From Lie Algebras to Chevalley Groups

Robert James Brown

Master of Science by Research

University of York
Mathematics
January 2018

Abstract

We follow Humphreys [4], studying the structure theory of semisimple Lie algebras (over algebraically closed fields of characteristic zero) in detail, proving the existence of a Chevalley basis and constructing Chevalley groups of adjoint type.

We provide elementary definitions and results about Lie algebras. We take the perspective of toral subalgebras to show the root space decomposition with respect to a maximal torus. We utilise representation theory to prove that the set of roots forms a root system. Studying root systems in their own right then gives us further structural results for semisimple Lie algebras. These enable us to prove the existence of a Chevalley basis, which allows us to transfer the Lie algebra structure to finite fields. We conclude by using this to construct Chevalley groups of adjoint type.

Contents

Preface

Finite simple groups have been completely classified [2] and the classification theorem is familiar to many. Such groups fall into one of the following categories: prime cyclic groups, alternating groups, Chevalley groups, twisted Chevalley groups and sporadic groups. Cyclic groups and alternating groups are straightforward to construct, but this is not the case for the other categories. It is possible to give relatively elementary constructions in some circumstances, for example the Suzuki [7] and Ree [8, 9] groups. However, such constructions are done with the aim of giving shortcuts to known destinations, so they lack the overarching mathematical backbone which connects all of these groups together.

It is important, then, for the standard constructions, and the machinery behind them, to be understood. This machinery is known as Lie theory (pronounced 'lee'), named after the Norwegian mathematician Sophus Lie. For us, this will be the study of Lie algebras; Lie groups are not mentioned. We primarily follow the path laid out by Humphreys [4], framing the root space decomposition in terms of toral subalgebras, as opposed to the more traditional Cartan subalgebras [1, 3]. Our goal is the construction of Chevalley groups (of adjoint type). We give a process for taking a finite field K and a semisimple Lie algebra L over an algebraically closed field of characteristic zero (the complex numbers, for example), and producing a finite group of matrices with entries in K, determined by the Lie algebra structure of L. This process is sufficient to construct all families of non-twisted Chevalley groups; the twisted Chevalley groups require further machinery [1].

This construction of Chevalley groups is our singular goal. As such we take a somewhat streamlined path through the literature, only discussing the results needed for this. In particular, the classification of Lie algebras using Dynkin diagrams, though important to Lie Theory as a whole, is not covered. This approach does allow us room to cover more detail, however. The elementary but essential results regarding simultaneous diagonalisibility, which are omitted in [4], are covered in full (Section 2). The discussions on the machinery behind the construction in sections 25.4 and 25.5 of [4] are expanded into formal results with proofs (Sections 18 and 19). Notably, a proof that $G(K)$ is an automorphism group of $L(K)$ (Theorem 19.16) and an explicit example of such a construction (example 19.17) are given. More explicit proofs have been given throughout, though it should be noted that certain results have been stated without proof (the exhaustive list is: Engel's Theorem (1.21), a corollary to Lie's Theorem (1.27), Weyl's Theorem (9.4), and an automorphism lemma (17.1)).

We begin with a brief introduction to some of the basic concepts (Section 1), followed by a discussion of some ideas from linear algebra that are of particular importance to us (Section 2). The notion of Jordan decomposition, though also very grounded in linear algebra, is the first key idea relating directly to Lie algebras which we discuss (Section 3). We prove Cartan's criterion (Section 4), which is necessary only for one crucial result: that the Killing form is nondegenerate (Section 7). Our study of Lie algebras begins in earnest with the introduction of toral subalgebras (Section 5), which leads into the central idea of the root space decomposition of a semisimple Lie algebra (Section 6). Our understanding of this decomposition will improve when we show that it revolves around the maximal torus (Section 8), then again when we begin to see the structure of $\mathfrak{sl}_2(\mathbb{F})$ appearing (Section 10), and finally when we discover that root spaces are 1-dimensional (Section 11). For this, we use the representation theory of $\mathfrak{sl}_2(\mathbb{F})$, which we develop beforehand (Section 9). We come to a temporary peak with the proof that the set of

roots has the structure of a root system (Section 12), which we follow with a look at root systems themselves (Sections 13 to 16). We are then set to prove the existence of a Chevalley basis (Section 17). This is the key we need to transfer our Lie algebra structure to a finite field (Section 18), over which we can construct automorphism groups generated by exponentiating the adjoint maps of (non-toral) elements of the Chevalley basis (Section 19). These are the Chevalley groups of adjoint type.

We write $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ to denote the natural numbers, integers, rationals and reals respectively. $\mathbb{Z}[T]$ denotes the set of polynomials in T with coefficients in $\mathbb{Z}$. The symbol $\subset$ denotes strict containment of sets; we always use $\subseteq$ when equality is a possibility. We use the shorthand $\mathbb{F}x$ to denote the $\mathbb{F}$-span of x.

Acknowledgements

I would like to thank my supervisors Michael Bate and Brent Everitt, whose support and guidance have been invaluable. Thanks are also due to all of the other fantastic lecturers at York. You have been a great inspiration to myself and many others.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

1. First Definitions

We begin by defining Lie algebras and laying out some of the basic results which we use in later sections. Certain results, such as Engel's Theorem (1.21) and a corollary to Lie's Theorem (1.27), will be taken as assumed (for proofs, see [4]).

Throughout this paper, $\mathbb{F}$ denotes some field and L denotes a finite dimensional Lie algebra (defined below) over $\mathbb{F}$. We place restrictions on L and $\mathbb{F}$ in subsequent sections.

1.1. **Definition.** A **Lie algebra**, L, is a vector space over some field $\mathbb{F}$, equipped with a multiplication

$$L \times L \to L : (x, y) \mapsto [xy],$$

called a **Lie bracket**, which is bilinear and satisfies

- $[xx] = 0$ for all $x \in L$,
- $[x[yz]]+[z[xy]]+[y[zx]] = 0$ for all $x, y, z \in L$. This is called the **Jacobi identity**.

1.2**.** It is useful to consider various reformulations of the Jacobi identity:

$$[x[yz]] + [z[xy]] + [y[zx]] = 0, \tag{1.2.1}$$

$$[x[yz]] = -[z[xy]] - [y[zx]], \tag{1.2.2}$$

$$[x[yz]] = [y[xz]] - [z[xy]] = [z[yx]] - [y[zx]]. \tag{1.2.3}$$

1.3. **Definition.** Let V be a vector space. We denote by $\mathrm{End}(V)$ the **space of linear maps** from V to V. Note that if L is a Lie algebra, then when we write $\mathrm{End}(L)$, we are still referring to vector space endomorphisms of the Lie algebra, not Lie algebra endomorphisms. The space $\mathrm{End}(V)$ is an associative algebra under the composition of functions operation, denoted $\circ$.

We can define a Lie bracket on $\mathrm{End}(V)$ by

$$[xy] = x \circ y - y \circ x,$$

for $x, y \in \mathrm{End}(V)$. We then obtain a Lie algebra called the **general linear Lie algebra** of V, denoted $\mathfrak{gl}(V)$. For $n \in \mathbb{N}$, we also write $\mathfrak{gl}_n(\mathbb{F})$ to denote the general linear Lie algebra of $\mathbb{F}^n$. The above method (defining a Lie bracket from the commutator of an algebra) can be used to generate a Lie algebra from any associative algebra. Such a Lie algebra is called a **linear Lie algebra**. It is also true that any finite dimensional Lie algebra is isomorphic to some linear Lie algebra (see [5, Chapter VI]).

Another example of a Lie algebra which will be useful to us is the **special linear Lie algebra**, denoted $\mathfrak{sl}(V)$. This is defined as the subalgebra of $\mathfrak{gl}(V)$ consisting of endomorphisms with trace 0. We also write $\mathfrak{sl}_n(\mathbb{F}) = \mathfrak{sl}(\mathbb{F}^n)$ for $n \in \mathbb{N}$.

1.4. **Definition.** A **homomorphism** of Lie algebras is a linear map $\phi : L \to L' : x \mapsto \phi(x)$ which satisfies $[\phi(x)\phi(y)] = \phi([xy])$ for all $x, y \in L$. An **isomorphism** of Lie algebras is a bijective Lie algebra homomorphism. An isomorphism from a Lie algebra to itself is called an **automorphism** and we denote the set of all automorphisms of L by $\mathrm{Aut}(L)$. This is a group under the composition of functions operation.

1.5. **Definition.** Let A, B be subspaces of L. We write $[AB] = \mathrm{span}_{\mathbb{F}} \{[ab] : a \in A, b \in B\}$ for the **product of subspaces**. Note that taking the span is necessary for this to be a subspace, because it is not in general possible to express a sum $[ab] + [a'b']$ as a single product.

We call A a **Lie subalgebra** of L if it is closed under the Lie bracket. That is, if $[AA] \subseteq A$, so A is a Lie algebra in its own right. The Lie subalgebra $[LL]$ of L is called the **derived subalgebra** of L.

Let I be some subspace of L. Then I is called an **ideal** of L if $[LI] \subseteq I$. Important examples of ideals are the derived subalgebra $[LL]$ and the centre of L (Definition 1.7).

1.6. **Lemma.** *The derived subalgebra of L is an ideal of L.*

Proof. Let $x, y, z \in L$. Then $[yz] \in L$ by definition, hence $[x[yz]] \in [LL]$. Therefore, $[L[LL]] \subseteq [LL]$. □

1.7. **Definition.** Write $Z(L) = \{z \in L : \forall x \in L : [xz] = 0\}$ and call this space the **centre** of L.

1.8. **Lemma.** *$Z(L)$ is an ideal of L.*

Proof. Let $x \in L$ and $z \in Z(L)$. Then $[xz] = 0$ by definition, where $0 \in Z(L)$ by linearity. □

1.9. **Definition.** Call L **simple** if $[LL] \neq \{0\}$ and the only ideals of L are $\{0\}$ and L. Throughout, we are primarily interested in semisimple Lie algebras (Definition 5.3), which is a weaker condition than simplicity. Semisimple Lie algebras are direct sums of simple ones.

1.10**.** Note that if L is simple, then as $[LL]$ is an ideal (Lemma 1.6), we must have that $[LL] = L$.

1.11. **Definition.** A map

$$\delta : L \to L : x \mapsto \delta(x),$$

is called a **derivation** of L if it is linear and satisfies

$$\delta([xy]) = [x\delta(y)] + [\delta(x)y]$$

for all $x, y \in L$. The space of derivations of L is denoted $\mathrm{Der}(L)$ and is a Lie subalgebra of $\mathfrak{gl}(L)$. The derivations we are interested in are the following:

Let $x \in L$. The map

$$\mathrm{ad}(x) : L \to L : y \mapsto [xy]$$

is called the **adjoint endomorphism** or **adjoint map** associated to x. It is important to note that $\mathrm{ad}(x)$ is a vector space endomorphism, but not in general a Lie algebra endomorphism. Derivations of this form are called **inner**.

When S is a subalgebra of L and $x \in S$, we can consider $\mathrm{ad}(x)$ to be acting on either L or S. To remove ambiguity, we write

$$\mathrm{ad}_L(x) : L \to L : y \mapsto [xy]$$

or

$$\mathrm{ad}_S(x) : S \to S : y \mapsto [xy]$$

in these situations.

1.12**.** As the Lie bracket is bilinear, it follows that $\mathrm{ad}(x) \in \mathrm{End}(L)$ for $x \in L$. However, $\mathrm{ad}(x)$ is not in general a Lie algebra homomorphism. We often use the notation of the Lie bracket and that of adjoint maps interchangeably, depending on what we want to emphasise.

The fact that adjoint maps are derivations will become useful in Section 19: Constructing Chevalley Groups, when we construct automorphisms by exponentiating certain adjoint maps.

1.13. **Lemma.** *Let $x \in L$. Then the map $\mathrm{ad}(x)$ is a derivation. That is, for $y, z \in L$,*

$$\mathrm{ad}(x)([yz]) = [y\,\mathrm{ad}(x)(z)] + [\mathrm{ad}(x)(y)z].$$

Proof. Immediate from reformulating the Jacobi identity (1.2.3). □

1.14. The notion of adjoint maps is important, because it gives us a very useful way of representing Lie algebras. We use the adjoint representation (1.15) directly, in order to prove many key structural results in Section 11: Integrality Properties. Additionally, many ideas, such as nilpotency (1.19), are much easier to phrase in terms of adjoint maps.

1.15. **Definition.** Let V be some vector space over $\mathbb{F}$. Then a Lie algebra homomorphism $\rho : L \to \mathfrak{gl}(L)$ is called a **representation** of L.

The representation we are interested in is called the **adjoint representation** of L, denoted ad. We define

$$\mathrm{ad} : L \to \mathfrak{gl}(L) : x \mapsto \mathrm{ad}(x).$$

As in the definition of $\mathrm{ad}(x)$ (1.11), we write ad_L when there is ambiguity as to the domain.

1.16. **Lemma.** *The map* ad *is a Lie algebra homomorphism (and hence is a representation).*

Proof. Let $x, y, z \in L$. The Lie bracket in $\mathfrak{gl}(L)$ is given by

$$[\mathrm{ad}(x)\,\mathrm{ad}(y)] = \mathrm{ad}(x) \circ \mathrm{ad}(y) - \mathrm{ad}(y) \circ \mathrm{ad}(x).$$

Therefore,

$$\begin{aligned}
[\mathrm{ad}(x)\,\mathrm{ad}(y)](z) &= (\mathrm{ad}(x) \circ \mathrm{ad}(y) - \mathrm{ad}(y) \circ \mathrm{ad}(x))(z) \\
&= (\mathrm{ad}(x) \circ \mathrm{ad}(y))(z) - (\mathrm{ad}(y) \circ \mathrm{ad}(x))(z) \\
&= \mathrm{ad}(x)(\mathrm{ad}(y)(z)) - \mathrm{ad}(y)(\mathrm{ad}(x)(z)) \\
&= [x[yz]] - [y[xz]].
\end{aligned}$$

By the Jacobi identity (1.2.3), this is equal to

$$[z[yx]] = -[[yx]z] = [[xy]z] = \mathrm{ad}([xy])(z).$$

As z was arbitrary, this implies that $\mathrm{ad}([xy]) = [\mathrm{ad}(x)\,\mathrm{ad}(y)]$. □

1.17. We can see that

$$\ker(\mathrm{ad}) = \{x \in L : \mathrm{ad}(x) = 0\} = \{x \in L : \forall y \in L : [xy] = 0\} = Z(L),$$

so $Z(L) = \{0\}$ if and only if ad is an injective homomorphism. That is,

$$L \cong \mathrm{ad}(L) \iff Z(L) = \{0\}.$$

1.18. **Lemma.** *Let $x, y \in L$ such that $[xy] = 0$. Then* $\mathrm{ad}(x)$ *and* $\mathrm{ad}(y)$ *commute.*

Proof. For all $z \in L$,

$$(\mathrm{ad}(x) \circ \mathrm{ad}(y))(z) = [x[yz],$$

which, applying the Jacobi identity (1.2.3) and the assumption that $[xy] = 0$, gives

$$[x[yz] = [y[xz]] - [z[xy]] = [y[xz]] = (\mathrm{ad}(y) \circ \mathrm{ad}(x))(z).$$

As this holds for all $z \in L$, we have

$$\mathrm{ad}(x) \circ \mathrm{ad}(y) = \mathrm{ad}(y) \circ \mathrm{ad}(x).$$

□

1.19. **Definition.** Let V be some vector space over $\mathbb{F}$ and $x \in \mathrm{End}(V)$. If $x^n = 0$ for some $n \in \mathbb{N}$, then x is called **nilpotent**. We say an element $x \in L$ is ad-nilpotent if $\mathrm{ad}(x)$ is nilpotent. L is called **nilpotent** if $\mathrm{ad}(L)^n(L) = \{0\}$ for some $n \in \mathbb{N}$.

1.20. Note that if all $x \in L$ are ad-nilpotent, then there exists an $n \in \mathbb{N}$ such that

$$\mathrm{ad}(x)^n(y) = 0$$

for all $x, y \in L$, whereas if L is nilpotent, then there exists an $n \in \mathbb{N}$ such that

$$(\mathrm{ad}(x_1) \circ \cdots \circ \mathrm{ad}(x_n))(y) = 0$$

for all $x_1, \cdots, x_n, y \in L$, so the latter appears to be a stronger condition. It is in fact true that the former implies the latter, so they are equivalent. This result is known as Engel's Theorem, and the proof nontrivial. We state Engel's Theorem here as an assumption, in addition to another result, about ideals of nilpotent Lie algebras. These will both be used in Section 8: The Maximal Torus.

1.21. **Theorem** (Engel's Theorem)**.** *Suppose that* $\mathrm{ad}(x)$ *is nilpotent for all* $x \in L$*. Then* L *is nilpotent.*

Proof. See Section 3.3 of [4]. □

1.22. **Lemma.** *Let* L *be nilpotent and let* K *be an ideal of* L*. Then* $K \neq \{0\}$ *implies that* $K \cap Z(L) \neq \{0\}$*.*

Proof. See Lemma 3.3 of [4]. □

1.23. **Definition.** The **derived series** of L is denoted

$$\left\{L^{(n)}\right\}_{n \geq 0},$$

where $L^{(0)} = L$ and $L^{(n)} = [L^{(n-1)}L^{(n-1)}]$ for $n > 0$. If $L^{(n)} = \{0\}$ for some $n \in \mathbb{N}$, then L is called **solvable**. An important example of a solvable Lie algebra is the centre $Z(L)$ of L.

1.24. **Lemma.** $Z(L)$ *is solvable.*

Proof. For all $x, y \in Z(L)$, we have $[xy] = 0$ by definition, hence $[Z(L)Z(L)] = \{0\}$. □

1.25. Solvability is a weaker property than nilpotency (Lemma 1.26) - a fact we use multiple times. Although it is also true that $[LL]$ is nilpotent if L is solvable. This result (Theorem 1.27) is a corollary to Lie's theorem and shall be taken as an assumption. It is used in Section 10: Orthogonality Properties.

1.26. **Lemma.** *Let* L *be nilpotent. Then* L *is solvable.*

Proof. Suppose $L^{(i)} \subseteq \mathrm{ad}(L)^i(L)$ for some $i \in \mathbb{N}$. Then, as $L^{(i)} \subseteq L$, we have

$$L^{(i+1)} = [L^{(i)}L^{(i)}] \subseteq [\mathrm{ad}(L)^i(L)L] = \mathrm{ad}(L)^{i+1}(L).$$

Further, $L^{(1)} = [LL] = \mathrm{ad}(L)(L)$. Therefore, by induction, $L^{(i)} \subseteq \mathrm{ad}(L)^i(L)$ for all $i \in \mathbb{N}$.

As L is nilpotent, $\mathrm{ad}(L)^n(L) = \{0\}$ for some $n \in \mathbb{N}$, hence $L^{(n)} \subseteq \mathrm{ad}(L)^n(L) = \{0\}$. Therefore $L^{(n)} = \{0\}$, so L is solvable. □

1.27. **Theorem** (Corollary to Lie's Theorem)**.** *Let* L *be solvable and let* $x \in [LL]$*. Then the map* $\mathrm{ad}_L(x)$ *is nilpotent. Further,* $[LL]$ *is nilpotent.*

Proof. See Corollary 4.1C of [4]. □

1.28. We have seen that the centre of L is solvable and an ideal. This gives $Z(L)$ a property (Proposition 1.29) which we need to prove a corollary to Cartan's Criterion in Section 4. We prove this result for an arbitrary solvable ideal.

1.29. **Proposition.** *Let* I *be a solvable ideal of* L*. Suppose* L/I *is solvable. Then* L *is solvable.*

Proof. Suppose $(L/I)^{(i)} = L^{(i)}/I$ for some $i \in \mathbb{N}$. Then

$$\begin{aligned}
(L/I)^{(i+1)} &= [(L/I)^{(i)}(L/I)^{(i)}] \\
&= [(L^{(i)}/I)(L^{(i)}/I)] \\
&= \operatorname{span}\left\{[(x+I)(y+I)] : x, y \in L^{(i)}\right\} \\
&= \operatorname{span}\left\{[xy] + I : x, y \in L^{(i)}\right\} \\
&= \operatorname{span}\left\{x + I : x \in [L^{(i)}L^{(i)}]\right\} \\
&= [L^{(i)}L^{(i)}]/I \\
&= L^{(i+1)}/I.
\end{aligned}$$

We have $(L/I)^0 = L/I = L^{(0)}/I$, so by induction we have that $(L/I)^{(i)} = L^{(i)}/I$ for all $i \in \mathbb{N}$.

As L/I is solvable, $(L/I)^{(n)}$ is zero for some $n \in \mathbb{N}$. Consider the canonical homomorphism

$$\pi : L \to L/I : x \mapsto x + I.$$

As $(L/I)^{(n)} = L^{(n)}/I$, we have

$$\pi(L^{(n)}) = \left\{x + I : x \in L^{(n)}\right\} = L^{(n)}/I = (L/I)^{(n)}.$$

Therefore, $L^{(n)} \subseteq \ker(\pi) = I$. As I is solvable, $I^{(m)} = \{0\}$ for some $m \in \mathbb{N}$, hence

$$L^{(n+m)} = (L^{(n)})^{(m)} \subseteq I^{(m)} = \{0\}.$$

That is, $L^{(n+m)} = \{0\}$, so L is solvable. □

1.30. **Example.** We use the toy example of $L = \mathfrak{sl}_2(\mathbb{F})$. By definition, this consists of traceless 2×2 matrices over $\mathbb{F}$. So elements of L can have arbitrary off-diagonal entries, and the diagonal entries must sum to zero. That is,

$$L = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} : a, b, c \in \mathbb{F} \right\}.$$

It would seem natural, then, to choose the basis

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

This turns out to be a good choice, as it fits well with the structure theory which we develop later. It is considered the standard basis for $\mathfrak{sl}_2(\mathbb{F})$.

As a subalgebra of $\mathfrak{gl}_2(\mathbb{F})$, the Lie bracket is given by $[xy] = x \circ y - y \circ x$. That is,

$$\begin{aligned}
[xy] &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = h, \\
[hx] &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = 2x, \\
[hy] &= \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -2 & 0 \end{pmatrix} = -2y.
\end{aligned}$$

2. Semisimplicity and Simultaneous Diagonalisation

Results about Jordan decomposition (Section 3) and root-space decomposition (Section 5) rely upon the ability to find a basis on which multiple diagonalisable maps are simultaneously diagonal. We prove results to that effect in this section.

Throughout this section, V denotes a finite dimensional vector space over $\mathbb{F}$. From this section onwards, we assume $\mathbb{F}$ is algebraically closed.

2.1. **Definition.** Let $x \in \mathrm{End}(V)$. We call x **semisimple** if the roots of its minimal polynomial are all distinct. As $\mathbb{F}$ is algebraically closed, this is equivalent to x being diagonalisable.

2.2. **Lemma.** *Let $x \in \mathfrak{gl}(V)$ be semisimple. Let $(v_1, \cdots, v_n)$ be a basis for V on which x is diagonal. Denote the standard basis vectors of $\mathfrak{gl}(V)$ relative to $(v_1, \cdots, v_n)$ by $e_{i,j}$ for $i, j = 1, \cdots, n$. That is, $e_{i,j}(v_k) = \delta_{i,k} v_j$. Then $xe_{i,j} = \lambda_i e_{i,j}$ and $e_{i,j}x = \lambda_j e_{i,j}$, where the values $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of x corresponding to $v_1, \cdots, v_n$ respectively.*

Proof. For each $k = 1, \cdots, m$, we have

$$(e_{i,j}x)(v_k) = \lambda_k e_{i,j}(v_k) = \lambda_k \delta_{i,k} v_j.$$

But when $k = i$, we have

$$\lambda_k \delta_{i,k} v_j = \lambda_i \delta_{i,i} v_j = \lambda_i e_{i,j}(v_k),$$

and when $k \neq i$, we have

$$\lambda_k \delta_{i,k} v_j = 0 = \lambda_i \delta_{i,k} v_j = \lambda_i e_{i,j}(v_k).$$

That is, $(e_{i,j}x)(v_k) = \lambda_i e_{i,j}(v_k)$ for all $k = 1, \cdots, m$, hence $e_{i,j}x = \lambda_i e_{i,j}$. Similarly, $xe_{i,j} = \lambda_j e_{i,j}$. □

2.3. **Proposition.** *Let $x_1, \cdots, x_m, y \in \mathrm{End}(V)$ commute and be diagonalisable. Suppose that $x_1, \cdots, x_m$ are simultaneously diagonal with respect to some basis $B = (e_1, \cdots, e_n)$ of V. Then $x_1, \cdots, x_m, y$ are all simultaneously diagonalisable.*

Proof. We are given that $x_i(e_j) = \lambda_{i,j} e_j$ for some scalars $\lambda_{i,j} \in \mathbb{F}$, for each $i = 1, \cdots, m$ and $j = 1, \cdots, n$. Then, as the maps commute,

$$x_i(y(e_j)) = y(x_i(e_j)) = \lambda_{i,j} y(e_j),$$

for each i and j. So y stabilizes the eigenspaces of all the x_i. Therefore, if V is decomposed into the intersections of the eigenspaces of all the x_i, say $V_1, \cdots, V_r$, then y stabilizes each V_j. Therefore, if we consider these maps as matrices with respect to B, then we can consider y and all the x_i as block diagonal matrices with blocks corresponding to the V_j. Then the x_i consist of scalar blocks. Specifically,

$$x_i = \begin{pmatrix} \Lambda_{i,1} & & \\ & \ddots & \\ & & \Lambda_{i,r} \end{pmatrix},$$

for each $i = 1, \cdots, m$, where the blocks $\Lambda_{i,j}$ are all scalar matrices (not necessarily distinct scalars), and $\Lambda_{1,j}, \cdots, \Lambda_{r,j}$ are all the same size for each j. As for y, we can write

$$y = \begin{pmatrix} Y_1 & & \\ & \ddots & \\ & & Y_r \end{pmatrix},$$

where each block Y_j is a square matrix of the same size as the $\Lambda_{i,j}$. If a block-diagonal matrix is diagonalisable, then each block must be diagonalisable, so each block $Y_1, \cdots, Y_r$ is diagonalisable.

Therefore, for each $j = 1, \cdots, r$, we have that $D_j = C_j Y_j C_j^{-1}$ is diagonal for some change of basis matrix C_j. Note that

$$C_j \Lambda_{i,j} C_j^{-1} = C_j(\lambda_{i,j} I) C_j^{-1} = \lambda_{i,j}(C_j C_j^{-1}) = \lambda_{i,j} I = \Lambda_{i,j},$$

for each $i = 1, \cdots, m$ and $j = 1, \cdots, r$, as $\Lambda_{i,j}$ is a scalar matrix for some eigenvalue $\lambda_{i,j} \in \mathbb{F}$ of x_i. So if we let

$$c = \begin{pmatrix} C_1 & & \\ & \ddots & \\ & & C_r \end{pmatrix},$$

then $c x_i c^{-1} = x_i$ for each $i = 1, \cdots, r$. Further,

$$cyc^{-1} = \begin{pmatrix} C_1 Y_1 C_1^{-1} & & \\ & \ddots & \\ & & C_r Y_r C_r^{-1} \end{pmatrix} = \begin{pmatrix} D_1 & & \\ & \ddots & \\ & & D_r \end{pmatrix},$$

which is diagonal. So conjugation by c preserves the diagonality of each x_i and takes y to a diagonal matrix. That is, if we apply c to the basis B, we obtain a basis on which $x_1, \cdots, x_m, y$ are all diagonal. □

2.4. **Corollary.** *Let $A \subseteq \mathrm{End}(V)$ be some finite set of commuting diagonalisable maps. Then all the maps in A are simultaneously diagonalisable.*

Proof. We can use induction on $|A|$ with Proposition 2.3 as the induction step. The base case of $|A| = 1$ is trivially true. □

2.5. **Lemma.** *Let $x_1, \cdots, x_r \in \mathrm{End}(V)$ be diagonal with respect to some basis $B = (e_1, \cdots, e_n)$ of V. Then any linear combination of $x_1, \cdots, x_r$ is also diagonal with respect to B.*

Proof. Let $y \in \mathrm{End}(V)$ be some linear combination of $x_1, \cdots, x_n$. That is,

$$y = \sum_{i=1}^{r} \mu_i x_i$$

for some set of scalars $\mu_i \in \mathbb{F}$. We have that $x_i(e_j) = \lambda_{i,j} e_j$ for some $\lambda_{i,j} \in \mathbb{F}$, for each $i = 1, \cdots, r$ and $j = 1, \cdots, n$. Therefore,

$$y(e_j) = \sum_{i=1}^{n} \mu_i x_i(e_j) = \sum_{i=1}^{n} \mu_i \lambda_{i,j} e_j = \left(\sum_{i=1}^{n} \mu_i \lambda_{i,j} \right) e_j,$$

for each $j = 1, \cdots, n$, hence y is diagonal with respect to B. □

2.6. **Theorem.** *Let $A \subseteq \mathrm{End}(V)$ be some set of commuting diagonalisable maps. Then all the maps in A are simultaneously diagonalisable.*

Proof. As $A \subseteq \mathrm{End}(V)$, we must have that $\mathrm{span}_{\mathbb{F}}(A) \subseteq \mathrm{End}(V)$ has a finite basis contained in A, say $x_1, \cdots, x_r$ (where $r \leq \dim(V)$). By Corollary 2.4, the x_i are simultaneously diagonalisable. Then by Lemma 2.5, any linear combination of $x_1, \cdots, x_r$ also diagonalises simultaneously with $x_1, \cdots, x_r$. That is, all maps in A are simultaneously diagonalisable. □

3. Jordan Decomposition

Jordan decomposition (or Jordan-Chevalley decomposition) is the idea that an endomorphism (or an element of a Lie algebra) can be split into a sum of two parts, one nilpotent and one semisimple (or ad-nilpotent and ad-semisimple for Lie algebras).

We begin with some results about semisimple and nilpotent endomorphisms. After the main result (Jordan Decomposition), we extend the notion to abstract Lie algebras through the adjoint representation, then show that the two correspond.

Throughout this section, V denotes a finite dimensional vector space over $\mathbb{F}$.

3.1. **Lemma.** *Let $x, y \in \mathrm{End}(V)$. Suppose x and y commute and are both semisimple. Then $x + y$ is semisimple.*

Proof. As x and y commute and are diagonalisable, Proposition 2.3 implies that x and y are simultaneously diagonalisable. Then Lemma 2.5 implies that $x + y$ is diagonalisable. □

3.2. **Lemma.** *Let $x, y \in \mathrm{End}(V)$. Suppose x and y commute and are both nilpotent. Then $x + y$ is nilpotent.*

Proof. As x and y are nilpotent, there exist $n_x, n_y \in \mathbb{N}$ such that $x^{n_x} = y^{n_y} = 0$. Let $n = \max\{n_x, n_y\}$. Then $x^n = y^n = 0$. Now consider

$$(x+y)^{2n} = \sum_{i=1}^{2n} \binom{2n}{i} x^i y^{2n-i}.$$

Note that whenever $i < n$, we have $2n - i > n$. So for each term in the sum, either $x^i = 0$ or $y^{2n-i} = 0$. Therefore, $(x + y)^{2n} = 0$, hence $x + y$ is nilpotent. □

3.3. **Lemma.** *Let $x \in \mathrm{End}(V)$. Suppose x is both nilpotent and semisimple. Then $x = 0$.*

Proof. As x is semisimple, there exists a basis for V, say $(v_1, \cdots, v_n)$, and a collection of scalars $\lambda_1, \cdots, \lambda_n \in \mathbb{F}$ such that $x(v_i) = \lambda_i v_i$ for all $i = 1, \cdots, n$. On the other hand, x is nilpotent, so $x^m = 0$ for some $m \in \mathbb{N}$. Therefore, $x^m(v_i) = \lambda_i^m v_i = 0$, hence $\lambda_i = 0$ and $x(v_i) = 0$ for all $i = 1, \cdots, n$. That is, $x = 0$. □

3.4. **Theorem** (Jordan Decomposition)**.** *Let $x \in \mathrm{End}(V)$. Then:*

- *There exist unique $x_s, x_n \in \mathrm{End}(V)$, such that $x = x_s + x_n$, where x_s is semisimple and x_n is nilpotent and x_s and x_n commute.*
- *There exist polynomials $p(X), q(X) \in \mathbb{F}[X]$, without constant term, such that $x_s = p(x)$ and $x_n = q(x)$. In particular, x_s and x_n commute with any endomorphism which commutes with x.*
- *Let A and B are subspaces of V where $A \subseteq B \subseteq V$. Suppose $x(B) \subseteq A$. Then $x_s(B) \subseteq A$ and $x_n(B) \subseteq A$.*

Proof. Let χ be the characteristic polynomial of x. This can be expressed as

$$\chi(X) = \prod_{i=1}^{n} (X - \lambda_i)^{m_i},$$

for some $n \in \mathbb{N}$, where the values $\lambda_1, \cdots, \lambda_n$ are the distinct eigenvalues of x with corresponding multiplicities $m_1, \cdots, m_n$. Let 1_V denote the identity in $\mathrm{End}(V)$. Define

$$V_i = \ker\left((x - \lambda_i 1_V)^{m_i}\right).$$

As $\mathbb{F}$ is algebraically closed, all roots of the polynomial $\chi(X)$ exist, hence V can be expressed as the direct sum of these subspaces:

$$V = \bigoplus_{i=1}^{n} V_i. \tag{3.4.1}$$

As x commutes with itself and 1_V, it commutes with $(x-\lambda_i 1_V)$ and hence $(x-\lambda_i 1_V)^{m_i}$. Therefore,

$$(x - \lambda_i 1_V)^{m_i}(x(v)) = x((x - \lambda_i 1_V)^{m_i}(v)) = x(0) = 0$$

and hence $x(v) \in V_i$ for all $v \in V_i$. That is, $x(V_i) \subseteq V_i$. Therefore, x can be considered as a linear endomorphism of each V_i. In this regard, x has characteristic polynomial $(X - \lambda_i)^{m_i}$. Applying the Chinese Remainder Theorem for the ring $\mathbb{F}[X]$ allows us to find a polynomial p satisfying $p(X) \equiv 0 \mod X$ and

$$p(X) \equiv \lambda_i \mod (X - \lambda_i)^{m_i} \tag{3.4.2}$$

for each $i = 1, \cdots, n$.

Set $q(X) = X - p(X)$. As $p(X) \equiv 0 \mod X$, there exists some $r \in \mathbb{F}[X]$ satisfying

$$p(X) = 0 + Xr(X),$$

hence $p \in X\mathbb{F}[X]$. Therefore, $q(X) = X - Xr(X) = X(1 - r(X))$, hence both p and q have no constant term.

Let $x_s = p(x)$ and $x_n = q(x)$. As x_s and x_n are polynomials in x, they commute with each other and any endomorphisms which commute with x, in addition to stabilizing any subspaces stabilized by x. In particular, they stabilize V_i for each $i = 1, \cdots, n$. By (3.4.2), we have that $p(X) = \lambda_i + r(X)(X - \lambda_i)^{m_i}$ for some $r \in \mathbb{F}[X]$, hence $x_s = p(x) = \lambda_i + r(x)(x - \lambda_i)^{m_i}$. Therefore, $(x_s - \lambda_i) = r(x)(x - \lambda_i)^{m_i}$. Let $v \in V_i$. Then $(x - \lambda 1_V)^{m_i}(v) = 0$ by definition, hence

$$(x_s - \lambda_i 1_V)(v) = (r(x)(x - \lambda_i 1_V)^{m_i})(v) = r(x)(0) = 0.$$

That is, $x_s(v) - (\lambda_i 1_V)(v) = 0$, hence $x_s(v) = \lambda_i v$. As $v \in V_i$ was arbitrary, this implies

$$x_s = \lambda_i 1_V. \tag{3.4.3}$$

Therefore, x_s acts diagonally on V_i with sole eigenvalue λ_i. Therefore, by (3.4.1), x_s acts diagonally on the whole of V, hence is semisimple. Further, (3.4.3) implies that $x_n = x - x_s = x - \lambda_i 1_V$, hence

$$x_n^{m_i}(v) = (x - \lambda_i 1_V)^{m_i}(v) = 0,$$

for all $v \in V_i$, hence x^n is nilpotent on V_i. This is true for all $i = 1, \cdots, n$, so by (3.4.1), x_n is nilpotent on the whole of V.

If $U \subseteq V$ is a subspace stabilized by x, then U is stabilized by x^k for any $k \in \mathbb{N}$. As U is a subspace, it is closed under addition and scalar multiplication and so is also stabilized by any sum or scalar multiple of stabilizing endomorphisms. That is, U is stabilised by any polynomial of x without constant term, in particular x_s and x_n.

Now it only remains to prove uniqueness. Suppose $x = s + n$ is another such decomposition of x. Then

$$s + n = x = x_s + x_n \implies x_s - s = n - x_n$$

and by part 2, all of these endomorphisms commute. Sums of commuting semisimple/nilpotent endomorphisms are also semisimple/nilpotent (Lemma 3.1 and 3.2). Further, if an endomorphism is both semisimple and nilpotent, then it is zero (Lemma 3.3). Therefore, $x_s - s$ is semisimple; $n - x_n$ is nilpotent; and $x_s - s = n - x_n$. Thus, $x_s - s$ is

both semisimple and nilpotent, hence zero. That is, $x_s - s = n - x_n = 0$, hence $x_s = s$ and $x_n = n$. Therefore the Jordan decomposition is unique. □

3.5. **Lemma.** *Let $x \in \mathfrak{gl}(V)$ be nilpotent. Then* $\mathrm{ad}(x)$ *is nilpotent.*

Proof. Consider the endomorphisms

$$\lambda_x : \mathfrak{gl}(V) \to \mathfrak{gl}(V) : y \mapsto x \circ y,$$
$$\rho_x : \mathfrak{gl}(V) \to \mathfrak{gl}(V) : y \mapsto y \circ x.$$

Let $y \in \mathfrak{gl}(V)$. As x is nilpotent, $x^n = 0$ for some $n \in \mathbb{N}$. Therefore,

$$(\lambda_x)^n(y) = x^n y = 0,$$

and

$$(\rho_x)^n(y) = y x^n = 0,$$

for some $n \in \mathbb{N}$, hence λ_x and ρ_x are both nilpotent. Further,

$$(\lambda_x \rho_x)(y) = \lambda_x(\rho_x(y)) = x \circ (y \circ x) = (x \circ y) \circ x = \rho_x(\lambda_x(y)) = (\rho_x \lambda_x)(y),$$

so λ_x and ρ_x commute. Sums of commuting nilpotent endomorphisms are nilpotent (Lemma 3.2), thus $\mathrm{ad}(x) = \lambda_x - \rho_x$ is nilpotent. □

3.6. **Lemma.** *Let $x \in \mathfrak{gl}(V)$ be semisimple. Then* $\mathrm{ad}(x)$ *is semisimple.*

Proof. As x is semisimple, there exists a basis, say $(v_1, \cdots, v_n)$, for V and some collection of scalars $\lambda_1, \cdots, \lambda_n \in \mathbb{F}$ such that $x(v_i) = \lambda_i v_i$ for each $i = 1, \cdots, n$. Denote the standard basis vectors of $\mathfrak{gl}(V)$ relative to $(v_1, \cdots, v_n)$ by $e_{i,j}$ for $i, j = 1, \cdots, n$. That is, $e_{i,j}(v_k) = \delta_{i,k} v_j$. Then Lemma 2.2 implies that

$$\mathrm{ad}(x)(e_{i,j}) = [x e_{i,j}] = x e_{i,j} - e_{i,j} x = \lambda_i e_{i,j} - \lambda_j e_{i,j} = (\lambda_i - \lambda_j) e_{i,j}.$$

Therefore, $\mathrm{ad}(x)$ acts diagonally on $\mathfrak{gl}(V)$, hence is semisimple. □

3.7. **Theorem.** *Let $x \in \mathrm{End}(V)$ and let $x = x_s + x_n$ be its Jordan decomposition. Then* $\mathrm{ad}(x) = \mathrm{ad}(x_s) + \mathrm{ad}(x_n)$ *is the Jordan decomposition of* $\mathrm{ad}(x)$ *in* $\mathrm{End}(\mathrm{End}(V))$.

Proof. We can write $\mathrm{ad}(x) = \mathrm{ad}(x_s + x_n) = \mathrm{ad}(x_s) + \mathrm{ad}(x_n)$, as ad is linear. Then, for $y \in \mathrm{End}(V)$, we have

$$\begin{aligned}[\mathrm{ad}(x_s)\,\mathrm{ad}(x_n)](y) &= \mathrm{ad}(x_s)\,\mathrm{ad}(x_n)(y) - \mathrm{ad}(x_n)\,\mathrm{ad}(x_s)(y) \\ &= [x_s[x_n y]] - [x_n[x_s y]].\end{aligned}$$

Applying the Jacobi identity (1.2.3), we get

$$[x_s[x_n y]] - [x_n[x_s y]] = [y[x_n x_s]] = -[[x_n x_s]y] = -\,\mathrm{ad}([x_n x_s])(y).$$

This implies that $[\mathrm{ad}(x_s)\,\mathrm{ad}(x_n)] = \mathrm{ad}([x_n x_s])$. By Theorem 3.4, x_n and x_s commute, hence $\mathrm{ad}([x_n x_s]) = 0$. Therefore, $[\mathrm{ad}(x_s)\,\mathrm{ad}(x_n)] = 0$, so $\mathrm{ad}(x_s)$ and $\mathrm{ad}(x_n)$ commute. As x_s is semisimple, $\mathrm{ad}(x_s)$ is semisimple (Lemma 3.6); as x_n is nilpotent, $\mathrm{ad}(x_n)$ is nilpotent (Lemma 3.5). Therefore, $\mathrm{ad}(x) = \mathrm{ad}(x_s) + \mathrm{ad}(x_n)$ is the Jordan decomposition of $\mathrm{ad}(x)$. □

3.8. If L is a Lie algebra and $x \in L$, then we can consider the Jordan decomposition of $\mathrm{ad}(x) \in \mathrm{End}(L)$. The above results show that this decomposition corresponds to the Jordan decomposition of x in L, if L is a linear Lie algebra. If L is an abstract Lie algebra - that is, not consisting of endomorphisms - then the Jordan decomposition of an element $x \in L$ is simply taken to be that corresponding to the Jordan decomposition of $\mathrm{ad}(x)$.

