorname{ad} x_s$ without constant term, $\operatorname{ad} y$ maps B into A. This means $y \in M$, so $Tr(xy) = 0$, or $\sum_i a_i f(a_i) = 0$. Applying f to this sum yields $\sum_i f(a_i)^2 = 0$, but the numbers $f(a_i)$ are rational so they are all zero. Since the a_i span E, f must be zero.

□

Note that we used the hypothesis that F had ch. 0 when we asserted that its prime subfield was isomorphic to $\mathbb{Q}$. We now prove a theorem known as Cartan's criterion, which is a criterion for solvability of a Lie algebra:

Theorem 6.6. (Cartan's Criterion) *Let L be a subalgebra of $\mathfrak{gl}(V)$, V a finite dimensional vector space over a typical field. If $Tr(xy) = 0$ for all $x \in [LL]$, $y \in L$, then L is solvable.*

Proof. Use the lemma: V is as above, $B = L$, $A = [LL]$. Then $M = \{z \in \mathfrak{gl}(V) | [zy] \in [LL] \text{ for all } y \in L\}$. We want to show that for $x \in [LL], z \in M$, $Tr(xz) = 0$. Then the theorem will give us x is nilpotent, so $\operatorname{ad}_{\mathfrak{gl}(V)} x$ is nilpotent, so $\mathrm{ad}_{[LL]}\, x$ is nilpotent. Then Engel's theorem gives us $[LL]$ nilpotent, which implies that L is solvable (since $L^{(i+1)} = ([LL])^{(i)} \subset ([LL])^{i}$).

It is enough to show that, for an arbitrary generator $[xy] \in [LL]$ $(x, y \in L)$, and some $z \in M$, $Tr([xy]z) = 0$. But we have $Tr([xy]z) = Tr(x[yz]) = Tr([yz]x)$. But since $z \in M$ and $y \in L$, $[yz] \in [LL]$, so by hypothesis $Tr([yz]x) = 0$.

□

Now, the Killing form enters the picture:

Corollary 6.7. *Let L be a typical Lie algebra, and S is the radical of the Killing form κ.*
(a) If $[LL] \subset S$, then L is solvable.
(b) If $I \subset L$ is an ideal and $[II] \subset S$, then I is solvable.

Proof. (a) Apply the theorem to $\mathrm{ad}\, L \subset \mathfrak{gl}(L)$: Since $\mathrm{ad}\,[LL] = [\,\mathrm{ad}\, L\, \mathrm{ad}\, L]$, if $\mathrm{ad}\, x \in [\,\mathrm{ad}\, L\, \mathrm{ad}\, L]$ we can assume $x \in [LL] \subset S$. So for $\mathrm{ad}\, y \in L$ we have $Tr(\,\mathrm{ad}\, x\, \mathrm{ad}\, y) = \kappa(x, y) = 0$. Hence $\mathrm{ad}\, L$ is solvable. But $\mathrm{ad}\, L \cong L/Z(L)$ and $Z(L)$ is solvable, so L is solvable.

(b) By Lemma 6.3, if S_I is the radical of the Killing form on I, $I \cap S \subset S_I$. Then the assertion is obvious from (a). □

Now all the pieces are in place for the big theorem of this section:

Theorem 6.8. *Let L be a typical Lie algebra. Then L is semisimple if and only if its Killing form is nondegenerate.*

Proof. S is an ideal and $[SS] \subset S$ so by Corollary 6.7(b), S is solvable. Since L is semisimple, this means $S = 0$.

Now assume $S = 0$ and let I be an abelian ideal of L. If $x \in I$ and $y \in L$, $\mathrm{ad}\, x\, \mathrm{ad}\, y$ maps L into I, and $(\,\mathrm{ad}\, x\, \mathrm{ad}\, y)^2$ maps L into $[II] = 0$. Hence $\mathrm{ad}\, x\, \mathrm{ad}\, y$ is nilpotent, so $\kappa(x, y) = Tr(\,\mathrm{ad}\, x\, \mathrm{ad}\, y) = 0$. Therefore $I \subset S = 0$. Now if J were any nonzero solvable ideal of L, the last nonzero $J^{(i)}$ would be an abelian ideal. This is a contradiction, so L is nilpotent. □

We are now in a position to prove that any semisimple typical Lie algebra can be written as the direct sum of simple Lie algebras (sometimes used as the definition of semisimplicity), but first a simple lemma:

Lemma 6.9. *If L is a typical Lie algebra and I is an ideal of L, let $I^{\perp} = \{x \in L | \kappa(x, y) = 0$ for all $y \in I\}$. Then $I^{\perp}$ is an ideal and $L = I \oplus I^{\perp}$ ($I^{\perp}$ is called the* ***ideal perpendicular to** I **(in** L**)***

Proof. $I^{\perp}$ is a subspace of L because of the bilinearity of κ. Now if $x \in I^{\perp}$, $z \in L$ we want to show $[zx] \in I^{\perp}$. For $y \in I$, note $[yz] \in I$. Then by the associativity of κ, $\kappa(y, [zx]) = \kappa([yz], x) = 0$. Then since $I \cap I^{\perp}$ is an ideal, the Killing form is trivial when restricted (Lemma 6.3) so by Corollary 6.7(a), $I \cap I^{\perp}$ is solvable and hence 0.

□

Theorem 6.10. (Decomposition of Semisimple Lie Algebras) *Let L be a semisimple typical Lie algebra. Then there exist ideals $L_1, \dots, L_t$ of L which are simple Lie algebras such that $L = L_1 \oplus \cdots \oplus L_t$, and any ideal of L is a direct sum of some of the L_i (in particular the simple ideals of L are the L_i).*

Proof. Use induction on the dimension of L. If L is already simple, we are done. If not let I be a proper ideal of L. Then (Lemma 6.9 above) we can write $L = I \oplus I^\perp$, where $I^\perp$ is also a (proper) ideal of L. Now any ideal of I is also an ideal of L, so that means I must be semisimple, and similarly $I^\perp$ is semisimple. By induction, we can write I as a direct sum of simple ideals, which are therefore also ideals of L, and likewise for $I^\perp$. This gives us a decomposition of L into a direct sum of simple ideals.

Now let I be a simple ideal of L, $L = L_1 \oplus L_2 \oplus \cdots \oplus L_n$. Then $[IL]$ is an ideal of I, and $[IL] \neq 0$ because $Z(L) = 0$. So $[IL] = I$. Then $I = [IL] = [IL_1] \oplus \cdots \oplus [IL_n]$, so all but one of the sums, say $[IL_i]$, equal zero. So $I = [IL_i]$, hence $I \subset L_i$ but this means $I = L_i$ since L_i is simple.

Finally, let I be any ideal of L. I is semisimple, so we can write I as a direct sum of simple ideals, which are also simple ideals of L. But by the above paragraph, each simple ideal of L is one of the L_i.

□

Corollary 6.11. *If L is semisimple, then $L = [LL]$ and all ideals and homomorphic images of L are semisimple.*

Proof. Clear from Theorem 6.10. □

7. Weyl's Theorem

In this section we are concerned with representations $\phi : L \to \mathfrak{gl}(V)$ of a semisimple typical Lie algebra. First, remember that a form on L is a function from $L \times L$ to F. We have the following helpful lemma:

Lemma 7.1. *If L is a finite dimensional Lie algebra, β any nongenerate bilinear form on L, and $(x_1, \dots, x_n)$ a basis of L, there is a uniquely determined dual basis $(y_1, \dots, y_n)$ relative to β such that $\beta(x_i, y_j) = \delta_{ij}$*

Proof. Let L^* denote the dual space of L. For $x \in L$, define $\lambda_x \in L^*$ by $\lambda_x(y) = \beta(x, y)$. Then by bilinearity, we have both that λ_x is actually in L^* as asserted, and that the map $\lambda : x \mapsto \lambda_x$ is a linear map. Since β is nondegenerate, λ has kernel 0, and since L is finite-dimensional this implies λ is an isomorphism. Then if $f_i \in L^*$ is the linear functional mapping x_j to δ_{ij}, we must have $y_i = \lambda^{-1}(f_i)$. □

Now let L be a Lie algebra over F, and $\phi : L \to \mathfrak{gl}(V)$ be a representation of L with V finite dimensional. If we have $\beta(x, y) = Tr(\phi(x)\phi(y))$, β is a symmetric associative bilinear form- simply use the argument of proposition 5.4 for the Killing form. Also, if F is typical, and ϕ is faithful (1-1), the radical S of β is isomorphicic to $\phi(S)$, and the argument from the previous section then implies S. solvable. Therefore if L is semisimple, β is nondegenerate.

Now pick a basis $x_1, \dots, x_n$ of L, and a basis $y_1, \dots, y_n$ which is dual to the x_i relative to β (as in the lemma). Write $c_\phi = \sum_i \phi(x_i)\phi(y_i)$, and call it the **Casimir element of** ϕ. Its trace is $\sum_i Tr(\phi(x_i)\phi(y_i)) = \sum_i \beta(x_i, y_i) = n$, where n is the dimension of L.

Proposition 7.2. *If $x \in L$, L a typical Lie algebra, and $\phi : L \to \mathfrak{gl}(V)$ is a faithful representation. Then c_ϕ commutes with $\phi(x)$.*

Proof. Write $[xx_i] = \sum_j a_{ij}x_j$ and $[xy_i] = \sum_j b_{ij}y_j$. Then $a_{ik} = \sum_j a_{ij}\beta(x_j, y_k) = \beta([xx_i], y_k) = \beta(-[x_ix], y_k) = -\beta(x_i, [xy_k]) = -\sum_j b_{kj}\beta(x_i, y_j) = -b_{ki}$ (using associativity of β). Now in $\text{End}(V)$ we have $[x, yz] = [x, y]z + y[x, z]$, therefore:

$$[\phi(x), c_\phi(\beta)] = \sum_i [\phi(x), \phi(x_i)\phi(y_i)] = \sum_i [\phi(x), \phi(x_i)]\phi(y_i) + \sum_i \phi(x_i)[\phi(x), \phi(y_i)]$$

We have $[\phi(x), \phi(x_i)] = \sum_j a_{ij}\phi(x_j)$ and $[\phi(x), \phi(y_i)] = \sum_j b_{ij}\phi(y_j)$. So:

$$[\phi(x), c_\phi(\beta)] = \sum_{i,j} a_{ij}\phi(x_j)\phi(y_i) + \sum_{i,j} b_{ij}\phi(x_i)\phi(y_j) = 0$$

Since $a_{ij} = -b_{ji}$.

□

Lemma 7.3. (Schur's Lemma) *L a Lie algebra over an algebraically closed field F, V a finite dimensional vector space over F. $\phi : L \to \mathfrak{gl}(V)$ an irreducible representation. If $T \in \mathfrak{gl}(V)$ commutes with all $\phi(x)$, $T = \lambda I$ for some $\lambda \in F$.*

Proof. Pick an eigenvalue λ of T (possible since V is finite dimensional and F is an ACF), and let $\overline{T} = T - \lambda I$. Then if $v \in \text{Ker } \overline{T}$, $x \in L$, since T commutes with $\phi(x)$ we have $\overline{T}(\phi(x).v) = T(\phi(x).v) - \lambda\phi(x).v = \phi(x).Tv - \lambda\phi(x).v = \phi(x).(Tv - \lambda v) = 0$. Hence $\ker \overline{T}$ is stable under action by L, it is nonzero since λ is an eigenvalue of T, and by irreducibility this implies $\ker \overline{T} = V$ so $\overline{T} = T - \lambda I = 0$. □

Corollary 7.4. *L, $\phi : L \to \mathfrak{gl}(V)$, c_ϕ as in proposition ??, V finite dimensional. If ϕ is faithful and irreducible, $c_\phi = \lambda I$ where $\lambda = \dim L / \dim V$.*

Proof. By Schur's Lemma, $c_\phi = \lambda I$ for some λ, and we get the value of λ using $Tr(c_\phi) = \dim L$. □

Lemma 7.5. *If L is a typical semisimple Lie algebra, $\phi : L \to \mathfrak{gl}(V)$ is a representation, and V is one dimensional, then L acts trivially on V (for any $x \in L$, $v \in V$, $\phi(x).v = 0$).*

Proof. □

Now we proceed with the main theorem of this section:

Theorem 7.6. (Weyl's Theorem) *If L is a typical semisimple Lie algebra and $\phi : L \to \mathfrak{gl}(V)$ a is a finite dimensional representation, then ϕ is completely reducible.*

Proof. Consult [1], page 28

□

8. Abstract Jordan Decomposition

In this section we prove an analogue for Jordan decomposition (section 5) for an arbitrary typical semisimple Lie algebra. The idea is: write $x \in L$ as $x = x_s + x_n$ where x_s is ad-semisimple and x_n is ad-nilpotent, and x_s and x_n commute. First we need a simple definition, motivated by the product rule for derivatives:

Definition 8.1. If L is a Lie algebra, a **derivation** of L is an element $\delta \in \mathfrak{gl}(L)$ such that $\delta([a,b]) = [a, \delta(b)] + [\delta(a), b]$. Denote by Der L the set of derivations of L. It is easily checked (left to the reader) that Der L is a subalgebra of $\mathfrak{gl}(L)$ (though not necessarily a subring).

Lemma 8.2. *L a typical Lie algebra, $\delta \in \mathfrak{gl}(L)$. If $\delta = \sigma + \nu$ is the Jordan decomposition of δ and $\delta \in$ Der L, then $\sigma, \nu \in$ Der L.*

Proof. Write $L_a = \{x \in L | (\delta - a)^k.x = 0$ for some k (depending on x)$\}$. Then L is the direct some of the nonzero L_a, and σ acts on L_a as scalar multiplication by a. Suppose $x \in L_a$ and $y \in L_b$ By induction on n, we can show $(\delta - (a+b))^n.[xy] = \sum_{i=0}^{n} \binom{n}{i}((\delta - a)^{n-i}.x) \cdot ((\delta - b)^i).y)$, hence for big enough n $(\delta - (a+b))^n.[xy] = 0$. This means $[xy] \in L_{a+b}$, so $\sigma([xy]) = (a+b)([xy]) = [(ax), y] + [x, (by)] = [\sigma(x), y] + [x, \sigma(y)]$. Since L is a direct sum of the L_a, it follows that $\sigma \in$ Der L, and therefore so is ν. □

One more lemma is necessary before proceeding with out main theorem:

Lemma 8.3. *If L is a typical, semisimple Lie algebra, then $\operatorname{ad} L =$ Der L.*

Proof. Let $M = \operatorname{ad} L$, $D =$ Der L. The first thing to notice is that M is an ideal of D: If $x \in L$ then by the Jacobi identity $\operatorname{ad} x([yz]) = [x[yz]] = [[xy]z] + [y[xz]] = [\operatorname{ad} x(y), z] + [y, \operatorname{ad} x(z)]$. This implies $M \subset D$. Then if $\delta \in D$, $[\delta, \operatorname{ad} x](y) = \delta([xy]) - \operatorname{ad} x(\delta.y) = [\delta.x, y] + [x, \delta.y] - [x, \delta.y] = [\delta.x, y] = \operatorname{ad} \delta.x(y)$.

By Corollary 6.11, M and D are semisimple. So we can write $D = M \oplus M^{\perp}$ (Lemma 6.9). Pick $\delta \in M^{\perp}$. Since M and $M^{\perp}$ are both ideals, if $x \in L$, $[\delta, \operatorname{ad} x] = \operatorname{ad} \delta.x \in M^{\perp} \cap M = 0$, so $\operatorname{ad} \delta.x = 0$ for all $x \in L$. But the kernel of ad is $Z(L) = 0$ (L being semisimple), so $\delta.x = 0$ for all $x \in L$, meaning $\delta = 0$. So: $M^{\perp} = 0$ and we have $D = M$. □

We are ready for the main theorem of this section:

Theorem 8.4. *Let L be a semisimple typical Lie algebra, and $x \in L$.*

*(a) There exist unique $x_s, x_n \in L$ with $x = x_s + x_n$, x_s ad-semisimple, x_n ad-nilpotent, $[x_s, x_n] = 0$. (This is called the **Abstract Jordan Decomposition** of x).*

(b) If L is a subalgebra of $\mathfrak{gl}(V)$ for a finite dimensional vector space V, the abstract and usual Jordan decompositions coincide.

(c) Let $\phi : L \to \mathfrak{gl}(V)$ be a finite dimensional representation of L, $x \in L$. If $x = x_s + x_n$ is the abstract Jordan decomposition of x and $\phi(x)_s + \phi(x)_n$ is the normal Jordan decomposition of $\phi(x)$, then $\phi(x)_s = \phi(x_s)$ and $\phi(x)_n = \phi(x_n)$. (Note that (b) is a special case of (c) when ϕ is the inclusion map).

Proof. (a): If $\operatorname{ad} x = (\operatorname{ad} x)_s + (\operatorname{ad} x)_n$ is the Jordan decomposition of $\operatorname{ad} x$, since $\operatorname{ad} x \in \operatorname{Der} L$, by Lemma 8.2 we have $(\operatorname{ad} x)_s, (\operatorname{ad} x)_n \in \operatorname{Der} L$. But then by Lemma 8.3, $(\operatorname{ad} x)_s, (\operatorname{ad} x)_n \in \operatorname{ad} L$, hence $(\operatorname{ad} x)_s = \operatorname{ad} x_s$ for some $x_s \in L$ (which is consequently ad-semisimple) and similarly for $(\operatorname{ad} x)_n$. Then $\operatorname{ad}[x_s, x_n] = [\operatorname{ad} x_s, \operatorname{ad} x_n] = 0$, and since ad is 1-1 this implies $[x_s, x_n] = 0$. Finally, if $x = s+n$ were another abstract decomposition, by uniqueness of the normal Jordan decomposition we have $\operatorname{ad} x_s = \operatorname{ad} s$, so $x_s = s$ (similarly $x_n = n$).

(b): If W is a subspace of V stabilized by L, let $L_W = \{y \in \mathfrak{gl}(V) | y$ stabilizes W and $Tr(y\,|_W) = 0\}$. Since $L = [LL]$ (Corollary 6.11), and $Tr([xy]) = 0$ for any $x, y \in L$, L s a subspace of each L_W. Let L' be the intersection of all the L_W with $N = N_{\mathfrak{gl}(V)}(L)$. If $x \in L$ and $x = x_s + x_n$ is the normal Jordan decomposition of x, we know x_n and x_s must both stabilize W for any W stabilized by L. As x_n is nilpotent, $Tr(x_n\,|_W) = 0$ so $x_n \in L_W$, so we must have $x_s = x - x_n \in L_W$ as well. Furthermore, $\operatorname{ad} x = \operatorname{ad} x_s + \operatorname{ad} x_n$ is the Jordan decomposition of $\operatorname{ad} x$ in $\operatorname{ad} \mathfrak{gl}(V)$ (Lemma 5.2), and $\operatorname{ad} x$ maps L into L so $\operatorname{ad} x_n, \operatorname{ad} x_s$ must map L into L, (Proposition 5.1(b)) ie $x_n, x_s \in N$. This means $x_s, x_n \in L'$ for any x.

We now show $L = L'$. We know L acts on L' via ad, and since L is semisimple by Weyl's theorem we can write $L' = L + M$ where the sum is direct and M is a subspace of L' stabilized by L. But since $L' \subset N$, $\operatorname{ad} x(y) \in L$ for $x \in L$, $y \in L'$, and this implies the action of L on M is trivial, ie for every $y \in M$, $x \in L$, $[xy] = 0$. Now let W be subspace of V such that the action of L on W is irreducible. Schur's Lemma implies that any $y \in M$ acts on W as a scalar. But since $y \in M \subset L' \subset L_W$, $Tr(y\,|_W) = 0$. So y acts on W as zero. But by Weyl's Theorem, V can be written as a direct sum of subspaces V_i stabilized by L such that the action of L on V_i is irreducible. Then the action of y on each V_i is zero, and so $y = 0$. This implies $M = 0$, so $L = L'$.

Now we have, for $x \in L$, $x_n, x_s \in L$. But x_s (resp. x_n) is semisimple (resp. nilpotent) and therefore ad-semisimple (resp. ad-nilpotent), and $[x_s, x_n] = 0$, so by the uniqueness clause in (a) this must also be the abstract Jordan decomposition of x.

(c): Write $x = x_s + x_n$. L is spanned by eigenvectors of $\operatorname{ad}_L x_s$ (x_s being ad-semisimple) and if y is an eigenvector of $\operatorname{ad}_L x_s$ we have $[\phi(x_s), \phi(y)] = \phi([x_s, y]) = \phi(\alpha y) = \alpha\phi(y)$, ie $\phi(y)$ is an eigenvector of $\operatorname{ad}_{\phi(L)} \phi(x_s)$. This means $\phi(L)$ is spanned by eigenvectors of $\operatorname{ad}_{\phi(L)} \phi(x_s)$, and so $\phi(x_s)$ is ad-semisimple (in $\phi(L)$). Similarly, if $\operatorname{ad}_L x_n^k = 0$, $\operatorname{ad}_{\phi(L)} \phi(x_n)^k = 0$ so $\phi(x_n)$ is ad-nilpotent. Since $\phi(L)$ is semisimple (Corollary 6.11), we have $\phi(x) = \phi(x_s) + \phi(x_n)$ is the abstract Jordan decomposition of $\phi(x)$ in $\phi(L)$, hence by (b) also the normal Joradn decomposition of $\phi(x)$ in $\mathfrak{gl}(V)$.

□

Part 3. Construction of the Chevalley group

9. The algebra $\mathfrak{sl}(2, F)$

Given a field F, $\mathfrak{sl}(2, F)$ is the Lie algebra of 2-by-2 matrices with entries in F that have zero trace. This is a three dimensional vector space which has a standard basis:

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \; y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \; h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

It is a straightforward matter to compute the bracket product of pairs of basis vectors: $[hx] = 2x$, $[hy] = -2y$, $[xy] = h$.

Lemma 9.1. *If F is typical, $\mathfrak{sl}(2, F)$ is semisimple.*

Proof. Theorem 6.8 says that $\mathfrak{sl}(2, F)$ is semisimple if its Killing form is non-degenerate. The Killing form is non-degenerate if the 3-by-3 matrix whose i, j entry is $\kappa(x_i, x_j)$ ($x_1 = x$, $x_2 = h$, $x_3 = y$) has non-zero determinant. To compute $\kappa(x_i, x_j)$, first we compute the matrices of the adjoint representations:

$$\operatorname{ad} x = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \; \operatorname{ad} h = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \; \operatorname{ad} y = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

From here it is straightforward to compute the matrix of κ:

$$\begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix}$$

Which has determinant -128, hence κ is non-degenerate. □

Now let F be typical, and $\phi : \mathfrak{sl}(2, F) \to \mathfrak{gl}(V)$ be a finite dimensional representation of $\mathfrak{sl}(2, F)$. Since h is semisimple, $\phi(h)$ is semisimple by Theorem ??, and we can write V as a direct sum of eigenspaces: $V_\lambda = \{v \in V | \phi(h).v = \lambda v\}$, for $\lambda \in F$. Whenever $V_\lambda \neq 0$, we say λ is a **weight** of h in V and we call V_λ a **weight space**.

We have the following elementary lemma:

Lemma 9.2. *If $v \in V_\lambda$, then $\phi(x).v \in V_{\lambda+2}$ and $\phi(y).v \in V_{\lambda-2}$.*

Proof.

$$\begin{aligned} \phi(h).(\phi(x).v) &= [\phi(h), \phi(x)].v + \phi(x).(\phi(h).v) \\ &= \phi([hx]).v + \phi(x).(\lambda v) \\ &= 2\phi(x).v + \lambda\phi(x).v \\ &= (2 + \lambda)\phi(x).v \end{aligned}$$

The proof for y is identical. □

We are now in position to prove the main result of this section: a classification of irreducible representations of $\mathfrak{sl}(2, F)$:

Theorem 9.3. *Let F be typical, and $\phi : \mathfrak{sl}(2, F) \to \mathfrak{gl}(V)$ be an irreducible finite dimensional representation. Then there exists a basis $v_0, v_1, \ldots, v_m$ of V such that the following formulas (which completely determine ϕ) hold:*

(a) $\phi(h).v_i = (m - 2i)v_i$
(b) $\phi(y).v_i = (i + 1)v_{i+1}$ $(\phi(y).v_m = 0)$
(c) $\phi(x).v_i = (m - i + 1)v_{i-1}$ $(\phi(x).v_0 = 0)$

In particular, the weights of h in V are the integers $m, m - 2, \ldots, -(m - 2), -m$, and the weight spaces V_i are all one-dimensional.