4. Cartan's Criterion

We begin with a result about fields of characteristic zero, which will be used immediately to prove a trace criterion for endomorphism nilpotency (Lemma 4.2), as well as later, in Section 12: Rationality Properties. We also prove an associativity identity on the trace function (Proposition 4.3). These results will be used to prove Cartan's Criterion (Theorem 4.4). It will, however, be the corollary to Cartan's Criterion which we will rely upon in Section 7: The Killing Form.

Throughout this section, V denotes a finite dimensional vector space over $\mathbb{F}$. From this section onwards, we assume that $\mathbb{F}$ has characteristic 0 (in addition to algebraic closure).

4.1. **Proposition.** *As $\mathbb{F}$ has characteristic zero, $\mathbb{Q} \subseteq \mathbb{F}$.*

Proof. As $\mathbb{F}$ is a field, by definition we have that $1 \in \mathbb{F}$, where 1 denotes the multiplicative identity. As $\mathbb{F}$ has characteristic zero, we have that $n = (1 + \cdots + 1) \in \mathbb{F}$ for all $n \in \mathbb{N}$. That is, $\mathbb{N} \subseteq \mathbb{F}$.

Also by definition, we have $0 \in \mathbb{F}$ and that $\mathbb{F}$ contains the additive inverses of all its elements. That is, $x \in \mathbb{F}$ implies $-x \in \mathbb{F}$. Therefore, $-\mathbb{N} \subseteq \mathbb{F}$. Thus we have

$$\mathbb{Z} = (\mathbb{N} \cup \{0\} \cup (-\mathbb{N})) \subseteq \mathbb{F}.$$

Further, $\mathbb{F}$ contains the multiplicative inverses of all its nonzero elements. That is, $x \in \mathbb{F} \setminus \{0\}$ implies $x^{-1} \in \mathbb{F}$. Let $q \in \mathbb{Q} \setminus \{0\}$. Then $q = z/n$ for some $z \in \mathbb{Z}$ and $n \in \mathbb{N}$. By the above, $z, n \in \mathbb{F}$, hence $n^{-1} \in \mathbb{F}$. Thus,

$$q = \frac{z}{n} = zn^{-1} \in \mathbb{F},$$

as $\mathbb{F}$ is closed under multiplication. That is, $\mathbb{Q} \subseteq \mathbb{F}$. □

4.2. **Lemma.** *Let A, B be subspaces of $\mathfrak{gl}(V)$ such that $A \subseteq B$. Let*

$$M = \{x \in \mathfrak{gl}(V) : [xB] \subseteq A\}$$

and $y \in M$. Suppose $\operatorname{trace}(yz) = 0$ *for all $z \in M$. Then y is nilpotent.*

Proof. Let $x = s + n$ be the Jordan decomposition of x. As s is semisimple, there exists a basis of V, say $(v_1, \cdots, v_m)$, and some collection of scalars $\lambda_1, \cdots, \lambda_m \in \mathbb{F}$, such that $s(v_i) = \lambda_i v_i$ for each $i = 1, \cdots, m$. The field $\mathbb{F}$ has characteristic zero, so by (Proposition 4.1), we have $\mathbb{Q} \subseteq \mathbb{F}$.

Let $E = \operatorname{span}_{\mathbb{Q}} \{\lambda_1, \cdots, \lambda_m\}$ be the vector space over $\mathbb{Q}$ spanned by the scalars λ_i. This is a vector subspace of $\mathbb{F}$, where $\mathbb{F}$ is considered as a vector space over $\mathbb{Q}$. Denote the dual space of E by E^*. Then the following statements are all equivalent: x is nilpotent; $x = n$; $s = 0$; the scalars $\lambda_1, \cdots, \lambda_m$ are all zero; $E = \{0\}$; $E^* = \{0\}$. In particular,

$$E^* = \{0\} \implies x \text{ nilpotent.} \tag{4.2.1}$$

Let $f \in E^*$. Let $y \in \mathfrak{gl}(V)$ be such that $y(v_i) = f(a_i)v_i$ for each $i = 1, \cdots, m$. Then y is semisimple by definition. Then Lemma 2.2 implies that $e_{i,j}s = \lambda_i e_{i,j}$ and $se_{i,j} = \lambda_j e_{i,j}$, as well as $e_{i,j}y = f(\lambda_i)e_{i,j}$ and $ye_{i,j}y = f(\lambda_j)e_{i,j}$. Therefore,

$$\operatorname{ad}(s)(e_{i,j}) = [se_{i,j}] = se_{i,j} - e_{i,j}s = (\lambda_j - \lambda_i)e_{i,j},$$
$$\operatorname{ad}(y)(e_{i,j}) = [ye_{i,j}] = ye_{i,j} - e_{i,j}y = (f(\lambda_i) - f(\lambda_j))e_{i,j}.$$

Note that, if $\lambda_i - \lambda_j = \lambda_k - \lambda_l$ for some values of i, j, k, l, then $f(\lambda_i - \lambda_j) = f(\lambda_k - \lambda_l)$, hence $f(\lambda_i) - f(\lambda_j) = f(\lambda_k) - f(\lambda_l)$. Further, if $\lambda_i - \lambda_j = 0$ for some values of i and j, then $f(\lambda_i - \lambda_j) = f(\lambda_i) - f(\lambda_j) = 0$. Therefore, for all pairs of i and j, the points $(0, 0)$ and $(\lambda_i - \lambda_j, f(\lambda_i) - f(\lambda_j))$ are either equal or have distinct first coordinates. We

can therefore use polynomial interpolation with these points to construct a polynomial $r \in \mathbb{F}[T]$ satisfying both $r(0) = 0$ and $r(\lambda_i - \lambda_j) = f(\lambda_i) - f(\lambda_j)$ for all pairs of i and j. Therefore,

$$\begin{aligned} r(\mathrm{ad}(s))(e_{i,j}) &= \left(\sum_k r_k \, \mathrm{ad}(s)^k \right) (e_{i,j}) \\ &= \sum_k r_k \, \mathrm{ad}(s)^k (e_{i,j}) \\ &= \sum_k r_k (\lambda_i - \lambda_j)^k e_{i,j} \\ &= \left(\sum_k r_k (\lambda_i - \lambda_j)^k \right) e_{i,j} \\ &= r(\lambda_i - \lambda_j) e_{i,j} \\ &= (f(\lambda_i - \lambda_j)) e_{i,j} \\ &= \mathrm{ad}(y)(e_{i,j}), \end{aligned}$$

for each pair of i and j, hence $r(\mathrm{ad}(s)) = \mathrm{ad}(y)$.

As $\mathrm{ad}(s)$ is the semisimple part of the Jordan decomposition of $\mathrm{ad}(x)$ (Theorem 3.7), $\mathrm{ad}(s)$ can be written as a polynomial in $\mathrm{ad}(x)$ with no constant term (Theorem 3.4). Therefore, $\mathrm{ad}(y)$ is also a polynomial in $\mathrm{ad}(x)$ without constant term. Then, by the assumption of the theorem, we have that $\mathrm{ad}(x)(B) \subseteq A$, hence $\mathrm{ad}(y)(B) \subseteq A$. But this implies that $y \in M$, hence $\mathrm{trace}(xy) = 0$ by assumption.

We can calculate

$$\begin{aligned} (xy)(v_i) &= x(f(\lambda_i) v_i) \\ &= f(\lambda_i) x(v_i) \\ &= f(\lambda_i)(s(v_i) + n(v_i)) \\ &= f(\lambda_i)\lambda_i v_i + f(\lambda_i) n(v_i), \end{aligned}$$

where $n(v_i)$ contains no terms in v_i, as n is nilpotent. Therefore,

$$0 = \mathrm{trace}(xy) = \sum_{i=1}^{m} f(\lambda_i)\lambda_i.$$

As $f(\lambda_i) \in \mathbb{Q}$ and $\lambda_i \in E$, this expression is a $\mathbb{Q}$-linear combination of elements of E, so we can apply f to get

$$\begin{aligned} f\left(\sum_{i=1}^{m} f(\lambda_i)\lambda_i \right) &= \sum_{i=1}^{m} f(f(\lambda_i)\lambda_i) \\ &= \sum_{i=1}^{m} f(\lambda_i) f(\lambda_i) \\ &= \sum_{i=1}^{m} f(\lambda_i)^2, \end{aligned} \tag{4.2.2}$$

which equals $f(0) = 0$. As $f \in E^*$, we have that $f(\lambda_i) \in \mathbb{Q}$ for each $i = 1, \cdots, m$, hence either $f(\lambda_i)^2 > 0$ or $f(\lambda_i) = 0$. This implies that $f(\lambda_i) = 0$ for all $i = 1, \cdots, m$, as expression (4.2.2) equals zero. That is, $f = 0$. As $f \in E^*$ was arbitrary, we have shown that $E^* = \{0\}$. Therefore, by (4.2.1), x is nilpotent. $\square$

4.3. **Proposition.** *Let $x, y, z \in \mathfrak{gl}(V)$. Then*

$$\operatorname{trace}([xy]z) = \operatorname{trace}(x[yz]).$$

Proof. From the definition of the Lie bracket in $\mathfrak{gl}(V)$, we have

$$[xy]z = (xy - yx)z = xyz - yxz$$

and

$$x[yz] = x(yz - zy) = xyz - xzy.$$

Therefore, using the fact that $\operatorname{trace}(ab) = \operatorname{trace}(ba)$ for all $a, b \in \mathfrak{gl}(V)$, in addition to the fact that trace is linear, we can calculate

$$\begin{aligned}
\operatorname{trace}([xy]z) &= \operatorname{trace}(xyz - yxz) \\
&= \operatorname{trace}(xyz) - \operatorname{trace}(y(xz)) \\
&= \operatorname{trace}(xyz) - \operatorname{trace}((xz)y) \\
&= \operatorname{trace}(xyz - xzy) \\
&= \operatorname{trace}(x[yz]).
\end{aligned}$$

□

4.4. **Theorem** (Cartan's Criterion)**.** *Let L be a Lie subalgebra of $\mathfrak{gl}(V)$. Suppose that $\operatorname{trace}(xy) = 0$ for all $x \in [LL]$ and $y \in L$. Then L is solvable.*

Proof. By Lemma 4.2, if we let $A = [LL]$ and $B = L$, then

$$M = \{x \in \mathfrak{gl}(V) : [xL] \subseteq [LL]\}$$

and all $x \in M$ which satisfy $\operatorname{trace}(xy) = 0$ for all $y \in M$ must be nilpotent.

Let $x \in [LL]$. Then $x = [x_1 x_2]$ for some $x_1, x_2 \in L$. Let $z \in M$. Then $[zL] \subseteq [LL]$ by definition, hence $[zx_1], [zx_2] \in [LL]$. Then the supposition implies that $\operatorname{trace}([zx_1]x_2) = 0$. Therefore (by Proposition 4.3),

$$0 = \operatorname{trace}([zx_1]x_2) = \operatorname{trace}(z[x_1x_2]) = \operatorname{trace}(zx) = \operatorname{trace}(xz).$$

As $x \in [LL] \subseteq L$, we have $[xL] \subseteq [LL]$, hence $x \in M$ by definition. Therefore, as $z \in M$ was arbitrary, we have that x is nilpotent. As $x \in [LL]$ was arbitrary, this implies that all $x \in [LL]$ are nilpotent, hence $[LL]$ is nilpotent by Engel's Theorem. Nilpotency implies solvability (Lemma 1.26), hence $[LL]$ is solvable, which implies that L is solvable. □

4.5. **Corollary.** *Suppose that $\operatorname{trace}(\operatorname{ad}(x)\operatorname{ad}(y)) = 0$ for all $x \in [LL]$ and $y \in L$. Then L is solvable.*

Proof. Consider the adjoint representation of L,

$$\operatorname{ad} : L \to \mathfrak{gl}(L) : x \mapsto \operatorname{ad}(x).$$

As ad is a homomorphism, we have that $\operatorname{ad}(L)$ is a Lie subalgebra of $\mathfrak{gl}(L)$, and that $[\operatorname{ad}(L)\operatorname{ad}(L)] = \operatorname{ad}([LL])$. That is, $x \in [LL]$ for all $\operatorname{ad}(x) \in [\operatorname{ad}(L)\operatorname{ad}(L)]$. Therefore, by assumption, $\operatorname{trace}(\operatorname{ad}(x)\operatorname{ad}(y)) = 0$ for all $\operatorname{ad}(x) \in [\operatorname{ad}(L)\operatorname{ad}(L)]$ and $\operatorname{ad}(y) \in \operatorname{ad}(L)$. Then (Theorem 4.4) implies that $\operatorname{ad}(L)$ is solvable.

We have that $\ker(\operatorname{ad}) = Z(L)$ (Remark 1.17). Thus $\ker(\operatorname{ad}) = Z(L)$ is a solvable ideal of L (Lemma 1.24 and Lemma 1.8). Further, $L/\ker(\operatorname{ad}) \cong \operatorname{ad}(L)$, which we have shown is solvable. Therefore, Proposition 1.29 implies that L is solvable. □

5. Toral Subalgebras

We introduce a type of Lie subalgebra called a torus. Tori play a central role in the structure of Lie algebras. We show that nonzero tori always exist in semisimple Lie algebras, and in the following section, we use the properties of tori to decompose semisimple L around a maximal torus.

5.1. **Definition.** Let T be a nonzero Lie subalgebra of L. If all elements of T are semisimple (Definition 2.1), then T is called a **toral subalgebra** or a **torus**.

5.2. **Proposition.** *Let T be a toral subalgebra of L. Then T is Abelian.*

Proof. T is Abelian if $[tt'] = 0$ for all $t, t' \in T$. That is, if $\mathrm{ad}_T(t) = 0$ for all $t \in T$. This is the case if $\mathrm{ad}_T(t)$ has all zero eigenvalues.

Let $t \in T$ and suppose (for a contradiction) $\mathrm{ad}_T(t)$ has a nonzero eigenvalue $\lambda \in \mathbb{F}$ with eigenvector $s \in T$. That is,

$$\mathrm{ad}_T(t)(s) = \lambda s \neq 0, \tag{5.2.1}$$

hence

$$\mathrm{ad}_T(s)(t) = -\mathrm{ad}_T(t)(s) = -\lambda s,$$

which implies that

$$\mathrm{ad}_T(s)([st]) = \mathrm{ad}_T(s)(-\lambda s) = -\lambda \,\mathrm{ad}_T(s)(s) = 0, \tag{5.2.2}$$

hence $[st] \in T$ is an eigenvector of $\mathrm{ad}_T(s)$ with eigenvalue 0.

On the other hand, $s \in T$, hence is semisimple, hence $\mathrm{ad}_T(s)$ is diagonalisable. Therefore there exists an $\mathrm{ad}_T(s)$-eigenbasis for T, say $e_1, \cdots, e_n$ with corresponding eigenvalues $\lambda_1, \cdots, \lambda_n$. t can then be expressed as

$$t = \sum_{i=1}^{n} \mu_i e_i,$$

for some set of coefficients $\mu_i \in \mathbb{F}$, with at least one $\mu_i \neq 0$.Then

$$[st] = \mathrm{ad}_T(s)\left(\sum_{i=1}^{n} \mu_i e_i\right) = \sum_{i=1}^{n} \mu_i \,\mathrm{ad}_T(s)(e_i) = \sum_{i=1}^{n} \mu_i \lambda_i e_i,$$

hence

$$\mathrm{ad}_T(s)([st]) = \mathrm{ad}_T(s)\left(\sum_{i=1}^{n} \mu_i \lambda_i e_i\right) = \sum_{i=1}^{n} \mu_i \lambda_i \,\mathrm{ad}_T(s)(e_i) = \sum_{i=1}^{n} \mu_i \lambda_i^2 e_i. \tag{5.2.3}$$

If $[st] = 0$, then $[ts] = 0$, which contradicts the supposition (5.2.1), hence $[st] \neq 0$. Therefore, for some $i = 1, \cdots, n$, the the coefficient of e_i is nonzero. That is, $\mu_i \lambda_i \neq 0$, which implies that both μ_i and λ_i are nonzero. Hence $\mu_i \lambda_i^2 \neq 0$, which, when applied to (5.2.3), implies that $\mathrm{ad}_T(s)([st]) \neq 0$. This contradicts (5.2.2), hence the supposition that $\mathrm{ad}_T(t)$ has a nonzero eigenvalue must be false. As $\mathrm{ad}_T(t)$ is diagonalisable and has all eigenvalues equal 0, $\mathrm{ad}_T(t) = 0$. As $t \in T$ was arbitrary, we have that $\mathrm{ad}_T(t) = 0$ for all $t \in T$. That is, $[tt'] = 0$ for all $t, t' \in T$. □

5.3. **Definition.** The **radical** of L, denoted $\mathrm{Rad}(L)$, is the maximal solvable ideal of L. This is unique, due to the fact that sums of solvable ideals are also solvable ideals. We call L **semisimple** (not to be confused with semisimple elements of L) if L is nonzero and $\mathrm{Rad}(L) = \{0\}$.

5.4. **Proposition.** *Let L be semisimple. Then $Z(L) = \{0\}$.*

Proof. $Z(L)$ is an ideal of L (Lemma 1.8) and is solvable (Lemma 1.24), hence is contained in $\mathrm{Rad}(L) = \{0\}$. □

5.5. If L is simple, then $\mathrm{Rad}(L)$ is either L or zero. If $\mathrm{Rad}(L) = L$, then L is solvable, hence $[LL] \neq L$. But this contradicts simplicity (Remark 1.10), hence $\mathrm{Rad}(L) = \{0\}$.

That is, if L is simple, then L is also semisimple.

5.6. **Lemma.** *Let L be semisimple. Then ad_L is injective.*

Proof. Apply Proposition 5.4 to Remark 1.17. □

5.7. **Lemma.** *Let L be simple. Then L is not nilpotent.*

Proof. As L is simple, we have that $\mathrm{ad}(L)(L) = [LL] = L$ (Remark 1.10), hence

$$\mathrm{ad}(L)^n(L) = \mathrm{ad}(L)^{n-1}(L) = \cdots = \mathrm{ad}(L)(L) = L \neq \{0\},$$

for all $n \in \mathbb{N}$, hence L cannot be nilpotent. □

5.8. **Lemma.** *Let L be semisimple. Then L is not nilpotent.*

Proof. Let I be a maximal ideal of L. As L is semisimple, $\mathrm{Rad}(L) = \{0\}$, hence either $I = \{0\}$ or I is not solvable. If $I = \{0\}$, then L is simple, hence Lemma 5.7 implies that L is not nilpotent.

Otherwise, I is not solvable. As I is an ideal, we have that $I \subseteq [LI] \subseteq [LL]$. Together these imply that $[LL]$ is not solvable. Nilpotency implies solvability (Lemma 1.26), hence non-solvability implies non-nilpotency. Therefore $[LL]$ is not nilpotent and so L must not be nilpotent. □

5.9. **Theorem.** *Let L be semisimple. Then L contains a nonzero toral subalgebra.*

Proof. Lemma 5.8 implies that L is not nilpotent. Therefore, by Engel's Theorem (1.21), not all $x \in L$ can be nilpotent. That is, there exists some $x \in L$, where if $x = x_n + x_s$ is its Jordan decomposition, then $x_s \neq 0$. As $x_s \in L$ is semisimple, this implies that $\mathrm{span}\{x_s\}$ is a nonzero subalgebra of L consisting of semisimple elements: a torus. □

5.10. **Definition.** If T is a toral subalgebra of L and there exists no other toral subalgebra T' such that $T \subset T'$, then T is called a **maximal toral subalgebra** or **maximal torus**.

5.11. **Example.** We return to our toy example of $L = \mathfrak{sl}_2(\mathbb{F})$. We want to prove that L is simple and find a maximal torus.

Suppose L has an ideal I. Recall the multiplication calculated in Example 1.30:

$$[xy] = h, \quad [hx] = 2x, \quad [hy] = -2y.$$

We then have

$$\begin{aligned}
[Lx] &= \mathrm{span}\{[xx], [yx], [hx]\} = \mathrm{span}\{-h, 2x\} = \mathbb{F}h \oplus \mathbb{F}x, \\
[Ly] &= \mathrm{span}\{[xy], [yy], [hy]\} = \mathrm{span}\{h, -2y\} = \mathbb{F}h \oplus \mathbb{F}y, \\
[Lh] &= \mathrm{span}\{[xh], [yh], [hh]\} = \mathrm{span}\{-2x, 2y\} = \mathbb{F}x \oplus \mathbb{F}y.
\end{aligned}$$

Therefore, if $h \in I$, we must have $x, y \in I$, hence $I = L$. Further, if either x or y is in I, then $h \in I$. That is, if any one of x, y or h (or any nonzero multiple of these) is contained in I, then $I = L$.

Suppose $ax + bh \in I$ for nonzero $a, b \in \mathbb{F}$. Then

$$[h(ax + bh)] = a[hx] = 2ax \in I,$$

hence $x \in I$. Similarly, if we suppose $ay + bh \in I$ for nonzero $a, b \in \mathbb{F}$, then

$$[h(ay + bh)] = a[hy] = -2ay \in I,$$

hence $y \in I$. Suppose $ax + by \in I$ for nonzero $a, b \in \mathbb{F}$. Then

$$[x(ax + by)] = b[xy] = bh \in I,$$

hence $h \in I$.

We have therefore shown that if $ax + by + ch \in I$ and any of a, b, or c are zero, then $I = L$. So suppose that $ax + by + ch \in I$ for nonzero $a, b, c \in \mathbb{F}$. Then

$$[h(ax + by + ch)] = a[hx] + b[hy] = 2ax - 2by \in I,$$

which by the above, implies that $I = L$ again.

Therefore, the only nonzero ideal of L is L itself. That is, L is simple.

We now want to find a maximal toral subalgebra of L. We can see that h is semisimple, as $\mathrm{ad}(h)$ maps x to $2x$ and y to $-2y$. Let T be the maximal toral subalgebra containing $\mathbb{F}h$. Tori are Abelian (Proposition 5.2), so the fact that x and y do not commute with h, implies that $x, y \notin T$. Further, as $[hx] = 2x$ and $[hy] = -2y$, there cannot exist nonzero scalars $a, b \in \mathbb{F}$ such that $ax+by$ commutes with h. Therefore, as $[h(ax+by+ch)] = [h(ax+by)]$, there cannot exist nonzero scalars $a, b, c \in \mathbb{F}$ such that $ax + by + ch$ commutes with h. That is, $T = \mathbb{F}h$ is a maximal toral subalgebra of L.

6. Root Space Decomposition

Tori are Abelian, hence their adjoint maps commute. Recall that commuting diagonalisable maps are simultaneously diagonalisable. So, given a torus, we can find a basis for L on which the entire torus acts diagonally. This enables us to decompose the Lie algebra into its maximal torus and a set of root spaces (defined below), parametrised by elements in the dual of this torus, called roots. These roots and their corresponding root spaces will turn out to have various interesting properties, and the roots will turn out to form a structure known as a root system (14.1).

In this section and all subsequent sections, we fix L to be a semisimple Lie algebra and T to be a maximal toral subalgebra of L, which exists by Theorem 5.9. We denote the dual space of T by T^*.

6.1. **Definition.** For $\alpha \in T^*$, we write

$$L_\alpha = \{x \in L : \forall t \in T : [tx] = \alpha(t)x\}.$$

If $\alpha \neq 0$ and $L_\alpha \neq \{0\}$, then we call α a **root** and L_α a **root space** of L. Note that L_α is indeed a subspace of L, as both the Lie bracket and α are linear.

The **set of roots** is denoted by Φ. That is,

$$\Phi = \{\alpha \in T^* \setminus \{0\} : L_\alpha \neq \{0\}\}.$$

6.2**.** Note that

$$L_0 = \{x \in L : \forall t \in T : [tx] = 0\} = C_L(T),$$

where $C_L(T)$ is the centraliser of T in L. As T is Abelian, we have $[tx] = 0$ for all $t, x \in T$, hence $T \subseteq L_0 = C_L(T)$.

6.3. **Proposition.** *Root spaces are all linearly independent subspaces. Specifically, if $\alpha, \beta \in \Phi \cup \{0\}$, then*

$$L_\alpha \cap L_\beta \neq \{0\} \iff \alpha = \beta.$$

Proof. Let $x_\beta \in L_\beta$ for some $\beta \in \Phi$. Suppose we have some set of linearly independent root spaces $L_{\alpha_1}, \cdots, L_{\alpha_n}$ with $\beta \notin \{\alpha_1, \cdots, \alpha_n\}$. Suppose also, that

$$x_\beta \in \bigoplus_{i=1}^{n} L_{\alpha_i}.$$

That is,

$$x_\beta = \sum_{i=1}^{n} \lambda_i x_i$$

for some $\lambda_i \in \mathbb{F}$ and some nonzero $x_i \in L_{\alpha_i}$. Then for all $t \in T$, we have $[tx_\beta] = \beta(t)x_\beta$, hence

$$\left[t\left(\sum_{i=1}^{n} \lambda_i x_i\right)\right] = \beta(t)\left(\sum_{i=1}^{n} \lambda_i x_i\right) = \sum_{i=1}^{n} \lambda_i \beta(t) x_i.$$

On the other hand, $[tx_i] = \alpha_i(t)x_i$, hence

$$\left[t\left(\sum_{i=1}^{n} \lambda_i x_i\right)\right] = \sum_{i=1}^{n} \lambda_i [tx_i] = \sum_{i=1}^{n} \lambda_i \alpha_i(t) x_i.$$

The x_i are all linearly independent by supposition, hence $\lambda_i \beta(t) = \lambda_i \alpha_i(t)$. As this holds for all $t \in T$, the supposition that $\beta \neq \alpha_i$ implies that $\lambda_i = 0$. This holds for all

$i = 1, \cdots, n$, hence $x_\beta = 0$. That is,

$$L_\beta \cap \bigoplus_{i=1}^{n} L_{\alpha_i} = \{0\},$$

hence $L_\beta, L_{\alpha_1}, \cdots, L_{\alpha_n}$ are all linearly independent.

Extending inductively from $n = 1$, we get that all root spaces are linearly independent. □

6.4. **Theorem.** *L can be expressed as*

$$L = \bigoplus_{\alpha \in T^*} L_\alpha = L_0 \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

Proof. Let $s, t \in T$. As T is Abelian (Proposition 5.2), $[st] = 0$. Then by the Jacobi identity, for all $x \in L$,

$$[s[tx]] + [x[st]] + [t[xs]] = 0,$$

hence

$$[s[tx]] = -[t[xs]] = [t[sx]].$$

That is,

$$(\mathrm{ad}(s) \circ \mathrm{ad}(t))(x) = (\mathrm{ad}(t) \circ \mathrm{ad}(s))(x),$$

hence the $\mathrm{ad}(t) \in \mathrm{End}(L)$ for $t \in T$ commute. We can therefore invoke Theorem 2.6: there exists a shared eigenbasis of L for the maps $\{\mathrm{ad}(t) : t \in T\}$, say $e_1, \cdots, e_n$. That is, a basis on which all these maps diagonalise simultaneously. Therefore, for each $i = 1, \cdots, n$,

$$\forall t \in T : \mathrm{ad}(t)(e_i) = \alpha_i(t) e_i,$$

where $\alpha_i : T \to \mathbb{F}$ is a function mapping $t \in T$ to the eigenvalue of $\mathrm{ad}(t)$ corresponding to the eigenvector e_i. As ad is linear, each α_i is linear. So each $\alpha_i \in T^*$ and satisfies $[te_i] = \alpha_i(t) e_i$ for all $t \in T$, hence $e_i \in L_{\alpha_i}$. Linearity of the Lie bracket then implies that $\mathbb{F}e_i \subseteq L_{\alpha_i}$. The α_i are not necessarily distinct, but each $\alpha \in T^*$ gives distinct L_α (Proposition 6.3), hence

$$L = \bigoplus_{\alpha \in T^*} L_\alpha.$$

By definition, the nonzero L_α in this sum either satisfy $\alpha = 0$ or $\alpha \in \Phi$. Therefore,

$$L = L_0 \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

□

6.5. The decomposition of L in Theorem 6.4 is called the **root space decomposition** of L. We have that $T \subseteq L_0$ by Remark 6.2. It will later turn out that $T = L_0$.

6.6. **Proposition.** *Let $\alpha, \beta \in T^*$. Then $[L_\alpha L_\beta] \subseteq L_{\alpha+\beta}$.*

Proof. Let $x \in L_\alpha$ and $y \in L_\beta$. For all $t \in T$, we then have $[tx] = \alpha(t)x$ and $[ty] = \alpha(t)y$. From the Jacobi identity (1.2.3), we have

$$\begin{aligned}
[t[xy]] &= [x[ty]] - [y[tx]] \\
&= [x\,(\beta(t)y)] - [y\,(\alpha(t)x)] \\
&= \beta(t)[xy] - \alpha(t)[yx] \\
&= \beta(t)[xy] + \alpha(t)[xy] \\
&= (\alpha(t) + \beta(t))[xy]
\end{aligned}$$

$$= (\alpha + \beta)(t)[xy],$$

which implies that $[xy] \in L_{\alpha+\beta}$. □

6.7. One specific case of the above result is when $\beta = -\alpha$. We then have $[L_\alpha L_{-\alpha}] \subseteq L_0$.

6.8. **Example.** We continue our toy example of $L = \mathfrak{sl}_2(\mathbb{F})$. Recall from Example 5.11 that $T = \mathbb{F}h$ was a maximal torus of L. We can see that $[hx] = 2x$, so if we let $\alpha \in T^*$ be the map sending h to 2, then we have $x \in L_\alpha$. Further, $[hy] = -2y = (-\alpha)(h)y$, hence $y \in L_{-\alpha}$. We thus have $\Phi = \{\alpha, -\alpha\}$ and

$$L = T \oplus L_\alpha \oplus L_{-\alpha},$$

where $T = \mathbb{F}h$, $L_\alpha = \mathbb{F}x$ and $L_{-\alpha} = \mathbb{F}y$.

7. The Killing Form

We introduce the Killing form: a symmetric bilinear form on L. This form will play a key role in all to come. Here, we will show that a semisimple Lie algebra has a nondegenerate Killing form. Further, we show the restriction of the Killing form to $L_0 = C_L(T)$ is also nondegenerate. This result underpins the following section. We conclude with a result about the non-orthogonality of certain root spaces.

We continue to take L to be semisimple, T to be a maximal torus of L and $\mathbb{F}$ to be algebraically closed with characteristic zero.

7.1. **Definition.** Let L be a Lie algebra over $\mathbb{F}$. The **Killing form**, κ, of L is a symmetric bilinear form

$$\kappa : L \times L \to \mathbb{F} : (x, y) \mapsto \langle x, y \rangle ,$$

where

$$\langle x, y \rangle = \operatorname{trace}(\operatorname{ad}(x) \operatorname{ad}(y)).$$

The **radical** of of the Killing form (not to be confused with the radical of L) is defined

$$\operatorname{Rad}(\kappa) = \{x \in L : \forall y \in L : \langle x, y \rangle = 0\} . \tag{7.1.1}$$

A bilinear form is called **nondegenerate** if its radical is zero.

7.2. **Proposition.** *The Killing form of a Lie algebra is associative with respect to the Lie bracket.*

Proof. Let $x, y, z \in L$. The maps $\operatorname{ad}(x), \operatorname{ad}(y), \operatorname{ad}(z)$ are linear, hence contained in $\mathfrak{gl}(L)$. Therefore, by Proposition 4.3, we have

$$\operatorname{trace}([\operatorname{ad}(x) \operatorname{ad}(y)] \operatorname{ad}(z)) = \operatorname{trace}(\operatorname{ad}(x)[\operatorname{ad}(y) \operatorname{ad}(z)]).$$

Further, the map $\operatorname{ad} : L \to \operatorname{Der} L$ is a Lie algebra homomorphism (Lemma 1.16), so

$$\operatorname{ad}([ab]) = [\operatorname{ad}(a) \operatorname{ad}(b)],$$

for all $a, b \in L$.

Therefore,

$$\begin{aligned} \langle [xy], z \rangle &= \operatorname{trace}(\operatorname{ad}([xy]) \operatorname{ad}(z)) \\ &= \operatorname{trace}([\operatorname{ad}(x) \operatorname{ad}(y)] \operatorname{ad}(z)) \\ &= \operatorname{trace}(\operatorname{ad}(x)[\operatorname{ad}(y) \operatorname{ad}(z)]) \\ &= \operatorname{trace}(\operatorname{ad}(x) \operatorname{ad}([yz])) \\ &= \langle x, [yz] \rangle . \end{aligned}$$

□

7.3. **Lemma.** *The radical of the Killing form is an ideal of L.*

Proof. $\operatorname{Rad}(\kappa)$ is an ideal if $[L \operatorname{Rad}(\kappa)] \subseteq \operatorname{Rad}(\kappa)$. So let $x \in L$ and $r \in \operatorname{Rad}(\kappa)$ be arbitrary. Then Proposition 7.2 implies that, for all $y \in L$,

$$\langle [xr], y \rangle = - \langle [rx], y \rangle = - \langle r, [xy] \rangle = 0,$$

hence $[xr] \in \operatorname{Rad}(\kappa)$, so we are done. □

7.4. **Proposition.** *Let $\alpha, \beta \in T^*$ such that $\alpha + \beta \neq 0$. Then $\langle L_\alpha, L_\beta \rangle = \{0\}$.*

Proof. Let $x \in L_\alpha$ and $y \in L_\beta$. As $\alpha + \beta \neq 0$, there must exist some $t \in T$ such that

$$(\alpha + \beta)(t) \neq 0. \tag{7.4.1}$$

Then $[tx] = \alpha(t)x$ and $[ty] = \beta(t)y$. We have that $\langle [xt], y\rangle = \langle x, [ty]\rangle$ (Proposition 7.2), hence

$$\begin{aligned} 0 &= \langle x, [ty]\rangle - \langle [xt], y\rangle \\ &= \langle x, [ty]\rangle + \langle [tx], y\rangle \\ &= \langle x, \beta(t)y\rangle + \langle \alpha(t)x, y\rangle \\ &= \beta(t)\langle x, y\rangle + \alpha(t)\langle x, y\rangle \\ &= (\alpha(t) + \beta(t))\langle x, y\rangle \\ &= (\alpha + \beta)(t)\langle x, y\rangle . \end{aligned}$$

Therefore either $(\alpha + \beta)(t) = 0$ or $\langle x, y\rangle = 0$; the former contradicts (7.4.1), hence the latter must be true. As $x \in L_\alpha$ and $y \in L_\beta$ were arbitrary, this implies $\langle L_\alpha, L_\beta\rangle = \{0\}$. □

7.5. **Proposition.** *Let L be semisimple. Then the Killing form, κ, is nondegenerate.*

Proof. By definition of semisimple,

$$\mathrm{Rad}(L) = \{0\}, \tag{7.5.1}$$

where $\mathrm{Rad}(L)$ is the maximal solvable subalgebra of L.

Let $S = \mathrm{Rad}(\kappa)$. Then by definition, for all $x \in S$ and $y \in L$,

$$\langle x, y\rangle = \mathrm{trace}(\mathrm{ad}(x)\,\mathrm{ad}(y)) = 0.$$

More specifically (as $[SS] \subseteq L$), for all $x \in S$ and $y \in [SS]$,

$$\mathrm{trace}(\mathrm{ad}(x)\,\mathrm{ad}(y)) = 0.$$

Therefore Corollary 4.5 implies S is solvable. S is an ideal of L (Lemma 7.3), hence is contained in the maximal ideal $\mathrm{Rad}(L)$. But $\mathrm{Rad}(L) = \{0\}$, so $S = \mathrm{Rad}(\kappa) = \{0\}$. Therefore the Killing form is nondegenerate. □

7.6. **Proposition.** *Let L be semisimple. Then the restriction of the Killing form to L_0 is nondegenerate.*

Proof. By Proposition 7.5, the Killing form is nondegenerate on L. By definition, $0 \notin \Phi$, so by Proposition 7.4, for all $\alpha \in \Phi$,

$$\langle L_0, L_\alpha\rangle = \{0\}. \tag{7.6.1}$$

Let κ_0 be the restriction of κ to L_0 and let $x \in \mathrm{Rad}(\kappa_0)$. Then, for all $y_0 \in L_0$,

$$\langle x, y_0\rangle = 0. \tag{7.6.2}$$

Further, by (7.6.1), for all $\alpha \in \Phi$ and $y_\alpha \in L_\alpha$,

$$\langle x, y_\alpha\rangle = 0. \tag{7.6.3}$$

Recall that we have the decomposition (Theorem 6.4)

$$L = L_0 \oplus \bigoplus_{\alpha\in\Phi} L_\alpha,$$

hence any $y \in L$ can be expressed as

$$y = y_0 + \sum_{\alpha\in\Phi} y_\alpha$$

for some $y_0 \in L_0$ and collection of $y_\alpha \in L_\alpha$. Therefore, by (7.6.2) and (7.6.3), for all $y \in L$,

$$\begin{aligned}\langle x, y\rangle &= \left\langle x, y_0 + \sum_{\alpha\in\Phi} y_\alpha \right\rangle \\ &= \langle x, y_0\rangle + \sum_{\alpha\in\Phi} \langle x, y_\alpha\rangle \\ &= 0.\end{aligned}$$

That is, $x \in \mathrm{Rad}(\kappa) = \{0\}$. As $x \in \mathrm{Rad}(\kappa_0)$ was arbitrary, we have $\mathrm{Rad}(\kappa_0) = \{0\}$, hence κ_0 is nondegenerate. □

7.7. **Proposition.** *Let $\alpha \in \Phi$. Then $\langle L_\alpha, L_{-\alpha}\rangle = \mathbb{F} \neq \{0\}$. Specifically, $\langle x, L_{-\alpha}\rangle = \mathbb{F}$ for all $x \in L_\alpha \setminus \{0\}$.*

Proof. Firstly, note that the first assertion follows from the second, as $\alpha \in \Phi$ implies that $L_\alpha \neq \{0\}$.

Let $x \in L_\alpha$. Suppose $\langle x, L_{-\alpha}\rangle = \{0\}$. By Proposition 7.4, we have that, for all $\beta \in T^*$,

$$\beta \neq -\alpha \implies \langle x, L_\beta\rangle = \{0\}.$$

But, by the supposition, we also have

$$\beta = -\alpha \implies \langle x, L_\beta\rangle = \{0\}.$$

Therefore $\langle x, L_\beta\rangle = \{0\}$ for all $\beta \in T^*$. But by (Theorem 6.4),

$$L = \bigoplus_{\beta\in T^*} L_\beta,$$

therefore $\langle x, L\rangle = \{0\}$. But as the Killing form is nondegenerate (Proposition 7.5), this implies that $x = 0$. As x was arbitrary, this shows that $L_\alpha = \{0\}$, which contradicts the assumption that $\alpha \in \Phi$. Therefore, the supposition that $\langle x, L_{-\alpha}\rangle = \{0\}$ must be false.

As $\langle x, L_{-\alpha}\rangle \neq \{0\}$, linearity implies that $\langle x, L_{-\alpha}\rangle = \mathbb{F}$: take some $y \in L_{-\alpha}$ such that $\langle x, y\rangle = \lambda \neq 0$, then

$$\mathbb{F} = \mathbb{F}\lambda = \mathbb{F}\langle x, y\rangle = \langle x, \mathbb{F}y\rangle \subseteq \langle x, L_{-\alpha}\rangle \subseteq \mathbb{F}.$$

□

7.8. **Corollary.** *If $\alpha \in \Phi$ then $-\alpha \in \Phi$.*

Proof. Suppose $-\alpha \notin \Phi$. Then $L_{-\alpha} = \{0\}$ by definition, hence $\langle L_\alpha, L_{-\alpha}\rangle = \{0\}$. But this contradicts Proposition 7.7, so the supposition is false and $\alpha \in \Phi$. □

8. The Maximal Torus

Our goal in this section is to prove that $T = L_0$. From this, we will obtain two useful results: that the Killing form is nondegenerate when restricted to T, and that the root space decomposition of L (Theorem 6.4) is centred around T.

We continue to take L to be semisimple, T to be a maximal torus of L and $\mathbb{F}$ to be algebraically closed with characteristic zero.

8.1. **Lemma.** *Let V be a finite dimensional vector space over $\mathbb{F}$ and $x, y \in \mathrm{End}(V)$ such that y is nilpotent and $xy = yx$. Then xy is nilpotent and* $\mathrm{trace}(xy) = 0$.

Proof. As y is nilpotent, there exists some $n \in \mathbb{N}$ such that $y^n = 0$. Therefore, the commutativity of x and y implies that $(xy)^n = x^n y^n = x^n \cdot 0 = 0$. That is, xy is nilpotent.

The trace of xy is equal to the constant term in the characteristic polynomial of xy. But as xy is nilpotent, this is zero. □

8.2. **Lemma.** *Let $x, y \in L$ such that y is nilpotent and $[xy] = 0$. Then $\langle x, y\rangle = 0$.*

Proof. By the definition of the Lie bracket on $\mathfrak{gl}(L)$, the fact that ad is a Lie algebra homomorphism (Lemma 1.16) and the assumption that $[xy] = 0$, we have

$$\begin{aligned}\mathrm{ad}(x)\,\mathrm{ad}(y) - \mathrm{ad}(y)\,\mathrm{ad}(x) &= [\mathrm{ad}(x)\,\mathrm{ad}(y)] \\ &= \mathrm{ad}([xy]) \\ &= \mathrm{ad}(0) \\ &= 0.\end{aligned}$$

Therefore, $\mathrm{ad}(x)\,\mathrm{ad}(y) = \mathrm{ad}(y)\,\mathrm{ad}(x)$. That is, the maps $\mathrm{ad}(x)$ and $\mathrm{ad}(y)$ commute. Further, as y is nilpotent, $\mathrm{ad}(y)$ is nilpotent, by definition.