Proof. Since V is finite-dimensional, there exists a maximal weight of h in V, call it λ. Now pick a nonzero $v_0 \in V_\lambda$, and for $i > 0$ write $v_i = (1/i!)\phi(y)^i.v_0$. We prove the following formulas:

(a) $\phi(h).v_i = (\lambda - 2i)v_i$
(b) $\phi(y).v_i = (i + 1)v_{i+1}$
(c) $\phi(x).v_i = (\lambda - i + 1)v_{i-1}$ $(i > 0)$

(a) follows from the above lemma. For (b), just use the definition of v_i. For (c), we make the following computation $(i > 0)$:

$$\begin{aligned} i\phi(x).v_i &= \phi(x).(\phi(y).v_{i-1}) \\ &= [\phi(x), \phi(y)].v_{i-1} + \phi(y).(\phi(x).v_{i-1}) \\ &= \phi([xy]).v_{i-1} + \phi(y).(\phi(x).v_{i-1}) \\ &= \phi(h).v_{i-1} + \phi(y).(\phi(x).v_{i-1}) \end{aligned}$$

Now we proceed by induction on i. Since λ is maximal, $\phi(x).v_0 \in V_{\lambda+2} = 0$, so by the above computation $\phi(x).v_1 = \phi(h).v_0 = \lambda v_0$ as desired (this is the base case). Now for $i > 1$ we have:

$$\begin{aligned} i\phi(x).v_i &= \phi(h).v_{i-1} + \phi(y).(\phi(x).v_{i-1}) \\ &= (\lambda - 2(i - 1))v_{i-1} + (\lambda - (i - 1) + 1)\phi(y).v_{i-2} \\ &= (\lambda - 2i + 2)v_{i-1} + (\lambda - i + 2)(i - 1)v_{i-1} \\ &= i(\lambda - i + 1)v_{i-1} \end{aligned}$$

And dividing by i gives the desired result. What does this tell us? First of all, the v_i span a vector subspace of V which is invariant under the action of $\mathfrak{sl}(2, F)$, and since ϕ is irreducible this means the v_i span V. Since by formula (a) $v_i \in V_{\lambda-2i}$ and V is the direct sum of weight spaces, the collection of nonzero v_i are linearly independent, so they form a basis (and therefore must be finite in number). Then we can find the largest integer m with $v_m \neq 0$, and it follows that $v_0, v_1, \ldots, v_m$ must all be nonzero and therefore form a basis of V. We showed that $\phi(x).v_0 = 0$, and formula (b) tells us $\phi(y).v_m = (m + 1)v_{m+1} = 0$.

So all that is left to show is $\lambda = m$. We use formula (c): $\phi(x).v_{m+1} = (\lambda - m)v_m$. But the left side is zero, hence the right side must be zero, but $v_m \neq 0$ so $\lambda - m = 0$. □

Corollary 9.4. *Let F be typical, and $\phi : \mathfrak{sl}(2,F) \to \mathfrak{gl}(V)$ be an irreducible finite dimensional representation. Then the eigenvalues of h on V are all integers, and dim V_n = dim V_{-n}. If we decompose V into a direct sum of subspaces such that the action on each subspace is irreducible (as in Weyl's theorem), the number of summands is dim V_0+ dim V_1.*

Proof. Use Weyl's theorem to write V as a direct sum of subspaces such that the action on each subspace is irreducible. The previous theorem completely describes the action of $\mathfrak{sl}(2,F)$ on each of these subspaces, which makes the first assertion clear. In addition, in each of these subspaces, either 0 is an eigenvalue or 1 is an eigenvalue but not both, which makes the second assertion clear. □

10. Root Space Decomposition

Engel's Theorem said that any finite dimensional Lie algebra consisting of ad-nilpotent elements is itself nilpotent (and vise-versa). Now we discuss algebras consisting of ad-semisimple elements:

Definition 10.1. If L is a finite-dimensional Lie algebra and T is a subalgebra of L, T is **toral** if it consists of ad-semisimple elements.

We then have the following lemma:

Lemma 10.2. *If T is a toral subalgebra of a finite-dimensional Lie algebra L over an algebraically closed field, T is abelian*

Proof. Pick $x \in T$. Since $\mathrm{ad}_L\, x$ is semisimple (therefore diagonalizable), so is $\mathrm{ad}_T\, x$. Pick a nonzero eigenvector y of $\mathrm{ad}_T\, x$, so $[xy] = ay$. We can find a basis of T which diagonalizes $\mathrm{ad}_T\, y : y_1, y_2, \ldots, y_n$, with eigenvalues $\alpha_1, \ldots, \alpha_n$. Then if $x = \beta_1 y_1 + \cdots + \beta_n y_n$, we have $(\mathrm{ad}_T\, y)^2(x) = \alpha_1^2 \beta_1 y_1 + \cdots + \alpha_n^2 \beta_n y_n$. On the other hand, $(\mathrm{ad}_T\, y)^2(x) = [y[yx]] = [y(-ay)] = -a[yy] = 0$. By linear independence, $\alpha_i^2 \beta_i = 0$ for all i, so $\alpha_i \beta_i = 0$ for all i, so $0 = \mathrm{ad}_T\, y(x) = [yx] = -ay$. Thus $a = 0$. Hence 0 is the only eigenvalue of $\mathrm{ad}_T\, x$, so $\mathrm{ad}_T\, x = 0$ for all $x \in T$. □

We are in need of a standard lemma of linear algebra:

Lemma 10.3. (Simultaneous Diagonalization) *Let W be a subspace of End(V), V a finite dimensional vector space over an algebraically closed field, consisting of commuting semisimple elements. Then we can find a basis of V that simultaneously diagonalizes every endomorphism in L.*

Proof. Work by induction on the dimension of W, dim $W = 1$ is obvious. If U is a subspace of W of codimension 1, we can form a basis $v_1, \ldots, v_n$ of V that simultaneously diagonalizes U. Now if $u \in U$, then $u(v_i)$ equals some scalar multiple of v_i, write $u(v_i) = \alpha(u)v_i$. Then α is a function from U to F, and it is easy to see that α is linear, that is $\alpha \in U^*$, the dual space of U. Now for $\alpha \in U^*$, let $V_\alpha = \{v \in V | u(v) = \alpha(u)v \text{ for all } u \in U\}$, and we can write V as a direct sum of the nonzero V_α.

Now if $v \in V_\alpha$, $u \in U$ and $w \in W - U$, by commutativity of W we have $u(w(v)) = w(u(v)) = w(\alpha(u)v) = \alpha(u)(w(v))$, hence $w(v) \in V_\alpha$. Then by proposition 3.3, $w\mid_{V_\alpha}$ is semisimple, so we can form a basis for V_α that diagonalizes w. Repeating this for each V_α we can construct a basis for V that diagonalizes w, but then it also diagonalizes each $u \in U$ (since each element of V_α is an eigenvector of each $u \in U$), so it diagonalizes $U + Fw = W$. □

Now for a typical semisimple Lie algebra L, pick a maximal toral subalgebra H. Since H is abelian, $\mathrm{ad}_L H$ consists of semisimple endomorphisms of L which commute. Following the above lemma, we can simultaneously diagonalize the elements of $\mathrm{ad}_L H$. In particular, we can consider subalgebras of the form $L_\alpha = \{x \in L | [hx] = \alpha(h)x \text{ for all } h \in H\}$, where $\alpha \in H^*$, and note that we can write L as a direct sum of the nonzero L_α. This is called the **root space decomposition** or **Cartan decomposition** of L. The set of nonzero $\alpha \in H^*$ with L_α nonzero is Φ, and the elements of Φ are called **roots** of L relative to H.

Proposition 10.4. *For all $\alpha, \beta \in H^*$, $[L_\alpha L_\beta] \subset L_{\alpha+\beta}$. If $x \in L_\alpha$, $\alpha \neq 0$, then $\mathrm{ad}\, x$ is nilpotent. If $\alpha, \beta \in H^*$ and $\alpha + \beta \neq 0$ then $\kappa(x, y) = 0$ for $x \in L_\alpha$, $y \in L_\beta$. (κ the Killing form of L)*

Proof. If $[xy]$ is a generator of $[L_\alpha L_\beta]$, we have $[h[xy]] = [[hx]y] + [x[hy]] = \alpha(h)[xy] + \beta(h)[xy] = (\alpha + \beta)(h)[xy]$ (this is the Jacobi identity). As for the second assertion, notice that L_α is nonzero for only finitely many α, on the other hand if $x \in L_\alpha$ and $y \in L_\beta$ the first assertion implies $(\mathrm{ad}\, x)^n(y) \in L_{n\alpha+\beta}$. Since $\alpha \neq 0$ and ch. $F = 0$, we can pick n big enough such that $L_{n\alpha+\beta}$ will be 0 for any $\beta \in \Phi$.

For the last assertion, find $h \in H$ such that $(\alpha + \beta)(h) \neq 0$. Then if $x \in L_\alpha$, $y \in L_\beta$, we have $\kappa([hx], y) = -\kappa([xh], y) = -\kappa(x, [hy])$, so $\alpha(h)\kappa(x, y) = -\beta(h)\kappa(x, y)$, so $(\alpha + \beta)(h)\kappa(x, y) = 0$. Thus $\kappa(x, y) = 0$. □

Corollary 10.5. *The restriction of the Killing form to $L_0 = C_L(H)$ is nondegenerate.*

Proof. If $x \in L_0$ is orthogonal to all $y \in L_0$ (that is, if $\kappa(x, y) = 0$), then by the proposition x is orthogonal to every element of L, hence x is contained in the radical of L. But since L is semisimple, theorem 5.10 says κ is nondegenerate. □

Proposition 10.6. $C_L(H) = H$

Proof. We prove this in seven steps:

Step 1:$C_L(H)$ contains the semisimple and nilpotent parts of its elements. If $x \in C_L(H)$, $\mathrm{ad}\, x$ maps the subspace H into the subspace 0. Then if $(\mathrm{ad}\, x)_s$ and $(\mathrm{ad}\, x)_n$ are the semisimple and nilpotent parts of $\mathrm{ad}\, x$, by Proposition 5.1(b) they both map H into 0. But by Theorem 8.4(c), $(\mathrm{ad}\, x)_s = \mathrm{ad}\, x_s$, hence $x_s \in C_L(H)$ and similarly for x_n.

Step 2: If $x \in C_L(H)$ is ad-semisimple (in L), $x \in H$. $H + Fx$ is a subalgebra of L (since x centralizes H) and since the sum of commuting semisimple elements is semisimple, $H + Fx$ is toral. By maximality of H, $H = H + Fx$ so $x \in H$.

Step 3: The restriction of κ to H is nondegenerate. Suppose $h \in H$ and $\kappa(h, x) = 0$ for all $x \in H$. Now pick $y \in C_L(H)$, $y = y_s + y_n$ the abstract Jordan decomposition, $y_n, y_s \in C_L(H)$ by step 1. Now $\mathrm{ad}\, h$ and $\mathrm{ad}\, y_n$ commute: $[\mathrm{ad}\, h, \mathrm{ad}\, y_n] = \mathrm{ad}\, [hy_n] = 0$, and since $\mathrm{ad}\, y_n$ is nilpotent, $\mathrm{ad}\, h\, \mathrm{ad}\, y_n$ is nilpotent.

So $0 = Tr(\operatorname{ad} h \operatorname{ad} y_n) = \kappa(h, y_n)$. But by step 2, $y_s \in H$ so $\kappa(h, y_s) = 0$. Hence $\kappa(h, y) = 0$ for all $y \in C_L(H)$, but the restriction of κ to $C_L(H)$ is nondegenerate (Corollary 10.5) so $h = 0$.

Step 4: $C_L(H)$ is nilpotent. Pick $x = x_n + x_s \in C_L(H)$. Then by step 1 $x_s, x_n \in C_L(H)$ and then by step 2 $x_s \in H$, so $\operatorname{ad}_{C_L(H)} x_s = 0$. So $\operatorname{ad}_{C_L(H)} x = \operatorname{ad}_{C_L(H)} x_n$ which is nilpotent since $\operatorname{ad}_L x_n$ is nilpotent. Then by Engel's Theorem, $C_L H$ is nilpotent.

Step 5: $H \cap [C_L(H), C_L(H)] = 0$. If $[xy]$ is a typical generator of $[C_L(H), C_L(H)]$ and $z \in H$, $\kappa([xy], z) = \kappa(x, [yz]) = \kappa(x, 0) = 0$. So $\kappa(x, z) = 0$ for $x \in [C_L(H), C_L(H)]$, $z \in H$. Now pick $h \in [C_L(H), C_L(H)] \cap H$. We have $\kappa(h, z) = 0$ for $z \in H$, and by step 3 this implies $h = 0$.