So $\mathrm{ad}(x)$ and $\mathrm{ad}(y)$ satisfy the criteria for Lemma 8.1, hence

$$\langle x, y\rangle = \mathrm{trace}(\mathrm{ad}(x)\,\mathrm{ad}(y)) = 0.$$

□

8.3. **Lemma.** *Let $x \in Z(L_0)$ be nilpotent. Then $x = 0$.*

Proof. Let $y \in L_0$. As $x \in Z(L_0)$, we have $[xy] = 0$. As x is nilpotent, Lemma 8.2 implies $\langle x, y\rangle = 0$. Thus $\langle x, L_0\rangle = \{0\}$, as $y \in L_0$ was arbitrary, hence x is in the radical of the Killing form restricted to L_0. This radical is zero (Proposition 7.6), hence $x = 0$. □

8.4. **Lemma.** *Let $x \in L_0$. If $x = x_n + x_s$ is its Jordan decomposition, then $x_n, x_s \in L_0$.*

Proof. Recall that $L_0 = \{y \in L : \forall t \in T : [yt] = 0\}$. So by definition,

$$\mathrm{ad}(x)(T) = \{0\}\,. \tag{8.4.1}$$

The Jordan decomposition of $\mathrm{ad}(x)$ in $\mathfrak{gl}(L)$ is

$$\mathrm{ad}(x) = \mathrm{ad}(x_n) + \mathrm{ad}(x_s).$$

One property of Jordan decomposition (Theorem 3.4) is that if $A \subseteq B \subseteq L$ are subspaces, then

$$\mathrm{ad}(x)(B) \subseteq A \implies \begin{cases}\mathrm{ad}(x_s)(B) \subseteq A, \\ \mathrm{ad}(x_n)(B) \subseteq A.\end{cases}$$

Applying this to $A = \{0\}$ and $B = T$, we obtain $\mathrm{ad}(x_s)(T) = \{0\}$ and $\mathrm{ad}(x_n)(T) = \{0\}$. Therefore $x_n, x_s \in L_0$. □

8.5. **Lemma.** *Let $x \in L_0$ be semisimple. Then $x \in T$.*

Proof. Let $t \in T$ and $y = \lambda x$ for some $\lambda \in \mathbb{F}$. Then both t and y are semisimple. Further, $[yt] = 0$, as $y \in L_0$.

Semisimplicity is, by definition, equivalent to ad-semisimplicity, hence $\mathrm{ad}(y)$ and $\mathrm{ad}(t)$ are semisimple. As $[yt] = 0$, the maps $\mathrm{ad}(y)$ and $\mathrm{ad}(t)$ commute (Lemma 1.18). Sums of commuting semisimple maps are semisimple (Lemma 3.1), thus $(\mathrm{ad}(y)+\mathrm{ad}(t)) = \mathrm{ad}(y+t)$ is semisimple. Therefore $y + t$ is semisimple.

As $t \in T$ and $y \in \mathbb{F}x$ were arbitrary, we have just shown that the space $T + \mathbb{F}x$ consists entirely of semisimple elements.

As $x \in L_0$, we have that $[xT] = \{0\}$. Therefore (and as T is Abelian),

$$\begin{aligned}[(T + \mathbb{F}x)(T + \mathbb{F}x)] &= [TT] + [T\,\mathbb{F}x] + [\mathbb{F}x\,T] + [\mathbb{F}x\,\mathbb{F}x] \\ &= [TT] + \mathbb{F}[xT] + \mathbb{F}[xx] \\ &= \{0\} + \mathbb{F}\,\{0\} + \mathbb{F}\,\{0\} \\ &= \{0\} \subseteq T + \mathbb{F}x.\end{aligned}$$

Therefore, $T + \mathbb{F}x$ is a Lie subalgebra of L. Further, as $T + \mathbb{F}x$ consists entirely of semisimple elements, it is a toral subalgebra. Clearly $T \subseteq T + \mathbb{F}x$. But as T is a maximal toral subalgebra, it cannot be strictly contained by $T + \mathbb{F}x$, which implies $T = T + \mathbb{F}x$ and hence $x \in T$. □

8.6. **Theorem.** *The restriction of the Killing form to T is nondegenerate.*

Proof. Let $x \in L_0$ and $t \in \mathrm{Rad}(\kappa|_T)$. That is, $t \in T$ and $\langle t, T\rangle = \{0\}$. Let $x = x_s + x_n$ be the Jordan decomposition of x. Then $x_s, x_n \in L_0$ by Lemma 8.4. Then as x_s is semisimple, we have $x_s \in T$ by Lemma 8.5. Therefore $\langle t, x_s\rangle = 0$. As $x_n \in L_0$, we have $[x_n T] = \{0\}$. Specifically $[x_n t] = 0$. As x_n is nilpotent, we have $\mathrm{ad}(x_n)$ is nilpotent. Therefore, we can apply Lemma 8.2 to get $\langle t, x_n\rangle = 0$. Therefore, we have

$$\begin{aligned}\langle t, x\rangle &= \langle t, x_s + x_n\rangle \\ &= \langle t, x_s\rangle + \langle t, x_n\rangle \\ &= 0.\end{aligned}$$

As $x \in L_0$ was arbitrary, we have shown that $\langle t, L_0\rangle = \{0\}$ and therefore that t is in the radical of the Killing form restricted to L_0. But this restriction of the Killing form is nondegenerate (Proposition 7.6), hence the radical is zero. Therefore $t = 0$. As t was arbitrary, we have shown that $\mathrm{Rad}(\kappa|_T) = \{0\}$. Therefore, the restriction of the Killing form to T is nondegenerate. □

8.7. **Lemma.** *L_0 is nilpotent.*

Proof. Let $x \in L_0$ and let $x = x_s + x_n$ be its Jordan decomposition. That is, $\mathrm{ad}(x) = \mathrm{ad}(x_s) + \mathrm{ad}(x_n)$ is the Jordan decomposition of $\mathrm{ad}(x)$. As L_0 is a Lie subalgebra of L, we can consider the restrictions of these adjoint maps to L_0:

$$\mathrm{ad}_{L_0}(x) = \mathrm{ad}_{L_0}(x_s) + \mathrm{ad}_{L_0}(x_n).$$

We have $x_s, x_n \in L_0$ by Lemma 8.4 and $x_s \in T$ by Lemma 8.5. Therefore, $[x_s L_0] = \{0\}$. That is, $\mathrm{ad}_{L_0}(x_s) = 0$. Therefore, $\mathrm{ad}_{L_0}(x) = \mathrm{ad}_{L_0}(x_n)$, which implies that $\mathrm{ad}_{L_0}(x)$ is nilpotent.

As $x \in L_0$ was arbitrary, we have shown that all elements of L_0 are ad-nilpotent. We can therefore apply Engel's Theorem (1.21) to get that L_0 is nilpotent. □

8.8. **Lemma.** $T \cap [L_0 L_0] = \{0\}$.

Proof. For $t \in T$ and $x \in L_0$ we have $[tx] = 0$. Therefore, utilizing the associativity of the Killing form (Proposition 7.2), we get

$$\begin{aligned}\langle T, [L_0 L_0]\rangle &= \{\langle t, [xy]\rangle : t \in T;\ x, y \in L_0\} \\ &= \{\langle [tx], y\rangle : t \in T;\ x, y \in L_0\} \\ &= \{\langle 0, y\rangle : y \in L_0\} \\ &= \{0\}.\end{aligned}$$

Let $t \in T \cap [L_0 L_0]$. Then $t \in [L_0 L_0]$, so by the above, $\langle t, T\rangle = \{0\}$. That is, $t \in \mathrm{Rad}(\kappa|_T)$. But $\mathrm{Rad}(\kappa|_T) = \{0\}$ by Theorem 8.6, so $t = 0$. As t was arbitrary, we have shown that $T \cap [L_0 L_0] = \{0\}$. □

8.9. **Lemma.** *L_0 is Abelian.*

Proof. Suppose that L_0 is nonAbelian. That is, $[L_0 L_0] \neq \{0\}$. By Lemma 8.7, L_0 is nilpotent. As $[L_0 L_0]$ is an ideal of L_0 (Lemma 1.6), we can apply Lemma 1.22 to get that $[L_0 L_0] \cap Z(L_0) \neq \{0\}$. So there exists a nonzero $z \in [L_0 L_0] \cap Z(L_0)$. This intersection is contained in L_0, so we have $z \in L_0$.

$$z \in L_0. \tag{8.9.1}$$

Suppose this z is semisimple. Then Lemma 8.5 implies $z \in T$, hence by Lemma 8.8,

$$z \in T \cap [L_0 L_0] \cap Z(L_0) \subseteq T \cap [L_0 L_0] = \{0\}.$$

But z is nonzero, so this is a contradiction. Therefore, the supposition that z is semisimple must be false. That is, if $z = z_n + z_s$ is the Jordan decomposition of z, then $z_n \neq 0$. As $z \in L_0$, Lemma 8.4 implies that $z_n \in L_0$. As $z \in Z(L_0)$, we have $[zL_0] = \{0\}$, hence $\mathrm{ad}(z)(L_0) = 0$. We can then apply Theorem 3.4: if $A \subseteq B \subseteq L$ are subspaces, then

$$\mathrm{ad}(z)(B) \subseteq A \implies \begin{cases}\mathrm{ad}(z_s)(B) \subseteq A, \\ \mathrm{ad}(z_n)(B) \subseteq A.\end{cases}$$

If we let $A = \{0\}$ and $B = L_0$, we have that $\mathrm{ad}(z_n)(L_0) = \{0\}$.

That is $[z_n L_0] = \{0\}$, hence $z_n \in Z(L_0)$. Therefore $z_n = 0$ by Lemma 8.3. But this contradicts the fact that $z_n \neq 0$, so the supposition that L_0 is nonAbelian must be false. □

8.10. **Theorem.** *$T = L_0$.*

Proof. Suppose that $T \neq L_0$. We have that $T \subseteq L_0$ (Remark 6.2), so the supposition implies that there exists $x \in L_0$ such that $x \notin T$. Let $x = x_n + x_s$ be the Jordan decomposition of x. Then $x_s, x_n \in L_0$ by Lemma 8.4, hence $x_s \in T$ by Lemma 8.5. If $x_n \in T$, then $x = x_s + x_n \in T$, contradicting the definition of x. Thus $x_n \notin T$ (specifically, $x_n \neq 0$).

L_0 is Abelian by Lemma 8.9, so $x_n \in Z(L_0) = L_0$. Then Lemma 8.3 implies $x_n = 0$. But this contradicts the fact that $x_n \neq 0$, so the supposition that $T \neq L_0$ must be false. □

8.11. **Corollary.** *$T = L_0$ and the root space decomposition of L can be expressed*

$$L = T \oplus \bigoplus_{\alpha \in \Phi} L_\alpha.$$

Proof. See Theorem 6.4 and Theorem 8.10. □

8.12. By Theorem 8.6, the Killing form is nondegenerate when restricted to T. Therefore, we can express a correspondence between T and T^* as follows: for each $\phi \in T^*$, there exists a unique $t_\phi \in T$ such that $\phi(t) = \langle t_\phi, t \rangle$ for all $t \in T$.

9. Modules and Weights

We introduce the basics of modules and present Weyl's Theorem (which we will take as an assumption: see [4] for a proof). We then introduce the notion of weights and discuss the representation theory of $\mathfrak{sl}_2(\mathbb{F})$. This is important to the larger theory of semisimple Lie algebras, because (it will turn out) these Lie algebras are built up from copies of $\mathfrak{sl}_2(\mathbb{F})$, so we can use results about $\mathfrak{sl}_2(\mathbb{F})$-modules to determine the structure of arbitrary semisimple Lie algebras.

Throughout the section, we take x, y and h to denote the standard basis vectors of $\mathfrak{sl}_2(\mathbb{F})$ (defined below).

9.1. **Definition.** Let L be a Lie algebra over some field $\mathbb{F}$. Then a vector space V over $\mathbb{F}$ is called an **L-module** if it is equipped with an operation

$$L \times V \to V : (a, v) \mapsto a \cdot v,$$

which satisfies

$$\begin{aligned}(\lambda a + \mu b) \cdot v &= \lambda(a \cdot v) + \mu(b \cdot v),\\ a \cdot (\lambda v + \mu w) &= \lambda(a \cdot v) + \mu(a \cdot w),\\ [ab] \cdot v &= a \cdot (b \cdot v) - b \cdot (a \cdot v),\end{aligned}$$

for all $\lambda, \mu \in \mathbb{F}$; $a, b \in L$; $v, w \in V$. Naturally, if S is a subspace of V and also an L-module under the same action, we say S is an **L-submodule** of V.

9.2. **Definition.** Let V be an L-module for some Lie algebra L. Then V is called **irreducible** if there exist precisely two L-submodules of V: itself and $\{0\}$. That is, if V is nontrivial and there exist no proper nontrivial L-submodules of V.

V is called **completely reducible** if V can be expressed as a direct sum of irreducible submodules.

9.3**.** It makes sense to talk about representations being irreducible and completely reducible: we apply the term to a representation $\rho : L \to \mathfrak{gl}(V)$ if the term applies to the L-module induced on V by ρ.

9.4. **Theorem** (Weyl's Theorem)**.** *Let $\phi : L \to \mathfrak{gl}(V)$ be a finite dimensional representation of a semisimple Lie algebra. Then ϕ is completely reducible.*

Equivalently, if V is a finite dimensional module of a semisimple Lie algebra, then V is completely reducible.

Proof. See Theorem 6.3 of [4]. □

9.5. **Lemma.** *Let L be a Lie algebra over some field $\mathbb{F}$ and let V be an L-module. Suppose L acts trivially on V. Then all subspaces of V are L-submodules, and submodules of V are irreducible if and only if they are 1-dimensional.*

Proof. Let $v \in V$. Then $[vL] = \{0\}$ by assumption, hence $\mathbb{F}v$ is an L-submodule of V. As $\mathbb{F}v$ is 1-dimensional, it has no proper nontrivial submodules, hence is irreducible. That is, all 1-dimensional subspaces (and hence submodules) of V are irreducible submodules.

If S is a subspace of V, then S is a direct sum of 1-dimensional subspaces, which are all submodules, hence S is a submodule of V and is irreducible if and only if $\dim(S) = 1$. All submodules are subspaces, so this also implies that all irreducible submodules are 1-dimensional. □

9.6. **Definition.** Recall our example of $\mathfrak{sl}_2(\mathbb{F})$ (Example 1.30). The representation theory of this Lie algebra is of particular importance to us, because we are able to utilise it in the study of arbitrary semisimple Lie algebras.

We have shown that $\mathfrak{sl}_2(\mathbb{F}) = \operatorname{span}\{x, y, h\}$, where

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The Lie bracket on $\mathfrak{sl}_2(\mathbb{F})$ is then given by

$$[xy] = h, \quad [hx] = 2x, \quad [hy] = -2y.$$

9.7. **Definition.** Let V be an $\mathfrak{sl}_2(\mathbb{F})$-module, $\lambda \in \mathbb{F}$ and $V_\lambda = \{v \in V : h \cdot v = \lambda v\}$. When $V_\lambda \neq \{0\}$, call λ a **weight** of h in V and call V_λ a **weight space**.

9.8. **Lemma.** *Let ϕ be a finite dimensional representation of L. Then ϕ maps semisimple elements to semisimple elements.*

Proof. See Corollary 6.4 of [4] □

9.9. **Lemma.** *Let V be an $\mathfrak{sl}_2(\mathbb{F})$-module. Then*

$$V = \bigoplus_{\lambda \in \mathbb{F}} V_\lambda,$$

where V_λ is the weight space of h in V with weight λ.

Proof. As h is semisimple, Lemma 9.8 implies that the representation inducing V maps h to a semisimple element of $\mathfrak{gl}(V)$, which gives a diagonal action of h on V. Therefore, V can be decomposed into eigenspaces of this action, which are just weight spaces of h. □

9.10. **Lemma.** *Let V be an $\mathfrak{sl}_2(\mathbb{F})$-module. Denote weight spaces of h in V by V_μ for $\mu \in \mathbb{F}$. Let $\lambda \in \mathbb{F}$ and $v \in V_\lambda$. Then $x \cdot v \in V_{\lambda+2}$ and $y \cdot v \in V_{\lambda-2}$.*

Proof. We have that $h \cdot v = \lambda v$. From the definition of the module action, we can calculate

$$h \cdot (x \cdot v) = [hx] \cdot v + x \cdot (h \cdot v) = (2x) \cdot v + x \cdot (\lambda v) = (\lambda + 2)x \cdot v,$$

and similarly

$$h \cdot (y \cdot v) = [hy] \cdot v + y \cdot (h \cdot v) = (-2y) \cdot v + y \cdot (\lambda v) = (\lambda - 2)y \cdot v,$$

hence $x \cdot v \in V_{\lambda+2}$ and $y \cdot v \in V_{\lambda-2}$. □

9.11. **Definition.** Let V be an $\mathfrak{sl}_2(\mathbb{F})$-module, $\lambda \in \mathbb{F}$ and V_λ a weight space of h in V. If $v \in V_\lambda$ is annihilated by x, we call v a **maximal vector of weight** λ. As V is finite dimensional, Lemma 9.9 and Lemma 9.10 imply such vectors must exist.

9.12. **Proposition.** *Let V be an irreducible $\mathfrak{sl}_2(\mathbb{F})$-module. Then the action of $\mathfrak{sl}_2(\mathbb{F})$ on V is given by*

- $h \cdot v_i = (\lambda - 2i)v_i$,
- $y \cdot v_i = (i + 1)v_{i+1}$,
- $x \cdot v_i = (\lambda - i + 1)v_{i-1}$, *for $i \geq 0$,*

where v_0 is a maximal vector of weight λ and

$$v_i = \frac{1}{i!} y^i \cdot v_0$$

for each $i \geq 0$.

Proof. Lemma 9.10 implies $y \cdot V_\lambda \subseteq V_{\lambda-2}$, hence $y^i \cdot V_\lambda \subseteq V_{\lambda-2i}$. Therefore, $v_i \in V_{\lambda-2i}$ for each i, hence

$$h \cdot v_i = (\lambda - 2i)v_i.$$

By definition,

$$y \cdot v_i = y \cdot \left(\frac{1}{i!}y^i \cdot v_0\right) = \left(\frac{i+1}{(i+1)!}y^{i+1} \cdot v_0\right) = (i+1)v_{i+1}.$$

We prove the third point using induction on i. Suppose $x \cdot v_i = (\lambda - i + 1)v_{i-1}$, for some i. Then

$$\begin{aligned}
(i+1)x \cdot v_{i+1} &= (i+1)x \cdot \left(\frac{1}{i+1}y \cdot v_i\right) \\
&= x \cdot y \cdot v_i \\
&= [xy] \cdot v_i + y \cdot (x \cdot v_i) \\
&= h \cdot v_i + y \cdot ((\lambda - i + 1)v_{i-1}) \\
&= (\lambda - 2i)v_i + (\lambda - i + 1)(y \cdot v_{i-1}) \\
&= (\lambda - 2i)v_i + (\lambda - i + 1)(iv_i) \\
&= ((\lambda - i) - i + i(\lambda - i) + i)v_i \\
&= (i+1)(\lambda - i)v_i,
\end{aligned}$$

hence $x \cdot v_{i+1} = (\lambda - (i+1) + 1)v_i$.

As v_0 is a maximal vector, $x \cdot v_0 = 0$ (this satisfies the formula under the convention that $v_{-1} = 0$), hence

$$\begin{aligned}
x \cdot v_1 &= x \cdot y \cdot v_0 \\
&= [xy] \cdot v_0 + y \cdot x \cdot v_0 \\
&= h \cdot v_0 + y \cdot (0) \\
&= \lambda v_0.
\end{aligned}$$

Therefore, the base case of $i = 1$ is satisfied and consequently the third point holds for all $i \geq 0$. □

9.13. **Theorem.** *Let V be an irreducible $\mathfrak{sl}_2(\mathbb{F})$-module. Then*

$$V = \bigoplus_{i=0}^{m} \mathbb{F}v_i,$$

where the v_i are defined as in Proposition 9.12.

Proof. By the first point of Proposition 9.12, the weight of h on v_i is $\lambda - 2i$. This implies that the v_i are linearly independent. As V is finite dimensional, there can only be finitely many nonzero v_i. Let $m \in \mathbb{N}$ be the minimum index such that $v_m \neq 0$ and $v_{m+1} = 0$. This implies that $v_i = 0$ for all $i > m$, by definition.

Consider the subspace M of V given by

$$M = \mathbb{F}v_0 \oplus \cdots \oplus \mathbb{F}v_m.$$

As $v_{-1} = 0 = v_{m+1}$, Proposition 9.12 implies that M is an L-submodule of V. Further, M is nonzero, as $v_0 \neq 0$. But V is irreducible, so we must have that $V = M$. □

9.14. **Corollary.** *Let V be an irreducible $\mathfrak{sl}_2(\mathbb{F})$-module. Then:*

- *Each nonzero weight space of V is 1-dimensional and spanned by v_i (as defined in (Proposition 9.12)).*

- *For some $\lambda \in \mathbb{F}$, there exists a maximal vector of weight λ in V. Further, all maximal vectors in V have weight λ.*
- *$\lambda = \dim V - 1$ (and is therefore a nonnegative integer).*
- *The weights of h on V are the values $\lambda, \lambda - 2, \cdots, -\lambda$. Consequently, if μ is a weight, then μ is an integer and $-\mu$ is also a weight.*
- *We can write $V = V_\lambda \oplus V_{\lambda-2} \oplus \cdots \oplus V_{-\lambda}$, where the V_μ are 1-dimensional weight spaces of h.*

Proof. The first point follows immediately from the theorem. This then implies that the maximal vectors in V are precisely the vectors in $\mathbb{F}v_0$, which gives us the second point.

The theorem gives us that $\dim V = m + 1$. By the third point of Proposition 9.12, we have that

$$x \cdot v_{m+1} = (\lambda - (m+1) + 1)v_m = (\lambda - m)v_m.$$

But $v_{m+1} = 0$ and $v_m \neq 0$, hence $\lambda = m = \dim V - 1$.

By the theorem, the weights of h on V are the values $\lambda, \lambda - 2, \cdots, \lambda - 2m$. But we can now express these as $\lambda, \lambda - 2, \cdots, -\lambda$. □

9.15. **Definition.** In light of (Corollary 9.14), there exists a unique positive integer λ for which vectors in V are maximal if and only if their weight is λ. Further, all other weights of h on V are integers less than λ. We therefore call λ the **highest weight of** V.

9.16. **Proposition.** *Let V be an $\mathfrak{sl}_2(\mathbb{F})$-module. Then V can be expressed as a direct sum,*

$$V = I_1 \oplus \cdots \oplus I_n, \tag{9.16.1}$$

of irreducible submodules I_i. If $\lambda \in \mathbb{F}$ is a weight of h on V, then $I_i \cap V_\lambda \neq \{0\}$ for some $i = 1, \cdots, n$, where V_λ is the weight space of h in V with weight λ. That is, there exists some nonzero $x \in I_i$ with weight λ.

Proof. As $\mathfrak{sl}_2(\mathbb{F})$ is simple, Weyl's Theorem (9.4) implies that V is completely reducible. That is, V can be expressed as in (9.16.1).

Suppose $\lambda \in \mathbb{F}$ is a weight of h on V. That is $V_\lambda \neq \{0\}$. Then there exists some nonzero $v \in V_\lambda$, so $h \cdot v = \lambda v$. By (9.16.1), there exist $v_i \in I_i$ for each $i = 1, \cdots, n$ such that

$$v = \sum_{i=1}^{n} v_i.$$

Then, as $h \cdot v = \lambda v$,

$$h \cdot v = \sum_{i=1}^{n} h \cdot v_i = \lambda v = \sum_{i=1}^{n} \lambda v_i.$$

As the I_i all intersect trivially, this implies that $h \cdot v_i = \lambda v_i$ for each $i = 1, \cdots, n$. That is, $v_i \in V_\lambda$. As v is nonzero, at least one of these v_i is nonzero, hence $I_i \cap V_\lambda \neq \{0\}$ for some $i = 1, \cdots n$. □

9.17. **Proposition.** *Let V be an irreducible $\mathfrak{sl}_2(\mathbb{F})$-module. Then each weight space of V can be generated from any other, by applying x or y.*

Proof. Let $S = \mathfrak{sl}_2(\mathbb{F})$. By Corollary 9.14, V is expressible as a direct sum of 1-dimensional weight spaces of h,

$$V = V_m \oplus V_{m-2} \oplus \cdots \oplus V_{-m}, \tag{9.17.1}$$

where $m = \dim(V) - 1$. Let Z denote the set of indices which appear in this expression. That is, $Z = \{m, m-2, \cdots, -m\}$.

Suppose that $x \cdot V_r = \{0\}$ for some $r \in Z \setminus \{m\}$. Let $V' = V_r \oplus \cdots \oplus V_{-m}$. Then $\{0\} \subset V' \subset V$. By Lemma 9.10, for all $\lambda \in \mathbb{F}$,

$$\begin{aligned} x \cdot V_\lambda &\subseteq V_{\lambda+2}, \\ y \cdot V_\lambda &\subseteq V_{\lambda-2}, \\ h \cdot V_\lambda &\subseteq V_\lambda. \end{aligned} \tag{9.17.2}$$

Therefore,

$$\begin{aligned} S \cdot V' &= S \cdot V_r \oplus \cdots S \cdot V_{-m} \\ &\subseteq x \cdot V_r \oplus V_r \oplus \cdots \oplus V_{-m-2} \\ &\subseteq V_r \oplus \cdots \oplus V_{-m} \\ &\subseteq V', \end{aligned}$$

as $V_{-m-2} = \{0\}$ by (9.17.1) and $x \cdot V_r = \{0\}$ by supposition. This implies that V' is an $\mathfrak{sl}_2(\mathbb{F})$-module. This contradicts the assumption that V is irreducible. Therefore, the supposition that $x \cdot V_r = \{0\}$ must be false. By (9.17.2), and as each nonzero weight space in the decomposition (9.17.1) is 1-dimensional, this implies that $x \cdot V_r = V_{r+2}$ for all $r \in Z$.

Using an almost identical argument to the above (replace x with y and reverse some signs), we can also conclude that $y \cdot V_r = V_{r-2}$ for all $r \in Z$. Together with the decomposition (9.17.1), these imply the result we are after. □

9.18. **Proposition.** *Let V be a finite dimensional $\mathfrak{sl}_2(\mathbb{F})$-module. If V is expressed as a direct sum*

$$V = I_1 \oplus \cdots \oplus I_n,$$

of n irreducible submodules I_i, then $n = \dim(V_0) + \dim(V_1)$.

Proof. Let $I = I_i$ for some irreducible submodule I_i of V. By Corollary 9.14,

$$I = W_m(I) \oplus W_{m-2}(I) \oplus \cdots \oplus W_{-m}(I),$$

for weight spaces $W_i(I)$, where $\dim(W_i(I)) = 1$ for each i occurring in the sum. As the indices in this sum are symmetric about 0, occur in intervals of 2 and are integers, precisely one of 0 or 1 occurs as an index. That is, either $W_0(I) \neq \{0\}$ or $W_1(I) \neq \{0\}$, but not both, hence

$$\dim(I \cap V_0) + \dim(I \cap V_1) = \dim(W_0(I)) + \dim(W_1(I)) = 1.$$

Therefore, as $V = I_1 \oplus \cdots \oplus I_n$, we have

$$\begin{aligned} \dim(V_0) + \dim(V_1) &= \dim(V \cap V_0) + \dim(V \cap V_1) \\ &= \dim\left(\bigoplus_{i=1}^{n} I_i \cap V_0\right) + \dim\left(\bigoplus_{i=1}^{n} I_i \cap V_1\right) \\ &= \sum_{i=1}^{n} \dim(I_i \cap V_0) + \dim(I_i \cap V_1) \\ &= \sum_{i=1}^{n} 1 \\ &= n. \end{aligned}$$

□

9.19. **Lemma.** *$\mathfrak{sl}_2(\mathbb{F})$ is an irreducible $\mathfrak{sl}_2(\mathbb{F})$-module (via the Lie bracket).*

Proof. Suppose $A \subset \mathfrak{sl}_2(\mathbb{F})$ is a nontrivial submodule of $\mathfrak{sl}_2(\mathbb{F})$. Then $[\mathfrak{sl}_2(\mathbb{F})A] \subseteq A$, hence A is an ideal of $\mathfrak{sl}_2(\mathbb{F})$. But $\mathfrak{sl}_2(\mathbb{F})$ is simple (Example 5.11), so has no proper nontrivial ideals. This is a contradiction, so the supposition must be false. Therefore $\mathfrak{sl}_2(\mathbb{F})$ has no nontrivial submodules. That is, $\mathfrak{sl}_2(\mathbb{F})$ is irreducible. □

10. Orthogonality Properties

We have been building up to a series of results about the root space decomposition of L, which will ultimately motivate the definition of a structure called a root system. This series of results is divided into three sets, each building upon the previous, the first of which we shall prove in this section.

We continue to take L to be semisimple, T to be a maximal torus of L and $\mathbb{F}$ to be algebraically closed with characteristic zero.

10.1. **Proposition.** *Φ spans T^*.*

Proof. Let $n = \dim(T) = \dim(T^*)$ and let $m = \dim(\operatorname{span}(\Phi))$. There exists a subset $A \subseteq \Phi$ such that $|A| = m$. Suppose that $\operatorname{span}(\Phi) \neq T^*$, hence $m < n$.

The kernel of any nonzero $\alpha \in T^*$ has codimension 1 in T. Thus $\dim(\ker(\alpha)) = n - 1$ for each $\alpha \in A$. As $m < n$, the intersection of m subspaces of dimension $n-1$ must have at least dimension 1. Therefore, there exists a nonzero $t_0 \in T$ such that

$$t_0 \in \bigcap_{\alpha \in A} \ker \alpha.$$

Any $\alpha \in \Phi \setminus A$ is a linear combination of elements of A, so has a kernel which is a linear combination of the kernels of A. Therefore, $t_0 \in \ker \alpha$ for all $\alpha \in \Phi$.

Then we have $[t_0 x] = \alpha(t_0)x = 0$ for all $x \in L_\alpha$ for $\alpha \in \Phi$. Further, $[t_0 t] = 0$ for all $t \in T$ as T is Abelian. Therefore, by the root space decomposition (Corollary 8.11), we have $[t_0 x] = 0$ for all $x \in L$, hence $t_0 \in Z(L)$. But as L is semisimple, $Z(L) = \{0\}$ (Proposition 5.4). This is a contradiction, so the supposition that $\operatorname{span}(\Phi) \neq T^*$ must be false. □

10.2. **Proposition.** *Let $\alpha \in \Phi$, $x \in L_\alpha$ and $y \in L_{-\alpha}$. Then $[xy] = \langle x, y \rangle t_\alpha$, where t_α is the element of T corresponding to α as in Remark 8.12.*

Proof. Let $t \in T$. We have $\langle t, [xy] \rangle = \langle [tx], y \rangle$ (Proposition 7.2). As $x \in L_\alpha$, we have $[tx] = \alpha(t)x$. Lastly, if t_α is the element of T corresponding to α (Remark 8.12), then $\alpha(t) = \langle t_\alpha, t \rangle$. Therefore,

$$\begin{aligned} \langle t, [xy] \rangle &= \langle \alpha(t)x, y \rangle \\ &= \langle \langle t_\alpha, t \rangle x, y \rangle \\ &= \langle t_\alpha, t \rangle \langle x, y \rangle \\ &= \langle \langle x, y \rangle t_\alpha, t \rangle \\ &= \langle t, \langle x, y \rangle t_\alpha \rangle . \end{aligned}$$

Rearranging the above equation gives

$$0 = \langle t, [xy] \rangle - \langle t, \langle x, y \rangle t_\alpha \rangle = \langle t, [xy] - \langle x, y \rangle t_\alpha \rangle . \tag{10.2.1}$$

The Killing form is nondegenerate on T (Theorem 8.6). Therefore, as (10.2.1) holds for all $t \in T$, we have that $[xy] - \langle x, y \rangle t_\alpha = 0$. Thus $[xy] = \langle x, y \rangle t_\alpha$. □

10.3. **Proposition.** *Let $\alpha \in \Phi$. Then $[L_\alpha L_{-\alpha}] = \mathbb{F} t_\alpha$.*

Proof. By Proposition 10.2, we have that

$$\begin{aligned} [L_\alpha L_{-\alpha}] &= \{[xy] : x \in L_\alpha;\ y \in L_{-\alpha}\} \\ &= \{\langle x, y \rangle t_\alpha : x \in L_\alpha;\ y \in L_{-\alpha}\} \\ &= \langle L_\alpha, L_{-\alpha} \rangle t_\alpha. \end{aligned}$$

We also have that $\langle L_\alpha, L_{-\alpha} \rangle = \mathbb{F}$ (Proposition 7.7), so we are done. □

10.4. **Lemma.** *Let $\alpha \in \Phi$. Let $x \in L_\alpha$ and $y \in L_{-\alpha}$ such that $\langle x, y\rangle \neq 0$. Let $S = \operatorname{span}\{x, y, [xy]\}$. Then $\dim(S) = 3$.*

Proof. The assumption that $\langle x, y\rangle \neq 0$ implies both x and y are nonzero. By Proposition 10.2, $[xy] = \langle x, y\rangle\, t_\alpha \neq 0$, as t_α corresponds to $\alpha \neq 0$. As x, y and $[xy]$ are nonzero elements of the linearly independent spaces L_α, $L_{-\alpha}$ and $T = L_0$ respectively (Proposition 6.3), their span has dimension 3. □

10.5. **Proposition.** *Let $\alpha \in \Phi$. Then $\alpha(t_\alpha) = \langle t_\alpha, t_\alpha\rangle \neq 0$*

Proof. Suppose $\alpha(t_\alpha) = 0$. Then, for all $x \in L_\alpha$ and $y \in L_{-\alpha}$, we have

$$[t_\alpha x] = [t_\alpha y] = 0. \tag{10.5.1}$$

We have that $\langle L_\alpha, L_{-\alpha}\rangle = \mathbb{F}$ (Proposition 7.7). Therefore, there exist $x \in L_\alpha$ and $y \in L_{-\alpha}$ such that $\langle x, y\rangle = 1$ (these elements must therefore be nonzero). Then $[xy] = t_\alpha$ by Proposition 10.2.

Let $S = \operatorname{span}\{x, y, t_\alpha\}$. By Lemma 10.4, we have $\dim(S) = 3$. Using (10.5.1), we can calculate

$$\begin{aligned}[SS] &= \operatorname{span}\{[xy], [t_\alpha x], [t_\alpha y]\} \\ &= \operatorname{span}\{t_\alpha, 0, 0\} \\ &= \mathbb{F}t_\alpha. \end{aligned} \tag{10.5.2}$$

Therefore

$$[[SS][SS]] = [(\mathbb{F}t_\alpha)(\mathbb{F}t_\alpha)] = \mathbb{F}[t_\alpha t_\alpha] = 0,$$

hence S is solvable. As L is semisimple, the adjoint representation of L is injective (Lemma 5.6). Therefore, $S \cong \operatorname{ad}_L(S) \subseteq \mathfrak{gl}(L)$, hence $\operatorname{ad}_L(S)$ is also solvable. Then Theorem 1.27 implies that $[\operatorname{ad}_L(S) \operatorname{ad}_L(S)] = \operatorname{ad}_L([SS])$ is nilpotent. By (10.5.2), $t_\alpha \in [SS]$, so $\operatorname{ad}_L(t_\alpha) \in \operatorname{ad}_L([SS])$, which is nilpotent. Therefore $\operatorname{ad}_L(t_\alpha)$ is nilpotent, but is also semisimple, as $t_\alpha \in T$. If a map is both nilpotent and semisimple, then it must be zero (Lemma 3.3). That is, $[t_\alpha L] = 0$, hence $t_\alpha \in Z(L)$. As L is semisimple, $Z(L) = \{0\}$ (Proposition 5.4), hence $t_\alpha = 0$. But this contradicts the definition of t_α, as $\alpha \neq 0$. Therefore, the supposition that $\alpha(t_\alpha) = 0$ must be false. □

10.6. **Proposition.** *Let $\alpha \in \Phi$ and $x_\alpha \in L_\alpha \setminus \{0\}$. Then there exists $y_\alpha \in L_{-\alpha}$ such that, for $h_\alpha = [x_\alpha y_\alpha]$, the span of $\{x_\alpha, y_\alpha, h_\alpha\}$ is a Lie subalgebra of L isomorphic to $\mathfrak{sl}_2(\mathbb{F})$. Further,*

$$h_\alpha = \frac{2t_\alpha}{\langle t_\alpha, t_\alpha\rangle}$$

and $h_{-\alpha} = -h_\alpha$.

Proof. We have that $\langle x_\alpha, L_{-\alpha}\rangle = \mathbb{F}$ (Proposition 7.7). Therefore, for any $\lambda \in \mathbb{F}$, we can find some $y_\alpha \in L_{-\alpha}$ such that $\langle x_\alpha, y_\alpha\rangle = \lambda$. Specifically, we can find a y_α satisfying

$$\langle x_\alpha, y_\alpha\rangle = \frac{2}{\langle t_\alpha, t_\alpha\rangle},$$

as this value is in $\mathbb{F}$ (Proposition 10.5).

Let

$$h_\alpha = \frac{2t_\alpha}{\langle t_\alpha, t_\alpha\rangle}.$$

Then, by Proposition 10.2,

$$[x_\alpha y_\alpha] = \langle x_\alpha, y_\alpha\rangle\, t_\alpha = \frac{2}{\langle t_\alpha, t_\alpha\rangle} t_\alpha = h_\alpha.$$

Further,

$$[h_\alpha x_\alpha] = \frac{2}{\langle t_\alpha, t_\alpha\rangle}[t_\alpha x_\alpha] = \frac{2}{\langle t_\alpha, t_\alpha\rangle}\alpha(t_\alpha)x_\alpha = 2\frac{\langle t_\alpha, t_\alpha\rangle}{\langle t_\alpha, t_\alpha\rangle}x_\alpha = 2x_\alpha.$$

Lastly,

$$[h_\alpha y_\alpha] = \frac{2}{\langle t_\alpha, t_\alpha\rangle}[t_\alpha y_\alpha] = \frac{2}{\langle t_\alpha, t_\alpha\rangle}(-\alpha)(t_\alpha)y_\alpha = -2\frac{\langle t_\alpha, t_\alpha\rangle}{\langle t_\alpha, t_\alpha\rangle}y_\alpha = -2y_\alpha.$$

Let $S = \operatorname{span}\{x_\alpha, y_\alpha, h_\alpha\}$. Lemma 10.4 implies that $\dim(S) = 3$; by the above, S has the same multiplication table as $\mathfrak{sl}_2(\mathbb{F})$.

For the final point, we can see that, for all $t \in T$,

$$(-\alpha)(t) = -\alpha(t) = -\langle t_\alpha, t\rangle = \langle -t_\alpha, t\rangle,$$

hence $t_{-\alpha} = -t_\alpha$. Therefore,

$$h_{-\alpha} = \frac{2}{\langle t_{-\alpha}, t_{-\alpha}\rangle}t_{-\alpha} = \frac{2}{\langle -t_\alpha, -t_\alpha\rangle}(-t_\alpha) = (-1)^3\frac{2}{\langle t_\alpha, t_\alpha\rangle}t_\alpha = -\frac{2}{\langle t_\alpha, t_\alpha\rangle}t_\alpha = -h_\alpha.$$

□

10.7. **Corollary.** *Let $\alpha \in \Phi$ and let $S_\alpha = \{x_\alpha, y_\alpha, h_\alpha\}$ be taken as in Proposition 10.6. Then $\alpha(h_\alpha) = 2$.*

Proof. We have $\alpha(h_\alpha) = \langle t_\alpha, h_\alpha\rangle$ (Remark 8.12). Therefore

$$\alpha(h_\alpha) = \langle t_\alpha, h_\alpha\rangle = \left\langle t_\alpha, \frac{2t_\alpha}{\langle t_\alpha, t_\alpha\rangle}\right\rangle = 2\frac{\langle t_\alpha, t_\alpha\rangle}{\langle t_\alpha, t_\alpha\rangle} = 2.$$

□

10.8. **Definition.** For $\alpha \in \Phi$, we write $S_\alpha = \operatorname{span}\{x_\alpha, y_\alpha, h_\alpha\}$, for some choice of $x_\alpha \in L_\alpha$ and $y_\alpha \in L_{-\alpha}$, and where $h_\alpha = [x_\alpha y_\alpha] \in T$. As in Proposition 10.6, we have that $S_\alpha \cong \mathfrak{sl}_2(\mathbb{F})$.

11. Integrality Properties

We continue the series of results on the root space structure from the previous section. Of note are the results that all root spaces are 1-dimensional and that $\pm\alpha$ are the only roots in the span of a root α. We also introduce the notion of a root string - roots of the form $\beta+i\alpha$ occur in unbroken sequences. This notion will be developed further in Section 15: Pairs of Roots. It will become important when we look at the Chevalley Basis of a Lie algebra in Section 17: Chevalley Basis.

For roots α, we use the notation S_α (Definition 10.8) to denote some 3 dimensional subalgebra of L, spanned by an element of L_α and $L_{-\alpha}$ and their product. These S_α are isomorphic to $\mathfrak{sl}_2(\mathbb{F})$ (Proposition 10.6), so we can take advantage of the representation theory of $\mathfrak{sl}_2(\mathbb{F})$ we discussed in the Section 9: Modules and Weights.

We continue to take L to be semisimple, T to be a maximal torus of L and $\mathbb{F}$ to be algebraically closed with characteristic zero.

11.1. **Lemma.** *Let $\alpha \in \Phi$ and let*

$$M = \sum_{\lambda\in\mathbb{F}} L_{\lambda\alpha}.$$

Then $M = S_\alpha + T = L_\alpha \oplus L_{-\alpha} \oplus L_0$.

Proof. Let $x_\alpha \in L_\alpha$. By Proposition 10.6, we can find $y_\alpha \in L_{-\alpha}$ such that $S_\alpha = \operatorname{span}\{x_\alpha, y_\alpha, h_\alpha\}$ (where $h_\alpha = [x_\alpha y_\alpha]$) is a 3 dimensional subalgebra of L isomorphic to $\mathfrak{sl}_2(\mathbb{F})$.