Step 6: $[C_L(H), C_L(H)] = 0$. By Lemma 4.9, if $[C_L(H), C_L(H)]$ is nonzero, since $C_L(H)$ is nilpotent (step 4), we can find $z \in Z(C_L(H)) \cap [C_L(H), C_L(H)]$, $z \neq 0$. By step 5, z is not in H, so by step 2 z is not ad-semisimple and hence has a nonzero nilpotent part z_n. Then since $\operatorname{ad} z$ maps $C_L(H)$ into 0, $(\operatorname{ad} z)_n = \operatorname{ad} z_n$ must also map $C_L(H)$ into 0, so $z_n \in Z(C_L(H))$. This implies for any $x \in C_L(H)$, $\operatorname{ad} z_n$ commutes with $\operatorname{ad} x$, so $\operatorname{ad} z_n \operatorname{ad} x$ is nilpotent, so $0 = Tr(\operatorname{ad} z_n \operatorname{ad} x) = \kappa(z_n, x)$. But this contracts the non-degeneracy of κ.

Step 7: $C_L(H) = H$. Otherwise $C_L(H)$ has a nonzero ad-nilpotent element x, by steps 1 and 2. For any $y \in C_L(H)$, $[\operatorname{ad} x, \operatorname{ad} y] = \operatorname{ad}[xy] = 0$ by step 6, so $\operatorname{ad} x \operatorname{ad} y$ is nilpotent and so $0 = Tr(\operatorname{ad} x \operatorname{ad} y) = \kappa(x, y)$. But this contradicts the non-degeneracy of κ.

□

Now since the restriction of κ to H is nondegenerate (step 3), the map $\lambda : H \to H^*$ defined by $\lambda : x \mapsto \lambda_x$ where $\lambda_x(y) = \kappa(x, y)$ is injective and therefore (H being finite-dimensional) bijective. Then for $\psi \in H^*$, let $t_\psi = \lambda^{-1}(\psi)$. Then $\kappa(t_\psi, h) = \lambda_{t_\psi}(h) = \psi(h)$. This property actually characterizes t_ψ: t_ψ is the unique element of H such that for all $h \in H$, $\kappa(t_\psi, h) = \psi(h)$. We can use this to define a symmetric bilinear product in H^*: $(\alpha, \beta) = \kappa(t_\alpha, t_\beta)$. (note that this also equals $\alpha(t_\beta)$ and $\beta(t_\alpha)$.)

11. Properties of the Root Space Decomposition

We are now in a position to introduce the connection to the Lie algebra $\mathfrak{sl}(2, F)$:

Proposition 11.1. *(a) Φ spans H^**
(b) If $\alpha \in \Phi$, $x \in L_\alpha$ nonzero, there exists $y \in L_{-\alpha}$ such that $\kappa(x, y) \neq 0$.
(c) If $\alpha \in \Phi$, $-\alpha \in \Phi$.
(d) If $\alpha \in \Phi$, $x \in L_\alpha$, $y \in L_{-\alpha}$, then $[xy] = \kappa(x, y) t_\alpha$.
(e) If $\alpha \in \Phi$, then $[L_\alpha, L_{-\alpha}]$ is nonzero and spanned by t_α.
(f) $(\alpha, \alpha) \neq 0$ for $\alpha \in \Phi$.
(g) If $\alpha \in \Phi$ and x_α is a nonzero element of L_α, there exists $y_\alpha \in L_{-\alpha}$ such that x_α, y_α, $h_\alpha = \frac{2t_\alpha}{(\alpha,\alpha)}$ span a three dimensional subalgebra of L isomorphic to $\mathfrak{sl}(2, F)$. (via $x_\alpha \mapsto x$, $y_\alpha \mapsto y$, $h_\alpha \mapsto h$.
(h) $h_\alpha = -h_{-\alpha}$.

Proof. (a): If not, pick $\psi \in H^*$ not in the span of Φ. Then there exists $\hat{h} \in H^{**}$ such that $\hat{h}(\psi) = 1$ but $\hat{h}(\alpha) = 0$ for $\alpha \in \Phi$. But by duality, there exists $h \in H$ such that $\alpha(h) = 0$ for all $\alpha \in \Phi$ and $\psi(h) = 1$ (so $h \neq 0$). Now for $x \in L_\alpha$, $[hx] = \alpha(h)x = 0$, and for $x \in L_0$, $[hx] = 0$. Since L is the direct sum of the L_α and L_0, $h \in Z(L) = 0$ (since L is semisimple), which is a contradiction.

(b): Let $\alpha \in \Phi$, and pick $x \in L_\alpha$, $x \neq 0$. Since the Killing form is nondegenerate and L is the direct sum of the root spaces, there exists $y \in L_\beta$ for some $\beta \in H^*$ such that $\kappa(x, y) \neq 0$. But by Proposition 10.4, if $\alpha + \beta \neq 0$ then $\kappa(x, y) = 0$, hence $\alpha + \beta = 0$, and $\beta = -\alpha$.

(c): If $\alpha \in \Phi$, pick $x \in L_\alpha$ nonzero. By part (b), there exists $y \in L_{-\alpha}$ such that $\kappa(x, y) \neq 0$, hence $y \neq 0$ and so $-\alpha \in \Phi$.

(d): Let $\alpha \in \Phi$, $x \in L_\alpha$, $y \in L_{-\alpha}$. Then for any $h \in H$ we have:

$$\begin{aligned}
\kappa(h, [xy]) &= \kappa([hx], y) \\
&= \alpha(h)\kappa(x, y) \\
&= \kappa(t_\alpha, h)\kappa(x, y) \\
&= \kappa(\kappa(x, y)t_\alpha, h) \\
&= \kappa(h, \kappa(x, y)t_\alpha)
\end{aligned}$$

So $\kappa(h, [xy] - \kappa(x, y)t_\alpha) = 0$ for all $h \in H$. Then $t_\alpha \in H$, and by Proposition 10.4 we have $[xy] \in L_0$ and by Proposition 10.6 $L_0 = C_L(H) = H$. Hence $[xy] - \kappa(x, y)t_\alpha \in H$, and by the nondegeneracy of the Killing form on H (step 3 of proposition 10.6) we have $[xy] - \kappa(x, y)t_\alpha = 0$.

(e): In light of (d), we only need to show $[L_\alpha L_{-\alpha}] \neq 0$. Pick $x \in L_\alpha$ nonzero, and using (b) pick $y \in L_{-\alpha}$ such that $\kappa(x, y) \neq 0$. Then by (d), $[xy] = \kappa(x, y)t_\alpha$, so $[xy] \neq 0$.

(f) Suppose $(\alpha, \alpha) = 0$ for some $\alpha \in \Phi$. Recall that $(\alpha, \alpha) = \kappa(t_\alpha, t_\alpha) = \alpha(t_\alpha)$. Using (b) pick $x \in L_\alpha$, $y \in L_{-\alpha}$ such that $\kappa(x, y) \neq 0$. Scaling x if necessary, we can assume $\kappa(x, y) = 1$. Then $[xy] = t_\alpha$ by (d), and $[t_\alpha x] = \alpha(t_\alpha)x = 0$, similarly $[t_\alpha y] = 0$, so x, y, t_α span a subspace S of L. Now suppose $\beta \in \Phi$ and pick $u \in L_\beta$. Then by the Jacobi identity, $\mathrm{ad}_L\, t_\alpha([xu]) = [x[t_\alpha u]] - [u[t_\alpha x]] = \beta(t_\alpha)[xu]$ and similarly for y, which implies S acts on L_β via ad.

But then $\mathrm{ad}\, t_\alpha$ acts diagonally on L_β with a constant eigenvalue $\beta(t_\alpha)$. However $\mathrm{ad}\, t_\alpha |_{L_\beta} = \mathrm{ad}\, [xy] |_{L_\beta} = [\, \mathrm{ad}\, x |_{L_\beta}, \mathrm{ad}\, y |_{L_\beta}]$, implying $\mathrm{ad}\, t_\alpha |_{L_\beta}$ has trace 0. So we must have $\beta(t_\alpha) = 0$ (F having characteristic 0). This holds for any $\beta \in \Phi$. But since Φ spans H^*, we must have $t_\alpha = 0$, contradicting the choice of t_α.

(g) Given nonzero $x_\alpha \in L_\alpha$, using (b) we can find $y_\alpha \in L_{-\alpha}$ such that $\kappa(x_\alpha, y_\alpha) \neq 0$. By part (f), scaling y_α if necessary, we can choose y_α so $\kappa(x_\alpha, y_\alpha) = \frac{2}{(\alpha,\alpha)}$. Using straightforward computation, the reader can verify that $[x_\alpha y_\alpha] = h_\alpha$ (using (d)),

$[h_\alpha x_\alpha] = \alpha(h_\alpha)x_\alpha = 2x_\alpha$ and $[h_\alpha y_\alpha] = -\alpha(h_\alpha)y_\alpha - 2y_\alpha$ (remember $\alpha(t_\alpha) = (\alpha, \alpha)$).

(h) We have $\kappa(t_\alpha + t_{-\alpha}, h) = \alpha(h) - \alpha(h) = 0$ for all $h \in H$, and by the nondegeneracy of κ we must have $t_\alpha = -t_{-\alpha}$. Then the assertion follows. □

Let us fix the definition of h_α as stated in the previous proposition. Furthermore, write $H_\alpha = [L_\alpha L_{-\alpha}]$. Now that we have a subalgebra isomorphic so $\mathfrak{sl}(2, F)$,we can use the results of the previous section for the following proposition.

Proposition 11.2. *(a) If $\alpha \in \Phi$, dim $L_\alpha = 1$. In particular, $S_\alpha = L_\alpha + L_{-\alpha} + H_\alpha$ (as vector spaces) is isomorphic to $\mathfrak{sl}(2, F)$, and for nonzero $x_\alpha \in L_\alpha$ there exists a unique $y_\alpha \in L_{-\alpha}$ satisfying $[x_\alpha y_\alpha] = h_\alpha$.*
(b) If $\alpha \in \Phi$, the only scalar multiples of α which are roots are α and $-\alpha$.
(c) If $\alpha, \beta \in \Phi$, then $\beta(h_\alpha) \in \mathbb{Z}$ and $\beta - \beta(h_\alpha)\alpha \in \Phi$ ($\mathbb{Z}$ is really the isomorphic copy of $\mathbb{Z}$ lying in F).
(d) Let $\alpha, \beta \in \Phi$, $\beta \neq \pm\alpha$. Let r, q be the largest integers for which $\beta - r\alpha$, $\beta + q\alpha$ are roots. Then $\beta + i\alpha \in \Phi$ for all $-r \leq i \leq q$, and $\beta(h_\alpha) = r - q$.
(e) If $\alpha, \beta, \alpha + \beta \in \Phi$ then $[L_\alpha L_\beta] = L_{\alpha+\beta}$.
(f) H is spanned by the h_α (hence L is generated as a Lie algebra by the L_α)

Proof. (a), (b): Suppose $\alpha \in \Phi$, and let S_α be the Lie algebra spanned by $x_\alpha, y_\alpha, h_\alpha$ as in the previous proposition, so S_α is isomorphic to $\mathfrak{sl}(2, F)$. Consider the subspace M of L spanned by H and root spaces of the form $L_{c\alpha}$ for $c \in F$ nonzero. By Proposition 10.4, S_α acts on M via ad. Now the weights of h_α on M are the integers 0 and $c\alpha(h_\alpha) = 2c$ where $c\alpha \in \Phi$.

Now (by duality) Ker α is a subspace of codimension 1 in H complementary to Fh_α. If $t \in$ Ker α, $[tx_\alpha] = \alpha(t)x_\alpha = 0$, $[ty_\alpha] = -\alpha(t)y_\alpha = 0$ and $[th_\alpha] = 0$ so S_α acts trivially on Ker α. Also, S_α acts on itself. But the elements which have weight zero are exactly those in H, which is contained in Ker $\alpha \oplus S_\alpha$. Now consider (by Weyl's theorem) breaking M up into subspaces M_i such that the action of S_α on M_i is irreducible. If M_i has 0 as a root, then either $M_i \subset$ Ker α (so S_α acts trivially on M_i) or $M_i = S_\alpha$. But by Theorem 9.3, each M_i has only even weights or only odd weights, this means the only even weights are $0, \pm 2$. Therefore, 2α (having weight 4) is not a root, so we have shown that twice a root is never a root. But then 1 cannot be a weight either (otherwise both $\alpha/2$ and α would be roots) and so $M =$ Ker $\alpha \oplus S_\alpha$. This means L_α must be spanned by x_α, $L_{-\alpha}$ by y_α, and $S_\alpha = L_\alpha + L_{-\alpha} + H_\alpha$ as asserted. Also, the only multiples of α which are roots are $\pm\alpha$.

(c), (d), (e): Now suppose $\beta \in \Phi$, $\beta \neq \pm\alpha$, and let K be the subspace of L spanned by $L_{\beta+i\alpha}$, where $i \in \mathbb{Z}$ and $\beta + i\alpha \in \Phi$. Then S_α acts on K via ad. Since $\beta + i\alpha \neq 0$ for any i (by (b)), K is a direct sum of one-dimensional subspaces $L_{\beta+i\alpha}$ having distinct integral weights $\beta(h_\alpha)+2i$. Then $\beta(h_\alpha) \in \mathbb{Z}$ (c). Furthermore, 0 and 1 cannot both be written in that form, meaning the action of S_α on K is irreducible. By Theorem 9.3, we conclude that each $\beta(h_\alpha) + 2i$ is a root between the maximum $\beta(h_\alpha) + 2q$ and minimum $\beta(h_\alpha) - 2r$, meaning $\beta + i\alpha \in \Phi$ for all $-r \leq i \leq q$. Furthermore, the highest and lowest roots are opposites: $\beta(h_\alpha) + 2q = -(\beta(h_\alpha) - 2r)$ so $\beta(h_\alpha) = r - q$. Then since $q \geq 0$, $r \geq \beta(h_\alpha)$ so $\beta - \beta(h_\alpha)\alpha \in \Phi$. Lastly, if $\beta + \alpha$ is a root then, since the action maps each weight space onto the adjacent weight spaces, the action maps L_β onto $L_{\beta+\alpha}$. This establishes (e).

(f): It is enough to show that the t_α span H. If not, there exists nonzero $\psi \in H^*$ such that $\psi(t_\alpha) = 0$ for all $\alpha \in \Phi$. But $\psi(t_\alpha) = \alpha(t_\psi)$, and since Φ span H^* this implies $t_\psi = 0$, hence $\psi = 0$, a contradiction.

□

Proposition 11.3. *Let $E_\mathbb{Q}$ be the $\mathbb{Q}$-subspace of H^* spanned by Φ, where H^* (ie, $E_\mathbb{Q}$ is rational linear combinations of Φ). Then any basis in Φ of H^* (over F) is also a basis of $E_\mathbb{Q}$ over $\mathbb{Q}$ in particular the dimension of H^* over F is the same as the dimension of $E_\mathbb{Q}$ over $\mathbb{Q}$. Furthermore, $(\alpha, \beta) = \kappa(t_\alpha, t_\beta)$ defines a nondegenerate, symmetric, bilinear form on $E_\mathbb{Q}$ which is positive definite: $(\lambda, \lambda) > 0$ for $\lambda \neq 0$.*

Proof. Pick a basis in Φ of H^*: $\alpha_1, \alpha_2, \ldots, \alpha_k$. To prove the first assertion, we need to show that any root $\beta \in \Phi$ is a rational linear combination of the α_i. Since the α_i span H^*, we have $\beta = \sum_{i=1}^k c_i\alpha_i$ where the $c_i \in F$. For $1 \leq j \leq k$, consider the equation $(\beta, \alpha_j) = \sum_{i=1}^k c_i(\alpha_i, \alpha_j)$. Multiplying by $2/(\alpha_j, \alpha_j)$ we get:

$$2\frac{(\beta, \alpha_j)}{(\alpha_j, \alpha_j)} = \sum_{i=1}^{k} \frac{2(\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)} c_i \text{ for all } 1 \leq j \leq k$$

Or: $\beta(h_{\alpha_j}) = \sum_{i=1}^k \alpha_i(h_{\alpha_j})c_i$. In view of Proposition 11.2(c), we have k equations with k unknowns and integral (importantly, rational) coefficients. Now since the α_i form a basis on H^* and the form is nondegenerate, the k by k matrix $A_{ij} = (\alpha_i, \alpha_j)$ is nonsingular, therefore the same holds for the coefficient matrix for this system. Hence the system has a solution in $\mathbb{Q}$, and that must be the unique solution.

For the second assertion, pick $h_1, h_2 \in H$. Since each of the L_α are one-dimensional, $\operatorname{ad} h_1$ (resp. h_2) is a diagonal matrix with one occurrence of $\alpha(h_1)$ for each $\alpha \in \Phi$ (and the rest of the entries are 0). It follows that $\kappa(h_1, h_2) = Tr(\operatorname{ad} h_1 \operatorname{ad} h_2) = \sum_{\alpha\in\Phi} \alpha(h_1)\alpha(h_2)$. So in particular for any $\lambda \in H^*$, $(\lambda, \lambda) = \kappa(t_\lambda, t_\lambda) = \sum_{\alpha\in\Phi} \alpha(t_\lambda)^2 = \sum_{\alpha\in\Phi}(\alpha, \lambda)^2$. This implies $(\lambda, \lambda) > 0$ unless $\lambda = 0$. It remains to show that $(\alpha, \beta) \in \mathbb{Q}$ for $\alpha, \beta \in \Phi$. Notice that $(\beta, \beta) = \sum_{\alpha\in\Phi}(\alpha, \beta)^2$. Dividing by $(\beta, \beta)^2$ we get $1/(\beta, \beta) = \sum_{\alpha\in\Phi}(\alpha, \beta)^2/(\beta, \beta)^2 = \sum_{\alpha\in\Phi} \alpha(h_\beta)^2/4$, and by Proposition 11.2(c) this means $(\beta, \beta) \in \mathbb{Q}$. Then $(\alpha, \beta) = (\beta, \beta)\alpha(h_\beta)/2$ is in $\mathbb{Q}$ as well.

□

12. Root Systems

If L is a typical semisimple Lie algebra and H is a maximal toral subalgebra, we showed in the previous section how to construct a vector space $E_\mathbb{Q} \subset H^*$ over $\mathbb{Q}$ with a positive definite symmetric bilinear form. It is a simple manner to extend the base field from $\mathbb{Q}$ to $\mathbb{R}$ to obtain a vector space E with a positive definite symmetric bilinear form, E is then called a **Euclidean space**. Then Φ is a subset of E which spans E (although note that E cannot necessarily be regarded as a subspace of H^*). The results from the previous section inspire the following definition:

Definition 12.1. If E is a Euclidean space and $\Phi \subset E$, Φ is called a **root system** if:

(R1) Φ is finite, spans E and does not contain 0.
(R2) The only multiples of $\alpha \in \Phi$ which lie in Φ are $\pm\alpha$.
(R3) If $\alpha, \beta \in \Phi$, $\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha \in \Phi$
(R4) If $\alpha, \beta \in \Phi$, $\frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}$

It is immediately clear from Propositions 11.1 and 11.2 that if (L, H) is a typical semisimple Lie algebra/maximal toral subalgebra pair, $\Phi \subset H^*$ is the set of roots and E is the Euclidean space described above, then Φ is a root system of E.

While the definition of a root system is motivated by the work done on root space decompositions, it is important to realize that the definition is purely geometric. That is, we can talk all about root systems without even mentioning Lie algebras once. This is what we plan to do for the rest of this section.

First we introduce some useful notation: Write $\langle\beta, \alpha\rangle$ for $2(\beta, \alpha)/(\alpha, \alpha)$, which if $\alpha, \beta \in \Phi$ is in $\mathbb{Z}$ by (R4). Now write $\sigma_\alpha(\beta) = \beta - \langle\beta, \alpha\rangle\alpha$. Then σ_α is a linear isomorphism of E of order 2, which geometrically can be viewed as reflection over the hyperplane perpendicular to α. If $\alpha, \beta \in \Phi$, $\sigma_\alpha(\beta) \in \Phi$ by (R3).

Lemma 12.2. *If $\alpha, \beta \in \Phi$, $\beta \neq \pm\alpha$ and $(\alpha, \beta) > 0$, $\alpha - \beta \in \Phi$.*

Proof. We have $\langle\beta, \alpha\rangle \cdot \langle\alpha, \beta\rangle = 4\frac{(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)}$. But the Cauchy-Schwartz inequality implies the fraction on the right is strictly less than 4 (strict since α, β not proportional), hence $\langle\beta, \alpha\rangle$ and $\langle\alpha, \beta$ are two positive integers whose product is less than 4. Therefore one of them must equal 1. If $\langle\beta, \alpha\rangle = 1$, then by (R3) $\sigma_\alpha(\beta) = \beta - \langle\beta, \alpha\rangle\alpha = \beta - \alpha \in \Phi$, so by (R1) $\alpha - \beta \in \Phi$. Similarly, if $\langle\alpha, \beta\rangle = 1$, $\sigma_\beta(\alpha) = \alpha - \beta \in \Phi$. □

We know that Φ spans E, so we can find a subset of Φ which is a basis of E. However, we are especially interested in a more special kind of subset of Φ:

Definition 12.3. A subset Δ of Φ is a **base** of Φ if:
(B1) Δ is a basis of E, and
(B2) Each root $\beta \in \Phi$ can be written $\sum_{\alpha\in\Delta} k_\alpha \alpha$ which each k_α is an integer, and they are either all nonnegative or all nonpositive.

Unfortunately, it is not obvious that such a set even exists. This is our next goal. First we prove a simple lemma:

Lemma 12.4. *Given a Euclidean space E and nonzero vectors $v_1, \dots, v_n$ in E, there exists $w \in E$ such that $(w, v_i) \neq 0$ for all i.*

Proof. Induct on n. If $n = 1$, we can simply choose $w = v_1$. If $n > 1$, by induction pick w_0 such that $(w_0, v_i) \neq 0$ for $1 \leq i \leq n - 1$. If $(w_0, v_n) \neq 0$ we are done. Otherwise, pick nonzero $a \in \mathbb{R}$ not equal to $(v_i, v_n)/(w_0, v_n)$ for $1 \leq i \leq n - 1$ (possible since $\mathbb{R}$ has infinitely many elements). Then setting $w = w_0 - av_n$, $(w, v_n) = -a(v_n, v_n) \neq 0$, and (w, v_i) is nonzero by our choice of a. □

Now if $\gamma \in E$ satisfies $(\gamma, \alpha) \neq 0$ for all $\alpha \in \Phi$, call γ **regular**. The lemma establishes the existence of regular elements of E. Now let $\Phi^+(\gamma) = \{\alpha \in \Phi | (\gamma, \alpha) > 0\}$ and $\Phi^-(\gamma) = \{\alpha \in \Phi | (\gamma, \alpha) < 0\}$. If γ is regular, call $\alpha \in \Phi^+(\gamma)$ **decomposable** if $\alpha = \beta_1 + \beta_2$, where $\beta_1, \beta_2 \in \Phi^+(\gamma)$. $\alpha \in \Phi^+(\gamma)$ is **indecomposable** otherwise. Finally, let $\Delta(\gamma)$ be the set of indecomposable roots. Then we have the following theorem:

Theorem 12.5. *$\Delta(\gamma)$ is a base of Φ, and any base of Φ can be written as $\Delta(\gamma)$ for some regular element γ.*

Proof. This is a proof in steps.

Step 1: If $\alpha, \beta \in \Delta(\gamma)$, $\alpha \neq \beta$, $(\alpha, \beta) \leq 0$. If $(\alpha, \beta) > 0$, then $\beta - \alpha$ and $\alpha - \beta$ are roots (Lemma 12.2). At least one of them is in $\Phi^+(\gamma)$, suppose $\alpha - \beta$. But then $\alpha = \beta + (\alpha - \beta)$, so α is decomposable, and in the other case β is decomposable. So $(\alpha, \beta) \leq 0$.

Step 2: If K is any finite subset of E such that $(\gamma, k) > 0$ for $k \in K$ and $(k_1, k_2) \leq 0$ for $k_1 \neq k_2$ in K, then K is a linearly independent set.

Suppose $\sum_i r_i k_i = 0$, and separate the positive coefficients from the negative coefficients to get $\sum s_i k_i = \sum t_j k_j = \varepsilon$, where each of the $s_i, t_j > 0$ and the k_i's and k_j's are distinct elements of K. Then $(\varepsilon, \varepsilon) = \sum_{i,j} s_i t_j (k_i, k_j) \leq 0$, meaning $\varepsilon = 0$. Now $0 = (\gamma, \varepsilon) = \sum_i s_i(\gamma, k_i)$ and since all the $(\gamma, k_i) > 0$ we know all the $s_i = 0$ (similarly, all $t_j = 0$).