We want M to be an S_α-module via the Lie bracket. This is automatically true, as long as we have closure: $[S_\alpha M] \subseteq M$. We have that

$$[L_\beta L_\gamma] \subseteq L_{\beta+\gamma} \tag{11.1.1}$$

for all $\beta, \gamma \in T^*$ (Proposition 6.6), and that $L_0 = T$ (Corollary 8.11), hence $[L_\alpha L_{-\alpha}] \subseteq T$. Therefore

$$S_\alpha \subseteq L_\alpha \oplus L_{-\alpha} \oplus L_0 \subseteq M. \tag{11.1.2}$$

Applying (11.1.1) again, we get

$$\begin{aligned}
[S_\alpha M] &= \sum_{\lambda\in\mathbb{F}} [S_\alpha L_{\lambda\alpha}] \\
&\subseteq \sum_{\lambda\in\mathbb{F}} [L_\alpha L_{\lambda\alpha}] + [L_{-\alpha} L_{\lambda\alpha}] + [L_0 L_{\lambda\alpha}] \\
&\subseteq \sum_{\lambda\in\mathbb{F}} L_{(\lambda+1)\alpha} + L_{(\lambda-1)\alpha} + L_{\lambda\alpha}.
\end{aligned}$$

For each $\lambda \in \mathbb{F}$, we have that $L_{(\lambda+1)\alpha}, L_{(\lambda-1)\alpha}, L_{\lambda\alpha} \subseteq M$ by definition, hence $[S_\alpha M] \subseteq M$ and M is an S_α-module. Specifically an S_α-submodule of L.

We now want to find the weights of h_α on M. Lemma 9.9 gives us that

$$M = \bigoplus_{\lambda\in\mathbb{F}} M_\lambda,$$

where $M_\lambda = \{x \in M : [h_\alpha x] = \lambda x\}$. The value $\lambda \in \mathbb{F}$ is a weight of h_α if $M_\lambda \neq \{0\}$. That is, λ is a weight if it is an eigenvalue of $\operatorname{ad}(h_\alpha)$ on M.

Let $\lambda \in \mathbb{F}$ and $x \in M_\lambda$. As $x \in M$, there exists an $x_\mu \in L_{\mu\alpha}$ for each $\mu \in \mathbb{F}$, such that

$$x = \sum_{\lambda\in\mathbb{F}} x_\mu.$$

Then, as $\alpha(h_\alpha) = 2$ (Corollary 10.7), we have

$$\begin{aligned}[h_\alpha x] &= \sum_{\mu\in\mathbb{F}} [h_\alpha x_\mu] \\ &= \sum_{\mu\in\mathbb{F}} (\mu\alpha)(h_\alpha) x_\mu \\ &= \sum_{\mu\in\mathbb{F}} (2\mu) x_\mu.\end{aligned}$$

But as $x \in M_\lambda$, we have $[h_\alpha x] = \lambda x$, hence

$$\sum_{\mu\in\mathbb{F}} 2\mu x_\mu = \lambda x = \sum_{\mu\in\mathbb{F}} \lambda x_\mu.$$

The x_μ are all linearly independent (Proposition 6.3). Therefore, for all $\mu \in \mathbb{F}$, we have $2\mu x_\mu = \lambda x_\mu$, hence either $x_\mu = 0$ or $2\mu = \lambda$. That is, for all $\mu \in \mathbb{F}$,

$$2\mu \neq \lambda \implies x_\mu = 0,$$

hence

$$x = x_{\frac{\lambda}{2}} \in L_{\frac{\lambda}{2}\alpha}.$$

We have therefore shown that

$$M_\lambda \subseteq L_{\frac{\lambda}{2}\alpha}$$

for all $\lambda \in \mathbb{F}$.

Now let $x \in L_{\lambda\alpha}$. Then $[h_\alpha x] = \lambda\alpha(h_\alpha)x = 2\mu x$ (Corollary 10.7). Therefore $x \in M_{2\lambda}$. We have therefore shown that $L_{\lambda\alpha} \subseteq M_{2\lambda}$ for all $\lambda \in \mathbb{F}$.

Putting these last two results together gives us that, for all $\lambda \in \mathbb{F}$,

$$M_\lambda = L_{\frac{\lambda}{2}\alpha} \tag{11.1.3}$$

This implies that λ is a weight of h_α on M if and only if $\frac{\lambda}{2}\alpha$ is a root. Further, by Corollary 9.14, these weights λ must be integers. That is,

$$M_{2\mu} \neq \{0\} \iff L_{\mu\alpha} \neq \{0\} \implies 2\mu \in \mathbb{Z}. \tag{11.1.4}$$

Let $t \in \ker(\alpha)$. Then

$$\begin{aligned}[x_\alpha t] &= -[t x_\alpha] = -\alpha(t) x_\alpha = 0, \\ [y_\alpha t] &= -[t y_\alpha] = -(-\alpha)(t) y_\alpha = 0, \\ [h_\alpha t] &\in [TT] = \{0\},\end{aligned}$$

hence $[S_\alpha t] = \{0\}$. Therefore, as t was arbitrary,

$$[S_\alpha \ker(\alpha)] = \{0\}. \tag{11.1.5}$$

By definition, $\mathrm{im}(\alpha) \subseteq \mathbb{F}$. As α is a root, it is nonzero. Therefore, $\mathrm{im}(\alpha) = \mathbb{F}$ and $\dim(\mathrm{im}(\alpha)) = 1$. This implies that $\ker(\alpha)$ has codimension 1 in T. We have $\alpha(h_\alpha) = 2$ (Corollary 10.7), hence $h_\alpha \notin \ker(\alpha)$. Therefore, $\ker(\alpha)$ is complementary to $\mathbb{F}h_\alpha$ in T. That is,

$$T = \ker(\alpha) \oplus \mathbb{F}h_\alpha. \tag{11.1.6}$$

By (11.1.3), the fact that $L_0 = T$ (Corollary 8.11), and (11.1.6) respectively,

$$M_0 = L_0 = T = \ker(\alpha) \oplus \mathbb{F}h_\alpha. \tag{11.1.7}$$

Further, $\mathbb{F}h_\alpha \subseteq S_\alpha$ by definition, hence

$$M_0 = T \subseteq \ker(\alpha) \oplus S_\alpha. \tag{11.1.8}$$

The sum is direct, because $x_\alpha, y_\alpha \notin T$.

Suppose (for a contradiction) that M has an even weight greater than 2 (or less than -2), say λ. Then by Proposition 9.16, there exists some irreducible submodule of M which intersects M_λ nontrivially, say I. By Corollary 9.14, we can express the weight space decomposition of I as

$$I = I_m \oplus I_{m-2} \oplus \cdots \oplus I_{-m}, \tag{11.1.9}$$

where $m = \dim(I) + 1$ and each weight space I_i has dimension 1. As $I \cap M_\lambda \neq \{0\}$, it contains elements of weight λ, which is even, hence I has a nonzero weight space of even weight. Therefore, one of the terms in (11.1.9) has even index, which implies all the terms have even index, as the indices are in intervals of 2. Therefore,

$$I_0 \neq \{0\}. \tag{11.1.10}$$

By definition, $I_0 \subseteq M_0$, hence $I_0 \subseteq \ker(\alpha) + S_\alpha$ by (11.1.8). Therefore, by (11.1.5) and as $[S_\alpha S_\alpha] = S_\alpha$,

$$\begin{aligned}
[S_\alpha I_0] &\subseteq [S_\alpha(\ker(\alpha) + S_\alpha)] \\
&\subseteq [S_\alpha \ker(\alpha)] + [S_\alpha S_\alpha] \\
&\subseteq \{0\} + S_\alpha \\
&\subseteq S_\alpha.
\end{aligned}$$

But then repeated application of $\mathrm{ad}(S_\alpha)$ to I_0 will still be contained in S_α. As S_α contains only elements of weight 0 and ± 2, only elements with these weights can be generated from I_0 by applying x_α and y_α. This contradicts Proposition 9.17, therefore the supposition that M has an even weight greater than 2 (or less than -2) must be false.

Specifically, 4 is not a weight of h_α on M. That is, $M_4 = \{0\}$, hence $L_{2\alpha} = \{0\}$ by (11.1.3), so 2α is not a root. As $\alpha \in \Phi$ was arbitrary, this shows that

$$\phi \in \Phi \implies 2\phi \notin \Phi,$$

or equivalently,

$$2\phi \in \Phi \implies \phi \notin \Phi.$$

As $\alpha \in \Phi$, we thus have $\frac{1}{2}\alpha \notin \Phi$. Therefore, by (11.1.3) again, we have that

$$M_1 = \{0\}. \tag{11.1.11}$$

Let $n = \dim(T)$. Then by (11.1.6), $\dim(\ker(\alpha)) = n - 1$. Let $(k_1, \cdots, k_{n-1})$ be a basis for $\ker(\alpha)$. That is,

$$\ker(\alpha) = \bigoplus_{i=1}^{n-1} \mathbb{F}k_i.$$

By (11.1.5), S_α acts trivially on $\ker(\alpha)$, so Lemma 9.5 implies that all 1-dimensional subspaces of $\ker(\alpha)$ are irreducible submodules. Specifically, $\mathbb{F}k_i$ is an irreducible submodule of $\ker(\alpha)$ for each $i = 1, \cdots, n-1$. Therefore, the above decomposition of $\ker(\alpha)$ is actually a decomposition into $n-1$ irreducible submodules.

By (11.1.8), $\ker(\alpha)$ and S_α are disjoint, so consider the direct sum

$$S_\alpha \oplus \ker(\alpha) = S_\alpha \oplus \left(\bigoplus_{i=1}^{n-1} \mathbb{F}k_i \right). \tag{11.1.12}$$

S_α is an irreducible S_α-module (Lemma 9.19). Both S_α and $\ker(\alpha)$ are contained in M by (11.1.2) and (11.1.7) respectively. Further, $\ker(\alpha)$ is an S_α-module by (11.1.5), so they

are also both S_α-submodules of M. Therefore, the expression (11.1.12) is a direct sum of n irreducible submodules of M.

However, by Proposition 9.18, any decomposition of M into irreducible submodules has $m = \dim(M_0) + \dim(M_1)$ terms. Thus by (11.1.11), (11.1.3) and the fact that $L_0 = T$ (Corollary 8.11) respectively,

$$m = \dim(M_0) = \dim(L_0) = \dim(T) = n.$$

That is, the expression (11.1.12) is, in fact, a decomposition of M, so

$$\begin{aligned} M &= S_\alpha \oplus \ker(\alpha) \\ &= S_\alpha + \mathbb{F}h_\alpha + \ker(\alpha) \\ &= S_\alpha + T, \end{aligned} \tag{11.1.13}$$

as $\mathbb{F}h_\alpha \subset S_\alpha$ and $\mathbb{F}h_\alpha + \ker(\alpha) = T$ by (11.1.6).

We have $S_\alpha \subseteq L_0 \oplus L_\alpha \oplus L_{-\alpha}$ and $L_0 = T$ (Corollary 8.11). Therefore (11.1.13) implies that

$$M = S_\alpha + T = S_\alpha + L_0 \subseteq L_0 \oplus L_\alpha \oplus L_{-\alpha},$$

which is contained in M by definition, hence

$$M = L_0 + S_\alpha = L_0 \oplus L_\alpha \oplus L_{-\alpha}.$$

□

11.2. **Theorem.** *Let $\alpha \in \Phi$. Then $\dim(L_\alpha) = 1$. That is, root spaces are 1-dimensional.*

Proof. Let

$$M = \sum_{\lambda \in \mathbb{F}} L_{\lambda\alpha}.$$

By Lemma 11.1, we have

$$M = L_0 + S_\alpha = L_0 \oplus L_\alpha \oplus L_{-\alpha}. \tag{11.2.1}$$

Let $n = \dim(T) = \dim(L_0)$. We know that $\dim(S_\alpha) = 3$ and that $\dim(S_\alpha \cap L_0) = \dim(\mathbb{F}h_\alpha) = 1$. Therefore,

$$\dim(L_0 + S_\alpha) = \dim(L_0) + \dim(S_\alpha) - \dim(S_\alpha \cap L_0) = n + 3 - 1 = n + 2,$$

whereas

$$\dim(L_0 \oplus L_\alpha \oplus L_{-\alpha}) = n + \dim(L_\alpha) + \dim(L_{-\alpha}).$$

Therefore, by (11.2.1),

$$\dim(L_\alpha) + \dim(L_{-\alpha}) = 2. \tag{11.2.2}$$

As $\alpha \in \Phi$, we also have $-\alpha \in \Phi$ (Corollary 7.8), hence both L_α and $L_{-\alpha}$ are nonzero. But equation (11.2.2) implies that if either root space has dimension 2, then the other must have dimension 0, hence they must both have dimension 1. That is,

$$\dim(L_\alpha) = 1.$$

□

11.3. **Lemma.** *Let $\alpha \in \Phi$. Then $S_\alpha = L_\alpha \oplus L_{-\alpha} \oplus [L_\alpha L_{-\alpha}]$.*

Proof. Theorem 11.2 implies both L_α and $L_{-\alpha}$ have dimension 1. That is,

$$[L_\alpha L_{-\alpha}] = [(\mathbb{F}x_\alpha)(\mathbb{F}y_\alpha)] = \mathbb{F}[xy] = \mathbb{F}h_\alpha,$$

hence $L_\alpha \oplus L_{-\alpha} \oplus [L_\alpha L_{-\alpha}]$ has dimension 3. By definition, $S_\alpha \subseteq L_\alpha \oplus L_{-\alpha} \oplus [L_\alpha L_{-\alpha}]$ and also has dimension 3, hence we have equality. □

11.4. **Proposition.** *Let α be a root. Then the only scalar multiples of α which are also roots are $\pm\alpha$.*

Proof. Let

$$M = \sum_{\lambda \in \mathbb{F}} L_{\lambda\alpha}.$$

Then $M = L_\alpha \oplus L_{-\alpha} \oplus L_0$ by Lemma 11.1. Therefore

$$L_{\lambda\alpha} \neq \{0\} \iff \lambda \in \{0, 1, -1\}. \tag{11.4.1}$$

By definition, if $\phi \in T^* \setminus \{0\}$, then $\phi \in \Phi$ if and only if $L_\phi \neq \{0\}$. So (11.4.1) implies that $\lambda\alpha$ is a root if and only if $\lambda = \pm 1$. □

11.5. **Lemma.** *Let $\alpha, \beta \in \Phi$ such that $\beta \neq \pm\alpha$. Let*

$$K = \sum_{i \in \mathbb{Z}} L_{\beta + i\alpha}.$$

Then K is an irreducible S_α-module via the Lie bracket and the weights of h_α on K are the values $\beta(h_\alpha) + 2i$ for which $\beta + i\alpha$ are roots, and these are all integers. Specifically, each weight space $K_{\beta(h_\alpha)+2i} = L_{\beta+i\alpha}$.

Proof. Let

$$K = \sum_{i \in \mathbb{Z}} L_{\beta + i\alpha}.$$

If $\beta + i\alpha = 0$ for any $i \in \mathbb{Z}$, then $\beta = -i\alpha$. But $\beta \in \Phi$, so $\beta = \pm\alpha$ (Proposition 11.4), which contradicts the assumption. Therefore, $\beta + i\alpha \neq 0$ for all $i \in \mathbb{Z}$.

For each $i \in \mathbb{Z}$, calculate

$$\begin{aligned}
[L_\alpha L_{\beta+i\alpha}] &\subseteq L_{\beta+(i+1)\alpha} \subseteq K, \\
[L_{-\alpha} L_{\beta+i\alpha}] &\subseteq L_{\beta+(i-1)\alpha} \subseteq K, \\
[L_0 L_{\beta+i\alpha}] &\subseteq L_{\beta+i\alpha} \subseteq K.
\end{aligned}$$

Therefore, $[S_\alpha K] \subseteq [(L_\alpha \oplus L_{-\alpha} \oplus L_0)K] \subseteq K$, hence K is an S_α-module. Then Lemma 9.9 implies that

$$K = \bigoplus_{\lambda \in \mathbb{F}} K_\lambda$$

for weight spaces K_λ. As in the proof of Lemma 11.1, we want to find a correspondence between weight spaces of K and root spaces of L.

Fix some $i \in \mathbb{Z}$ and let $x \in L_{\beta+i\alpha}$. Then $[tx] = (\beta + i\alpha)(t)x$ for all $t \in T$. Therefore, as $\alpha(h_\alpha) = 2$ (Corollary 10.7), we have

$$\begin{aligned}
[h_\alpha x] &= (\beta + i\alpha)(h_\alpha)x \\
&= (\beta(h_\alpha) + i\alpha(h_\alpha))x \\
&= (\beta(h_\alpha) + 2i)x,
\end{aligned}$$

hence $x \in K_{\beta(h_\alpha)+2i}$. This implies that

$$L_{\beta+i\alpha} \subseteq K_{\beta(h_\alpha)+2i}. \tag{11.5.1}$$

Now let $x \in K_{\beta(h_\alpha)+2i}$. As $x \in K$, there exists an $x_j \in L_{\beta+j\alpha}$ for each $j \in \mathbb{Z}$, such that

$$x = \sum_{j \in \mathbb{Z}} x_j.$$

Therefore,

$$[h_\alpha x] = \sum_{j\in\mathbb{Z}} [h_\alpha x_j] = \sum_{j\in\mathbb{Z}} (\beta + j\alpha)(h_\alpha)x_j.$$

On the other hand, as $\alpha(h_\alpha) = 2$ (Corollary 10.7),

$$\begin{aligned}
[h_\alpha x] &= (\beta(h_\alpha) + 2i)x \\
&= (\beta(h_\alpha) + i\alpha(h_\alpha))x \\
&= (\beta + i\alpha)(h_\alpha)x \\
&= (\beta + i\alpha)(h_\alpha)\sum_{j\in\mathbb{Z}} x_j \\
&= \sum_{j\in\mathbb{Z}} (\beta + i\alpha)(h_\alpha)x_j.
\end{aligned}$$

Put together, these two equations give

$$\sum_{j\in\mathbb{Z}} (\beta + i\alpha)(h_\alpha)x_j = \sum_{j\in\mathbb{Z}} (\beta + j\alpha)(h_\alpha)x_j,$$

which, as $\alpha(h_\alpha) = 2$ by (Corollary 10.7), is equivalent to

$$\sum_{j\in\mathbb{Z}} (\beta(h_\alpha) + 2i)x_j = \sum_{j\in\mathbb{Z}} (\beta(h_\alpha) + 2j)x_j.$$

Root spaces are linearly independent (Proposition 6.3), hence for each $j \in \mathbb{Z}$ we have either $x_j = 0$ or $\beta(h_\alpha) + 2i = \beta(h_\alpha) + 2j$. That is,

$$\beta(h_\alpha) + 2i \neq \beta(h_\alpha) + 2j \implies x_j = 0,$$

which simplifies to

$$i \neq j \implies x_j = 0.$$

In other words, $x = x_i \in L_{\beta+i\alpha}$. As x was arbitrary, we have shown that

$$K_{\beta(h_\alpha)+2i} \subseteq L_{\beta+i\alpha}. \tag{11.5.2}$$

Putting (11.5.1) and (11.5.2) together, we get

$$K_{\beta(h_\alpha)+2i} = L_{\beta+i\alpha}. \tag{11.5.3}$$

Therefore, the weights of h_α on K are the values $\beta(h_\alpha) + 2i$ for which $\beta + i\alpha$ are roots.

Suppose that both 0 and 1 are weights of h_α on K. Then $\beta(h_\alpha)+2i = 0$ and $\beta(h_\alpha)+2j = 1$ for some $i, j \in \mathbb{Z}$. Therefore,

$$1 = 1 - 0 = (\beta(h_\alpha) + 2j) - (\beta(h_\alpha) + 2i) = 2j - 2i = 2(j - i),$$

so $(j - i) = \frac{1}{2} \notin \mathbb{Z}$, which contradicts the fact the i and j are integers. Therefore, not both 0 and 1 can be weights of h_α on K. That is,

$$K_0 \neq \{0\} \iff K_1 = \{0\} \tag{11.5.4}$$

Consider K expressed as a direct sum of irreducible submodules. Proposition 9.18 states that the number of irreducible submodules in such an expression must me equal to $n = \dim(K_0) + \dim(K_1)$. We have that $n \neq 0$, otherwise K would be zero, so K_0 and K_1 cannot both be zero. By (11.5.4), K_0 and K_1 cannot both be nonzero. Therefore, either $n = \dim(K_0)$ or $n = \dim(K_1)$. But (11.5.3) implies the weight spaces are all 1-dimensional, as root spaces are 1-dimensional (Theorem 11.2), so either way, $n = 1$. That is, K is irreducible. Then Corollary 9.14 implies that the weights of h_α on K are all integers. □

11.6. **Corollary.** *Let $\alpha, \beta \in \Phi$. Then $\beta(h_\alpha) \in \mathbb{Z}$.*

Proof. If $\beta \neq \pm\alpha$, then Lemma 11.5 implies the result. Otherwise, $\alpha(h_\alpha) = 2$ (Corollary 10.7) and $(-\alpha)(h_\alpha) = -\alpha(h_\alpha) = -2$. □

11.7. **Definition.** For $\alpha, \beta \in \Phi$, we refer to the values $\beta(h_\alpha)$ as **Cartan integers**.

11.8. **Proposition.** *Let $\alpha, \beta \in \Phi$ such that $\beta \neq \pm\alpha$. Let r and q be the largest integers for which $\beta - r\alpha$ and $\beta + q\alpha$ are roots. Then $\beta(h_\alpha) = r - q$ and we have an unbroken sequence of roots $\{\beta - r\alpha, \cdots, \beta, \cdots, \beta + q\alpha\} \subseteq \Phi$.*

Proof. Define K as in Lemma 11.5 and consider the highest and lowest weights of h_α on K. By Lemma 11.5, these must be $\beta(h_\alpha) + 2q$ and $\beta(h_\alpha) - 2r$ respectively, where the values $q, r \in \mathbb{Z}^+$ are the largest integers for which $\beta + q\alpha$ and $\beta - r\alpha$ are roots. By Corollary 9.14, the weights on K occur as an arithmetic progression with difference 2. Therefore, the weights on K are precisely

$$\beta(h_\alpha) - 2r, \cdots, \beta(h_\alpha), \cdots, \beta(h_\alpha) + 2q,$$

and due to the correspondence between weights and roots in Lemma 11.5, the roots of the form $\beta + i\alpha$ are precisely

$$\beta - r\alpha, \cdots, \beta, \cdots, \beta + q\alpha.$$

Further, Corollary 9.14 also implies that the lowest weight is the negative of the highest weight. That is, $\beta(h_\alpha) - 2r = -(\beta(h_\alpha) + 2q)$, hence $2\beta(h_\alpha) = 2r - 2q$, so

$$\beta(h_\alpha) = r - q.$$

□

11.9. **Definition.** By Proposition 11.8, we can precisely list all roots of the form $\beta + i\alpha$ for $i \in \mathbb{Z}$. We write

$$S_\alpha^\beta = \{\beta - r\alpha, \cdots, \beta, \cdots, \beta + q\alpha\}$$

denote the set of such roots. We call such a set a **root string**. Specifically, S_α^β is called the **α-string through β**.

11.10. **Proposition.** *Let $\alpha, \beta \in \Phi$ such that $\beta \neq \pm\alpha$. Then $(\beta - \beta(h_\alpha)\alpha) \in \Phi$.*

Proof. Let $Z = \{i \in \mathbb{Z} : \beta + i\alpha \in \Phi\}$. By Proposition 11.8, the roots of the form $\beta + i\alpha$ are precisely

$$\beta - r\alpha, \cdots, \beta, \cdots, \beta + q\alpha,$$

for some $r, q \in \mathbb{Z}^+$, hence $Z = \{-r, \cdots, q\}$. Therefore $q - r \in Z$. That is, $\beta + (q-r)\alpha \in \Phi$, so as $\beta(h_\alpha) = r - q$ (Proposition 11.8), we have

$$\beta - \beta(h_\alpha)\alpha = \beta - (r - q)\alpha = \beta + (q - r)\alpha \in \Phi.$$

□

11.11. **Proposition.** *Let $\alpha, \beta \in \Phi$ such that $\beta \neq \pm\alpha$. Then $[L_\alpha L_\beta] = L_{\alpha+\beta}$.*

Proof. Let K be as in Lemma 11.5. Suppose that $\beta + \alpha$ is also a root. Then Lemma 11.5 implies that $\beta(h_\alpha)$ and $\beta(h_\alpha) + 2$ are weights of h_α on K, as well as that K is irreducible. Therefore, by Proposition 9.17, all its weight spaces can be generated from any single one by applying x_α and y_α. Specifically, if we apply Lemma 9.10, we get

$$[x_\alpha K_{\beta(h_\alpha)}] = K_{\beta(h_\alpha)+2}.$$

Therefore, by Lemma 11.5 again, we have

$$[x_\alpha L_\beta] = L_{\beta+\alpha},$$

and as $x_\alpha \in L_\alpha$, this implies that

$$[L_\alpha L_\beta] = L_{\alpha+\beta}.$$

□

11.12. **Theorem.** *L is generated as a Lie algebra by the root spaces L_α for $\alpha \in \Phi$.*

Proof. Φ spans T^* (Proposition 10.1), so there exists some subset $\Phi' \subseteq \Phi$ which is a basis for T^*. We have a bijection between T^* and T, sending each $\alpha \in \Phi$ to t_α (Remark 8.12). Thus $\{t_\alpha : \alpha \in \Phi'\}$ must be a basis for T. That is,

$$T = \bigoplus_{\alpha \in \Phi'} \mathbb{F} t_\alpha. \tag{11.12.1}$$

By Lemma 11.3, we have that $S_\alpha = L_\alpha \oplus L_{-\alpha} \oplus [L_\alpha L_{-\alpha}]$, where $[L_\alpha L_{-\alpha}] = \mathbb{F} h_\alpha$, for

$$h_\alpha = \frac{2t_\alpha}{\langle t_\alpha, t_\alpha \rangle} \in \mathbb{F} t_\alpha.$$

Therefore, $[L_\alpha L_{-\alpha}] = \mathbb{F} t_\alpha$. So by (11.12.1),

$$T = \bigoplus_{\alpha \in \Phi'} [L_\alpha L_{-\alpha}]. \tag{11.12.2}$$

We have the root space decomposition of L (Corollary 8.11):

$$L = T \oplus \left(\bigoplus_{\alpha \in \Phi} L_\alpha \right).$$

By (11.12.2), this can be expressed:

$$L = \left(\bigoplus_{\alpha \in \Phi'} [L_\alpha L_{-\alpha}] \right) \oplus \left(\bigoplus_{\alpha \in \Phi} L_\alpha \right).$$

□

11.13. **Corollary.** *The root space decomposition of L can now be expressed:*

$$L = \left(\bigoplus_{\alpha \in \Phi'} \mathbb{F} h_\alpha \right) \oplus \left(\bigoplus_{\alpha \in \Phi} \mathbb{F} x_\alpha \right),$$

where $\Phi' \subseteq \Phi$ is a basis for T^* and $h_\alpha = [x_\alpha x_{-\alpha}]$ for each $\alpha \in \Phi$.

12. Rationality Properties

In this section, we complete the set of properties for semisimple Lie algebras that we have been working towards. The results which we are interested in are summarised in Theorem 12.12. The crux of this is that Φ forms a root system, which we will define and study in upcoming sections. This is significant, because root systems live in Euclidean space, so we can use Euclidean geometry to obtain results about L, despite the fact that L is over $\mathbb{F} \neq \mathbb{R}$. We begin by recalling the definition of a Euclidean space, which will be used in (Corollary 12.11) as well as in subsequent sections.

We continue to take L to be semisimple, T to be a maximal torus of L and $\mathbb{F}$ to be algebraically closed with characteristic zero.

12.1. **Definition.** A **Euclidean space** is a (finite dimensional) vector space over $\mathbb{R}$ with a positive definite symmetric bilinear form.

12.2. **Definition.** Let $\alpha, \beta \in T^*$. We have corresponding elements $t_\alpha, t_\beta \in T$ (Remark 8.12), so we can define a form on T^* by

$$\langle \alpha, \beta \rangle = \langle t_\alpha, t_\beta \rangle .$$

This form is nondegenerate, as the Killing form is nondegenerate on T (Theorem 8.6).

12.3. **Lemma.** *Let $\alpha, \beta \in \Phi$. Then*

$$\frac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} = \beta(h_\alpha) \in \mathbb{Z}.$$

Proof. We have that $\langle t_\alpha, t_\alpha \rangle \neq 0$ (Proposition 10.5) and $\beta(h_\alpha) \in \mathbb{Z}$ (Corollary 11.6). Therefore,

$$\frac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} = \frac{2\langle t_\beta, t_\alpha \rangle}{\langle t_\alpha, t_\alpha \rangle} = \left\langle t_\beta, \frac{2t_\alpha}{\langle t_\alpha, t_\alpha \rangle} \right\rangle = \langle t_\beta, h_\alpha \rangle = \beta(h_\alpha) \in \mathbb{Z}.$$

□

12.4. **Lemma.** *Let $\beta \in \Phi$. If β is expressed as a sum over some basis $\Phi' \subseteq \Phi$ of T^*, then the coefficients of each $\alpha \in \Phi'$ are rational numbers. Additionally, such a basis exists.*

Proof. Let $(\alpha_1, \cdots, \alpha_n)$ be a basis for T^*, where each $\alpha_i \in \Phi$. This exists because Φ spans T^* (Proposition 10.1). Then, for $\beta \in \Phi$, we can express β as

$$\beta = \sum_{i=1}^{n} \lambda_i \alpha_i,$$

for some collection of $\lambda_i \in \mathbb{F}$.

Then for each $j = 1, \cdots, n$, we have

$$\langle \beta, \alpha_j \rangle = \left\langle \sum_{i=1}^{n} \lambda_i \alpha_i, \alpha_j \right\rangle = \sum_{i=1}^{n} \lambda_i \langle \alpha_i, \alpha_j \rangle .$$

Therefore,

$$\frac{2\langle \beta, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle} = \sum_{i=1}^{n} \frac{2\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle} \lambda_i \tag{12.4.1}$$

We thus have a system of n equations (one for each $j = 1, \cdots, n$), each in n unknowns $(\lambda_1, \cdots, \lambda_n)$, in which the coefficients are integers (Lemma 12.3). The coefficient matrix

for this system of equations is

$$C = \begin{pmatrix} \frac{2\langle\alpha_1,\alpha_1\rangle}{\langle\alpha_1,\alpha_1\rangle} & \cdots & \frac{2\langle\alpha_n,\alpha_1\rangle}{\langle\alpha_1,\alpha_1\rangle} \\ \vdots & \ddots & \vdots \\ \frac{2\langle\alpha_1,\alpha_n\rangle}{\langle\alpha_n,\alpha_n\rangle} & \cdots & \frac{2\langle\alpha_n,\alpha_n\rangle}{\langle\alpha_n,\alpha_n\rangle} \end{pmatrix}.$$

Let

$$M = \begin{pmatrix} \langle\alpha_1,\alpha_1\rangle & \cdots & \langle\alpha_1,\alpha_n\rangle \\ \vdots & \ddots & \vdots \\ \langle\alpha_n,\alpha_1\rangle & \cdots & \langle\alpha_n,\alpha_n\rangle \end{pmatrix}.$$

Suppose that M is singular. Then there exists some linear dependence between the rows of the matrix. That is, there exists a set of coefficients $\lambda_1, \cdots, \lambda_n \in \mathbb{F}$ - not all zero - such that, for each $j = 1, \cdots, n$,

$$\sum_{i=1}^{n} \lambda_i \langle\alpha_i, \alpha_j\rangle = 0.$$

Then, for each $j = 1, \cdots, n$,

$$\langle\beta, \alpha_j\rangle = \left\langle \sum_{i=1}^{n} \lambda_i\alpha_i, \alpha_j \right\rangle = \sum_{i=1}^{n} \lambda_i \langle\alpha_i, \alpha_j\rangle = 0.$$

As $(\alpha_1, \cdots, \alpha_n)$ is a basis for T^*, this implies that $\langle\beta, T^*\rangle = \{0\}$. Therefore, $\beta = 0$, as the form is nondegenerate by definition. But this contradicts the definition of β. Therefore, the supposition that M is singular must be false. That is, M is nonsingular.

For each $i = 1, \cdots, n$, let $M_i = (\langle\alpha_i, \alpha_1\rangle, \cdots, \langle\alpha_i, \alpha_n\rangle)$. Then we can express M as

$$M = \begin{pmatrix} M_1 \\ \vdots \\ M_n \end{pmatrix}.$$

Then

$$C = \begin{pmatrix} \frac{2}{\langle\alpha_1,\alpha_1\rangle} M_1 \\ \vdots \\ \frac{2}{\langle\alpha_n,\alpha_n\rangle} M_n \end{pmatrix},$$

hence

$$\det(C) = \left(\prod_{i=1}^{n} \frac{2}{\langle\alpha_i, \alpha_i\rangle} \right) \det(M).$$

As M is nonsingular, $\det(M) \neq 0$. Therefore, $\det(C) \neq 0$ and C is nonsingular. As C is the matrix describing the system of equations (12.4.1), this implies that there exists a unique solution over $\mathbb{Q}$. That is, $\beta \in \operatorname{span}_{\mathbb{Q}} \{\alpha_1, \cdots, \alpha_n\}$. □

12.5. **Definition.** We write $E_{\mathbb{Q}} = \operatorname{span}_{\mathbb{Q}}(\Phi)$. It makes sense to consider rational multiples of elements of T^*, because $\mathbb{Q} \subseteq \mathbb{F}$ (Proposition 4.1), so in fact $E_{\mathbb{Q}} \subseteq T^*$.

12.6. **Lemma.** *There exists a subset $\Phi' \subseteq \Phi$ which is a basis for both T^* and $E_{\mathbb{Q}}$.*

Proof. By Lemma 12.4, such a Φ' exists as a basis for T^*, where $\Phi \subseteq \operatorname{span}_{\mathbb{Q}}(\Phi')$. Therefore, for each $\alpha \in \Phi$,

$$\alpha = \sum_{\beta \in \Phi'} \mu_{\alpha,\beta}\beta.$$

If $\gamma \in E_{\mathbb{Q}}$, then for some collection of $\lambda_\alpha \in \mathbb{Q}$,

$$\gamma = \sum_{\alpha\in\Phi} \lambda_\alpha \alpha = \sum_{\alpha\in\Phi} \lambda_\alpha \sum_{\beta\in\Phi'} \mu_{\alpha,\beta}\beta = \sum_{\beta\in\Phi'} \left(\sum_{\alpha\in\Phi} \lambda_\alpha \mu_{\alpha,\beta} \right) \beta \in \operatorname{span}_{\mathbb{Q}}(\Phi').$$

That is, $E_{\mathbb{Q}} \subseteq \operatorname{span}_{\mathbb{Q}}(\Phi')$. Further $E_{\mathbb{Q}} = \operatorname{span}_{\mathbb{Q}}(\Phi) \supseteq \operatorname{span}_{\mathbb{Q}}(\Phi')$, as $\Phi' \subseteq \Phi$, hence we have equality: $E_{\mathbb{Q}} = \operatorname{span}_{\mathbb{Q}}(\Phi')$. Lastly, Φ' is $\mathbb{F}$-linearly independent (as it is a basis for T^*), so it must also be $\mathbb{Q}$-linearly independent, as $\mathbb{Q} \subseteq \mathbb{F}$ (Proposition 4.1). Therefore, Φ' is a basis for $E_{\mathbb{Q}}$. □

12.7. **Lemma.** *The form on T^* can be restricted to $E_{\mathbb{Q}}$. This form on $E_{\mathbb{Q}}$ is positive definite.*

Proof. Let $\gamma, \delta \in T^*$. Their corresponding elements in T are t_γ and t_δ respectively. Recall the root space decomposition (Theorem 8.11):

$$L = T \oplus \left(\bigoplus_{\alpha\in\Phi} L_\alpha \right),$$

where $L_\alpha = \mathbb{F}x_\alpha$ (Theorem 11.2). Therefore $\{t_1, \cdots, t_n\} \cup \{x_\alpha : \alpha \in \Phi\}$ is a basis for L, where $\{t_1, \cdots, t_n\}$ is a basis for T.

By definition of L_α, $\operatorname{ad}(t)(x_\alpha) = \alpha(t)x_\alpha$ for all $t \in T$ and $\alpha \in \Phi$. Therefore,

$$\begin{aligned}(\operatorname{ad}(t_\gamma) \circ \operatorname{ad}(t_\delta))(x_\alpha) &= \operatorname{ad}(t_\gamma)(\operatorname{ad}(t_\delta)(x_\alpha)) \\ &= \operatorname{ad}(t_\gamma)(\alpha(t_\delta)x_\alpha) \\ &= \alpha(t_\delta)\operatorname{ad}(t_\gamma)(x_\alpha) \\ &= \alpha(t_\delta)\alpha(t_\gamma)x_\alpha.\end{aligned}$$

Further, t_γ and t_δ act trivially on T, as T is Abelian (Proposition 5.2), hence

$$\operatorname{trace}(\operatorname{ad}(t_\gamma)\operatorname{ad}(t_\delta)) = \sum_{\alpha\in\Phi} \alpha(t_\gamma)\alpha(t_\delta).$$

Therefore,

$$\begin{aligned}\langle \gamma, \delta \rangle &= \langle t_\gamma, t_\delta \rangle \\ &= \operatorname{trace}(\operatorname{ad}(t_\gamma)\operatorname{ad}(t_\delta)) \\ &= \sum_{\alpha\in\Phi} \alpha(t_\gamma)\alpha(t_\delta) \\ &= \sum_{\alpha\in\Phi} \langle \alpha, \gamma \rangle \langle \alpha, \delta \rangle,\end{aligned}$$

that is,

$$\langle \gamma, \delta \rangle = \sum_{\alpha\in\Phi} \langle \alpha, \gamma \rangle \langle \alpha, \delta \rangle. \tag{12.7.1}$$

Let $\beta \in \Phi$. By (12.7.1), we have that

$$\langle \beta, \beta \rangle = \sum_{\alpha\in\Phi} \langle \alpha, \beta \rangle^2.$$

Therefore,

$$\frac{1}{\langle \beta, \beta \rangle} = \frac{\langle \beta, \beta \rangle}{\langle \beta, \beta \rangle^2}$$

$$= \sum_{\alpha\in\Phi} \frac{\langle\alpha,\beta\rangle^2}{\langle\beta,\beta\rangle^2}$$
$$= \sum_{\alpha\in\Phi} \left(\frac{\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle}\right)^2. \tag{12.7.2}$$

By Lemma 12.3, for each $\alpha,\beta\in\Phi$, we have

$$\frac{2\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle}\in\mathbb{Z}, \tag{12.7.3}$$

hence

$$\frac{\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle}\in\mathbb{Q}.$$

Which, when applied to (12.7.2), gives

$$\frac{1}{\langle\beta,\beta\rangle}\in\mathbb{Q},$$

hence we also have that $\langle\beta,\beta\rangle\in\mathbb{Q}$.

Let $\alpha,\beta\in\Phi$. By (12.7.3) again, we have

$$\langle\alpha,\beta\rangle = \frac{1}{2}\langle\beta,\beta\rangle\left(\frac{2\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle}\right)\in\frac{1}{2}\mathbb{Q}\mathbb{Z}=\mathbb{Q}.$$

That is,

$$\langle\Phi,\Phi\rangle\subseteq\mathbb{Q}.$$

Therefore, as $E_{\mathbb{Q}}=\mathrm{span}_{\mathbb{Q}}(\Phi)$,

$$\langle E_{\mathbb{Q}},E_{\mathbb{Q}}\rangle = \left\langle\sum_{\alpha\in\Phi}\mathbb{Q}\alpha,\sum_{\beta\in\Phi}\mathbb{Q}\beta\right\rangle = \sum_{\alpha,\beta\in\Phi}\mathbb{Q}\langle\alpha,\beta\rangle\subseteq\sum_{\alpha,\beta\in\Phi}\mathbb{Q}\mathbb{Q}=\mathbb{Q}. \tag{12.7.4}$$

We can therefore consider this form restricted to $E_{\mathbb{Q}}$.

Suppose $\gamma\in E_{\mathbb{Q}}$ such that $\langle\gamma,E_{\mathbb{Q}}\rangle=\{0\}$. We can express γ as

$$\gamma=\sum_{\alpha\in\Phi}\lambda_\alpha\alpha,$$

for some collection of scalars $\lambda_\alpha\in\mathbb{Q}$. By the supposition, $\langle\gamma,\alpha\rangle=0$ for each $\alpha\in\Phi$, hence

$$\langle\gamma,T^*\rangle=\langle\gamma,\mathrm{span}_{\mathbb{F}}(\Phi)\rangle=\sum_{\alpha\in\Phi}\mathbb{F}\langle\gamma,\alpha\rangle=0.$$

As $\gamma\in E_{\mathbb{Q}}\subseteq T^*$ and the form on T^* is nondegenerate, γ must be zero. This shows that the form is still nondegenerate when restricted to $E_{\mathbb{Q}}$.

Let $\gamma\in E_{\mathbb{Q}}$. By (12.7.4), $\langle\alpha,\gamma\rangle\in\mathbb{Q}$ for all $\alpha\in\Phi$, as $\Phi\subseteq E_{\mathbb{Q}}$. By (12.7.1), we have that

$$\langle\gamma,\gamma\rangle=\sum_{\alpha\in\Phi}\langle\alpha,\gamma\rangle^2.$$

That is, $\langle\gamma,\gamma\rangle$ is a sum of squares of rational numbers, hence is either positive or zero, and is zero only if all terms are zero. That is, if $\langle\gamma,\gamma\rangle=0$, then $\langle\alpha,\gamma\rangle=0$ for all $\alpha\in\Phi$. But as Φ spans $E_{\mathbb{Q}}$ by definition, this implies that $\langle E_{\mathbb{Q}},\gamma\rangle=\{0\}$, which in turn implies that $\gamma=0$, as the form is nondegenerate on $E_{\mathbb{Q}}$. That is, the form is positive definite on $E_{\mathbb{Q}}$. □

12.8. **Definition.** As $\mathbb{Q}$ is a subfield of $\mathbb{R}$ and $E_{\mathbb{Q}}$ is a $\mathbb{Q}$-vector space, both $\mathbb{R}$ and $E_{\mathbb{Q}}$ are $\mathbb{Q}$-modules. We can therefore define the tensor product $E = \mathbb{R} \otimes_{\mathbb{Q}} E_{\mathbb{Q}}$. That is,

$$E = \left\{ \sum_{i=1}^{n} \lambda_i \otimes \gamma_i : \lambda_i \in \mathbb{R};\ \gamma_i \in E_{\mathbb{Q}} \right\},$$

where $(\lambda \otimes \gamma) = 1_{\mathbb{R}} \otimes (\lambda\gamma)$ for all $\lambda \in \mathbb{Q}$ and $\gamma \in E_{\mathbb{Q}}$.