Step 3: Any element in $\Phi^+(\gamma)$ can be written as a linear combination of elements in $\Delta(\gamma)$ where the coefficients are nonnegative integers

If not, pick $\alpha \in \Phi^+(\gamma)$ which cannot be written in this way such that (γ, α) is minimal. Then α is not in $\Delta(\gamma)$, so $\alpha = \beta_1 + \beta_2$, $\beta_1, \beta_2 \in \Phi^+(\gamma)$. Then $(\gamma, \alpha) = (\gamma, \beta_1) + (\gamma, \beta_2)$, and both the (γ, β_i) are positive and hence less than (γ, α). This means each of the β_i can be written as a linear combination in the desired way, hence so can α.

Step 4: $\Delta(\gamma)$ is a base.

By steps 1 and 2, $\Delta(\gamma)$ is linearly independent. Then if $\beta \in \Phi$, if $\beta \in \Phi^+(\gamma)$ we know β can be written as a linear combination of elements in $\Delta(\gamma)$ with non-negative integral coefficients (step 3). Then if $\beta \in \Phi^-(\gamma)$, $-\beta \in \Phi^+(\gamma)$ so β can be written as a linear combination with nonpositive integral coefficients. Since γ is regular, $\Phi = \Phi^+(\gamma) \cup \Phi^-(\gamma)$, so this establishes property (B2). Then the fact that $\Delta(\gamma)$ spans follows from the fact that Φ is a spanning set.

Step 5: If Δ is a base, there exists γ such that $(\gamma, \alpha) > 0$ for all $\alpha \in \Delta$.

Let $\Delta = \alpha_1, \ldots, \alpha_k$. Then for $1 \leq i \leq k$, we can find nonzero u_i such that $(\alpha_j, u_i) = 0$ for $i \neq j$: Just pick a vector orthogonal to the subspace spanned by the α_j with $j \neq i$. Then $(\alpha_i, u_i) \neq 0$ (this would force $u_i = 0$). Now let $\delta_i = \langle \alpha_i, u_i \rangle u_i$, and $\gamma = \sum_i \delta_i$. Then $(\alpha_i, \gamma) = \langle \alpha_i, u_i \rangle (\alpha_i, u_i) = \frac{(\alpha_i, u_i)^2}{(u_i, u_i)} > 0$.

Step 6: If Δ is a base, $\Delta = \Delta(\gamma)$ for some γ.

Pick γ such that $(\gamma, \alpha) > 0$ for all $\alpha \in \Delta$ (by step 5). Now if $\beta \in \Phi$, $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$ where (by (B2)) the k_α are all either nonnegative or nonpositive. In the first case $(\gamma, \beta) > 0$, in the second case $(\gamma, \beta) < 0$. So γ is regular. Then $\Delta \subset \Phi^+(\gamma)$. Suppose $\alpha \in \Delta$ is decomposable: $\alpha = \beta_1 + \beta_2$, each $\beta_i \in \Phi^+(\gamma)$.

Then both β_i can be written as a nonnegative integral linear combination of the elements in Δ, so the sum of the coefficients in each combination is at least one. That means α can be written as a nonnegative linear combination of the elements in Δ with the sum of the coefficients at least 2 (but can also be written $\alpha = 1 \cdot \alpha$, which contradicts the linear independence of Δ. That means each element in Δ is indecomposable, so $\Delta \subset \Delta(\gamma)$. But since each is a basis of the same vector space, $\Delta = \Delta(\gamma)$.

□

Now we have some simple lemmas about bases that will help us out in the next section:

Lemma 12.6. *If Δ is a base of Φ, then $(\alpha, \beta) \leq 0$ for $\alpha \neq \beta$ in Δ, and $\alpha - \beta$ is not a root.*

Proof. $\alpha - \beta$ cannot be a root, as this violates (B2) (the coefficients are not all nonnegative or all nonpositive). But by Lemma 12.2, if $(\alpha, \beta) > 0$, since $\alpha \neq \beta$ (and obviously $\alpha \neq -\beta$) we have $\alpha - \beta$ is a root. □

We say an element in Φ is **positive** (relative to a base Δ) if the coefficients (as in (B2)) are all nonnegative, and **negative** otherwise. Note that if $\Delta = \Delta(\gamma)$, then the positive elements are exactly $\Phi^+(\gamma)$, and the negative elements $\Phi^-(\gamma)$.

Lemma 12.7. *If α is a positive root (relative to Δ) and $\alpha \notin \Delta$, there exists $\beta \in \Delta$ such that $\alpha - \beta$ is a root (and $\alpha - \beta$ is necessarily positive).*

Proof. By Theorem 12.5, $\Delta = \Delta(\gamma)$ for some $\gamma \in E$. Then $(\gamma, \alpha) > 0$. If $(\alpha, \beta) \leq 0$ for all $\beta \in \Delta$, Step 2 of Theorem 12.5 would imply $\Delta \cup \{\alpha\}$ is a linearly independent set, which is a contradiction. So pick $\beta \in \Delta$ with $(\alpha, \beta) > 0$, Lemma 12.2 implies $\alpha - \beta \in \Phi$. Then $\alpha - \beta$ is positive, otherwise $\beta = (\beta - \alpha) + \alpha$ would be a decomposition of β would be a decomposition of β into a sum of positive roots. □

Corollary 12.8. *Each $\beta \in \Phi^+$ can be written in the form $\alpha_1 + \cdots + \alpha_k$, each $\alpha \in\in \Delta$ (not necessarily distinct), such that each partial sum $\alpha_1 + \cdots + \alpha_i$ is a root.*

Proof. Use the lemma and induct on the sum of the coefficients of β when written as a nonnegative integral linear combination of elements in Δ. □

Here we introduce the fairly straightforward idea of a root system isomorphism:

Definition 12.9. Φ and Φ' (in E, E' respectively) are said to be **isomorphic** if there exists $\phi : E \to E'$, a vector space isomorphism, such that $\langle \phi(\beta), \phi(\alpha) \rangle = \langle \beta, \alpha \rangle$ for each pair of roots β, α. ϕ is called a root system **isomorphism**.

The last major notion of root systems is the idea of reducibility:

Definition 12.10. A root system Φ is called **irreducible** if it cannot be partitioned into the union of two proper subsets such that each root in one is orthogonal to each root in the other.

A useful proposition follows:

Proposition 12.11. *If Δ is a base of Φ, Φ is irreducible iff Δ cannot be partitioned in the same way, ie into two proper subsets such that each root in one is orthogonal to each root in the other.*

Proof. First suppose $\Phi = \Phi_1 \cup \Phi_2$ is such a partition. Then $\Delta = (\Delta \cap \Phi_1) \cup (\Delta \cap \Phi_2)$. This is a valid partition unless either of those is empty. But if, say $\Delta \cap \Phi_1 = \emptyset$, then every root in Δ is orthogonal to every root in Φ_1, implying $\Phi_1 = \emptyset$.

Now suppose $\Delta = \Delta_1 \cup \Delta_2$ is such a partition. Let Φ_1 be the intersection of Φ with the span of Δ_1 and Φ_2 be the intersection of Φ with the span of Δ_2. Then clearly every root in Φ_1 is orthogonal to every root in Φ_2 and both these sets are nonempty, it remains to show that $\Phi = \Phi_1 \cup \Phi_2$. Suppose not. Pick $\beta \in \Phi - (\Phi_1 \cup \Phi_2)$, WLOG β is positive. By Corollary 12.8, write $\beta = \alpha_1 + \alpha_2 + \cdots + \alpha_k$ such that each partial sum is a root. WLOG, $\alpha_1 \in \Delta_1$. Since $\beta \notin \Phi_1$, there exists a first $\alpha_i \in \Delta_2$, call it γ. So $\alpha_1 + \alpha_2 + \cdots + \alpha_j + \gamma \in \Phi$, and $\alpha_i \in \Delta_1$ for $1 \leq i \leq j$. Call this sum δ.

Now consider $\sigma_\gamma(\delta)$. Since for $1 \leq i \leq j$, $(\alpha_i, \gamma) = 0$ we have $\sigma_\gamma(\alpha_i) = \alpha_i$. Also $\sigma_\gamma(\gamma) = -\gamma$. So $\sigma_\gamma(\delta) = \alpha_1 + \cdots + \alpha_j - \gamma$ is a root (R3). But since γ is not equal to any of the α_i, we have just written a root as a linear combination of roots in Δ with some coefficients positive and some negative, contradicting (B2). Hence $\Phi = \Phi_1 \cup \Phi_2$ as desired. □

We end with a useful feature of irreducible root systems. First, call $\beta \in \Phi$ **maximal** if β is positive and for any positive root α, $\beta + \alpha$ is not a root.

Theorem 12.12. *If Φ is an irreducible root system, there exists a unique maximal root.*

Proof. Existence of a maximal root is easy: simply start with a positive root α and keep adding positive roots to get new roots. You will never repeat a root (since the sum of positive roots can't be zero) so eventually you will reach a root and not be able to add a positive root. Then you will have a maximal root.

For uniqueness, let β be a maximal root, and write $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$. Let Δ_1 be the set of α with $k_\alpha > 0$, and Δ_2 be the set of α with $k_\alpha = 0$ (then $\Delta = \Delta_1 \cup \Delta_2$). Then for $\alpha \in \Delta_2$, $(\beta, \alpha) \leq 0$ (Lemma 12.6) and since Φ is irreducible $\alpha \in \Delta_2$ must be nonorthogonal to some $\alpha' \in \Delta_1$, hence $(\alpha, \alpha') < 0$. This forces $(\alpha, \beta) < 0$, so $\alpha + \beta$ is a root (Lemma 12.2) which is a contradiction. Hence Δ_2 is empty and all $k_\alpha > 0$ for all α. We also $(\beta, \alpha) \geq 0$ for all $\alpha \in \Delta$.

Now let β' be another maximal root. The same facts must be true of β', and there must be at least one $\alpha \in \Delta$ for which $(\alpha, \beta) > 0$, hence $(\beta, \beta') > 0$. Then (Lemma 12.2) $\beta - \beta'$ is a root. WLOG it is positive (otherwise use $\beta' - \beta$), but then $\beta = (\beta - \beta') + \beta'$ which contradicts the maximality of β'. □

13. Isomorphism Theorem

Suppose L is a typical semisimple Lie algebra, H a maximal toral subalgebra and $\Phi \subset H^*$ the set of roots. Then we showed that we can consider Φ as a subset of a Euclidean space E, by first considering the rational span of Φ in H^* and then extending the base field from $\mathbb{Q}$ to $\mathbb{R}$. Then Φ is a root system in E.

One might ask whether, if two semisimple typical Lie algebra/maximal toral subalgebra pairs yield isomorphic root systems, are the original Lie algebras/toral subalgebras isomorphic? The answer is, in fact, yes. That is the goal of this section.

For the rest of this section, all the terminology is as before: L is a typical semisimple Lie algebra, H is a maximal toral subalgebra, Φ the set of nonzero roots of L, $L = H + \sum_{\alpha\in\Phi} L_\alpha$ the root space decomposition. For $\alpha \in H^*$, $t_\alpha \in H$ is defined such that for all $h \in H$, $\kappa(t_\alpha, h) = \alpha(h)$. Then for $\alpha \in \Phi$, we define $h_\alpha = \frac{2t_\alpha}{(\alpha,\alpha)}$, where the inner product on H^* is defined by $(\alpha, \beta) = \kappa(t_\alpha, t_\beta)$. We showed that if $x_\alpha \in L_\alpha \neq 0$ there exists a unique $y_\alpha \in L_{-\alpha}$ such that $[x_\alpha y_\alpha] = h_\alpha$.