12.9. **Proposition.** $E \cong \operatorname{span}_{\mathbb{R}}(\Phi)$.

Proof. Consider a pure tensor $\lambda \otimes \gamma \in E$ for some $\lambda \in \mathbb{R}$ and $\gamma \in E_{\mathbb{Q}}$. We have

$$\gamma = \sum_{\alpha \in \Phi} \mu_\alpha \alpha$$

for some collection of scalars $\mu_\alpha \in \mathbb{Q}$. Then

$$\lambda \otimes \gamma = \lambda \otimes \sum_{\alpha \in \Phi} \mu_\alpha \alpha = \sum_{\alpha \in \Phi} \lambda \otimes (\mu_\alpha \alpha) = \sum_{\alpha \in \Phi} (\mu_\alpha \lambda) \otimes \alpha = \sum_{\alpha \in \Phi} \nu_\alpha \otimes \alpha,$$

where $\nu_\alpha = \mu_\alpha \lambda \in \mathbb{Q}\mathbb{R} = \mathbb{R}$. We identify $\nu_\alpha \otimes \alpha$ with $\nu_\alpha(1_{\mathbb{R}} \otimes \alpha)$ and $1_{\mathbb{R}} \otimes \alpha$ with α, hence $\nu_\alpha \otimes \alpha$ with $\nu_\alpha \alpha$. Therefore,

$$\lambda \otimes \gamma = \sum_{\alpha \in \Phi} \nu_\alpha \otimes \alpha = \sum_{\alpha \in \Phi} \nu_\alpha \alpha \in \operatorname{span}_{\mathbb{R}}(\Phi).$$

That is, the pure tensors in E are contained in $\operatorname{span}_{\mathbb{R}} \{\Phi\}$. As E consists of sums of pure tensors and $\operatorname{span}_{\mathbb{R}} \{\Phi\}$ is closed under addition, we have shown that

$$E \subseteq \operatorname{span}_{\mathbb{R}}(\Phi). \tag{12.9.1}$$

Now let $\gamma \in \operatorname{span}_{\mathbb{R}}(\Phi)$. That is,

$$\gamma = \sum_{\alpha \in \Phi} \mu_\alpha \alpha = \sum_{\alpha \in \Phi} \mu_\alpha \otimes \alpha$$

for some collection of scalars $\mu_\alpha \in \mathbb{R}$. Then $\mu_\alpha \otimes \alpha \in E$ for each $\alpha \in \Phi$, hence $\gamma \in E$. That is,

$$E \supseteq \operatorname{span}_{\mathbb{R}}(\Phi).$$

Which, with (12.9.1), implies equality. □

12.10. **Lemma.** *The form on $E_{\mathbb{Q}}$ can be extended to E and it remains positive definite.*

Proof. Let $\gamma, \delta \in E$. In light of Proposition 12.9, we consider E as $\operatorname{span}_{\mathbb{R}}(\Phi)$, so

$$\gamma = \sum_{\alpha \in \Phi} \lambda_\alpha \alpha; \quad \delta = \sum_{\beta \in \Phi} \mu_\beta \beta,$$

for some collection of scalars $\lambda_\alpha, \mu_\beta \in \mathbb{R}$.

We extend the form on $E_{\mathbb{Q}}$ to E via

$$\langle \gamma, \delta \rangle_E = \sum_{\alpha \in \Phi} \sum_{\beta \in \Phi} \lambda_\alpha \mu_\beta \langle \alpha, \beta \rangle_{E_{\mathbb{Q}}} \in \mathbb{R}.$$

This is consistent: if each $\lambda_\alpha, \mu_\beta \in \mathbb{Q}$ (and hence $\gamma, \delta \in E_{\mathbb{Q}}$), then

$$\langle \gamma, \delta \rangle_E = \sum_{\alpha \in \Phi} \sum_{\beta \in \Phi} \lambda_\alpha \mu_\beta \langle \alpha, \beta \rangle_{E_{\mathbb{Q}}} = \left\langle \sum_{\alpha \in \Phi} \lambda_\alpha \alpha, \sum_{\beta \in \Phi} \mu_\beta \beta \right\rangle_{E_{\mathbb{Q}}} = \langle \gamma, \delta \rangle_{E_{\mathbb{Q}}}.$$

We can thus write $\langle \gamma, \delta \rangle$ without having to specify whether this is the form on $E_{\mathbb{Q}}$ or E.

As E and $E_{\mathbb{Q}}$ have the same basis, the form must stay positive definite when extended from $E_{\mathbb{Q}}$ to E. □

12.11. **Corollary.** *$E \cong \operatorname{span}_{\mathbb{R}}(\Phi)$ is a Euclidean space.*

12.12. **Theorem.** *Let L be a simple Lie algebra over an algebraically closed field of characteristic zero. Let Φ denote the set of roots of L. Then:*

- *Φ spans E and does not contain 0.*
- *If $\alpha \in \Phi$, then the scalar multiples of α contained in Φ are precisely $\pm\alpha$.*
- *If $\alpha, \beta \in \Phi$, then*
$$\frac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} \in \mathbb{Z}.$$
- *If $\alpha, \beta \in \Phi$, then*
$$\beta - \frac{2\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle}\alpha \in \Phi.$$

Proof. Φ spans E by Proposition 12.9 and does not contain zero by definition. The second point is from Proposition 11.4. The third point is from Lemma 12.3, which with Proposition 11.10 also implies the final point. □

12.13. This theorem is equivalent to the statement that Φ is a root system in E. Root systems will be defined and discussed in Section 14: Root Systems.

13. Reflections In Euclidean Space

Before we start looking at root systems, we need some results about reflections. We outline some basic properties and prove a criterion (Theorem 13.7) which will be useful when we look at root systems in the following section.

Throughout this section, we take E to be a Euclidean space (Definition 12.1). We denote the form on E by $\langle \alpha, \beta \rangle$ for $\alpha, \beta \in E$.

13.1. **Definition.** A **hyperplane** in E is a subspace of codimension 1. For a nonzero vector $\alpha \in E$, we write P_α to denote the corresponding **hyperplane orthogonal to** α, defined

$$P_\alpha = \{\beta \in E : \langle \beta, \alpha \rangle = 0\}.$$

A **reflection** in E is an invertible linear transformation of E which fixes some hyperplane and sends any vector orthogonal to that hyperplane to its negative. For nonzero $\alpha \in E$, denote by σ_α the reflection in the hyperplane P_α.

13.2. **Lemma.** *Let $\alpha, \beta \in E$ and P be some hyperplane in E. If $\alpha \notin P$, then $\beta = \rho + a\alpha$ for some $\rho \in P$ and $a \in \mathbb{R}$. Specifically, if $\alpha \neq 0$, then $\beta = \rho + a\alpha$ for some $\rho \in P_\alpha$ and $a \in \mathbb{R}$.*

Proof. As P has codimension 1, we have that $E = P \oplus \mathbb{R}\alpha$. We can therefore write $\beta = \rho + a\alpha$ for some $\rho \in P$ and $a \in \mathbb{R}$.

If $\alpha \neq 0$, then we have that $\langle \alpha, \alpha \rangle \neq 0$, as the form is positive definite. Therefore $\alpha \notin P_\alpha$, so we can apply the above for $P = P_\alpha$. □

13.3. **Definition.** For $\alpha, \beta \in E$, we will use the notation (linear in the first variable only)

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

Note that this means

$$(\alpha, \alpha) = \frac{2\langle \alpha, \alpha \rangle}{\langle \alpha, \alpha \rangle} = 2.$$

13.4. **Lemma.** *Let $\alpha \in E$. Then the reflection σ_α is given by*

$$\sigma_\alpha(\beta) = \beta - (\beta, \alpha)\alpha,$$

for all $\beta \in E$.

Proof. Let $\beta \in E$. By Lemma 13.2, there exists $\rho_\beta \in P_\alpha$ and $a_\beta \in \mathbb{R}$ such that $\beta = \rho_\beta + a_\beta\alpha$. Therefore,

$$\begin{aligned} \sigma_\alpha(\beta) &= \sigma_\alpha(\rho_\beta + a_\beta\alpha) \\ &= \sigma_\alpha(\rho_\beta) + a_\beta\sigma_\alpha(\alpha) \\ &= \rho_\beta - a_\beta\alpha \\ &= \beta - 2a_\beta\alpha. \end{aligned} \tag{13.4.1}$$

Further, $\langle \rho_\beta, \alpha \rangle = 0$ by definition, hence

$$\begin{aligned} \langle \beta, \alpha \rangle &= \langle \rho_\beta + a_\beta\alpha, \alpha \rangle \\ &= \langle \rho_\beta, \alpha \rangle + a_\beta \langle \alpha, \alpha \rangle \\ &= a_\beta \langle \alpha, \alpha \rangle. \end{aligned}$$

Therefore,

$$a_\beta = \frac{\langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle},$$

which, when applied to (13.4.1), gives

$$\sigma_\alpha(\beta) = \beta - 2\frac{\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha.$$

□

13.5. **Lemma.** *For all $\alpha \in E$, the reflection σ_α is an involution. That is, $\sigma_\alpha = \sigma_\alpha^{-1}$.*

Proof. Let $\beta \in E$. Then, by Lemma 13.4,

$$\begin{aligned}\sigma_\alpha^2(\beta) &= \sigma_\alpha(\beta - (\beta,\alpha)\alpha)\\ &= \sigma_\alpha(\beta) - (\beta,\alpha)\sigma_\alpha(\alpha)\\ &= \beta - (\beta,\alpha)\alpha + (\beta,\alpha)\alpha\\ &= \beta.\end{aligned}$$

As β was arbitrary, this shows that $\sigma_\alpha^2 = 1$, which proves the result. □

13.6. **Definition.** We write $\mathrm{GL}(E)$ to denote the **general linear group of** E. This is defined as the group (under the composition of maps operation) of invertible linear maps in $\mathrm{End}(E)$.

13.7. **Theorem.** *Let Φ be a finite set (not containing zero) which spans E and where all reflections $\{\sigma_\alpha : \alpha \in \Phi\}$ leave Φ invariant. Suppose $\sigma \in \mathrm{GL}(E)$ leaves Φ invariant, pointwise fixes a hyperplane P of E and sends some $\alpha \in \Phi$ to its negative. Then $\sigma = \sigma_\alpha$ and $P = P_\alpha$.*

Proof. Let $\tau = \sigma\sigma_\alpha$. Then $\tau = \sigma\sigma_\alpha^{-1}$ by Lemma 13.5. By definition, $\tau(\Phi) = \Phi$ and $\tau(\alpha) = \alpha$, hence τ fixes $\mathbb{R}\alpha$.

Let $\beta \in E$. As σ pointwise fixes P, but not α, we have that $\alpha \notin P$. Therefore, there exist $\rho \in P$ and $a \in \mathbb{R}$ such that $\beta = \rho + a\alpha$ (Lemma 13.2). Then

$$\begin{aligned}\sigma(\beta) &= \sigma(\rho + a\alpha)\\ &= \sigma(\rho) + a\sigma(\alpha)\\ &= \rho - a\alpha\\ &= \rho + a\alpha - 2a\alpha\\ &= \beta - 2a\alpha,\end{aligned}$$

hence

$$\begin{aligned}\tau(\beta + \mathbb{R}\alpha) &= \tau(\beta) + \mathbb{R}\tau(\alpha)\\ &= \sigma(\sigma_\alpha(\beta)) + \mathbb{R}\alpha\\ &= \sigma(\beta - (\beta,\alpha)\alpha) + \mathbb{R}\alpha\\ &= \sigma(\beta) - (\beta,\alpha)\sigma(\alpha) + \mathbb{R}\alpha\\ &= \beta - 2a\alpha + (\beta,\alpha)\alpha + \mathbb{R}\alpha\\ &= \beta + \mathbb{R}\alpha.\end{aligned}$$

So τ pointwise fixes both $\mathbb{R}\alpha$ and $E/\mathbb{R}\alpha$.

Let $(\varepsilon_1, \cdots, \varepsilon_{n-1})$ be a basis for P_α and let $\varepsilon_n = \alpha$. Then $(\varepsilon_1, \cdots, \varepsilon_n)$ is a basis for E. We then have, for some $e_i \in \mathbb{R}$,

$$\tau : \varepsilon_i \mapsto \begin{cases}\varepsilon_i + e_i\varepsilon_n, & (i = 1, \cdots, n-1),\\ \varepsilon_n, & (i = n).\end{cases}$$

Consider the representation of E where

$$\varepsilon_1 \mapsto \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \varepsilon_n \mapsto \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Then the matrix representation A_τ of τ is given by

$$A_\tau = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \ddots & 0 \\ e_1 & \cdots & \cdots & e_{n-1} & 1 \end{pmatrix}.$$

Thus we can write $A_\tau = N_\tau + I_n$, where I_n is the identity matrix and

$$N_\tau = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ e_1 & \cdots & e_{n-1} & 0 \end{pmatrix}.$$

As N_τ is strictly lower triangular, ${N_\tau}^n = 0$, hence $(A_\tau - I_n)^n = {N_\tau}^n = 0$. Therefore, the minimal polynomial of A_τ (and hence that of τ) divides $(X-1)^n \in \mathbb{R}[X]$.

As $\tau(\Phi) = \Phi$ and Φ is finite, for each $\beta \in \Phi$, we have that $\tau^{k_\beta}(\beta) = \beta$ for some $k_\beta \in \mathbb{N}$. The value

$$\prod_{\beta \in \Phi} k_\beta$$

exists in $\mathbb{N}$, as Φ is finite; further, as k_β divides k, we have that τ^k fixes β for all $\beta \in \Phi$. Therefore, as Φ spans E, we have that τ^k fixes all of E, hence $\tau^k = 1$. Therefore the minimal polynomial of τ divides $X^k - 1$.

The minimal polynomial of τ must therefore divide the greatest common divisor of $(X-1)^k$ and $X^k - 1$, which is $X - 1$. That is, $\tau = 1$. Therefore, $\sigma\sigma_\alpha^{-1} = 1$, which implies that $\sigma = \sigma_\alpha$. $\square$

14. Root Systems

We look at root systems, discussing the Weyl group and duals/inverses of root systems. We conclude by relating the notion of root duals/inverses back to our Lie algebra setting.

We continue to take E to be a Euclidean space. We also take Φ to be a root system in E (defined below).

14.1. **Definition.** Let Φ be a subset of the Euclidean space E. Then Φ is called a **root system** in E if the following axioms hold:

- Φ is finite, spans E and does not contain 0.
- If $\alpha \in \Phi$, the scalar multiples of α in Φ are precisely $\pm\alpha$.
- If $\alpha \in \Phi$, the reflection σ_α leaves Φ invariant.
- If $\alpha, \beta \in \Phi$, then $(\beta, \alpha) \in \mathbb{Z}$.

Let Φ and Φ' be root systems in E and E' respectively. We call Φ and Φ' **isomorphic** as root systems, if there exists a vector space isomorphism, $\phi : E \to E'$, which satisfies:

- $\phi(\Phi) = \Phi'$,
- $(\phi(\beta), \phi(\alpha)) = (\beta, \alpha)$ for all $\alpha, \beta \in \Phi$.

14.2. **Lemma.** *Suppose $\sigma \in \mathrm{GL}(E)$ leaves Φ invariant. Then, for all $\alpha, \beta \in \Phi$:*

- $\sigma\sigma_\alpha\sigma^{-1} = \sigma_{\sigma(\alpha)}$,
- $(\beta, \alpha) = (\sigma(\beta), \sigma(\alpha))$.

Proof. Both σ and σ_α leave Φ invariant (by supposition and by definition of Φ). Therefore,

$$\sigma\sigma_\alpha\sigma^{-1}(\Phi) = \sigma\sigma_\alpha\sigma^{-1}(\sigma(\Phi)) = \sigma\sigma_\alpha(\Phi) = \Phi,$$

so $\sigma\sigma_\alpha\sigma^{-1}$ leaves Φ invariant.

For all $\beta \in E$, we have

$$\begin{aligned}\sigma\sigma_\alpha\sigma^{-1}(\sigma(\beta)) &= \sigma\sigma_\alpha(\beta)\\ &= \sigma(\beta - (\beta, \alpha)\alpha)\\ &= \sigma(\beta) - (\beta, \alpha)\sigma(\alpha). \qquad (14.2.1)\end{aligned}$$

Further, if $\beta \in P_\alpha$, then $\langle \beta, \alpha \rangle = 0$, hence $(\beta, \alpha) = 0$; therefore,

$$\begin{aligned}\sigma\sigma_\alpha\sigma^{-1}(\sigma(\beta)) &= \sigma(\beta) - (\beta, \alpha)\sigma(\alpha)\\ &= \sigma(\beta),\end{aligned}$$

so $\sigma\sigma_\alpha\sigma^{-1}$ fixes $\sigma(P_\alpha)$, which is still a hyperplane, because σ is invertible so preserves dimension.

Lastly,

$$\sigma\sigma_\alpha\sigma^{-1}(\sigma(\alpha)) = \sigma\sigma_\alpha(\alpha) = \sigma(-\alpha) = -\sigma(\alpha).$$

Therefore, $\sigma\sigma_\alpha\sigma^{-1}$ leaves Φ invariant, pointwise fixes $\sigma(P_\alpha)$ and sends $\sigma(\alpha)$ to its negative. Therefore, Theorem 13.7 implies that $\sigma\sigma_\alpha\sigma^{-1} = \sigma_{\sigma(\alpha)}$.

Therefore, for all $\beta \in E$,

$$\sigma_{\sigma(\alpha)}(\sigma(\beta)) = \sigma\sigma_\alpha\sigma^{-1}(\sigma(\beta)) = \sigma(\beta) - (\beta, \alpha)\sigma(\alpha),$$

by (14.2.1), whereas by Lemma 13.4,

$$\sigma_{\sigma(\alpha)}(\sigma(\beta)) = \sigma(\beta) - (\sigma(\beta), \sigma(\alpha))\sigma(\alpha).$$

Therefore,

$$\sigma(\beta) - (\beta, \alpha)\sigma(\alpha) = \sigma(\beta) - (\sigma(\beta), \sigma(\alpha))\sigma(\alpha),$$

which simplifies to

$$(\beta, \alpha) = (\sigma(\beta), \sigma(\alpha)).$$

□

14.3. **Corollary.** *Let $\phi \in \mathrm{GL}(E)$. Suppose ϕ leaves Φ invariant. Then ϕ is a root system automorphism of Φ.*

Proof. The definition of a root system isomorphism is already satisfied by ϕ, bar one point: that $(\beta, \alpha) = (\phi(\beta), \phi(\alpha))$ for all $\alpha, \beta \in \Phi$. This is implied by Lemma 14.2. □

14.4. **Lemma.** *Let (E, Φ) and (E', Φ') be isomorphic root systems with isomorphism ϕ. Then*

$$\sigma_{\phi(\alpha)}(\phi(\beta)) = \phi(\sigma_\alpha(\beta))$$

for all $\alpha, \beta \in \Phi$.

Proof. Let $\alpha, \beta \in \Phi$. Then,

$$\begin{aligned}\sigma_{\phi(\alpha)}(\phi(\beta)) &= \phi(\beta) - (\phi(\beta), \phi(\alpha))\phi(\alpha) \\ &= \phi(\beta) - (\beta, \alpha)\phi(\alpha) \\ &= \phi(\beta - (\beta, \alpha)\alpha) \\ &= \phi(\sigma_\alpha(\beta))\end{aligned}$$

by Lemma 13.4 and the definition of root system isomorphisms. □

14.5. **Definition.** Denote by $\mathcal{W}$ the subgroup of $\mathrm{GL}(E)$ generated by the reflections $\{\sigma_\alpha : \alpha \in \Phi\}$. This is called the **Weyl group** of Φ.

14.6. **Proposition.** *$\mathcal{W}$ is a subgroup of the symmetric group on Φ and is finite.*

Proof. By definition, Φ is left invariant by the reflections generating $\mathcal{W}$. As Φ is a finite set, these generators are permutations of Φ, hence are elements of the symmetric group on Φ. □

14.7. **Proposition.** *Let (E, Φ) and (E', Φ') be isomorphic root systems with isomorphism ϕ. Let $\mathcal{W}$ and $\mathcal{W}'$ denote their respective Weyl groups. Then the map*

$$\phi^* : \mathcal{W} \to \mathcal{W}' : \sigma \mapsto \phi \circ \sigma \circ \phi^{-1}$$

is a group isomorphism.

Proof. Let $\sigma_1, \sigma_2 \in \mathcal{W}$. Then

$$\begin{aligned}\phi^*(\sigma_1\sigma_2) &= \phi \circ (\sigma_1 \circ \sigma_2) \circ \phi^{-1} \\ &= \phi \circ \sigma_1 \circ (\phi^{-1} \circ \phi) \circ \sigma_2 \circ \phi^{-1} \\ &= (\phi \circ \sigma_1 \circ \phi^{-1}) \circ (\phi \circ \sigma_2 \circ \phi^{-1}) \\ &= \phi^*(\sigma_1)\phi^*(\sigma_2),\end{aligned}$$

hence ϕ^* is a group homomorphism.

$\mathcal{W}$ and $\mathcal{W}'$ are generated by the sets $R = \{\sigma_\alpha : \alpha \in \Phi\}$ and $R' = \{\sigma_\beta : \beta \in \Phi'\}$ respectively. As $\sigma_\alpha = \sigma_\beta$ if and only if $\beta \in \mathbb{R}\alpha$, and the scalar multiples of $\alpha \in \Phi$ are precisely $\pm\alpha$ by definition, we have that $|R| = \frac{1}{2}|\Phi|$ and $|R'| = \frac{1}{2}|\Phi'|$. Further, $|\Phi| = |\Phi'|$, as the root systems are isomorphic, and this value is finite by definition. Therefore,

$$|R| = |R'| \tag{14.7.1}$$

and is finite.

Let $\sigma_\beta \in P'$. That is, $\beta \in \Phi'$. As $\phi(\Phi) = \Phi'$, we have $\beta = \phi(\alpha)$ for some $\alpha \in \Phi$. As ϕ is an isomorphism, for all $\gamma \in E'$, there exists some $\delta \in E$ such that $\gamma = \phi(\delta)$ and $\delta = \phi^{-1}(\gamma)$. Further, $\sigma_{\phi(\alpha)}(\phi(\delta)) = \phi(\sigma_\alpha(\delta))$ by Lemma 14.4, hence

$$\sigma_\beta(\gamma) = \sigma_{\phi(\alpha)}(\phi(\delta)) = \phi(\sigma_\alpha(\delta)) = \phi(\sigma_\alpha(\phi^{-1}(\gamma))) = (\phi \circ \sigma_\alpha \circ \phi^{-1})(\gamma) = \phi^*(\sigma_\alpha)(\gamma).$$

As $\gamma \in E$ was arbitrary, we have $\sigma_\beta = \phi^*(\sigma_\alpha)$, where $\sigma_\alpha \in R$. Therefore, as $\sigma_\beta \in R'$ was arbitrary, we have $R' \subseteq \phi^*(R)$. Thus, by (14.7.1),

$$|R'| \leq |\phi^*(R)| \leq |R| = |R'|,$$

hence $|R'| = |\phi^*(R)|$ and $R' = \phi^*(R)$. As these sets are finite and of equal size, we have that ϕ^* acts as a bijection between R and R'. As these are the generators of $\mathcal{W}$ and $\mathcal{W}'$, we have that ϕ^* is a bijection. □

14.8. **Proposition.** *The Weyl group of* Φ *is a subgroup of* $\mathrm{Aut}(\Phi)$.

Proof. Let $\sigma \in \mathcal{W}$. Then $\sigma \in \mathrm{GL}(E)$ and $\sigma(\Phi) = \Phi$. Therefore, by Theorem 14.3, σ is a root system automorphism of Φ.

So $\mathcal{W} \subseteq \mathrm{Aut}(\Phi)$, hence is a subgroup, as the group operation (composition of functions) is the same. □

14.9. **Definition.** Let $\alpha \in \Phi$. We define

$$\alpha^\vee = \frac{2\alpha}{\langle \alpha, \alpha \rangle}$$

and call

$$\Phi^\vee = \{\alpha^\vee : \alpha \in \Phi\}$$

the **dual** or **inverse** of Φ.

14.10. **Lemma.** *Let* $\alpha \in \Phi$. *Then* $\sigma_{\alpha^\vee} = \sigma_\alpha$.

Proof. As $\alpha^\vee$ is a scalar multiple of α, the hyperplanes orthogonal to α and $\alpha^\vee$ are the same, hence the reflections in those hyperplanes are the same. □

14.11. **Theorem.** $\Phi^\vee$ *is a root system in* E.

Proof. By definition, $|\Phi^\vee| \leq |\Phi|$, so is finite. The form is positive definite, so

$$\alpha^\vee = \frac{2\alpha}{\langle \alpha, \alpha \rangle}$$

exists and, as $\alpha \neq 0$, is nonzero for all $\alpha \in \Phi$. Therefore, $0 \notin \Phi^\vee$. Further,

$$\begin{aligned}
\mathrm{span}(\Phi^\vee) &= \sum_{\alpha^\vee \in \Phi^\vee} \mathbb{R}\alpha^\vee \\
&= \sum_{\alpha \in \Phi} \mathbb{R}\alpha^\vee \\
&= \sum_{\alpha \in \Phi} \mathbb{R}\left(\frac{2}{\langle \alpha, \alpha \rangle}\alpha\right) \\
&= \sum_{\alpha \in \Phi} \mathbb{R}\alpha \\
&= \mathrm{span}(\Phi) \\
&= E,
\end{aligned}$$

hence $\Phi^\vee$ spans E.

Note that, for $\alpha \in \Phi$,

$$(-\alpha)^\vee = \frac{2(-\alpha)}{\langle(-\alpha),(-\alpha)\rangle} = -\left(\frac{2\alpha}{\langle\alpha,\alpha\rangle}\right) = -(\alpha^\vee). \tag{14.11.1}$$

Suppose $\alpha^\vee$ and $\beta^\vee$ in $\Phi^\vee$ are scalar multiples. That is, $\beta^\vee = \lambda\alpha^\vee$ for some $\lambda \in \mathbb{R}$. Then

$$\frac{2\beta}{\langle\beta,\beta\rangle} = \lambda\frac{2\alpha}{\langle\alpha,\alpha\rangle},$$

which implies

$$\beta = \lambda\frac{\langle\beta,\beta\rangle}{\langle\alpha,\alpha\rangle}\alpha.$$

That is, β and α are scalar multiples. But as β and α are in Φ, we have that $\beta = \pm\alpha$. Therefore, either $\beta^\vee = \alpha^\vee$ or $\beta^\vee = (-\alpha)^\vee = -(\alpha^\vee)$, by (14.11.1). That is, the scalar multiples of $\alpha^\vee \in \Phi^\vee$ are precisely $\pm\alpha^\vee$.

Let $\alpha^\vee, \beta^\vee \in \Phi^\vee$. Then

$$(\sigma_\alpha(\beta))^\vee = \frac{2\sigma_\alpha(\beta)}{\langle\sigma_\alpha(\beta),\sigma_\alpha(\beta)\rangle}.$$

By Lemma 13.4, we have that

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

where we can express

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

Putting these together then gives

$$\langle\sigma_\alpha(\beta),\sigma_\alpha(\beta)\rangle = \langle\beta,\beta\rangle - 2(\beta,\alpha)\langle\beta,\alpha\rangle + 2(\beta,\alpha)\langle\beta,\alpha\rangle = \langle\beta,\beta\rangle.$$

Therefore, applying Lemma 13.4 again,

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

where the last step uses Lemma 14.10. Φ is a root system, so $\sigma_\alpha(\beta) \in \Phi$, hence $(\sigma_\alpha(\beta))^\vee = \sigma_{\alpha^\vee}(\beta^\vee) \in \Phi^\vee$ for all $\alpha, \beta \in \Phi$. That is, the reflections $\sigma_{\alpha^\vee}$ preserve $\Phi^\vee$.

Finally, let $\alpha^\vee, \beta^\vee \in \Phi^\vee$. Then,

$$(\beta^\vee,\alpha^\vee) = \frac{2\langle\beta^\vee,\alpha^\vee\rangle}{\langle\alpha^\vee,\alpha^\vee\rangle} = \frac{2\left\langle\frac{2\beta}{\langle\beta,\beta\rangle},\frac{2\alpha}{\langle\alpha,\alpha\rangle}\right\rangle}{\left\langle\frac{2\alpha}{\langle\alpha,\alpha\rangle},\frac{2\alpha}{\langle\alpha,\alpha\rangle}\right\rangle} = 2\frac{\left\langle\frac{\beta}{\langle\beta,\beta\rangle},\frac{\alpha}{\langle\alpha,\alpha\rangle}\right\rangle}{\left\langle\frac{\alpha}{\langle\alpha,\alpha\rangle},\frac{\alpha}{\langle\alpha,\alpha\rangle}\right\rangle}.$$

We can express the denominator as

$$\left\langle \frac{\alpha}{\langle \alpha, \alpha \rangle}, \frac{\alpha}{\langle \alpha, \alpha \rangle} \right\rangle = \frac{\langle \alpha, \alpha \rangle}{\langle \alpha, \alpha \rangle^2} = \frac{1}{\langle \alpha, \alpha \rangle},$$

hence

$$(\beta^\vee, \alpha^\vee) = 2 \left\langle \frac{\beta}{\langle \beta, \beta \rangle}, \frac{\alpha}{\langle \alpha, \alpha \rangle} \right\rangle \langle \alpha, \alpha \rangle = 2 \frac{\langle \beta, \alpha \rangle \langle \alpha, \alpha \rangle}{\langle \beta, \beta \rangle \langle \alpha, \alpha \rangle} = \frac{2 \langle \beta, \alpha \rangle}{\langle \alpha, \alpha \rangle} = (\beta, \alpha),$$

which is an integer, as Φ is a root system. □

14.12. **Proposition.** *The root systems $\Phi^\vee$ and Φ have the same Weyl group.*

Proof. Let $\mathcal{W}^\vee$ denote the Weyl group of $\Phi^\vee$. This is generated by the set $R^\vee = \{\sigma_{\alpha^\vee} : \alpha \in \Phi\}$ and $\mathcal{W}$ is generated by $R = \{\sigma_\alpha : \alpha \in \Phi\}$. By Lemma 14.10, for each $\alpha \in \Phi$, the reflections $\sigma_{\alpha^\vee}$ and σ_α are the same, hence $R^\vee = R$ and consequently $\mathcal{W}^\vee = \mathcal{W}$. □

14.13. **Corollary** (to Theorem 12.12)**.** *Let L be a simple Lie algebra over an algebraically closed field of characteristic zero. Then the set of roots is a root system.*

Proof. Now that we have defined what a root system is, we can see that this result is equivalent to Theorem 12.12. □

14.14. **Lemma.** *Let L be a semisimple Lie algebra over an algebraically closed field of characteristic zero. Let T denote a maximal torus of L, and Φ denote the set of roots. Then in the correspondence between T^* and T (Remark 8.12), the element of T which corresponds to $\alpha^\vee$ is*

$$h_\alpha = \frac{2t_\alpha}{\langle t_\alpha, t_\alpha \rangle}.$$

Proof. For all $t \in T$, we have that $\alpha(t) = \langle t_\alpha, t \rangle$, hence

$$\alpha^\vee(t) = \left(\frac{2\alpha}{\langle \alpha, \alpha \rangle} \right)(t) = \frac{2}{\langle t_\alpha, t_\alpha \rangle} \alpha(t) = \frac{2}{\langle t_\alpha, t_\alpha \rangle} \langle t_\alpha, t \rangle = \left\langle \frac{2t_\alpha}{\langle t_\alpha, t_\alpha \rangle}, t \right\rangle = \langle h_\alpha, t \rangle .$$

□

15. Pairs of Roots

We study root systems further, demonstrating restrictions on angles and length ratios between pairs of roots. We use this to prove an interesting result about root strings: they have length at most 4. Further, the α-string through β is reversed by the reflection σ_α (that is, the midpoint of the string lies on the hyperplane P_α).

We continue to take E to be a Euclidean space and Φ to be a root system in E.

15.1. **Theorem.** *Let $\alpha, \beta \in \Phi$ be nonproportional roots, labelled such that $||\alpha|| \leq ||\beta||$. Let θ be the angle between them. Then θ is one of the values in this table:*

(α,β)	(β,α)	θ	$\|\|\beta\|\|^2/\|\|\alpha\|\|^2$
0	0	$\pi/2$	-
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
1	2	$\pi/4$	2
-1	-2	$3\pi/4$	2
1	3	$\pi/6$	3
-1	-3	$5\pi/6$	3

Proof. As $||\alpha|| \leq ||\beta||$, we have that $\langle\alpha,\alpha\rangle \leq \langle\beta,\beta\rangle$, hence

$$\frac{1}{\langle\alpha,\alpha\rangle} \geq \frac{1}{\langle\beta,\beta\rangle}.$$

Therefore,

$$(\beta,\alpha) = \frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} \geq \frac{2\langle\beta,\alpha\rangle}{\langle\beta,\beta\rangle} = \frac{2\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle} = (\alpha,\beta).$$

As E is a Euclidean space, we can use the standard identity for the angle, θ, between α and β:

$$||\alpha||\,||\beta||\cos(\theta) = \langle\alpha,\beta\rangle.$$

This implies that,

$$(\beta,\alpha) = \frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} = 2\frac{||\beta||\,||\alpha||\cos(\theta)}{||\alpha||\,||\alpha||\cos(0)} = 2\frac{||\beta||}{||\alpha||}\cos(\theta),$$

which in turn gives

$$(\beta,\alpha)(\alpha,\beta) = \left(2\frac{||\beta||}{||\alpha||}\cos(\theta)\right)\left(2\frac{||\alpha||}{||\beta||}\cos(\theta)\right) = 4\cos^2(\theta). \tag{15.1.1}$$

As $\cos(\theta) \in [-1,1]$, we have that $\cos^2(\theta) \in [0,1]$, hence

$$(\beta,\alpha)(\alpha,\beta) \in [0,4].$$

Further, as Φ is a root system, both (β,α) and (α,β) are integers, hence

$$(\beta,\alpha)(\alpha,\beta) \in [0,4] \cap \mathbb{Z} = \{0,1,2,3,4\}.$$

If $(\beta,\alpha)(\alpha,\beta) = 4$, then by (15.1.1), $\cos(\theta) = 1$, hence $\theta = 0$, which implies that α and β are proportional. This contradicts the assumption, hence $(\beta,\alpha)(\alpha,\beta) \neq 4$. If either one of (β,α) or (α,β) are zero, they must both be zero. The rest of the table exhausts the possibilities for two integers multiplying together to give 1, 2 or 3.

For the ratios of norms, we can see that if $(\beta,\alpha) = n(\alpha,\beta)$ for some $n = 1,2,3$, we get

$$\frac{2\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} = n\frac{2\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle},$$

which simplifies to $\langle \beta, \beta \rangle = n \langle \alpha, \alpha \rangle$, hence

$$\frac{||\beta||^2}{||\alpha||^2} = \frac{\langle \beta, \beta \rangle}{\langle \alpha, \alpha \rangle} = \frac{n \langle \alpha, \alpha \rangle}{\langle \alpha, \alpha \rangle} = n.$$

□

15.2. **Lemma.** *Let $\alpha, \beta \in \Phi$ be nonproportional.*

- *If $\langle \alpha, \beta \rangle < 0$, then $\alpha + \beta \in \Phi$.*
- *If $\langle \alpha, \beta \rangle > 0$, then $\alpha - \beta \in \Phi$.*

Proof. Suppose $\langle \alpha, \beta \rangle > 0$. Then $(\alpha, \beta) > 0$, as the form is positive definite. By Theorem 15.1, we have at least one of $(\alpha, \beta) = 1$ or $(\beta, \alpha) = 1$.

By definition, Φ is closed under the reflections σ_α and σ_β. Suppose $(\alpha, \beta) = 1$. Then by Lemma 13.4,

$$\sigma_\beta(\alpha) = \alpha - (\alpha, \beta)\beta = \alpha - \beta \in \Phi.$$

Now suppose $(\beta, \alpha) = 1$. Then, again by Lemma 13.4,

$$\sigma_\alpha(\beta) = \beta - (\beta, \alpha)\alpha = \beta - \alpha \in \Phi,$$

which implies that $\alpha - \beta \in \Phi$, as Φ is closed under negation.

Now suppose that $\langle \alpha, \beta \rangle < 0$. Then $\langle \alpha, -\beta \rangle > 0$. We have that $-\beta \in \Phi$, so we have already shown that this implies $\alpha - (-\beta) = \alpha + \beta \in \Phi$. □

15.3. **Theorem.** *Let $\alpha, \beta \in \Phi$ be nonproportional. Then the α-string through β is reversed by σ_α, is unbroken, and has length at most 4.*

Proof. Let $r, q \in \mathbb{Z}$ be the largest integers for which $\beta - r\alpha \in \Phi$ and $\beta + q\alpha \in \Phi$ respectively. As $\beta \in \Phi$, these values exist and are nonnegative.

Suppose there exists some $i \in \mathbb{Z}$ strictly between r and q, such that $\beta + i\alpha \notin \Phi$. Consider the substring of non-roots surrounding $\beta + i\alpha$. Specifically, the end points of this string. For some $p \in \{-r, \cdots, i-1\}$, we have

$$\beta + p\alpha \in \Phi; \quad \beta + (p+1)\alpha \notin \Phi. \tag{15.3.1}$$

Similarly, for some $s \in \{i+1, \cdots, q\}$, we have

$$\beta + (s-1)\alpha \notin \Phi; \quad \beta + s\alpha \in \Phi. \tag{15.3.2}$$

By Lemma 15.2,

$$\langle \alpha, \beta + p\alpha \rangle < 0 \implies \alpha + (\beta + p\alpha) \in \Phi,$$

and

$$\langle \alpha, \beta + s\alpha \rangle > 0 \implies \alpha - (\beta + s\alpha) \in \Phi.$$

But by (15.3.1), $\beta + (p+1)\alpha \in \Phi$, hence $\langle \alpha, \beta + p\alpha \rangle \geq 0$. Additionally, by (15.3.2), $-(\beta + (s-1)\alpha) \notin \Phi$, hence $\beta + (s-1)\alpha \notin \Phi$, as Φ is closed under negation. Therefore, $\langle \alpha, \beta + s\alpha \rangle \leq 0$. That is,

$$\langle \alpha, \beta + s\alpha \rangle \leq 0 \leq \langle \alpha, \beta + p\alpha \rangle .$$

Therefore,

$$\langle \alpha, \beta \rangle + s \langle \alpha, \alpha \rangle \leq \langle \alpha, \beta \rangle + p \langle \alpha, \alpha \rangle ,$$

which can be simplified to

$$s \leq p,$$

because $\langle \alpha, \alpha \rangle > 0$, as the form is positive definite. But $p < s$ by definition, so this is a contradiction. Therefore, the supposition that $\beta + i\alpha \notin \Phi$ must be false. That is, the root string is unbroken.

We use the notation S_α^β to denote the set of roots in the α-string through β. Let $L = \{\beta + t\alpha : t \in \mathbb{R}\}$ denote the line through S_α^β. Let $L_s \subset L$ denote the closed line segment from $\beta - r\alpha$ to $\beta + q\alpha$. That is,

$$S_\alpha^\beta = L \cap \Phi \subset L_s \subset L.$$

The line L is, by definition, parallel to α, hence orthogonal to P_α. So the reflection σ_α reverses L about the point where L intersects P_α. Call this point χ. Therefore,

$$\sigma_\alpha(S_\alpha^\beta) \subseteq \sigma_\alpha(L) = L.$$

As Φ is a root system, it is preserved by σ_α. As $S_\alpha^\beta \subseteq \Phi$, we have that $\sigma_\alpha(S_\alpha^\beta) \subseteq \Phi$. Therefore,

$$\sigma_\alpha(S_\alpha^\beta) \subseteq L \cap \Phi = S_\alpha^\beta \tag{15.3.3}$$

Suppose σ_α does not reverse L_s. Then χ is not the midpoint of L_s (that is, P_α does not intersect L_s at its midpoint). Therefore, one of the endpoints of L_s lies further from χ than the other. Therefore, σ_α sends this endpoint outside L_s. That is, we have at least one of:

- $\sigma_\alpha(\beta - r\alpha) \notin L_s$;
- $\sigma_\alpha(\beta + q\alpha) \notin L_s$.

But as $S_\alpha^\beta \subset L_s$, this implies that σ_α sends at least one of $\beta - r\alpha$ or $\beta + q\alpha$ outside of S_α^β, which contradicts (15.3.3). Therefore, the supposition that σ_α does not reverse L_s must be false. That is, σ_α reverses L_s and, because the endpoints of L_s and S_α^β coincide, also reverses S_α^β.

Let $\gamma = \beta - r\alpha$. Let $S_\alpha^\gamma = \{\gamma - r'\alpha, \cdots, \gamma + q'\alpha\}$ denote the α-string through γ. This is the same as $S_\alpha^\beta = \{\beta - r\alpha, \cdots, \beta + q\alpha\}$, hence $r' = 0$ and $q' = r + q$. As σ_α reverses S_α^β, it maps the endpoints to one another. That is,

$$\sigma_\alpha(\beta') = \beta' + q'\alpha.$$

Therefore, by Lemma 13.4

$$\beta' - (\beta', \alpha)\alpha = \beta' + q'\alpha,$$

hence

$$(\beta', \alpha) = -q'.$$

Theorem 15.1 limits the possible values of (β', α), giving $|-q'| \leq 3$. As q' is nonnegative, this implies $q' \leq 3$. The length of the root string $S_\alpha^\beta = S_\alpha^{\beta'}$ is given by

$$|S_\alpha^\beta| = |\{\beta', \beta' + \alpha, \cdots, \beta' + q'\alpha\}| = |\{0, 1, \cdots, q'\}| = q' + 1.$$

Therefore, the length of the α-string through β is $q' + 1 \leq 4$. □

16. Bases of Root Systems

We define a base of a root system and prove that every root system has a base (Corollary 16.19) and that every base occurs in a certain form (Theorem 16.21). We will use bases to define a Chevalley basis in the following section, which will be instrumental in constructing Chevalley groups.

We continue to take E to be a Euclidean space and Φ to be a root system in E.