Proposition 13.1. *If Δ is a base of Φ, L is generated (as a Lie algebra) by arbitrary nonzero root vectors $x_\alpha \in L_\alpha$, $y_\alpha \in L_{-\alpha}$.*

Proof. Let β be an arbitrary positive root (relative to Δ). By Corollary 12.8, β can be written as $\alpha_1 + \cdots + \alpha_s$ where each partial sum is also a root and all $\alpha_i \in \Delta$. We also know that if $\gamma, \delta \in \Phi$ and $\gamma + \delta \in \Phi$, $[L_\gamma L_\delta] = L_{\gamma+\delta}$ (Proposition 11.2(e)). Then by induction on s we see that L_β must lie in any subalgebra of L containing each L_α for $\alpha \in \Delta$. Similarly if β is negative, L_β lies in any subalgebra of L containing each $L_{-\alpha}$ ($\alpha \in \Delta$). Finally, each $[x_\alpha y_\alpha]$ is some nonzero multiple of h_α, and the h_α span H (Proposition 11.2(f)). This proves the proposition. □

This leads naturally into the following definition:

Definition 13.2. If Δ is a base of Φ, $0 \neq x_\alpha \in L_\alpha$, $0 \neq y_\alpha \in L_{-\alpha}$ for $\alpha \in \Delta$, and $[x_\alpha, y_\alpha] = h_\alpha$, we call the set $\{x_\alpha, y_\alpha | \alpha \in \Delta\}$ a **standard set of generators** for L.

Now we can relate the notion of irreducible root systems to simple Lie algebras:

Proposition 13.3. *If L is simple, Φ is an irreducible root system.*

Proof. Suppose $\Phi = \Phi_1 \cap \Phi_2$ is a partition of Φ into nonempty orthogonal components. Consider the subalgebra K of L generated by all the L_α for $\alpha \in \Phi_1$. Then for $\beta \in \Phi_2$, $(\alpha + \beta, \alpha) \neq 0$ and $(\alpha + \beta, \beta) \neq 0$ so $\alpha + \beta$ cannot be a root. Hence $[L_\alpha L_\beta] = 0$, so the L_β centralize K. Since $Z(L) = 0$, K cannot be all of L. Furthermore, the L_α for $\alpha \in \Phi_1$ must normalize K, and therefore all L_α for $\alpha \in \Phi$, and therefore all of L (Proposition 13.1). □

Now we state a useful theorem that allows us to restrict our attention to simple Lie algebras:

Theorem 13.4. *Let L be a semisimple typical Lie algebra with maximal toral subalgebra H and root system Φ. If $L = L_1 \oplus L_2 \oplus \cdots \oplus L_n$ is a the decomposition of L into simple ideals, and $H_i = L_i \cap H$, then H_i is a maximal toral subalgebra of L_i with (irreducible) root system Φ_i. Then the Φ_i can be thought of as subsets of Φ such that $\Phi = \Phi_1 \cup \cdots \cup \Phi_n$ is the decomposition of Φ into irreducible components.*

Proof. Each H_i is toral in L_i. Suppose T were a larger toral algebra. Then since any element in L_i acts (via ad) trivially on any L_j with $j \neq i$, T is toral in L. But then the direct sum of T and the H_j with $j \neq i$ would be a toral subalgebra of L containing H (since T centralizes each of the H_j).

Now if $\alpha \in \Phi_i$ (so $\alpha \in H_i^*$) we can think of $\alpha \in H^*$ by simply saying $\alpha(H_j) = 0$ for $j \neq i$. Then α is clearly a root of L relative to H, and $L_\alpha \subset L_i$. On the other hand, if $\alpha \in \Phi$, since $[HL_\alpha] \neq 0$ we must have $[H_i L_\alpha] \neq 0$ for some i, and then $L_\alpha \subset L_i$ so $\alpha\,|_{H_i}$ is a root of L_i. □

Theorem 13.5. *Let L, L' be simple typical Lie algebras with respective maximal toral subalgebras H, H' and root systems Φ, Φ'. Suppose $\phi : \Phi \to \Phi'$ is a root system isomorphism. Furthermore, pick a base $\Delta \subset \Phi$, so $\Delta' = \phi(\Delta) \subset \Phi'$ is a base of Φ'. For each $\alpha \in \Delta$, choose arbitrary nonzero $x_\alpha, x_{\phi(\alpha)}$ in $L_\alpha, L_{\phi(\alpha)}$ respectively. Then there exists a unique isomorphism $\pi : L \to L'$ such that $\pi(h_\alpha) = h_{\phi(\alpha)}$ and $\pi(x_\alpha) = x_{\phi(\alpha)}$.*

Now we prove the isomorphism theorem, which essentially says that isomorphic irreducible root systems arise from isomorphic Lie algebra/maximal toral subalgebra pairs:

Proof. Let $\alpha \in \Delta$ and $\alpha' = \phi(\alpha) \in \Delta'$. Then there exist unique $y_\alpha \in L_{-\alpha}$, $y_{\alpha'} \in L_{-\alpha'}$ such that $[x_\alpha y_\alpha] = h_\alpha$ and $[x_{\alpha'} y_{\alpha'}] = h_{\alpha'}$ But if such a π exists, we would have $[x_{\alpha'}, \pi(y_\alpha)] = [\pi(x_\alpha), \pi(y_\alpha)] = \pi([x_\alpha y_\alpha]) = \pi(h_\alpha) = h_{\alpha'}$, hence $\pi(y_\alpha) = y_{\alpha'}$. Then since these x_α, y_α generate L and the value of π on these elements is completely determined, π is unique.

Now we proceed with existence. Consider $L \oplus L'$ be the direct sum of L and L', it is therefore a typical semisimple Lie algebra with unique simple ideals L, L'. Let D be the subalgebra generated by the elements $\overline{x}_\alpha = (x_\alpha, x_{\alpha'})$, $\overline{y}_\alpha = (y_\alpha, y_{\alpha'})$, $\overline{h}_\alpha = (h_\alpha, h_{\alpha'})$.

Now since L and L' are simple, Φ and Φ' are irreducible and so have unique maximal roots β, β' which must be mapped to each other by ϕ. Choose $x \in L_\beta$, $x' \in L_{\beta'}$ nonzero, and set $\overline{x} = (x, x') \in K$. Let M be the subspace of $L \oplus L'$ spanned by all $\operatorname{ad} \overline{y}_{\alpha_1} \operatorname{ad} \overline{y}_{\alpha_2} \cdots \operatorname{ad} \overline{y}_{\alpha_m}(\overline{x})$ where $\alpha_i \in \Delta$ (not necessarily distinct). Now consider the L "coordinate" of this expression: We have $y_{\alpha_m} \in L_{-\alpha_m}$ hence $\operatorname{ad} y_{\alpha_m}(x) \in L_{\beta - \alpha_m}$ (??). Continuing this, we see that $\operatorname{ad} y_{\alpha_1} \operatorname{ad} y_{\alpha_2} \cdots \operatorname{ad} y_{\alpha_m}(x) \in L_{\beta - \sum_i \alpha_i}$. Repeating this logic for L', we see the original expression lies in $L_{\beta - \sum_i \alpha_i} \oplus L'_{\beta - \sum_i \alpha_i}$. It follows that $M \cap (L_\beta \oplus L'_\beta)$ is only one dimensional (and so M is not all of $L \oplus L'$).

We now claim that the action of D on $L \oplus L'$ via ad stabilizes M. By definition, $\operatorname{ad} \overline{y}_\alpha$ stabilizes M for $\alpha \in \Delta$. For $\overline{h}_\alpha$ we proceed by induction on m in the expression $\operatorname{ad} \overline{y}_{\alpha_1} \operatorname{ad} \overline{y}_{\alpha_2} \cdots \operatorname{ad} \overline{y}_{\alpha_m}(\overline{x})$. First note that $\operatorname{ad} \overline{h}_\alpha(\overline{x}) = 2\overline{x}$ (this is the case $m = 0$. Now for the case $m > 0$, write the expression as $\operatorname{ad} \overline{y}_{\alpha_m}(\overline{u})$ where by induction $\operatorname{ad} \overline{h}_\alpha(\overline{u}) \in M$. Then:

$$\begin{aligned}
\operatorname{ad} \overline{h}_\alpha \operatorname{ad} \overline{y}_{\alpha_m}(\overline{u}) &= [\overline{h}_\alpha[\overline{y}_{\alpha_m}\overline{u}]] \\
&= [[\overline{h}_\alpha \overline{y}_{\alpha_m}]\overline{u}] + [\overline{y}_{\alpha_m}[\overline{h}_\alpha \overline{u}]] \\
&= -2[\overline{y}_{\alpha_m}\overline{u}] + [\overline{y}_{\alpha_m}[\overline{h}_\alpha \overline{u}]]
\end{aligned}$$

Which is clearly in M. Similarly for $\overline{x}_\alpha$: $\operatorname{ad}\overline{x}_\alpha(\overline{x}) = 0$ since it is an element of $L_{\alpha+\beta}$ which is 0 by maximality of β (this is the case $m = 0$). Now for the case $m = 0$ we write $\overline{u}$ as above. Then:

$$\begin{aligned}\operatorname{ad}\overline{x}_\alpha \operatorname{ad}\overline{y}_{\alpha_m}(\overline{u}) &= [\overline{x}_\alpha[\overline{y}_{\alpha_m}\overline{u}]] \\ &= [[\overline{x}_\alpha\overline{y}_{\alpha_m}]\overline{u}] + [\overline{y}_{\alpha_m}[\overline{x}_\alpha\overline{u}]]\end{aligned}$$

Now in the last expression, the term on the right is in M by induction. For the term on the right, if $\alpha \neq \alpha_m$ we have $[\overline{x}_\alpha\overline{y}_{\alpha_m}] \in L_{\alpha-\alpha_m}$, but by (??), since $\alpha, \alpha_m \in \Delta$, $\alpha - \alpha_m$ is not a root, so $[\overline{x}_\alpha\overline{y}_{\alpha_m}] = 0$. If $\alpha = \alpha_m$, then we get $[\overline{x}_\alpha\overline{y}_\alpha] = \overline{h}_\alpha$, and we have already established that $\overline{h}_\alpha$ stabilizes M.

Then if $D = L \oplus L'$, since D stabilizes M we have M is a proper nonzero ideal of $L \oplus L'$. But by the simplicity of L and L', the only such ideals are L and L', but clearly $M \neq L$ and $M \neq L'$.

Now consider the projections of D onto its first and second coordinates, π_1 and π_2. The projections are Lie algebra homomorphisms, and onto since the x_α and y_α generate L. Now let $I = \operatorname{Ker} \pi_2$, ie elements in D whose second coordinate is 0. Since π_1 is onto, $\pi_1(I)$ is an ideal of L. If $\pi_1(I)$ is nonzero, by simplicity it equals L, which means D contains all elements of the form $(x, 0)$ for $x \in L$. But then D contains $(x_\alpha, 0)$ for all $\alpha \in \Delta$, so $\overline{x} - (x_\alpha, 0) = (0, x_{\alpha'}) \in D$. Similarly, $(0, y_{\alpha'}) \in D$, and so (since these elements generate L') D contains all elements of the form $(0, x)$ with $x \in L'$. Then we have $D = L \oplus L'$ which we just showed could not be the case. That means $I = 0$, so π_2 is an isomorphism (similarly π_1 is an isomorphism).

Now the isomorphism $L \to L'$ obtained by D is exactly the isomorphism we want to show exists. □

We end with a straightforward result using the isomorphism theorem:

Proposition 13.6. *L, H, Φ, etc. as stated at the beginning of this section. Fix a base Δ of Φ and pick $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$ with $[x_\alpha y_\alpha] = h_\alpha$ for $\alpha \in \Delta$. Then there is an automorphism σ of order 2 satisfying $\sigma(x_\alpha) = -y_\alpha$, $\sigma(y_\alpha) = -x_\alpha$, $\sigma(h) = -h$.*

Proof. Let L_i be a simple subalgebra of L with maximal toral subalgebra H_i and root system Φ_i. Then the map sending Φ_i to $-\Phi_i$ is a root system isomorphism inducing $\pi : H_i \to H_i$ which sends h to $-h$. Furthermore for each α with $L_\alpha \subset L_i$, we can say that x_α is sent to $-y_\alpha$, then we have an automorphism $\sigma_i : L_i \to L_i$ by Theorem 13.3. Then this allows us to define σ on all of L and it must satisfy the desired properties. □

14. Construction of a Chevalley Group

We are in the same boat as previous sections: L a semisimple Lie algebra, H a maximal toral subalgebra, Φ the root system, and all other notions as previously established.

Proposition 14.1. *Let $\alpha, \beta \in \Phi$ be linearly independent roots. Suppose r is the greatest integer such that $\beta - r\alpha$ is a root, and q is the greatest integer with $\beta + q\alpha$ a root (call $\beta - r\alpha, \ldots, \beta + q\alpha$ the $\alpha-$string through β. Then:*

(a) $\langle \beta, \alpha \rangle = r - q$
(b) If $\alpha + \beta \in \Phi$*, then* $r + 1 = \frac{q(\alpha+\beta,\alpha+\beta)}{\beta,\beta}$.