16.1. **Definition.** Let $\Delta \subseteq \Phi$. We call Δ a **base** of Φ if it is a basis for E and each $\beta \in \Phi$ can be expressed as

$$\beta = \sum_{\alpha \in \Delta} \lambda_\alpha \alpha, \tag{16.1.1}$$

for some collection of $\lambda_\alpha \in \mathbb{Z}$, either all nonnegative or all nonpositive. With respect to a given base, the roots contained in that base are called **simple**.

The **height** of a root β relative to Δ, denoted $\mathrm{ht}(\beta)$, is defined

$$\mathrm{ht}(\beta) = \sum_{\alpha \in \Delta} \lambda_\alpha,$$

where λ_α is the coefficient of α in the expression for β (16.1.1). These values are unique for each $\beta \in \Phi$, as Δ is a basis for E.

16.2. **Lemma.** *Let $\alpha \in \Delta$. Then $\mathrm{ht}(\alpha) = 1$.*

Proof. This is immediate from the definition. □

16.3. **Lemma.** *Let $\beta \in \Phi$. Express β as*

$$\beta = \sum_{\alpha \in \Delta} \lambda_\alpha \alpha.$$

If $\mathrm{ht}(\beta) > 0$, then $\lambda_\alpha \geq 0$ for all $\alpha \in \Delta$. If $\mathrm{ht}(\beta) < 0$, then $\lambda_\alpha \leq 0$ for all $\alpha \in \Delta$. As $0 \notin \Phi$, the value $\mathrm{ht}(\beta)$ is never zero.

Proof. By definition, the coefficients λ_α are either all nonnegative or all nonpositive. □

16.4. **Definition.** Let $\beta \in \Phi$. If $\mathrm{ht}(\beta) > 0$, call β **positive**; if $\mathrm{ht}(\beta) < 0$, call β **negative**. Denote the sets of positive and negative roots in Φ by Φ^+ and Φ^- respectively.

16.5. **Lemma.** *Φ can be expressed as a disjoint union $\Phi = \Phi^+ \cup \Phi^-$.*

Proof. By Lemma 16.3, we have that $\mathrm{ht}(\beta) \neq 0$ for all $\beta \in \Phi$. By definition Φ^+ and Φ^- do not intersect, so the union is disjoint. □

16.6. **Lemma.** *Let $\alpha, \beta \in \Delta$ be distinct. Then $\langle \alpha, \beta \rangle \leq 0$ and $(\alpha - \beta) \notin \Phi$.*

Proof. Let λ_γ be a collection of scalars indexed by $\gamma \in \Delta$, where all $\lambda_\gamma = 0$ except $\lambda_\alpha = 1$ and $\lambda_\beta = -1$. Then

$$\alpha - \beta - \sum_{\gamma \in \Delta} \lambda_\gamma \gamma.$$

This collection of scalars is not all nonpositive or all nonnegative, hence by definition of Δ, we have that $\alpha - \beta \notin \Phi$. The roots α and β are distinct elements of a basis, hence nonproportional. Therefore, by Lemma 15.2,

$$\langle \alpha, \beta \rangle > 0 \implies \alpha - \beta \in \Phi.$$

We have $\alpha - \beta \notin \Phi$, hence $\langle \alpha, \beta \rangle \leq 0$. □

16.7. **Definition.** Let $\gamma \in E$. Denote the subset of roots which lie on the same side of the hyperplane P_γ as γ by

$$\Phi^+(\gamma) = \{\alpha \in \Phi : \langle \alpha, \gamma \rangle > 0\},$$

and the subset of roots which lie on the opposite side of P_γ from γ by

$$\Phi^-(\gamma) = \{\alpha \in \Phi : \langle \alpha, \gamma \rangle < 0\}.$$

16.8. **Lemma.** *For all $\gamma \in E$, we have $\Phi^-(\gamma) = -\Phi^+(\gamma)$.*

Proof. Let $\alpha \in \Phi^-(\gamma)$. That is, $\langle \alpha, \gamma \rangle < 0$, hence $\langle -\alpha, \gamma \rangle > 0$. Therefore $-\alpha \in \Phi^+(\gamma)$, or equivalently $\alpha \in -\Phi^+(\gamma)$. This implies that $\Phi^-(\gamma) \subseteq -\Phi^+(\gamma)$.

Now let $\alpha \in -\Phi^+(\gamma)$. That is, $-\alpha \in \Phi^+(\gamma)$, hence $\langle -\alpha, \gamma \rangle > 0$. Therefore, $\langle \alpha, \gamma \rangle < 0$, hence $\alpha \in \Phi^-(\gamma)$. This implies that $-\Phi^+(\gamma) \subseteq \Phi^-(\gamma)$. □

16.9. **Definition.** Let $\gamma \in E$. Call γ **regular** if

$$\gamma \in E \setminus \bigcup_{\alpha \in \Phi} P_\alpha,$$

and **singular** otherwise.

16.10. **Lemma.** *Let $\gamma \in E$ be regular. Then $\Phi = \Phi^+(\gamma) \cup \Phi^-(\gamma)$.*

Proof. Suppose $\Phi \neq \Phi^+(\gamma) \cup \Phi^-(\gamma)$. Then there exists some $\alpha \in \Phi$ which is not an element of $\Phi^+(\gamma)$ or $-\Phi^+(\gamma)$. Then both $\langle \alpha, \gamma \rangle > 0$ and $-\langle \alpha, \gamma \rangle > 0$ are false. That is, both $\langle \alpha, \gamma \rangle \leq 0$ and $\langle \alpha, \gamma \rangle \geq 0$ are true, hence $\langle \alpha, \gamma \rangle = 0$. Therefore, $\gamma \in P_\alpha$. But as γ is regular, $\gamma \notin P_\alpha$, so we have a contradiction. Therefore the supposition must be false. □

16.11. **Definition.** Let $\gamma \in E$ be regular. Then $\alpha \in \Phi^+(\gamma)$ is called **decomposable** if $\alpha = \beta_1 + \beta_2$ for some $\beta_1, \beta_2 \in \Phi^+(\gamma)$, and **indecomposable** otherwise. Denote the **set of indecomposable roots in** $\Phi^+(\gamma)$ by $\Delta(\gamma)$.

16.12. **Lemma.** *Let $\gamma \in E$ be regular. Then each root in $\Phi^+(\gamma)$ is a nonnegative $\mathbb{Z}$-linear combination of roots in $\Delta(\gamma)$.*

Proof. Suppose otherwise. Then $\Phi^+(\gamma) \setminus Z \neq \emptyset$, where $Z = \operatorname{span}_{\mathbb{Z}^+}(\Delta(\gamma))$. Let $\alpha \in \Phi^+(\gamma) \setminus Z$ such that the value of $\langle \gamma, \alpha \rangle$ is minimal in this set. As $1 \in \mathbb{Z}^+$, we have that $\Delta(\gamma) \subseteq Z$, hence $\alpha \notin \Delta(\gamma)$. That is, α is decomposable, so can be expressed as $\alpha = \beta_1 + \beta_2$ for some $\beta_1, \beta_2 \in \Phi^+(\gamma)$. By definition of $\Phi^+(\gamma)$, both $\langle \gamma, \beta_1 \rangle$ and $\langle \gamma, \beta_2 \rangle$ are strictly positive. Further, $\langle \gamma, \alpha \rangle = \langle \gamma, \beta_1 \rangle + \langle \gamma, \beta_2 \rangle$, hence both $\langle \gamma, \beta_1 \rangle < \langle \gamma, \alpha \rangle$ and $\langle \gamma, \beta_2 \rangle < \langle \gamma, \alpha \rangle$. Therefore, $\beta_1, \beta_2 \notin \Phi^+(\gamma) \setminus Z$, as α was chosen to have minimal $\langle \gamma, \alpha \rangle$ in this set. That is, $\beta_1, \beta_2 \in \Phi^+(\gamma) \cap Z$. The set Z is closed under addition by definition, hence $\alpha + \beta_1 + \beta_2 \in Z$. But this contradicts the definition of α, hence such α cannot exist. This implies that $\Phi^+(\gamma) \setminus Z = \emptyset$ (as nonempty finite sets always contain their minimums). This contradicts the supposition, hence we have the result. □

16.13. **Lemma.** *Let $\gamma \in E$ be regular and let $\alpha, \beta \in \Delta(\gamma)$. Suppose $\alpha \neq \beta$. Then $\langle \alpha, \beta \rangle \leq 0$.*

Proof. Suppose $\langle \alpha, \beta \rangle > 0$. By definition, if $\alpha \in \Phi^+(\gamma)$, then $-\alpha \notin \Phi^+(\gamma)$. Therefore $\beta \neq -\alpha$, hence α and β are nonproportional. Therefore by Lemma 15.2, $\alpha - \beta \in \Phi$, hence by Lemma 16.10, either $\alpha - \beta \in \Phi^+(\gamma)$ or $\beta - \alpha \in \Phi^+(\gamma)$.

Suppose $\alpha - \beta \in \Phi^+(\gamma)$. Then we can express $\alpha = (\alpha - \beta) + \beta$, where both components of α are elements of $\Phi^+(\gamma)$. That is, α is decomposable, which is a contradiction, as $\alpha \in \Delta(\gamma)$. Therefore, the supposition must be false.

Now suppose $\beta - \alpha \in \Phi^+(\gamma)$. Then we can express $\beta = (\beta - \alpha) + \alpha$, where both components of β are elements of $\Phi^+(\gamma)$. That is, β is decomposable, which again is a contradiction, as $\beta \in \Delta(\gamma)$. Therefore, the supposition must be false.

Together, these points give a further contradiction, which implies that the original supposition must be false. That is, $\langle \alpha, \beta \rangle \leq 0$. □

16.14. **Lemma.** *Let $S \subset E$ be a nonempty set of vectors lying strictly on one side of some hyperplane. Let*

$$\alpha = \sum_{\beta \in S} \beta.$$

Then α lies on the same side of the hyperplane as the vectors in S.

Proof. Let P_γ denote the hyperplane in question, where $\gamma \in E$ is some nonzero vector orthogonal to it and on the same side as the vectors in S. Then each $\beta \in S$ satisfies $\langle \beta, \gamma \rangle > 0$ and $\beta \notin P_\gamma$. Thus we can write $\beta = \rho_\beta + \lambda_\beta \gamma$, for some $\rho_\beta \in P_\gamma$ and $\lambda_\gamma \in \mathbb{R}$ (Lemma 13.2). We have that $\langle \rho_\beta, \gamma \rangle = 0$, hence

$$0 < \langle \beta, \gamma \rangle = \langle \rho_\beta + \lambda_\beta \gamma, \gamma \rangle = \langle \rho_\beta, \gamma \rangle + \lambda_\beta \langle \gamma, \gamma \rangle = \lambda_\beta \langle \gamma, \gamma \rangle ,$$

where $\langle \gamma, \gamma \rangle > 0$ as the form is positive definite. Therefore, $\lambda_\beta > 0$. Now we can express α as

$$\alpha = \sum_{\beta \in S} \beta = \sum_{\beta \in S} (\rho_\beta + \lambda_\beta \gamma) = \left(\sum_{\beta \in S} \rho_\beta \right) + \left(\sum_{\beta \in S} \lambda_\beta \right) \gamma = \rho_\alpha + \lambda_\alpha \gamma,$$

where $\rho_\alpha \in P_\gamma$ and λ_α is a sum of strictly positive reals, hence is strictly positive. Therefore,

$$\langle \alpha, \gamma \rangle = \langle \rho_\alpha + \lambda_\alpha \gamma, \gamma \rangle = \langle \rho_\alpha, \gamma \rangle + \lambda_\alpha \langle \gamma, \gamma \rangle = \lambda_\alpha \langle \gamma, \gamma \rangle > 0,$$

again as the form is positive definite. That is, α lies on the positive side of P_γ. □

16.15. **Lemma.** *Let $S \subset E$ be a set of vectors lying strictly on one side of some hyperplane. Suppose that $\alpha \neq \beta$ implies $\langle \alpha, \beta \rangle \leq 0$ for all $\alpha, \beta \in S$. Then S is a linearly independent set.*

Proof. Suppose there exists some set $R = \{r_\alpha : \alpha \in S\} \subset \mathbb{R}$ which satisfies

$$\sum_{\alpha \in S} r_\alpha \alpha = 0.$$

Let $S^+ = \{\alpha \in S : r_\alpha > 0\}$ and $S^- = \{\alpha \in S : r_\alpha < 0\}$. Then $\alpha \in S \setminus (S^+ \cup S^-)$ implies $r_\alpha = 0$. Therefore,

$$0 = \sum_{\alpha \in S} r_\alpha \alpha = \sum_{\alpha \in S^+} r_\alpha \alpha + \sum_{\alpha \in S^-} r_\alpha \alpha.$$

For each $\alpha \in S^-$, let $t_\alpha = -r_\alpha$. Then each $t_\alpha > 0$ and we have

$$\sum_{\alpha \in S^+} r_\alpha \alpha = - \sum_{\alpha \in S^-} r_\alpha \alpha = \sum_{\alpha \in S^-} -r_\alpha \alpha = \sum_{\alpha \in S^-} t_\alpha \alpha.$$

Let

$$\varepsilon = \sum_{\alpha \in S^+} r_\alpha \alpha = \sum_{\alpha \in S^-} t_\alpha \alpha. \tag{16.15.1}$$

Then

$$\langle \varepsilon, \varepsilon \rangle = \left\langle \sum_{\alpha \in S^+} r_\alpha \alpha, \sum_{\beta \in S^-} t_\beta \beta \right\rangle = \sum_{\alpha \in S^+} \sum_{\beta \in S^-} r_\alpha t_\beta \langle \alpha, \beta \rangle .$$

As S^+ and S^- intersect trivially by definition, each pair of α and β in the above sum satisfy $\alpha \neq \beta$, which implies $\langle \alpha, \beta \rangle \leq 0$ by assumption. Further, each r_α and t_β appearing in the above sum is positive. Therefore, the sum must be negative or zero. That is, $\langle \varepsilon, \varepsilon \rangle \leq 0$, hence $\langle \varepsilon, \varepsilon \rangle = 0$ and $\varepsilon = 0$, as the form is positive definite.

As all the coefficients r_α and t_α appearing in (16.15.1) are positive, the vectors ($r_\alpha \alpha$ and $t_\alpha \alpha$) in each sum lie on the same side of the hyperplane as the vectors in S. Supposing that these sums are nonempty, we can apply Lemma 16.14 to get that ε also lies on that side of the hyperplane. But $\varepsilon = 0$, so cannot lie strictly on one side of any hyperplane, which is a contradiction. Therefore, the supposition that the sums are nonempty must be false. That is, $S^+ = S^- = \emptyset$, hence $r_\alpha = 0$ for all $\alpha \in S$.

We have shown that in an arbitrary expression of linear dependence on S, all coefficients are zero. That is, S is a linearly independent set. □

16.16. **Proposition.** *Let $\gamma \in E$ be regular. Then $\Delta(\gamma)$ is a base of Φ.*

Proof. Let $\beta \in \Phi$. By Lemma 16.10, we can express $\Phi = \Phi^+(\gamma) \cup \Phi^-(\gamma)$, so either $\beta \in \Phi^+(\gamma)$ or $\beta \in \Phi^-(\gamma)$. In the latter case, $-\beta \in \Phi^+(\gamma)$. Therefore, by Lemma 16.12, one of β or $-\beta$ can be expressed as a nonnegative $\mathbb{Z}$-linear combination of elements of $\Delta(\gamma)$. That is, β can be expressed as either a nonnegative or nonpositive $\mathbb{Z}$-linear combination of elements of $\Delta(\gamma)$.

Therefore, $\Delta(\gamma)$ spans Φ. As Φ spans E by definition, we have that $\Delta(\gamma)$ also spans E. By definition, $\Phi^+(\gamma)$ lies strictly on one side of some hyperplane, hence so does $\Delta(\gamma) \subseteq \Phi^+(\gamma)$. Further, by Lemma 16.13, $\alpha \neq \beta$ implies $\langle \alpha, \beta \rangle \leq 0$ for all $\alpha, \beta \in \Delta(\gamma)$. Therefore, by Lemma 16.15, $\Delta(\gamma)$ is linearly independent. Therefore, $\Delta(\gamma)$ is a basis for E. □

16.17. **Lemma.** *There exists $\gamma \in E$ which satisfies $\langle \gamma, \alpha \rangle > 0$ for all $\alpha \in \Delta$.*

Proof. Label the elements in Δ such that $\Delta = \{\alpha_1, \cdots, \alpha_n\}$, where $n = \dim(E)$. For each $k = 1, \cdots, n$, let $E_k = \operatorname{span}\{\alpha_1, \cdots, \alpha_k\}$.

Suppose $\gamma \in E_k$ satisfies

$$\langle \gamma, \alpha_i \rangle > 0 \tag{16.17.1}$$

for each $i = 1, \cdots, k$. Let

$$P_k = E_k^\perp = \{\delta \in E : \langle \delta, E_k \rangle = \{0\}\}.$$

Then $\dim(P_k) = n - k$ and $\dim(E_{k+1}) = k + 1$. We have that

$$\begin{aligned} \dim(P_k \oplus E_{k+1}) &= \dim(P_k) + \dim(E_{k+1}) - \dim(P_k \cap E_{k+1}) \\ &= (n - k) + (k + 1) - \dim(P_k \cap E_{k+1}) \\ &= n + 1 - \dim(P_k \cap E_{k+1}). \end{aligned}$$

But as these spaces are contained in E, we must have $\dim(P_k \oplus E_{k+1}) \leq n$. Thus $\dim(P_k \cap E_{k+1}) \geq 1$.

So there exists a nonzero $\delta \in P_k \cap E_{k+1}$. As $\delta \in P_k$, we have that

$$\langle \delta, \alpha_i \rangle = 0 \tag{16.17.2}$$

for each $i = 1, \cdots, k$.

Suppose that $\langle \delta, \alpha_{k+1} \rangle = 0$. With (16.17.2), this implies that $\langle \delta, E_{k+1} \rangle = \{0\}$. But as $\delta \in E_{k+1}$, this implies that $\langle \delta, \delta \rangle = 0$, which in turn implies that $\delta = 0$, as the form is positive definite. This contradicts the definition of δ, hence the supposition that $\langle \delta, \alpha_{k+1} \rangle = 0$ must be false. That is, $\langle \delta, \alpha_{k+1} \rangle \neq 0$.

For each $\lambda \in \mathbb{R}$, let $\gamma_\lambda = \gamma + \lambda\delta$. Then

$$\langle \gamma_\lambda, \alpha_{k+1} \rangle = \langle \gamma + \lambda\delta, \alpha_{k+1} \rangle = \langle \gamma, \alpha_{k+1} \rangle + \lambda \langle \delta, \alpha_{k+1} \rangle .$$

As $\langle \delta, \alpha_{k+1} \rangle \neq 0$, there exists some $\lambda \in \mathbb{R}$ for which $\langle \gamma_\lambda, \alpha_{k+1} \rangle > 0$. Fix λ to be such a value.

Then, by (16.17.2), we have

$$\langle \gamma_\lambda, \alpha_i \rangle = \langle \gamma + \lambda\delta, \alpha_i \rangle = \langle \gamma, \alpha_i \rangle + \lambda \langle \delta, \alpha_i \rangle = \langle \gamma, \alpha_i \rangle ,$$

for each $i = 1, \cdots, k$. By (16.17.1), we have $\langle \gamma, \alpha_i \rangle > 0$, hence $\langle \gamma_\lambda, \alpha_i \rangle > 0$.

We have shown that if there exists a $\gamma \in E_k$ satisfying $\langle \gamma, \alpha_i \rangle > 0$ for all $i = 1, \cdots, k$, then there must also exist a $\gamma_\lambda \in E_{k+1}$ satisfying $\langle \gamma_\lambda, \alpha_i \rangle > 0$ for all $i = 1, \cdots, k+1$.

Further, $\alpha_1 \in E_1$ and $\langle \alpha_1, \alpha_1 \rangle > 0$ as the form is positive definite. Therefore, by induction, there exists a $\gamma \in E = E_n$ satisfying $\langle \gamma, \alpha_i \rangle > 0$ for all $i = 1, \cdots, n$. □

16.18. **Lemma.** *Let $\gamma \in E$ satisfy $\langle \gamma, \alpha \rangle > 0$ for all $\alpha \in \Delta$. Then γ is regular.*

Proof. Let $\beta \in \Phi$. As Δ is a base, β can be expressed

$$\beta = \sum_{\alpha \in \Delta} \lambda_\alpha \alpha,$$

where the coefficients are either all nonnegative or all nonpositive. We have $\gamma \in P_\beta$ only if

$$\begin{aligned} 0 &= \langle \gamma, \beta \rangle \\ &= \left\langle \gamma, \sum_{\alpha \in \Delta} \lambda_\alpha \alpha \right\rangle \\ &= \sum_{\alpha \in \Delta} \lambda_\alpha \langle \gamma, \alpha \rangle . \end{aligned}$$

But by assumption, $\langle \gamma, \alpha \rangle > 0$ for each $\alpha \in \Delta$ and the coefficients all have the same sign, so $\langle \gamma, \beta \rangle \neq 0$ and $\gamma \notin P_\beta$.

As β was arbitrary, $\gamma \notin P_\beta$ for all $\beta \in \Phi$, hence γ is regular. □

16.19. **Corollary.** *There exists a base for Φ. Specifically, there exists a regular $\gamma \in E$ and $\Delta(\gamma)$ is a base.*

Proof. Together, Lemma 16.17 and Lemma 16.18 imply that there exists a regular $\gamma \in E$. Then Proposition 16.16 implies that $\Delta(\gamma)$ is a base □

16.20. **Lemma.** *Let $\gamma \in E$ satisfy $\langle \gamma, \alpha \rangle > 0$ for all $\alpha \in \Delta$. Then $\Phi^+ = \Phi^+(\gamma)$ and $\Phi^- = \Phi^-(\gamma)$.*

Proof. Let $\beta \in \Phi^+$. Then $\beta \in \Phi$ and can be expressed

$$\beta = \sum_{\alpha \in \Delta} \lambda_\alpha \alpha,$$

where $\lambda_\alpha \in \mathbb{R}^+$ for all $\alpha \in \Delta$. Therefore,

$$\langle \gamma, \beta \rangle = \left\langle \gamma, \sum_{\alpha \in \Delta} \lambda_\alpha \alpha \right\rangle = \sum_{\alpha \in \Delta} \lambda_\alpha \langle \gamma, \alpha \rangle > 0,$$

as $\langle \gamma, \alpha \rangle > 0$ for each $\alpha \in \Delta$ by assumption (we have $\langle \gamma, \beta \rangle \neq 0$ as $\beta \neq 0$), hence $\beta \in \Phi^+(\gamma)$. As β was arbitrary, this implies that $\Phi^+ \subseteq \Phi^+(\gamma)$.

Now let $\beta \in \Phi^-$. Then $-\beta \in \Phi^+$, so by the above, we have $-\beta \in \Phi^+(\gamma)$. By Lemma 16.8, $\Phi^-(\gamma) = -\Phi^+(\gamma)$, hence $\beta \in \Phi^-(\gamma)$. This implies that $\Phi^- \subseteq \Phi^-(\gamma)$. We have that $\Phi = \Phi^+ \cup \Phi^-$ (Lemma 16.5) and that $\Phi = \Phi^+(\gamma) \cup \Phi^-(\gamma)$ (Lemma 16.10). Both of these unions are disjoint. Therefore,

$$\Phi^+ \cup \Phi^- = \Phi^+(\gamma) \cup \Phi^-(\gamma),$$

hence

$$\begin{aligned}\Phi^+ &= (\Phi^+(\gamma) \cup \Phi^-(\gamma)) \setminus \Phi^- \\ &= \Phi^+(\gamma) \cup (\Phi^-(\gamma) \setminus \Phi^-),\end{aligned}$$

as $\Phi^- \subseteq \Phi^-(\gamma)$. But then the fact that the union is disjoint, along with $\Phi^+ \subseteq \Phi^+(\gamma)$, implies both that $\Phi^+ = \Phi^+(\gamma)$ and that $\Phi^-(\gamma) \setminus \Phi^- = \emptyset$, hence $\Phi^- = \Phi^-(\gamma)$. □

16.21. **Theorem.** *There exists a regular $\gamma \in E$ such that $\Delta = \Delta(\gamma)$.*

Proof. By Lemma 16.17, there exists a $\gamma \in E$ which satisfies $\langle \gamma, \alpha \rangle > 0$ for all $\alpha \in \Delta$.

Let $\beta \in \Delta$. Then $\beta \in \Phi^+ = \Phi^+(\gamma)$ by Lemma 16.2 and Lemma 16.20.

Suppose that β is decomposable with respect to γ. That is, $\beta = \beta_1 + \beta_2$ for some $\beta_1, \beta_2 \in \Phi^+(\gamma)$. As Δ is a base, we can express this as

$$\beta = \left(\sum_{\alpha \in \Delta} \lambda_\alpha \alpha \right) + \left(\sum_{\alpha \in \Delta} \mu_\alpha \alpha \right), \tag{16.21.1}$$

where each collection of scalars $\lambda_\alpha \in \mathbb{Z}$ and $\mu_\alpha \in \mathbb{Z}$ is either all nonnegative or all nonpositive. As $\beta_1, \beta_2 \in \Phi^+(\gamma) = \Phi^+$ by Lemma 16.2, these values are in fact all nonnegative.

We can rearrange (16.21.1) to give

$$(\lambda_\beta + \mu_\beta - 1)\beta + \sum_{\alpha \in \Delta \setminus \{\beta\}} (\lambda_\alpha + \mu_\alpha)\alpha = 0.$$

This is an expression of linear dependence on Δ, hence all the coefficients are zero, as Δ is a basis. That is, $\lambda_\alpha + \mu_\alpha = 0$ for all $\alpha \in \Delta \setminus \{\beta\}$ and $\lambda_\beta + \mu_\beta - 1 = 0$. As $\lambda_\alpha, \mu_\alpha \geq 0$ for all $\alpha \in \Delta$, this implies that $\lambda_\alpha, \mu_\alpha = 0$ for all $\alpha \in \Delta \setminus \{\beta\}$ and that one of λ_β or μ_β is zero (as $\lambda_\beta, \mu_\beta \in \mathbb{Z}$). But at least one λ_α and one μ_α must be nonzero, as both β_1 and β_2 are nonzero. We therefore have a contradiction, hence the supposition that β is decomposable must be false. That is, $\beta \in \Delta(\gamma)$.

Thus, we have shown that $\Delta \subseteq \Delta(\gamma)$. By Proposition 16.16, $\Delta(\gamma)$ is a base of Φ, and by definition Δ is a base of Φ. Therefore, $|\Delta| = \dim(E) = |\Delta(\gamma)|$, hence the inclusion $\Delta \subseteq \Delta(\gamma)$ implies that $\Delta = \Delta(\gamma)$. □

16.22. **Corollary.** *The root space decomposition of L can now be expressed:*

$$L = \left(\bigoplus_{\alpha \in \Delta} \mathbb{F} h_\alpha \right) \oplus \left(\bigoplus_{\alpha \in \Phi} \mathbb{F} x_\alpha \right),$$

where Δ is a base of Φ and $h_\alpha = [x_\alpha x_{-\alpha}]$ for each $\alpha \in \Phi$.

Proof. By Theorem 16.21, Φ has some base Δ. Then $\Phi \subseteq \mathrm{span}_{\mathbb{Z}}(\Delta)$ by definition. This implies that $\Phi \subseteq \mathrm{span}_{\mathbb{F}}(\Delta)$, as $\mathbb{F}$ has characteristic zero. Further, as Φ spans T^*, we have that $T^* = \mathrm{span}_{\mathbb{F}}(\Delta)$. The result then follows from Corollary 11.13. □

17. Chevalley Basis

In this section, we return to the Lie algebra setting, armed with the results about root systems from the previous sections, which we can apply to the set of roots, Φ, as this is a root system (Theorem 12.12). We will define and construct a special kind of basis for L - one in which all structure constrants are integers - called a Chevalley basis. For this, we will need a result about a certain automorphism of L (Lemma 17.1), which we will take as assumed (for a proof, see [4]).

It should be noted that the statement and proof of (Lemma 17.4) is slightly different from the corresponding result in [4] (Proposition 25.1b). The statement of the result differs so as to make its use in the proof of (Lemma 17.5) clear and accurate. The proof differs so that it (and hence Lemma 17.5 (corrispondingly Proposition 25.1c of [4])) do not depend on the classification theorem.

We continue to take L to be semisimple, T to be a maximal torus of L and $\mathbb{F}$ to be algebraically closed with characteristic zero.

17.1. **Lemma.** *There exists an automorphism σ of L, of order 2, which maps L_α to $L_{-\alpha}$ for each $\alpha \in \Phi$ and maps t to $-t$ for all $t \in T$.*

Proof. See Proposition 14.3 of [4]. □

17.2. Let $\alpha, \beta \in E$ and let θ be the angle between these two vectors. Then:

- $\theta < \pi/2 \iff \langle \alpha, \beta \rangle > 0$.
- $\theta = \pi/2 \iff \langle \alpha, \beta \rangle = 0$.
- $\theta > \pi/2 \iff \langle \alpha, \beta \rangle < 0$.

Consider the hyperplane P_α: this consists of points in E orthogonal to α, hence points with an angle of $\pi/2$ with α. We also have that $\langle \alpha, P_\alpha \rangle = \{0\}$ by definition.

If β lies on the same side of P_α as α, then $\theta < \pi/2$. This is the positive side of P_α, so $\langle \alpha, \beta \rangle > 0$. Similarly, if β is on the opposite side of P_α from α, then $\theta > \pi/2$ and $\langle \alpha, \beta \rangle < 0$.

17.3. **Lemma.** *Let $\alpha, \beta \in E$ be nonzero and let θ be the angle between them. Then the angle between α and $\sigma_\alpha(\beta)$ is $\pi - \theta$.*

Proof. Reflections preserve angles, so the angle between $\sigma_\alpha(\alpha) = -\alpha$ and $\sigma_\alpha(\beta)$ is also θ. Further, the angle between α and $-\alpha$ is π. As we are working in the 2 dimensional Euclidean plane spanned by α and β, the angles between α and $\sigma_\alpha(\beta)$ and $-\alpha$ must add up to π. Therefore, the angle between α and $\sigma_\alpha(\beta)$ is $\pi - \theta$. □

17.4. **Proposition.** *Let $\alpha, \beta \in \Phi$ be nonproportional. Let $S_\alpha^\beta = \{\beta - r\alpha, \cdots, \beta + q\alpha\}$ denote the α-string through β. Then at most two distinct root lengths occur in $S_\alpha^\beta \cup \{\alpha\}$.*

Proof. Without loss of generality, choose β in this root string such that $\langle \beta, \alpha \rangle$ is minimal. We have that

$$\langle \beta - \alpha, \alpha \rangle = \langle \beta, \alpha \rangle - \langle \alpha, \alpha \rangle < \langle \beta, \alpha \rangle ,$$

as the form is positive definite, hence $\beta - \alpha$ cannot be a root. Therefore, we can express the root string as

$$S_\alpha^\beta = \{\beta, \cdots, \beta + q\alpha\} . \tag{17.4.1}$$

The reflection σ_α reverses this root string (Theorem 15.3), hence

$$\sigma_\alpha(\beta + i\alpha) = \beta + (q - i)\alpha. \tag{17.4.2}$$

Let θ be the angle between α and β. Suppose $\theta < \pi/2$. By Lemma 17.3, the angle between α and $\sigma_\alpha(\beta)$ is $\pi - \theta$. By (17.4.2), we have $\sigma_\alpha(\beta) = \beta + q\alpha$, thus by the

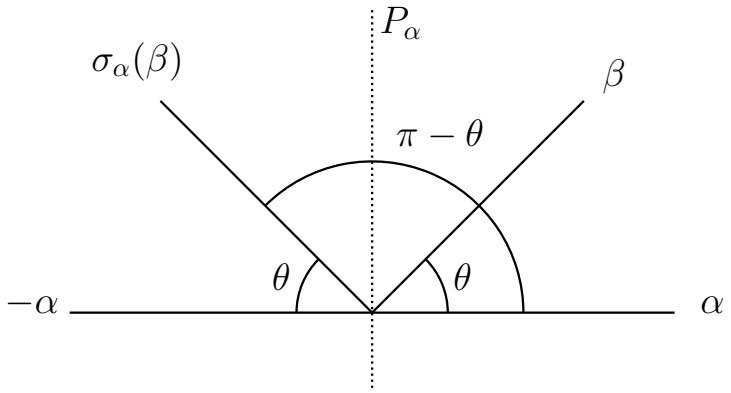

FIGURE 1.

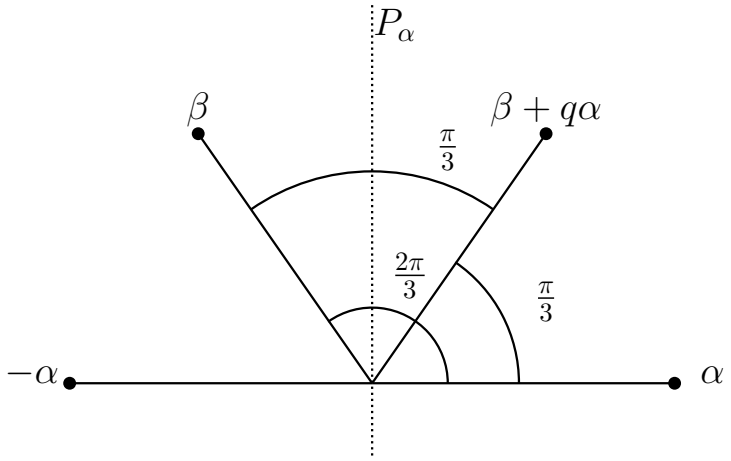

FIGURE 2.

supposition, the angle between α and $\beta + q\alpha$ is greater than $\pi/2$ (see Figure 1). That is, in view of Remark 17.2,

$$\langle \alpha, \beta + q\alpha \rangle < 0 < \langle \alpha, \beta \rangle ,$$

which contradicts the minimality of $\langle \alpha, \beta \rangle$. Therefore, $\theta \geq \pi/2$, hence by Theorem 15.1, we are restricted to

$$\theta \in \left\{ \frac{\pi}{2}, \frac{2\pi}{3}, \frac{3\pi}{4}, \frac{5\pi}{6} \right\}.$$

Suppose $\theta = \pi/2$. Then $(\beta, \alpha) = 0$ (Theorem 15.1). Suppose $q > 0$. Then by Lemma 13.4,

$$\sigma_\alpha(\beta + q\alpha) = \sigma_\alpha(\beta) + q\sigma_\alpha(\alpha) = \beta - (\beta, \alpha)\alpha - q\alpha = \beta - q\alpha$$

is also a root. But this contradicts (17.4.1), which implies there are no roots of this form. We conclude that $q = 0$ and $S_\alpha^\beta = \{\beta\}$.

Now suppose that $\theta = 2\pi/3$ (see Figure 2). By Lemma 17.3 and (17.4.2), the angle between α and $\beta + q\alpha$ is $\pi - \theta = \pi/3$. As $\beta + q\alpha$ lies in the arc between α and β, this implies that the angle between β and $\beta + q\alpha$ is $\theta - \pi/3 = \pi/3$. Further, the roots β and $\beta + q\alpha$ have the same length, as reflection preserves length. Therefore the points 0, β and $\beta + q\alpha$ form the vertices of an equilateral triangle (see Figure 3), hence the length of the edge from β to $\beta + q\alpha$ is equal to the length of β. These lengths are given by

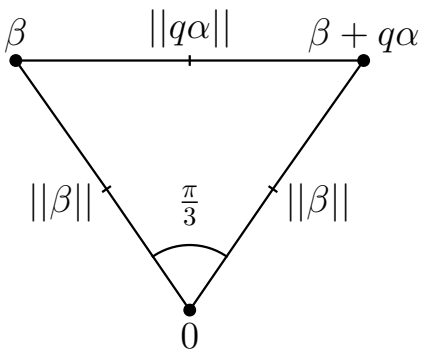

FIGURE 3.

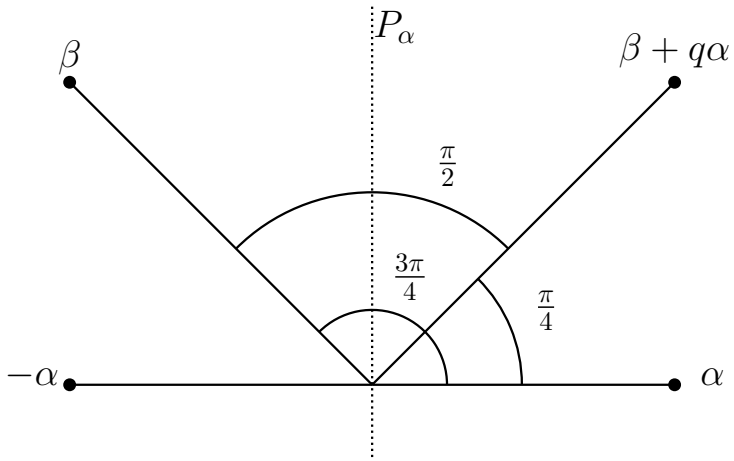

FIGURE 4.

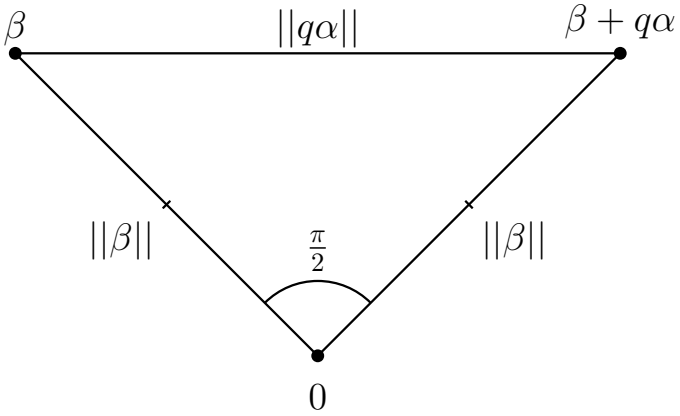

FIGURE 5.

$||\beta + q\alpha - \beta|| = ||q\alpha||$ and $||\beta||$ respectively. Therefore,

$$||\beta|| = ||q\alpha|| = q||\alpha||,$$

where $||\alpha|| = ||\beta||$ by Theorem 15.1, hence $q = 1$. That is, $S_\alpha^\beta = \{\beta, \beta + \alpha\}$, where both these roots have the same length.

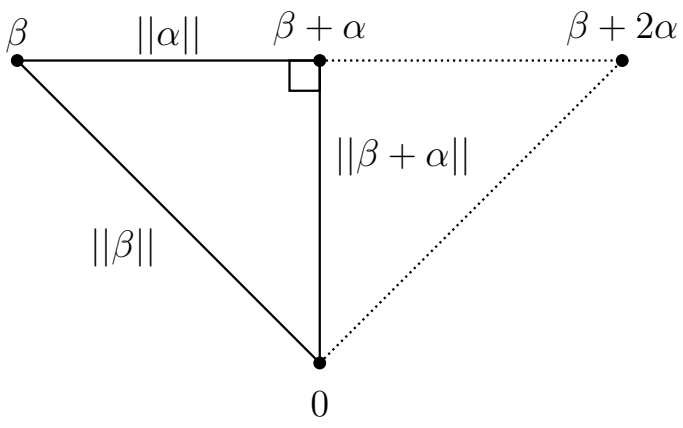

FIGURE 6.

Now suppose that $\theta = 3\pi/4$ and $||\alpha|| \leq ||\beta||$ (see Figure 4). By Lemma 17.3 and (17.4.2), the angle between α and $\beta + q\alpha$ is $\pi - 3\pi/4 = \pi/4$. Further, $||\beta|| = ||\beta + q\alpha||$, as reflections preserve length. The angles between α and $\beta + q\alpha$ and β must add to $3\pi/4$ (the angle between α and β), hence the angle between $\beta + q\alpha$ and β must equal $\pi/2$. We can then apply Pythagoras' Theorem on the triangle with vertices 0, β and $\beta + q\alpha$ (see Figure 5), giving

$$||q\alpha||^2 = ||\beta||^2 + ||\beta + q\alpha||^2 = 2||\beta||^2,$$

which implies that

$$q^2 = 2\frac{||\beta||^2}{||\alpha||^2} = 2 \cdot 2 = 4$$

by Theorem 15.1. Therefore, as q is a nonnegative integer, $q = 2$. This implies that $\beta + \alpha$ is the midpoint of the line segment from β to $\beta + q\alpha$, hence 0, β and $\beta + \alpha$ are the vertices of a right-angled triangle (see Figure 6). Therefore, we obtain

$$||\beta||^2 = ||\alpha||^2 + ||\beta + \alpha||^2,$$

which, as $||\beta||^2 = 2||\alpha||^2$ by Theorem 15.1, implies

$$||\beta + \alpha||^2 = ||\alpha||^2,$$

hence $||\beta + \alpha|| = ||\alpha||$. So we have that $S_\alpha^\beta = \{\beta, \beta + \alpha, \beta + 2\alpha\}$, where $||\beta|| = ||\beta + 2\alpha||$ and $||\beta + \alpha|| = ||\alpha||$.

Now suppose that $\theta = 3\pi/4$ and $||\alpha|| > ||\beta||$ (Figures 4 and 5 still apply). As in the previous case, the angle between β and $\beta + q\alpha$ is $\pi/2$ and $||\beta|| = ||\beta + q\alpha||$. This gives us

$$||q\alpha||^2 = ||\beta||^2 + ||\beta + q\alpha||^2 = 2||\beta||^2,$$

which implies that

$$q^2 = 2\frac{||\beta||^2}{||\alpha||^2} = \frac{2||\beta||^2}{2||\beta||^2} = 1,$$

as $||\alpha||^2 = 2||\beta||^2$ by Theorem 15.1. That is, $S_\alpha^\beta = \{\beta, \beta + \alpha\}$, where $||\beta|| = ||\beta + \alpha||$.

Now suppose that $\theta = 5\pi/6$ and $||\alpha|| \leq ||\beta||$ (see Figure 7). By Lemma 17.3 and (17.4.2), the angle between α and $\beta + q\alpha$ is $\pi - 5\pi/5 = \pi/6$. Therefore, the angle between β and $\beta + q\alpha$ is $2\pi/3$; further, these roots have the same length, as $\beta + q\alpha$ is the reflection of β in P_α. Consider the midpoint of the line segment between these roots:

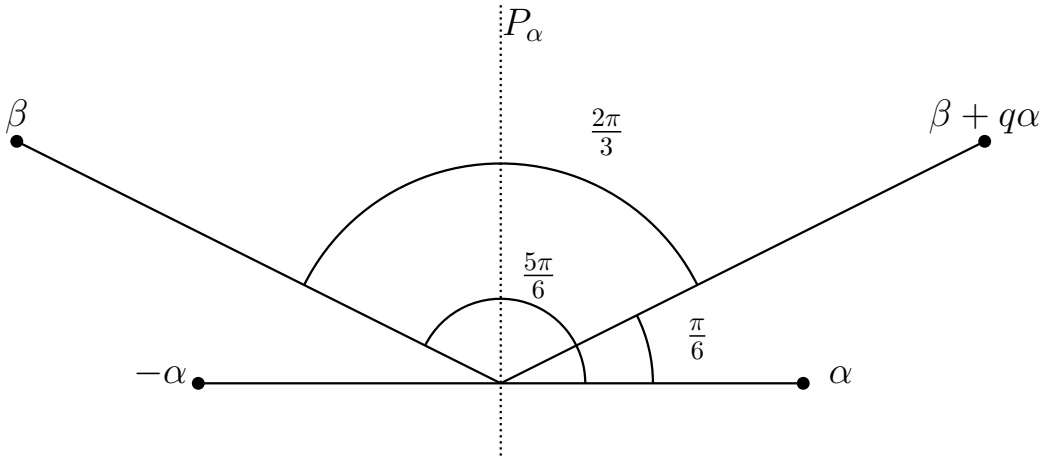

FIGURE 7.

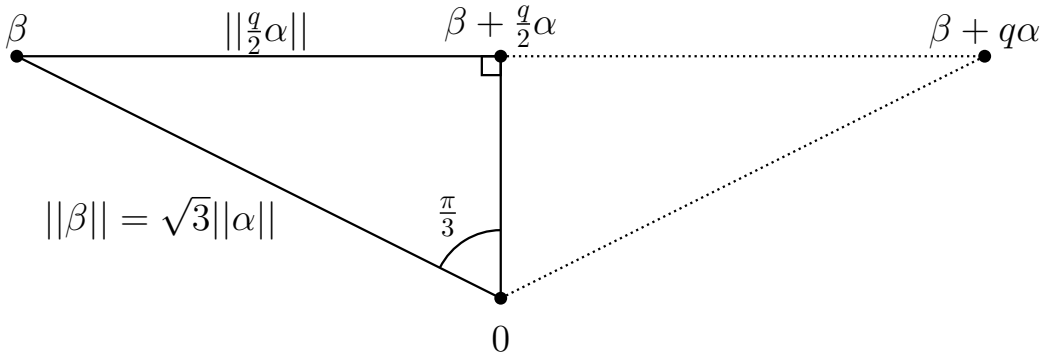

FIGURE 8.

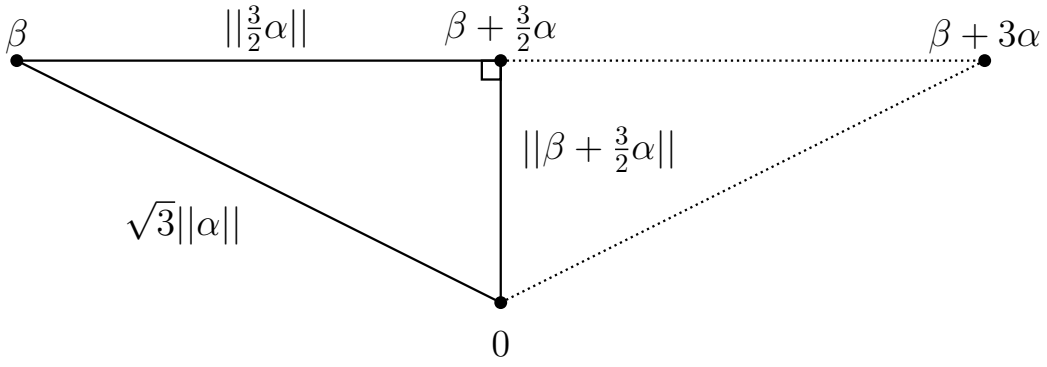

FIGURE 9.

$\beta + \frac{q}{2}\alpha$. This point, with β and 0, form the vertices of a right angled triangle with angle $\pi/3$ at the vertex 0 (see Figure 8). Therefore, applying trigonometry gives us

$$\sin\left(\frac{\pi}{3}\right) = \frac{||\frac{q}{2}\alpha||}{||\beta||} = \frac{q}{2} \cdot \frac{||\alpha||}{||\beta||} = \frac{q}{2} \cdot \frac{||\alpha||}{\sqrt{3}||\alpha||} = \frac{\sqrt{3}}{6}q,$$

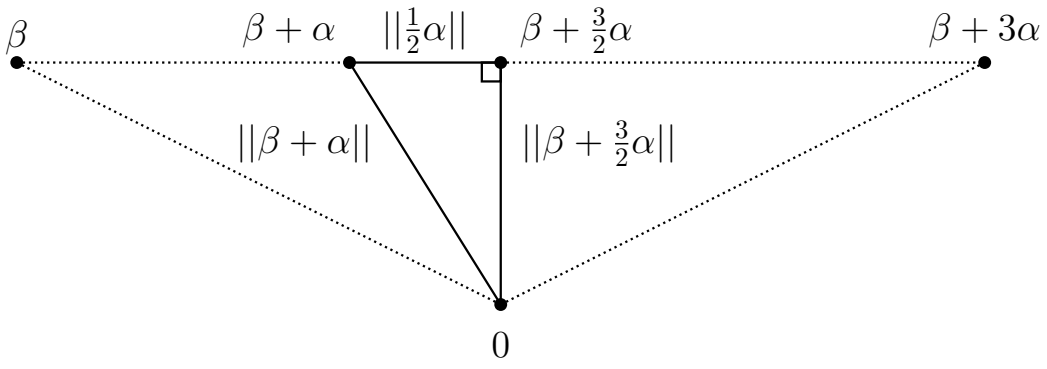

FIGURE 10.

as $||\beta|| = \sqrt{3}||\alpha||$ by Theorem 15.1, hence

$$\frac{\sqrt{3}}{2} = \frac{\sqrt{3}}{6}q,$$

thus $q = 3$. We can now find $||\beta + \frac{3}{2}\alpha||$ (see Figure 9). We have

$$||\beta||^2 = ||\frac{3}{2}\alpha||^2 + ||\beta + \frac{3}{2}\alpha||^2,$$

which, as $||\beta|| = \sqrt{3}||\alpha||$, implies

$$\begin{aligned} ||\beta + \frac{3}{2}\alpha||^2 &= 3||\alpha||^2 - (\frac{3}{2}||\alpha||)^2 \\ &= (3 - \frac{9}{4})||\alpha||^2 \\ &= \frac{3}{4}||\alpha||^2. \end{aligned}$$

Now consider the right angled triangle with vertices 0, $\beta + \alpha$ and $\beta + \frac{3}{2}\alpha$ (see Figure 10). From this, we get

$$||\beta + \alpha||^2 = ||\frac{1}{2}\alpha||^2 + ||\beta + \frac{3}{2}\alpha||^2 = \frac{1}{4}||\alpha||^2 + \frac{3}{4}||\alpha||^2 = ||\alpha||^2,$$

hence $||\beta + \alpha|| = ||\alpha||$. Further, by (17.4.2), we have $\sigma_\alpha(\beta + \alpha) = \beta + 2\alpha$, hence these two roots also have the same length. So we have $S_\alpha^\beta = \{\beta, \beta + \alpha, \beta + 2\alpha, \beta + 3\alpha\}$, where $||\alpha|| = ||\beta + \alpha|| = ||\beta + 2\alpha||$ and $||\beta|| = ||\beta + 3\alpha||$.

Now suppose that $\theta = 5\pi/6$ and $||\alpha|| > ||\beta||$. As in the previous case, the angle between β and $\beta + q\alpha$ is $2\pi/3$, and these roots have the same length, hence the points 0, β and $\beta + \frac{q}{2}\alpha$ form a right angled triangle with angle $\pi/3$ at the vertex 0. We again use trigonometry to get

$$\sin\left(\frac{\pi}{3}\right) = \frac{\frac{q}{2}||\alpha||}{||\beta||},$$

which, as $||\alpha|| = \sqrt{3}||\beta||$ by Theorem 15.1, implies

$$\frac{\sqrt{3}}{2} = \frac{\frac{q}{2}\sqrt{3}||\beta||}{||\beta||} = \frac{\sqrt{3}}{2}q,$$

hence $q = 1$. That is, $S_\alpha^\beta = \{\beta, \beta + \alpha\}$, where $||\beta|| = ||\beta + \alpha||$. □

17.5. **Lemma.** *Let $\alpha, \beta, \alpha+\beta \in \Phi$. Let $S_\alpha^\beta = \{\beta - r\alpha, \cdots, \beta + q\alpha\}$ denote the α-string through β. Then*

$$r + 1 = \frac{q\langle \alpha+\beta, \alpha+\beta\rangle}{\langle \beta, \beta\rangle}.$$

Proof. We have that $(\beta, \alpha) = r - q$ (Proposition 11.8), hence $r = q + (\beta, \alpha)$. Therefore,

$$\begin{aligned}
(r+1) - \frac{q\langle \alpha+\beta, \alpha+\beta\rangle}{\langle \beta, \beta\rangle} &= (q + (\beta, \alpha)) + 1 - \frac{q\langle \alpha+\beta, \alpha+\beta\rangle}{\langle \beta, \beta\rangle} \\
&= q + (\beta, \alpha) + 1 - \frac{q\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle} - \frac{2q\langle \alpha, \beta\rangle}{\langle \beta, \beta\rangle} - \frac{q\langle \beta, \beta\rangle}{\langle \beta, \beta\rangle} \\
&= (\beta, \alpha) + 1 - \frac{q\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle} - \frac{2q\langle \alpha, \beta\rangle}{\langle \beta, \beta\rangle} \\
&= (\beta, \alpha) + 1 - \frac{q\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle} - \frac{2q\langle \alpha, \beta\rangle\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle\langle \alpha, \alpha\rangle} \\
&= (\beta, \alpha) + 1 - \frac{q\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle} - \left(\frac{2\langle \beta, \alpha\rangle}{\langle \alpha, \alpha\rangle}\right)\left(\frac{q\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle}\right) \\
&= (\beta, \alpha) + 1 - \frac{q\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle} - (\beta, \alpha)\frac{q\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle} \\
&= ((\beta, \alpha) + 1)\left(1 - \frac{q\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle}\right) \\
&= AB,
\end{aligned}$$

where we let

$$A = (\beta, \alpha) + 1$$

and

$$B = 1 - \frac{q\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle}.$$

Now we need only show that $AB = 0$.

Suppose $\langle \alpha, \alpha\rangle \geq \langle \beta, \beta\rangle$. Then $|(\beta, \alpha)| \leq |(\alpha, \beta)|$, hence by Theorem 15.1, we have $(\beta, \alpha) \in \{-1, 0, 1\}$. If $(\beta, \alpha) = -1$, then $A = 0$. Otherwise, $(\beta, \alpha) \geq 0$, hence $\langle \beta, \alpha\rangle \geq 0$. Therefore,

$$\langle \beta+\alpha, \beta+\alpha\rangle = \langle \alpha, \alpha\rangle + \langle \beta, \beta\rangle + 2\langle \beta, \alpha\rangle \geq \langle \alpha, \alpha\rangle + \langle \beta, \beta\rangle .$$

As the form is positive definite, $\langle \alpha, \alpha\rangle, \langle \beta, \beta\rangle > 0$, hence $\langle \beta+\alpha, \beta+\alpha\rangle$ is strictly greater than both of these. That is, $||\beta+\alpha|| \neq ||\beta||, ||\alpha||$. As $\beta + \alpha \in \Phi$, we have $\beta + \alpha \in S_\alpha^\beta$, hence $||\alpha|| = ||\beta||$, because only two distinct root lengths occur in S_α^β (Proposition 17.4). That is,

$$\langle \alpha, \alpha\rangle = \langle \beta, \beta\rangle . \tag{17.5.1}$$

Further, $\langle \beta+2\alpha, \beta+2\alpha\rangle > \langle \beta+\alpha, \beta+\alpha\rangle$, hence $||\beta+2\alpha|| \neq ||\beta+\alpha||$. Therefore, $\beta + 2\alpha \notin \Phi$ (Proposition 17.4). That is, $q = 1$, hence by (17.5.1),

$$B = 1 - \frac{q\langle \alpha, \alpha\rangle}{\langle \beta, \beta\rangle} = 1 - \frac{\langle \alpha, \alpha\rangle}{\langle \alpha, \alpha\rangle} = 0.$$

Now suppose $\langle \alpha, \alpha\rangle < \langle \beta, \beta\rangle$. Then $||\alpha|| \neq ||\beta||$. As $\alpha + \beta \in \Phi$, Proposition 17.4 implies that $||\alpha+\beta|| \in \{||\alpha||, ||\beta||\}$, or equivalently,

$$\langle \alpha+\beta, \alpha+\beta\rangle \in \{\langle \alpha, \alpha\rangle, \langle \beta, \beta\rangle\}$$

We have that $\langle \alpha, \alpha \rangle, \langle \beta, \beta \rangle > 0$ as the form is positive definite, hence

$$\langle \beta + \alpha, \beta + \alpha \rangle \leq \langle \alpha, \alpha \rangle + \langle \beta, \beta \rangle .$$

But

$$\langle \beta + \alpha, \beta + \alpha \rangle = \langle \alpha, \alpha \rangle + \langle \beta, \beta \rangle + 2 \langle \beta, \alpha \rangle ,$$

hence

$$\langle \beta, \alpha \rangle < 0. \tag{17.5.2}$$

Therefore,

$$\langle \beta - \alpha, \beta - \alpha \rangle = \langle \beta, \beta \rangle + \langle \alpha, \alpha \rangle - 2 \langle \beta, \alpha \rangle > \langle \beta, \beta \rangle + \langle \alpha, \alpha \rangle > \langle \beta, \beta \rangle > \langle \alpha, \alpha \rangle .$$

Therefore, $\beta - \alpha \notin \Phi$ (Proposition 17.4). That is, $r = 0$. As $||\alpha|| < ||\beta||$, Theorem 15.1 implies $(\alpha, \beta) \in \{-1, 0, 1\}$. But by (17.5.2),

$$(\alpha, \beta) = \frac{\langle \alpha, \beta \rangle}{\langle \beta, \beta \rangle} < 0,$$

hence $(\alpha, \beta) = -1$. We have that $r = 0$ and $q = r - (\beta, \alpha)$. Therefore,

$$q = -(\beta, \alpha) = \frac{(\beta, \alpha)}{-1} = \frac{(\beta, \alpha)}{(\alpha, \beta)} = \frac{\langle \beta, \alpha \rangle \langle \beta, \beta \rangle}{\langle \alpha, \beta \rangle \langle \alpha, \alpha \rangle} = \frac{\langle \beta, \beta \rangle}{\langle \alpha, \alpha \rangle},$$

hence

$$B = 1 - \frac{q \langle \alpha, \alpha \rangle}{\langle \beta, \beta \rangle} = 1 - \frac{\langle \beta, \beta \rangle \langle \alpha, \alpha \rangle}{\langle \beta, \beta \rangle \langle \alpha, \alpha \rangle} = 0.$$

□

17.6. **Lemma.** *Let $\alpha, \beta \in \Phi$ be nonproportional. Choose $x_\alpha \in L_\alpha$ and $x_{-\alpha} \in L_{-\alpha}$ for which*

$$[x_\alpha x_{-\alpha}] = h_\alpha = \frac{2t_\alpha}{\langle t_\alpha, t_\alpha \rangle} \in T.$$

Let $x_\beta \in L_\beta$ be arbitrary. If $S_\alpha^\beta = \{\beta - r\alpha, \cdots, \beta, \cdots, \beta + q\alpha\}$ is the α-string through β, then

$$[x_{-\alpha}[x_\alpha x_\beta]] = q(r+1)x_\beta.$$

Proof. If $x_\beta = 0$, then the result holds. So suppose $x_\beta \neq 0$. If $\beta + \alpha \notin \Phi$, then $[L_\alpha L_\beta] = L_{\alpha+\beta} = \{0\}$ (Proposition 11.11), hence $[x_\alpha x_\beta] = 0$. This implies that $q = 0$, so the result holds.

So suppose also, that $\alpha + \beta \in \Phi$. Therefore, $L_{\alpha+\beta} \neq 0$, hence $L_{\alpha+\beta}$, L_α and L_β are all 1-dimensional (Theorem 11.2). Therefore, we can consider the subalgebra S_α. Recall that L is an S_α-module under the Lie bracket. Note that

$$[S_\alpha x_\beta] = [\operatorname{span}_{\mathbb{F}} \{x_\alpha, x_{-\alpha}, h_\alpha\} x_\beta] = [\operatorname{span}_{\mathbb{F}} \{x_\alpha, x_{-\alpha}, h_\alpha\} \operatorname{span}_{\mathbb{F}} \{x_\beta\}] = [S_\alpha L_\beta]$$

and

$$\begin{aligned} [S_\alpha L_{\beta+i\alpha}] &= [L_\alpha L_{\beta+i\alpha}] + [L_{-\alpha} L_{\beta+i\alpha}] + [\mathbb{F} h_\alpha L_{\beta+i\alpha}] \\ &= L_{\beta+(i+1)\alpha} + L_{\beta+(i-1)\alpha} + L_{\beta+i\alpha}, \end{aligned}$$

by Proposition 11.11. Therefore the S_α-module generated by x_β is

$$K = \sum_{i \in \mathbb{Z}} L_{\beta+i\alpha}.$$

Then by Lemma 11.5, we have that

$$K = L_{\beta-r\alpha} \oplus \cdots L_\beta \oplus \cdots \oplus L_{\beta+q\alpha},$$

is irreducible, and the weight spaces of K correspond to root spaces via

$$K_{\beta(h_\alpha)+2i} = L_{\beta+i\alpha}. \tag{17.6.1}$$

The highest weight on K is thus

$$\begin{aligned}\beta(h_\alpha) + 2q &= \left\langle t_\beta, \frac{2t_\alpha}{\langle t_\alpha, t_\alpha\rangle} \right\rangle + 2q \\ &= \frac{2\langle \beta, \alpha\rangle}{\langle \alpha, \alpha\rangle} + 2q \\ &= (\beta, \alpha) + 2q \\ &= (r - q) + 2q \\ &= r + q. \end{aligned} \tag{17.6.2}$$

By Corollary 9.14, there exists a maximal vector k_0 in K_{r+q}. By Proposition 9.12, there exist vectors

$$k_i = \frac{1}{i!}\operatorname{ad}^i(x_{-\alpha})(k_0),$$

and $\operatorname{ad}(x_{-\alpha})(k) \in K_{\lambda-2}$ for all $k \in K_\lambda$ and $\lambda \in \mathbb{F}$. Therefore,

$$\operatorname{ad}^q(x_{-\alpha})(k_0) \in K_{r+q-2q} = K_{r-q}.$$

That is, $k_q \in K_{r-q}$. As $x_\beta \in L_\beta$, we have that $x_\beta \in K_{\beta(h_\alpha)}$ by (17.6.1). As these root/weight spaces are 1-dimensional, by Theorem 11.2, x_β and k_q must be proportional. That is,

$$x_\beta = \mu k_q$$

for some $\mu \in \mathbb{F}$. Finally, $r+q$ is the highest weight on K by (17.6.2), hence by Proposition 9.12 we have

$$\begin{aligned}[x_\alpha k_q] &= ((r+q) - q + 1)k_{i-1} = (r+1)k_{i-1}, \\ [x_{-\alpha} k_{q-1}] &= ((q-1)+1)k_{q-1+1} = qk_q\end{aligned}$$

so we can calculate

$$\begin{aligned}[x_{-\alpha}[x_\alpha x_\beta]] &= [x_{-\alpha}[x_\alpha(\mu k_q)]] \\ &= \mu[x_{-\alpha}((r+1)k_{q-1})] \\ &= \mu(r+1)[x_{-\alpha}k_{q-1}] \\ &= \mu(r+1)(qk_q) \\ &= q(r+1)x_\beta.\end{aligned}$$

□

17.7. **Proposition.** *There exists a choice of $x_\alpha \in L_\alpha$ for each $\alpha \in \Phi$ which satisfies:*

- $[x_\alpha x_{-\alpha}] = h_\alpha$ *for all* $\alpha \in \Phi$*;*
- *If* $\alpha, \beta, \alpha+\beta \in \Phi$ *and* $[x_\alpha x_\beta] = c_{\alpha,\beta}x_{\alpha+\beta}$*, then* $c_{\alpha,\beta} = -c_{-\alpha,-\beta}$.

Proof. Let $\alpha \in \Phi$. By Lemma 17.1, there exists an automorphism σ of L of order 2, for which $\sigma(L_\alpha) = L_{-\alpha}$ and $\sigma(t) = -t$ for all $t \in T$.

Let $x_\alpha \in L_\alpha \setminus \{0\}$ and let $x_{-\alpha} = \sigma(x_\alpha)$. Then $x_{-\alpha} \in L_{-\alpha}$ by definition, and $x_{-\alpha} \neq 0$, otherwise σ maps a nonzero element to zero, which contradicts the bijectivity of σ.

Suppose $\langle x_\alpha, x_{-\alpha}\rangle = 0$. By Proposition 10.2, $[x_\alpha x_{-\alpha}] = \langle x_\alpha, x_{-\alpha}\rangle\, t_\alpha$, hence $[x_\alpha x_{-\alpha}] = 0$. But root spaces are 1-dimensional (Theorem 11.2), so this implies that $[L_\alpha L_{-\alpha}] = \{0\}$, which contradicts Proposition 10.3. Therefore, $\langle x_\alpha, x_{-\alpha}\rangle \neq 0$. We also have that $\langle \alpha, \alpha\rangle \neq$

0, as the form on E is positive definite. Therefore, as $\mathbb{F}$ is algebraically closed, we can define

$$c = \sqrt{\frac{2}{\langle x_\alpha, x_{-\alpha}\rangle \langle \alpha, \alpha\rangle}} \in \mathbb{F}.$$

Therefore, $cx_\alpha \in L_\alpha$ and $-\sigma(cx_\alpha) = cx_{-\alpha} \in L_{-\alpha}$. Then

$$\begin{aligned}\langle cx_\alpha, cx_{-\alpha}\rangle &= c^2 \langle x_\alpha, x_{-\alpha}\rangle \\ &= \frac{2}{\langle x_\alpha, x_{-\alpha}\rangle \langle \alpha, \alpha\rangle} \langle x_\alpha, x_{-\alpha}\rangle \\ &= \frac{2}{\langle \alpha, \alpha\rangle}.\end{aligned}$$

Therefore, by Proposition 10.2,

$$[(cx_\alpha)(cx_{-\alpha})] = \langle cx_\alpha, cx_{-\alpha}\rangle t_\alpha = \frac{2t_\alpha}{\langle \alpha, \alpha\rangle} = h_\alpha.$$

For each pair $(\alpha, -\alpha)$, we take this choice of root vectors for x_α and $x_{-\alpha}$, which exhausts all $\alpha \in \Phi$.

So fix a choice of x_α for $\alpha \in \Phi$ satisfying the first point. Let $\alpha, \beta \in \Phi$ and suppose $\alpha + \beta \in \Phi$. We have that $[x_\alpha x_\beta] \in L_{\alpha+\beta}$ (Proposition 6.6), and as $L_{\alpha+\beta}$ is 1-dimensional (Theorem 11.2), we also have that $[x_\alpha x_\beta] = c_{\alpha,\beta} x_{\alpha+\beta}$ for some $c_{\alpha,\beta} \in \mathbb{F}$. As roots occur with their negations, we must also have that $-\alpha, -\beta, -\alpha - \beta \in \Phi$ and $[x_{-\alpha} x_{-\beta}] = c_{-\alpha,-\beta} x_{-\alpha-\beta}$ for some $c_{-\alpha,-\beta} \in \mathbb{F}$. We have defined $x_{-\alpha-\beta} = -\sigma(x_{\alpha+\beta})$, thus

$$\sigma([x_\alpha x_\beta]) = \sigma(c_{\alpha,\beta} x_{\alpha+\beta}) = c_{\alpha,\beta}\sigma(x_{\alpha+\beta}) = -c_{\alpha,\beta} x_{-\alpha-\beta}.$$

On the other hand, as σ is a morphism,

$$\sigma([x_\alpha x_\beta]) = [\sigma(x_\alpha)\sigma(x_\beta)] = [(-x_{-\alpha})(-x_{-\beta})] = [x_{-\alpha} x_{-\beta}] = c_{-\alpha,-\beta} x_{-\alpha-\beta}.$$

Therefore, $c_{-\alpha,-\beta} = -c_{\alpha,\beta}$. □

17.8. **Lemma.** *Let $x_\alpha \in L_\alpha$ for each $\alpha \in \Phi$ such that Proposition 17.7 is satisfied. Then for all $\alpha, \beta \in \Phi$ for which $\alpha + \beta \in \Phi$, we have*

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

where $[x_\alpha x_\beta] = c_{\alpha,\beta} x_{\alpha+\beta}$ and $S_\alpha^\beta = \{\beta - r\alpha, \cdots, \beta, \cdots, \beta + q\alpha\}$.

Proof. Let $\alpha, \beta \in \Phi$ such that $\alpha + \beta \in \Phi$. Then α and β must be nonproportional (Theorem 12.12), hence t_α and t_β must be nonproportional and satisfy $t_\alpha + t_\beta = t_{\alpha+\beta}$ (Remark 8.12). We can calculate

$$\begin{aligned}[(c_{\alpha,\beta} x_{\alpha+\beta})(c_{\alpha,\beta} x_{-\alpha-\beta})] &= c_{\alpha,\beta}^2 [x_{\alpha+\beta} x_{-\alpha-\beta}] \\ &= c_{\alpha,\beta}^2 h_{\alpha+\beta} \\ &= c_{\alpha,\beta}^2 \frac{2t_{\alpha+\beta}}{\langle \alpha+\beta, \alpha+\beta\rangle} \\ &= c_{\alpha,\beta}^2 \frac{2(t_\alpha + t_\beta)}{\langle \alpha+\beta, \alpha+\beta\rangle}. \qquad (17.8.1)\end{aligned}$$

On the other hand, we have $c_{\alpha,\beta} x_{\alpha+\beta} = [x_\alpha x_\beta]$ and $c_{\alpha,\beta} x_{-\alpha-\beta} = -[x_{-\alpha} x_{-\beta}]$, hence

$$\begin{aligned}[(c_{\alpha,\beta} x_{\alpha+\beta})(c_{\alpha,\beta} x_{-\alpha-\beta})] &= -[[x_\alpha x_\beta][x_{-\alpha} x_{-\beta}]] \\ &= -[x_\alpha [x_\beta [x_{-\alpha} x_{-\beta}]]] + [x_\beta [x_\alpha [x_{-\alpha} x_{-\beta}]]]\end{aligned}$$

$$= [x_\alpha[x_\beta[x_{-\beta}x_{-\alpha}]]] + [x_\beta[x_\alpha[x_{-\alpha}x_{-\beta}]]]. \tag{17.8.2}$$

Consider these root strings involving $\pm\alpha$ and $\pm\beta$:

$$\begin{aligned}
S_\alpha^\beta &= \{\beta - r\alpha, \cdots, \beta + q\alpha\}, \\
S_{-\alpha}^{-\beta} &= \{-\beta - r(-\alpha), \cdots, -\beta + q(-\alpha)\} \\
S_\beta^\alpha &= \{\alpha - r'\beta, \cdots, \alpha + q'\beta\}, \\
S_{-\beta}^{-\alpha} &= \{-\alpha - r'(-\beta), \cdots, -\alpha + q'(-\beta)\},
\end{aligned}$$

for some nonnegative integers r' and q' (we define S_β^α with r' and q' then calculate the negative root strings using the fact that $\beta+i\alpha \in \Phi$ if and only if $-(\beta+i\alpha) = -\beta+i(-\alpha) \in \Phi$). By Lemma 17.6, we have that

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

Therefore, applying these to (17.8.2), we get

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

Then applying this to (17.8.1) gives

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

hence

$$c_{\alpha,\beta}^2\frac{2t_\alpha}{\langle\alpha+\beta,\alpha+\beta\rangle} + c_{\alpha,\beta}^2\frac{2t_\beta}{\langle\alpha+\beta,\alpha+\beta\rangle} = q'(r'+1)\frac{2t_\alpha}{\langle\alpha,\alpha\rangle} + q(r+1)\frac{2t_\beta}{\langle\beta,\beta\rangle}.$$

Which, as t_α and t_β are linearly independent, implies

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

hence

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

□

17.9. **Definition.** A **Chevalley basis** of a Lie algebra L is a basis $\{x_\alpha : \alpha \in \Phi\} \cup \{h_1, \cdots, h_n\}$ for which the set of x_α satisfies Proposition 17.7, and for which all $h_i = h_{\alpha_i}$ for some base $\Delta = \{\alpha_1, \cdots, \alpha_n\}$ of Φ.

17.10. **Theorem.** *Let $\{x_\alpha : \alpha \in \Phi\} \cup \{h_1, \cdots, h_n\}$ be a Chevalley basis of L. Then the associated structure constants are all integers. Specifically,*

- $[h_i h_j] = 0$ *for all* $i, j = 1, \cdots, n$.
- $[h_i x_\alpha] = (\alpha, \alpha_i)x_\alpha$ *for all* $\alpha \in \Phi$ *and* $i = 1, \cdots, n$.
- $[x_\alpha x_{-\alpha}] = h_\alpha$ *for all* $\alpha \in \Phi$, *where* $h_\alpha \in \operatorname{span}_{\mathbb{Z}}\{h_1, \cdots, h_n\}$.

- *For $\alpha, \beta \in \Phi$, then either $[x_\alpha x_\beta] = 0$ if $\alpha + \beta \in \Phi$ or $[x_\alpha x_\beta] = \pm(r+1)x_{\alpha+\beta}$ if $\alpha + \beta \in \Phi$, where $S_\alpha^\beta = \{\beta - r\alpha, \cdots, \beta + q\alpha\}$.*

Proof. For each $i, j = 1, \cdots, n$, we have that $h_i, h_j \in T$, hence $[h_i h_j] = 0$ as T is Abelian (Proposition 5.2).

Let $\alpha \in \Phi$ and $i = 1, \cdots, n$. By definition, $x_\alpha \in L_\alpha$, so $[t x_\alpha] = \alpha(t) x_\alpha$ for all $t \in T$. Specifically, $[h_i x_\alpha] = \alpha(h_i) x_\alpha$. We also have $\alpha(h_i) = \langle t_\alpha, h_i \rangle$ (Remark 8.12), hence

$$[h_i x_\alpha] = \langle t_\alpha, h_i \rangle x_\alpha = \left\langle t_\alpha, \frac{2t_{\alpha_i}}{\langle t_{\alpha_i}, t_{\alpha_i} \rangle} \right\rangle x_\alpha = \frac{2 \langle t_\alpha, t_{\alpha_i} \rangle}{\langle t_{\alpha_i}, t_{\alpha_i} \rangle} x_\alpha = (\alpha, \alpha_i) x_\alpha.$$

Let $\alpha \in \Phi$. By Theorem 14.11, $\Phi^\vee = \{\alpha^\vee : \alpha \in \Phi\}$ is a root system with base $\Delta^\vee = \{\alpha_1^\vee, \cdots, \alpha_n^\vee\}$, where

$$\alpha^\vee = \frac{2\alpha}{\langle \alpha, \alpha \rangle}.$$

By Lemma 14.14, $\alpha^\vee(t) = \langle h_\alpha, t \rangle$ for all $t \in T$. As $\Delta^\vee$ is a base for $\Phi^\vee$, there exist scalars $\mu_1, \cdots, \mu_n \in \mathbb{Z}$ such that

$$\alpha^\vee = \sum_{i=1}^n \mu_i \alpha_i^\vee.$$

Therefore, for all $t \in T$,

$$\langle h_\alpha, t \rangle = \alpha^\vee(t) = \sum_{i=1}^n \mu_i \alpha_i^\vee(t) = \sum_{i=1}^n \mu_i \langle h_i, t \rangle = \left\langle \sum_{i=1}^n \mu_i h_i, t \right\rangle,$$

hence

$$\left\langle h_\alpha - \sum_{i=1}^n \mu_i h_i, t \right\rangle = 0.$$

As the Killing form is nondegenerate, this implies that

$$h_\alpha = \sum_{i=1}^n \mu_i h_i.$$

Finally, let $\alpha, \beta \in \Phi$ be nonproportional and write $S_\alpha^\beta = \{\beta - r\alpha, \cdots, \beta + q\alpha\}$ for the α-string through β. Suppose that $\alpha + \beta \notin \Phi$, then $L_{\alpha+\beta} = \{0\}$. We have that $[x_\alpha x_\beta] \in L_{\alpha+\beta}$ (Proposition 6.6), hence $[x_\alpha x_\beta] = 0$. Now suppose that $\alpha + \beta \in \Phi$. Then by Lemma 17.5,

$$r + 1 = \frac{q \langle \alpha + \beta, \alpha + \beta \rangle}{\langle \beta, \beta \rangle}. \tag{17.10.1}$$

As $\{x_\alpha : \alpha \in \Phi\}$ is part of a Chevalley basis, it satisfies Proposition 17.7, hence $[x_\alpha x_\beta] = c_{\alpha,\beta} x_{\alpha+\beta}$, where

$$c_{\alpha,\beta}^2 = q(r+1) \frac{\langle \alpha + \beta, \alpha + \beta \rangle}{\langle \beta, \beta \rangle} \tag{17.10.2}$$

by Lemma 17.8. Rearranging the equations (17.10.1) and (17.10.2) gives

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

which implies that

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

hence $c_{\alpha,\beta} = \pm(r+1)$. □

18. Passing to a Finite Field

Up to this point, we have been working with a Lie algebra over an algebraically closed field of characteristic zero. We want to capture the Lie algebra structure of L over an arbitrary finite field. For a given field K of characteristic p, we reduce the vectors of L modulo p to get a Lie algebra over $\mathbb{F}_p$, then we extend to K. We can do this because, with respect to a Chevalley basis, all the structure constants are integers, so the Lie bracket remains closed.

We continue to take L to be semisimple and $\mathbb{F}$ to be algebraically closed with characteristic zero.

18.1. **Definition.** Let $B = \{x_\alpha : \alpha \in \Phi\} \cup \{h_1, \cdots, h_n\}$ be a Chevalley basis for L. Define $L(\mathbb{Z}) = \operatorname{span}_{\mathbb{Z}}(B)$.

18.2. Given some choice of x_α for each $\alpha \in \Phi$ which satisfies (Proposition 17.7), a Chevalley basis is then determined by a choice of Δ. The following lemma is to show that this choice of Δ does not affect the $L(\mathbb{Z})$ obtained from the resulting Chevalley basis. That is, $L(\mathbb{Z})$ is determined only by the choice of x_α.

18.3. **Lemma.** *$L(\mathbb{Z})$ is independent of the choice of Δ.*

Proof. Let Δ, Δ' be bases for Φ, with respective Chevalley basis B and B'. By Lemma 14.14, the map taking $\alpha \in T^*$ to $h_\alpha \in T$ is a bijective linear map. Write $h(\Delta) = \{h_\alpha : \alpha \in \Delta\} = \{h_1, \cdots, h_n\}$ and $h(\Delta') = \{h_\alpha : \alpha \in \Delta'\} = \{h'_1, \cdots, h'_n\}$. Then

$$B = \{x_\alpha : \alpha \in \Phi\} \cup h(\Delta), \tag{18.3.1}$$

$$B' = \{x_\alpha : \alpha \in \Phi\} \cup h(\Delta'). \tag{18.3.2}$$

By the definition of a base, we have that $\Phi \supseteq \Delta, \Delta'$ and $\Phi \subseteq \operatorname{span}_{\mathbb{Z}}(\Delta), \operatorname{span}_{\mathbb{Z}}(\Delta')$. The former implies that $\operatorname{span}_{\mathbb{Z}}(\Phi) \supseteq \operatorname{span}_{\mathbb{Z}}(\Delta), \operatorname{span}_{\mathbb{Z}}(\Delta')$; whereas the latter implies that $\operatorname{span}_{\mathbb{Z}}(\Phi) \subseteq \operatorname{span}_{\mathbb{Z}}(\operatorname{span}_{\mathbb{Z}}(\Delta)) = \operatorname{span}_{\mathbb{Z}}(\Delta)$ and similarly, $\operatorname{span}_{\mathbb{Z}}(\Phi) \subseteq \operatorname{span}_{\mathbb{Z}}(\Delta')$. Therefore,

$$\operatorname{span}_{\mathbb{Z}}(\Delta) = \operatorname{span}_{\mathbb{Z}}(\Phi) = \operatorname{span}_{\mathbb{Z}}(\Delta'). \tag{18.3.3}$$

Now we calculate

$$h(\operatorname{span}_{\mathbb{Z}}(\Delta)) = h\left(\sum_{\delta \in \Delta} \mathbb{Z}\delta\right) = \sum_{\delta \in \Delta} \mathbb{Z}h(\delta) = \sum_{h_\delta \in h(\Delta)} \mathbb{Z}h_\delta = \operatorname{span}_{\mathbb{Z}}(h(\Delta)).$$

Similarly, we have that $h(\operatorname{span}_{\mathbb{Z}}(\Delta')) = \operatorname{span}_{\mathbb{Z}}(h(\Delta'))$. Therefore, by (18.3.3), we have

$$\operatorname{span}_{\mathbb{Z}}(h(\Delta)) = h(\operatorname{span}_{\mathbb{Z}}(\Delta)) = h(\operatorname{span}_{\mathbb{Z}}(\Delta')) = \operatorname{span}_{\mathbb{Z}}(h(\Delta')).$$

Therefore, by (18.3.1) and (18.3.2), we have

$$\begin{aligned}
\operatorname{span}_{\mathbb{Z}}(B) &= \operatorname{span}_{\mathbb{Z}}(\{x_\alpha : \alpha \in \Phi\} \cup h(\Delta)) \\
&= \operatorname{span}_{\mathbb{Z}}(\{x_\alpha : \alpha \in \Phi\}) \oplus \operatorname{span}_{\mathbb{Z}}(h(\Delta)) \\
&= \operatorname{span}_{\mathbb{Z}}(\{x_\alpha : \alpha \in \Phi\}) \oplus \operatorname{span}_{\mathbb{Z}}(h(\Delta')) \\
&= \operatorname{span}_{\mathbb{Z}}(\{x_\alpha : \alpha \in \Phi\} \cup h(\Delta')) \\
&= \operatorname{span}_{\mathbb{Z}}(B').
\end{aligned}$$

□

18.4. **Lemma.** *The space $L(\mathbb{Z})$ is closed under the Lie bracket inherited from L.*

Proof. Let B denote the Chevalley basis for L from which $L(\mathbb{Z})$ is constructed. By Theorem 17.10, $[b_1 b_2] \in \operatorname{span}_{\mathbb{Z}}(B) = L(\mathbb{Z})$ for any $b_1, b_2 \in B$. Thus,

$$[L(\mathbb{Z})L(\mathbb{Z})] = \left[\left(\sum_{b\in B} \mathbb{Z}b\right)\left(\sum_{b\in B} \mathbb{Z}b\right)\right] = \sum_{b_1,b_2\in B} \mathbb{Z}[b_1 b_2],$$

which is contained in

$$\sum_{b_1,b_2\in B} \mathbb{Z}\operatorname{span}_{\mathbb{Z}}(B) = \mathbb{Z}\operatorname{span}_{\mathbb{Z}}(B) = \operatorname{span}_{\mathbb{Z}}(B) = L(\mathbb{Z}),$$

hence $[L(\mathbb{Z})L(\mathbb{Z})] \subseteq L(\mathbb{Z})$. □

18.5. **Definition.** Fix some prime p and define $L(\mathbb{Z})$ as above for some Chevalley basis B of L. Let $\mathbb{F}_p$ denote the prime field of characteristic p. We have that $L(\mathbb{Z})$ and $\mathbb{F}_p$ are $\mathbb{Z}$-modules under the actions

$$x \cdot n := \sum_{i=1}^{n} x \quad (x \in L(\mathbb{Z}), n \in \mathbb{Z}),$$

and

$$n \cdot \lambda := \sum_{i=1}^{n} \lambda \quad (n \in \mathbb{Z}, \lambda \in \mathbb{F}_p)$$

respectively. We can therefore define the tensor product

$$L(\mathbb{F}_p) = L(\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{F}_p = \left\{ \sum_{i=1}^{n} x_i \otimes \lambda_i : x_i \in L(\mathbb{Z});\ \lambda_i \in \mathbb{F}_p \right\},$$

where we identify $(x \cdot n) \otimes \lambda = x \otimes (n \cdot \lambda)$ for all $x \in L(\mathbb{Z})$, $n \in \mathbb{Z}$ and $\lambda \in \mathbb{F}_p$.

18.6. **Proposition.** $L(\mathbb{F}_p) \cong \operatorname{span}_{\mathbb{F}_p}(B)$.

Proof. Consider a pure tensor $x \otimes \lambda \in L(\mathbb{F}_p)$. We can write

$$x = \sum_{b\in B} n_b b,$$

for some $n_b \in \mathbb{Z}$. Therefore,

$$x \otimes \lambda = \left(\sum_{b\in B} n_b b\right) \otimes \lambda = \sum_{b\in B} (n_b b) \otimes \lambda = \sum_{b\in B} b \otimes (n_b \cdot \lambda),$$

where $n_b \cdot \lambda$ is just some element of $\mathbb{F}_p$. Arbitrary elements of $L(\mathbb{F}_p)$ can therefore be expressed as sums of such pure tensors, which can then also be expressed in the form

$$\sum_{b\in B} b \otimes \lambda_b,$$

for some $\lambda_b \in \mathbb{F}_p$, because $(b \otimes \lambda_1) + (b \otimes \lambda_2) = b \otimes (\lambda_1 + \lambda_2)$. For each element in $L(\mathbb{F}_p)$, the values λ_b are unique, and for any choice of λ_b, the above expression is an element of $L(\mathbb{F}_p)$. Therefore, mapping $b \otimes \lambda$ in $L(\mathbb{F}_p)$ to λb in $\operatorname{span}_{\mathbb{F}_p}(B)$ gives the desired isomorphism. □

18.7. **Definition.** We define a Lie bracket on $L(\mathbb{F}_p)$ by

$$[(x \otimes \lambda)(y \otimes \mu)]_{L(\mathbb{F}_p)} = [xy]_{L(\mathbb{Z})} \otimes \lambda\mu,$$

for each $(x \otimes \lambda), (y \otimes \mu) \in L(\mathbb{F}_p)$.