Proof. (a): This is 11.2(d).
(b): Consult [1] (page 146). □

We have a simple lemma:

Lemma 14.2. *Let* α, β *be linearly independent roots. Choose* $x_\alpha \in L_\alpha$*,* $y_\alpha \in L_{-\alpha}$ *for which* $[x_\alpha y_\alpha] = h_\alpha$ *and let* $x_\beta \in L_\beta$ *be arbitrary. Then if* $\beta - r\alpha, \dots, \beta + q\alpha$ *is the* $\alpha-$*string through* β*,* $[y_\alpha[x_\alpha x_\beta]] = q(r+1)x_\beta$.

Proof. If $\alpha + \beta \notin \Phi$ then $q = 0$ and $[x_\alpha x_\beta] = 0$ so both sides are 0. Otherwise we can consider S_α acting on $L_{\beta - r\alpha} \oplus \cdots \oplus L_{\beta+q\alpha}$. Then (using the notation as in section 9) the highest weight is $r+q$ and x_β is a nonzero multiple of v_q, so applying x_α followed by y_α results in $q(r+1)x_\beta$. □

Now our task is to construct a **Chevalley basis** of L: A set $\{x_\alpha, \alpha \in \Phi\} \cup \{h_\alpha, \alpha \in \Delta\}$ where Δ is some base of Φ, and whenever $\alpha, \beta, \alpha + \beta \in \Phi$ with $[x_\alpha x_\beta] = c_{\alpha,\beta} x_{\alpha+\beta}$, $c_{\alpha,\beta} = -c_{-\alpha,-\beta}$.

Proposition 14.3. *There exists a Chevalley basis.*

Proof. Using Proposition 13.6, take the automorphism σ which sends L_α to $L_{-\alpha}$ and acts on H by multiplication by -1. If $x_\alpha \in L_\alpha$, write $x_{-\alpha} = -\sigma(x_\alpha)$, which is nonzero. Then $\kappa(x_\alpha, x_{-\alpha}) \neq 0$ (proposition 11.1(b)). By scaling x_α if necessary, we can make $\kappa(x_\alpha, x_{-\alpha}) = \frac{2}{(\alpha,\alpha)}$. Then $[x_\alpha x_{-\alpha}] = h_\alpha$ (proposition 11.1(d)). For each pair of roots $\{\alpha, -\alpha\}$ we fix a pair $x_\alpha, x_{-\alpha}$ satisfying this relation.

Now let $\alpha, \beta, \alpha + \beta \in \Phi$ so $[x_\alpha x_\beta] = c_{\alpha,\beta}$ for some $c_{\alpha,\beta} \in F$. Applying σ, $[-x_{-\alpha}, -x_{-\beta}] = -c_{\alpha,\beta} x_{-\alpha-\beta}$. So this choice of $x_\alpha, x_{-\alpha}$ satisfy the conditions of being a Chevalley basis. □

The important fact about a Chevalley basis is the following theorem:

Theorem 14.4. *Let* $\{x_\alpha, \alpha \in \Phi\ h_\alpha, \alpha \in \Delta\}$ *be a Chevalley basis of* L*. Then the structure constants lie in* $\mathbb{Z}$*:*
(a) $[h_\alpha h_\beta] = 0$
(b) $[h_\alpha x_\beta] = \langle \beta, \alpha \rangle x_\beta$
(c) $[x_\beta x_{-\beta}] = h_\beta$ *is a linear combination of* $h_\alpha, \alpha \in \Delta$ *where the coefficients are integers.*
(d) α, β *linearly independent roots,* $\beta - r\alpha, \dots, \beta + q\alpha$ *the* $\alpha-$*string through* β*, then* $[x_\alpha x_\beta] = 0$ *if* $q = 0$ *or* $\pm(r+1)x_{\alpha+\beta}$ *otherwise.*

Proof. (a) is clear, since H is abelian. (b) follows from the fact that $\beta(h_\alpha) = \langle \beta, \alpha \rangle$.

(c): For $\alpha \in \Phi$, consider $\alpha^v = \frac{2\alpha}{(\alpha,\alpha)}$. Then:

$$\langle \beta^v, \alpha^v \rangle = 2\frac{(\alpha^v, \beta^v)}{(\alpha^v, \alpha^v)} = 2\frac{(\alpha, \beta)}{(\beta, \beta)} = \langle \alpha, \beta \rangle$$

.

Also, for $\alpha \in \Phi$, recall σ_α is an isometry, so:

$$\begin{aligned}\sigma_{\beta^v}(\alpha^v) = \alpha^v - \langle \alpha^v, \beta^v \rangle \beta^v &= \frac{2\alpha}{(\alpha,\alpha)} - \langle \beta, \alpha \rangle \frac{2\beta}{(\beta,\beta)} \\ &= \frac{2(\alpha - \langle \alpha, \beta \rangle \beta)}{(\alpha,\alpha)} \\ &= \frac{2\sigma_\beta(\alpha)}{(\sigma_\beta(\alpha), \sigma_\beta(\alpha))} \\ &= (\sigma_\beta(\alpha))^v\end{aligned}$$

This means that, if we let $\Phi^v = \{\alpha^v | \alpha \in \Phi\}$, Φ^v is a root system (called the system **dual** to Φ). We will show that if Δ is a base of Φ, Δ^v is a base of Φ^v. We know $\Delta = \Delta(\gamma)$ for some γ, and $(\alpha, \gamma) > 0$ iff $(\alpha^v, \gamma) > 0$, so consider $\Delta^v(\gamma)$, the nondecomposable elements of Φ^v relative to gamma. We will show $\Delta^v \subset \Delta^v(\gamma)$, and considering cardinalities $\Delta^v = \Delta^v(\gamma)$ is a base.

Suppose $\alpha \in \Delta$, but α^v is decomposable, ie $\alpha^v = \beta_1^v + \beta_2^v$ where β_i^v are positive. So β_i are positive, and we have $\alpha = \frac{(\alpha,\alpha)}{(\beta_1,\beta_1)}\beta_1 + \frac{(\alpha,\alpha)}{(\beta_2,\beta_2)}\beta_2$. But then both β_i are linear combinations of elements of Δ with nonnegative coefficients, meaning each β_i can only have a positive coefficient associated with α. But then each β_i is a multiple of α, meaning $\beta_1 = \beta_2 = \alpha$. But this gives us $\alpha^v = \alpha^v + \alpha^v$, which is absurd.

Now note that the linear map $\lambda : H^* \to H$ sending α to t_α sends α^v to h_α. We have showed that if Δ is a base of Φ, Δ^v is a base of Φ^v, ie each β^v can be written as an integral linear combination of elements $\alpha^v \in \Delta^v$. Applying λ, we see that each h_β can be written as an integral linear combination of elements h_α, $\alpha \in \Delta$.

(d): If $q = 0$ then $\alpha + \beta$ is not a root and the result is clear. Otherwise, since $t_{\alpha+\beta} = t_\alpha + t_\beta$, we get:

$$[c_{\alpha,\beta}x_{\alpha+\beta}, c_{\alpha,\beta}x_{-\alpha-\beta}] = c_{\alpha,\beta}^2 h_{\alpha+\beta} = \frac{2c_{\alpha,\beta}^2}{(\alpha+\beta, \alpha+\beta)}(t_\alpha + t_\beta)$$

But the left side also equals $-[[x_\alpha x_\beta][x_{-\alpha}x_{-\beta}]] = [x_\alpha[x_\beta[x_{-\beta}x_{-\alpha}]]] + [x_\beta[x_\alpha[x_{-\alpha}x_{-\beta}]]]$. If the $\beta-$string through α is $\alpha - r'\beta, \dots, \alpha + q'\beta$, then we can use Lemma 14.2 to compute this (note that replacing α with $-\alpha$, β with $-\beta$ does not change q, q', r, r'):

$$\begin{aligned}[x_\alpha[x_\beta[x_{-\beta}x_{-\alpha}]]] + [x_\beta[x_\alpha[x_{-\alpha}x_{-\beta}]]] &= q'(r'+1)[x_\alpha x_{-\alpha}] + q(r+1)[x_\beta x_{-\beta}] \\ &= \frac{2q'(r'+1)}{(\alpha,\alpha)}t_\alpha + \frac{2q(r+1)}{(\beta,\beta)}t_\beta\end{aligned}$$

Now we can compare the t_β coefficients (since the t_α, t_β are linearly independent) and use Proposition 4.1(c) to get:

$$c_{\alpha,\beta}^2 = \frac{q(r+1)}{(\beta,\beta)}(\alpha+\beta, \alpha+\beta) = (r+1)^2$$

And the result follows. □

We can now consider, for a semisimple Lie algebra L with Chevalley basis $\{x_\alpha, \alpha \in \Phi; h_\alpha, \alpha \in \Delta\}$, the set $L(\mathbb{Z})$, the integral span of the basis elements. We have just showed that $L(\mathbb{Z})$ is closed under the bracket product. We are on the verge of being able to construct Chevalley groups. First a proposition:

Proposition 14.5. *Let $\alpha \in \Phi$, $m \in \mathbb{Z}^+$. Then $(\operatorname{ad} x_\alpha)^m/m!$ leaves $L(\mathbb{Z})$ invariant.*

Proof. It is enough to consider the action on the elements of the Chevalley basis. We have $(\operatorname{ad} x_\alpha)(h_\beta) = [x_\alpha h_\beta] = -\langle \alpha, \beta \rangle x_\alpha \in L(\mathbb{Z})$, and for $m > 1$ we have $(\operatorname{ad} x_\alpha)^m/m!(h_\beta) = 0$. Also, $\operatorname{ad} x_\alpha(x_\alpha) = 0$. Now $(\operatorname{ad} x_\alpha)(x_{-\alpha}) = h_\alpha \in L(\mathbb{Z})$, $(\operatorname{ad} x_\alpha)^2/2(x_{-\alpha}) = \frac{1}{2}[x_\alpha h_\alpha] = -x_\alpha \in L(\mathbb{Z})$, and $(\operatorname{ad} x_\alpha)^m/m!(x_{-\alpha}) = 0$ for $m > 2$. Finally consider $(\operatorname{ad} x_\alpha)^m/m!(x_\beta)$ where β is not $\pm\alpha$. We know $[x_\alpha x_\beta] = \pm(r+1)$, and the numbers which play the role of r for $\beta, \beta+\alpha, \ldots, \beta+(m-1)\alpha$ are $r, r+1, \ldots, r+(m-1)$, if each of these are roots. Then $(\operatorname{ad} x_\alpha)^m/m!(x_\beta) = \pm\frac{(r+1)(r+2)\cdots(r+m)}{m!} x_{\beta+m\alpha} = \binom{m+r}{m} x_{\beta+m\alpha} \in L(\mathbb{Z})$. If $\beta + m\alpha$ is not a root, then the expression is just equal to zero. □

Now, each x_α being nilpotent, we can consider the endomorphisms $\exp\ \operatorname{ad} x_\alpha = 1 + \operatorname{ad} x_\alpha = (\operatorname{ad} x_\alpha)^2/2! + \cdots + (\operatorname{ad} x_\alpha)^m/m! + \cdots$, which is a finite sum. (This definition is based off the Taylor expansion for exp in the real numbers). It follows that $\exp\ \operatorname{ad} x_\alpha$ stabilizes $L(\mathbb{Z})$. Also, it is a simple theorem from ring theory that if u is nilpotent, $1+u$ is invertible. Hence $\exp\ \operatorname{ad} x_\alpha$ is invertible, and so together the $\exp\ \operatorname{ad} x_\alpha$ generate a matrix group (relative to the Chevalley basis) that acts on $L(\mathbb{Z})$, hence has integral coefficients.

More generally, we can consider $\exp\ \operatorname{ad} Tx_\alpha$, where T is an indeterminant, which is an invertible matrix with coefficients in $\mathbb{Z}[T]$. These matrices generate a matrix group G, reducing the coefficients mod p and considering T to be an element of an arbitrary extension field K of $\mathbb{F}_p$ we get a matrix group $G(K)$ over K. This is called the **Chevalley group (of adjoint type)**.

This concludes the work of this paper.

Acknowledgments. I thank my mentor, Tam, very much for all of her help getting my thoughts straight and helping me create a paper I am proud of.

References

[1] J.E. Humphreys Introduction to Lie Algebras and Representation Theory
[2] The Bruhat-Tits building of a p-adic Chevalley group and... www.math.harvard.edu/ rabinoff/misc/building.pdf