18.8. The Lie algebra $L(\mathbb{F}_p)$ has the same Lie bracket structure as L, unless p divides one of the structure constants of L. In that case, some of the nonzero structure constants in L become zero in $L(\mathbb{F}_p)$.

18.9. **Definition.** Let K be some field extension of $\mathbb{F}_p$. Then K is an $\mathbb{F}_p$-module under multiplication. As $L(\mathbb{F}_p)$ is a vector space over $\mathbb{F}_p$, it is an $\mathbb{F}_p$-module under scalar multiplication. We can therefore define the tensor product

$$L(K) = L(\mathbb{F}_p) \otimes_{\mathbb{F}_p} K = \left\{ \sum_{i=1}^{n} x_i \otimes \kappa_i : x_i \in L(\mathbb{F}_p);\ \kappa_i \in K \right\},$$

where we identify $(\lambda x) \otimes \kappa = x \otimes (\lambda \kappa)$ for all $x \in L(\mathbb{F}_p)$, $\lambda \in \mathbb{F}_p$ and $\kappa \in K$. By a similar argument to that in (Proposition 18.6), we have that

$$L(K) \cong \mathrm{span}_K(B).$$

Again, we define a Lie bracket on $L(K)$ by

$$[(x \otimes \lambda)(y \otimes \mu)]_{L(K)} = [xy]_{L(\mathbb{F}_p)} \otimes \lambda\mu,$$

for each $(x \otimes \lambda), (y \otimes \mu) \in L(K)$.

19. Constructing Chevalley Groups

The thrust of the previous section is that we can take any finite field K and construct a Lie algebra $L(K)$ over K with the same multiplication table as L (apart from the exceptions noted in Remark 18.8). Now, we construct an automorphism group over each $L(K)$ by exponentiating the adjoint maps of elements of the Chevalley basis (recall that, by construction, this is a basis for L, $L(\mathbb{F}_p)$ and $L(K)$). These groups are called Chevalley groups (of adjoint type). We also use some ideas from [1] in this final section.

It should be noted that, in this section, we give a substantially more thorough treatment of the results than appears in the literature [1, 4]. In particular, we prove that Chevalley groups are in fact automorphism groups of $L(K)$ (Proposition 19.14 and Theorem 19.16) and prove a concrete example of constructing a Chevalley group from a given Lie algebra (example 19.17).

We continue to take L to be semisimple, T to be a maximal torus of L and $\mathbb{F}$ to be algebraically closed with characteristic zero.

19.1. Recall that for $x \in L$, the map $\mathrm{ad}(x)$ is a derivation (Lemma 1.13). It therefore satisfies an identity known as the Leibniz rule. We prove the Leibniz rule for arbitrary derivations.

19.2. **Lemma.** *Let δ be a derivation of L and $a, b \in L$. Then for all $n \in \mathbb{N}$:*

$$\frac{\delta^n([ab])}{n!} = \sum_{i=0}^{n} \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n-i}(b)}{(n-i)!} \right) \right]. \tag{19.2.1}$$

Proof. We use induction on n. The base case is satisfied:

$$\begin{aligned} \frac{\delta^0([ab])}{0!} = [xy] &= \sum_{i=0}^{0} [ab] \\ &= \sum_{i=0}^{0} \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{0-i}(b)}{(0-i)!} \right) \right]. \end{aligned}$$

The inductive assumption is that (19.2.1) is satisfied for some $n \in \mathbb{N}$. This implies that

$$\begin{aligned} \frac{\delta^{n+1}([ab])}{(n+1)!} &= \frac{1}{(n+1)} \delta \left(\frac{\delta^n([ab])}{n!} \right) \\ &= \frac{1}{(n+1)} \delta \left(\sum_{i=0}^{n} \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n-i}(b)}{(n-i)!} \right) \right] \right) \\ &= \frac{1}{(n+1)} \sum_{i=0}^{n} \delta \left(\left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n-i}(b)}{(n-i)!} \right) \right] \right) \\ &= \frac{1}{(n+1)} \sum_{i=0}^{n} \left[\left(\frac{\delta^i(a)}{i!} \right) \delta \left(\frac{\delta^{n-i}(b)}{(n-i)!} \right) \right] + \left[\delta \left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n-i}(b)}{(n-i)!} \right) \right] \\ &= \frac{1}{(n+1)} \sum_{i=0}^{n} \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n-i)!} \right) \right] + \left[\left(\frac{\delta^{i+1}(a)}{i!} \right) \left(\frac{\delta^{n-i}(b)}{(n-i)!} \right) \right] \\ &= \frac{1}{(n+1)} \left(\sum_{i=0}^{n} \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n-i)!} \right) \right] + \sum_{i=0}^{n} \left[\left(\frac{\delta^{i+1}(a)}{i!} \right) \left(\frac{\delta^{n-i}(b)}{(n-i)!} \right) \right] \right) \\ &= \frac{1}{(n+1)} (A + B) \end{aligned}$$

We shall deal with the sums A and B separately. Firstly,

$$\begin{aligned}
A &= \sum_{i=0}^{n} \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n-i)!} \right) \right] \\
&= \sum_{i=0}^{n} \frac{n+1-i}{n+1-i} \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n-i)!} \right) \right] \\
&= \sum_{i=0}^{n} (n+1-i) \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right] \\
&= \sum_{i=0}^{n+1} (n+1-i) \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right],
\end{aligned}$$

as the $(n+1)$'th term is zero.

Secondly,

$$\begin{aligned}
B &= \sum_{i=0}^{n} \left[\left(\frac{\delta^{i+1}(a)}{i!} \right) \left(\frac{\delta^{n-i}(b)}{(n-i)!} \right) \right] \\
&= \sum_{i=1}^{n+1} \left[\left(\frac{\delta^{i}(a)}{(i-1)!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right] \\
&= \sum_{i=1}^{n+1} \frac{i}{i} \left[\left(\frac{\delta^{i}(a)}{(i-1)!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right] \\
&= \sum_{i=1}^{n+1} i \left[\left(\frac{\delta^{i}(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right] \\
&= \sum_{i=0}^{n+1} i \left[\left(\frac{\delta^{i}(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right],
\end{aligned}$$

as the 0'th term is zero. The sum of A and B is then

$$\begin{aligned}
A + B &= \sum_{i=0}^{n+1} (n+1-i) \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right] + i \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right] \\
&= \sum_{i=0}^{n+1} (n+1) \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right] = C.
\end{aligned}$$

Continuing from where we left off, we have

$$\begin{aligned}
\frac{\delta^{n+1}([ab])}{(n+1)!} &= \frac{1}{(n+1)} (A + B) \\
&= \frac{1}{(n+1)} C \\
&= \frac{1}{(n+1)} \sum_{i=0}^{n+1} (n+1) \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right] \\
&= \sum_{i=0}^{n+1} \left[\left(\frac{\delta^i(a)}{i!} \right) \left(\frac{\delta^{n+1-i}(b)}{(n+1-i)!} \right) \right].
\end{aligned}$$

□

19.3. **Lemma.** *Let $\alpha \in \Phi$ and $x \in L_\alpha$. Then* $\mathrm{ad}(x)$ *is nilpotent.*

Proof. We calculate

$$\mathrm{ad}(x)(L_0) \subseteq [L_\alpha L_0] \subseteq L_\alpha$$

(Proposition 6.6), which gives

$$\mathrm{ad}(x)^2(L_0) \subseteq \mathrm{ad}(x)(L_\alpha) \subseteq [L_\alpha L_\alpha] = \{0\}.$$

Now, for $\beta \in \Phi$ and $i \in \mathbb{N}$, we calculate

$$\mathrm{ad}(x)^i(L_\beta) \subseteq \mathrm{ad}(L_\alpha)^i(L_\beta) \subseteq L_{\beta+i\alpha}.$$

Root strings are finite (Theorem 15.3). Therefore, for some $i \in \mathbb{N}$, we have $L_{\beta+i\alpha} = \{0\}$ and hence $\mathrm{ad}(x)^i(L_\beta) = \{0\}$.

As L is finite dimensional, we can take $n \in \mathbb{N}$ to be the maximum of these i (and 2). Then, using the root space decomposition (Theorem 6.4)

$$L = L_0 \oplus \bigoplus_{\beta \in \Phi} L_\beta,$$

we have that $\mathrm{ad}(x)^n = 0$. □

19.4**.** For $\alpha \in \Phi$, we can express $\exp(\mathrm{ad}(x_\alpha))$ as a finite sum:

$$\exp(\mathrm{ad}(x_\alpha)) = \sum_{i=0}^{m} \frac{\mathrm{ad}(x_\alpha)^i}{i!}.$$

This is because $\mathrm{ad}(x_\alpha)$ is nilpotent by Lemma 19.3.

19.5. **Lemma.** *Let δ be a nilpotent derivation of L. Then* $\exp(\delta)$ *is a Lie algebra homomorphism. That is,* $\exp(\delta)$ *is linear and, for all $a, b \in L$,*

$$\exp(\delta)([ab]) = [(\exp(\delta))(a)\,(\exp(\delta))(b)].$$

Proof. The linearity of $\exp(\delta)$ follows from the linearity of δ.

Let $a, b \in L$. As δ is nilpotent, let $n \in \mathbb{N}$ be such that $m > n \implies \delta^m = 0$. Then

$$\begin{aligned}[(\exp(\delta))(a)\,(\exp(\delta))(b)] &= \left[\left(\sum_{i=0}^{n} \frac{\delta^i(a)}{i!}\right)\left(\sum_{i=0}^{n} \frac{\delta^i(b)}{i!}\right)\right] \\ &= \sum_{i,j=0}^{n} \left[\left(\frac{\delta^i(a)}{i!}\right)\left(\frac{\delta^j(b)}{j!}\right)\right]. \qquad (19.5.1)\end{aligned}$$

We want to put this sum into a form where we can simplify using the Leibniz rule (19.2.1). Notice that we are summing over $\{(i,j) \in \mathbb{Z}^2 : 0 \le i, j \le n\}$:

$$\begin{array}{ccc} (0,0) & \cdots & (0,n) \\ \vdots & \ddots & \vdots \\ (n,0) & \cdots & (n,n) \end{array}$$

Any method of iteration which exhausts this list will give us an equivalent expression for (19.5.1). We can rearrange this list such that each row corresponds to each upwards diagonal of the previous:

$$\begin{array}{ccc} (0,0) & & \\ \vdots & \ddots & \\ (n,0) & \cdots & (n,0) \\ & \ddots & \vdots \\ & & (n,n) \end{array}$$

As $\delta^m = 0$ for all $m > n$, we can include terms (i, j) with either of $i, j > n$ and retain equality with (19.5.1):

$$\begin{array}{ccccccc} (0,0) & & & & & & \\ \vdots & \ddots & & & & & \\ (n,0) & \cdots & (n,0) & & & & \\ \vdots & \ddots & \vdots & \ddots & & & \\ (2n,0) & \cdots & (n,n) & \cdots & (0,2n) & & \end{array}$$

In this arrangement, the ith row contains the terms

$$\{(i,0), \cdots, (0,i)\} = \{(j, i-j) : j = 0, \cdots, i\}.$$

This gives us the expression for (19.5.1):

$$\sum_{i=0}^{2n} \sum_{j=0}^{i} \left[\left(\frac{\delta^j(a)}{j!} \right) \left(\frac{\delta^{i-j}(b)}{(i-j)!} \right) \right].$$

From here we can apply (19.2.1) to give

$$\begin{aligned} [(\exp(\delta))(a)(\exp(\delta))(b)] &= \sum_{i=0}^{2n} \frac{\delta^i([ab])}{i!} \\ &= \sum_{i=0}^{n} \frac{\delta^i([ab])}{i!} \\ &= (\exp(\delta))([ab]), \end{aligned}$$

again using the fact that $\delta^m = 0$ for $m > n$. □

19.6. **Proposition.** *Let δ be a nilpotent derivation of some Lie algebra L. Then $\exp(\delta)$ is an automorphism of L, with inverse*

$$(\exp(\delta))^{-1} = \sum_{i=0}^{n} (-1)^i \eta^i,$$

where $\eta = \exp(\delta) - 1$ and n is the largest power for which $\delta \neq 0$.

Proof. By Lemma 19.5, we have that $\exp(\delta)$ is a homomorphism, so we need only show that the product with the claimed inverse equals 1.

$$\begin{aligned} \left(\sum_{i=0}^{n} (-1)^i \eta^i \right) (\exp(\delta)) &= \left(\sum_{i=0}^{n} (-1)^i \eta^i \right) (1 + \eta) \\ &= \left(\sum_{i=0}^{n} (-1)^i \eta^i \right) + \eta \left(\sum_{i=0}^{n} (-1)^i \eta^i \right) \\ &= \left(\sum_{i=0}^{n} (-1)^i \eta^i \right) + \left(\sum_{i=0}^{n} -(-1)^{i+1} \eta^{i+1} \right) \\ &= \left(\sum_{i=0}^{n} (-1)^i \eta^i \right) + \left(\sum_{i=1}^{n+1} -(-1)^i \eta^i \right) \\ &= (-1)^0 \eta^0 + \left(\sum_{i=1}^{n} (-1)^i \eta^i - (-1)^i \eta^i \right) - (-1)^{n+1} \eta^{n+1} \end{aligned}$$

$$
\begin{aligned}
&= 1 + \left(\sum_{i=1}^{n} 0\right) \pm \eta^{n+1} \\
&= 1 \pm \eta^{n+1} \\
&= 1 \pm (\exp(\delta) - 1)^{n+1} \\
&= 1 \pm \left(\left(\sum_{i=0}^{n} \frac{\delta^i}{i!}\right) - 1\right)^{n+1} \\
&= 1 \pm \left(1 + \left(\sum_{i=1}^{n} \frac{\delta^i}{i!}\right) - 1\right)^{n+1} \\
&= 1 \pm \left(\sum_{i=1}^{n} \frac{\delta^i}{i!}\right)^{n+1} \\
&= 1 \pm \left(\sum_{i=1}^{n} \frac{\delta^i}{i!}\right)^{n+1}.
\end{aligned}
$$

The lowest power of δ in the above sum is 1, hence the lowest power occurring in the expression as a whole is $n+1$. However, $\delta^{n+1} = 0$, hence

$$1 \pm \left(\sum_{i=1}^{n} \frac{\delta^i}{i!}\right)^{n+1} = 1 \pm 0 = 1.$$

□

19.7. **Lemma.** *Let $\alpha \in \Phi$ and $m \in \mathbb{N}$. Then*

$$\frac{\operatorname{ad}(x_\alpha)^m}{m!}(L(\mathbb{Z})) \subseteq L(\mathbb{Z}).$$

Proof. We can express the Chevalley basis as $B = \{x_\alpha : \alpha \in \Phi\} \cup \{h_1, \cdots, h_n\}$. Fix some $\alpha \in \Phi$ and let $f_m = \operatorname{ad}(x_\alpha)^m / m!$.

Suppose $f_m(B) \subseteq L(\mathbb{Z})$. Let $x \in L(\mathbb{Z})$. Then

$$x = \sum_{b \in B} \lambda_b b$$

for some collection of scalars $\lambda_b \in \mathbb{Z}$. Therefore,

$$f_m(x) = f_m\left(\sum_{b \in B} \lambda_b b\right) = \sum_{b \in B} \lambda_b f_m(b),$$

hence $f(x) \in \operatorname{span}_{\mathbb{Z}}(f_m(B))$. By the supposition, this implies that

$$f_m(x) \in \operatorname{span}_{\mathbb{Z}}(f_m(B)) \subseteq \operatorname{span}_{\mathbb{Z}}(L(\mathbb{Z})) = L(\mathbb{Z}).$$

As x was arbitrary, we have $f_m(L(\mathbb{Z})) \subseteq L(\mathbb{Z})$. That is,

$$f_m(B) \subseteq L(\mathbb{Z}) \implies f_m(L(\mathbb{Z})) \subseteq L(\mathbb{Z}). \tag{19.7.1}$$

By definition, for $m \geq 1$,

$$f_m(x) = \frac{1}{m} \operatorname{ad}(x_\alpha)(f_{m-1}(x)). \tag{19.7.2}$$

Therefore, if $x \in L$ and $l \geq m$, then

$$f_m(x) = 0 \implies f_l(x) = 0. \tag{19.7.3}$$

For $m = 1$, we have

$$f_m(x) = \mathrm{ad}(x_\alpha)(x_\alpha) = [x_\alpha x_\alpha] = 0,$$

hence by (19.7.3), we have $f_m(x_\alpha) = 0$ for all $m \in \mathbb{N}$.

By Theorem 17.10,

$$[h_i x_\alpha] = (\alpha, \alpha_i) x_\alpha \in \mathbb{Z} x_\alpha,$$

hence for $m = 1$, we also have

$$f_m(h_i) = \mathrm{ad}(x_\alpha)(h_i) = [x_\alpha h_i] = -[h_i x_\alpha] = -(\alpha, \alpha_i) x_\alpha \in L(\mathbb{Z}),$$

for each $i = 1, \cdots, n$. Then for $m = 2$, (19.7.2) implies that

$$f_m(h_i) = \frac{1}{2}\,\mathrm{ad}(x_\alpha)(f_1(h_i)) = \frac{1}{2}\,\mathrm{ad}(x_\alpha)(-(\alpha, \alpha_i) x_\alpha) = \frac{1}{2}(\alpha, \alpha_i)[x_\alpha x_\alpha] = 0.$$

Therefore, by (19.7.3), we have $f_m(h_i) \in L(\mathbb{Z})$ for all $m \in \mathbb{N}$.

By Theorem 17.10 again, we have $[x_\alpha x_{-\alpha}] = h_\alpha$, where $h_\alpha \in \mathrm{span}_{\mathbb{Z}} \{h_1, \cdots, h_n\} \subset L(\mathbb{Z})$, hence for $m = 1$, we have

$$f_m(x_{-\alpha}) = \mathrm{ad}(x_\alpha)(x_{-\alpha}) = h_\alpha.$$

Then, for $m = 2$, (19.7.2) implies that

$$f_m(x_{-\alpha}) = \frac{1}{2}\,\mathrm{ad}(x_\alpha)(f_1(x_{-\alpha})) = \frac{1}{2}\,\mathrm{ad}(x_\alpha)(h_\alpha) = -\frac{1}{2}[h_\alpha x_\alpha] = -\frac{1}{2}\alpha(h_\alpha) x_\alpha = -x_\alpha,$$

as $\alpha(h_\alpha) = 2$ by (Corollary 10.7). Then for $m = 3$, we have

$$f_m(x_{-\alpha}) = \frac{1}{3}\,\mathrm{ad}(x_\alpha)(f_2(x_{-\alpha})) = \frac{1}{3}[x_\alpha(-x_\alpha)] = -\frac{1}{3}[x_\alpha x_\alpha] = 0.$$

Therefore, by (19.7.3), $f_m(x_{-\alpha}) \in L(\mathbb{Z})$ for all $m \in \mathbb{N}$.

Now let $\beta \in \Phi$ such that $\beta \neq \pm\alpha$. Denote the α-string through β by S_α^β. That is,

$$S_\alpha^\beta = \{\beta - r\alpha, \cdots, \beta + q\alpha\}.$$

For each $i \in \mathbb{Z}$, consider $S_\alpha^{\beta+i\alpha}$, the α-string through $\beta + i\alpha$. We write

$$S_\alpha^{\beta+i\alpha} = \{(\beta + i\alpha) - r_i\alpha, \cdots, (\beta + i\alpha) + q_i\alpha\}.$$

Therefore, as $S_\alpha^\beta = S_\alpha^{\beta+i\alpha}$, we have

$$\begin{aligned} S_\alpha^\beta &= \{\beta - r\alpha, \cdots, \beta + q\alpha\} \\ &= \{(\beta + i\alpha) - (r + i)\alpha, \cdots\}, \end{aligned}$$

hence $r_i = r + i$. Therefore, by (Theorem 17.10),

$$[x_\alpha x_{\beta+i\alpha}] = \pm(r_i + 1) x_{\beta+(i+1)\alpha} = \pm(r + i + 1) x_{\beta+(i+1)\alpha}, \tag{19.7.4}$$

hence we can calculate

$$\begin{aligned} f_m(x_\beta) &= \left(\frac{\mathrm{ad}(x_\alpha)^m}{m!}\right)(x_\beta) \\ &= \frac{1}{m!}\,\mathrm{ad}(x_\alpha)^{m-1}([x_\alpha x_\beta]). \end{aligned}$$

With repeated application of (19.7.4), we obtain

$$\begin{aligned} f_m(x_\beta) &= \frac{1}{m!}\left(\prod_{i=1}^{m}(r + i)\right) x_{\beta+m\alpha} \\ &= \frac{1}{m!}\left(\frac{(r + m)!}{r!}\right) x_{\beta+m\alpha} \end{aligned}$$

$$= \frac{(r+m)!}{m!r!} x_{\beta+m\alpha}$$
$$= \binom{r+m}{m} x_{\beta+m\alpha},$$

where the last line uses the binomial coefficient and we take $x_{\beta+m\alpha} = 0$ if $\beta + m\alpha \notin \Phi$. That is, $f_m(x_\beta) \in L(\mathbb{Z})$ for al $m \in \mathbb{N}$.

We have shown that $f_m(B) \subseteq L(\mathbb{Z})$ for all $m \in \mathbb{N}$. Therefore, applying (19.7.1), we have that $f_m(L(\mathbb{Z})) \subseteq L(\mathbb{Z})$ for all $m \in \mathbb{N}$. □

19.8. **Proposition.** *Let $\alpha \in \Phi$. Then $\exp(\mathrm{ad}(x_\alpha))$ leaves $L(\mathbb{Z})$ invariant.*

Proof. By Lemma 19.7,

$$\forall i \in \mathbb{N} : \left(\frac{\mathrm{ad}(x_\alpha)^i}{i!}\right)(L(\mathbb{Z})) \subseteq L(\mathbb{Z}).$$

We can express $\exp(\mathrm{ad}(x_\alpha))$ as a finite sum, by (Remark 19.4). Therefore,

$$\begin{aligned}
\exp(\mathrm{ad}(x_\alpha))(L(\mathbb{Z})) &= \left(\sum_{i=1}^{m} \frac{\mathrm{ad}(x_\alpha)^i}{i!}\right)(L(\mathbb{Z})) \\
&= \sum_{i=1}^{m} \left(\frac{\mathrm{ad}(x_\alpha)^i}{i!}\right)(L(\mathbb{Z})). \\
&\subseteq \sum_{i=1}^{m} L(\mathbb{Z}) \\
&= L(\mathbb{Z}),
\end{aligned}$$

hence $\exp(\mathrm{ad}(x_\alpha))$ leaves $L(\mathbb{Z})$ invariant. □

19.9. **Lemma.** *For $\lambda \in \mathbb{F}$, we have that $\exp(\mathrm{ad}(\lambda x_\alpha))$ is an automorphism of L.*

Proof. We have that $\mathrm{ad}(x)$ is a derivation of L for any $x \in L$. By Lemma 19.3, $\mathrm{ad}(x_\alpha)$ is nilpotent. Thus $\mathrm{ad}(\lambda x_\alpha) = \lambda \,\mathrm{ad}(x_\alpha)$ is a nilpotent derivation of L. Therefore, by Proposition 19.6, $\exp(\mathrm{ad}(\lambda x_\alpha))$ is an automorphism of L. □

19.10**.** It turns out that $\exp(\mathrm{ad}(\lambda x_\alpha))$ only depends on λ is a specific way: elements of the Chevalley basis B are sent to $\mathbb{Z}[\lambda]$-linear combinations in B. Further, the coefficients in these polynomials in $\mathbb{Z}[\lambda]$ do not depend on λ. Therefore, if we replace λ with an indeterminate T, we obtain a map taking elements of B to $\mathbb{Z}[T]$-linear combinations in B. These maps are the automorphisms of $L(K)$ from which we construct our Chevalley groups.

19.11. **Proposition.** *Let T be an indeterminate. Let $e_1, \cdots, e_n$ denote the elements of the Chevalley basis B (in no particular order). Then, for each $\alpha \in \Phi$, $i \in \mathbb{N}$ and $j = 1, \cdots, n$,*

$$\left(\frac{\mathrm{ad}(Tx_\alpha)^i}{i!}\right)(e_j) = \sum_{k=1}^{n} \mu(\alpha, i, j, k) T^i e_k,$$

for some $\mu(\alpha, i, j, k) \in \mathbb{Z}$ and

$$\exp(\mathrm{ad}(Tx_\alpha))(e_j) = \sum_{k=1}^{n} f_{\alpha,j,k}(T) e_k,$$

for some $f_{\alpha,j,k} \in \mathbb{Z}[T]$.

Proof. By Lemma 19.7, for $e_j \in B$ and $i \in \mathbb{N}$, we have

$$\left(\frac{\operatorname{ad}(x_\alpha)^i}{i!}\right)(e_j) = \sum_{k=1}^{n} \mu(\alpha, i, j, k) e_k,$$

where $\mu(\alpha, i, j, k) \in \mathbb{Z}$. Therefore,

$$\begin{aligned}\left(\frac{\operatorname{ad}(Tx_\alpha)^i}{i!}\right)(e_j) &= T^i \left(\frac{\operatorname{ad}(x_\alpha)^i}{i!}\right)(e_j) \\ &= T^i \sum_{k=1}^{n} \mu(\alpha, i, j, k) e_k \\ &= \sum_{k=1}^{n} \mu(\alpha, i, j, k) T^i e_k.\end{aligned}$$

Thus,

$$\begin{aligned}\exp(\operatorname{ad}(Tx_\alpha))(e_j) &= \sum_{i=1}^{m} \left(\frac{\operatorname{ad}(Tx_\alpha)^i}{i!}\right)(e_j) \\ &= \sum_{i=1}^{m} \sum_{k=1}^{n} \mu(\alpha, i, j, k) T^i e_k \\ &= \sum_{k=1}^{n} \left(\sum_{i=1}^{m} \mu(\alpha, i, j, k) T^i\right) e_k \\ &= \sum_{k=1}^{n} f_{\alpha,j,k}(T) e_k.\end{aligned}$$

As each $\mu(\alpha, i, j, k) \in \mathbb{Z}$, we have that

$$f_{\alpha,j,k}(T) = \sum_{i=1}^{m} \mu(\alpha, i, j, k) T^i \in \mathbb{Z}[T].$$

□

19.12. **Definition.** For each $\alpha \in \Phi$, denote the matrix representation (with respect to B) of $\exp(\operatorname{ad}(Tx_\alpha))$ by $M_\alpha(T)$.

19.13. **Corollary.** *For each $\alpha \in \Phi$, the entries of the matrix $M_\alpha(T)$ are elements of $\mathbb{Z}[T]$.*

19.14. **Proposition.** *For each $\alpha \in \Phi$, the matrix $M_\alpha(T)$ is invertible.*

Proof. Fix $\alpha \in \Phi$ and write $M = M_\alpha(T)$. As $\operatorname{ad}(Tx_\alpha) = T \operatorname{ad}(x_\alpha) = 0$ if $\operatorname{ad}(x_\alpha) = 0$, we can extend Remark 19.4 to give that

$$\exp(\operatorname{ad}(Tx_\alpha)) = \sum_{i=0}^{m} \frac{\operatorname{ad}(Tx_\alpha)^i}{i!}.$$

Therefore, $M = M_0 + M_1 + \cdots + M_m$, where each M_i is the matrix representation of

$$\frac{\operatorname{ad}(Tx_\alpha)^i}{i!}.$$

Note that this implies that M_0 is the identity matrix, I_n.

Let $e_1, \cdots, e_n$ denote the elements of the Chevalley basis B. By Proposition 19.11, for each $i \in \mathbb{N}$ and $j = 1, \cdots, n$,

$$\left(\frac{\mathrm{ad}(Tx_\alpha)^i}{i!}\right)(e_j) = \sum_{k=1}^{n} \mu(\alpha, i, j, k) T^i e_k,$$

for some $\mu(\alpha, i, j, k) \in \mathbb{Z}$. As $\mathrm{ad}(x_\alpha)$ is nilpotent (Lemma 19.3), M_i is nilpotent for each $i > 0$. Further,

$$\left(\frac{\mathrm{ad}(Tx_\alpha)^i}{i!}\right)\left(\frac{\mathrm{ad}(Tx_\alpha)^j}{j!}\right) = \frac{\mathrm{ad}(Tx_\alpha)^{i+j}}{i!j!} = \left(\frac{\mathrm{ad}(Tx_\alpha)^j}{j!}\right)\left(\frac{\mathrm{ad}(Tx_\alpha)^i}{i!}\right),$$

so these M_i commute.

Let $N = M_1 + \cdots + M_m$. Sums of commuting nilpotent endomorphisms are nilpotent (Lemma 3.2), thus N is nilpotent. Therefore, for some change of basis, N is strictly upper triangular, hence $M = I_n + N$ is upper triangular with each diagonal entry equal to 1:

$$M = \begin{pmatrix} 1 & M_{2,1} & \cdots & M_{n,1} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & M_{n,n-1} \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Thus M has determinant 1 and is invertible. □

19.15. **Definition.** Let $G(K) = \langle M_\alpha(\kappa) : \alpha \in \Phi, \kappa \in K \rangle$. This group is called a **Chevalley group (of adjoint type)**.

19.16. **Theorem.** *The group $G(K)$ acts on $L(K)$ as a group of automorphisms.*

Proof. Let $M_\alpha(\kappa) \in G(K)$. Firstly, we need to show that $M_\alpha(\kappa)$ acts on $L(K)$. By Proposition 19.11, we have that

$$M_\alpha(\kappa) e_i = \sum_{j=1}^{n} f_{\alpha,i,j}(\kappa) e_j,$$

for $f_{\alpha,i,j}(\kappa) \in \mathbb{Z}[\kappa]$. As K is a field, it is a natural $\mathbb{Z}$-module via

$$a\kappa = \sum_{i=1}^{a} \kappa,$$

hence we can interpret $f_{\alpha,i,j}(\kappa)$ as an element of K. Therefore,

$$M_\alpha(\kappa) e_i \in \mathrm{span}_K(B) = L(K),$$

and by extension,

$$M_\alpha(\kappa)(L(K)) \subseteq L(K).$$

By Proposition 19.14, $M_\alpha(\kappa)$ is invertible, hence is a bijection. So for $M_\alpha(\kappa)$ to be an automorphism, we only need to show that is is a Lie algebra homomorphism. That is, for $x, y \in L(K)$, we need to show that

$$M_\alpha(\kappa)([xy]) = [M_\alpha(\kappa)(x) M_\alpha(\kappa)(y)].$$

As $L(K) = \mathrm{span}_K(B)$, we need only show that, for $e_i, e_j \in B$,

$$M_\alpha(\kappa)([e_i e_j]) = [M_\alpha(\kappa)(e_i) M_\alpha(\kappa)(e_j)]$$

and the desired result will follow by linearity.

By Proposition 19.11,

$$\exp(\mathrm{ad}(\lambda x_\alpha))(e_j) = \sum_{k=1}^{n} f_{\alpha,j,k}(\lambda) e_k,$$

for some $f_{\alpha,j,k}(\lambda) \in \mathbb{Z}[\lambda]$, with coefficients not dependent on λ. Therefore,

$$\begin{aligned}
\exp(\mathrm{ad}(\lambda x_\alpha))([e_i e_j]) &= \exp(\mathrm{ad}(\lambda x_\alpha)) \left(\sum_{k=1}^{n} c_{i,j}^k e_k \right) \\
&= \sum_{k=1}^{n} c_{i,j}^k \exp(\mathrm{ad}(\lambda x_\alpha))(e_k) \\
&= \sum_{k=1}^{n} c_{i,j}^k \sum_{l=1}^{n} f_{\alpha,k,l}(\lambda) e_l \\
&= \sum_{l=1}^{n} \left(\sum_{k=1}^{n} c_{i,j}^k f_{\alpha,k,l}(\lambda) \right) e_l \\
&= \sum_{l=1}^{n} g_{\alpha,i,j,l}(\lambda) e_l, \qquad (19.16.1)
\end{aligned}$$

where the coefficient of e_l, $g_{\alpha,i,j,l}(\lambda) \in \mathbb{Z}[\lambda]$, is a polynomial with coefficients not dependent of λ. Further,

$$\begin{aligned}
[\exp(\mathrm{ad}(\lambda x_\alpha))(e_i) \exp(\mathrm{ad}(\lambda x_\alpha))(e_j)] &= \left[\left(\sum_{k=1}^{n} f_{\alpha,i,k}(\lambda) e_k \right) \left(\sum_{l=1}^{n} f_{\alpha,j,l}(\lambda) e_l \right) \right] \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n} f_{\alpha,i,k}(\lambda) f_{\alpha,j,l}(\lambda) [e_k e_l] \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n} f_{\alpha,i,k}(\lambda) f_{\alpha,j,l}(\lambda) \sum_{t=1}^{n} c_{k,l}^t e_t \\
&= \sum_{t=1}^{n} \left(\sum_{k=1}^{n} \sum_{l=1}^{n} c_{k,l}^t f_{\alpha,i,k}(\lambda) f_{\alpha,j,l}(\lambda) \right) e_t \\
&= \sum_{t=1}^{n} h_{\alpha,i,j,t}(\lambda) e_t, \qquad (19.16.2)
\end{aligned}$$

where the coefficient of e_t, $h_{\alpha,i,j,t}(\lambda) \in \mathbb{Z}[\lambda]$, is - again - a polynomial with coefficients not dependent on λ. By Lemma 19.9, $\exp(\mathrm{ad}(\lambda x_\alpha))$ is an automorphism of L, hence (19.16.1) and (19.16.2) are equal for all $\lambda \in \mathbb{F}$. Therefore they must be coordinatewise equal, hence $g_{\alpha,i,j,t}(\lambda) = h_{\alpha,i,j,t}(\lambda)$ for each $t = 1, \cdots, n$. As the coefficients of these polynomials do not depend on λ, and they are equal for all $\lambda \in \mathbb{F}$, they must be equal as polynomials. Therefore they are still equal when taking values of $\kappa \in K$, hence

$$\exp(\mathrm{ad}(\kappa x_\alpha))([e_i e_j]) = [\exp(\mathrm{ad}(\kappa x_\alpha))(e_i) \exp(\mathrm{ad}(\kappa x_\alpha))(e_j)].$$

That is,

$$M_\alpha(\kappa)([e_i e_j]) = [M_\alpha(\kappa)(e_i) M_\alpha(\kappa)(e_j)].$$

□

19.17. **Example.** We shall use our toy example of $L = \mathfrak{sl}_2(\mathbb{F})$ to demonstrate the process of constructing a Chevalley group. Let $K = \mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$. For concreteness, we could take $\mathbb{F} = \mathbb{C}$, however, as the structure constants are integers with respect to the Chevalley basis, any algebraically closed field of characteristic zero gives the same Lie structure.

Recall the standard basis (x, h, y) of L. As $[xy] = h$ and there are no triples of roots $\alpha, \beta, \alpha + \beta \in \Phi$, this basis satisfies Proposition 17.7, hence is a Chevalley basis.

The root system Φ of L has two roots, α and $-\alpha$, where $x = x_\alpha$ and $y = x_{-\alpha}$. So the generators of $G(K)$ will be the matrices $M_\alpha(\kappa)$ and $M_{-\alpha}(\kappa)$ for $\kappa \in K$, where $M_\alpha(\kappa)$ is the matrix of $\exp(\mathrm{ad}(\kappa x))$ and $M_{-\alpha}(\kappa)$ is the matrix of $\exp(\mathrm{ad}(\kappa y))$. That is, we have $|\Phi| \cdot |K \setminus \{0\}| = 2 \cdot 2 = 4$ generators (when $\kappa = 0$ we obtain the identity). These generators are $\dim L \times \dim L = 3 \times 3$ matrices over $K = \mathbb{F}_3$.

Firstly, we need to compute the adjoint maps of x and y,

$$\mathrm{ad}(x) : \begin{cases} x \mapsto 0, \\ h \mapsto -2x, \\ y \mapsto h, \end{cases} \qquad \mathrm{ad}(y) : \begin{cases} x \mapsto -h, \\ h \mapsto 2y, \\ y \mapsto 0. \end{cases}$$

This gives us the matrices

$$X = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix},$$

respectively. Further, we have

$$X^2 = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Y^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix},$$

and $X^3 = Y^3 = 0$, Therefore, as $(\mathrm{ad}(Tx))^n = T^n(\mathrm{ad}(x))^n$, we have

$$M_\alpha(T) = I + TX + \frac{1}{2}T^2X^2 = \begin{pmatrix} 1 & -2T & -T^2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$M_{-\alpha}(T) = I + TY + \frac{1}{2}T^2Y^2 = \begin{pmatrix} 1 & 0 & 0 \\ -T & 1 & 0 \\ -T^2 & 2T & 1 \end{pmatrix}.$$

The generators for $G(K)$ are then

$$X_1 = M_\alpha(1) = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad X_2 = M_\alpha(2) = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix},$$

$$Y_1 = M_{-\alpha}(1) = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix}, \qquad Y_2 = M_{-\alpha}(2) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix},$$

and $I = M_\alpha(0) = M_{-\alpha}(0)$. From these, we obtain the elements

$$O_1 = X_1Y_1 = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{pmatrix}, \qquad O_2 = X_2Y_2 = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 0 & 2 \\ 2 & 1 & 1 \end{pmatrix},$$

$$L_1 = X_1Y_2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}, \qquad L_2 = X_2Y_1 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 2 & 2 \\ 2 & 2 & 1 \end{pmatrix},$$

$$U_1 = Y_2X_1 = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix}, \qquad U_2 = Y_1X_2 = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix},$$

and

$$D = O_1O_2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

The multiplication table is given by:

	I	X_1	X_2	Y_1	Y_2	O_1	L_1	L_2	O_2	U_2	U_1	D
I	I	X_1	X_2	Y_1	Y_2	O_1	L_1	L_2	O_2	U_2	U_1	D
X_1	X_1	X_2	I	O_1	L_1	L_2	O_2	Y_1	Y_2	U_1	D	U_2
X_2	X_2	I	X_1	L_2	O_2	Y_1	Y_2	O_1	L_1	D	U_2	U_1
Y_1	Y_1	O_2	U_2	Y_2	I	X_2	L_2	D	U_1	O_1	X_1	L_1
Y_2	Y_2	U_1	O_1	I	Y_1	U_2	D	L_1	X_1	X_2	O_2	L_2
O_1	O_1	Y_2	U_1	L_1	X_1	I	Y_1	U_2	D	L_2	X_2	O_2
L_1	L_1	D	L_2	X_1	O_1	U_1	U_2	O_2	X_2	I	Y_2	Y_1
L_2	L_2	L_1	D	O_2	X_2	X_1	O_1	U_1	U_2	Y_2	I	Y_2
O_2	O_2	U_2	Y_1	X_2	L_2	D	U_1	Y_2	I	X_1	L_1	O_1
U_2	U_2	Y_1	O_2	D	U_1	Y_2	I	X_2	L_2	L_1	O_1	X_1
U_1	U_1	O_1	Y_2	U_1	D	L_1	X_1	I	Y_1	O_2	L_2	X_2
D	D	L_2	L_1	U_2	U_1	O_2	X_2	X_1	O_1	Y_2	Y_1	I

From this, we can read off that $G(K)$ is a group of order 12, has 3 elements of order 2, namely O_1, O_2 and D, and 8 elements of order 3. This is the alternating group A_4, also known in the literature as $PSL_2(3)$. It should be noted that this is not a simple group. However, for larger fields K and larger Lie algebras, this process does give simple groups.

19.18**.** This process - taking a finite field K and generating a matrix group over K determined by L - is how the non-twisted Chevalley groups are constructed. For the twisted Chevalley groups, we need to take advantage of symmetries in the Dynkin diagrams. That is, we need the machinery behind the classification, which is covered in any good introductory text on Lie algebras [1, 3, 4, 6].

The classification is based upon a correspondence between the pairs (L, T) and (Φ, E). We have seen how a root system Φ in E is associated to the maximal toral subalgebra T of L (Theorem 12.12). We have also seen how L is determined by its roots, in that it is generated by the 1-dimensional root spaces L_α for $\alpha \in \Phi$ (Theorem 11.12). There are still some important questions to answer, however. For any given L, we may have a choice of T - does this freedom of choice affect the resulting root system? The answer is no: given two maximal tori T and T' of L, there exists an automorphism of L which maps T onto T' [4].

Another question is whether the group $G(K)$ is affected by the choice of Chevalley basis. We know (Lemma 18.3) that $L(\mathbb{Z})$, and hence $G(K)$, is unaffected by the choice of Δ, but there is still some freedom when choosing the x_α vectors in the Chevalley basis. It turns out that this freedom of choice actually reduces down to merely a choice of sign [4], which therefore does not ultimately affect $G(K)$.

References

[1] R. Carter, *Lectures on Lie Groups and Lie Algebras* (London Mathematical Society, 1995).

[2] J. Conway, *Atlas of Finite Groups* (Oxford : Clarendon, 1985).

[3] K. Erdmann, *Introduction to Lie Algebras* (Springer, 2006).

[4] J. Humphreys, *Introduction to Lie Algebras and Representation Theory* (Springer, 1972).

[5] N. Jacobson, *Lie Algebras* (New York : Interscience Publishers, 1962).

[6] H. Samelson, *Notes on Lie Algebras* (Springer, 1990).

[7] R. Wilson, A New Approach to the Suzuki Groups. *Math. Proc. Camb. Phil. Soc.* **148** (2010), 425–428.

[8] R. Wilson, A Simple Construction of the Ree Groups of Type ${}^2F_4$. *J. Algebra* **323** (2010), 1468-1481.

[9] R. Wilson, A New Construction of the Ree Groups of Type ${}^2G_2$. *Proc. Edinb. Math. Soc.* **53** (2010), 531–542.